

Properties of mappings between subsets of \mathbb{R}^n

Sean Fitzpatrick

June 13, 2018

When we want to discuss a change of variables for a multiple integral, it becomes necessary to consider mappings (that is, functions) $T : D \subset \mathbb{R}^n \rightarrow E \subset \mathbb{R}^n$ used to define one set of variables in terms of another; for example, $(x, y, z) = T(u, v, w)$, where $(x, y, z) \in E$, $(u, v, w) \in D$, and each of the variables x, y and z is viewed as a function of the variables u, v, w . We can think of this as a vector-valued function

$$\langle x, y, z \rangle = T(u, v, w) = \langle x(u, v, w), y(u, v, w), z(u, v, w) \rangle.$$

Similar statements hold for $n = 1$ or $n = 2$ (or for that matter, $n > 3$). Of course, if $n = 1$ we just have $x = T(u)$ and there are no vectors involved. Since we want to discuss properties that apply to all such maps T regardless of dimension, let's let $\mathbf{x} \in \mathbb{R}^n$ represent $x \in \mathbb{R}$, or $(x, y) \in \mathbb{R}^2$, or $(x, y, z) \in \mathbb{R}^3$, depending on the value of n , and let $\mathbf{u} \in \mathbb{R}^n$ play the same role for u , (u, v) or (u, v, w) , respectively.

We will discuss properties such as being one-to-one or onto, differentiability, etc. Usually one-to-one functions are introduced in a single variable calculus course, but in case you haven't seen the definition before, it's included below. In each of the definitions below, we will assume that T is a function whose domain is $D \subset \mathbb{R}^n$ and whose range is contained in $E \subset \mathbb{R}^n$, where D and E are closed and bounded.

Definition 1. We say that T is **one-to-one** if no two distinct elements $\mathbf{u}_1, \mathbf{u}_2 \in D$ get mapped to the same value $\mathbf{x} \in E$. That is, if $T(\mathbf{u}_1) = T(\mathbf{u}_2)$, then $\mathbf{u}_1 = \mathbf{u}_2$.

Note that this is not the same as stating that T is a function, which means that no single element $\mathbf{u} \in D$ gets mapped to two different values $\mathbf{x}_1 \neq \mathbf{x}_2 \in E$. One-to-one functions usually come up in the context of inverse functions: since every $\mathbf{u} \in D$ corresponds to a *unique* value $\mathbf{x} = T(\mathbf{u})$, we can define a function T^{-1} by $T^{-1}(\mathbf{x}) = \mathbf{u}$. However, notice that the range of T might not be all of the set E . The range (or image) of T is defined by

$$\text{range}(T) = T(D) = \{T(\mathbf{u}) | \mathbf{u} \in D\}.$$

That is, $T(D) \subset E$ is the set of all $\mathbf{x} \in E$ such that $\mathbf{x} = T(\mathbf{u})$ for some $\mathbf{u} \in D$. We would like to consider the situation where every point in E corresponds to a unique point in D , and vice-versa. (This is often phrased as the statement that T defines a “one-to-one correspondence” between D and E .) The additional property required is the following:

Definition 2. We say that T is **onto** if $T(D) = E$; that is, if for each $\mathbf{x} \in E$ there exists a $\mathbf{u} \in D$ such that $T(\mathbf{u}) = \mathbf{x}$.

If $T : D \rightarrow E$ is one-to-one and onto, then we can define the inverse mapping $T^{-1} : E \rightarrow D$ according to

$$T^{-1}(\mathbf{x}) = \mathbf{u} \quad \text{if and only if} \quad \mathbf{x} = T(\mathbf{u}).$$

Notice that the onto condition guarantees that the domain of T^{-1} is all of E . When considering a changes of variables for a multiple integral over a region E , we would ideally like to have a one-to-one and onto mapping from D to E to ensure that when we convert to an integral over D , each point in E only gets “counted once”. For example, consider the mapping $T(u, v) = (u^2, v)$ defined on $[-1, 1] \times [0, 1]$. (That is, $x = u^2$ with $-1 \leq u \leq 1$ and $y = v$, with $0 \leq v \leq 1$.) The image of T is the square $[0, 1] \times [0, 1]$, but each point (x, y) corresponds to two points $(\pm\sqrt{x}, \sqrt{y})$ in D , so integrating over D would be the same as integrating over E *twice*!

