Partial Differentiation and Multiple Integrals

6 lectures, 1MA Series
Dr D W Murray
Michaelmas 1994

Textbooks

Most mathematics for engineering books cover the material in these lectures. Stephenson, "Mathematical Methods for Science Students" (Longman) is reasonable introduction, but is short of diagrams. Do look at other texts — they may cover something in a way that is more illuminating for you, and don't just rely on these notes. Other texts are E Kreysig "Advanced Engineering Mathmetics", 5th ed, Sokolnikoff and Redheffer, "Mathematics and Physics of Modern Engineering" (McGraw-Hill), K F Riley "Mathematical methods for the Physical Sciences" (CUP). The Schaum Series book "Calculus" contains all the worked examples you could wish for.

Contents

1. The partial derivative. First and higher partial derivatives. Total and partial differentials, and their use in estimating errors. Testing for total (or perfect) differentials. Integrating total differentials to recover original function.

2. Relationships involving first order partial derivatives. Function of a function. Composite functions, the Chain Rule and the Chain Rule for Partials. Implicit Functions.

3. Transformations from one set of variables to another. Transformations as "old in terms of new" and "new in terms of old". Jacobians. Transformations to Plane, spherical and polar coordinates. What makes a good transformation? Functional dependence. Shape.

4. Applications. Taylor's Theorem for two variables. Stationary points. Plotting functions. Constrained extrema using Lagrange multipliers.

5. Double, triple (and higher) integrals using repeated integration. Transformations, the use of the Jacobian. Plane, spherical and polar coordinates revisited.

6. Volumes of revolution. Line integrals and surface integrals (not involving vector fields!). A note on vector field theory.

Chapter 1

Total and Partial Derivatives and Differentials

So far in your explorations of the differential and integral calculus, it is most likely that you have only considered functions of one variable — $y = f(x)$ and all that. The independant variable is x; y depends on x via the function f. You will know, for example, how to find

$$\frac{d}{dx}[\ln[\cosh[\sin^{-1}[1 - \frac{1}{x}]]]] \qquad : x > 1$$

and

$$\int \frac{x}{(x^2+1)^n} dx \ .$$

But you do not have to look too hard to find quantities which depend on more than one variable. So how can we extend our notions of differentiation and integration to cover such cases?

1.1 Revision of continuity and the derivative for one variable

1.1.1 Functions and ranges of validity

Suppose we want to make a real function f of some real variable x. We certainly need the function "recipe", but we should also specify a range E of admissible values for x for which x is mapped onto y by $y = f(x)$. We might have $y = f(x) = x^2$, with E being the interval $-1 \leq x \leq 1$. Note that with this definition it is meaningless to ask "what is $f(-2)$?". Often, however, the range is not specified: the assumption then is that the range is such as to make y real. For example, if $y = (1+x)^{-1/2}$ we would assume that $-1 < x$. (Those of you experienced in writing functions in a computer language will know the importance of setting proper ranges for input variables. Without them, a program may crash.)

1.1.2 Continuity for a function of one variable

Suppose we have a value $x = a$. We can define a *neighbourhood* of x-values around a by requiring $|x-a| < \delta$ or, equivalently, $(x-a)^2 < \delta^2$. A function $y = f(x)$ is said to be continuous at $x = a$ if, for every positive number ϵ (however small), one can find a neighbourhood

$$|x - a| < \delta \tag{1.1}$$

in which

$$|f(x) - f(a)| < \epsilon \quad . \tag{1.2}$$

Another way of expressing this is:

$$\lim_{x \to a} f(x) = f(a) \quad . \tag{1.3}$$

Note that this requires both the limit to exist and to equal $f(a)$. Also it should not matter how x tends to a. Some useful facts about continuous functions are:

- The sum, difference and product of two continuous functions is continuous.
- The quotient of two continuous functions is continous at every point where the denominator is not zero.

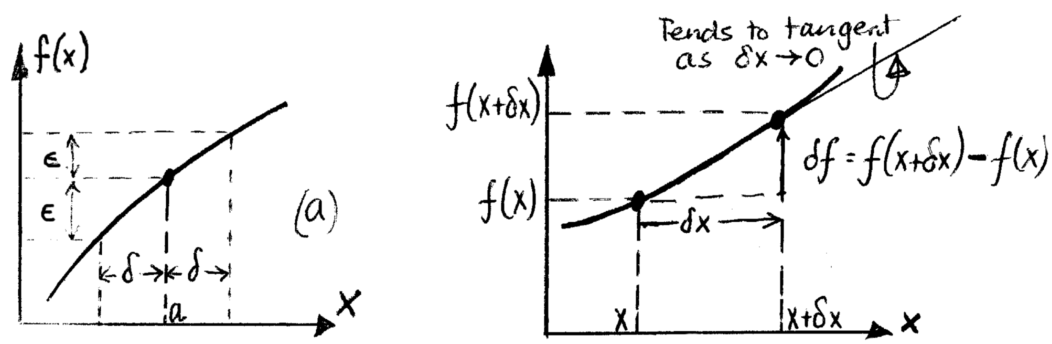

Figure 1.1: (a) Neighbourhood for continuity for a function of one variable and (b) geometrical interpretation of the derivative.

1.1.3 The derivative for a function of one variable

Recall from previous lectures that the *derivative* is defined as

$$\frac{d}{dx} f(x) = \lim_{\delta x \to 0} \left[\frac{f(x + \delta x) - f(x)}{\delta x} \right] \quad . \tag{1.4}$$

$f(x)$ is differentiable if this limit exists, and exists independent of how $\delta x \to 0$. **Note** that

- A differentiable function *is* continuous.
- A continuous function *is not necessarily* differentiable. $f(x) = |x|$ is an example of a function which is continous but not differentiable (at $x = 0$).

1.2 Moving to more than one variable

The big changes we will encounter take place in going from $n = 1$ variables to $n > 1$ variables. So for much of the time we can keep the page uncluttered by dealing with functions of only two variables

$$z = f(x, y) \quad .$$

Functions of two variables are conveniently represented graphically using the Cartesian axes $Oxyz$. The function representation is a *surface*, as opposed to a *plane curve* for a one variable function. It is a good deal harder to represent functions of more than two variables – you might ask yourself why.
Let us look at the graphical representations of some functions. We will return to discuss how to make sketches of functions later. What is the admissible range E for the last function?

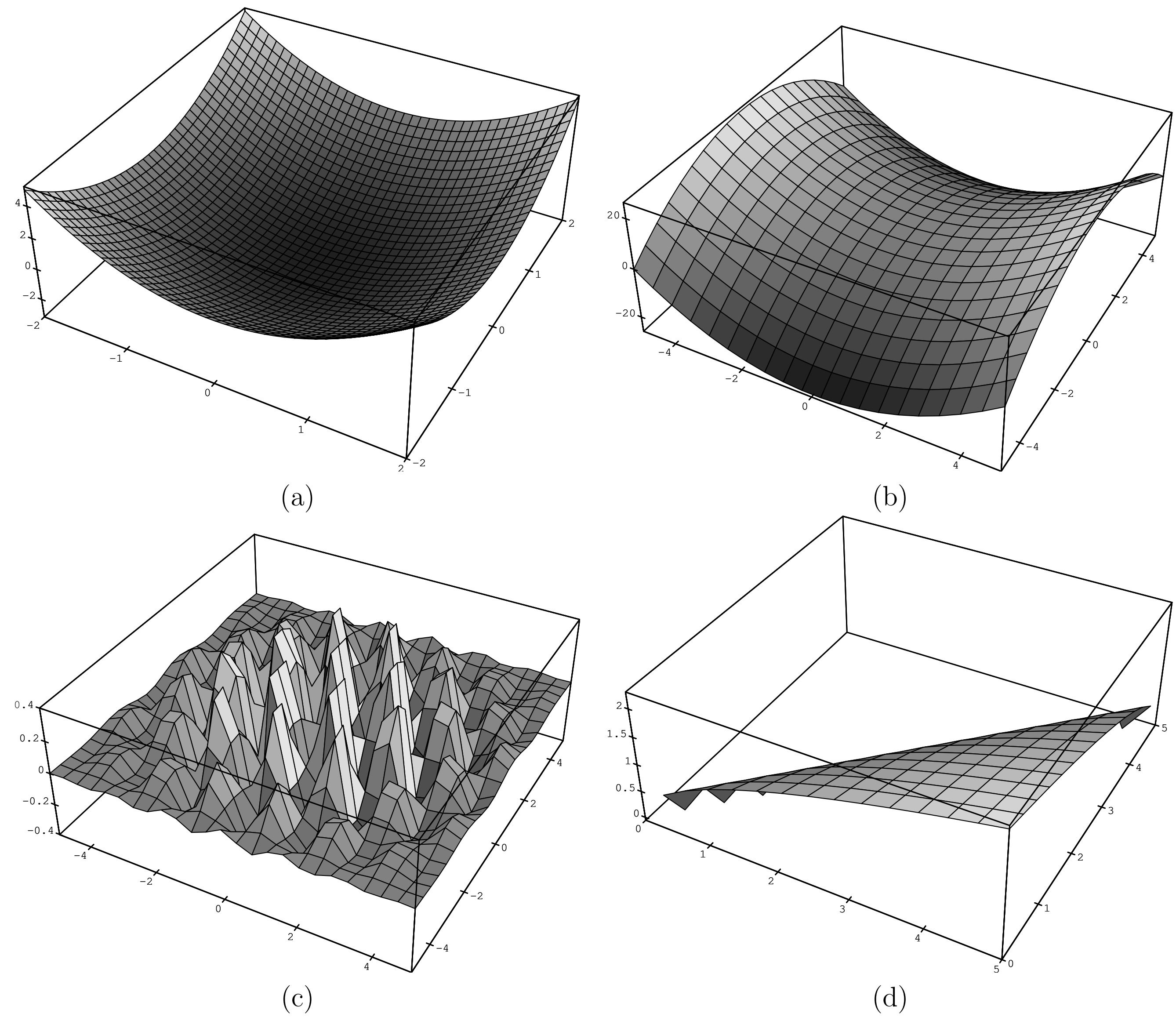

Figure 1.2: Surface plots of (a) $x^2 + y^2 - 3$, (b) $(x + y)(x - y)$, (c) $\exp[-(x^2 + y^2)/10] \sin(2x) \cos(4y)$, and (d) $(x - y)^{1/2}$.

1.2.1 Continuity for functions of several variables

A function $z = f(x, y)$, defined in some region E, is continuous at a point $(x, y) = (a, b)$ in E if, for every positive number ϵ (however small), it is possible to find a positive δ such that for all points in the neighbourhood defined by

$$(x-a)^2 + (y-b)^2 < \delta^2 \tag{1.5}$$

we have

$$|f(x, y) - f(a, b)| < \epsilon. \tag{1.6}$$

Or equivalently, as before,

$$\lim_{(x,y)\longrightarrow(a,b)} f(x, y) = f(a, b). \tag{1.7}$$

Note that for functions of more variables $f(x_1, x_2, x_3, ...)$ the neighbourhood would be defined by $(x_1 - a)^2 + (x_2 - b)^2 + (x_3 - c)^2 + \ldots < \delta^2$.

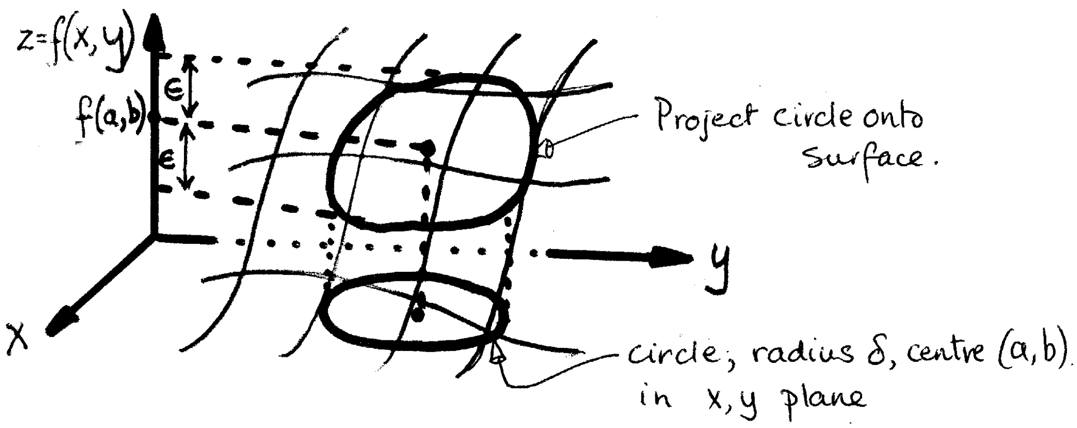

Figure 1.3: Neighbourhood for continuity for a function of 2 variables

1.3 The partial derivative

The extension of the idea of continuity to functions of several variables was direct. Extending the notion of the derivative is not quite as simple — the "slope" of the $f(x, y)$ surface at (x, y) depends on which direction you move off in. So we have to think about slope in a particular direction. The obvious directions are those along the $x-$ and $y-$axes.

Now, **if one wants to move off from (x, y) in the x direction one has to keep y fixed**. This is the key to defining the *partial derivative* of the function with respect to x:

$$\left(\frac{\partial f}{\partial x}\right)_y = f_x = \lim_{\delta x\longrightarrow 0}\left[\frac{f(x+\delta x, y) - f(x, y)}{\delta x}\right] \tag{1.8}$$

The subscript y indicates that y is being kept constant. If we are dealing with a function of more variables, we keep all but the one variable constant. Eg for $f(x_1, x_2, x_3, ...)$ we have

$$\begin{aligned} f_{x_3} &= \left(\frac{\partial f}{\partial x_3}\right)_{x_1 x_2 x_4 \ldots} \qquad (1.9) \\ &= \lim_{\delta x_3 \longrightarrow 0} \left[\frac{f(x_1, x_2, x_3 + \delta x_3, x_4, \ldots) - f(x_1, x_2, x_3, x_4, \ldots}{\delta x_3}\right] \end{aligned}$$

Given a list of the variables and the one being varied, the "held constant" subscripts are superfluous and are often omitted. Leave them in there until you are *au fait* with the techniques.

1.3.1 Geometrical interpretation of the pd

Figure 1.4 shows the geometrical interpretation of the partial derivatives of a function of two variables.

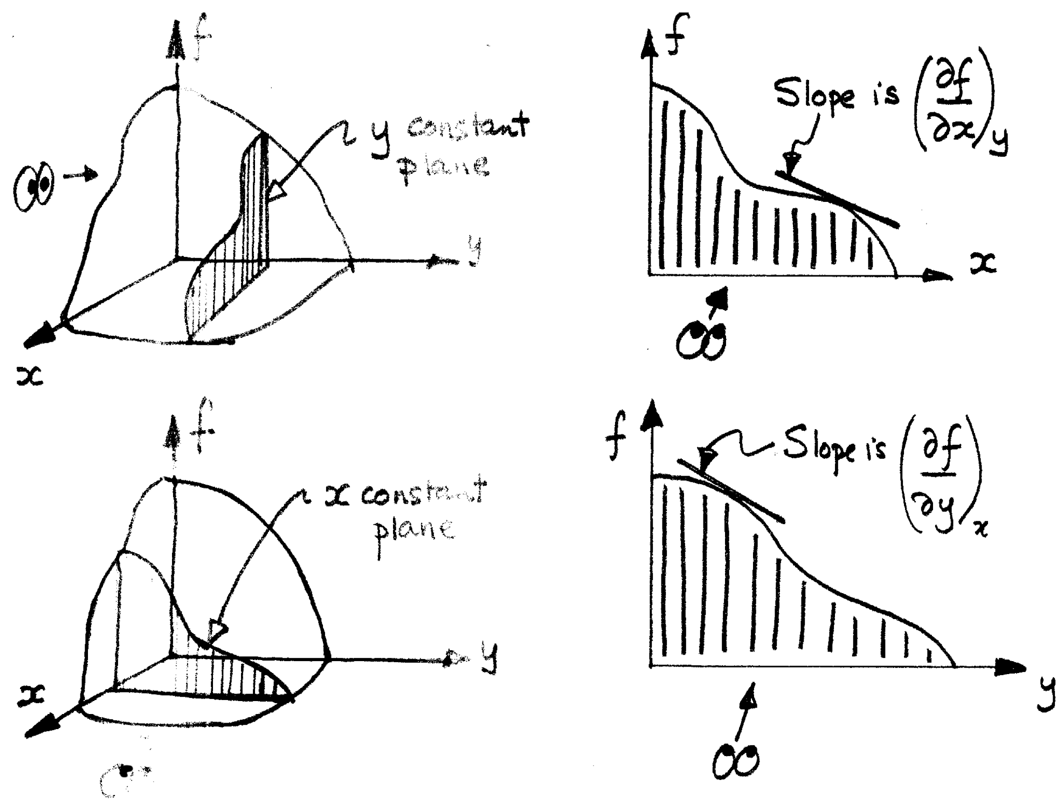

Figure 1.4: Interpreting partial derivatives as the slopes of slices through the function

1.3.2 The mechanics of evaluating partial derivatives

The definition of the partial derivative indicates that operationally partial differentiation is exactly the same as normal differentiation with respect to one variable, with all the others treated as constants.

