MTH4101 Calculus II

Lecture notes for Week 2
Derivatives IV and V

Thomas' Calculus, Sections 14.2 to 14.5

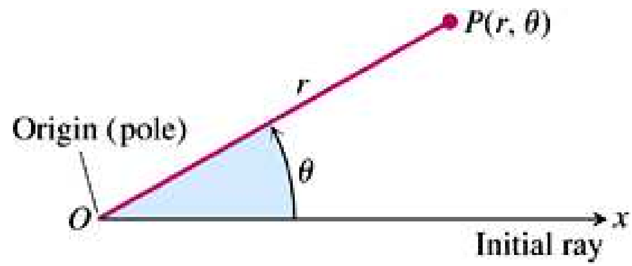
Rainer Klages

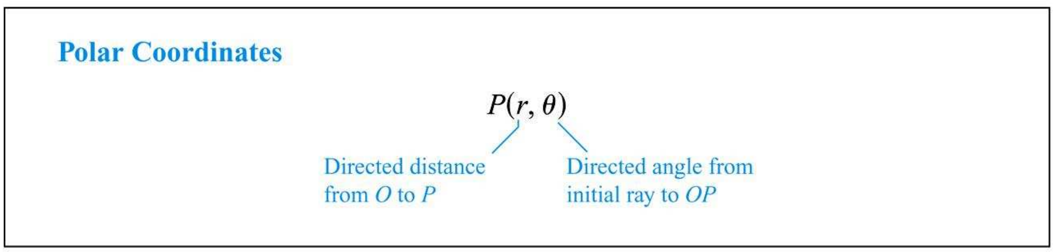
School of Mathematical Sciences
Queen Mary, University of London

Spring 2013

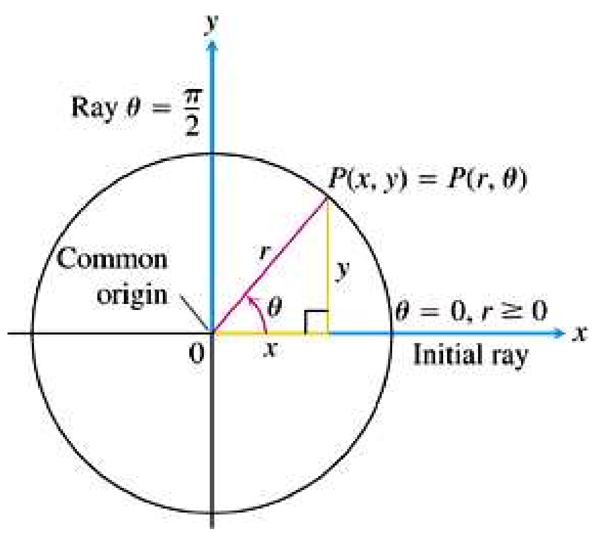
Sometimes it is useful to use polar coordinates.

Reminder (or perhaps not?): Polar coordinates
As an alternative to **Cartesian coordinates** (x, y), we can describe a point P in the plane by using **polar coordinates**:





These coordinates are particularly useful if a function, or a problem, has some circular symmetry. Typically, we restrict ourselves to $0 \leq r$ and $0 \leq \theta < 2\pi$ (why?). Polar and Cartesian coordinates can be converted into each other:



For the direction polar to Cartesian coordinates we easily derive

$$x = r\cos\theta\ , \quad y = r\sin\theta$$

That is, given (r, θ), we can compute (x, y). The direction Cartesian to polar coordinates is left to you as an exercise.[1]

[1] If you have not encountered polar coordinates before in sufficient detail, I highly recommend that you familiarize yourself with Thomas' Calculus, Section 11.3.

Example:
Determine the continuity of the function defined by

$$f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

In polar coordinates, i.e., by using $x = r\cos\theta$, $y = r\sin\theta$, the function can be written as

