April 1, 2019

LECTURE 16: CHANGING VARIABLES IN INTEGRATION.

110.211 HONORS MULTIVARIABLE CALCULUS
PROFESSOR RICHARD BROWN

Synopsis. Here, we focus on the idea of changing the coordinates in an integral. In Calculus I, the Substitution Method was an actual change of coordinates used usually to make the integrand easier to play with. Here, and in more generality, changing the coordinate system on a region is used more to make the region easier to integrate over. Of course, it must be true that the value of the definite integral should be the same no matter the coordinates used. Hence one must be careful to properly account for the change, precisely as in the Substitution Method, where a change of variable creates a new variable corresponding to the "inside function" of the composition of functions in the integrand (this is a function of the old variable). The extra piece was the derivative of the inside function. This generalizes as the 1-dimensional version of a similar phenomena in higher dimensions. We detail this today.

Helpful Documents. PDF: `SphereVolume`.

Parameterization. Placing new coordinates on a space involves again a function: Let $T : X \subset \mathbb{R}^2 \to \mathbb{R}^2$, $T(u, v) = (x(u, v), y(u, v))$ be a C^1-map. Then, any subset $D \subset X$, in the uv-coordinates, is mapped to its image $T(D)$, another region (in the xy-coordinates). See Figure 16.1. To do this, one needs to write the old coordinates as functions of the new coordinates. In functional form, this involves a function whose domain is in the new coordinates and whose codomain uses the old coordinates.

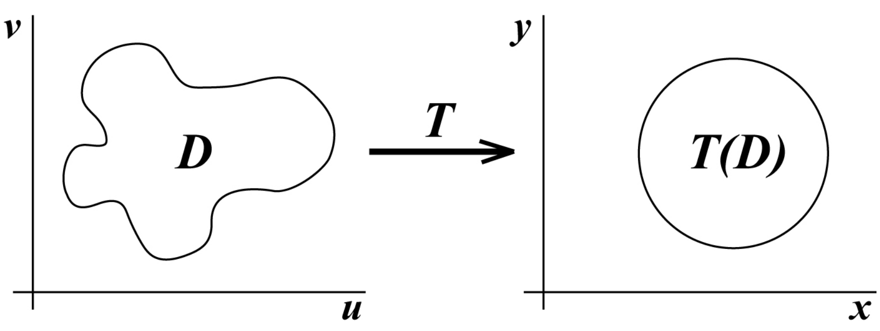

Figure 16.1. A coordinate change in the plane.

Example 16.1. ***Polar Coordinates.*** Working in polar coordinates in the plane involves switching the variables from x and y, to r and θ, representing, respectively, range (Euclidean distance from the origin) and angle (from a reference line in the xy-plane, often the positive x-axis. The resulting coordinate equations are $x = r\cos\theta$ and $y = r\sin\theta$, which takes the domain $D = [0,1] \times [0, 2\pi]$ in the $r\theta$-plane to the unit disk $T(D)$ in the xy-plane. See Figure 16.2 below.

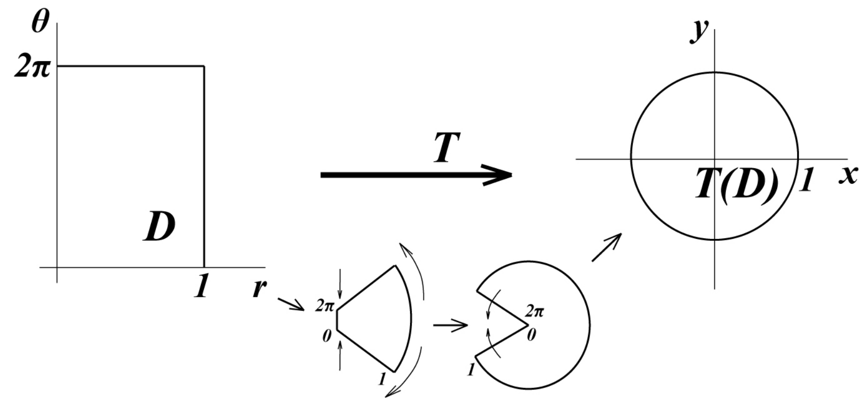

FIGURE 16.2. Here, $T(r, \theta) = (x(r,\theta), y(r,\theta)) = (r\cos\theta, r\sin\theta) = (x, y)$.