♣ **Examples.**

1. Suppose

$$f(x,y) = x^2y^3 - 2y^2 \tag{1.10}$$

First assume y is a constant:

$$f_x = 2xy^3 \tag{1.11}$$

Then x is a constant:

$$f_y = 3x^2y^2 - 4y \tag{1.12}$$

2.

$$\begin{aligned} f(x,y) &= e^{-(x^2+y^2)}\sin(xy^2) && (1.13)\\ \Rightarrow f_x &= e^{-(x^2+y^2)}[-2x\sin(xy^2) + y^2\cos(xy^2)] && (1.14)\\ f_y &= e^{-(x^2+y^2)}[-2y\sin(xy^2) + 2xy\cos(xy^2)] && (1.15) \end{aligned}$$

3. If $f(x,y) = \ln(xy)$, derive an expression for $f_x f_y$ in terms of f.

$$\begin{aligned} f(x,y) &= \ln(x) + \ln(y) && (1.16)\\ \Rightarrow f_x &= 1/x && (1.17)\\ f_y &= 1/y && (1.18)\\ \Rightarrow f_x f_y &= 1/xy = e^{-f(x,y)}\ . && (1.19) \end{aligned}$$

1.3.3 Higher partial derivatives?

Why not? f_x and f_y are (well, probably are) perfectly good functions of (x,y). An example of the notation used is:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} f_x = f_{xx} \tag{1.20}$$

♣ **Example**.

$$\begin{aligned} f(x,y) &= x^2y^3 - 2y^2 && (1.21)\\ f_x &= 2xy^3 && (1.22)\\ f_y &= 3x^2y^2 - 4y && (1.23)\\ f_{xx} &= 2y^3 && (1.24)\\ f_{yy} &= 6x^2y - 4\ . && (1.25) \end{aligned}$$

But we should also consider

$$\frac{\partial}{\partial y} f_x = f_{yx} = \frac{\partial^2 f}{\partial y \partial x}\ :\ \text{in this case, } 6xy^2 \tag{1.26}$$

and

$$\frac{\partial}{\partial x} f_y = f_{xy} = \frac{\partial^2 f}{\partial x \partial y}\ :\ \text{in this case, } 6xy^2. \tag{1.27}$$

Thus in this case $f_{xy} = f_{yx}$ — is that always true?
♣ **A meaner example**:

$$\begin{aligned}
f(x,y) &= e^{-(x^2+y^2)}\sin(xy^2) && (1.28)\\
f_x &= e^{-(x^2+y^2)}[-2x\sin(xy^2)+y^2\cos(xy^2)] && (1.29)\\
f_y &= e^{-(x^2+y^2)}[-2y\sin(xy^2)+2xy\cos(xy^2)] && (1.30)\\
& && (1.31)\\
\Rightarrow\ f_{yx} &= e^{-(x^2+y^2)}[-4x^2y\cos()+2y\cos()-2y^3x\sin() && (1.32)\\
&\quad -\ 2y[-2x\sin()+y^2\cos()]]\\
&= e^{-(x^2+y^2)}[\sin()[-2y^3x+4xy]+\cos()[-4x^2y+2y-2y^3]]\\
\text{and}\ \ f_{xy} &= e^{-(x^2+y^2)}[-2y^3\cos()+2y\cos()-2xy^3\sin() && (1.33)\\
&\quad -\ 2x[-2y\sin()+2xy\cos()]]\\
&= e^{-(x^2+y^2)}[\sin()[-2xy^3+4xy]+\cos()[-2y^3+2y-4x^2y]]
\end{aligned}$$

So they are equal in this case too.
In fact,

> when both sides exist, and are continuous at the point of interest, then the operators are equivalent. Ie:
>
> $$\frac{\partial^2}{\partial x\partial y} = \frac{\partial^2}{\partial y\partial x} \tag{1.34}$$

This result has an interesting consequence for higher partial derivatives.
♣ **Example**. Show that

$$\frac{\partial^3 f}{\partial^2 x\partial y}\left[\frac{\partial^3 f}{\partial y\partial x^2}\right]^4 - \left[\frac{\partial^3 f}{\partial x\partial y\partial x}\right]^5 = 0 \tag{1.35}$$

for an appropriate function f. Now using ordering result

$$\frac{\partial^3 f}{\partial^2 x\partial y} = \frac{\partial}{\partial x}\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial^3 f}{\partial x\partial y\partial x} = \frac{\partial^3 f}{\partial y\partial x\partial x}\ , \tag{1.36}$$

and thus the equation is $p.p^4 - p^5$ — which is indeed zero.

So the order of higher partials is unimportant — but a sensible ordering can save time!
♣ **Example**. Find

$$\frac{\partial^3}{\partial t\partial y\partial x}\left[(y^5+xy)\cosh(\cosh(x^2+1/x))+y^2tx\right]. \tag{1.37}$$

Silly Method. Grind away blindly differentiating with respect to x then y then t. This may take a fortnight because the functions of x and y are moderately unpleasant. You will be cross when you find that the result of $2y$ could have been obtained much more quickly by differentiating with respect to t first. (Of course this result is obvious enough if you had thought ahead and noticed that the first expression was some $f(x,y)$, ie independent of t. Now $\partial^2 f(x,y)/\partial y\partial x$ must be some other $g(x,y)$ and thus $\partial g/\partial t$ must be zero.)

1.3.4 A Warning. Partial derivatives are not fraction-like

Although one should be careful about thinking of total derivatives in terms of fractions, they do have fraction-like qualities. It is worth stressing early on that one must be much more cautious with partial derivatives.

♣ **Example**. Suppose we were given $y = u^{1/3}$, $u = v^3$ and $v = x^2$ and asked to find dy/dx. We would write

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dv}\frac{dv}{dx} = \frac{1}{3}u^{-2/3}.3v^2.2x = 2x \tag{1.38}$$

which you can check by finding $y = x^2$ explicitly.

But! Suppose we were given the perfect gas law $pV = RT$ and asked "what is $\frac{\partial p}{\partial V}.\frac{\partial V}{\partial T}.\frac{\partial T}{\partial p}$?" Would you answer +1?

♣ **Example**. If $pV = RT$ then $p = RT/V$, $V = RT/p$ and $T = Vp/R$. Thus

$$\frac{\partial p}{\partial V}.\frac{\partial V}{\partial T}.\frac{\partial T}{\partial p} = \frac{-RT}{V^2}\frac{R}{p}\frac{V}{R} = -1. \tag{1.39}$$

Alas, not +1, as one might have guessed.

In fact, we will be able to show after studying implicit functions that if we have any function $f(x, y, z) = 0$, then $\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -1$.

1.4 Total and partial differentials

First, note that a *differential* is a different beastie from a *derivative*.

Differentials are about the following. Suppose that we have a continuous function $f(x, y)$ in some region, and both f_x and f_y are continuous in that region. How much does the value of the function change as one moves infinitesimal amounts dx and dy in the $x-$ and $y-$directions? The amount, df, is the total differential — how do we express it?

Given small changes in x and y it is easy enough to write the change in f as:

$$\delta f = f(x + \delta x, y + \delta y) - f(x, y) \ . \tag{1.40}$$

By adding in two cancelling terms this can be rewritten as

$$\delta f = [f(x + \delta x, y + \delta y) - f(x, y + \delta y)] + [f(x, y + \delta y) - f(x, y)] \ . \tag{1.41}$$

But recall that

$$f_x = \lim_{\delta x \to 0}\left[\frac{f(x + \delta x, y) - f(x, y)}{\delta x}\right] \tag{1.42}$$

so that

$$\delta f = [\frac{\partial}{\partial x}f(x, y + \delta y)\delta x + \alpha\delta x] + [\frac{\partial}{\partial y}f(x, y)\delta y + \beta\delta y] \tag{1.43}$$

where

$$\lim_{\delta x \to 0} \alpha = 0 \ ; \qquad \lim_{\delta y \to 0} \beta = 0. \tag{1.44}$$

(Note that the α and β are rather like constants of integration, which vanish as we take the differential to the limit.) Also

$$\frac{\partial}{\partial x} f(x, y+\delta y) = \frac{\partial}{\partial x} f(x,y) + \gamma \tag{1.45}$$

where

$$\lim_{\delta y \to 0} \gamma = 0 \tag{1.46}$$

So,

$$\delta f = \frac{\partial f}{\partial x}\delta x + \gamma \delta x + \alpha \delta x + \frac{\partial f}{\partial y}\delta y + \beta \delta y \quad , \tag{1.47}$$

and in the limit

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \tag{1.48}$$

where we neglect $(\alpha + \gamma)\delta x$ and $\beta \delta y$, which go to zero as $O(\delta x^2)$, $O(\delta x \delta y)$ and $O(\delta y^2)$.

The *total differential* df is the sum of the *partial differentials* $\frac{\partial f}{\partial x} dx$ and $\frac{\partial f}{\partial y} dy$.

Note that our expression 1.48 is an exact one for df *in the limit as* δx *and* δy *tend to zero.* Later on we will develop Taylor's expansion for a function of two or more variables and will see a better approximation for δf, better that is than

$$\delta f = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y \quad . \tag{1.49}$$

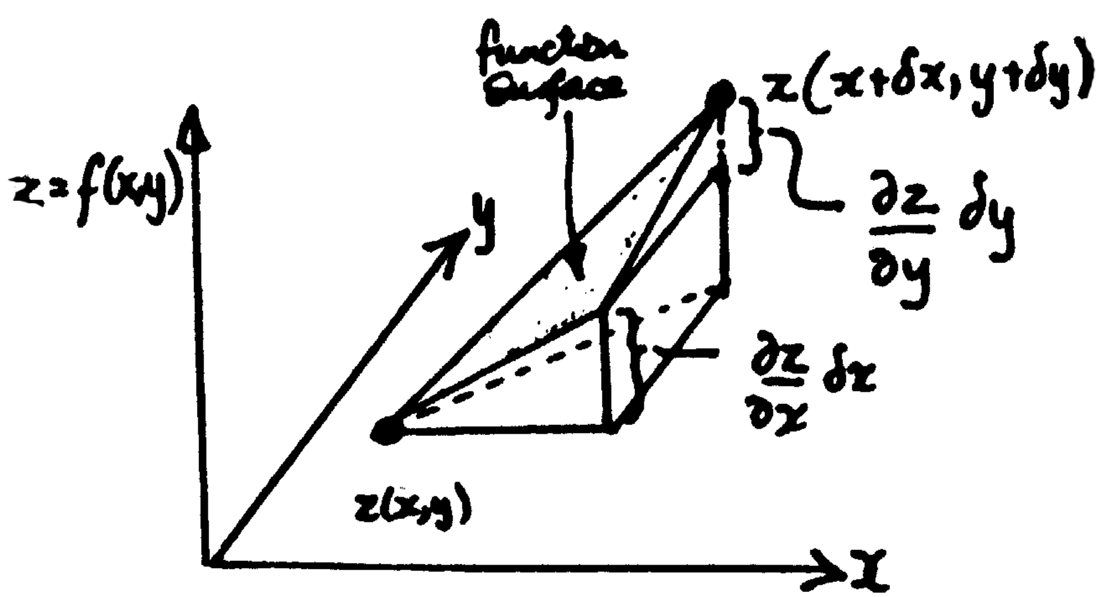

Figure 1.5: Total differential as the sum of partial differentials

1.4.1 Using the total differential

♣ **Example.** A material with a temperature coefficient of α is made into a block of sides x, y, z measured at some temperature T. The temperature is raised by ΔT. Derive the new volume of the block (a) exactly and (b) using the total differential.

(a) Exactly:

$$V + dV = x(1+\alpha\Delta T)y(1+\alpha\Delta T)z(1+\alpha\Delta T) = V(1+\alpha\Delta T)^3 \ . \tag{1.50}$$

(b) The volume of the block is $V = xyz$. Using the expression for the total differential

$$\begin{aligned} dV &= yzdx + xzdy + xydz & (1.51)\\ &= yzx(\alpha\Delta T) + xzy(\alpha\Delta T) + xyz(\alpha\Delta T) & (1.52)\\ &= 3V\alpha\Delta T & (1.53)\end{aligned}$$

Thus

$$V + dV = V(1+3\alpha\Delta T). \tag{1.54}$$

Note that the answer using the total differential is to the first order in small quantities. The exact expression

$$\begin{aligned} (a) &= V(1+3\alpha\Delta T+3(\alpha\Delta T)^2+(\alpha\Delta T)^3) & (1.55)\\ &\approx V(1+3\alpha\Delta T) = (b) & (1.56)\end{aligned}$$

1.4.2 When is an expression a total differential?

Suppose we are given some expression $p(x,y)dx + q(x,y)dy$. Can we deetrmine if it is the total differential of some function $f(x,y)$?

Now, if it is,

$$df = p(x,y)dx + q(x,y)dy \ . \tag{1.57}$$

But then we must have that $p(x,y) = f_x$ and $q(x,y) = f_y$; and using $\partial^2/\partial x\partial y = \partial^2/\partial y\partial x$ the condition for the total differential is simply

$$p_y = q_x \ . \tag{1.58}$$

♣ **Example.** Show that there is no function having continous second partial derivatives whose total differential is $xydx + 2x^2dy$.

Set $p = xy$ and $q = 2x^2$. Then $p_y = x \neq q_x = 4x$. Hence there is no such function.

1.4.3 Recovering the function from its total differential

Suppose we found $p(x,y)dx + q(x,y)dy$ to be total differential using the above test. Could we recover the function f? To recover f we must perform the reverse of partial differentiation. As $f_x = p(x,y)$:

$$f = \int p(x,y)dx + g(y) + K_1 \tag{1.59}$$

where g is a function of y alone and K_1 is a constant. You can see that we need the $g(y)$ because when we differentiate with respect to x it vanishes. As far as x is concerned, it is a constant of integration. Similarly,

$$f = \int q(x, y)dy + h(x) + K_2 \tag{1.60}$$

We now need to resolve the two expressions for f, and this is possible, up to a constant K, as the following example shows.

♣ **Example.** Let $f = xy^3 + \sin x \sin y + 6y + 10$ — but pretend we do not know it. Instead we are asked whether

$$(y^3 + \cos x \sin y)dx + (3xy^2 + \sin x \cos y + 6)dy$$

is a perfect differential and, if it is, of what function f?
Using the test just described we find that

$$\frac{\partial}{\partial y}(y^3 + \cos x \sin y) = 3y^2 + \cos x \sin y = \frac{\partial}{\partial x}(3xy^2 + \sin x \cos y + 6)$$

so it is a perfect differential. Integrating $(y^3 + \cos x \sin y)$ over x and $(3xy^2 + \sin x \cos y + 6)$ over y we find:

$$f = y^3 x + \sin x \sin y + g(y) + K_1$$

and

$$f = xy^3 + \sin x \sin y + 6y + h(x) + K_2 \quad .$$

Comparing and resolving these expressions we have $g(y) = 6y$ and $h(x) = 0$ and $K_1 = K_2$. Thus

$$f = xy^3 + \sin x \sin y + 6y + K_1 \quad .$$

Thus we have indeed recovered the original function, up to a constant of integration. We would need some extra piece of information to recover this — say the value of the function at a particular point.

This page is blank intentionally.

Chapter 2

Partial derivatives and function character

We now consider a whole series of relationships which can be derived if the function of several variables has a certain form. There is a danger that the various cases will become an unconnected jumble. It seems helpful to always ask the question "of what is f a function, exactly?"

