# Lecture 22

# $u$ substitution

We saw last time that to find areas (i.e., to compute integrals) we must find antiderivatives.

$u$ substitution is a trick for finding antiderivatives.

## 22.1 Basic laws of integration

Before we get to $u$ substitution, here are some basic properties about integrals.

(I) **You can concatenate intervals.**

$$\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx. \tag{22.1.1}$$

You actually used the above fact last time: To find the area of $f$ over an interval $[a, c]$, you can divide the interval into two pieces–$[a, b]$ and $[b, c]$—and computed the area over each of those smaller intervals, then add the result. Note that $b$ doesn't need to be a midpoint or anything; it's just any point between $a$ and $c$. Here, for example, is a consequence of the above fact:

$$\int_a^b f(x)dx = \int_a^c f(x)dx - \int_b^c f(x)dx. \tag{22.1.2}$$

(II) **Integrals scale.** Here is another fact. If $m$ is any real number, we have

$$\int_a^b mf(x)dx = m\int_a^b f(x)dx. \tag{22.1.3}$$

That is, area scales (by $m$) if you scale $f$ (by $m$). Intuitively, if you think of $f(x)$ as describing the height of a curvy fence at position $x$, if you make the fence $m$ times taller everywhere, the area also grows by $m$ times. Here, $m$ is any real number. It could be zero or negative or positive.

(III) **If you add the integrands, you add the integrals.** Another fact that will be useful is

$$\int_a^b f(x) + g(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx. \tag{22.1.4}$$

Informally, if you have something of height $f(x)$ at $x$, then you add on a height of $g(x)$ at every $x$, then the area of the resulting figure is obtained by adding the areas given by $f$ and given by $g$.

(IV) **Reverse, reverse.** The final fact that will be useful is

$$\int_a^b f(x)dx = \int_b^a -f(x)dx \tag{22.1.5}$$

This is is a somewhat strange rule, but it will come up. So far, we've always had $\int_a^b$ consist of numbers such that $a < b$; but as we compute, we will sometimes end up with $\int_b^a$ with $b > a$. If this happens, you just reverse the sign of $f$ and reverse the roles of $b$ and $a$. If you like, $\int_a^b$ tells you to take the area "from $a$ to $b$," which usually has meant "from left to right." If the integral tells us to reverse directions by going from right to left (e.g., from $b$ to $a$), we think of the contributions from $f$ negatively.

That's it. Three rules. As a student, you might feel good that there's "less to memorize," but actually, having fewer rules means you have fewer guideposts to rely on, so you need to be more creative about how to compute integrals!

## 22.2 Some notation

For today you've practiced taking integrals. For example, to compute

$$\int_1^4 \frac{1}{x}dx,$$

you must

1. find an antiderivative $F$ to the integrand, and then

2. compute $F(b) - F(a)$ (which in this case if $F(4) - F(1)$).

We know that $\ln(x)$ has derivative given by $1/x$, so we can take $F(x) = \ln x$. We thus find

$$\int_1^4 \frac{1}{x} dx = \ln(4) - \ln(1) \tag{22.2.1}$$
$$= \ln(4) - 0 \tag{22.2.2}$$
$$= \ln 4. \tag{22.2.3}$$

Let me introduce the following:

**Notation 22.2.1.** We write

$$F(x)\Big|_a^b$$

to mean

$$F(b) - F(a).$$

In other words,

$$F(x)\Big|_a^b = F(b) - F(a).$$

**Example 22.2.2.** So for example, the above work could have been written

$$\int_1^4 \frac{1}{x} dx = \ln(x)\Big|_1^4 \tag{22.2.4}$$
$$= \ln(4) - \ln(1) \tag{22.2.5}$$
$$= \ln(4) - 0 \tag{22.2.6}$$
$$= \ln 4. \tag{22.2.7}$$

As usual, this notation is meant to *help* you. I promise you'll begin to use this notation in no time, because you'll always first the antiderivative first, and then you'll plug in $b$ and $a$. This "vertical bar" notation will help you keep track of those steps as you write.

## 22.3 More terms

Let me also review some things that you learned in lab.

**Definition 22.3.1.** The *indefinite integral* is an old terminology that survives mostly in calculus textbooks. The notation

$$\int f(x)dx$$

is called an "indefinite" integral because the integral does not specify *where* we are integrating—i.e., there is no $a$ and no $b$.

A lot of calculus textbooks ask you to "solve the indefinite integral." I am not a fan—at all—of this terminology. Regardless, because it is so prevalent in calculus culture, I have to tell you about it to prevent you from experiencing confusion or frustration should some other calculus professor torture you with this old terminology.

