5 Partial Differentiation

5.1 The partial derivative [see Riley et al, Sec. 5.1]

So far we have considered functions of a single variable ie $f = f(x)$ and the slope or gradient at x have been given by $\frac{df(x)}{dx}$ where

$$\frac{df(x)}{dx} = \lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x} \ .$$

We now consider a function of two (or more) variables $f(x, y)$, which for two variables represents a surface (see below), the z axis representing the value of the function $f(x, y)$.

Definition of the partial derivatives

It is clear that a function $f(x, y)$ of two variables will have a gradient in all directions in the xy plane. These rates of change/slopes/gradients are defined as partial derivatives w.r.t the x and y axes. For the positive x direction, holding y constant

$$\left(\frac{\partial f}{\partial x}\right)_y = \lim_{\delta x \to 0} \frac{f(x+\delta x, y) - f(x, y)}{\delta x} = f_x \ .$$

Similarly for the positive y direction, holding x constant

$$\left(\frac{\partial f}{\partial y}\right)_x = \lim_{\delta y \to 0} \frac{f(x, y+\delta y) - f(x, y)}{\delta y} = f_y \ .$$

We can also define second and higher partial derivatives ie

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

Provided the second partial derivatives are continuous then $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$.

Examples

If $s = t^u$ find $\frac{\partial s}{\partial t}$ and $\frac{\partial s}{\partial u}$

$$\frac{\partial s}{\partial t} = u t^{u-1}$$

$$\frac{\partial s}{\partial u} = \frac{\partial t^u}{\partial u} = \frac{\partial}{\partial u}(e^{u \ln t}) = \ln t e^{u \ln t} = t^u \ln t \ .$$

5.2 The total differential and total derivative. [See Riley et al, Sec. 5.2]

For a function of one varaiable, $f(x)$,

$$df = \frac{df}{dx} dx = f(x+\delta x) - f(x)$$

is the differential (change) in f when x is changed infinitesimally by dx.
Having defined the partial derivatives, we now ask what is the change df in $f(x, y)$ if the coordinates (x, y) are changed to $(x + dx, y + dy)$

$$\text{We have } \; df = f(x + \delta x, y + \delta y) - f(x, y)$$

$$= f(x + \delta x, y + \delta y) - f(x, y + \delta y) + f(x, y + \delta y) - f(x, y)$$

$$\text{and } \; \text{df} = \left(\frac{f(x + \delta x, y + \delta y) - f(x, y + \delta y)}{\delta x} \delta x \right)$$

$$+ \left(\frac{f(x, y + \delta y) - f(x, y)}{\delta y} \delta y \right) \; .$$

$$\text{Thus } \; df = \left(\frac{\partial f}{\partial x} \right)_y dx + \left(\frac{\partial f}{\partial y} \right)_x dy.$$

As an example find the total differential of the function

$$f(x, y) = ye^{x+y}$$

$$\text{We have } \left(\frac{\partial f}{\partial x} \right)_y = ye^{x+y} \; \text{ and } \; \left(\frac{\partial f}{\partial y} \right)_x = ye^{x+y} + e^{x+y} \; .$$

$$\text{Thus } \; df = ye^{x+y} dx + (1 + y)e^{x+y} dy \; .$$

<u>Total derivative</u>
When $x = x(t)$ and $y = y(t)$ then $f(x, y)$ is essentially a function of one variable, t. to get the total derivative df/dt, instead of substituting $x(t)$ and $y(t)$ into f, we can proceed using

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

so to obtain

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

Moreover, we note that if f has an explicit dependence on t, then we rewrite the above as:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt} \; .$$

<u>Example</u>
If $f(x, y, t) = \ln t + xe^{-y}$ and $x = 1 + at$, $y = bt^3$ (a,b constants), find df/dt.
We calculate the partial derivatives of f

$$\frac{\partial f}{\partial t} = \frac{1}{t}, \qquad \frac{\partial f}{\partial x} = e^{-y}, \qquad \frac{\partial f}{\partial y} = -xe^{-y}$$

and the derivatives of x and y with respect to t

$$\frac{dx}{dt} = a, \qquad \frac{dy}{dt} = 3bt^2$$

Thus, the total derivative is given by

$$\frac{df}{dt} = \frac{1}{t} + e^{-y}a - xe^{-y}3bt^2 = \frac{1}{t} + e^{-bt^3}\left[a - (1+at)3bt^2\right]$$

5.3 Exact and inexact differentials [see Riley et al, Sec. 5.3]

In the last section we obtained the total differential df by determining the partial derivatives from $f(x, y)$. We now address the inverse problem.

Consider the differential:

$$df = A(x, y)dx + B(x, y)dy \ .$$

Can we go back to the function $f(x, y)$? If we can, this is an exact differential, and if not it is an inexact differential.