Notice here that, in this coordinate change, the edges of the box D under T do not behave so well. For instance, the entire left edge is mapped to the origin (all points $(0, \theta)$ in the $r\theta$-plane represent the origin no matter the value of θ), and the top and bottom edges of D under T are *identified* (mapped point by point to the same image). This represents the fact that $(r, \theta + 2\pi) = (r, \theta)$. However, on the inside of the box D, the map T is injective. This will be vitally important.

Example 16.2. ***Translations.*** In the process of studying the properties of an equilibrium solution in a system of ordinary differential equations, one often changes coordinates by a translation, moving the equilibrium solution (really a point in Euclidean space) to the origin, before analyzing. For example, the coordinate change $u = x - 1$ and $v = y - 2$ is a transformation of the plane that moves the origin down 2 and to the left 1. Then, writing x and y in terms of the new variables u and v, we see that the unit square $= [0,1] \times [0,1]$ in the uv-plane is moved by T to $T(D) = [1,2] \times [2,3]$. Under this transformation, if, for example, one were to integrate a function over $[1,2]$ in x and $[2,3]$ in y, then by changing coordinates, one could instead integrate over the unit square in the uv-plane.

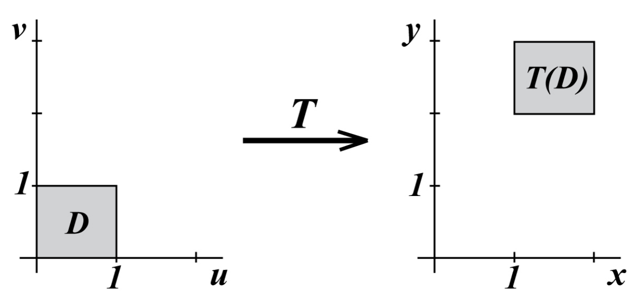

FIGURE 16.3. Here, $T(u, v) = (x(u,v), y(u,v))) = (u + 1, v + 2)$.

Example 16.3. ***A Linear Map of*** $\mathbb{R}^2$**.** Domains which are parallelograms have the quality that a linear transformation can take them to squares (or at least rectabgles). Consider T the linear planar transformation $x = u + v$ and $y = u - v$. As a transformation, T takes the unit square D to a bigger square $T(D)$, though here it is one that is not oriented so that its sides are parallel to the coordinate axes. See Figure 16.4

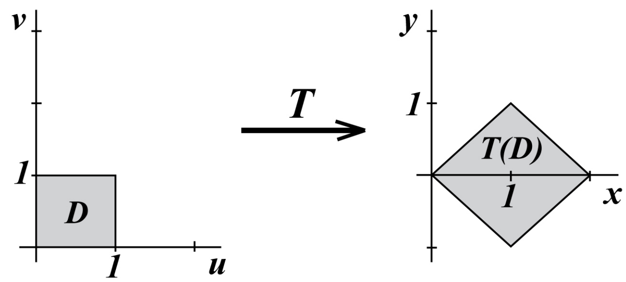

FIGURE 16.4. Here, $T(u, v) = (u(x, y), v(x, y))) = (x + 1, y + 2)$.

Notes:

- In general, it takes practive to "see" a transformation, and to construct one.
- One good practice to well-understand how to transform the boundary first, or at least any vertices.
- Example 16.3 is an example of a transformation that is often easily constructible.

