Section 5: The Jacobian matrix and applications.

S1: Motivation

S2: Jacobian matrix $+$ differentiability

S3: The chain rule

S4: Inverse functions

Images from "Thomas' calculus" by Thomas, Wier, Hass & Giordano, 2008, Pearson Education, Inc.

S1: Motivation. Our main aim of this section is to consider "general" functions and to define a general derivative and to look at its properties.

In fact, we have slowly been doing this. We first considered vector–valued functions of one variable $\mathbf{f} : \mathbb{R} \to \mathbb{R}^n$

$$\mathbf{f}(t) = (f_1(t), \ldots, f_n(t))$$

and defined the derivative as

$$\mathbf{f}'(t) = (f_1'(t), \ldots, f_n'(t)).$$

We then considered real–valued functions of two and three variables $f : \mathbb{R}^2 \to \mathbb{R}$, $f : \mathbb{R}^3 \to \mathbb{R}$ and (as we will see later) we may think of the derivatives of these functions, respectively, as

$$\begin{aligned} \nabla f &= (\partial f / \partial x, \partial f / \partial y) \\ \nabla f &= (\partial f / \partial x, \partial f / \partial y, \partial f / \partial z). \end{aligned}$$

There are still more general functions than those two or three types above. If we combine the elements of each, then we can form "vector–valued functions of many variables".

A function $\mathbf{f} : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ $(n > 1)$ is a vector–valued function of m variables.

Example 1

$$\mathbf{f} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y + z \\ xyz \end{pmatrix}$$

defines a function from $\mathbb{R}^3$ to $\mathbb{R}^2$.

When it comes to these vector–valued functions, we should write vectors as column vectors (essentially because matrices act on column vectors), however, we will use both vertical columns and horizontal m–tuple notation. Thus, for example, for the vector $\mathbf{x} \in \mathbb{R}^3$ we will write both

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad \text{or} \quad (x, y, z) \quad \text{(and } x\mathbf{i} + y\mathbf{j} + z\mathbf{k}\text{)}$$

and so we could write $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ as

$$\mathbf{f}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} f_1(x, y, z) \\ f_2(x, y, z) \end{pmatrix} \quad \text{and}$$

$$\mathbf{f}(x, y, z) = (f_1(x, y, z), f_2(x, y, z))$$
$$= f_1(x, y, z)\mathbf{i} + f_2(x, y, z)\mathbf{j}$$

or combinations of columns and m-tuples.

In Example 1, the real–valued functions

$$f_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + y + z \quad \text{and}$$

$$f_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = xyz$$

are called the **co–ordinate** or **component functions** of $\mathbf{f}$, and we may write

$$\mathbf{f} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}.$$

Generally, any $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ is determined by n co–ordinate functions $f_1, \ldots, f_n$ and we write

$$\mathbf{f} = \begin{pmatrix} f_1(x_1, \ldots, x_m) \\ f_2(x_1, \ldots, x_m) \\ \vdots \\ f_n(x_1, \ldots, x_m) \end{pmatrix}. \tag{1}$$

We shall be most interested in the cases where $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ or $\mathbf{f} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ because this is where the most applications occur and because it will prove to be extremely useful in our topic on multiple integration.

For these special cases we can use the following notation

$$
\begin{aligned}
\mathbf{f}(\mathbf{x}) &= \mathbf{f}(x, y) \\
&= (f_1(x, y), f_2(x, y)) \\
&= f_1(x, y)\mathbf{i} + f_2(x, y)\mathbf{j}.
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{f}(\mathbf{x}) &= \mathbf{f}(x, y, z) \\
&= (f_1(x, y, z), f_2(x, y, z), f_3(x, y, z)) \\
&= f_1(x, y, z)\mathbf{i} + f_2(x, y, z)\mathbf{j} + f_3(x, y, z)\mathbf{k}.
\end{aligned}
$$

One way of visualizing $\mathbf{f}$, say, $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is to think of $\mathbf{f}$ as a transformation between co–ordinate planes.