2.1 A function of a function. $f = f(u)$ where $u = u(x, y)$.

Suppose

$$f(x,y) = xy\sin(xy) \quad . \tag{2.1}$$

We could find f_x etc in the usual way as

$$f_x = y\sin(xy) + xy^2\cos(xy) \tag{2.2}$$
$$f_y = x\sin(xy) + x^2 y\cos(xy) \tag{2.3}$$

BUT we might notice that $f = f(u) = u\sin u$ where $u = xy$. So f is a function of a single variable, so df/du exists, and u is in turn is a function of more than one variable. Then

$$\frac{\partial f}{\partial x} = \frac{df}{du}\frac{\partial u}{\partial x} \tag{2.4}$$
$$\frac{\partial f}{\partial y} = \frac{df}{du}\frac{\partial u}{\partial y} \tag{2.5}$$

which you can easily check gives the same result for the given example.
We can prove this result as follows. We know that

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy, \text{and} \tag{2.6}$$
$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy \quad . \tag{2.7}$$

If we consider df and du when y is fixed we would have

$$(df)_y = \frac{\partial f}{\partial x}dx \tag{2.8}$$
$$(du)_y = \frac{\partial u}{\partial x}dx \tag{2.9}$$

But the following must be true,

$$(df)_y = \frac{df}{du}(du)_y \quad , \tag{2.10}$$

so that

$$\frac{df}{du} = \frac{\partial f/\partial x}{\partial u/\partial x} \ . \tag{2.11}$$

Similarly by fixing x we would have found

$$\frac{df}{du} = \frac{\partial f/\partial y}{\partial u/\partial y} \ . \tag{2.12}$$

The result quoted follows immediately.

♣ **Examples**.

1. Find f_x and f_y when $f = \tan^{-1}(y/x)$.

   Set $f = \tan^{-1}(u)$ with $u = y/x$. Then

$$\frac{df}{du} = \frac{x^2}{x^2+y^2}; \quad u_x = -y/x^2; \quad u_y = 1/x \tag{2.13}$$

$$\Rightarrow f_x = \frac{-y}{x^2+y^2}; \quad f_y = \frac{x}{x^2+y^2} \tag{2.14}$$

2. This is a rather harder example. Show that $z = f(x-ct) + g(x+ct)$ where c is a constant is a solution of the wave equation $z_{xx} = \frac{1}{c^2} z_{tt}$.

   Write $u = x - ct$ and $v = x + ct$, so that $z = f(u) + g(v)$. Then note that z is the sum of two functions, each of a single variable, so that

$$\frac{\partial z}{\partial t} = \frac{df}{du}\frac{\partial u}{\partial t} + \frac{dg}{dv}\frac{\partial v}{\partial t} \tag{2.15}$$

$$= \frac{df}{du}(-c) + \frac{dg}{dv}(c) \quad . \tag{2.16}$$

   But because df/du is another function of u, $df/du = \xi(u)$ say, and because df/dv a function of v, $dg/dv = \eta(u)$ say, we have

$$\frac{\partial^2 z}{\partial t^2} = -c\frac{d^2 f}{du^2}\frac{\partial u}{\partial t} + c\frac{d^2 g}{dv^2}\frac{\partial v}{\partial t} \tag{2.17}$$

$$= c^2\left(\frac{d^2 f}{du^2} + \frac{d^2 g}{dv^2}\right) \tag{2.18}$$

   Similarly

$$\frac{\partial z}{\partial x} = \frac{df}{du}\frac{\partial u}{\partial x} + \frac{dg}{dv}\frac{\partial v}{\partial x} \tag{2.19}$$

$$= \frac{df}{du} + \frac{dg}{dv} \tag{2.20}$$

and so

$$\frac{\partial^2 z}{\partial x^2} = \frac{d^2 f}{du^2}\frac{\partial u}{\partial x} + \frac{d^2 g}{dv^2}\frac{\partial v}{\partial x} \tag{2.21}$$

$$= \left(\frac{d^2 f}{du^2} + \frac{d^2 g}{dv^2}\right) \tag{2.22}$$

$$= \frac{1}{c^2}\frac{\partial^2 z}{\partial t^2}. \tag{2.23}$$

Note that we have not had to say anything about the functions f and g, and you might care to check the result for any pair of arbitrary functions.

2.2 Composite functions and the Chain Rule

There are various cases of composite functions, and we deal with them in turn. This section will also introduce the chain rule, which you will use over and over again.

2.2.1 What if $z = f(x, y)$ and $x = x(t)$ and $y = y(t)$?

Function f is said to be a *composite function*. Note that f is effectively a function of t alone, so that total df/dt exists.

There is an important rule, the **Chain Rule**, which states that for such functions the total derivative is given by

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \quad . \tag{2.24}$$

To prove this, recall the expression for the total differential (before the limit was taken):

$$\delta f = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + (\alpha + \gamma)\delta x + \beta\delta y \ . \tag{2.25}$$

Because $x = x(t)$ and $y = y(t)$, as $\delta t \to 0$ we find that $\delta x \to 0$, $\delta y \to 0$, $(\alpha + \gamma) \to 0$ and $\beta \to 0$. Now, divide through by δt

$$\frac{\delta f}{\delta t} = \frac{\partial f}{\partial x}\frac{\delta x}{\delta t} + \frac{\partial f}{\partial y}\frac{\delta y}{\delta t} + \alpha\frac{\delta x}{\delta t} + (\beta + \gamma)\frac{\delta y}{\delta t} \tag{2.26}$$

and then take the limit $\delta t \to 0$

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \qquad [+\{\alpha\frac{dx}{dt} + (\beta + \gamma)\frac{dy}{dt}\} = 0] \tag{2.27}$$

which is the required result.

Now if we have a function of many variables $f(x_1, x_2, x_3, \ldots, x_n)$ with $x_i = x_i(t)$ then the Chain Rule becomes:

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \ldots \tag{2.28}$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{dx_i}{dt} \quad . \tag{2.29}$$

♣ **Example**. Find du/dt when

$$u = u(x, y, z) = xy + yz + zx; \quad x = t, y = e^{-t}, z = \cos t \tag{2.30}$$

Thus

$$\frac{du}{dt} = \frac{\partial u}{\partial y}\frac{dy}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} + \frac{\partial u}{\partial z}\frac{dz}{dt} + \tag{2.31}$$

$$= (y+z)\frac{dx}{dt} + (x+z)\frac{dy}{dt} + (x+y)\frac{dz}{dt} \tag{2.32}$$

$$= e^{-t}(1 - t - \cos t - \sin t) - t\sin t. \tag{2.33}$$

2.2.2 What if $z = f(x, y)$ and $x = x(t_1, t_2, ...)$ and $y = y(t_1, t_2, ...)$?

This is like the previous composite function, but now f is effectively a function of several variables $t_1, t_2, \ldots$. Suppose we fix all but one of the t_i. We would end up with a composite function as above, but now instead of *total derivatives* df/dt and dx_i/dt we must have *partial derivatives* $\partial f/\partial t_j$ and $\partial x_i/\partial t_j$ Thus

$$\frac{\partial f}{\partial t_j} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t_j} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t_j}\ , \qquad j = 1, 2, ... \tag{2.34}$$

We can now easily generalize this result to a function $f(x_1, x_2, x_3, ..., x_n)$ with $x_i = x_i(t_1, t_2, ..., t_m)$. After chomping your way through the indices you should find that

$$\frac{\partial f}{\partial t_j} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t_j} \ : j = 1, \ldots, m. \tag{2.35}$$

It turns out that the *chain rule for partials* is very commonly used in transforming from one set of coordinates to another, and we shall return to it again.

♣ **Example**. $x = r\cos\phi$ and $y = r\sin\phi$ defines the transformation between Cartesian and plane polar coordinates. Find f_r and f_ϕ when $f(x, y) = x^2 + y^2$.

Using the chain rule for partials:

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} \tag{2.36}$$

$$= 2x\cos\phi + 2y\sin\phi = 2r \tag{2.37}$$

$$\frac{\partial f}{\partial \phi} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \phi} \tag{2.38}$$

$$= 2x(-r\sin\phi) + 2y(r\cos\phi) = 0 \tag{2.39}$$

(Because $f = x^2 + y^2 = r^2$, this can be checked directly! As a bonus, note that $df = \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial \phi}d\phi = 2rdr$, which is again consistent.)

2.2.3 What if $z = f(x, y)$ and $y = y(x)$?

Clearly z is a composite function of x alone. So this is like t being x. (Effectively $x = x(x)$, but one wouldn't bother to write this down.)

So,

$$\begin{aligned} \frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ \downarrow & \quad\;\; \downarrow \;\; \downarrow \qquad \downarrow \;\; \downarrow \\ \frac{df}{dx} &= \frac{\partial f}{\partial x}[1] + \frac{\partial f}{\partial y}\frac{dy}{dx} \end{aligned} \tag{2.40}$$

Ie,

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} \quad . \tag{2.41}$$

♣ **Example**. Suppose $z = xy + x/y$ and $y = \sqrt{x}$. Find dz/dx.

Using the results just obtained,

$$\frac{dz}{dx} = y + \frac{1}{y} + x(1 - \frac{1}{y^2})\frac{1}{2}x^{-1/2} \tag{2.42}$$

$$= x^{1/2} + x^{-1/2} + \frac{1}{2}\left(x^{1/2} - x^{-1/2}\right) \tag{2.43}$$

$$= \frac{3}{2}x^{1/2} + \frac{1}{2}x^{-1/2} \tag{2.44}$$

a result which can be checked directly in this case, using $z = x^{3/2} + x^{1/2}$.

This result leads us on to discuss an important class of functions, *implicit functions*.

2.3 Implicit Functions

In calculus and analytic geometry it is common to see the equations of plane curves written as

$$f(x, y) = 0. \tag{2.45}$$

For example the equation of a circle of radius $\sqrt{2}$ with centre at $(0, 1)$ is

$$x^2 + (y - 1)^2 - 2 = 0 \tag{2.46}$$

In this case it is easy to write an equivalent equation $y = y(x) = 1 \pm (2 - x^2)^{1/2}$ — but often it is the case that although one can write down $f(x, y) = 0$ one *cannot* solve for $y = y(x)$ *explicitly*. But if we were to solve for y numerically we might trace out a curve $y = y(x)$ which was single valued and differentiable. In other words, $f(x, y) = 0$ may define a $y = y(x)$ *implicitly*. Such a $y = y(x)$ is an *implicit function*.

It turns out that for many purposes we don't need an explicit expression for the function. It is enough to characterize the function by its derivatives. So how do we compute them?

2.3.1 Some cases of implicit functions

At the end of §2.2.3 we saw that for $z = f(x, y)$ with $y = y(x)$,

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}. \tag{2.47}$$

Now suppose that we have $f(x, y) = 0$. We have just argued that this gives us an implicit $y = y(x)$, so our previous result should hold. Indeed it does, but now we have $z = 0$ *always*, so that $dz/dx = 0$ *always*. Hence:

$$0 = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx}. \tag{2.48}$$

so that:

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y} \tag{2.49}$$

(provided $\partial f/\partial y$ is not zero at the point of interest).
♣ **Examples**.

1. First consider an example where the implicit function can be derived explicitly. This gives us a check on our results. Find dy/dx when $f(x, y) = x - x^2y^3 = 0$.

   First use the result for implicit functions just derived:

   $$\frac{dy}{dx} = -\left(\frac{1 - 2xy^3}{-3x^2y^2}\right). \tag{2.50}$$

   Because we can find $y = y(x)$ explicitly, we can substitute for y and so $\frac{dy}{dx} = -x^{-4/3}/3$.

   As a check, find y explicitly. Obviously $y = x^{-1/3}$, so that $\frac{dy}{dx} = -x^{-4/3}/3$. So the result is sound.

   But remember! The whole point about implicit functions is that you do not need an explicit $y = y(x)$ to get information about the derivatives. This is made clear in the next example.

2. Find dy/dx when $f(x, y) = ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$.

   This f is an conic. Now we can't find y as a function of x, but using the result for implicit functions we have

   $$\frac{dy}{dx} = -\frac{2ax + 2hy + 2g}{2by + 2hx + 2f}. \tag{2.51}$$

2.3.2 More on implicit functions

With a function of two variables, one equation $f(x, y) = 0$ was sufficient to determine $y = y(x)$. Suppose now that we have functions of 3 variables. Two of these, say

$$f(x, y, z) = 0 \qquad g(x, y, z) = 0 \tag{2.52}$$

are required to define implicitly $y = y(x)$ and $z = z(x)$. (Think about simultaneous equations. From the $f = 0$ we get $y = y(x, z)$. From the $g = 0$ we get $z = z(x, y)$. Thus we have $y = y(x, z(x, y))$ or $y = y(x)$. Putting this in the expression for z we get $z = z(x, y(x))$ or $z = z(x)$.)

Now if we use the chain rule:

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx} = 0 \tag{2.53}$$

and

$$\frac{dg}{dx} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y}\frac{dy}{dx} + \frac{\partial g}{\partial z}\frac{dz}{dx} = 0\ . \tag{2.54}$$

So we have two simultaneous equations in dy/dx and dz/dx which can be solved

$$\frac{dy}{dx} = \left(\frac{\partial f}{\partial z}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial g}{\partial z}\right) / \left(\frac{\partial f}{\partial y}\frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial y}\right) \tag{2.55}$$

and

$$\frac{dz}{dx} = \left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right) / \left(\frac{\partial f}{\partial y}\frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial y}\right). \tag{2.56}$$

♣ **Example.**

1. First try an example where the implicit functions can in fact be derived explicitly. Find dy/dx and dz/dx when

$$f(x, y, z) = x + y + z = 0 \qquad g(x, y, z) = x - y + 2z = 0. \tag{2.57}$$

   Using the result obtained above: $f_x = f_y = f_z = 1$ and $g_x = 1$, $g_y = -1$ and $g_z = 2$. Thus $\frac{dy}{dx} = -\frac{1}{3}$ and $\frac{dz}{dx} = -\frac{2}{3}$

   In this case we can solve explicitly and check the result. Clearly $2x + 3z = 0$ And $\Rightarrow dz/dx = -2/3$. Also $-x - 3y = 0$ And $\Rightarrow dy/dx = -1/3$. So the result is verified.

2.4 Differentiation of implicit functions

Suppose now that rather than having 2 functions of 3 variables, we have only one

$$f(x, y, z) = 0 \tag{2.58}$$

This may define $z = z(x, y)$ implicitly. (*May* because not every function will: for example $x^2 + y^2 + z^2 + 1 = 0$ is never satisfied for real x, y, z.)

But suppose we can find an (x_0, y_0, z_0) which satisfies $f(x_0, y_0, z_0) = 0$ and that, near this point, f and its first partial derivatives are continuous and $\partial f/\partial z = 0$. Then an existence theorem states that in the region around (x_0, y_0) there is precisely one differentiable function $z = \phi(x, y)$ which satisfies $f(x, y, z) = 0$ and is such that $z_0 = z(x_0, y_0)$.

Suppose now we fix y at $y = y_0$; then $f(x, y_0, z) = 0$. But this implicitly defines $z = z(x)$ in the neighbourhood of x_0.

So, with y fixed at y_0 we have $f(x, y_0, z) = 0$ and $z = z(x)$. Recall that earlier we saw that for $f(x, y) = 0, y = y(x)$ we showed that:

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y}. \tag{2.59}$$

Rewriting this for z instead we have

$$\frac{dz}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial z}. \tag{2.60}$$

This is nearly right — but not quite. We have fixed y, so that dz/dx should be a partial not a total derivative. Thus

$$\frac{\partial z}{\partial x} = -\frac{\partial f/\partial x}{\partial f/\partial z}. \tag{2.61}$$

We could reverse the roles of x and y, so that

$$\frac{\partial z}{\partial y} = -\frac{\partial f/\partial y}{\partial f/\partial z}. \tag{2.62}$$

Indeed, for a function of several variables $f = f(x_1, x_2, \ldots, x_n) = 0$ we have

$$\frac{\partial x_i}{\partial x_j} = -\frac{\partial f/\partial x_j}{\partial f/\partial x_i} \tag{2.63}$$

(provided $i \neq j$ and $\partial f/\partial x_i \neq 0$ at the point of interest).

