# 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, i.e.

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

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

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

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

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

Gradient

The partial derivatives can be interpreted as (cartesian) components of a vector called the gradient,

$$\underline{\nabla} f(x, y) = \frac{\partial f}{\partial x}\underline{i} + \frac{\partial f}{\partial y}\underline{j} \equiv \begin{pmatrix} f_x \\ f_y \end{pmatrix} .$$

The notation for the gradient uses the so called Nabla operator $\underline{\nabla}$ applied to a function $f(x, y)$. Altogether, $\underline{\nabla} f$ combines the properties of a partial derivative and a vector.

The gradient vector will in general depend both in magnitude and direction on the position $(x, y)$, as the partial derivatives are functions of the independent variables $x$ and $y$. This is the first example of a so called vector field, which we will discuss below in more detail in the context of line integrals. The gradient has a very useful interpretation: its direction at a given point in the $(x, y)$ variable plane coincides with the direction where the function $f(x, y)$ rises most strongly (maximal slope). The magnitude of the gradient at this point then gives this maximal slope.

Example: Determine the gradient vector field for the function $f(x, y) = x^2 + y^2$. Evaluate and describe the gradient at the point $x = 1, y = 1$.

The partial derivatives are given by

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y.$$

The gradient vector field is therefore described by

$$\underline{\nabla} f(x, y) = 2x\underline{i} + 2y\underline{j} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

At the point $x = 1, y = 1$ it is given by $\underline{v} = \underline{\nabla} f(1, 1) = 2\underline{i} + 2\underline{j}$. The maximum slope of the function $f$ at $(1, 1)$ is the magnitude of this vector $|\underline{v}| = 2\sqrt{2}$. The direction of steepest ascent can for example be determined from the unit vector in the direction of $\underline{v}$, $\hat{\underline{v}} = \frac{1}{\sqrt{2}}\underline{i} + \frac{1}{\sqrt{2}}\underline{j}$. It points radially outward from the origin with an angle of $\pi/4$ to the positive $x$ axis.

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

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

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

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 $(x + dx, y + dy)$

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

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

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

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

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

Using the gradient vector and defining the infinitesimal displacement vector $d\underline{r} = dx\underline{i} + dy\underline{j}$, the total differential can also be written as

$$df = (\underline{\nabla} f)\cdot d\underline{r},$$

where the dot signifies that we take the scalar product between the gradient and the displacement vector.

<u>Example:</u> 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:

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

Using the definition of the gradient, we can write the total derivative in a compact way,

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + (\underline{\nabla} f)\cdot\frac{d\underline{r}}{dt} \ ,$$

where $\frac{dr}{dt}$ is the derivative of the vector $\underline{r}(t) = x(t)\underline{i} + y(t)\underline{j}$ describing the path in terms of a parameter $t$.

<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

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

Derivative in a direction

As a special case of the total derivative we can consider the following simple path,

$$\underline{r}(t) = \begin{pmatrix} x(t) \\ y(y) \end{pmatrix} = \begin{pmatrix} x_0 + n_x t \\ y_0 + n_y t \end{pmatrix},$$

originating at a point $(x_0, y_0)$ and with the components $n_x$ and $n_y$ of a normalized unit vector $\underline{\hat{e}} = n_x \underline{i} + n_y \underline{j}$, $n_x^2 + n_y^2 = 1$. The total derivative at the point $(x_0, y_0)$ is then given by

$$\left.\frac{df}{dt}\right|_{0,\underline{\hat{e}}} = (\underline{\nabla} f)_0 \cdot \underline{\hat{e}}.$$

Using the properties of vector scalar products, several useful geometric interpretations can be inferred from this result:

1. $\left.\frac{df}{dt}\right|_{0,\underline{\hat{e}}}$ gives the slope of the function $f$ at the point $(x_0, y_0)$ in the direction of $\underline{\hat{e}}$.

2. If $\underline{\hat{e}}$ is pointing in the same direction as the gradient $(\underline{\nabla} f)_0$, the slope is maximal (this coincides with our earlier interpretation of the gradient).

3. If $\underline{\hat{e}}$ is pointing in the opposite direction of the gradient, the slope is negative and given by $-(\underline{\nabla} f)_0$ (maximal descent).

4. If $\underline{\hat{e}}$ is perpendicular to the gradient $(\underline{\nabla} f)_0$, the slope is zero, i.e. the function $f$ remains constant when (slightly) moving in this direction.

## 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 Line integrals

Scalar and vector fields

A field assigns a quantity to each point in a region of space.

