Math 3204: Calculus IV[1] Mikhail Lavrov

# Lecture 3: Substitution with two variables

*August 21, 2023* *Kennesaw State University*

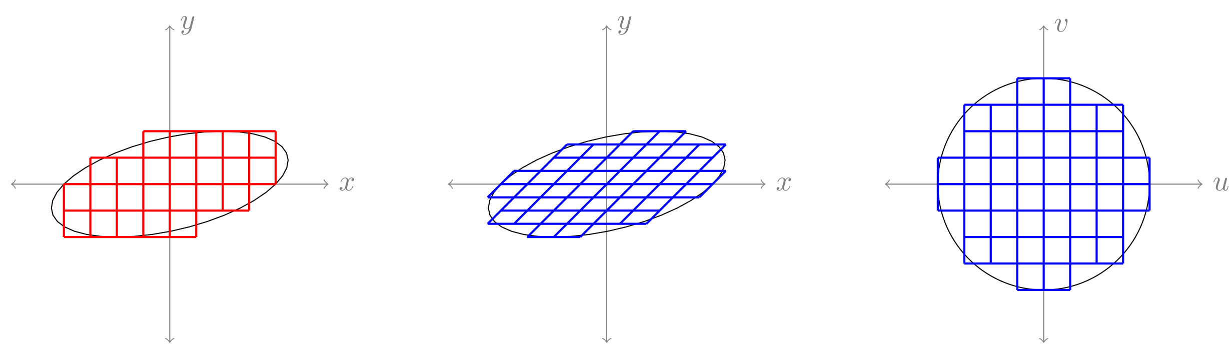
## 1 A motivational example

Suppose we want to know the area of the ellipse defined by $x^2 - 2xy + 5y^2 \le 1$. We could set this up as a rather horrific integral:

$$\int_{x=-\sqrt{5}/2}^{\sqrt{5}/2} \int_{y=(x-\sqrt{5-4x^2})/5}^{(x+\sqrt{5-4x^2})/5} \mathrm{d}y\,\mathrm{d}x.$$

This integral is defined by a Riemann sum. The idea behind that Riemann sum is that we approximate the ellipse by a number of $\Delta y \times \Delta x$ squares (as in Figure 1a), multiply the number of squares it takes by the area of each square, and then take the limit as $\Delta x$ and $\Delta y$ both go to 0. (If they're squares, then really, $\Delta x$ and $\Delta y$ are the same quantity, but it's convenient to allow them to be different.)

We can make the ellipse look much nicer with a substitution: $x^2 - 2xy + 5y^2$ can be rewritten as $(x-y)^2 + (2y)^2$, so by setting $u = x - y$ and $v = 2y$, we can obtain the equation $u^2 + v^2 \le 1$. How do we use this to help us find the area?




(a) The ellipse $x^2 - 2xy + 5y^2 \le 1$ approximated by square cells

(b) The same ellipse approximated by parallelogram cells

(c) The same diagram in the $uv$-plane rather than the $xy$-plane

Figure 1: Approximating the area of an ellipse ($\Delta x = \Delta y = \Delta u = \Delta v = \frac{1}{4}$ in all examples)

One way to visualize what's going on is by drawing a skew lattice on our ellipse. Rather than drawing lines for many $x$-values and many $y$-values, we draw lines for many $u$-values (which are $\Delta u$ apart) and many $v$-values (which are $\Delta v$ apart). This divides our ellipse into many parallelograms (as in Figure 1b). To find the area, we can count the number of parallelograms, multiply by the area of each parallelogram, and then take the limit as $\Delta u$ and $\Delta v$ both go to 0.

[1]This document comes from the Math 3204 course webpage: http://facultyweb.kennesaw.edu/mlavrov/courses/3204-fall-2023.php

Another way to visualize $u$ and $v$ is to draw a diagram in the $uv$-plane rather than the $xy$-plane (as in Figure 1c). Here, the shape of $u^2 + v^2 \leq 1$ is much nicer: it's a circle of radius 1. The area is *still* given by a limit as $\Delta u$ and $\Delta v$ go to 0 of the number of $\Delta u \times \Delta v$ squares times the area of each square. However, we know the area in the $uv$-plane directly: it's just $\pi$, the area of a circle of radius 1.

We can also relate the approximations we see to each other. The number of small cells (call it $N$) is exactly the same in both diagrams, because the ellipse and the circle are divided up in exactly the same way. The area of the ellipse is approximately $N$ times the area of a parallelogram. The area of the circle is approximately $N$ times the area of a square. So the ratio in areas between the ellipse and the circle is exactly the same as the ratio in areas between the parallelogram and the square.