Proposition 16.4. *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, *where* $\boldsymbol{det}\, A = ad - bc \neq 0$. *Then the planar transformation* $T(u, v) = A \begin{bmatrix} u \\ v \end{bmatrix}$ *is one-to-one, onto, takes parallelograms to parallelograms (vertices to vertices), and*

$$\boldsymbol{area}(T(D)) = |\boldsymbol{det}\, A|\, \boldsymbol{area}(D).$$

Some notes:

- As a linear transformation, note that it is always the case that the origin goes to the origin, so no translations here.
- Of course, this proposition generalizes to n-dimensions, for $n \in \mathbb{N}$.
- The proof is really just linear algebra with a bit of geometry thrown in.

So here is the new question: Why would someone interested in integration over a domain want to change the shape and/or size of the domain? The answer is that rectangles are easier to integrate over than more general domains, and often a nonelementary domain can be made elementary via a coordinate change.

Special Note: Notice in the above examples that the more complicated regions are in the codomain of the transformation $T(u, v) = (x, y)$. This is because one writes the given variables x and y in terms of the new variables u and v, so $x = x(u, v)$ and $y = y(u, v)$. Then, by composition, any function $f(x, y)$ can be composed with T to generate a version of f written in terms of u and v:

$$(f \circ T)(u, v) = f(T(u, v)) = f(x(u, v), y(u, v)) = f(u, v).$$

You have seen this before in single variable calculus in the form of the *Substitution Method* for the evaluation of an integral. If $g'(x)$ is continuous on $[a, b]$, and f is continuous on the range of $u = g(x)$, then

$$\int_a^b f\left(g(x)\right) g'(x)\, dx = \int_{g(a)}^{g(b)} f(u)\, du.$$

Here, one changed the variable x to u in order to make the integrans easier to integrate. Part of the simplifying effect of the change in variables was the required term given by the relationship between differentials $du = g'(x)\, dx$. We will also need this in our new multi-dimensional case.

Let $T(u,v) = (x(u,v), y(u,v))$ be a C^1-transformation of $\mathbb{R}^2$. Then

$$DT(u,v) = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}, \quad \text{and}$$

$$\begin{aligned} \mathbf{Jac}(T) = \mathbf{det}(DT) &= \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \\ &= \left| \frac{\partial(x,y)}{\partial(u,v)} \right| = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}. \end{aligned}$$

We have

Theorem 16.5. *Let D and D^* be elementary regions in the xy-plane and uv-plane, respectively, and suppose $T : \mathbb{R}^2 \to \mathbb{R}^2$ is C^1, $D = T(D^*)$, and T is injective on te interior of D^*. Then for any integrable $f : D \to \mathbb{R}$,*

$$\iint_D f(x,y)\, dx\, dy = \iint_{D^*} f\left(x(u,v), y(u,v)\right) \left| \frac{\partial(x,y)}{\partial(u,v)} \right| du\, dv.$$

Here is a fun fact: By this theorem, you actually learned the Substitution Method backwards!

In 1-dimension, the theorem says

$$\int_I f(x)\, dx = \int_{I^*} f(x(u)) x'(u)\, du,$$

where $I = [a,b]$ and $I^* = [u(a), u(b)]$, and the 1-dimensional Jacobian was $x'(u)$ for the coordinate change $T(u) = x(u)$. Some notes:

- You learned it backwards because you prupose back then was to simplify the integrand.
- The purpose in more generality (in multivariable calculus) is to simplfy the integrating region.
- This was never stressed in Calculus I, but even in 1-dimension, the coordinate change to switch to the new variable does need to be 1-1 for the transformation to work correctly. Funny how the examples you worked on always did work out that way...!