A scalar field assigns a single value (usually a real number) to a point, so it is nothing but a multi-variate function $f = f(x, y)$ that we have been discussing in this chapter. Examples of scalar fields in physics are height of land above sea level, air pressure or temperature (in 3D space), potential energy of a particle in a gravitational field.

A vector field assigns a vector, i.e. a quantity with magnitude and direction, to each point, $\underline{v} = \underline{v}(\underline{r})$. Examples of vector fields in physics are: velocity of air in the atmosphere, electric and magnetic field strength, the gravitational force field. We have already encountered an important example of a vector field: the gradient of a function, $\underline{\nabla} f(\underline{r})$. A physical application of the gradient is the determination of a force field $\underline{F}(\underline{r})$ (the force a particle feels as a function of $x$ and $y$) from the potential energy $U(\underline{r})$, $\underline{F} = -\underline{\nabla} U$ (the minus sign applies here as the force will be in a direction such that the potential energy *decreases*.)

Definition of a line integral
A line integral takes a vector field $\underline{F}(\underline{r})$ and a path $C$, defined by a parametrization $C : \underline{r} = \underline{r}(t)$, and 'integrates' over their scalar product,

$$I = \int_C \underline{F} \cdot d\underline{r}.$$

The notation used in the above equation reflects how the line integral is mathematically defined: The path is subdivided into infinitesimally small segments $d\underline{r}$ and the scalar product $dI = \underline{F} \cdot d\underline{r}$ is calculated for each segment. Finally, one sums over all pieces and the integral is achieved by taking the limit such that the lengths of the segments go to zero, $\lim_{|d\underline{r}|\to 0} \sum_C dI$. As we are taking the scalar product between the vector field (at a given point of the path) and the path vector segment, the end result of the integration is a number.

The above procedure may sound complicated but if we have a parametrization of the path, i.e. functions $x = x(t)$, $y = y(t)$ that determine the path $\underline{r} = x(t)\underline{i} + y(t)\underline{j}$, we can write the line integral as

$$I = \int_C \underline{F} \cdot d\underline{r} = \int_{t_A}^{t_B} \left[ \underline{F}(\underline{r}(t)) \cdot \frac{d\underline{r}}{dt} \right] dt.$$

Here we have used the differential $d\underline{r} = \frac{d\underline{r}}{dt} dt$. The right-hand side of the above equation is nothing but a normal integral over $t$ with the integrand given in the square brackets. The integral limits $t_A$ and $t_B$ represent the start and end points of the path $\underline{r}$ in the given parametrization: $\underline{r}_A = \underline{r}(t_A)$ (start point), $\underline{r}_B = \underline{r}(t_B)$ (end point).

In general a line integral will depend on both the vector field and the choice of path. Even if the vector field and the start / end point are fixed, the line integral will depend on the choice of path from $\underline{r}_A$ to $\underline{r}_B$. (Note: the line integral will be independent of the specific *parametrization* of a given path, though).

Example: A vector field $\underline{F}$ is defined by $\underline{F} = xy\underline{i} - y^2\underline{j}$. Evaluate the line integral for the path defined by a straight line from the start point $\underline{r}_A = (0, 0)$ to the end point $\underline{r}_B = (2, 1)$.

A parametrization of the path is given by

$$x(t) = 2t, \quad y(t) = t \quad \Rightarrow \quad \underline{r}(t) = 2t\underline{i} + t\underline{j},$$

with the values $t_A = 0$ and $t_B = 1$ of the parameter $t$ corresponding to the start and end point, respectively.

The line integral can then be calculated as

$$\begin{aligned}
I &= \int_C \underline{F} \cdot d\underline{r} \\
&= \int_{t_A}^{t_B} \underline{F}(\underline{r}(t)) \cdot \frac{d\underline{r}}{dt} dt \\
&= \int_0^1 (x(t)y(t)\underline{i} - y^2(t)\underline{j}) \cdot \frac{d(2t\underline{i} + t\underline{j})}{dt} dt \\
&= \int_0^1 (2t^2\underline{i} - t^2\underline{j}) \cdot (2\underline{i} + \underline{j}) dt \\
&= \int_0^1 3t^2 dt = \left[x^3\right]_0^1 = 1.
\end{aligned}$$

For sufficiently simple paths, it is often possible to use one of the coordinates as the parameter. This speeds up the calculation as we can take a few short cuts. In the above example, we could have simply chosen $y$ as parameter and expressed $x$ in terms of $y$: $x = 2y$.

Example: For the same vector field $\underline{F} = xy\underline{i} - y^2\underline{j}$, calculate the line integral over the parabolic path $y = x^2/4$ from $(0, 0)$ to $(2, 1)$.