♣ **Examples.**

1. This is one we can check by explicit evaluation. Find z_x and z_y when $f(x, y, z) = x + 2y^2 + 3e^{-z} = 0$.

   Using the expressions just derived we have:

   $$z_x = \frac{1}{3}e^z \qquad z_y = \frac{4}{3}ye^z. \tag{2.64}$$

   (As a check, write

   $$z = -\ln(-x - 2y^2) + \ln(3) \tag{2.65}$$

   Then

   $$\begin{aligned} z_x &= \frac{1}{-x-2y^2} = \frac{1}{3e^{-z}} = \frac{1}{3}e^z && (2.66)\\ z_y &= \frac{4y}{-x-2y^2} = \frac{4y}{3e^{-z}} = \frac{4}{3}ye^z && (2.67) \end{aligned}$$

2. If $x^2 - xy + z^2 + yz = 4$, find z_x and z_y at $(x, y) = (1, 3)$.

   It is worth giving two methods of solution, to verify the formulism given above.

   (a) Using

   $$\frac{\partial x_i}{\partial x_j} = -\frac{\partial f/\partial x_j}{\partial f/\partial x_i} \tag{2.68}$$

   we have

   $$z_x = -\frac{2x - z}{2z + y - x} \qquad z_y = \frac{-z}{2z + y - x}. \tag{2.69}$$

   When $x = 1, y = 3$, then $z^2 + 2z - 3 = 0$ and $z = 1, -3$.

   $$z = 1 : \qquad z_x = \frac{-1}{4}\ ,\ z_y = \frac{-1}{4} \tag{2.70}$$

   $$z = -3 : \qquad z_x = \frac{5}{4}\ ,\ z_y = \frac{-3}{4}. \tag{2.71}$$

   (b) This is a slightly different method of solution. $f(x, y, z) = x^2 - xy + z^2 + yz - 4 = 0$. Assume z is defined implicitly as $z = z(x, y)$. Differentiate wrt x and y respectively:

   $$2x - z - x\frac{\partial z}{\partial x} + 2z\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial x} = 0 \tag{2.72}$$

   $$-x\frac{\partial z}{\partial y} + 2z\frac{\partial z}{\partial y} + z + y\frac{\partial z}{\partial y} = 0 \tag{2.73}$$

   Hence

   $$z_x = \frac{z - 2x}{2z + y - x} \qquad z_y = \frac{-z}{2z + y - x}. \tag{2.74}$$

   Then carry on as before.

3. The final example goes back to the perfect gas law — now written as $f(p, V, T) = 0$. Recall that we asked what was the value of

   $$\frac{\partial p}{\partial V}\cdot\frac{\partial V}{\partial T}\cdot\frac{\partial T}{\partial p} \quad .$$

   Using the result just obtained we could now write:

   $$\frac{\partial p}{\partial V}\cdot\frac{\partial V}{\partial T}\cdot\frac{\partial T}{\partial p} = \left[-\frac{\frac{\partial f}{\partial V}}{\frac{\partial f}{\partial p}}\right]\left[-\frac{\frac{\partial f}{\partial T}}{\frac{\partial f}{\partial V}}\right]\left[-\frac{\frac{\partial f}{\partial p}}{\frac{\partial f}{\partial T}}\right] \tag{2.75}$$

   $$= -1 \quad . \tag{2.76}$$

   Note this result is *independent* of the form the perfect gas law.

2.5 Euler's Theorem

Finally, a theorem by Euler. This is it, *his theorem*.
If $u = f(x_1, x_2, x_3, \ldots)$ is a homogeneous function of degree n, then

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_3 \frac{\partial f}{\partial x_3} + \ldots = nf \quad . \tag{2.77}$$

The proof is as follows. First note that a homogeneous function of degree n is such that

$$f(\gamma x_1, \gamma x_2, \ldots) = \gamma^n f(x_1, x_2, \ldots) \tag{2.78}$$

Now Write $x_i = \gamma X_i,\ i = 1, 2, \ldots$. Then

$$u = f(\gamma X_1, \gamma X_2, \gamma X_3, \ldots) = \gamma^n f(X_1, X_2, X_3, \ldots) \tag{2.79}$$

Using the chain rule:

$$\frac{du}{d\gamma} = \frac{\partial f}{\partial x_1}\frac{dx_1}{d\gamma} + \frac{\partial f}{\partial x_2}\frac{dx_2}{d\gamma} + \ldots = n\gamma^{n-1} f(X_1, X_2, X_3, \ldots) \tag{2.80}$$

$$= \frac{\partial f}{\partial x_1} X_1 + \frac{\partial f}{\partial x_2} X_2 + \ldots \tag{2.81}$$

Ie

$$\frac{\partial f}{\partial x_1} X_1 + \frac{\partial f}{\partial x_2} X_2 + \ldots = n\gamma^{n-1} f(X_1, X_2, X_3, \ldots) \tag{2.82}$$

Multiply the lhs and rhs by γ and find

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + x_3 \frac{\partial f}{\partial x_3} + \ldots = nf \quad . \tag{2.83}$$

Chapter 3

Changing variables, transformations and Jacobians

In Lecture 1 it was pointed out that, because the slope of the $f(x, y)$ surface depends on the direction one moves in, the partial derivative must be defined in terms of the change in the function along a particular direction or axis. For a function $f(x, y)$, the obvious axes or directions to choose are the x and y axes, keeping y and x fixed, respectively.

The question now is *are these really the obvious directions or axes?*

Certainly given a function $f(x, y)$ it is *operationally* easiest to find $\partial f/\partial x$ and $\partial f/\partial y$, but these directions may not fit very well at all with the symmetry of the function being considered. For example, consider the function

$$f(x, y) = e^{-(x^2+y^2)} \cos(4(x^2 + y^2))$$

shown in the figure. Does it really make sense to impose a square "x-constant, y-constant" mesh onto this function?
Probably not! It would be in more sympathy with the function to use a mesh with radial symmetry. To effect this we need to make an appropriate transformation to a new set of variables. This raises the questions of

- How to choose the new variables and thus describe the transformation.
- How to describe the function in the new variables.
- How to find the partial derivatives with respect to these new variables.

3.1 Choosing the new variables

There is no standard grind for doing this, but usually the problem symmetry drops very large hints as to the transformation you might choose. Later on we will consider some of the most standard transformations, viz between Cartesian and

- Plane Polar Coordinates (2D): for radial symmetry
- Spherical Polar Coordinates (3D): for spherical symmetry, and

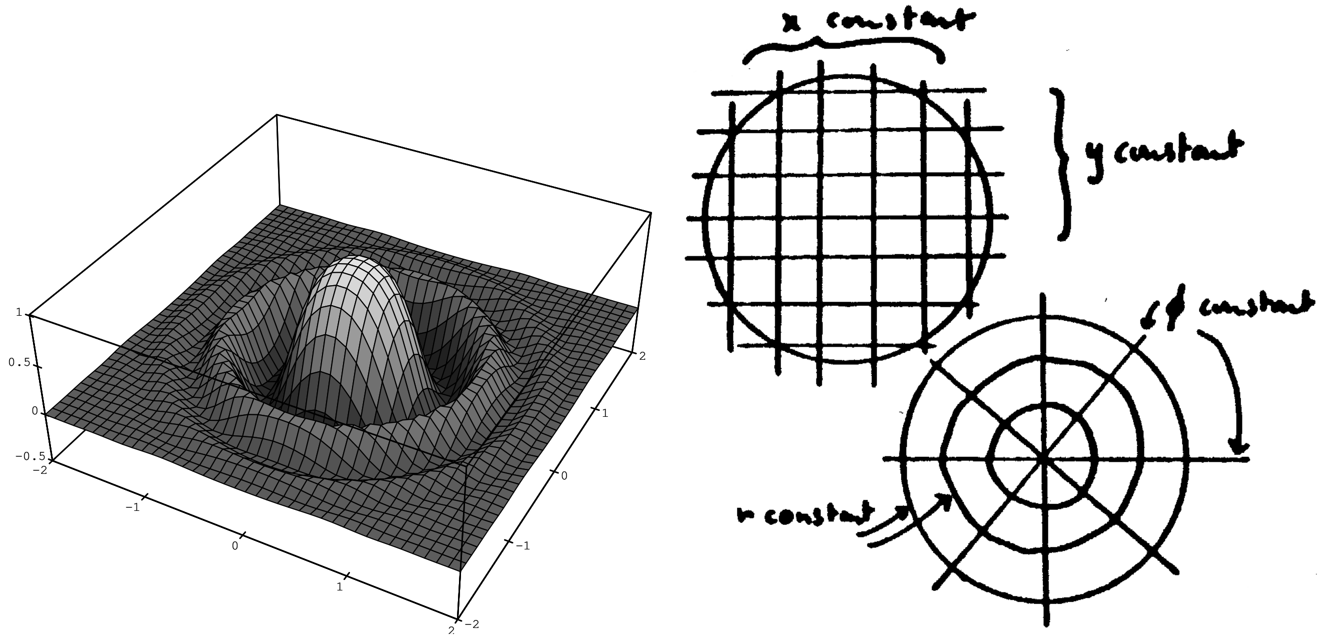

Figure 3.1: The function $\exp -(x^2+y^2)\cos(4(x^2+y^2))$. A radial mesh reflects the symmetry rather better than the x, y constant cake rack.

- Cylindrical Polars Coordinates (3D): for cylindrical symmetry.

We also consider working up a transformation from scratch.
First though we should consider the general problems of rewriting the function and finding the partial derivatives with respect to the new variables.

3.2 Rewriting the function and find the derivatives

We look at

1. the straightforward case: "old in terms of new", and
2. the more complicated case: "new in terms of old".

3.2.1 Case 1

We wish to transform the function $z = f(x, y)$ to (u, v) coordinates. Suppose the transformation is given as "old in terms of new" variables, that is:

$$x = x(u, v) \quad \text{and} \quad y = y(u, v) \quad . \tag{3.1}$$

If we actually know what the function is explicity (eg $f(x, y) = x/y + \sin(x)$ or whatever) it is easy to replace x and y in the function to get the new function as

$$z = f(x(u, v), y(u, v)) = F(u, v) \quad . \tag{3.2}$$

Then one can get the partial derivatives with respect to u and v directly from the new function $z = F(u, v)$.

But suppose we don't wish to consider the function $F(u, v)$ explicity. Recall (from lecture 2) that we can still get the partial derivatives using the Chain Rule for partials:

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} \quad (3.3)$$

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v} \quad (3.4)$$

Note carefully that Chain Rule for partials gives you the partial derivatives with respect to to u and v *without reference to the function* $F(u, v)$.

3.2.2 Case 2

We wish still wish to transform the function $z = f(x, y)$ to (u, v) coordinates. But now suppose the transformation is given as "new in terms of old" variables, that is:

$$u = u(x, y) \quad \text{and} \quad v = v(x, y) \quad . \quad (3.5)$$

To find $F(u, v)$ we might be able use the expressions for $u = u(x, y)$ and $v = v(x, y)$ as simultaneous equations from which we could find $x = x(u, v)$ and $y = y(u, v)$. (Eg if

$$u(x, y) = x + y \quad \text{and} \quad v(x, y) = x - y$$

then

$$x(u, v) = (u + v)/2 \quad \text{and} \quad y(u, v) = (u - v)/2 \quad .)$$

We could then go on to find the partials either directly or using the chain rule for partials just as Case 1 above.

But what happens if we cannot solve for $x = x(u, v)$ and $y = y(u, v)$? Obviously we cannot find the function $F(u, v)$ explicitly, but does that stop us find the partial derivatives?
Fortunately not. The Chain rule for partials comes to the rescue — but not from equations 3.3 because we cannot find $\partial x/\partial u$, $\partial y/\partial u$, $\partial x/\partial v$, and $\partial y/\partial v$.
What we *can* find are $\partial u/\partial x$, $\partial u/\partial y$, $\partial v/\partial x$, $\partial v/\partial y$; and $\partial z/\partial x$, $\partial z/\partial y$. So it makes sense to look at

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial x} \quad (3.6)$$

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial z}{\partial v}\frac{\partial v}{\partial y} \quad . \quad (3.7)$$

Now treat these equations as simultaneous equations for $\partial z/\partial u$ and $\partial z/\partial v$. Ie:

$$\frac{\partial z}{\partial u} = \left(\frac{\partial z}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial z}{\partial y}\frac{\partial v}{\partial x}\right) / \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right) \quad (3.8)$$

$$\frac{\partial z}{\partial v} = \left(\frac{\partial z}{\partial y}\frac{\partial u}{\partial x} - \frac{\partial z}{\partial x}\frac{\partial u}{\partial y}\right) / \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right) \quad (3.9)$$

So we can recover the partial derivatives with respect to the new variables without an explicit expression for the function in the new variables.

3.2.3 Summary

To summarize:

1. If you have the transformation as OLD variables in terms of NEW either
   (a) find the function explicity and find the partials from it directly, or
   (b) find the partials using the Chain-rule-for-partials directly.

2. If you have the transformation as NEW variables in terms of OLD either
   (a) use the given transformation as simultaneous equations for OLD variables in terms of NEW and use case 1, or
   (b) use the Chain-rule-for-partials for the OLD variables to give simultaneous equations for the partial derivatives with respect to to the NEW variables.

Note that Case 1a will fail if you do not know the function explicitly. Case 2a will fail if you cannot invert the transformation, or if you do not know the function explicitly.

♣ **Examples.**

1. Given a function $V = F(r, \phi)$, where $x = r\cos\phi, y = r\sin\phi$, find V_x and V_y.

   Here our transformation is NEW variables in terms of OLD. Also we know nothing explicit about the function so we must use the Chain rule for partials indirectly.

$$\frac{\partial V}{\partial r} = \frac{\partial V}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial r} \tag{3.10}$$

$$= \frac{\partial V}{\partial x}\cos\phi + \frac{\partial V}{\partial y}\sin\phi \tag{3.11}$$

$$\frac{\partial V}{\partial \phi} = \frac{\partial V}{\partial x}\frac{\partial x}{\partial \phi} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial \phi} \tag{3.12}$$

$$= -\frac{\partial V}{\partial x} r\sin\phi + \frac{\partial V}{\partial y} r\cos\phi \tag{3.13}$$

   Solving the simultaneous equations we have

$$\frac{\partial V}{\partial x} = \cos\phi\frac{\partial V}{\partial r} - \frac{\sin\phi}{r}\frac{\partial V}{\partial \phi} \tag{3.14}$$

$$\frac{\partial V}{\partial y} = \sin\phi\frac{\partial V}{\partial r} + \frac{\cos\phi}{r}\frac{\partial V}{\partial \phi} \tag{3.15}$$

2. Laplace's equation is $\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0$ where $V = V(x, y)$. Express this equation in plane polar coordinates (r, ϕ) where $x = r\cos\phi$, $y = r\sin\phi$.

The first part is as (1) above, but it is useful to express $\partial/\partial x$ etc as operators:

$$\frac{\partial}{\partial x} = \cos\phi\frac{\partial}{\partial r} - \frac{\sin\phi}{r}\frac{\partial}{\partial\phi} \tag{3.16}$$

$$\frac{\partial}{\partial y} = \sin\phi\frac{\partial}{\partial r} + \frac{\cos\phi}{r}\frac{\partial}{\partial\phi} \tag{3.17}$$

So

$$\frac{\partial^2}{\partial x^2} = \left(\cos\phi\frac{\partial}{\partial r} - \frac{\sin\phi}{r}\frac{\partial}{\partial\phi}\right)\left(\cos\phi\frac{\partial}{\partial r} - \frac{\sin\phi}{r}\frac{\partial}{\partial\phi}\right) \tag{3.18}$$

$$= \cos\phi^2\frac{\partial^2}{\partial r^2} - \cos\phi\sin\phi\left[-\frac{2}{r^2}\frac{\partial}{\partial\phi} + \frac{1}{r}\frac{\partial^2}{\partial\phi\partial r}\right] \tag{3.19}$$

$$-\frac{1}{r}\sin\phi\left[-\sin\phi\frac{\partial}{\partial r} + \cos\phi\frac{\partial^2}{\partial\phi\partial r}\right] + \frac{1}{r^2}\sin\phi^2\frac{\partial^2}{\partial\phi^2} \tag{3.20}$$

Note how the operators move through to the right, operating as they go.