Next we want to consider integrability and differentiability. Recall that a vector-valued function

$$\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle$$

is continuous if and only if each of the component functions $x(t), y(t), z(t)$ is continuous, and similarly, $\mathbf{r}(t)$ is differentiable if and only if each of the component functions is differentiable, and

$$\mathbf{r}'(t) = \langle x'(t), y'(t), z'(t) \rangle.$$

Similarly, a function $T : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is continuous if and only if each of its components is continuous (as a function of several variables), and the partial derivatives of T can be viewed as the vector-valued functions

$$\begin{aligned} \mathbf{r}_u(u, v, w) &= \frac{\partial T}{\partial u}(u, v, w) = \left\langle \frac{\partial x}{\partial u}(u, v, w), \frac{\partial y}{\partial u}(u, v, w), \frac{\partial z}{\partial u}(u, v, w) \right\rangle, \\ \mathbf{r}_v(u, v, w) &= \frac{\partial T}{\partial v}(u, v, w) = \left\langle \frac{\partial x}{\partial v}(u, v, w), \frac{\partial y}{\partial v}(u, v, w), \frac{\partial z}{\partial v}(u, v, w) \right\rangle, \\ \mathbf{r}_w(u, v, w) &= \frac{\partial T}{\partial w}(u, v, w) = \left\langle \frac{\partial x}{\partial w}(u, v, w), \frac{\partial y}{\partial w}(u, v, w), \frac{\partial z}{\partial w}(u, v, w) \right\rangle, \end{aligned}$$

with similar formulas for $n = 2$. (For $n = 1$ we have only the single derivative $T'(u)$.) If each of the components of each of the partial derivatives is continuous (that is, if the partial derivative of each of the \mathbf{x} variables with respect to each of the \mathbf{u} variables is continuous) we say that T is C^1 , or continuously differentiable.

Remark 3. Recall that a general function $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ is *differentiable* if

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - (D_{\mathbf{a}}f)\mathbf{h}\|}{\|\mathbf{h}\|} = 0,$$

where $D_{\mathbf{a}}f$ is the matrix of partial derivatives of f at \mathbf{a} , and $(D_{\mathbf{a}}f)\mathbf{h}$ denotes matrix multiplication, with \mathbf{h} viewed as a column vector.

If a function $T : D \subset \mathbb{R}^n \rightarrow E \subset \mathbb{R}^n$ is C^1 , then as with real-valued functions, being continuously differentiable implies that T is differentiable (in the sense of the above definition),

and therefore continuous. The derivative of T is then an $n \times n$ matrix. For example, when $n = 2$, if $T(u, v) = (x(u, v), y(u, v))$, we get

$$D_{(u,v)}T = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}.$$

Notice that, while the gradients $\nabla x(u, v), \nabla y(u, v)$ make up the rows of the derivative matrix $D_{\mathbf{a}}T$, the *columns* of $D_{\mathbf{a}}T$ are the partial derivative vectors \mathbf{r}_u and \mathbf{r}_v .

Given our function $T : D \subset \mathbb{R}^n \rightarrow E \subset \mathbb{R}^n$, let us denote by DT the matrix of partial derivatives, as in the remark above. Since the dimension of the domain and range are the same, DT is a square ($n \times n$) matrix, so we can compute its determinant.

Definition 4. For any C^1 function $T : D \subset \mathbb{R}^n \rightarrow E \subset \mathbb{R}^n$, the real-valued function $J_T : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$J_T(\mathbf{u}) = \det(D_{\mathbf{u}}T)$$

is called the **Jacobian** of T . (For $n = 1$ the Jacobian is simply $J_T(u) = T'(u)$.) The Jacobian is also denoted by

$$J_T(u, v) = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| \text{ if } n = 2, \text{ or } J_T(u, v, w) = \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| \text{ if } n = 3.$$

Remark 5. The latter notation is sometimes used for the derivative matrix: if $(x, y, z) = T(u, v, w)$, we write $D_{(u,v,w)}T = \frac{\partial(x, y, z)}{\partial(u, v, w)}$. The notation is intended to be reminiscent of the Leibnitz form of the chain rule. Recall that the general chain rule can be stated in terms of matrix multiplication: given maps $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $S : \mathbb{R}^k \rightarrow \mathbb{R}^n$, the composition $T \circ S$ is defined, and

$$D_{\mathbf{a}}(T \circ S) = D_{S(\mathbf{a})}T \cdot D_{\mathbf{a}}S = \frac{\partial(y_1, \dots, y_m)}{\partial(x_1, \dots, x_n)} \cdot \frac{\partial(x_1, \dots, x_n)}{\partial(u_1, \dots, u_k)}.$$

This notation can also help us remember how area and volume elements transform under a change of variables:

$$dx = \frac{dx}{du} du, \quad dx dy = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv, \quad dx dy dz = \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du dv dw$$

for a change of variables in single, double, and triple integrals, respectively.