Cartesian xy-plane   Cartesian uv-plane
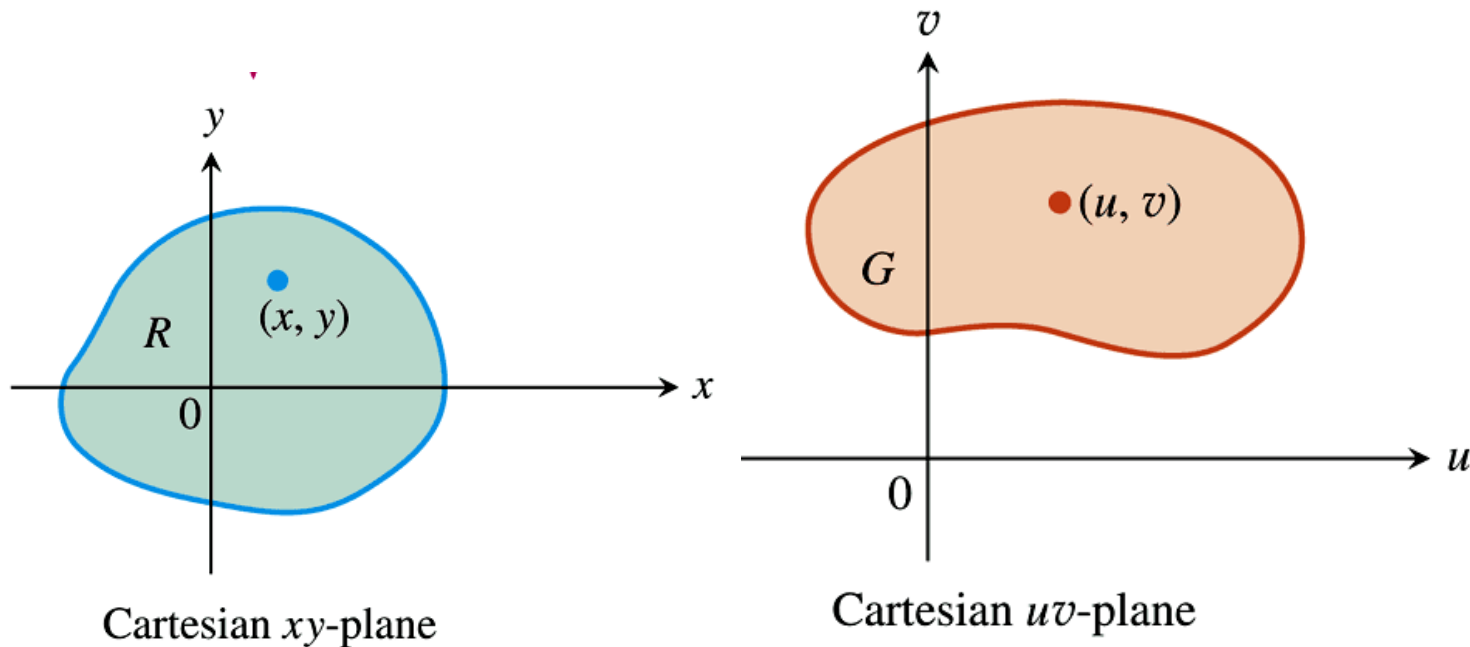
So that $\mathbf{f}$ may stretch, compress, rotate etc sets in its domain.

The above be particularly useful when dealing with multiple integration and change of variables.

S2: Jacobian matrix + differentiability.

Our first problem is how we define the derivative of a vector–valued function of many variables. Recall that if $f : \mathbb{R}^2 \to \mathbb{R}$ then we can form the directional derivative, i.e.,

$$D_{\mathbf{u}} f = u_1 \frac{\partial f}{\partial x} + u_2 \frac{\partial f}{\partial y} = \nabla f \cdot \mathbf{u}$$

where $\mathbf{u} = (u_1, u_2)$ is a unit vector. Thus, knowledge of the gradient of f gives information about all directional derivatives. Therefore it is reasonable to assume

$$\nabla_{\mathbf{p}} f = \left(\frac{\partial f}{\partial x}(\mathbf{p}), \frac{\partial f}{\partial y}(\mathbf{p}) \right)$$

is the derivative of f at $\mathbf{p}$. (The story is more complicated than this but when we say f is "differentiable" we mean ∇f represents the derivative, to be discussed a little later.)