To "solve an indefinite integral" means to "find all antiderivatives of $f(x)$." (You saw last time—via the Fundamental Theorem of Calculus (FTC)—why finding an antiderivative helps you find the integral.)

If $F$ is an antderivative of $f(x)$, it is customary to write the following as a "correct answer" to the calculus textbook's problem of finding an indefinite integral:

$$\int f(x)dx = F(x) + C.$$

The $+C$ is a perennially confusing notation. Let me explain it. Recall from the lecture about the mean value theorem that we discovered the following fact: If $G$ and $F$ have the same derivative (meaning $F' = G'$) then $F - G$ must be a constant function. For example, it could be that $F - G = \pi$ or $F - G = 0$ or $F - G = 3$. The important part is that it could be any constant. So if $F$ is an antiderivative, *every* other antiderivative can be expressed as $F+$ something, where something is a constant. that's what the $+C$ above means. The $C$ stands for "constant," and the notation $\int f(x)dx = F(x) + C$ is a lazy way of saying "any antiderviative of $f$ can be obtained from $F(x)$ by adding some constant $C$."

**Example 22.3.2.** The indefinite integral $\int \sin(x)dx$ is given by

$$-\cos(x) + C.$$

I really dislike the language of "indefinite integral," because we already have a perfectly good word for the underlying idea one is exploring: Antiderivative. The "$+C$" warning really just gives you a heads up that a lot of integral tables and encyclopedia articles have $+C$ because an antiderivative really could be obtained from $F$ by adding some constant, and a particular constant may be very important for a particular engineering problem. But if some calculus person ever asks you to find an "indefinite" integral, just find the antiderivative and write "$+C$." And know that the only reason you're writing $+C$ is to indicate that any other antiderivative can be obtained by adding a constant to the one you found.

## 22.4 $u$ substitution

You can think of $u$ substitution as like a "reverse chain rule." Let me say what I mean.

Suppose $F(x) = g(h(x))$. That is, $F = g \circ h$, so that $F$ is a composition the functions $g$ and $h$. Then you know that

$$F'(x) = g'(h(x)) \cdot h'(x). \tag{22.4.1}$$

That's the chain rule.

In the last lecture, we saw the importance of being able to *work backwards*—that is, can you recognize when you see something like $g'(h(x)) \cdot h'(x)$? If so, all you need to do to find the antiderivative is

- Recognize $h$, and
- Take the antiderivative of $g'$. Then to conclude, just
- Set $F = g \circ h$.

**Exercise 22.4.1.** Find an antiderivative for the following functions:

(a) $f(x) = 2x\cos(x^2)$.

(b) $f(x) = \frac{2x}{x^2+3}$

(c) $f(x) = -\sin(\sin(x)) \cdot \cos(x)$.

I am not exactly sure of why—perhaps because it is hard to recognize two derivatives ($g'$ and $h'$) at once—calculus textbooks teach us a technique called *$u$ substitution* to find antiderivatives in situations like this. It can sometimes be confusing, and though I am not a huge fan of $u$ substitution, I will teach it to you in case you find it easier than eye-balling the chain rule.

The way $u$ substitution works is by identifying the $h$ in the equation (22.4.1). For example, consider the indefinite integral

$$\int \cos(x)\sqrt{\sin(x)}dx. \tag{22.4.2}$$

You might recognize a "function within a function," i.e., a composition, in $\sqrt{\sin(x)}$. You might recognize that the "inside function"—$\sin(x)$—has a derivative given by the factor outside the $\sqrt{--}$ symbol, namely the $\cos(x)$ factor. Thus you can verify that the inside function $h(x) = \sin(x)$ says that our integrand is of the form $g'(h(x)) \cdot h'(x)$. In this case, then, we see that $g'$ must be the square root function.

But, rather than thinking this all through, $u$ substitution encourages you to stop thinking and try to do algebra instead. (I am not a fan.) Here is how you do it:

Step One: One substitutes the inside function by a variable $u$. You should think of $u$ as a function of $x$. So, for example, a naive re-writing of (22.4.2) gives

$$\int \cos(x)\sqrt{u}dx. \tag{22.4.3}$$

Things look worse right now—there is a $u$ and an $x$ and who knows what in the world this means. Here is the (useful?) confusing part:

**Notation 22.4.2** ($du$)**.** Because $u$ is a function of $x$, we can introduce a new symbol called

$$du$$

that is *defined* to satisfy the following property:

$$du = \frac{du}{dx}dx. \tag{22.4.4}$$

I warn you that $du$ and $dx$ are just symbols—they are *not* numbers—so the fraction notation is more misleading than it is useful. You can't just cancel symbols willy-nilly without knowing what they mean. Regardless, $du$—as a symbol—is defined precisely a way that encourages such dangerous (and, in this case, correct) cancellation. Indeed, note that the lefthand side of (22.4.4) can be obtained from the righthand side by "cancelling" the $dx$.