The observant reader might notice that the single integral is the only one where we do not take the absolute value of the Jacobian. This is because double and triple integrals do not have a property analogous to the one in a single integral which states that reversing the order of integration introduces a sign change. The sign of the Jacobian has to do with whether or not it reverses *orientation*, and double and triple integrals do not keep track of orientation. To do this, we need the added machinery of vector calculus, or more generally, differential forms.

We can now give the definition of a *transformation*, which is the type of function used to define a change of variables. (Note that our definition below includes all the conditions needed for the change of variables formula, while Stewart uses “transformation” as a synonym for “function” in this context.)

Definition 6. Let D and E be closed, bounded subsets of \mathbb{R}^n . A function $T : D \subset \mathbb{R}^n \rightarrow E \subset \mathbb{R}^n$ is said to be a **transformation** from D onto E if

1. T is C^1 on D ; that is, all the partial derivatives in the matrix DT exist and are continuous.
2. T is onto E .
3. T is one-to-one, except possibly on the boundary of D .
4. $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$, except possibly on the boundary of D .

If we let $d\mathbf{x}$ denote either dx , dA , or dV , depending on whether $n = 1, 2$ or 3 , and doing the same for $d\mathbf{u}$, the change of variables formula for a transformation $T : D \rightarrow E$ can be written as

$$\int_E f(\mathbf{x})d\mathbf{x} = \int_D f(T(\mathbf{u}))|J_T(\mathbf{u})|d\mathbf{u},$$

where the integral sign represents a single, double, or triple integral, depending on the value of n . Note that the properties required for T to be a transformation tell us that every point of E corresponds to a point in D , and integrating over D is the same as integrating over E , once we account for the “stretch factor” of the transformation given by the Jacobian $J_T(\mathbf{u})$. A rigorous proof of the change of variables formula is very difficult, but we will give an argument in class similar to the one we considered for the polar and spherical coordinate transformations which, although not a complete proof, is at least a plausible explanation!

Remark 7. The general inverse function theorem, which is not stated in Stewart’s text, (probably in part because the statement requires defining the matrix DT of partial derivatives and explaining what the inverse of a matrix is), states that if $T : D \rightarrow E$ is one-to-one and onto, then T^{-1} exists, and moreover, if T is C^1 and $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$, then T^{-1} is also a C^1 function, and

$$DT^{-1}(\mathbf{x}) = (DT(T(\mathbf{x})))^{-1},$$

where the -1 on the right-hand side denotes a matrix inverse¹. (Compare this to the result $(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}$ in one variable.) A useful consequence of this is obtained by taking the determinant of both sides of the above equation:

$$J_{T^{-1}}(\mathbf{x}) = \frac{1}{J_T(T^{-1}(\mathbf{x}))}.$$

This result can come in handy in cases where it’s easy to come up with the inverse mapping $\mathbf{u} = T^{-1}(\mathbf{x})$, but hard to solve for \mathbf{x} in terms of \mathbf{u} to obtain T .

¹A basic result from linear algebra tells us that a matrix is invertible if and only if its determinant is non-zero.

We end with a theorem that can be very useful when trying to determine the transformation to use for a change of variables: the boundary of E must correspond to the boundary of D . This is useful because we usually would like D to be as simple as possible, ideally a rectangle (or box, if $n = 3$). Since the sides of the rectangle are given by setting u or v equal to a constant, we look at the curves that define the boundary of E . If the boundary of E can be expressed in terms of level curves for two functions $f(x, y)$ and $g(x, y)$, we can define $u = f(x, y)$ and $v = g(x, y)$, which allows us to define $T^{-1}(x, y) = (f(x, y), g(x, y))$. From there, we can try to compute T from T^{-1} , which is a matter of solving for x and y in terms of u and v .