A similar expression can be developed for y

$$\frac{\partial^2}{\partial y^2} = \sin\phi^2\frac{\partial^2}{\partial r^2} - \cos\phi\sin\phi\left[\frac{2}{r^2}\frac{\partial}{\partial\phi} - \frac{1}{r}\frac{\partial^2}{\partial\phi\partial r}\right] \tag{3.21}$$

$$+\frac{1}{r}\cos\phi\left[\cos\phi\frac{\partial}{\partial r} + \sin\phi\frac{\partial^2}{\partial\phi\partial r}\right] + \frac{1}{r^2}\cos\phi^2\frac{\partial^2}{\partial\phi^2} \tag{3.22}$$

Adding them up:

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\phi^2}\right]. \tag{3.23}$$

So Laplace's equation in plane polar coordinates is

$$\frac{\partial^2 V}{\partial r^2} + \frac{1}{r}\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial\phi^2} = 0. \tag{3.24}$$

3.3 Jacobians

Recall equations 3.8. They can be written using determinants as:

$$\frac{\partial z}{\partial u} = \begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial z}{\partial y} & \frac{\partial v}{\partial y} \end{vmatrix} \Big/ \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} \end{vmatrix} \tag{3.25}$$

Similarly:

$$\frac{\partial z}{\partial v} = -\begin{vmatrix} \frac{\partial z}{\partial x} & \frac{\partial u}{\partial x} \\ \frac{\partial z}{\partial y} & \frac{\partial u}{\partial y} \end{vmatrix} \Big/ \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} \end{vmatrix} \tag{3.26}$$

The determinant in the denominator is called a *Jacobian* and has two special notations:

$$J\frac{(u,v)}{(x,y)} = \frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} \end{vmatrix} = \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial v}{\partial x}\frac{\partial u}{\partial y}\right) \tag{3.27}$$

This Jacobian is of special interest, because it contains all the information about the transformation between one set of coordinates (x,y) and another (u,v). As we shall see later, this same Jacobian is especially useful when changing variables from (x,y) to (u,v) in multiple integrals.

WARNING! You may see two definitions of the Jacobian. One is the one we have used (it concurs with Stephenson's book):

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial v}{\partial x} \\ \frac{\partial u}{\partial y} & \frac{\partial v}{\partial y} \end{vmatrix} \tag{3.28}$$

The other (eg Solkolnikoff & Redheffer, or Riley) is

$$\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \quad . \tag{3.29}$$

If you know about determinants, you will see that these are **IDENTICAL**. I have chosen the first definition because its layout appears the same as the matrix notation for transformations which we outline later. (You will learn about matrices & determinants later in the course, so remember to revisit this.)

3.4 Some standard transformations

There are certain transformations which occur very frequently, viz:

- Cartesian to plane polar coordinates.
- Cartesian to spherical polar coordinates.
- Cartesian to cylindrical polar coordinates.

3.4.1 Cartesian to plane polars

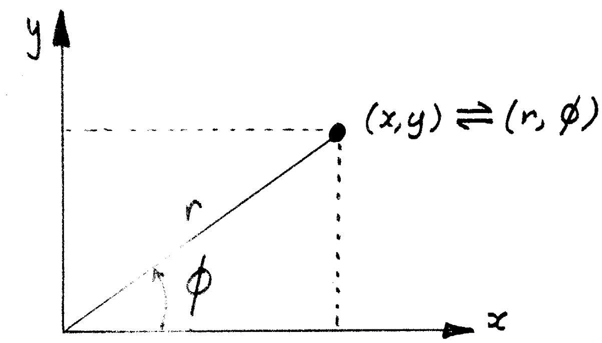

Figure 3.2: Cartesian to plane polars

$$x = r\cos\phi; \quad y = r\sin\phi \quad . \tag{3.30}$$

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y} \tag{3.31}$$

$$\frac{\partial}{\partial \phi} = \frac{\partial x}{\partial \phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \phi}\frac{\partial}{\partial y} \tag{3.32}$$

$\Rightarrow$

$$\frac{\partial}{\partial r} = \cos\phi\frac{\partial}{\partial x} + \sin\phi\frac{\partial}{\partial y} \tag{3.33}$$

$$\frac{\partial}{\partial \phi} = -r\sin\phi\frac{\partial}{\partial x} + r\cos\phi\frac{\partial}{\partial y} \tag{3.34}$$

Hence

$$\frac{\partial}{\partial x} = \cos\phi\frac{\partial}{\partial r} - \frac{\sin\phi}{r}\frac{\partial}{\partial \phi} \tag{3.35}$$

$$\frac{\partial}{\partial y} = \sin\phi\frac{\partial}{\partial r} + \frac{\cos\phi}{r}\frac{\partial}{\partial \phi} \tag{3.36}$$

$$\tag{3.37}$$

3.4.2 Cartesian to spherical polars

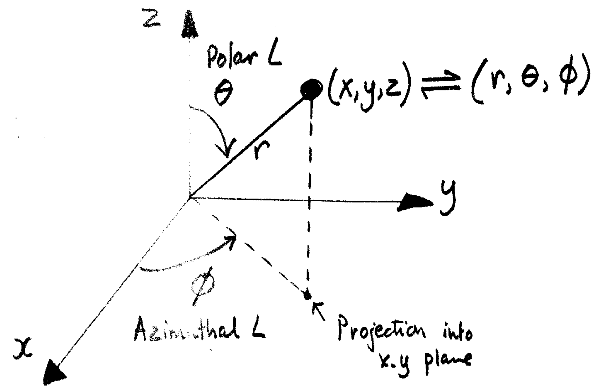

Figure 3.3: Cartesian to spherical polars

$$x = r\sin\theta\cos\phi; \quad y = r\sin\theta\sin\phi; \quad z = r\cos\theta. \tag{3.38}$$

$$\frac{\partial}{\partial r} = \sin\theta\cos\phi\frac{\partial}{\partial x} + \sin\theta\sin\phi\frac{\partial}{\partial y} + \cos\theta\frac{\partial}{\partial z} \tag{3.39}$$

$$\frac{\partial}{\partial \theta} = r\cos\theta\cos\phi\frac{\partial}{\partial x} + r\cos\theta\sin\phi\frac{\partial}{\partial y} - r\sin\theta\frac{\partial}{\partial z} \tag{3.40}$$

$$\frac{\partial}{\partial \phi} = -r\sin\theta\sin\phi\frac{\partial}{\partial x} + r\sin\theta\cos\phi\frac{\partial}{\partial y} \tag{3.41}$$

$$\tag{3.42}$$

Hence

$$\frac{\partial}{\partial x} = \sin\theta\cos\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\phi\frac{\partial}{\partial \theta} - \frac{\sin\phi}{r\sin\theta}\frac{\partial}{\partial \phi} \tag{3.43}$$

$$\frac{\partial}{\partial y} = \sin\theta\sin\phi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\phi\frac{\partial}{\partial \theta} + \frac{\cos\phi}{r\sin\theta}\frac{\partial}{\partial \phi} \tag{3.44}$$

$$\frac{\partial}{\partial z} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta} \tag{3.45}$$

$$\tag{3.46}$$

3.4.3 Cartesian to cylindrical polars

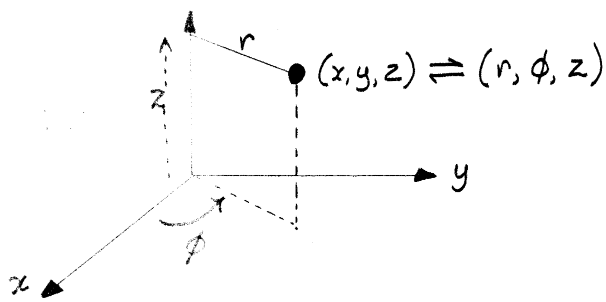

Figure 3.4: Cartesian to cylindrical polars

$$x = r\cos\phi; \quad y = r\sin\phi; \quad z = z. \tag{3.47}$$

So

$$\frac{\partial}{\partial r} = \cos\phi\frac{\partial}{\partial x} + \sin\phi\frac{\partial}{\partial y} \tag{3.48}$$

$$\frac{\partial}{\partial \phi} = -r\sin\phi\frac{\partial}{\partial x} + r\cos\phi\frac{\partial}{\partial y} \tag{3.49}$$

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial z} \tag{3.50}$$

These are very similar to plane polars.

3.5 Matrix notation for transformations

When you are *au fait* with transformations (and matrices!) you may care to think about the following as a convenient notation. There is nothing new about partial differentiation here.
In matrix notation, the transformation from Cartesian to plane polar coordinates, for example, can be written as:

$$\begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} = [\mathbf{J}] \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \phi} \end{pmatrix} \tag{3.51}$$

where the matrix is

$$[\mathbf{J}] = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial \phi}{\partial x} \\ \frac{\partial r}{\partial y} & \frac{\partial \phi}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\phi & -\frac{\sin\phi}{r} \\ \sin\phi & \frac{\cos\phi}{r} \end{pmatrix} \tag{3.52}$$

Notice anything? The layout of the $[\mathbf{J}]$ is exactly the same as the Jacobian $\partial(r,\phi)/\partial(x,y)$.
The inverse operation is then

$$\begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial \theta} \end{pmatrix} = [\mathbf{K}] \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial y}{\partial r} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} \end{pmatrix} \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{pmatrix} \tag{3.53}$$

But matrix theory indicates that

$$[\mathbf{K}] = [\mathbf{J^{-1}}]\ , \tag{3.54}$$

the *inverse* of $[\mathbf{J}]$. We can check this by working out $[\mathbf{K}][\mathbf{J}]$. It should give the identity matrix. For the particular example,

$$\begin{pmatrix} \cos\phi & \sin\phi \\ -r\sin\phi & r\cos\phi \end{pmatrix} \begin{pmatrix} \cos\phi & -\frac{\sin\phi}{r} \\ \sin\phi & \frac{\cos\phi}{r} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.55}$$

which is as expected.

3.6 Setting up a transformation

♣ A simple example.

Suppose that the turbine blades of a jet engine are attached to the main shaft by a piece of metal with the cross section in the xy plane shown in the figure. You have measured (see the exhibit by the front door on how it is done!!) the steady state temperature T at points throughout the section, but because of the positions of the points it proves much easier to compute partial derivatives of temperature along the directions of constant u and v where

$$u = x + y; \qquad v = x - y \ .$$

Rolls-Royce want to know the Whizz Function of the blades, which is given by $\partial^2 T/\partial x^2 - \partial^2 T/\partial y^2 + \partial T/\partial x + \partial T/\partial y$. How is this evaluated in terms of partials with respect to u and v?

You have NEW variables in terms of OLD, and do not have an explicit form of the function. Thus you must use the Chain rule for partials indirectly, setting up simultaneous equations in the operators. Ie:

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial}{\partial v} = \frac{\partial}{\partial u} + \frac{\partial}{\partial v} \tag{3.56}$$

$$\frac{\partial}{\partial y} = \frac{\partial u}{\partial y}\frac{\partial}{\partial u} + \frac{\partial v}{\partial y}\frac{\partial}{\partial v} = \frac{\partial}{\partial u} - \frac{\partial}{\partial v} \tag{3.57}$$

So that

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial u^2} + 2\frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v^2} \tag{3.58}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 f}{\partial u^2} + -2\frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v^2} \ . \tag{3.59}$$

Thus

$$\frac{\partial^2 f}{\partial x^2} - \frac{\partial^2 f}{\partial y^2} + \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = 4\frac{\partial^2 f}{\partial u \partial v} + 2\frac{\partial f}{\partial u} \ . \tag{3.60}$$

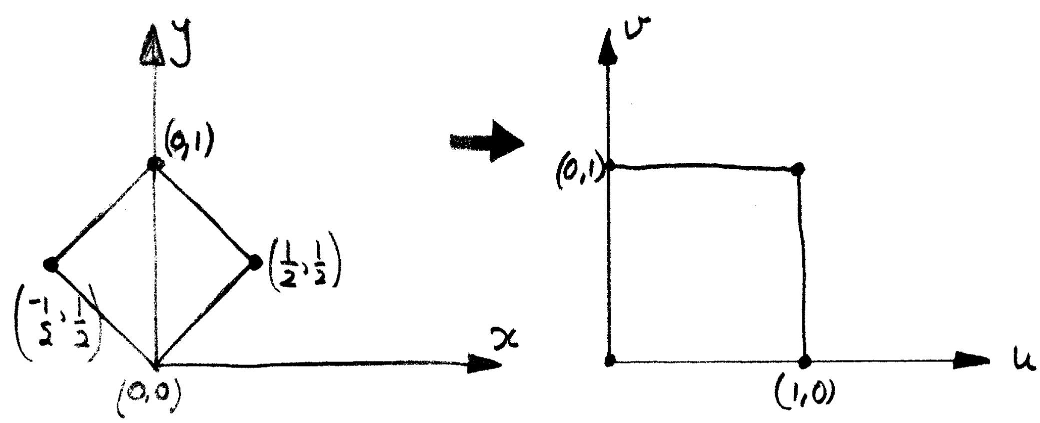

Figure 3.5: The blade's basal cross section in the xy plane

3.7 A "good" transformation

You may still be equivocal about what makes a good transformation. Here is one hint and one hard fact.

3.7.1 A hint on shape and transformation

If we asked what is the nicest shape to fit into Cartesian coordinates xy you would answer "a square". Similarly if you drew your uv axes at right angles you would want a shape in that to appear as a square. So given some arbitrary shape in the xy plane, the best sort of transformation is one which maps it onto a square in the uv plane. That was true in the example above. Suppose we were given the shape in the figure — a transformation to plane polars would turn it into a square in $r\phi$ coordinates.

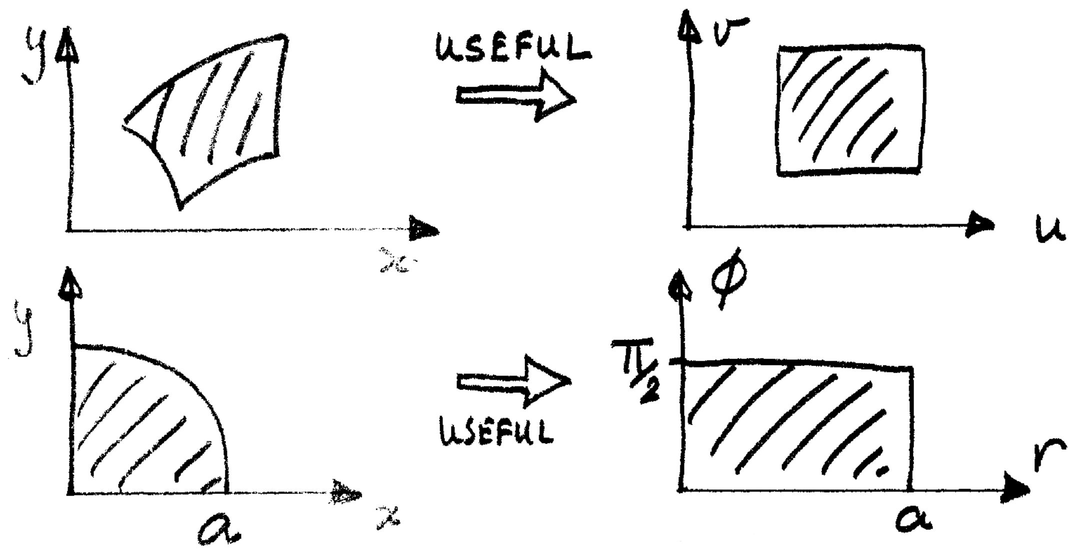

Figure 3.6:

3.7.2 Hard fact: Jacobians and Functional Dependence

We have just been discussing transformations from one set of coordinates (x, y) to another (u, v). Will any transformation work?

For example, suppose we had some function $f(x, y)$ and were considering a transformation to a new set of coordinates, (u, v), given by

$$u = x^2 + y + 1 \qquad v = x^4 + 2x^2y + y^2 + x^2 - y. \tag{3.61}$$

Could we be sure that this would make a good transformation?
In fact, this would not make a good transformation because there is a relationship between u and v

$$v = (u-1)^2 - (u-1) = u^2 - 3u + 2, \tag{3.62}$$

and, far from being independent, u and v are *functionally dependent.* How could we test for this?

You will remember that the transformation involves a Jacobian as denominator. For a transform from $f(x, y)$ to $F(u, v)$ the Jacobian involved in the denominator is

$$\frac{\partial(u, v)}{\partial(x, y)}.$$

and you might guess that chaos would ensue if the Jacobian were zero. In fact, this is just the test for functional dependence, as we prove now.