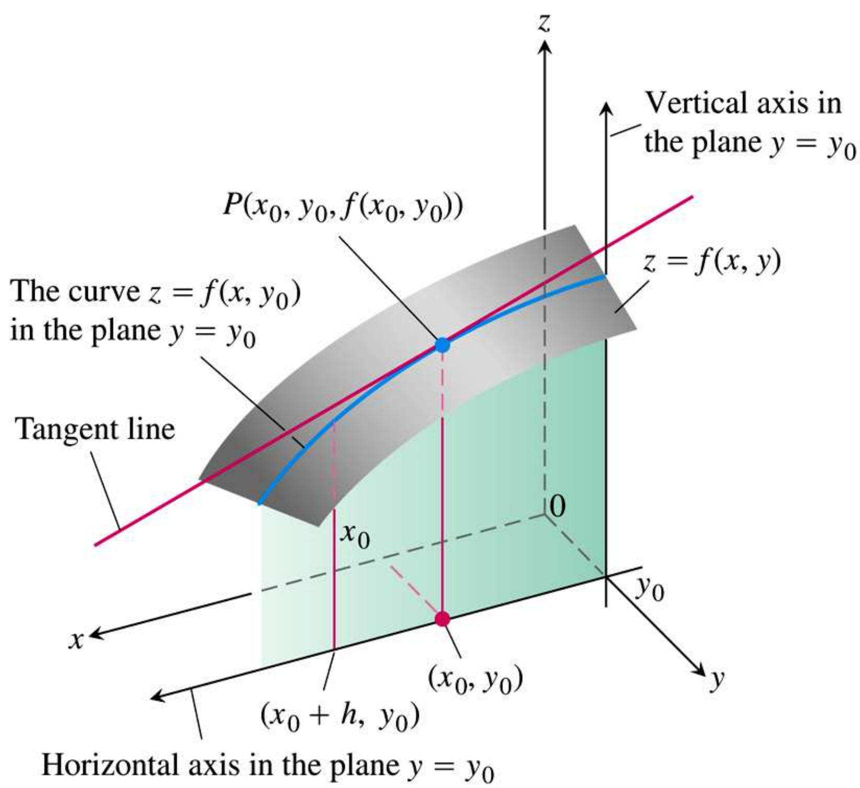
$$f(r,\theta) = \frac{2r^2\cos\theta\sin\theta}{r^2(\cos^2\theta + \sin^2\theta)} = \sin 2\theta$$

provided we are not at the origin (i.e. provided $r \neq 0$). Therefore, as $r \to 0$, the outcome depends on the angle θ. For example, along $\theta = \pi/4$, $f = \sin 2\theta = \sin \pi/2 = 1$ everywhere along the line. Therefore the function is not continuous.

Partial Derivatives

Reminder: Derivative
For functions of one variable, $y = f(x)$, the *derivative* at a point is the gradient of the tangent to the curve at that point.

But for functions of two variables, $z = f(x,y)$, an infinite number of tangents exist at a point. However, if we fix $y = y_0$ in $f(x,y)$ and let x vary, then $f(x,y_0)$ depends only on x:



That is, we can reduce the problem of the many-variable derivative effectively to the one-variable case by holding all but one of the independent variables constant.

Definition:
The **partial derivative** of $f(x,y)$ with respect to x at the point (x_0,y_0) is

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} = \lim_{h\to 0}\frac{f(x_0+h,y_0)-f(x_0,y_0)}{h} = f_x(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0)$$

provided the limit exists.

In complete analogy, the partial derivative of $f(x,y)$ with respect to y at the point (x_0,y_0) is

$$\left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} = \lim_{h\to 0}\frac{f(x_0,y_0+h)-f(x_0,y_0)}{h} = f_y(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0)$$

provided the limit exists.

For example, if $f(x,y)=x^2+y^2$ then $f_x=2x$, $f_y=2y$.

Note how we treat the other variables as constants when we do partial differentiation!

We can extend this to three (or more) dimensions. For example, if $f(x,y,z)=xy^2z^3$ then $f_x=y^2z^3$, $f_y=2xyz^3$, $f_z=3xy^2z^2$.

Example:
Find $\partial f/\partial x$ and $\partial f/\partial y$ at the point $(4,-5)$ for the function $f(x,y)=x^2+3xy+y-1$.

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{\partial}{\partial x}(x^2+3xy+y-1) = 2x+3y\\ \frac{\partial f}{\partial y} &= \frac{\partial}{\partial y}(x^2+3xy+y-1) = 3x+1\,.\end{aligned}$$

At the point $(4,-5)$ we have

$$\left.\frac{\partial f}{\partial x}\right|_{(4,-5)} = -7\,, \qquad \left.\frac{\partial f}{\partial y}\right|_{(4,-5)} = 13\,.$$

Example:
Find $\partial z/\partial x$ if the equation $yz-\ln z = x+y$ (implicitly) defines $z=z(x,y)$.

$$\frac{\partial}{\partial x}(yz-\ln z) = \frac{\partial}{\partial x}(x+y)\,.$$

Hence

$$y\frac{\partial z}{\partial x} - \frac{1}{z}\frac{\partial z}{\partial x} = 1+0\,.$$

This gives

$$\left(y-\frac{1}{z}\right)\frac{\partial z}{\partial x} = 1\,; \qquad \Rightarrow \qquad \frac{\partial z}{\partial x} = \frac{z}{yz-1}\,.$$

We can also obtain higher order derivatives.