New question: So why does the Jacobian arise is the way that it does when changing variables? Here is the answer:

Under the transformation T, a small rectangular region R centered at (u_0, v_0) in D^* given by $\Delta u = u - u_0$ and $\Delta v = v - v_0$ is taken to a new region $R = T(R^*) \subset D$: Here,

aera$(R^*) = \Delta u \Delta v > 0$, and, since T is C^1 and bijective, **area**$(R) > 0$ also. We cannot calculate **area**(R) directly, but we can approximate it: A linear approximation to $T(u,v)$, near (u_0, v_0) is the function

$$h(u,v) = T(u_0,v_0) + DT(u_0,v_0)\begin{bmatrix} u-u_0 \\ v-v_0 \end{bmatrix}.$$

This linear approximation maps $(u_0,v_0) \mapsto T(u_0,v_0) = (x_0,y_0)$, and takes R^* to a parallelogram $h(R^*)$ which is close to $R = T(R^*)$, so that

$$\mathbf{area}(R) \approx \mathbf{area}\left(h(R^*)\right).$$

We will focus on this later in the course, but it is true that the area of a parallelogram in the plane can be computed by using a form of the cross product adapted to vectors in $\mathbb{R}^2$. Indeed, given two vectors $\mathbf{a}$ and $\mathbf{b}$ in $\mathbb{R}^2$, embed them as vectors in $\mathbb{R}^3$ simply by giving each a 0 in the last coordinate. Then, the area of the parallelogram that has these two vectors as sides is given by the quantity $||\mathbf{a} \times \mathbf{b}||$. This is a general feature of the cross product in R^3. So here, then,

$$\begin{aligned} \mathbf{area}(R^*) &= \left|\left| \begin{bmatrix} \Delta u \\ 0 \end{bmatrix} \times \begin{bmatrix} 0 \\ \Delta v \end{bmatrix} \right|\right| = \Delta u \Delta v, \quad \text{and} \\ \mathbf{area}(h(R^*)) &\approx \left|\left| h\left(\begin{bmatrix} \Delta u \\ 0 \end{bmatrix}\right) \times h\left(\begin{bmatrix} 0 \\ \Delta v \end{bmatrix}\right) \right|\right| \\ &= |\mathbf{det}\, DT(u_0,v_0)| \cdot \mathbf{area}(R^*) \\ &= \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \Delta u \Delta v. \end{aligned}$$

Now consider a cahnge of coordinates from polar to Cartesian: $x = r\cos\theta$ and $y = r\sin\theta$. Then

$$dA = \, dx\, dy = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| dr\, d\theta = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} dr\, d\theta = r\, dr\, d\theta.$$

Hence

$$\iint_D f(x,y)\, dx\, dy = \iint_{D^*} f(r\cos\theta, r\sin\theta) r\, dr\, d\theta,$$

where the region D^* is rectangular in the $r\theta$-plane. This is precisely where the extra r in the integrand comes from when converting to polar coordinates.

Note again, here, though, that the transformation T from polar to Cartesian is not 1-1 on D^*. But it is on the interior, and hence is fine. So why is this the case when integrating??

Recall the famous formula from single variable calculus: Given a function $r = f(\theta)$, one can find the area of the region D formed from the origin to $f(\theta)$ between two values θ_1 and θ_2 as

$$\mathbf{area}(D) = \int_{\theta_1}^{\theta_2} \frac{1}{2}\left[f(\theta)\right]^2\, d\theta.$$

So why is this true?

The region D is actually written in the Cartesian plane (the xy-plane), as in the right side of Figure 16.5. Then, under the polar coordinate transformation, D^* looks like the left side.

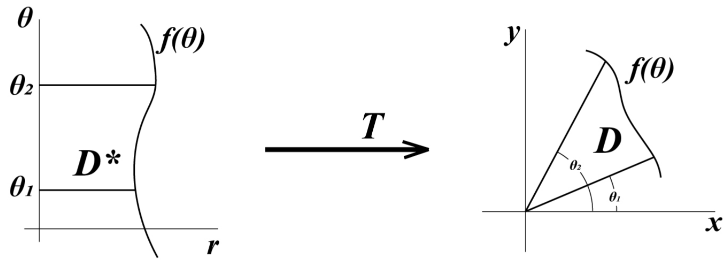

FIGURE 16.5. The classic formula from Calculus I on the area of a sector capped by $r = f(\theta)$.