At this point, I am going to invoke a fact from linear algebra, without proof (because it will take us too far off track). Let $\mathbf{f}(u, v) = (u + \frac{1}{2}v, \frac{1}{2}v)$ be the linear transformation that turns our $(u, v)$ coordinates back into $(x, y)$ coordinates. (In other words, $\mathbf{f}$ is the inverse of the transformation that turns $(x, y)$ into $u = x - y$ and $v = 2y$.) Then:

**Claim 1.1.** *A linear transformation $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ scales all areas by the absolute value of the determinant of the matrix representing $\mathbf{f}$.*

In this particular case,

$$\mathbf{f}(u, v) = \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix}, \quad \text{and} \qquad \det \begin{bmatrix} 1 & 1/2 \\ 0 & 1/2 \end{bmatrix} = 1 \cdot \frac{1}{2} - \frac{1}{2} \cdot 0 = \frac{1}{2}.$$

So the area of each parallelogram in the $xy$-plane is $\frac{1}{2}$ the area of each square in the $uv$-plane, and the area of the ellipse is $\frac{\pi}{2}$: half the area of the circle.[2]

We will take Claim 1.1 as given for today. The only thing we need to know about determinants for now is the rule for the determinant of a $2 \times 2$ matrix:

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$

## 2 The Jacobian determinant

The idea we used in the example in the previous section can be applied to any change of variables $(x, y) \rightsquigarrow (u, v)$, turning an integral $\mathrm{d}x\,\mathrm{d}y$ into an integral $\mathrm{d}u\,\mathrm{d}v$. There are only two changes that happen when we generalize:

- First of all, in a general Riemann sum, we multiply the area of each small cell by the function we're integrating, evaluated at a point in that cell.

  This just means that in general, our new integral should also have a corresponding expression inside it, just written in terms of $u$ and $v$ rather than $x$ and $y$.

- Second, most changes of variables are not linear. For a nonlinear change of variables, the ratio of areas between the small cells will vary over our region.

[2] In this example, we could have used Claim 1.1 directly, without dividing the regions up into cells, but the division into cells will be important later.

To deal with the second bullet point, we need derivatives. Suppose we have a transformation $\mathbf{f}(u,v) = (x(u,v), y(u,v))$ which gives the $(x,y)$ point corresponding to $(u,v)$. Then a linear *approximation* to $\mathbf{f}$ near a point $(u_0, v_0)$ is

$$\mathbf{f}(u,v) \approx \mathbf{f}(u_0,v_0) + \left(\frac{\partial x}{\partial u}(u_0,v_0), \frac{\partial y}{\partial u}(u_0,v_0)\right)(u-u_0) + \left(\frac{\partial x}{\partial v}(u_0,v_0), \frac{\partial y}{\partial v}(u_0,v_0)\right)(v-v_0).$$

(As a reminder, an expression like $\frac{\partial x}{\partial u}$ is a **partial derivative**: we take the derivative of $x(u,v)$ with respect to $u$, treating $v$ as a constant.)

By Claim 1.1, the area of a cell in the $xy$-plane near $\mathbf{f}(u_0,v_0)$ is approximately the are of a cell in the $uv$-plane near $(u_0,v_0)$, multiplied by

$$\left|\det \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}\right|.$$

This matrix of partial derivatives is called the **Jacobian matrix**, and its determinant is called the **Jacobian determinant**. We write it so often that we will give it special notation: $\frac{\partial(x,y)}{\partial(u,v)}$. (Note: other sources use this notation for the matrix itself, not just its determinant, but we will only ever need the determinant of the matrix.)

Let's see how this works on another example. Suppose we are working with the integral

$$I := \int_{x=1}^{2} \int_{y=1/x}^{2/x} \frac{\sqrt{x+1}}{y} \, \mathrm{d}y \, \mathrm{d}x.$$

To make the $\sqrt{x+1}$ easier to integrate, we substitute $u = x+1$; to make the bounds $\frac{1}{x} \le y \le \frac{2}{x}$ (or equivalently $1 \le xy \le 2$) simpler, we substitute $v = xy$.