Example:
If $f(x,y) = x\cos y + y\,e^x$, find

$$f_{xx} = \frac{\partial^2 f}{\partial x^2}\,, \quad f_{yx} = \frac{\partial^2 f}{\partial x \partial y}\,, \quad f_{yy} = \frac{\partial^2 f}{\partial y^2} \quad \text{and} \quad f_{xy} = \frac{\partial^2 f}{\partial y \partial x}\,.$$

The first step is to find the first partial derivatives:

$$\begin{aligned} \frac{\partial f}{\partial x} &= \cos y + y\,e^x \\ \frac{\partial f}{\partial y} &= -x\sin y + e^x\,. \end{aligned}$$

Now we take the partial derivatives of the first partial derivatives. This gives:

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= y\,e^x \\ \frac{\partial^2 f}{\partial y \partial x} &= -\sin y + e^x \\ \frac{\partial^2 f}{\partial x \partial y} &= -\sin y + e^x \\ \frac{\partial^2 f}{\partial y^2} &= -x\cos y\,. \end{aligned}$$

This illustrates the following Theorem:

Theorem: *Mixed Derivative Theorem*
If $f(x,y)$ and its partial derivatives f_x, f_y, f_{xy} and f_{yx} are *defined* throughout an open region containing a point (a,b) and are *all continuous* at (a,b) then

$$f_{xy}(a,b) = f_{yx}(a,b)\,.$$

(An example where $f_{xy}(a,b) \neq f_{yx}(a,b)$ is provided by the function discussed on p.9/10 of the lecture notes in week 1.)
The theorem can be extended to higher orders, provided the derivatives are continuous.

Example:
Find f_{yxyz} if $f(x,y,z) = 1 - 2xy^2z + x^2y$.

$$f_y = -4xyz + x^2\,, \quad f_{yx} = -4yz + 2x\,, \quad f_{yxy} = -4z\,, \quad f_{yxyz} = -4\,.$$

Reminder:
For functions of a single variable it holds that if $y = f(x)$ is differentiable at $x = x_0$, then the change in the value of f that results from changing x from x_0 to $x_0 + \Delta x$ is given by the *differential approximation*

$$\Delta y = f'(x_0)\Delta x + \epsilon \Delta x$$

in which $\epsilon \to 0$ as $\Delta x \to 0$ (see Thomas' Calculus Section 3.9). For functions of two variables, the analogous property yields the *definition* of differentiability:

DEFINITION **Differentiable Function**

A function $z = f(x, y)$ is **differentiable at** (x_0, y_0) if $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exist and Δz satisfies an equation of the form

$$\Delta z = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_1 \Delta x + \epsilon_2 \Delta y,$$

in which each of $\epsilon_1, \epsilon_2 \to 0$ as both $\Delta x, \Delta y \to 0$. We call f **differentiable** if it is differentiable at every point in its domain.

Note in particular that for $z = f(x, y)$, *differentiability is more than the existence of the partial derivatives*, as becomes also clear from the following statement:

If f_x and f_y are *continuous* throughout an open region R, then f is *differentiable* at every point of R.

It also holds, in analogy to functions of a single variable:

If a function $f(x, y)$ is *differentiable* at a point (x_0, y_0) then f is *continuous* at (x_0, y_0).

If you are interested in the details underlying the above statements, like the *Increment Theorem*, please check out Thomas' Calculus p.771/772.

