The Chain Rule

The Problem

You already routinely use the one dimensional chain rule

$$\tfrac{d}{dt} f\big(x(t)\big) = \tfrac{df}{dx}\big(x(t)\big)\tfrac{dx}{dt}(t)$$

in doing computations like

$$\tfrac{d}{dt}\sin(t^2) = \cos(t^2)2t$$

In this example, $f(x) = \sin(x)$ and $x(t) = t^2$.

We now generalize the chain rule to functions of more than one variable. For concreteness, we consider the case in which all functions are functions of two variables. That is, we find the partial derivatives $\frac{\partial g}{\partial s}$ and $\frac{\partial g}{\partial t}$ of a function $g(s,t)$ that is defined as a composition

$$g(s,t) = f\big(x(s,t), y(s,t)\big)$$

We assume that $f(x,y)$, $x(s,t)$ and $y(s,t)$ are all differentiable.

The Solution

Recall that

$$\begin{aligned}\frac{\partial g}{\partial s}(s_0,t_0) &= \lim_{\Delta s\to 0}\frac{g(s_0+\Delta s,t_0) - g(s_0,t_0)}{\Delta s}\\ &= \lim_{\Delta s\to 0}\frac{f\big(x(s_0+\Delta s,t_0), y(s_0+\Delta s,t_0)\big) - f\big(x(s_0,t_0), y(s_0,t_0)\big)}{\Delta s}\end{aligned} \tag{1}$$

To help evaluate this limit, I now introduce some notation, called the "little oh" notation. Roughly speaking, $o(h)$ is used to stand for any function that tends to zero as $h \to 0$ faster than h does. Precisely, writing $G(h) = o(h)$ means that

$$\lim_{h\to 0}\tfrac{G(h)}{h} = 0$$

One particular consequence of "$G(h) = o(h)$" is that, given any constant $C > 0$, we have that $|G(h)| \le C|h|$ for all sufficiently small h. (To see this, choose $\varepsilon = C$ in the δ–ε definition of $\lim_{h\to 0}\frac{G(h)}{h} = 0$. Then there is a $\delta > 0$ such that $\big|\frac{G(h)}{h}\big| < C$ whenever $0 < |h| < \delta$.)

We may rewrite the definition of the derivative $F'(s_0)$ using the little oh notation. The original definition is

$$F'(s_0) = \lim_{\Delta s\to 0}\frac{F(s_0+\Delta s) - F(s_0)}{\Delta s}$$

It is equivalent to

$$\lim_{\Delta s\to 0}\frac{F(s_0+\Delta s) - F(s_0) - F'(s_0)\Delta s}{\Delta s} = 0 \quad\text{or}\quad F(s_0+\Delta s) - F(s_0) - F'(s_0)\Delta s = o(\Delta s)$$

This is also written

$$F(s_0+\Delta s) = F(s_0) + F'(s_0)\Delta s + o(\Delta s) \tag{2}$$

Applying (2) with $F(s) = x(s,t_0)$ and $F(s) = y(s,t_0)$ (with t_0 held fixed),

$$x(s_0+\Delta s,t_0) = x(s_0,t_0) + \frac{\partial x}{\partial s}(s_0,t_0)\Delta s + o(\Delta s) \tag{3}$$

$$y(s_0+\Delta s,t_0) = y(s_0,t_0) + \frac{\partial y}{\partial s}(s_0,t_0)\Delta s + o(\Delta s) \tag{4}$$

Similarly, the definition of "f is differentiable at (x_0,y_0)" is

$$f(x_0+\Delta x, y_0+\Delta y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)\Delta x + \frac{\partial f}{\partial y}(x_0,y_0)\Delta y + o\big(\,|(\Delta x,\Delta y|\,\big) \tag{5}$$

To compute the partial derivative $\frac{\partial g}{\partial s}$ in (1) we use (5) with x_0, y_0, Δx and Δy chosen so that (3) and (4) become

$$x(s_0+\Delta s,t_0) = x_0+\Delta x \qquad y(s_0+\Delta s,t_0) = y_0+\Delta y$$

We choose

$$\begin{aligned} x_0 &= x(s_0,t_0) \quad & \Delta x &= \tfrac{\partial x}{\partial s}(s_0,t_0)\Delta s + o(\Delta s) \\ y_0 &= y(s_0,t_0) \quad & \Delta y &= \tfrac{\partial y}{\partial s}(s_0,t_0)\Delta s + o(\Delta s) \end{aligned}$$

