THE CHAIN RULE AND POLAR COORDINATES

In this note we write out some of the chain rule calculations we did, in the lecture of
February 16, for the transformations between Cartesian and polar coordinates. We also
give a motivating example.

Let us start with the usual expressions for \( x \) and \( y \) as functions of \( r \) and \( \theta \), and from
these calculate the partial derivatives \( \frac{\partial x}{\partial r} \), \( \frac{\partial x}{\partial \theta} \), \( \frac{\partial y}{\partial r} \), and \( \frac{\partial y}{\partial \theta} \):

\[
x = x(r, \theta) = r \cos \theta, \quad \frac{\partial x}{\partial r} = \cos \theta, \quad \frac{\partial x}{\partial \theta} = -r \sin \theta, \\
y = y(r, \theta) = r \sin \theta, \quad \frac{\partial y}{\partial r} = \sin \theta, \quad \frac{\partial y}{\partial \theta} = r \cos \theta.
\]

(1)

It is important that in (1) we are treating \( r \) and \( \theta \) as the independent variables, so that par-
tial derivatives with respect to \( r \) are taken while holding \( \theta \) constant, and partial derivatives
with respect to \( \theta \) are taken while holding \( r \) constant.

On the other hand, we can think of \( r \) and \( \theta \) as functions of \( x \) and \( y \), leading to:

\[
r = r(x, y) = \sqrt{x^2 + y^2}, \quad \frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} = \cos \theta, \quad \frac{\partial r}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}} = \sin \theta, \\
\theta = \theta(x, y) = \tan^{-1} \frac{y}{x}, \quad \frac{\partial \theta}{\partial x} = \frac{1}{1 + (y/x)^2} \left(-\frac{y}{x^2}\right) = -\frac{y}{x^2 + y^2} = -\frac{\sin \theta}{r}, \quad \frac{\partial \theta}{\partial y} = \frac{1}{1 + (y/x)^2} \left(\frac{1}{x}\right) = \frac{x}{x^2 + y^2} = \frac{\cos \theta}{r}.
\]

(2)

In (2), partial derivatives with respect to \( x \) are taken at constant \( y \), and partial derivatives
with respect to \( y \) are taken at constant \( x \). Notice that we have written these partial
derivatives both in terms of \( x \) and \( y \) and in terms of \( r \) and \( \theta \); the first is in some sense
more natural here, since \( x \) and \( y \) are the fundamental variables, but the second will be
useful below. We could have done the same thing in (1). In fact, it is best not to think
of either \( x \) and \( y \) as fundamental, or of \( r \) and \( \theta \) as fundamental, but to go back and forth
freely from one set to the other. However, it would be confusing to write formulas mixing
\( x \) or \( y \) with \( r \) or with \( \theta \).

There is a surprise buried in (1)–(2): notice that

\[
\frac{\partial x}{\partial r} = \cos \theta = \frac{\partial r}{\partial x}.
\]

(3)

Of course we know that for ordinary derivatives of functions of one variable the rule is,
for example, \( \frac{du}{dt} = (\frac{dt}{du})^{-1} \), but the analogous rule for partial derivatives clearly does not
hold. The reason is that different variables are being held constant in the calculation of
the two partial derivatives in (3): on the left side \( \theta \), on the right side \( y \). The equality in
(3) is just an accident; in general there is no \textit{a priori} relation between \( \frac{\partial u}{\partial t} \) and \( \frac{\partial t}{\partial u} \).
Now we can use (1) and/or (2), with the chain rule, to express partial derivatives of some function \( f \) with respect to \( x \) and \( y \) in terms of partial derivatives with respect to \( r \) and \( \theta \), or vice versa. This is important in many problems in physics, engineering, and mathematics; we give an example below. For the moment let us just suppose that we start with some function of the polar variables, \( f(r, \theta) \). We can then use (2) to express \( r \) and \( \theta \) in terms of the Cartesian variables and thus introduce the composite function \( F(x, y) = f(r(x, y), \theta(x, y)) \). We want to express the partial derivatives of \( F \) with respect to \( x \) and \( y \) in terms of the partial derivatives of \( f \) with respect to \( r \) and \( \theta \). We use the chain rule and (2) to find

\[
\frac{\partial F}{\partial x} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial x} = \frac{\partial f}{\partial r} \cos \theta + \frac{\partial f}{\partial \theta} \left( -\frac{\sin \theta}{r} \right) = \cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta},
\]

(4a)

\[
\frac{\partial F}{\partial y} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial y} = \frac{\partial f}{\partial r} \sin \theta + \frac{\partial f}{\partial \theta} \left( \frac{\cos \theta}{r} \right) = \sin \theta \frac{\partial f}{\partial r} + \frac{\cos \theta}{r} \frac{\partial f}{\partial \theta}.
\]

(4b)

For completeness let us write out (4a) a little more carefully, paying attention to the arguments of the functions:

\[
\left( \frac{\partial F}{\partial x} \right)_{(x=r \cos \theta, y=r \sin \theta)} = \cos \theta \left( \frac{\partial f}{\partial r} \right)_{(r, \theta)} - \frac{\sin \theta}{r} \left( \frac{\partial f}{\partial \theta} \right)_{(r, \theta)}.
\]

We could of course do the same for (4b).