The Chain Rule

Reminder: Chain Rule for Function of One Variable

If $w = f(x)$ is a differentiable function of x and $x = g(t)$ is a differentiable function of t, then

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\mathrm{d}w}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}\,.$$

Similarly:

Theorem: Chain Rule for Functions of Two Variables

If $w = f(x, y)$ is differentiable and if $x = x(t)$, $y = y(t)$ are differentiable functions of t, then $w = f(x(t), y(t))$ is a differentiable function of t and

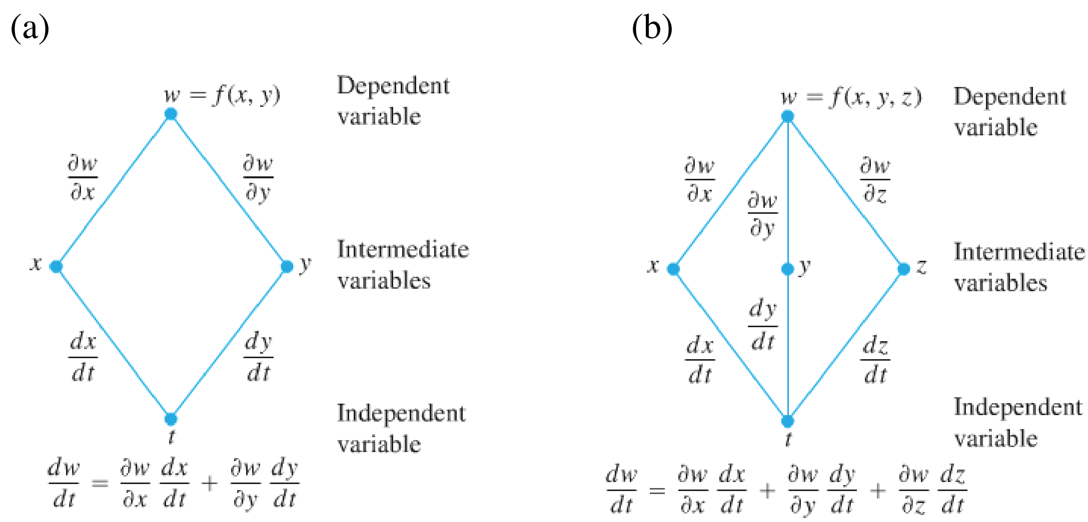
$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}\,.$$

This straightforwardly follows from the above definition of differentiability.

We can easily extend this theorem to functions $w = f(x, y, z)$ of three variables:

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\partial w}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}t}\,.$$

We can use **tree diagrams** to illustrate the application of the Chain Rule:



(a) To find $\mathrm{d}w/\mathrm{d}t$, start at w and read down each route to t, multiplying derivatives along the way; then add the products. (b) For functions of three variables there are three routes from w to t instead of two, but finding $\mathrm{d}w/\mathrm{d}t$ is still the same: read down each route, multiplying derivatives along the way; then add.

Example:
Use the Chain Rule to find the derivative of $w = xy$ with respect to t along the path $x = \cos t$, $y = \sin t$.

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial w}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} = y(-\sin t) + x(\cos t) = -\sin^2 t + \cos^2 t = \cos 2t\,.$$

Note that we could have done this more directly by noting that

$$w = xy = \cos t \sin t = \frac{1}{2}\sin 2t\,; \qquad \frac{\mathrm{d}w}{\mathrm{d}t} = \frac{1}{2}\cdot 2\cos 2t = \cos 2t\,.$$

If $w = f(x, y)$ where $x = g(r, s)$ and $y = h(r, s)$ then

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} \quad \text{and} \quad \frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s}$$

and in analogy for functions $w = f(x, y, z)$. Also, if $w = f(x)$ and $x = g(r, s)$ then

$$\frac{\partial w}{\partial r} = \frac{\mathrm{d}w}{\mathrm{d}x}\frac{\partial x}{\partial r} \quad \text{and} \quad \frac{\partial w}{\partial s} = \frac{\mathrm{d}w}{\mathrm{d}x}\frac{\partial x}{\partial s}\,.$$

Example:
For $u = w(x, y, z)$, express $\partial w/\partial r$ and $\partial w/\partial s$ in terms of r and s if

$$w = x + 2y + z^2\,, \quad x = \frac{r}{s}\,, \quad y = r^2 + \ln s\,, \quad z = 2r\,.$$

We have

$$\begin{aligned}\frac{\partial w}{\partial r} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r} \\ &= (1)\left(\frac{1}{s}\right) + (2)(2r) + (2z)(2) = \frac{1}{s} + 12r\end{aligned}$$

and

$$\begin{aligned}\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s} \\ &= (1)\left(\frac{-r}{s^2}\right) + (2)\left(\frac{1}{s}\right) + (2z)(0) = \frac{2}{s} - \frac{r}{s^2}\,.\end{aligned}$$

Suppose that $w = F(x, y)$ is differentiable and that $F(x, y) = 0$ defines y (implicitly) as a differentiable function of x. Then

$$0 = \frac{dw}{dx} = F_x\frac{dx}{dx} + F_y\frac{dy}{dx} = F_x + F_y\frac{dy}{dx}\,.$$

Hence, at any point where $F_y \neq 0$,

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{F_x}{F_y}\,.$$

This is the **Formula for Implicit Differentiation**.