Note that

- here Δx and Δy are implicitly functions of Δs. (We are thinking of s_0 and t_0 as constants.)
- in $\Delta x(\Delta s) = \frac{\partial x}{\partial s}(s_0,t_0)\Delta s + o(\Delta s)$, $o(\Delta s)$ stands for some function $E_1(\Delta s)$ that obeys $\lim\limits_{\Delta s\to 0}\frac{E_1(\Delta s)}{\Delta s} = 0$.
- there is some constant C_1 such that $|\Delta x(\Delta s)| \le C_1|\Delta s|$ for all small Δs. (We may take C_1 to be any number strictly bigger than $\big|\frac{\partial x}{\partial s}(s_0,t_0)\big|$.)
- in $\Delta y(\Delta s) = \frac{\partial y}{\partial s}(s_0,t_0)\Delta s + o(\Delta s)$, $o(\Delta s)$ stands for some function $E_2(\Delta s)$ that obeys $\lim\limits_{\Delta s\to 0}\frac{E_2(\Delta s)}{\Delta s} = 0$. *The two functions $E_1(\Delta s)$ and $E_2(\Delta s)$ need not be the same.*
- there is some constant C_2 such that $|\Delta y(\Delta s)| \le C_2|\Delta s|$ for all small Δs. (We may take C_2 to be any number strictly bigger than $\big|\frac{\partial y}{\partial s}(s_0,t_0)\big|$.)
- there is some constant C_3 such that $|(\Delta x,\Delta y)| \le C_3|\Delta s|$ for all small Δs. (For example, we could take $C_3 = \sqrt{C_1^2+C_2^2}$.) So $o\big(\,|(\Delta x,\Delta y|\,\big) = o(\Delta s)$.

Now let's return to the computation of the partial derivative $\frac{\partial g}{\partial s}$ in (1). Using the notation above

$$\begin{aligned} g(s_0+\Delta s,t_0) - g(s_0,t_0) &= f\big(x(s_0+\Delta s,t_0), y(s_0+\Delta s,t_0)\big) - f\big(x(s_0,t_0),y(s_0,t_0)\big) \\ &= f\big(x_0+\Delta x, y_0+\Delta y\big) - f\big(x_0,y_0\big) \\ &= \tfrac{\partial f}{\partial x}\big(x_0,y_0\big)\Delta x + \tfrac{\partial f}{\partial y}\big(x_0,y_0\big)\Delta y + o\big(\,|(\Delta x,\Delta y|\,\big) \\ &= \tfrac{\partial f}{\partial x}\big(x(s_0,t_0),y(s_0,t_0)\big)\tfrac{\partial x}{\partial s}(s_0,t_0)\Delta s + \tfrac{\partial f}{\partial y}\big(x(s_0,t_0),y(s_0,t_0)\big)\tfrac{\partial y}{\partial s}(s_0,t_0)\Delta s + o(\Delta s) \end{aligned}$$

Dividing by Δs and taking the limit, we get the upper formula in

$$\begin{aligned} \frac{\partial g}{\partial s}(s_0,t_0) &= \frac{\partial f}{\partial x}\big(x(s_0,t_0),y(s_0,t_0)\big)\frac{\partial x}{\partial s}(s_0,t_0) + \frac{\partial f}{\partial y}\big(x(s_0,t_0),y(s_0,t_0)\big)\frac{\partial y}{\partial s}(s_0,t_0) \\ \frac{\partial g}{\partial t}(s_0,t_0) &= \frac{\partial f}{\partial x}\big(x(s_0,t_0),y(s_0,t_0)\big)\frac{\partial x}{\partial t}(s_0,t_0) + \frac{\partial f}{\partial y}\big(x(s_0,t_0),y(s_0,t_0)\big)\frac{\partial y}{\partial t}(s_0,t_0) \end{aligned}$$