To find $\frac{\partial(x,y)}{\partial(u,v)}$, we must first find $x$ and $y$ in terms of $u$ and $v$. From $u = x+1$, we have $x = u-1$, so $v = (u-1)y$, and therefore $y = \frac{v}{u-1}$. This yields

$$\frac{\partial(x,y)}{\partial(u,v)} = \det \begin{bmatrix} 1 & 0 \\ -\frac{v}{(u-1)^2} & \frac{1}{u-1} \end{bmatrix} = \frac{1}{u-1}.$$

Now onto the bounds. If $1 \le x \le 2$, then $1 \le u-1 \le 2$, so $2 \le u \le 3$. If $\frac{1}{x} \le y \le 2x$, then $1 \le xy \le 2$, so $1 \le v \le 2$. Finally, we replace $x$ and $y$ by $u-1$ and $(u-1)y$ inside the integral. We get

$$I = \int_{u=2}^{3} \int_{v=1}^{2} \frac{\sqrt{u}}{v/(u-1)} \cdot \frac{1}{u-1} \, \mathrm{d}v \, \mathrm{d}u = \int_{u=2}^{3} \int_{v=1}^{2} \frac{\sqrt{u}}{v} \, \mathrm{d}v \, \mathrm{d}u.$$

(In general, $\frac{1}{u+1}$ would need to be $|\frac{1}{u+1}|$, but when $2 \le u \le 3$, $\frac{1}{u+1}$ is already always positive.)

We can factor this double integral into two single integrals:

$$I = \left(\int_{u=2}^{3} \sqrt{u} \, \mathrm{d}u\right)\left(\int_{v=1}^{2} \frac{\mathrm{d}v}{v}\right).$$

From here, simplifying $I$ to $\frac{2}{3}(3^{3/2} - 2^{3/2}) \ln 2$ is just single-variable calculus.

Before we go on to the final topic, here is a word of caution: these 2-variable substitutions only work if $\mathbf{f}(u,v) = (x(u,v), y(u,v))$ is a bijection between the region in the $uv$-plane and the region in the $xy$-plane. If parts of the $uv$-region "fold over" onto each other in the $xy$-region, then some parts of the original integral will end up overweighted in the integral we write after the substitution.

## 3 Oriented double integrals

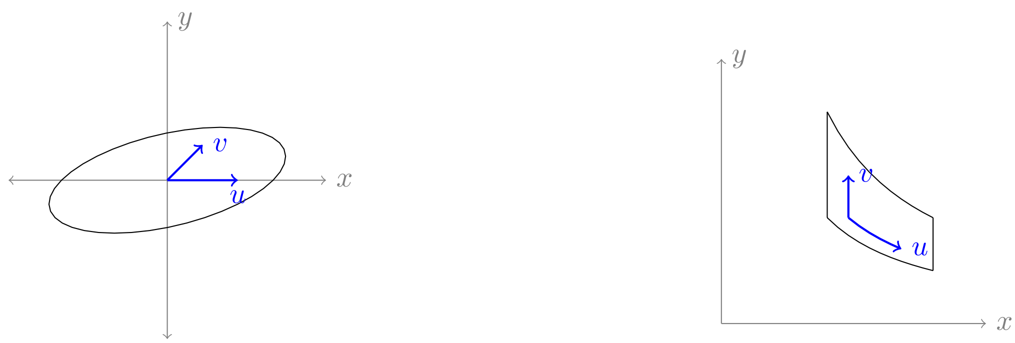
It is said: "Only a Sith deals in absolutes." The meaning of this proverb is that in mathematics, whenever you encounter an expression inside an absolute value, you should ask whether it has meaning without that absolute value as well. What if we replaced the rule

$$\mathrm{d}y\,\mathrm{d}x = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \mathrm{d}u\,\mathrm{d}v$$

by a rule that did not take the absolute value of $\frac{\partial(x,y)}{\partial(u,v)}$?

This would be a weird rule that sometimes made even our areas negative, so we definitely wouldn't want to use it all the time. However, there *is* a meaning attached to the negative signs. If the determinant of a linear transformation $\mathbf{f} \colon \mathbb{R}^2 \to \mathbb{R}^2$ is negative, that means $\mathbf{f}$ is orientation-reversing: it turns shapes into their mirror images!

One way to track this is to compare how the axes are oriented. The way we draw the $xy$-plane, the positive $y$-axis is counterclockwise from the positive $x$-axis. So we say that a substitution is **orientation-preserving** at a point if the direction of increasing $v$ is clockwise from the direction increasing $u$, rather than counterclockwise. Otherwise, it is **orientation-reversing**.




(a) The orientation of our first substitution

(b) The orientation of our second substitution

Figure 2: Geometrically, we can see that both of our $uv$-substitutions are orientation-preserving

A better way to say this is that an **oriented region** in $\mathbb{R}^2$ is a region which has a notion of "clockwise" and "counterclockwise" at each point. A region $R$ whose notion of "clockwise" and "counterclockwise" matches the way we normally draw $\mathbb{R}^2$ is positively oriented. It has an evil twin: a region $R'$ which is identical as a set, but has the reverse orientation at each point.