Or, rearranging (22.4.4), we find

$$dx = \frac{1}{\frac{du}{dx}} \; du. \tag{22.4.5}$$

End of notation.

Step Two: We plug in $u(x) = \sin(x)$, so that $\frac{du}{dx} = \cos(x)$. Then we can continue to simplify (22.4.3):

$$\int \cos(x)\sqrt{u}dx = \int \cos(x)\sqrt{u}\frac{1}{\frac{du}{dx}} \; du \tag{22.4.6}$$

$$= \int \cos(x)\sqrt{u}\frac{1}{\cos(x)} \; du \tag{22.4.7}$$

$$= \int \sqrt{u} \; du. \tag{22.4.8}$$

Notice that we have used the definition of $du$ to get rid of the $dx$.

Step Three: Take the integral in terms of $u$. What the indefinite integral in (22.4.8) is asking is: Can you find the antiderivative of the square root function? Yes, you can! Moreover, the integral is no longer viewing the integrand as a function of $x$; the "$du$" symbol is telling you to think of the integrand as a function of $u$. Well,

$$\frac{d}{du}(u^{3/2}) = \frac{3}{2}u^{1/2},$$

so we find that

$$\frac{d}{du}\frac{2}{3}(u^{3/2}) = u^{1/2}.$$

In other words, we can solve the indefinite integral in (22.4.8) to find

$$\int \sqrt{u}\ \ du = \frac{2}{3}u^{3/2}. \tag{22.4.9}$$

And now let's plug back in what $u$ equals; we defined $u$ to be $u(x) = \sin(x)$, so the righthand side of (22.4.9) becomes

$$\frac{2}{3}u^{3/2} = \frac{2}{3}(\sin(x))^{3/2} = \frac{2}{3}\sqrt{\sin(x)^3}.$$

Indeed, you can check that this function of $x$ is an antiderivative of our original function $\cos(x)\sqrt{\sin(x)}$.

Here is the **summary of $u$ substitution**:

$$\int g'(h(x))h'(x)dx = \int g'(u)du.$$

In the end, if you find the integral $\int g'(u)du = g(u)$, make sure you substitute back in $h(x) = u(x)$ to get

$$\int g'(h(x))h'(x)dx = g(h(x)).$$

Here is (what I think is) a good application of $u$ substitution.

**Exercise 22.4.3.** Find

$$\int \tan(x)dx.$$

## 22.5 The integral of $\tan(x)$

Let's note

$$\int \tan(x)dx = \int \frac{\sin x}{\cos x}dx = \int \sin(x) \cdot \frac{1}{\cos(x)}dx.$$

We note that $\sin(x)$ is (almost) the derivative of $\cos(x)$—it's off by a sign. But it almost looks like we can take

$$g(x) = \frac{1}{x}, \qquad h(x) = \cos(x),$$

for then

$$g(h(x))h'(x) = \frac{1}{\cos(x)} \cdot (-\sin(x)).$$

So we have that

$$\int \sin(x) \cdot \frac{1}{\cos(x)}dx = -\int (-\sin(x)) \cdot \frac{1}{\cos(x)}dx. \tag{22.5.1}$$

Letting $u = \cos(x)$, we have that

$$du = -\sin(x)dx, \qquad dx = \frac{du}{-\sin(x)}.$$

Hence (22.5.1) becomes

$$-\int (-\sin(x)) \cdot \frac{1}{\cos(x)}dx = -\int (-\sin(x)) \cdot \frac{1}{u} \cdot \frac{du}{-\sin(x)} \tag{22.5.2}$$

$$= -\int \frac{1}{u}du \tag{22.5.3}$$

But you know how to integrate $\frac{1}{u}$; the antiderivative is $\ln(|u|)$. Hence we have

$$-\int \frac{1}{u}du = -\ln(|u|) + C.$$

Now, let's remember that $u(x) = \cos(x)$, so plugging this in, we have

$$\int \tan(x)dx = -\int \frac{1}{u}du = -\ln(|u|) + C = -\ln(|\cos(x)|) + C. \tag{22.5.4}$$