Theorem 8. *Let $D, E \subset \mathbb{R}^n$ be closed, bounded regions. If $T : D \rightarrow E$ is a transformation, then the boundary of E is the image under T of the boundary of D ; that is, if $T(\mathbf{u}) = \mathbf{x}$ is on the boundary of E , then \mathbf{u} is on the boundary of D .*

We will prove this result in the case that T is one-to-one, with $J_T(\mathbf{u}) \neq 0$, on all of D , including the boundary. Note that if this property fails on some portion on the boundary, this will not affect the integral. For example, if $n = 2$, the boundary of D consists of a finite union of continuous curves, so any portion of the boundary is a continuous curve, and we know that we can neglect the graphs of finitely many continuous curves when carrying out an integral. We begin by first proving a lemma:

Lemma 9. *If $f : A \rightarrow B$ is a continuous, one-to-one, and onto mapping from A to B with continuous inverse $f^{-1} : B \rightarrow A$, then f maps open sets to open sets. That is, if $U \subset A$ is an open subset of A , then the image $f(U) = \{f(\mathbf{u}) \in B \mid \mathbf{u} \in U\}$ is an open subset of B .*

Proof. Let $U \subset A$ be open, and let $\mathbf{x} \in f(U)$. We need to show that there exists some $\delta > 0$ such that $N_\delta(\mathbf{x}) = \{\mathbf{y} \in A \mid \|\mathbf{x} - \mathbf{y}\| < \delta\}$ is a subset of $f(U)$. (By definition, $f(U)$ is open if each element of $f(U)$ has a δ -neighbourhood completely contained in $f(U)$.) Since f is one-to-one and onto, there exists a unique $\mathbf{v} = f^{-1}(\mathbf{x}) \in U$ such that $f(\mathbf{v}) = \mathbf{x}$. (We must have $\mathbf{v} \in U$ since $f(\mathbf{v}) \in f(U)$.) Since U is open, there exists an $\epsilon > 0$ such that $N_\epsilon(\mathbf{v}) \subset U$.

Now, since f^{-1} is continuous, there exists a $\delta > 0$ such that if $\mathbf{y} \in N_\delta(\mathbf{x})$, then $f^{-1}(\mathbf{y}) \in N_\epsilon(\mathbf{v})$. But if $f^{-1}(\mathbf{y}) \in N_\epsilon \subset U$, then $f(f^{-1}(\mathbf{y})) = \mathbf{y} \in f(U)$, by definition of $f(U)$. Thus, $N_\delta(\mathbf{x}) \subset f(U)$, which is what we needed to show. \square

Using the above lemma, we can now give a proof of our theorem:

Proof of Theorem 8. Let $T : D \rightarrow E$ be the given transformation, which is one-to-one and onto, and such that $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$. Since T is one-to-one and onto, we can find an inverse function $T^{-1} : E \rightarrow D$. Since T is C^1 and $J_T(\mathbf{u}) \neq 0$ for all $\mathbf{u} \in D$, the inverse function theorem tells us that T^{-1} must be C^1 on E . Since T and T^{-1} are both C^1 , they are differentiable and therefore continuous.

Now, let $\mathbf{x} \in E$ be a boundary point. We need to show that \mathbf{x} is the image of a boundary point in D . Recall that \mathbf{x} is a boundary point if and only if every neighbourhood of \mathbf{x} contains both points in E and points not in E . Let $\mathbf{u} = T^{-1}(\mathbf{x}) \in D$ be the element of D that is mapped to \mathbf{x} by T . For the sake of contradiction, suppose that \mathbf{u} is not a boundary point of D . Then since $\mathbf{u} \in D$ it must be an interior point of D , and therefore, there exists

some $\delta > 0$ such that $N_\delta(\mathbf{u}) \subset D$. (That is, there is a neighbourhood of \mathbf{u} that is completely contained in D .)

However, since T satisfies the conditions of Lemma 9, we know that T must map open sets to open sets. In particular, since $N_\delta(\mathbf{u})$ is an open subset of D , $T(N_\delta(\mathbf{u}))$ must be an open subset of E . But since $\mathbf{u} \in N_\delta(\mathbf{u})$, we must have $\mathbf{x} = T(\mathbf{u}) \in T(N_\delta(\mathbf{u}))$, and thus $T(N_\delta(\mathbf{u}))$ is an open subset of E that contains \mathbf{x} , which contradicts the fact that \mathbf{x} is a boundary point. Thus, it must be the case that \mathbf{u} is a boundary point of D . \square

Note that since $T^{-1} : E \rightarrow D$ is also a transformation with the same properties as T , the converse to this result is valid as well: if \mathbf{u} belongs to the boundary of D , then $T(\mathbf{u})$ belongs to the boundary of E .