More generally if $f : \mathbb{R}^m \to \mathbb{R}$ then we take the derivative at $\mathbf{p}$ to be the row vector

$$\left(\frac{\partial f}{\partial x_1}(\mathbf{p}), \frac{\partial f}{\partial x_2}(\mathbf{p}), \ldots, \frac{\partial f}{\partial x_m}(\mathbf{p}) \right) = \nabla_{\mathbf{p}} f$$

Now take $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ where $\mathbf{f}$ is as in equation (1), then the natural candidate for the derivative of $\mathbf{f}$ at $\mathbf{p}$ is

$$J_{\mathbf{p}}\mathbf{f} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_m} \\[2ex] \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_m} \\[1ex] \vdots & \vdots & \ddots & \vdots \\[1ex] \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_m} \end{pmatrix}$$

where the partial derivatives are evaluated at $\mathbf{p}$. This $n \times m$ matrix is called the **Jacobian matrix of** $\mathbf{f}$. Writing the function $\mathbf{f}$ as a column helps us to get the rows and columns of the Jacobian matrix the right way round. Note the "Jacobian" is usually the determinant of this matrix when the matrix is square, i.e., when $m = n$.

Example 2 *Find the Jacobian matrix of* $\mathbf{f}$ *from Example 1 and evaluate it at* $(1, 2, 3)$.

Most of the cases we will be looking at have $m = n = $ either 2 or 3. Suppose $u = u(x, y)$ and $v = v(x, y)$. If we define $f : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\mathbf{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix} \equiv \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}$$

then the Jacobian matrix is

$$J\mathbf{f} = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2ex] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix}$$

and the Jacobian (determinant)

$$\det(J\mathbf{f}) = \begin{vmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\[2ex] \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{vmatrix} = \frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}.$$

We often denote $\det(J\mathbf{f})$ by $\dfrac{\partial(u, v)}{\partial(x, y)}$.

Example 3 *Consider the transformation from polar to Cartesian co-ordinates, where*

$$x = r \cos\theta \quad \text{and} \quad y = r \sin\theta.$$

We have

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\[2mm] \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r.$$

We have already noted that if $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ then the Jacobian matrix at each point $\mathbf{a} \in \mathbb{R}^m$ is an $m \times n$ matrix. Such a matrix $J_{\mathbf{a}}\mathbf{f}$ gives us a **linear map** $D_{\mathbf{a}}\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ defined by

$$\left(D_{\mathbf{a}}\mathbf{f}\right)(\mathbf{x}) := J_{\mathbf{a}}\mathbf{f} \cdot \mathbf{x} \quad \text{for all } \mathbf{x} \in \mathbb{R}^n.$$

Note that $\mathbf{x}$ is a column vector.

When we say $\mathbf{f} : \mathbb{R}^m \to \mathbb{R}^n$ is **differentiable** at $\mathbf{q}$ we mean that, the affine function $\mathbf{A}(\mathbf{x}) := \mathbf{f}(\mathbf{q}) + (J_{\mathbf{q}}\mathbf{f}) \cdot (\mathbf{x} - \mathbf{q})$, is a "good" approximation to $\mathbf{f}(\mathbf{x})$ near $\mathbf{x} = \mathbf{q}$ in the sense that

$$\lim_{\mathbf{x} \to \mathbf{q}} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{q}) - (J_{\mathbf{q}}\mathbf{f}) \cdot (\mathbf{x} - \mathbf{q})\|}{\|\mathbf{x} - \mathbf{q}\|} = 0$$

where

$$\|\mathbf{x} - \mathbf{q}\| = \sqrt{(x_1 - q_1)^2 + \ldots + (x_m - q_m)^2}.$$

You should compare this to the one variable case: a function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at a if $\lim\limits_{h \to 0} \dfrac{f(a+h) - f(a)}{h}$ exists, and we call this limit $f'(a)$. But we could equally well say this as $f : \mathbb{R} \to \mathbb{R}$ is differentiable at a if there is a number, written $f'(a)$, for which

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - f'(a) \cdot h|}{|h|} = 0,$$

because a linear map $L : \mathbb{R} \to \mathbb{R}$ can only operate by multiplication with a number.

How do we easily recognize a differentiable function? If all of the component functions of the Jacobian matrix of $\mathbf{f}$ are continuous, then $\mathbf{f}$ is differentiable.