The lower formula is derived similarly. This is true for all evaluation points (s_0, t_0), so

$$\boxed{\begin{aligned}\frac{\partial g}{\partial s}(s,t) &= \frac{\partial f}{\partial x}\big(x(s,t),y(s,t)\big)\,\frac{\partial x}{\partial s}(s,t) + \frac{\partial f}{\partial y}\big(x(s,t),y(s,t)\big)\,\frac{\partial y}{\partial s}(s,t)\\ \frac{\partial g}{\partial t}(s,t) &= \frac{\partial f}{\partial x}\big(x(s,t),y(s,t)\big)\,\frac{\partial x}{\partial t}(s,t) + \frac{\partial f}{\partial y}\big(x(s,t),y(s,t)\big)\,\frac{\partial y}{\partial t}(s,t)\end{aligned}}$$

The Memory Aid

To help remember these formulae, it is useful to pretend that our variables are physical quantities with f, g having units of grams, x, y having units of meters and s, t having units of seconds. Note that

a) The function g appears once in the numerator on the left. The function f, which is gotten from g by a change of variables, appears once in the numerator of each term on the right.

b) The variable in the denominator on the left appears once in the denominator of each term on the right.

c) f is a function of two independent variables, x and y. There are two terms on the right, one for each independent variable of f. Each term contains the partial derivative of f with respect to a different independent variable. That independent variable appears once in the denominator and once in the numerator, so that its units cancel out. Thus all terms on the right hand side have the same units as that on the left hand side. Namely, grams per second.

d) The left hand side is a function of s and t. Hence the right hand side must also be a function of s and t. As f is a function of x and y this is achieved by evaluating f at $x = x(s,t)$ and $y = y(s,t)$.

I recommend strongly that you use the following procedure, without leaving out any steps, the first couple of dozen times that you use the chain rule.

Step 1 List **explicitly** all the functions involved and what each is a function of. Ensure that all different functions have different names. Invent new names for some of the functions if necessary. In the example on the previous page, the list would be

$$f(x,y) \qquad x(s,t) \qquad y(s,t) \qquad g(s,t) = f\big(x(s,t),y(s,t)\big)$$

While the functions f and g are closely related, they are not the same. One is a function of x and y while the other is a function of s and t.

Step 2 Write down the template

$$\frac{\partial g}{\partial s} = \frac{\partial f}{}\frac{}{\partial s}$$

Note that the template satisfies a) and b) above.

Step 3 Fill in the blanks with everything that makes sense. In particular, since f is a function of x and y, it may only be differentiated with respect to x and y.

$$\frac{\partial g}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

Note that x and y are functions of s so that the derivatives $\frac{\partial x}{\partial s}$ and $\frac{\partial y}{\partial s}$ make sense. Also note that the units work out right. See c) above.

Step 4 Put in the functional dependence **explicitly**. Fortunately, there is only one functional dependence that makes sense. See d) above.

$$\frac{\partial g}{\partial s}(s,t) = \frac{\partial f}{\partial x}\big(x(s,t),y(s,t)\big)\frac{\partial x}{\partial s}(s,t) + \frac{\partial f}{\partial y}\big(x(s,t),y(s,t)\big)\frac{\partial y}{\partial s}(s,t)$$

Example 1: Find $\frac{d}{dt}f\big(x(t),y(t)\big)$, for $f(x,y) = x^2 - y^2$, $x(t) = \cos(t)$ and $y(t) = \sin(t)$.

Define $g(t) = f\big(x(t),y(t)\big)$. The appropriate chain rule for this example is

$$\frac{dg}{dt}(t) = \frac{\partial f}{\partial x}\big(x(t),y(t)\big)\frac{dx}{dt}(t) + \frac{\partial f}{\partial y}\big(x(t),y(t)\big)\frac{dy}{dt}(t)$$

For the given functions

$$\begin{aligned}
&f(x,y) = x^2 - y^2 \\
&\tfrac{\partial f}{\partial x}(x,y) = 2x \qquad \tfrac{\partial f}{\partial x}(x(t),y(t)) = 2x(t) \quad = 2\cos t \\
&\tfrac{\partial f}{\partial y}(x,y) = -2y \qquad \tfrac{\partial f}{\partial y}(x(t),y(t)) = -2y(t) = -2\sin t
\end{aligned}$$

so that

$$\frac{dg}{dt}(t) = (2\cos t)(-\sin t) + (-2\sin t)(\cos t) = -4\sin t\cos t$$

Of course, in this example we can compute $g(t)$ explicitly

$$g(t) = f\big(x(t),y(t)\big) = x(t)^2 - y(t)^2 = \cos^2 t - \sin^2 t$$

and then differentiate

$$g'(t) = 2(\cos t)(-\sin t) - 2(\sin t)(\cos t) = -4\sin t\cos t$$

Example 2: Find $\frac{\partial}{\partial t}f(x+ct)$.