From the above general expression, we can identify the partial derivatives

$$\frac{\partial f}{\partial x} = A(x, y), \qquad \frac{\partial f}{\partial y} = B(x, y)$$

then using the property

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$$

we can derive the condition for the differential to be exact:

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}$$

Example

Show that $x^2 dy - (y^2 + xy)dx$ is an inexact differential, but if you multiply by $(xy^2)^{-1}$, it is exact.

$$A = -(y^2 + xy), \qquad B = x^2$$

$$\frac{\partial A}{\partial y} = -2y - x, \qquad \frac{\partial B}{\partial x} = 2x$$

which shows that it is not an exact differential.

Now multiply by $(xy^2)^{-1}$:

$$\frac{x^2}{xy^2}dy - \frac{y^2 + xy}{xy^2}dx = \frac{x}{y^2}dy - \frac{x+y}{xy}dx$$

$$A = -\frac{x+y}{xy} = -\frac{1}{x} - \frac{1}{y}, \qquad B = \frac{x}{y^2}$$

$$\frac{\partial A}{\partial y} = \frac{1}{y^2}, \qquad \frac{\partial B}{\partial x} = \frac{1}{y^2}$$

which shows that it is an exact differential.
[NB: The function is $f = -x/y - \ln x$]

5.4 Change of variables [See Riley, section 5.6]

We have a function $f(x, y)$ and $x = x(t, s)$ and $y = y(t, s)$. We want to change variable to determine $\frac{\partial f}{\partial t}$ and $\frac{\partial f}{\partial s}$.
From previously

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

And, since x, y are functions of t, s:

$$\begin{aligned} dx &= \frac{\partial x}{\partial t}dt + \frac{\partial x}{\partial s}ds \\ dy &= \frac{\partial y}{\partial t}dt + \frac{\partial y}{\partial s}ds \end{aligned}$$

Thus,

$$\begin{aligned} df &= \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial t}dt + \frac{\partial x}{\partial s}ds\right) + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial t}dt + \frac{\partial y}{\partial s}ds\right) \\ &= \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}\right)dt + \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}\right)ds \end{aligned}$$

But f is also a function of t and s, so

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial s}ds$$

Comparing the last two equations:

$$\begin{aligned} \frac{\partial f}{\partial t} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \\ \frac{\partial f}{\partial s} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \end{aligned}$$

Example
If $z(x, y) = xy$ and $x(s, t) = s - t$ and $y(s, t) = \sin(s + t)$, find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

First use:

$$
\begin{aligned}
\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\
&= y \cdot 1 + x \cdot \cos(s+t) = \\
&= \sin(s+t) + (s-t)\cos(s+t)
\end{aligned}
$$

and similarly for $\frac{\partial z}{\partial t}$.
Or we could have expressed z in terms of s and t

$$z = xy = (s-t)\sin(s+t)$$

from which we can immediately derive the partial derivative

$$\frac{\partial z}{\partial s} = \sin(s+t) + (s-t)\cos(s+t)$$

5.5 Stationary points of multivariate functions [see Riley, section 5.8]

In multivariate calculus, stationary points are determined by the condition

$$\left(\frac{\partial f}{\partial x}\right)_y = \left(\frac{\partial f}{\partial y}\right)_x = 0 \ .$$

To determine their nature, we consider the following conditions:

- A minimum if the following three conditions are satisfied

  $$f_{xx} > 0, f_{yy} > 0 \text{ and } f_{xx}f_{yy} > f_{xy}^2 \ .$$

- A maximum if

  $$f_{xx} < 0, f_{yy} < 0 \text{ and } f_{xx}f_{yy} > f_{xy}^2 \ .$$

  This last part of this condition turns out the same as for a minimum.

- A saddle point if

  $$f_{xy}^2 > f_{xx}f_{yy} \ .$$

  If $f_{xy}^2 = f_{xx}f_{yy}$, further investigation is required by Taylor-expanding (see Chapter 6) the function to higher orders. This includes the case $f_{xx} = f_{yy} = f_{xy} = 0$.

Example
Find the critical point of the function

$$f(x,y) = x^2 - 2xy + 2y^2 - 2y + 2$$

and show that this critical point is a local minimum.

We have

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x - 2y \\ \frac{\partial f}{\partial y} &= -2x + 4y - 2\end{aligned}$$

By setting the first partial derivatives to zero we find:

$$\begin{aligned}2x - 2y &= 0 \Rightarrow x = y \\ -2x + 4y - 2 &= 0 \Rightarrow (\text{replacing } x = y) 2x - 2 = 0\end{aligned}$$

which gives $x = y = 1$. Now, we calculate the higher order derivative:

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= 2 \\ \frac{\partial^2 f}{\partial y^2} &= 4 \\ \frac{\partial^2 f}{\partial x \partial y} &= -2\end{aligned}$$

Since $f_{xx} > 0$, $f_{yy} > 0$ and $f_{xx} f_{yy} - f_{xy}^2 > 0$, f has a local minimum at $(1, 1)$.