Example 4 *Write the derivative of the function in Example 1 at* $(1, 2, 3)$ *as a linear map.*

Suppose $\mathbf{f}$ and $\mathbf{g}$ are two differentiable functions from $\mathbb{R}^m$ to $\mathbb{R}^n$. It is easy to see that the derivative of $\mathbf{f} + \mathbf{g}$ is the sum of the derivatives of $\mathbf{f}$ and $\mathbf{g}$. We can take the dot product of $\mathbf{f}$ and $\mathbf{g}$ and get a function from $\mathbb{R}^m$ to $\mathbb{R}$, and then differentiate that. The result is a sort of product rule, but I'll leave you to work out what happens. Since we cannot divide vectors, there cannot be a quotient rule, so of the standard differentiation rules, that leaves the chain rule.

S3: The chain rule. Now suppose that $\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^s$ and $\mathbf{f} :$ $\mathbb{R}^s \to \mathbb{R}^n$. We can now form the composition $\mathbf{f} \circ \mathbf{g}$ by mapping with $\mathbf{g}$ first and then following with $\mathbf{f}$:

$$\mathbf{x} \to \mathbf{g}(\mathbf{x}) \to \mathbf{f}(\mathbf{g}(\mathbf{x})) \tag{2}$$

$$(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) := \mathbf{f}(\mathbf{g}(\mathbf{x})) \quad \text{for all } \mathbf{x} \in \mathbb{R}^m.$$

Example 5 *Let* $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^2$ *and* $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^3$ *be defined, respectively, by*

$$\mathbf{g}\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} x + y \\ xy \end{pmatrix} \quad \text{and} \quad \mathbf{f}\begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} \sin x \\ x - y \\ xy \end{pmatrix}.$$

Then $\mathbf{f} \circ \mathbf{g}$ *is defined by*

$$(\mathbf{f} \circ \mathbf{g})\begin{pmatrix} x \\ y \end{pmatrix} = f\left(g\begin{pmatrix} x \\ y \end{pmatrix}\right) = f\begin{pmatrix} x + y \\ xy \end{pmatrix} = \begin{pmatrix} \sin(x + y) \\ x + y - xy \\ (x + y)(xy) \end{pmatrix}.$$

Let $\mathbf{b} = \mathbf{g}(\mathbf{p}) \in \mathbb{R}^s$. If $\mathbf{f}$ and $\mathbf{g}$ in (2) above are differentiable then the maps $J_{\mathbf{p}}\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^s$ and $J_{\mathbf{b}}\mathbf{f} : \mathbb{R}^s \to \mathbb{R}^n$ are defined, and we have the following general result.

Theorem 1 (The Chain Rule) *Suppose that* $\mathbf{g} : \mathbb{R}^m \to \mathbb{R}^s$ *and* $\mathbf{f} : \mathbb{R}^s \to \mathbb{R}^n$ *are differentiable. Then*

$$J_{\mathbf{p}}(\mathbf{f} \circ \mathbf{g}) = J_{\mathbf{g}(\mathbf{p})}\mathbf{f} \cdot J_{\mathbf{p}}\mathbf{g}.$$

This is again just like the one variable case, except now we are multiplying *matrices* (see below).

Example 6 *Consider Example 5:*

$$\mathbf{g}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ xy \end{pmatrix} \quad \text{and} \quad \mathbf{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sin x \\ x-y \\ xy \end{pmatrix}.$$

Find $J_{\mathbf{p}}(\mathbf{f} \circ \mathbf{g})$ *where* $\mathbf{p} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$. *We have*

$$J_{\mathbf{p}}\mathbf{g} = \begin{pmatrix} 1 & 1 \\ y & x \end{pmatrix}_{\mathbf{p}} = \begin{pmatrix} 1 & 1 \\ a_2 & a_1 \end{pmatrix}.$$

Also

$$J_{\mathbf{g(p)}}\mathbf{f} = \begin{pmatrix} \cos x & 0 \\ 1 & -1 \\ y & x \end{pmatrix}_{x=a_1+a_2, y=a_1 a_2}$$

$$= \begin{pmatrix} \cos(a_1+a_2) & 0 \\ 1 & -1 \\ a_1 a_2 & a_1+a_2 \end{pmatrix}.$$