Theorem If $u(x, y)$ and $v(x, y)$ are functionally dependent, then

$$\frac{\partial(u, v)}{\partial(x, y)} = 0.$$

Proof. If $u(x,y)$ and $v(x,y)$ are functionally dependendent, then there is some $z = F(u,v) = 0$. This implicitly defines $u = u(v)$. Differentiating z with respect to x and y gives:

$$F_u u_x + F_v v_x = 0 \tag{3.63}$$
$$F_u u_y + F_v v_y = 0 \tag{3.64}$$

For consistency between these two equations we must have $u_x = \beta u_y$ and $v_x = \beta v_y$, where β is some constant. Thus

$$u_x v_y - v_x u_y = \begin{vmatrix} u_x & v_x \\ u_y & v_y \end{vmatrix} = \frac{\partial(u,v)}{\partial(x,y)} = 0. \tag{3.65}$$

3.8 Transformations using implicit functions

This section will test your understanding.

We have seen that to define an implicit function in 2 variables we need 1 function of 3 variables. To define two functions in two variables we require 2 functions of 4 variables. In other words

$$f(x,y,u,v) = 0 \qquad g(x,y,u,v) = 0 \tag{3.66}$$

are sufficient to define $u = u(x,y)$ and $v = v(x,y)$ implicitly.

Now differentiate each function with respect to x and y:

$$f_x + f_u u_x + f_v v_x = 0 \tag{3.67}$$
$$f_y + f_u u_y + f_v v_y = 0 \tag{3.68}$$
$$g_x + g_u u_x + g_v v_x = 0 \tag{3.69}$$
$$g_y + g_u u_y + g_v v_y = 0. \tag{3.70}$$

These are analogous to equations 2.53 and 2.54, but involve partial derivatives because the implicit functions are of more than one variable.
Solving, we find

$$u_x = -\frac{\partial(f,g)}{\partial(x,v)} \Big/ \frac{\partial(f,g)}{\partial(u,v)}. \tag{3.71}$$

Similarly:

$$u_y = -\frac{\partial(f,g)}{\partial(y,v)} \Big/ \frac{\partial(f,g)}{\partial(u,v)}. \tag{3.72}$$

Check these for yourselves, and try to remember the *principals* on which the results were obtained, not the results themselves.

This page is blank intentionally

Chapter 4

Applications of Differentiation

4.1 Taylor's Theorem for two or more variables

First, let us recall that Taylor's theorem in one variable is: If $f(x)$ is a continuous single-valued function of x with continuous derivatives $f'(x), f''(x), \ldots,$ upto $f^{(n)}(x)$, in an interval $a \leq x \leq a + \alpha$, and if $f^{(n+1)}(x)$ exists in $a < x < \alpha$, then:

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \ldots + \frac{(x-a)^n}{n!}f^{(n)}(a) + R_n(x) \tag{4.1}$$

where the remainder is

$$R_n(x) = \frac{(x-a)^{n+1}}{(n+1)!}f^{n+1}(\epsilon), \quad a < \epsilon < x \tag{4.2}$$

Before we get overwhelmed, what are we trying to do when we extend Taylor's theorem to two variables? We have some function $f(x, y)$, and we know the value of f and its derivatives at (a, b). How can we use these to estimate the function at a point (x, y) displaced slightly from (a, b)?

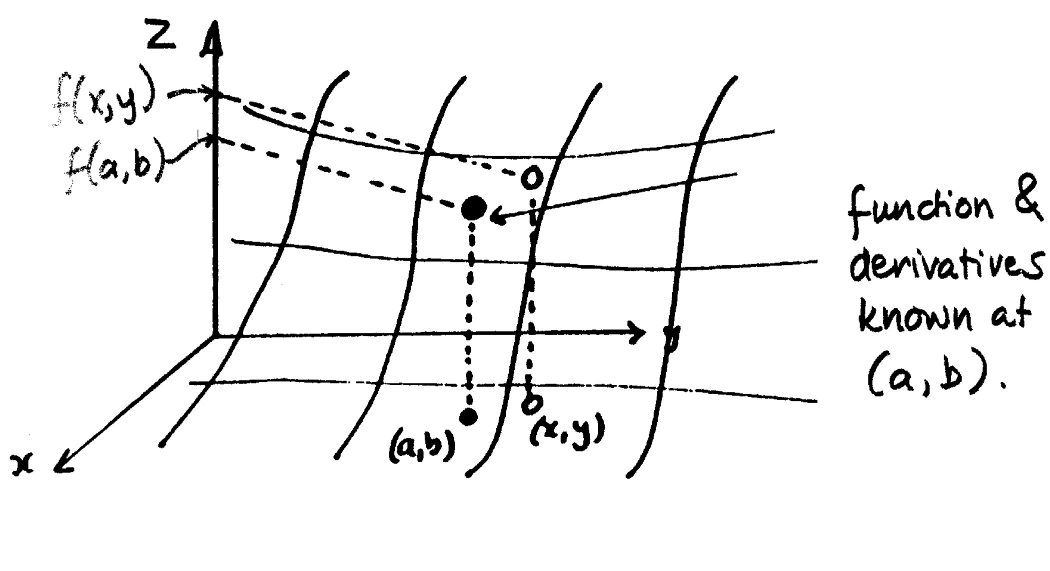

Figure 4.1: Geometry for Taylor's theorem in two variables

Let us first assume that the relevant derivatives exist in the region $(a, b) \to (a + \alpha, b + \beta)$. Let the displacement be expressed as $(\alpha t, \beta t)$ so that $x = a + \alpha t$ and $y = b + \beta t$, with $0 \leq t \leq 1$. Then

$$f(x, y) = f(a + \alpha t, b + \beta t) = w(t). \tag{4.3}$$

Differentiating w we have

$$\begin{aligned} \frac{dw}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} && (4.4) \\ &= f_x \alpha + f_y \beta && (4.5) \end{aligned}$$

We can carry on to find d^2w/dt^2 and so on. The expression in operator form for the nth derivative is

$$w^{(n)} = \frac{d^n w}{dt^n} = \left(\alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y}\right)^n f \quad . \tag{4.6}$$

But MacLaurin's series (in one variable!) tells us that:

$$w(t) = w(0) + tw'(0) + \frac{t^2}{2!}w''(0) + \ldots + \frac{t^n}{n!}w^{(n)}(0) + R_n(t) \tag{4.7}$$

where the remainder is

$$R_n(t) = \frac{t^{n+1}}{(n+1)!}w^{n+1}(\epsilon), \quad 0 < \epsilon < t \tag{4.8}$$

Thus inserting the expressions for $w^{(n)}$ we find

$$\begin{aligned} f(x, y) = f(a, b) &+ [(x - a)f_x(a, b) + (y - b)f_y(a, b)] && (4.9) \\ &+ \frac{1}{2!}[(x - a)^2 f_{xx} + 2(x - a)(y - b)f_{xy} + (y - b)^2 f_{yy}] \\ &+ \ldots + R_n \end{aligned}$$

♣ **Example.** At some instant, the temperature distribution in a plate lying in the xy-plane is given by $T(x, y) = C + e^{-(x^2+y^2)}$. Assuming that measurements of T and its partial derivatives up to the second order are available at the point $(x, y) = (2, 1)$, estimate the temperature at $(x, y) = (2.1, 1.1)$. Compare this with the actual value from the function. How would you estimate the temperature gradient in the $x-$direction at this point?

First construct values for the "measurements" at $(2, 1)$.

Quantity				At $x = 2, y = 1$
$T(x, y)$	$=$	$C + \exp[-(x^2 + y^2)]$	$=$	$C + \exp[-5]$
T_x	$=$	$-2x \exp[-(x^2 + y^2)]$	$=$	$-4 \exp[-5]$
T_y	$=$	$-2y \exp[-(x^2 + y^2)]$	$=$	$-2 \exp[-5]$
T_{xx}	$=$	$(4x^2 - 2) \exp[-(x^2 + y^2)]$	$=$	$14 \exp[-5]$
T_{yy}	$=$	$(4y^2 - 2) \exp[-(x^2 + y^2)]$	$=$	$2 \exp[-5]$
T_{xy}	$=$	$4xy \exp[-(x^2 + y^2)]$	$=$	$8 \exp[-5]$

Using Taylor's thm to 2nd order:

$$\begin{aligned} T(2.1,1.1) &\approx T(2,1)+[-(0.1)4\exp[-5]+-(0.1)2\exp[-5]] \\ &+ \frac{1}{2}[(0.01)14\exp[-5]+2(0.01)8\exp[-5]+(0.01)2\exp[-5]] \\ &= C+0.56\exp[-5]. \end{aligned} \tag{4.10}$$

Using the exact function:

$$\begin{aligned} T(2.1,1.1) &= C+\exp[-(2.1^2+1.1^2)] \\ &= C+\exp[-5.62] \\ &= C+\exp[-0.62]\exp[-5] \\ &= C+0.538\exp[-5] \end{aligned} \tag{4.11}$$

Note. You will have been struck by the similarity between Taylor's Theorem to estimate the function at some displacement and using the Total Differential. If you look at the Taylor expansion, you will see that taking the total differential is like expanding to first order. The total differential is appropriate in the limit as $\Delta x = x - a$ and $\Delta y = y - a$ both tend to zero.

4.2 Finding maxima, minima and saddle points

Back in Lecture 1 we saw some surface plots of functions of two variables. Armed with Taylor's Theorem, we can start to explore some special points of those functions, viz.

- maxima,
- minima, and
- saddle points.

4.2.1 Maxima and Minima

A function has a maximum at (a,b) if

$$f(a+\delta x, b+\delta y) - f(a,b) < 0 \tag{4.12}$$

for arbitrarily small δx and δy. Geometrically we see that slices through the function in both $x-$ and $y-$directions exhibit maxima.
Similarly a function has a minimum at (a,b) if

$$f(a+\delta x, b+\delta y) - f(a,b) > 0 \tag{4.13}$$

for arbitrarily small δx and δy. Geometrically we see that slices through the function in both $x-$ and $y-$directions exhibit minima.
How do we compute the stationary values. One might guess by looking for points where $\partial f/\partial x$ and $\partial f/\partial y$ are both zero.
Indeed this is the case. For $f(x,y)$ to be stationary we require the total differential to be zero, ie:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = 0. \tag{4.14}$$

Figure 4.2:

Because dx and dy are independent, this gives rise to the condition that $\partial f/\partial x = 0$ and $\partial f/\partial y = 0$.

But how to decide whether the point is a maximum or minimum? Again one might guess that this will have something to do with second derivatives — but what exactly?
Using Taylor's Theorem we have

$$\begin{aligned} f(a+\delta x, b+\delta y) = f(a,b) &+ [(\delta x)f_x(a,b) + (\delta y)f_y(a,b)] \\ &+ \frac{1}{2!}[(\delta x)^2 f_{xx} + 2(\delta x)(\delta y)f_{xy} + (\delta y)^2 f_{yy}] \\ &+ \ldots \end{aligned}$$

But the first order derivativens $f_x,\ f_y$ are zero, so that, to the second order in small quantities:

$$f(a+\delta x, b+\delta y) - f(a,b) = \frac{1}{2}[(\delta x)^2 f_{xx} + 2(\delta x)(\delta y)f_{xy} + (\delta y)^2 f_{yy}]. \tag{4.15}$$

Now the conditions for maxima and minima above (eqs 4.12 and 4.13) tell us that we should be interested in the **sign** of the rhs of equation 4.15 or, equivalently, the sign of:

$$[(\delta x)^2 f_{xx} + 2(\delta x)(\delta y)f_{xy} + (\delta y)^2 f_{yy}]$$

This expression can be rewritten in two ways:

1. $\frac{1}{f_{xx}}\left\{[(\delta x)f_{xx} + (\delta y)f_{xy}]^2 - (\delta y)^2[{f_{xy}}^2 - f_{xx}f_{yy}]\right\}$
2. $\frac{1}{f_{yy}}\left\{[(\delta x)f_{xy} + (\delta y)f_{yy}]^2 - (\delta x)^2[{f_{xy}}^2 - f_{xx}f_{yy}]\right\}$

It is clear that in general the sign depends on the actual δx and δy under consideration, that is, on how you move off from the point where $f_x = f_y = 0$.

But what we can say unequivocally is that if $Q = [{f_{xy}}^2 - f_{xx}f_{yy}] < 0$ then the term in $\{\ldots\}$ is positive. So then the sign depends on f_{xx}, or equivalently f_{yy}. (In fact, we have just shown that, when $Q < 0$, $SIGN(f_{xx}) = SIGN(f_{yy})$.)
So,

- when $Q < 0$ and $f_{xx} < 0$ (or equivalently $f_{yy} < 0$) the point is a maximum
- when $Q < 0$ and $f_{xx} > 0$ (or equivalently $f_{yy} > 0$) the point is a minimum.

What about when $Q > 0$? Then the sign does depend on how you move off from the point where $f_x = f_y = 0$. This is a **saddle point**. The function appears to have a maximum if you move in one direction and a minimum if you move in another.

The tests are summed up in the following table.

$f_x = f_y$	$Q = [{f_{xy}}^2 - f_{xx}f_{yy}]$	f_{xx} (or f_{yy})	Type
$= 0$	< 0	> 0	Minimum
$= 0$	< 0	< 0	Maximum
$= 0$	> 0	Irrelevant	Saddle

♣ **Example.**

1. Find the stationary points of $f(x, y) = x^4 + 4x^2y^2 - 2x^2 + 2y^2 - 1$ and indicate their character.

$$f_x = 4x^3 + 8xy^2 - 4x = 0 \quad \text{or} \quad 4x(x^2 + 2y^2 - 1) = 0 \tag{4.16}$$
$$f_y = 8x^2y + 4y = 0 \quad \text{or} \quad 4y(2x^2 + 1) = 0 \tag{4.17}$$

Solving for the real roots:

- From (4.16): $x = 0$ and then from (4.17) $y = 0$.
- From (4.17): $y = 0$ and then from (4.16) $x = \pm 1$.

Hence the roots are $(0, 0)$, $(1, 0)$ and $(-1, 0)$.

Now

$$f_{xx} = 12x^2 + 8y^2 - 4 \tag{4.18}$$
$$f_{yy} = 8x^2 + 4 \tag{4.19}$$
$$f_{xy} = 16xy \tag{4.20}$$
$$\tag{4.21}$$

Point	$Q = [{f_{xy}}^2 - f_{xx}f_{yy}]$	f_{xx} (or f_{yy})	Type
$(0, 0)$	$0 - (-4)(4) > 0$	Irrelevant	Saddle
$(1, 0)$	$0 - (8)(12) < 0$	$8 > 0$	Minimum
$(-1, 0)$	$0 - (8)(12) < 0$	$8 > 0$	Minimum

4.3 Plotting Functions of several variables

In lecture 1 we noted that functions of 2 variables were relatively easy to visualize, whereas those in more variables wer considerably more difficult, simply because we live in a 3D world. (Time is an exception — 3 variable where one is time is visualized as a moving surface.)

4.3.1 3D Surface plots

Scattered in the notes are various examples of three dimensional surface plots. These are created by using the $z = f(x, y)$ surface like a landscape which can be viewed from various positions. The algorithm used works out which surfaces are visible and which are hidden. An extra help in visualizing the surface is to shade it as if a directional light source were placed above the surface.

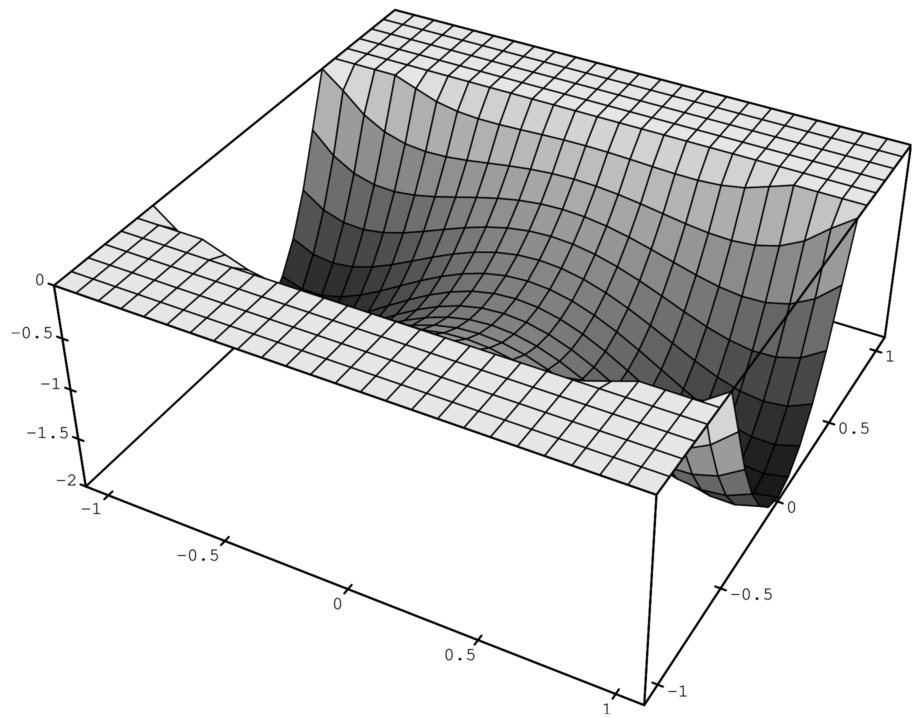

Figure 4.3: The function $f(x,y) = x^4 + 4x^2y^2 - 2x^2 + 2y^2 - 1$.