With care we can go on and calculate higher order derivatives; we will demonstrate the method by expressing \( \frac{\partial^2 F}{\partial x^2} \) in terms of \( r \) and \( \theta \) derivatives. To save space we will use subscript notation for partial derivatives: \( \frac{\partial^2 F}{\partial x^2} = F_{xx} \), etc. Then:

\[
F_{xx} = \frac{\partial F_x}{\partial x} = \frac{\partial}{\partial x} \left( \frac{\partial F}{\partial x} \right) = \frac{\partial}{\partial r} \left( \frac{\partial F}{\partial x} \right) \frac{\partial r}{\partial x} + \frac{\partial}{\partial \theta} \left( \frac{\partial F}{\partial x} \right) \frac{\partial \theta}{\partial x}.
\]

(5)

Now from (4a) we have

\[
\frac{\partial}{\partial r} \left( \frac{\partial F}{\partial x} \right) = \frac{\partial}{\partial r} \left( \cos \theta f_r - \frac{\sin \theta}{r} f_\theta \right) = \cos \theta f_{rr} + \frac{\sin \theta}{r^2} f_\theta - \frac{\sin \theta}{r} f_{r\theta},
\]

(6)

where we have used the product rule and remembered that \( r \)-derivatives are taken with constant \( \theta \). Similarly,

\[
\frac{\partial}{\partial \theta} \left( \frac{\partial F}{\partial x} \right) = \frac{\partial}{\partial \theta} \left( \cos \theta f_r - \frac{\sin \theta}{r} f_\theta \right) = -\sin \theta f_r + \cos \theta f_{r\theta} - \frac{\cos \theta}{r} f_\theta - \frac{\sin \theta}{r} f_{\theta\theta}.
\]

(7)

We now put (6) and (7), together with the formulas for \( \frac{\partial r}{\partial x} \) and \( \frac{\partial \theta}{\partial x} \) from (2), into (5), to obtain

\[
F_{xx} = \left( \cos \theta f_{rr} + \frac{\sin \theta}{r^2} f_\theta - \frac{\sin \theta}{r} f_{r\theta} \right) \cos \theta
\]

\[
+ \left( -\sin \theta f_r + \cos \theta f_{r\theta} - \frac{\cos \theta}{r} f_\theta - \frac{\sin \theta}{r} f_{\theta\theta} \right) \left( -\frac{\sin \theta}{r} \right)
\]

\[
= \cos^2 \theta f_{rr} - \frac{2 \sin \theta \cos \theta}{r} f_{r\theta} + \frac{\sin^2 \theta}{r^2} f_{\theta\theta} + \frac{\sin^2 \theta}{r} f_r + \frac{2 \sin \theta \cos \theta}{r^2} f_\theta.
\]

(8)
We have used Clairaut’s Theorem: \( f_{r\theta} = f_{\theta r} \).

By an almost identical calculation we obtain

\[
F_{yy} = \sin^2 \theta f_{rr} + \frac{2 \sin \theta \cos \theta}{r} f_{r\theta} + \frac{\cos^2 \theta}{r^2} f_{\theta\theta} + \frac{\sin^2 \theta}{r} f_r - \frac{2 \sin \theta \cos \theta}{r^2} f_\theta,
\]

(9)

and adding (8) and (9) gives

\[
F_{xx} + F_{yy} = f_{rr} + \frac{1}{r^2} f_{\theta\theta} + \frac{1}{r} f_r.
\]

(10)

The quantity \( F_{xx} + F_{yy} \) is called the \text{Laplacian} of \( F \) and plays an important role in various physical problems. We illustrate with an example.

**Note:** In applied problems a quantity such as \( f(r, \theta) \) would represent some physical quantity, say temperature \( T \). For this reason physicists, engineers, and applied mathematicians will often not make a notational distinction between the functions \( f \) and \( F \), as we did, but simply think of the physical quantity \( T \) as something that can be expressed in terms of either Cartesian or polar coordinates, and writing both \( T(x, y) \) and \( T(r, \theta) \). We will use this convention in the example below.

**Example: Steady state temperature distribution.** We consider first a rectangular metal plate \( a \) units long by \( b \) units wide, and describe a point in this plate by coordinates \((x, y)\), where \(0 \leq x \leq a \) and \(0 \leq y \leq b \). We hold the edges of the plate at some temperatures which are fixed in time but may vary with position, and are interested in the value of the temperature \( T(x, y) \) of the point \((x, y)\) in the interior of the plate. We wait a long time before asking this question, so that the temperature reaches a \text{steady state} and we can ignore any further time dependence. For example, we might suppose that the left and right ends of the plate are held at a temperature \( T_1 \) and the top and bottom at temperature \( T_2 \):

\[
T(0, y) = T_1, \quad T(a, y) = T_1, \quad T(x, 0) = T_2, \quad T(x, b) = T_2, \quad 0 < x < a, \quad 0 < y < b.
\]

(11)

Then it is known that \( T(x, y) \) is determined by the boundary conditions (11) and the fact that it satisfies \text{Laplace’s equation}:

\[
T_{xx} + T_{yy} = 0.
\]

(12)

Laplace’s equation (12) is an example of a \text{partial differential equation}. The heat equation, discussed in Example 11 of Section 14.3 of our text, is another example.

But now suppose that our plate has the shape, not of a rectangle, but a disk of radius \( R \); we can introduce coordinates so that this disk fills the region \( x^2 + y^2 \leq R^2 \). It is natural to use polar rather than Cartesian coordinates to study the temperature, so that the disk is the region \( r \leq R, \quad 0 \leq \theta \leq 2\pi \) and the temperature is written as \( T(r, \theta) \). Boundary conditions would now be imposed on the circular boundary of the disk:

\[
T(R, \theta) = g(\theta), \quad 0 \leq \theta \leq 2\pi,
\]

(13)

for some given function \( g(\theta) \). The temperature function would still satisfy Laplace’s equation, which according to (10) would now be

\[
T_{rr} + \frac{1}{r^2} T_{\theta\theta} + \frac{1}{r} T_r = 0.
\]

(14)

Problems like this are the reason that calculation such as we did above are so important.