Define $u(x,t) = x + ct$ and $w(x,t) = f(x+ct) = f\big(u(x,t)\big)$. Then

$$\tfrac{\partial}{\partial t}f(x+ct) = \tfrac{\partial w}{\partial t}(x,t) = \tfrac{df}{du}\big(u(x,t)\big)\tfrac{\partial u}{\partial t}(x,t) = cf'(x+ct)$$

Example 3: Find $\frac{\partial^2}{\partial t^2}f(x+ct)$.

Define $W(x,t) = \frac{\partial w}{\partial t}(x,t) = cf'(x+ct) = F\big(u(x,t)\big)$ where $F(u) = cf'(u)$. Then

$$\tfrac{\partial^2}{\partial t^2}f(x+ct) = \tfrac{\partial W}{\partial t}(x,t) = \tfrac{dF}{du}\big(u(x,t)\big)\tfrac{\partial u}{\partial t}(x,t) = cf''(x+ct)c = c^2f''(x+ct)$$

Example 4: Suppose that $F(P,V,T) = 0$. Find $\frac{\partial P}{\partial T}$.

Before we can solve this problem, we first have to decide what it means. This happens regularly in applications. In fact, this particular problem comes from thermodynamics. The variables $P,\ V,\ T$ are the pressure, volume and temperature, respectively, of some gas. These three variables are not independent. They are related by an equation of state, here denoted $F(P,V,T)=0$. Given values for any two of $P,\ V,\ T$, the third can be found by solving $F(P,V,T)=0$. We are being asked to find $\frac{\partial P}{\partial T}$. This implicitly instructs us to treat P, in this problem, as the dependent variable. So a careful wording of this problem (which you will never encounter in the "real world") would be the following. The function $P(V,T)$ is defined by $F\big(P(V,T),V,T\big)=0$. Find $\big(\frac{\partial P}{\partial T}\big)_V$, that is the rate of change of pressure as the temperature is varied, while holding the volume fixed.

Since we are not told explicitly what F is, we cannot solve explicitly for $P(V,T)$. So, instead we differentiate both sides of

$$F\big(P(V,T),V,T\big)=0$$

with respect to T, while holding V fixed. Think of the left hand side, $F\big(P(V,T),V,T\big)$, as being $F\big(P(V,T),Q(V,T),R(V,T)\big)$ with $Q(V,T)=V$ and $R(V,T)=T$. By the chain rule,

$$\tfrac{\partial}{\partial T}F\big(P(V,T),Q(V,T),R(V,T)\big)=F_1\tfrac{\partial P}{\partial T}+F_2\tfrac{\partial Q}{\partial T}+F_3\tfrac{\partial R}{\partial T}=0$$

with F_j referring to the partial derivative of F with respect to its $j^{\rm th}$ argument. Experienced chain rule users never introduce Q and R. Instead, they just write

$$\tfrac{\partial F}{\partial P}\tfrac{\partial P}{\partial T}+\tfrac{\partial F}{\partial V}\tfrac{\partial V}{\partial T}+\tfrac{\partial F}{\partial T}\tfrac{\partial T}{\partial T}=0$$

Recalling that V and T are the independent variables and that, in computing $\frac{\partial}{\partial T}$, V is to be treated as a constant,

$$\tfrac{\partial V}{\partial T}=0 \qquad\qquad \tfrac{\partial T}{\partial T}=1$$

Now putting in the functional dependence

$$\tfrac{\partial F}{\partial P}\big(P(V,T),V,T\big)\tfrac{\partial P}{\partial T}(V,T)+\tfrac{\partial F}{\partial T}\big(P(V,T),V,T\big)=0$$

and solving

$$\tfrac{\partial P}{\partial T}(V,T)=-\frac{\frac{\partial F}{\partial T}\big(P(V,T),V,T\big)}{\frac{\partial F}{\partial P}\big(P(V,T),V,T\big)}$$

Example 5: Suppose that $f(x,y)=0$. Find $\frac{d^2y}{dx^2}$.