4.3.2 Contour plots

To make 3D surface plots you really do need a computer. But contour plots are rather easier to sketch. There are two slightly different types.
In the first, one looks directly down the z-axis onto the function surface, and you plot *level contours*, or contours of constant z. So these are the things cartographers use in a standard topographical map. The contours are spaced evenly in z.

Let us look the level contour plot of the function $f(x,y) = x^4 + 4x^2y^2 - 2x^2 + 2y^2 - 1$.
The other sort of contour plot is (for want of a better name) a slice plot. Here one traces out the function $f(x, y_o)$ for several fixed values of y (Eg, $y_o = 1, 2, ...$), and/or the function $f(x_o, y)$ for several fixed values of x_o.
This sort of plot is commonly used to display transistor characteristics.

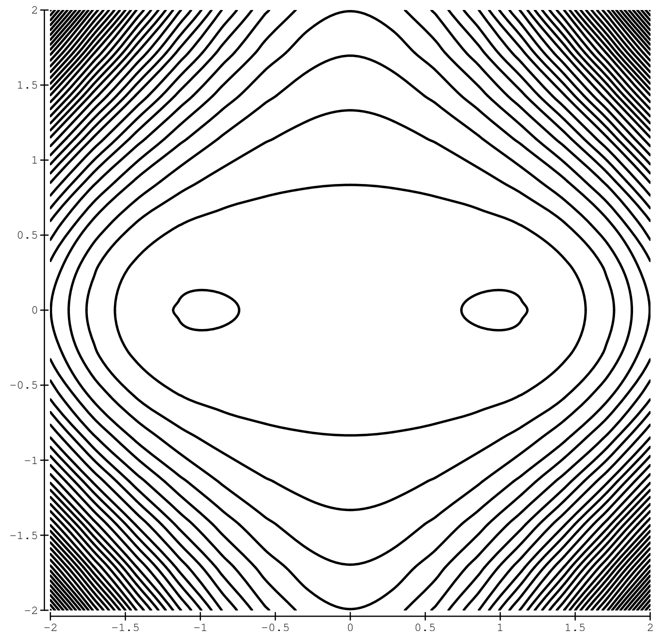

Figure 4.4: Level contours in $f(x, y) = x^4 + 4x^2y^2 - 2x^2 + 2y^2 - 1$.

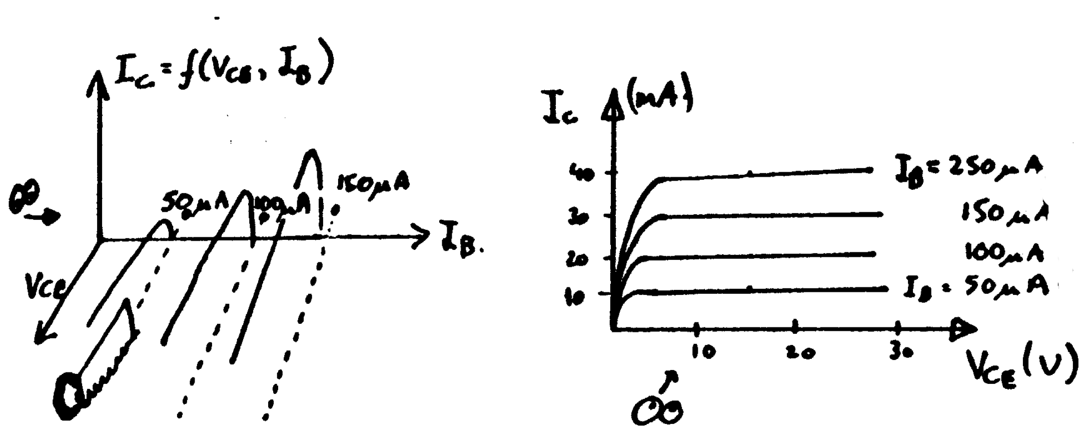

Figure 4.5: A slice plot through transistor characteristics

4.4 Constrained Maxima and Minima: Lagrange Multipliers

We are all familiar with the notion of maximizing and minimizing a function. But sometimes it is necesary to maximize or minimize a function subject to some constraint.

Suppose that some $F(x, y)$ is to be examined for stationary points, subject to the constraint that $G(x, y) = 0$. For $F(x, y)$ to be stationary we require:

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0 \tag{4.22}$$

which earlier gave rise to the condition that $\frac{\partial F}{\partial x} = 0$ and $\frac{\partial F}{\partial y} = 0$ because dy and dx were independent. But now dx and dy are NOT independent, because they are implicitly related via $G(x, y) = 0$. (In other words $G(x, y) = 0$ defines $y = y(x)$ implicitly and therefore dy/dx exists.)
Remember that we have shown in this case that

$$dG = \frac{\partial G}{\partial x}dx + \frac{\partial G}{\partial y}dy = 0 \tag{4.23}$$

(from where we get $dy/dx = -\frac{\partial G}{\partial x}/\frac{\partial G}{\partial y}$).
So we can add in any amount of dG to dF and still get zero. Thus:

$$d(F + \lambda G) = \left(\frac{\partial F}{\partial x} + \lambda\frac{\partial G}{\partial x}\right)dx + \left(\frac{\partial F}{\partial y} + \lambda\frac{\partial G}{\partial y}\right)dy = 0 \tag{4.24}$$

For the constrained maxima then we require λ such that

$$\left(\frac{\partial F}{\partial x} + \lambda\frac{\partial G}{\partial x}\right) = 0 \tag{4.25}$$

$$\left(\frac{\partial F}{\partial y} + \lambda\frac{\partial G}{\partial y}\right) = 0 \tag{4.26}$$

These two equations, along with $G(x, y) = 0$ are sufficient to determine the stationary points of interest (and λ).
♣ **Example.** Find the maximum distance from the origin to the curve $3x^2 + 3y^2 + 4xy - 2 = 0$. The distance from the origin to any point (x, y) is $l = \sqrt{(x^2 + y^2)}$. We want to minimize this, but subject to (x, y) lying on the curve – ie, subject to the constraint $3x^2 + 3y^2 + 4xy - 2 = 0$. In fact, things are made a bit easier if we minimize not l but l^2: it obviously amounts to the same thing.
The Lagrange equations are:

$$2x + \lambda(6x + 4y) = 0 \tag{4.27}$$

$$2y + \lambda(6y + 4x) = 0 \tag{4.28}$$

and these must be solved with

$$3x^2 + 3y^2 + 4xy - 2 = 0 \ . \tag{4.29}$$

From $y[4.27] - x[4.28]$ we have $4\lambda(y^2 - x^2) = 0$ and so $y = \pm x$.
If $y = +x$ then eq 4.29 gives $10x^2 - 2 = 0$ or $x = \pm\frac{1}{\sqrt{5}}$.
If $y = -x$ then eq 4.29 gives $2x^2 - 2 = 0$ or $x = \pm 1$.
Thus the stationary points are

$$\left(\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right) \quad \left(-\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right) \quad (1, -1) \quad (-1, 1) \quad .$$

For the first pair, we have $l^2 = 2/5$ and for the second pair, $l^2 = 2$. Thus the maximum distance is $l = \sqrt{2}$.

Chapter 5

Double and Triple Integrals: Repeated Integration

5.1 The integral for a single variable

First, we review the definition of the integral in one variable. Consider a function $f(x)$ which is defined in some bounded linear region R of x. Let R be divided up into n subregions, where δx_i denotes the length of the ith subregion. Let f_i be the associated function value. If the sum

$$\sum_{i=1}^{n} f_i \delta x_i \tag{5.1}$$

exists and is finite as $n \to \infty$ and $\delta x_i \to 0$, then that limit is the integral: that is,

$$\int f(x)dx = \lim_{\substack{n\to\infty \\ \delta x_i \to 0}} \sum_{i=1}^{n} f_i \delta x_i . \tag{5.2}$$

We assume here that it is irrelevant how the region is subdivided and the the x_i chosen. This is always true for a continuous function f.
The integral is often thought of as the "area" under the curve. Remember that when the curve is below the x-axis the contribution to the area is negative. (In what circumstances would this be a genuine area? Make sure you understand the difference between this and the area integral using a double integral.)

5.1.1 Double Integral

The extension to two variables is direct.
Consider a function $f(x, y)$ which is defined in some bounded region R of the (x, y) plane. Let R be divided up into n subregions, where δA_i denotes the area of the ith subregion. Let f_i be the function value associated with the ith subregion. If the sum

$$\sum_{i=1}^{n} f_i \delta A_i \tag{5.3}$$

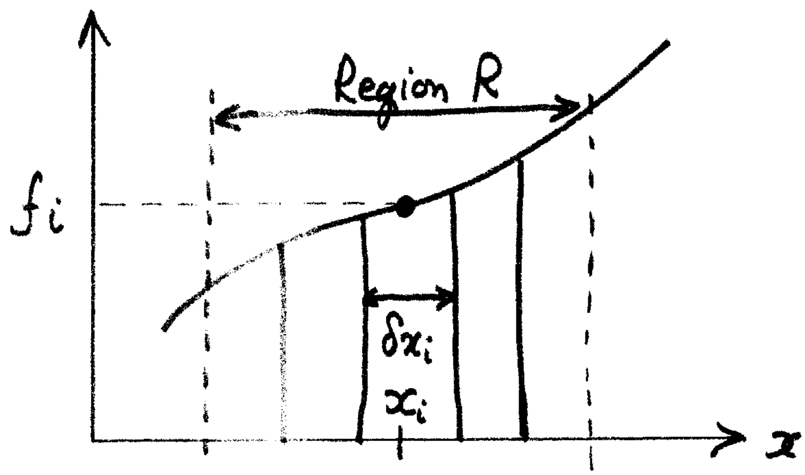

Figure 5.1: The integral for one variable

exists and is finite as $n\to\infty$ and $\delta A_i\to 0$, then that limit is the *double integral*:

$$\iint f(x,y)dA = \lim_{\substack{n\to\infty \\ \delta A_i\to 0}} \sum_{i=1}^{n} f_i \delta A_i. \tag{5.4}$$

Again we assume here that it is irrelevant how the region is subdivided and the the x_i, y_i chosen. This is always true for a continuous function $f(x,y)$. Later we indicate how to extend integration to discontinuous functions.
Rather that the "area" under a plane curve, the double integral could be thought of as the "volume" under the function surface. (Again, under what circumstances would this be a genuine volume? Distingush this from the volume integral obtained using triple integration.)

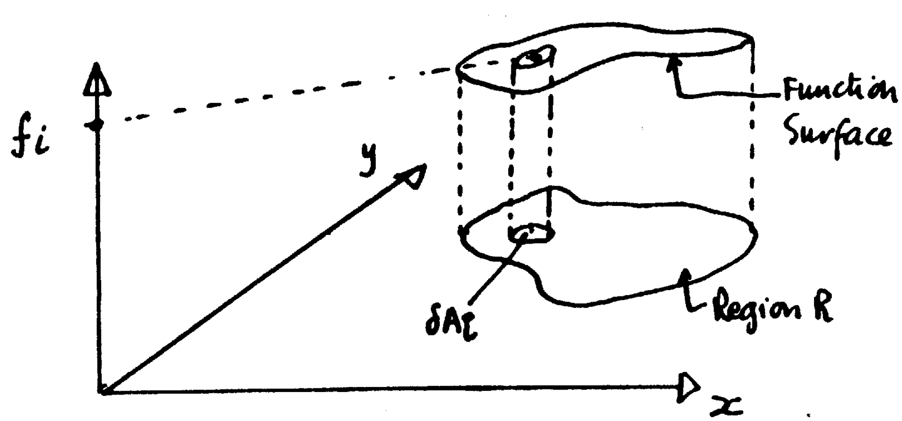

Figure 5.2: The double integral

Later we shall consider a variety of physical properties computable using the double integral. For now we consider just two. Suppose a plate or lamina of uniform thickness and surface density $\sigma(x,y)$ (ie mass per unit area) occupies the region R in the $xy-$plane. Then

1. The plate's Area is $\iint_R dA$.

2. The plate's Mass is $\iint_R \sigma(x,y)dA$.

Item (2) gives us an example of the general form of the double integral.

5.2 Performing the integration: repeated integration

Earlier we noted that it should not matter how one subdivides the region R for a continuous function. However, to be able to produce a symbolic result for the integral, it is important to produce an appropriate subdivision of the region. As noted in Lecture 3, it is most useful if the subdivision or mesh or tiling is in sympathy with the shape of the region of integration.

To produce a symbolic result for the integral, where the integrand is a function $f(x,y)$, the xy plane must be divided by cake rack of x-constant, y-constant lines. The tiles are rectangles of sides δx and δy. Then, obviously,

$$\delta A = \delta x \delta y \quad \text{and, in the limit} \quad dA = dxdy\ . \tag{5.5}$$

The integral for the mass ((2) above) becomes

$$\text{Mass} = \iint_R \sigma(x,y)dxdy. \tag{5.6}$$

5.2.1 Carrying on

Suppose we first summed the little $dxdy$ tiles into a strip parallel to the $y-$axis. For each strip this involves holding x constant and summing up the length along y from one boundary intersection to the other. Note that we are assuming that the $x = const.$ line intersects the boundary just twice. More on this later.

The mass of the thin strip is

$$\delta M = \left[\int_{y=y_1(x)}^{y=y_2(x)} \sigma(x,y)dy\right]\delta x \tag{5.7}$$

Then we want to sum these up from the smallest to the largest values of x, so that the total integral mass is

$$\text{Mass of plate} = \int_{x=x_1}^{x_2}\left[\int_{y=y_1(x)}^{y=y_2(x)} \sigma(x,y)dy\right]dx \tag{5.8}$$

How do we evaluate the integral? In two stages, as a *repeated integral.*

- First, consider the integral $\left[\int_{y=y_1(x)}^{y=y_2(x)} \sigma(x,y)dy\right]$. Because x is held constant, simply treat it no differently from any other constant. (Recall that we already used this method in Lecture 1 when integrating total or perfect differentials.) We will then be left with some new function of (x,y) which is evaluated between the limits $y = y_1(x)$ and $y = y_2(x)$. This yields some function of x alone, $g(x)$ say.

- We then integrate the function $g(x)$ over x and evaluate it between the limits x_1 and x_2.

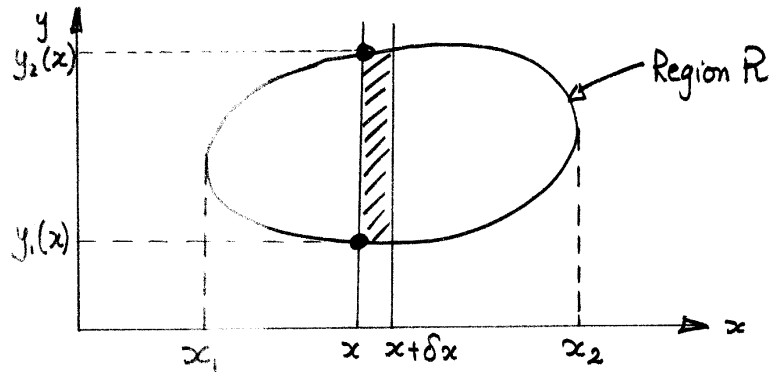

Figure 5.3: Summing tiles into a strip parallel to the y-axis.