Example:
Find $\mathrm{d}y/\mathrm{d}x$ if $y^2 - x^2 - \sin xy = 0$.

$$\begin{aligned}F(x, y) &= y^2 - x^2 - \sin xy \\ \frac{\mathrm{d}y}{\mathrm{d}x} &= -\frac{F_x}{F_y} = -\frac{(-2x - y\cos xy)}{(2y - x\cos xy)} = \frac{2x + y\cos xy}{2y - x\cos xy}\,.\end{aligned}$$

You may wish to compare this method with the one that you have learned in Calculus 1, i.e., differentiating the whole equation with respect to x and then solving for dy/dx.

Directional Derivatives and Gradient Vectors

We now investigate the derivative of a function $f(x, y)$ at a point *in a particular direction*:

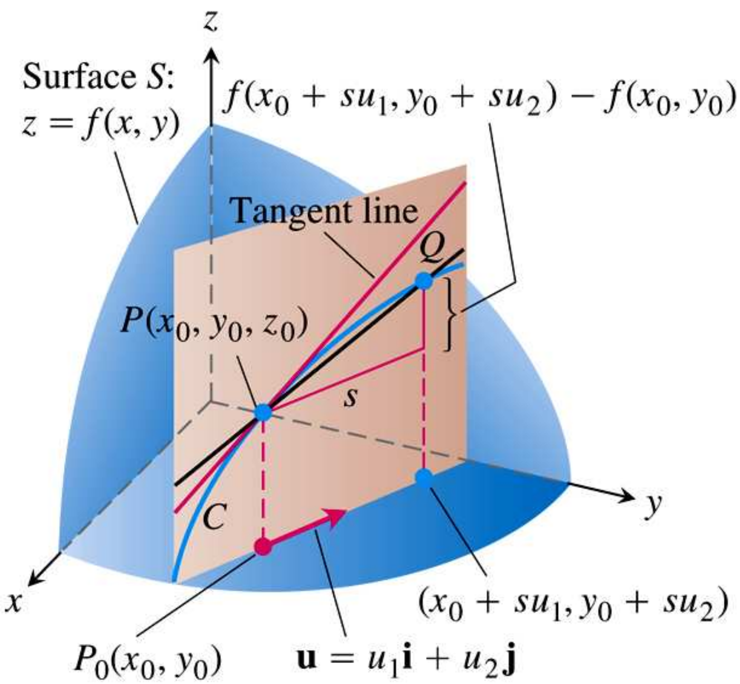
DEFINITION Directional Derivative

The **derivative of f at $P_0(x_0, y_0)$ in the direction of the unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$** is the number

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = \lim_{s\to 0}\frac{f(x_0 + su_1, y_0 + su_2) - f(x_0, y_0)}{s}, \qquad (1)$$

provided the limit exists.

It is also denoted by $(D_{\mathbf{u}}f)_{P_0}$ as the derivative of f at the point P_0 in the direction of the unit vector $\mathbf{u}$. The meaning is illustrated in the following figure:



We can develop a more efficient formula for the directional derivative by considering the line

$$x = x_0 + su_1\,, \qquad y = y_0 + su_2$$

through the point $P_0(x_0, y_0)$, parametrised with the arc length parameter s increasing in the direction of the unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$. Then

$$\begin{aligned}
\left(\frac{\mathrm{d}f}{\mathrm{d}s}\right)_{\mathbf{u},P_0} &= \left(\frac{\partial f}{\partial x}\right)_{P_0}\frac{\mathrm{d}x}{\mathrm{d}s} + \left(\frac{\partial f}{\partial y}\right)_{P_0}\frac{\mathrm{d}y}{\mathrm{d}s} \qquad \text{(via the Chain Rule)}\\
&= \left(\frac{\partial f}{\partial x}\right)_{P_0} u_1 + \left(\frac{\partial f}{\partial y}\right)_{P_0} u_2\\
&= \left[\left(\frac{\partial f}{\partial x}\right)_{P_0}\mathbf{i} + \left(\frac{\partial f}{\partial y}\right)_{P_0}\mathbf{j}\right]\cdot[u_1\mathbf{i} + u_2\mathbf{j}]
\end{aligned}$$