Once again, x and y are not independent variables. Given a value for either x or y, the other is determined by solving $f(x,y)=0$. Since we are asked to find $\frac{d^2y}{dx^2}$, it is y that is to be viewed as a function of x, rather than the other way around. So $f(x,y)=0$ really means, in this problem, $f\big(x,y(x)\big)=0$ for all x. Differentiating both sides of this equation with respect to x,

$$\begin{aligned} &f\big(x,y(x)\big)=0 \qquad \text{for all } x\\ \Longrightarrow\ &\tfrac{d}{dx}f\big(x,y(x)\big)=0 \end{aligned}$$

Note that $\frac{d}{dx}f\big(x,y(x)\big)$ is not the same as $f_x\big(x,y(x)\big)$. The former is, by definition, the rate of change with respect to x of $g(x)=f\big(x,y(x)\big)$. Precisely,

$$\begin{aligned}\frac{dg}{dx}&=\lim_{\Delta x\to 0}\frac{g(x+\Delta x)-g(x)}{\Delta x}\\&=\lim_{\Delta x\to 0}\frac{f\big(x+\Delta x,y(x+\Delta x)\big)-f\big(x,y(x)\big)}{\Delta x}\end{aligned}\tag{1}$$

On the other hand

$$\begin{aligned}f_x(x,y)&=\lim_{\Delta x\to 0}\frac{f(x+\Delta x,y)-f(x,y)}{\Delta x}\\\implies f_x\big(x,y(x)\big)&=\lim_{\Delta x\to 0}\frac{f\big(x+\Delta x,y(x)\big)-f\big(x,y(x)\big)}{\Delta x}\end{aligned}\tag{2}$$

The two right hand sides (1) and (2) are not the same. In (1), as Δx varies the value of y that is substituted into the first $f(\cdots)$, namely $y(x+\Delta x)$ varies. That is, we are computing the rate of change of f along the (curved) path $y=y(x)$. In (2), the corresponding value of y is $y(x)$ and is independent of Δx. That is, we are computing the rate of change of f along a horizontal straight line. As a concrete example, suppose that $f(x,y)=x+y$. Then, $y(x)=-x$ so that

$$\tfrac{d}{dx}f\big(x,y(x)\big)=\tfrac{d}{dx}f(x,-x)=\tfrac{d}{dx}[x+(-x)]=\tfrac{d}{dx}0=0$$

But $f(x,y)=x+y$ implies that $f_x(x,y)=1$ for all x and y so that

$$f_x(x,y(x))=f_x(x,y)\big|_{y=-x}=1\big|_{y=-x}=1$$

Now back to

$$\begin{aligned}&f\big(x,y(x)\big)=0\qquad\text{for all }x\\\implies&\tfrac{d}{dx}f\big(x,y(x)\big)=0\\\implies&f_x\big(x,y(x)\big)\tfrac{dx}{dx}+f_y\big(x,y(x)\big)\tfrac{dy}{dx}(x)=0\qquad\text{by the chain rule}\\\implies&\frac{dy}{dx}(x)=-\frac{f_x\big(x,y(x)\big)}{f_y\big(x,y(x)\big)}\\\implies&\frac{d^2y}{dx^2}(x)=-\frac{d}{dx}\Big[\frac{f_x\big(x,y(x)\big)}{f_y\big(x,y(x)\big)}\Big]\\&\qquad=-\frac{f_y\big(x,y(x)\big)\frac{d}{dx}[f_x\big(x,y(x)\big)]-f_x\big(x,y(x)\big)\frac{d}{dx}[f_y\big(x,y(x)\big)]}{f_y\big(x,y(x)\big)^2}\end{aligned}\tag{3}$$

by the quotient rule. Now it suffices to substitute in $\frac{d}{dx}[f_x\big(x,y(x)\big)]$ and $\frac{d}{dx}[f_y\big(x,y(x)\big)]$. For the former apply the chain rule to $h(x)=u\big(x,y(x)\big)$ with $u(x,y)=f_x\big(x,y\big)$.