Here is one more simplification we can make: Remember the formula

$$a\ln(b) = \ln(b^a).$$

(If you don't remember it, you should verify it using what you know about exponent laws and the definition of ln!) In particular,

$$-\ln(b) = \ln(\frac{1}{b}).$$

Thus, we can further modify (22.5.4) to become

$$\int \tan(x)dx = \ln(\Big|\frac{1}{\cos(x)}\Big|) + C.$$

Or, if you like secant, which is defined by $\sec(x) = 1/\cos(x)$, you can rewrite this as

$$\int \tan(x)dx = \ln(\Big|\sec(x)\Big|) + C.$$

**Remark 22.5.1.** If you prefer the "eyeball" method, you could have recognized that $\tan(x)$ is of the form $\sin(x) \times$ something, and that this something has $\cos(x)$ in it. Thus you could be inspired to use the (reverse) chain rule.

$$\sin(x) \cdot \frac{1}{\cos(x)} = h'(x) \cdot g'(h(x)).$$

You recognize now that $g'(x)$ has to be $\frac{1}{x}$, so that $g(x)$ has to be $\ln|x|$. Then, by the (reverse) chain rule,

$$\int g'(h(x))h'(x)dx = g(h(x)) + C = \ln|\frac{1}{\cos x}| + C.$$

I much prefer this method, but there are uses for $u$ substition in one's life, so if you prefer to solve problems using $u$ substitution (which will require you to get used to manipulating equations like $du = \frac{du}{dx}\ dx$), go for it!

## 22.6 Preparation for next lecture

$u$ substitution isn't just for computing antiderivatives; it also allows you to compute integrals! Here is the new fact you'll practice for next quiz: If $u(x) = h(x)$, then

$$\int_a^b g'(h(x))h(x)dx = \int_{u(a)}^{u(b)} g'(u)du. \tag{22.6.1}$$

**Example 22.6.1.** Let's evaluate

$$\int_1^4 \frac{2x}{1+x^2} dx.$$

If I want to use $u$ substitution, I recognize that $2x$ is the derivative of $1+x^2$. So I will set $u(x) = 1 + x^2$, so that $du = 2xdx$. Then

$$\int \frac{2x}{1+x^2}\, dx = \int \frac{2x}{u} \cdot \frac{1}{2x}\, du = \int \frac{1}{u} du.$$

What the fact (22.6.1) tells us is that we can evaluate the definite integral in using the $u$ variable form of the integral:

$$\int_1^4 \frac{2x}{1+x^2} dx = \int_{u(1)}^{u(4)} \frac{1}{u} du.$$

So we find

$$\begin{aligned}
\int_{u(1)}^{u(4)} \frac{1}{u} du &= \ln|u| \Big|_{u(1)}^{u(4)} && (22.6.2)\\
&= \ln|u| \Big|_{1+1^2}^{1+4^2} && (22.6.3)\\
&= \ln|u| \Big|_{2}^{17} && (22.6.4)\\
&= \ln|17| - \ln|2| && (22.6.5)\\
&= \ln \frac{|17|}{|2|} && (22.6.6)\\
&= \ln \frac{17}{2}. && (22.6.7)
\end{aligned}$$

If we want, we could have computed this without using $u$ substitution. Again recognizing that if $h(x) = 1 + x^2$, then $h'(x)$, we have that the integrand is equal to $h'(x) \cdot \frac{1}{h(x)}$. Thus we want $g'(x) = \frac{1}{x}$, which has integral $g(x) = \ln|x|$. We conclude

$$\begin{aligned}
\int_1^4 g'(h(x))h'(x)\, dx &= g(h(x)) \Big|_1^4 && (22.6.8)\\
&= \ln|1+4^2| - \ln|1+1^2| && (22.6.9)\\
&= \ln|17| - \ln|2| && (22.6.10)\\
&= \ln \frac{17}{2}. && (22.6.11)
\end{aligned}$$

For next time, I expect you to be able to compute the following integrals (using $u$ substitution, or otherwise):

(a)
$$\int_0^1 x(x^2-1)^5 dx.$$

(b)
$$\int_0^{1/12} \frac{1}{\sqrt[3]{1-6x}} dx.$$

(c)
$$\int_2^3 x e^{x^2}\, dx$$

(d)
$$\int_0^1 x(x^2-1)^5 dx.$$

(e)
$$\int_{\pi/4}^{\pi/2} \frac{\cos(x)}{\sin^2(x)}\, dx.$$