Note that D^* is elementary in θ. Then the area of D can be calculated:

$$\begin{aligned}
\mathbf{area}(D) &= \iint_D 1\,dx\,dy = \iint_{D^*} r\,dr\,d\theta \\
&= \int_{\theta_1}^{\theta_2} \int_0^{f(\theta)} r\,dr\,d\theta = \int_{\theta_1}^{\theta_2} \left[\frac{r^2}{2}\bigg|_0^{f(\theta)} \right] d\theta \\
&= \int_{\theta_1}^{\theta_2} \frac{1}{2} \left[f(\theta) \right]^2 \,d\theta.
\end{aligned}$$

Ans lastly, a quick comment: ALl of this discussion was in the context of functions defined in a region in the plane. But all of this is readily generalizable to $\mathbb{R}^n$: For example, a C^1-transformation in three space would look like

$$T(u, v, w) = \left(x(u, v, w), y(u, v, w), z(u, v, w) \right),$$

and the resulting Jacobian would look like

$$\mathbf{Jac}(T) = \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}.$$

Example 16.6. ***Cartesian-to-Spherical Coordinate Jacobian.*** The *spherical coordinate system* is one of the natural generalizations of polar coordinates in the plane. One set of equations constructing the transformation is $x = \rho \sin\varphi \cos\theta$, $y = \rho \sin\varphi \sin\theta$, and $z = \rho \cos\varphi$. Given a point $\mathbf{p} \in \mathbb{R}^3$ in this coordinate system, ρ is the distance from $\mathbf{p}$ to the origin in the Euclidean metric, φ is angle the ray from $\mathbf{0}$ to $mathbfp$ makes with respect to the positive z-axis, and θ is the angle between the positive x-axis and the ray formed inside the xy-plane between the origin and the projection of $\mathbf{p}$ into the xy-plane inside $\mathbb{R}^3$. If we were to switch from Cartesian coordinates to spherical coordinates in order to integrate, we would need the Jacobian. We have

$$\left|\frac{\partial(x,y,z)}{\partial(\rho,\varphi,\theta)}\right| d\rho\, d\varphi\, d\theta = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \varphi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \varphi} & \frac{\partial z}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \sin\varphi\cos\theta & \rho\cos\varphi\cos\theta & -\rho\sin\varphi\sin\theta \\ \sin\varphi\sin\theta & \rho\cos\varphi\sin\theta & \rho\sin\varphi\cos\theta \\ \cos\varphi & -\rho\sin\varphi & 0 \end{vmatrix}$$

$$\begin{aligned} &= \cos\varphi \begin{vmatrix} \rho\cos\varphi\cos\theta & -\rho\sin\varphi\sin\theta \\ \rho\cos\varphi\sin\theta & \rho\sin\varphi\cos\theta \end{vmatrix} \\ &\quad - \left(-\rho\sin\varphi \begin{vmatrix} \sin\varphi\cos\theta & -\sin\varphi\sin\theta \\ \rho\sin\varphi\sin\theta & \rho\sin\varphi\cos\theta \end{vmatrix}\right) \\ &= \rho^2\cos\varphi\left(\sin\varphi\cos\varphi\cos^2\theta + \sin\varphi\cos\varphi\sin^2\theta\right) \\ &\quad + \rho^2\sin\varphi\left(\sin^2\varphi\cos^2\theta + \sin^2\varphi\sin^2\theta\right) \\ &= \rho^2\sin\varphi\cos^2\varphi + \rho^2\sin\varphi\sin^2\varphi \\ &= \rho^2\sin\varphi. \end{aligned}$$

Hence the integration becomes

$$\iiint_{W^*} f(x,y,z)\, dx\, dy\, dz = \iiint_W f\left(\rho\sin\varphi\cos\theta, \rho\sin\varphi\sin\theta, \rho\cos\varphi\right)\rho^2\sin\varphi\, d\rho\, d\varphi\, d\theta.$$