5.6 Stationary points when there is a constraint [see Riley 5.9]

We may have a situation where not all variables are independent, as has been the case so far. So, we may have have a constraint of the form $\phi(x, y, z)$ =constant. Then one of the variables, say z is not independent, it depends on x and y. We could in fact use $\phi(x, y, z) = c$ to eliminate z from f, but this can be difficult or even impossible. The method of the Lagrange multiplier is an elegant way of handling this problem.

So, for a function of three variables $f(x, y, z)$ and the constraint $\phi(x, y, z) = c$ we have

$$df = 0 = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$$

$$\text{and } d\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy + \frac{\partial \phi}{\partial z} dz = 0$$

multiplying $d\phi$ by λ and adding to df

$$df + \lambda d\phi = \left(\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} \right) dx + \left(\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} \right) dy + \left(\frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} \right) dz = 0 \ .$$

where λ is the Lagrange multiplier.

Then since dx, dy, dz are independent and $df + \lambda d\phi = 0$, we must choose λ such that

$$\begin{aligned}
\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} &= 0 \\
\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} &= 0 \\
\frac{\partial f}{\partial z} + \lambda \frac{\partial \phi}{\partial z} &= 0
\end{aligned}$$

Example

Find the rectangle of maximum area which can be placed with its sides parallel lo the x and y axes inside the ellipse $x^2 + 4y^2 = 1$.

So we want to maximize the area

$$A = 2x \cdot 2y = 4xy$$

under the constraint

$$x^2 + 4y^2 = 1 \ .$$

We identify f and ϕ as

$$f = 4xy$$

$$\phi = x^2 + 4y^2$$

and derive:

$$\begin{aligned}
\frac{\partial f}{\partial x} + \lambda \frac{\partial \phi}{\partial x} &= 4y + 2\lambda x = 0 \Rightarrow 2y + \lambda x = 0 \Rightarrow \lambda = -\frac{2y}{x} \\
\frac{\partial f}{\partial y} + \lambda \frac{\partial \phi}{\partial y} &= 4x + 8\lambda y = 0 \Rightarrow x + 2\lambda y = 0 \\
&\Rightarrow \text{(substituting for lambda)} x + 2y(-)\frac{2y}{x} = 0 \Rightarrow x^2 - 4y^2 = 0 \Rightarrow x = \pm 2y
\end{aligned}$$

but $x > 0, y > 0$ so $x = 2y$. And replacing in the original equation for the ellipse:

$$x^2 + 4y^2 = 1 \Rightarrow 4y^2 + 4y^2 = 1 \Rightarrow y = \frac{1}{2\sqrt{2}}, \quad x = \frac{1}{\sqrt{2}}$$

and the maximum area is

$$A = 4xy = 4\frac{1}{\sqrt{2}}\frac{1}{2\sqrt{2}} = 1 \ .$$

Example

Find the values of x and y that maximise the function

$$f(x, y) = xy^{3/2}$$

subject to the constraint

$$x + 2y = 100 \ .$$

The first order conditions are:

$$\begin{aligned} y^{3/2} + \lambda &= 0 \\ \frac{3}{2} x y^{1/2} + \lambda \cdot 2 &= 0 \end{aligned}$$

from which

$$\frac{3}{2} x y^{1/2} - 2 y^{3/2} = 0 \Rightarrow y = \frac{3}{4} x$$

and replacing in the constraint:

$$x + 2y = 100 \Rightarrow x + 2\frac{3}{4}x = 100 \Rightarrow \frac{5}{2}x = 100 \Rightarrow x = 40$$

and

$$y = \frac{3}{4}x = \frac{3}{4} \cdot 40 = 30$$

5.7 Polar coordinates in two dimensions

Consider polar coordinates in two dimensions. The position vector is

$$\underline{r} = x\underline{i} + y\underline{j} \ ,$$

with

$$\begin{aligned} x &= r\cos\theta && (1) \\ y &= r\sin\theta \ . && (2) \end{aligned}$$