$$\begin{aligned}\tfrac{d}{dx}[f_x\big(x,y(x)\big)]&=\tfrac{dh}{dx}(x)\\&=u_x\big(x,y(x)\big)\tfrac{dx}{dx}+u_y\big(x,y(x)\big)\tfrac{dy}{dx}(x)\\&=f_{xx}\big(x,y(x)\big)\tfrac{dx}{dx}+f_{xy}\big(x,y(x)\big)\tfrac{dy}{dx}(x)\\&=f_{xx}\big(x,y(x)\big)-f_{xy}\big(x,y(x)\big)\Big[\frac{f_x\big(x,y(x)\big)}{f_y\big(x,y(x)\big)}\Big]\end{aligned}$$

Substituting this and

$$\begin{aligned}\tfrac{d}{dx}[f_y(x,y(x))] &= f_{yx}(x,y(x))\tfrac{dx}{dx} + f_{yy}(x,y(x))\tfrac{dy}{dx}(x)\\ &= f_{yx}(x,y(x)) - f_{yy}(x,y(x))\Big[\frac{f_x(x,y(x))}{f_y(x,y(x))}\Big]\end{aligned}$$

into the right hand side of (3) gives the final answer.

Example 6: Find the gradient of a function given in polar coordinates.

Once again, figuring out what the question means is half the battle. The gradient is a quantity that frequently appears in applications (e.g. the temperature gradient). By definition, the gradient of a function $g(x,y)$ is the vector $\big(g_x(x,y), g_y(x,y)\big)$. In this question we are told that we are given some function $f(r,\theta)$ of the polar coordinates r and θ. We are supposed to convert this function to Cartesian coordinates. This means that we are to consider the function

$$g(x,y) = f\big(r(x,y), \theta(x,y)\big)$$

with

$$\begin{aligned}r(x,y) &= \sqrt{x^2+y^2}\\ \theta(x,y) &= \arctan\tfrac{y}{x}\end{aligned}$$

Then we are to compute the gradient of $g(x,y)$ and express the answer in terms of r and θ. By the chain rule

$$\begin{aligned}\frac{\partial g}{\partial x} &= \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x}\\ &= \frac{\partial f}{\partial r}\frac{1}{2}\frac{2x}{\sqrt{x^2+y^2}} + \frac{\partial f}{\partial \theta}\frac{-y/x^2}{1+(y/x)^2}\\ &= \frac{\partial f}{\partial r}\frac{x}{\sqrt{x^2+y^2}} - \frac{\partial f}{\partial \theta}\frac{y}{x^2+y^2}\\ &= \frac{\partial f}{\partial r}\frac{r\cos\theta}{r} - \frac{\partial f}{\partial \theta}\frac{r\sin\theta}{r^2}\\ &= \frac{\partial f}{\partial r}\cos\theta - \frac{\partial f}{\partial \theta}\frac{\sin\theta}{r}\end{aligned}$$

Similarly

$$\begin{aligned}\frac{\partial g}{\partial y} &= \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial y}\\ &= \frac{\partial f}{\partial r}\frac{1}{2}\frac{2y}{\sqrt{x^2+y^2}} + \frac{\partial f}{\partial \theta}\frac{1/x}{1+(y/x)^2}\\ &= \frac{\partial f}{\partial r}\frac{y}{\sqrt{x^2+y^2}} - \frac{\partial f}{\partial \theta}\frac{x}{x^2+y^2}\\ &= \frac{\partial f}{\partial r}\sin\theta + \frac{\partial f}{\partial \theta}\frac{\cos\theta}{r}\end{aligned}$$

So

$$(g_x, g_y) = f_r\ (\cos\theta, \sin\theta) + \tfrac{1}{r} f_\theta\ (-\sin\theta, \cos\theta)$$

or, with all the arguments put in

$$\begin{aligned}\big(g_x(x,y),g_y(x,y)\big) =& f_r\big(r(x,y),\theta(x,y)\big)\ \big(\cos\theta(x,y),\sin\theta(x,y)\big)\\ &+\tfrac{1}{r(x,y)} f_\theta\big(r(x,y),\theta(x,y)\big)\ \big(-\sin\theta(x,y),\cos\theta(x,y)\big)\end{aligned}$$