We choose $x$ as the path parameter ranging from 0 to 2. Therefore $dy = \frac{1}{2}xdx$ and the line integral is given by

$$\begin{aligned}
I &= \int_C \underline{F} \cdot d\underline{r} \\
&= \int_C (xy\underline{i} - y^2\underline{j}) \cdot (dx\underline{i} + dy\underline{j}) \\
&= \int_C (\frac{1}{4}x^3\underline{i} - \frac{1}{16}x^4\underline{j}) \cdot (dx\underline{i} + \frac{1}{2}xdx\underline{j}) \\
&= \int_0^2 (\frac{1}{4}x^3 - \frac{1}{32}x^5) dx = \left[\frac{1}{16}x^4 - \frac{1}{192}x^6\right]_0^2 = 1 - \frac{1}{3} = \frac{2}{3}.
\end{aligned}$$

It can be seen that the line integrals in the two examples differ as a consequence of the different paths chosen.

Physics application: Work done in a force field

Consider a particle in a force field $\underline{F}$, e.g. a satellite in the gravitational force field of the earth. At a given point $\underline{r}$, the particle will feel the force $\underline{F}(\underline{r})$ and moving it to an infinitesimally near point $\underline{r} + d\underline{r}$ will require the mechanical work

$$dW = -\underline{F} \cdot d\underline{r}.$$

The scalar product applies to ensure that only the component of the force along the displacement vector takes effect. Summing over such infinitesimal steps along a given path $C$ will therefore correspond to taking the line integral

$$W_C = -\int_C \underline{F} \cdot d\underline{r}.$$

Loop integrals

A loop integral is a line integral for which the path forms a closed loop, i.e. the start and end points coincide, $\underline{r}_A = \underline{r}_B$. Such loop integrals are sometimes denoted using the special $\oint$ notation,

$$I = \oint_C \underline{F} \cdot d\underline{r}.$$

In order to practically calculate a loop integral requires an appropriate parametrization such that $t_A$ is different from $t_B$.

Example: Consider the vector field $\underline{F} = (y + 2xy)\underline{i} + (x + x^2)\underline{j}$. Calculate the loop integral formed by the following path: Starting at $A : (1,1)$, go straight to $B : (2,1)$, then straight to $C : (2,2)$, then straight to $D : (1,2)$ and finally straight back to $A : (1,1)$.

We calculate the loop integral by splitting it into four parts:

$$I = \oint_C \underline{F} \cdot d\underline{r} = I_{A\to B} + I_{B\to C} + I_{C\to D} + I_{D\to A}.$$

The individual parts are calculated as conventional line integrals where we use either $x$ or $y$ as parameter as appropriate:

$$\begin{aligned}
I_{A\to B} &= \int \underline{F} \cdot dx\underline{i} \qquad \text{with } dy = 0 \text{ as } y = 1 = \text{const} \\
&= \int_1^2 (1 + 2x1)dx = [x + x^2]_1^2 = 4, \\
I_{B\to C} &= \int \underline{F} \cdot dy\underline{j} \qquad \text{with } dx = 0 \text{ as } x = 2 = \text{const} \\
&= \int_1^2 (2 + 2^2)dy = 6, \\
I_{C\to D} &= \int \underline{F} \cdot dx\underline{i} \qquad \text{with } dy = 0 \text{ as } y = 2 = \text{const} \\
&= \int_2^1 (2 + 4x)dx \quad \text{(note the order of the limits)} \\
&= \left[2x + 2x^2\right]_2^1 = 4 - 12 = -8, \\
I_{D\to A} &= \int \underline{F} \cdot dy\underline{j} \qquad \text{with } dx = 0 \text{ as } x = 1 = \text{const} \\
&= \int_2^1 (1 + 1^2)dy = -2.
\end{aligned}$$

Summing all parts yields

$$I = \oint_C \underline{F} \cdot d\underline{r} = 4 + 6 - 8 - 2 = 0.$$

The loop integral is zero as $\underline{F}$ is the gradient vector field of the function $f(x,y) = x^2y + xy$. The example illustrates the technique of splitting up a path into two or more parts parts and summing the individual results afterward.

## 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 $\lambda$ 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 $\lambda$ is the Lagrange multiplier.
Then since $dx$, $dy$, $dz$ are independent and $df + \lambda d\phi = 0$, we must choose $\lambda$ 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 $\phi$ 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

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

<u>Example:</u> 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:

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

from which

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

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 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

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

## 5.8 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

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

## 5.9 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 $\rho$, $\phi$, $z$, 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:

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

Spherical polar coordinates

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 $r$,$\theta$,$\phi$ respectively are

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