The unit vectors are $\underline{\hat{r}}$ and $\underline{\hat{\theta}}$ and are not constant because their directions change. If a vector, $\underline{r}$, depends on a parameter u, then a vector that points in the direction determined by an infinitesimal increase in u is defined by

$$\underline{e_u} = \frac{\partial \underline{r}}{\partial u}$$

and the unit vector pointing in the same direction is

$$\underline{\hat{e}_u} = \frac{\underline{e_u}}{|\underline{e_u}|}$$

In terms of $\underline{i}$ and $\underline{j}$

$$\underline{e_r} = \frac{\partial \underline{r}}{\partial r}, \qquad \underline{\hat{e}_r} \equiv \underline{\hat{r}} = \cos\theta \underline{i} + \sin\theta \underline{j}$$

$$\underline{e_\theta} = \frac{\partial \underline{r}}{\partial \theta}, \qquad \underline{\hat{e}_\theta} \equiv \underline{\hat{\theta}} = -\sin\theta \underline{i} + \cos\theta \underline{j}$$

from which we can derive the derivatives:

$$\frac{d\underline{\hat{r}}}{d\theta} = -\sin\theta\underline{i} + \cos\theta\underline{j} = \underline{\hat{\theta}}$$

$$\frac{d\underline{\hat{\theta}}}{d\theta} = -\cos\theta\underline{i} - \sin\theta\underline{j} = -\underline{\hat{r}}$$

The velocity $\underline{v}$ is

$$\begin{aligned}
\underline{v} = \frac{d\underline{r}}{dt} = \frac{dr\underline{\hat{r}}}{dt} &= \frac{dr}{dt}\underline{\hat{r}} + r\frac{d\underline{\hat{r}}}{dt} \\
&= \frac{dr}{dt}\underline{\hat{r}} + r\frac{d\underline{\hat{r}}}{d\theta}\frac{d\theta}{dt} \\
&= \frac{dr}{dt}\underline{\hat{r}} + r\frac{d\theta}{dt}\underline{\hat{\theta}} \\
&= v_r\underline{\hat{r}} + v_\theta\underline{\hat{\theta}}
\end{aligned}$$

and the acceleration is given by (show it as an exercise)

$$\begin{aligned}
\underline{a} = \frac{d\underline{v}}{dt} &= \frac{d}{dt}\left(\frac{dr}{dt}\underline{\hat{r}} + r\frac{d\theta}{dt}\underline{\hat{\theta}}\right) = \\
&= \left(\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right)\underline{\hat{r}} + \left(r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt}\right)\underline{\hat{\theta}}
\end{aligned}$$

or in another notation

$$\underline{a} = \left(\ddot{r} - r\dot{\theta}^2\right)\underline{\hat{r}} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\underline{\hat{\theta}}$$

5.8 Cylindrical and spherical polar coordinates [see Riley, section 10.9]

Cylindrical polar coordinates

The position of a point P in cylindrical polar coordinates is

$$\begin{aligned}
x &= \rho\cos\phi \\
y &= \rho\sin\phi \\
z &= z
\end{aligned}$$

and the position vector $\underline{r}$ is

$$\underline{r} = \rho\cos\phi\underline{i} + \rho\sin\phi\underline{j} + z\underline{k}\ .$$

The unit vectors, $\underline{\hat{\rho}}$, $\underline{\hat{\phi}}$, $\underline{\hat{k}}$, are in the directions of increasing ρ, ϕ, z, i.e.:

$$\begin{aligned}
\underline{e}_\rho &= \frac{\partial\underline{r}}{\partial\rho} \\
\underline{e}_\phi &= \frac{\partial\underline{r}}{\partial\phi} \\
\underline{e}_z &= \frac{\partial\underline{r}}{\partial z}
\end{aligned}$$

and after normalization we have for the unit vectors:

$$\begin{aligned}
\underline{\hat{\rho}} &= \underline{i}\cos\phi + \underline{j}\sin\phi \\
\underline{\hat{\phi}} &= -\underline{i}\sin\phi + \underline{j}\cos\phi \\
\underline{\hat{k}} &= \underline{k}
\end{aligned}$$

<u>Spherical polar coordinates</u>

The position of point P in spherical polar coordinates is

$$\begin{aligned}
x &= r\cos\phi\sin\theta \\
y &= r\sin\phi\sin\theta \\
z &= r\cos\theta \ .
\end{aligned}$$

The position vector is

$$\underline{r} = r\cos\phi\sin\theta\underline{i} + r\sin\phi\sin\theta\underline{j} + r\cos\theta\underline{k}$$

and the unit vectors $\underline{\hat{r}}$,$\underline{\hat{\theta}}$,$\underline{\hat{\phi}}$ in directions of increasing r,θ,ϕ respectively are

$$\begin{aligned}
\underline{\hat{r}} &= \underline{i}\sin\theta\cos\phi + \underline{j}\sin\theta\sin\phi + \underline{k}\cos\theta \\
\underline{\hat{\theta}} &= \underline{i}\cos\theta\cos\phi + \underline{j}\cos\theta\sin\phi - \underline{k}\sin\theta \\
\underline{\hat{\phi}} &= -\underline{i}\sin\phi + \underline{j}\cos\phi
\end{aligned}$$