When we write oriented integrals, we will use slightly different notation to emphasize this: rather than writing "$\mathrm{d}x\,\mathrm{d}y$" at the end, we will write "$\mathrm{d}x \wedge \mathrm{d}y$". The $\wedge$ (wedge) symbol has other meanings we will not discuss now. For now, all we need to know is that

$$\iint_R g(x,y)\,\mathrm{d}x \wedge \mathrm{d}y$$

is an oriented integral which is equal to the ordinary integral of $g(x, y)$ over $R$ when $R$ is positively oriented, and to the negative of that integral when $R$ is negatively oriented.

Suppose that our substitution $(x(u, v), y(u, v))$ is a bijection between a region $S$ in the $uv$-plane and a region $R$ in the $xy$-plane, *as sets*. Then, depending on the sign of $\frac{\partial(x,y)}{\partial(u,v)}$, the substitution could either be a bijection from $S$ to $R$, or from $S$ to $R'$, *as oriented regions*. If $\frac{\partial(x,y)}{\partial(u,v)} < 0$, then clockwise movement in $S$ will correspond to counterclockwise movement in $R$, so $S$ really corresponds to $R'$, not to $R$. As a result, with oriented integrals, we have a simpler substitution rule:

$$\iint_R g(x,y)\,\mathrm{d}x \wedge \mathrm{d}y = \iint_S g(x(u,v), y(u,v))\frac{\partial(x,y)}{\partial(u,v)}\,\mathrm{d}u \wedge \mathrm{d}v.$$

The sign change potentially introduced by $\frac{\partial(x,y)}{\partial(u,v)}$ will exactly match the sign change potentially introduced by $R$ and $S$ having opposite orientations.

Is this useful? Maybe not so much, right now. It's a way to dip our toes into the world of integrals that come with orientations; we'll see more use out of that world later.

## 4 Wedge products

These oriented integrals come with a bonus feature: a different, equivalent way to compute the factor in the substitution rule! Here is how it is done:

1. In the differential $\mathrm{d}x \wedge \mathrm{d}y$, replace $\mathrm{d}x$ by $\frac{\partial x}{\partial u}\,\mathrm{d}u + \frac{\partial x}{\partial v}\,\mathrm{d}v$ and $\mathrm{d}y$ by $\frac{\partial y}{\partial u}\,\mathrm{d}u + \frac{\partial y}{\partial v}\,\mathrm{d}v$: this is another form of the multivariate chain rule.
2. Distribute $\wedge$ over addition and pull out non-differential factors, so that you get an expression in terms of $\mathrm{d}u \wedge \mathrm{d}u$, $\mathrm{d}u \wedge \mathrm{d}v$, $\mathrm{d}v \wedge \mathrm{d}u$, and $\mathrm{d}v \wedge \mathrm{d}v$.
3. Simplify these with two rules. First, $\mathrm{d}v \wedge \mathrm{d}u = -(\mathrm{d}u \wedge \mathrm{d}v)$: this makes sense, since these coordinate systems have opposite orientations. Second, $\mathrm{d}u \wedge \mathrm{d}u = \mathrm{d}v \wedge \mathrm{d}v = 0$: a "coordinate system" with the same variable repeated twice cannot measure any nonzero area.

Let's see this on two examples. First, our transformation of the ellipse, with $x = u + \frac{1}{2}v$ and $y = \frac{1}{2}v$. Here,

$$\mathrm{d}x \wedge \mathrm{d}y = (\mathrm{d}u + \tfrac{1}{2}\,\mathrm{d}v) \wedge (\tfrac{1}{2}\,\mathrm{d}v) = \tfrac{1}{2}\,\mathrm{d}u \wedge \mathrm{d}v + \tfrac{1}{4}\,\mathrm{d}v \wedge \mathrm{d}v = \tfrac{1}{2}\,\mathrm{d}u \wedge \mathrm{d}v.$$

Next, let's return to our previous substitution example, with $x = u - 1$ and $y = \frac{v}{u-1}$. Here, $\mathrm{d}x = \mathrm{d}u$ and $\mathrm{d}y = -\frac{v}{(u-1)^2}\,\mathrm{d}u + \frac{1}{u-1}\,\mathrm{d}v$, so $\mathrm{d}x \wedge \mathrm{d}y$ becomes

$$\mathrm{d}u \wedge \left(-\frac{v}{(u-1)^2}\,\mathrm{d}u + \frac{1}{u-1}\,\mathrm{d}v\right) = -\frac{v}{(u-1)^2}\,\mathrm{d}u \wedge \mathrm{d}u + \frac{1}{u-1}\,\mathrm{d}u \wedge \mathrm{d}v = \frac{1}{u-1}\,\mathrm{d}u \wedge \mathrm{d}v.$$

You should feel free to use either method—the Jacobian determinant, or the wedge product—to perform substitutions. They will give the same answer, as long as you remember that for a non-oriented integral, you need to take the absolute value of the scaling factor on the differential!