(Ex cont.) and

$$J_{\mathbf{p}}(\mathbf{f} \circ \mathbf{g}) = \begin{pmatrix} \cos(x+y) & \cos(x+y) \\ 1-y & 1-x \\ 2xy + y^2 & x_2 + 2xy \end{pmatrix}_{\mathbf{p}}$$

We observe that

$$\begin{pmatrix} \cos(a_1 + a_2) & \cos(a_1 + a_2) \\ 1 - a_2 & 1 - a_1 \\ 2a_1 a_2 + a_2{}^2 & a_1{}^2 + 2a_1 a_2 \end{pmatrix}$$

$$= \begin{pmatrix} \cos(a_1 + a_2) & 0 \\ 1 & -1 \\ a_1 a_2 & a_1 + a_2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ a_2 & a_1 \end{pmatrix}$$

The one variable chain rule is a special case of the chain rule that we've just met — the same can be said for the chain rules we saw in earlier sections.

Let $x : \mathbb{R} \to \mathbb{R}$ be a differentiable function of t and and let $u : \mathbb{R} \to \mathbb{R}$ a differentiable function of x. Then $(u \circ x) : \mathbb{R} \to \mathbb{R}$ is given by $(u \circ x)(t) = u(x(t))$. In the notation of this chapter

$$J_t(u \circ x) = J_{x(t)}u \cdot J_t x$$

i.e. $\qquad \left[\dfrac{d}{dt}(u \circ x)\right]_t = \left[\dfrac{du}{dx}\right]_{x(t)} \left[\dfrac{dx}{dt}\right]_t .$

We usually write this as

$$\frac{du}{dt} = \frac{du}{dx}\frac{dx}{dt}$$

keeping in mind that when we write $\dfrac{du}{dt}$ we are thinking of u as a function of t, i.e., $u(x(t))$ and when we write $\dfrac{du}{dx}$ we are thinking of u as a function of x.

Now suppose we have $x = x(t), y = y(t)$ and $z = f(x, y)$. Then

$$J_t(f \circ \mathbf{x}) = J_{\mathbf{x}(t)}f \cdot J_t\mathbf{x}$$

Therefore

$$\frac{d}{dt}\left(f(x(t), y(t))\right) = \begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{pmatrix} \cdot \begin{pmatrix} \dfrac{dx}{dt} \\[2ex] \dfrac{dy}{dt} \end{pmatrix}$$

so that

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt},$$

which is just what we saw in earlier sections.

S4: Inverse functions.

In first year (or earlier) you will have met the inverse function theorem, which says essentially that if $f'(a)$ is not zero, then there is a differentiable inverse function f^{-1} defined near $f(a)$ with

$$\left[\frac{d}{dt}(f^{-1})\right]_{f(a)} = \frac{1}{f'(a)}.$$

What happens in the multi–variable case?

Let us consider a case where we can write down the inverse. For polar coordinates we have

$$x = r\cos\theta, \quad y = r\sin\theta$$
$$r = \sqrt{x^2 + y^2}, \quad \theta = \arctan\left(\frac{y}{x}\right).$$

Now differentiating we obtain

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} = \frac{r\cos\theta}{r} = \cos\theta \quad \text{and} \quad \frac{\partial x}{\partial r} = \cos\theta$$

$$\text{i.e.,} \quad \frac{\partial r}{\partial x} \neq \frac{1}{\dfrac{\partial x}{\partial r}}.$$

We see that the one variable inverse function theorem does not apply to partial derivatives. However, there is a simple generalisation if we use the multivariable derivative, that is, the Jacobian matrix.

To continue with the polar coordinate example, define

$$\mathbf{f}\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} x(r,\theta) \\ y(r,\theta) \end{pmatrix} = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \tag{3}$$

and

$$\mathbf{g}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r(x,y) \\ \theta(x,y) \end{pmatrix} = \begin{pmatrix} \sqrt{x^2+y^2} \\ \arctan\left(\frac{y}{x}\right) \end{pmatrix}. \tag{4}$$

Consider

$$(\mathbf{f}\circ\mathbf{g})\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{f}\left(\mathbf{g}\begin{pmatrix} x \\ y \end{pmatrix}\right) = \mathbf{f}\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = Id\begin{pmatrix} x \\ y \end{pmatrix}.$$

Therefore $\mathbf{f}\circ\mathbf{g} = Id$, the identity operator on $\mathbb{R}^2$. Similarly $\mathbf{g}\circ\mathbf{f} = Id$.

Recall

$$Id\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{so that} \quad J(Id) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \equiv 2\times 2 \text{ identity matrix.}$$

Thus by the chain rule

$$Jf \cdot Jg = J(Id) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = Jg \cdot Jf$$

so that $(Jf)^{-1} = Jg$. Note for simplicity the points of evaluation have been left out. Therefore

$$\begin{pmatrix} \dfrac{\partial r}{\partial x} & \dfrac{\partial r}{\partial y} \\[2ex] \dfrac{\partial \theta}{\partial x} & \dfrac{\partial \theta}{\partial y} \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\[2ex] \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{pmatrix}.$$

We can check this directly by substituting $\dfrac{\partial r}{\partial x} = \dfrac{x}{\sqrt{x^2 + y^2}} = \cos \theta$

etc.

The same idea works in general:

Theorem 2 (The Inverse Function Theorem) *Let* $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$
be differentiable at $\mathbf{p}$. *If* $J_{\mathbf{p}}\,\mathbf{f}$ *is an invertible matrix then there is an*
inverse function $\mathbf{f}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ *defined in some neighbourhood of*
$\mathbf{b} = \mathbf{f}(\mathbf{p})$ *and*

$$(J_{\mathbf{b}}\,\mathbf{f}^{-1}) = (J_{\mathbf{p}}\,\mathbf{f})^{-1}.$$

Note that the inverse function may only exist in a small region around
$\mathbf{b} = \mathbf{f}(\mathbf{p})$.

Example 7 *We earlier saw that for polar coordinates, with the notation*
of equation (3)

$$J\mathbf{f} = \begin{pmatrix} \cos\theta & -r\,\sin\theta \\ \sin\theta & r\,\cos\theta \end{pmatrix},$$

with determinant r. *So it follows from the inverse function theorem that*
the inverse function $\mathbf{g}$ *is differentiable if* $r \neq 0$.

Example 8 *The function $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is given by*

$$\mathbf{f}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x^2 - y^2 \\ x^2 + y^2 \end{pmatrix}.$$

Where is $\mathbf{f}$ invertible? Find the Jacobian matrix of $\mathbf{f}^{-1}$ where $\mathbf{f}$ is invertible.

SOLN: $J\mathbf{f} = \begin{pmatrix} 2x & -2y \\ 2x & 2y \end{pmatrix}$ and $\det J\mathbf{f} = 8xy$, so $\mathbf{f}$ is invertible everywhere except the axes.

$$J\mathbf{f}^{-1} = \begin{pmatrix} 2x & -2y \\ 2x & 2y \end{pmatrix}^{-1} = \frac{1}{8xy}\begin{pmatrix} 2y & 2y \\ -2x & -2x \end{pmatrix} = \frac{1}{4}\begin{pmatrix} x^{-1} & x^{-1} \\ -y^{-1} & -y^{-1} \end{pmatrix}.$$

Translate to (u, v) coordinates and this is

$$J\mathbf{f}^{-1} = \frac{\sqrt{2}}{4}\begin{pmatrix} (u+v)^{-1/2} & (u+v)^{-1/2} \\ -(v-u)^{-1/2} & -(v-u)^{-1/2} \end{pmatrix}.$$

Finally let us apply the inverse function theorem to the Jacobian determinants. We recall that

$$\frac{\partial(r, \theta)}{\partial(x, y)} = \det J\mathbf{g} = \begin{vmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{vmatrix} \quad \text{and}$$

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \det J\mathbf{f} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix}.$$

Since $J\mathbf{g}$ and $J\mathbf{f}$ are inverse matrices, their determinants are inverses:

$$\frac{\partial(r, \theta)}{\partial(x, y)} = \frac{1}{\frac{\partial(x,y)}{\partial(r,\theta)}}.$$

This sort of result is true for any change of variable — in any number of dimensions — and will prove very useful in integration.