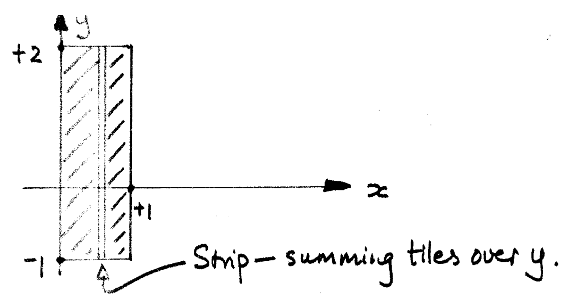

Figure 5.4: Always sketch the region of integration

Thus the double integral is broken down into two single integrals.

♣ **Example.** Find the mass of a lamina whose surface density varies as $\sigma = 1 + x^2y^2(x^2 - y^3)$ and which occupies the region $0 \le x \le 1$, $-1 \le y \le 2$.

The integral is

$$\int_{x=0}^{1} \left(\int_{y=-1}^{2} (1 + x^2y^2(x^2 - y^3))dy \right) dx. \tag{5.9}$$

In practice, there is a step which you should carry out before the others, and that is

- **sketch the region of integration**! See Figure 5.4
- Considering x as a constant in the $(\ldots)$ then:

$$(\ldots) = \left[y + \frac{1}{3}x^4y^3 - \frac{1}{6}x^2y^6 \right|_{y=-1}^{y=2} \tag{5.10}$$

$$= 3 + 3x^4 - \frac{21}{2}x^2 \ . \tag{5.11}$$

- So we have our $g(x)$. The final stage is:

$$\text{Mass} = \int_{x=0}^{1} (3 + 3x^4 - \frac{21}{2}x^2)dx = 3 + \frac{3}{5} - \frac{21}{6} = \frac{1}{10} \ . \tag{5.12}$$

5.2.2 The order of integration

Because x and y are independent variables, one could just as well exchange the order of integration, deriving first the area of a strip parallel to the $x-$axis, then summing these along y, as shown in the figure.

♣ **Example.** Do the previous example, with the order of integration reversed.
The integral becomes

$$\text{Mass} = \int_{y=-1}^{2}\left(\int_{x=0}^{1}(1+x^2y^2(x^2-y^3))dx\right)dy \tag{5.13}$$

$$= \int_{y=-1}^{2}\left[x+\frac{1}{5}x^5y^2-\frac{1}{3}x^3y^5\right]_{x=0}^{1}dy \tag{5.14}$$

$$= \int_{y=-1}^{2} 1+\frac{y^2}{5}-\frac{y^5}{3}dy \tag{5.15}$$

$$= \left[y+\frac{y^3}{15}-\frac{y^6}{18}\right]_{-1}^{2} \tag{5.16}$$

$$= 3+\frac{9}{15}-\frac{64}{18}+\frac{1}{18}=\frac{1}{10} \quad . \tag{5.17}$$

Note, rather like higher partial differentiation, that although the order of integration is unimportant absolutely, it may affect the ease with which you can perform the integrals, as we will see below.

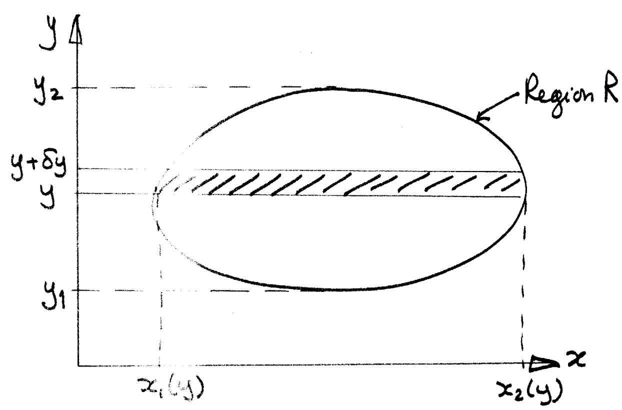

Figure 5.5: Reversing the order of integration. Summing tiles over x, then summing strips over y.

5.2.3 Non-constant limits

In the examples we just looked at, the limits of integration were constants, but we must consider cases where the limits are functions of the other variable(s).

♣ **Examples**

1. Derive the area of the hatched region of the $xy-$plane by integration.

   The answer is obviously $\frac{1}{2}$, but using repeated integration we have:

   Order A: y then x.

$$\begin{aligned} A &= \int_{x=0}^{1}\int_{y=x}^{1} dydx && (5.18)\\ &= \int_{x=0}^{1} y\Big|_{y=x}^{1}\, dx && (5.19)\\ &= \int_{x=0}^{1}(1-x)dx && (5.20)\\ &= 1-\frac{x^2}{2}\Big|_{x=0}^{1} && (5.21)\\ &= 1/2. && (5.22)\end{aligned}$$

   Order B: x then y. Note that we can change the order of integration, but that we have to think a little about the limits of integration. Reversed the integral is:

$$A = \int_{y=0}^{1}\int_{x=0}^{x=y} dxdy \qquad (5.23)$$

   which you should easily find to be $\frac{1}{2}$.

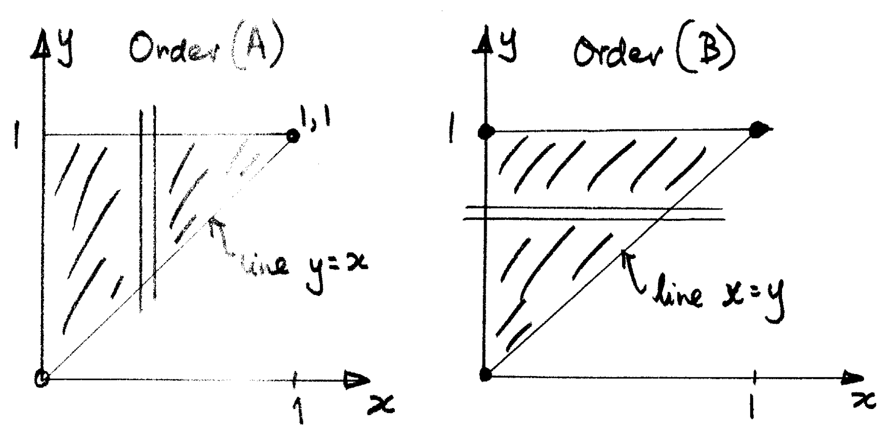

Figure 5.6: Figure for Example 1

2. Find the mass of a thin plate whose surface density varies as $\sigma(x, y) = x^2y$ and which occupies the hatched part of the positive quadrant shown, using both orders for repeated integration.

Order A: y first then x The integral we require is

$$\begin{aligned}
M &= \int_{x=0}^{a/2}\int_{y=0}^{(a^2-x^2)^{1/2}} \sigma(x,y)dydx && (5.24)\\
&= \int_{x=0}^{a/2}\left[\frac{1}{2}x^2y^2\right]_0^{(a^2-x^2)^{1/2}} dx && (5.25)\\
&= \int_{x=0}^{a/2}\frac{x^2}{2}[a^2-x^2]dx && (5.26)\\
&= \left[\frac{x^3a^2}{6}-\frac{x^5}{10}\right]_0^{a/2} && (5.27)\\
&= \frac{17a^5}{960}. && (5.28)
\end{aligned}$$

Order B: x first then y Now with the order of integration reversed, we hit a problem. The rightmost boundary is not defined by one equation, but but two: $x=\sqrt{a^2-y^2}$ and $x=a/2$ (see above figure). We have to take the integral in two bites therefore.

$$\begin{aligned}
M &= \int_{\sqrt{3}a/2}^{a}\int_0^{\sqrt{a^2-y^2}} x^2ydxdy + \int_0^{\sqrt{3}a/2}\int_0^{a/2} x^2ydxdy && (5.29)\\
&= \int_{\sqrt{3}a/2}^{a}\frac{y(a^2-y^2)^{3/2}}{3}dy + \int_0^{\sqrt{3}a/2}\frac{ya^3}{24}dy && (5.30)\\
&= \left[-\frac{1}{15}(a^2-y^2)^{5/2}\right|_{\sqrt{3}a/2}^{a} + \left[\frac{a^3}{24}\frac{y^2}{2}\right|_0^{\sqrt{3}a/2} && (5.31)\\
&= \frac{1}{15}\frac{a^5}{32}+\frac{a^5}{64} = \frac{17a^5}{960} \quad . && (5.32)
\end{aligned}$$

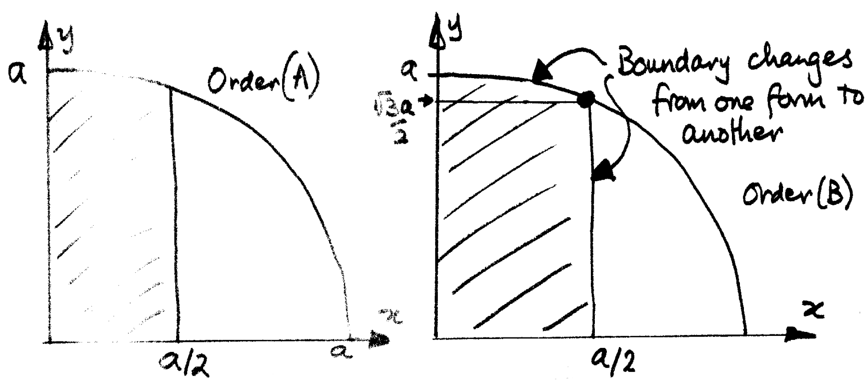

Figure 5.7: Figure for example 2

5.2.4 Complicated regions

There are several types of complicated regions.

A. Early on we said that for a continuous function it does not matter how you subdivide the region. However suppose we wish to find the mass, say, of a plate covering a region R comprised of two sub-regions R_1 having a density function $\sigma_1(x, y)$ and R_2 having a density function $\sigma_2(x, y)$. Now it does matter how you subdivide R. No little element δA should straddle the boundary of R_1 and R_2. The obvious solution is to perform two separate double integrations and add up the results:

$$\iint_R \sigma dA = \iint_{R_1} \sigma_1 dA + \iint_{R_2} \sigma_1 dA \quad .$$

Note that the functionsσ_1 and σ_2 can be discontinuous at the boundary.

B. Any closed smooth contour has two points where the tangent is parallel to the x axis and another two parallel to the y axes which coincide with the extremal values of x and y. Problems are caused when there are more parallel tangent points which do not coincide with the extremal has tangents which are parallel to either the x or y axes, but which do not coincide with the extremal values. In this case, there will some lines of constant x and/or constant y which cut the boundary more than twice (but note that the number of crossings is always even). But integrals only have two limits! The solution is to subdivide the region R into subregions not troubled by this shape problem, integrate over the subregions and then sum the results.

C. Another problem is the one we met in the previous example. The boundary is not defined by one function. Again we must subdivide into regions which do have a single description.

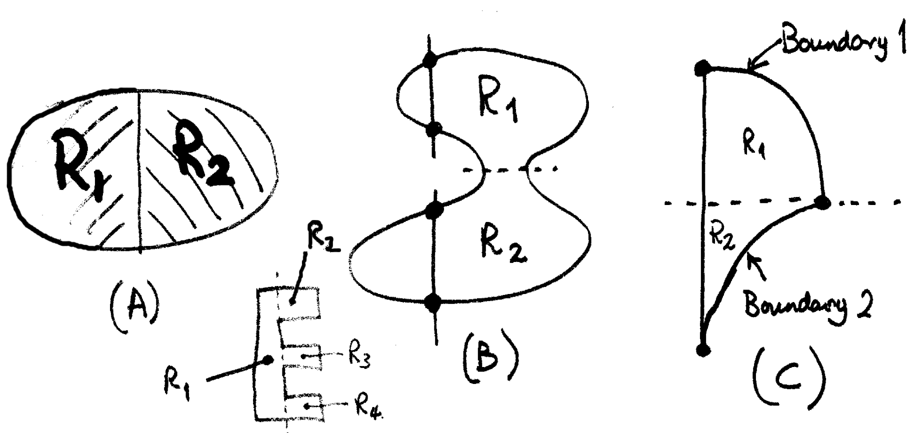

Figure 5.8: Three region problems, A, B C, discussed above

5.3 Volume under a surface

Usefully when we require a volume integral, we turn immediately to triple integrals. However, on introducing double integrals, we noted that they could be thought of as the volume under the surface. The volume is a genuine volume if x, y and $z = f(x, y)$ have the same units.
♣ **Example** Find the volume bounded by the cylinder $x^2 + y^2 = 4$ and the planes $y + z = 4$; $z = -1$.

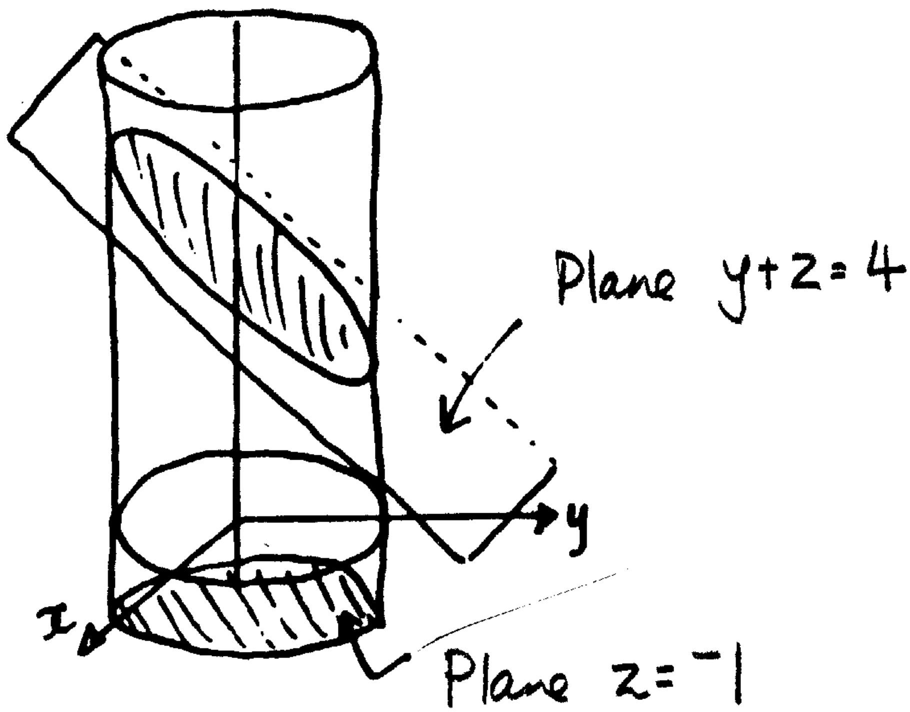

Figure 5.9: Volume under a surface

From the figure it is evident that $z = (4 - y) + 1$ is to be integrated over the circle $x^2 + y^2 = 4$ in the xy-plane. Hence

$$V = \int_{-2}^{2} \int_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} (5-y) dx dy \tag{5.33}$$

$$= \int_{-2}^{2} (5-y) x \Big|_{-\sqrt{4-y^2}}^{\sqrt{4-y^2}} dy \tag{5.34}$$

$$= 2\int_{-2}^{2} (5-y)\sqrt{4-y^2} dy \quad \text{Set } y = 2\sin\theta \tag{5.35}$$

$$= 2\int_{-\pi/2}^{\pi/2} (5 - 2\sin\theta)\sqrt{4 - 4\sin^2\theta}.2\cos\theta d\theta \tag{5.36}$$

$$= 8\int_{-\pi/2}^{\pi/2} (5 - 2\sin\theta)\cos^2\theta d\theta \tag{5.37}$$

$$= 40\int_{-\pi/2}^{\pi/2} \cos^2\theta d\theta - \frac{16}{3}\cos^3\theta \Big|_{-\pi/2}^{\pi/2} \tag{5.38}$$

$$= 40\pi/2 \quad = \quad 20\pi \quad . \tag{5.39}$$

5.4 Change of variables in double integrals

We started out by defining the integral in terms of an arbitrary subdivision into tiles dA, and then for a function $f(x,y)$ pointed out that a convenient and obvious tiling was that with $dA = dxdy$. There is however nothing particularly special about chopping the region of integration into little rectangles. Indeed, as we discussed in Lecture 3, often the shape and symmetry of a problem suggest a different tiling or mesh, which is introduced by a transformation to a new set of variables.
Suppose we specify a transformation to u, v coordinates as

$$x = x(u,v) \quad ; \quad y = y(u,v) \quad . \tag{5.40}$$

Let us keep the function in xy space and consider the new tiling (middle diagram below). The tiling is no longer one of lines of constant x separated by δx and constant y separated by δy, but rather one of lines of constant u separated by δu and constant v separated by δv. Typically, in the xy plane these will curves. (Indeed (u,v) are commonly called *curvilinear coordinates*.)

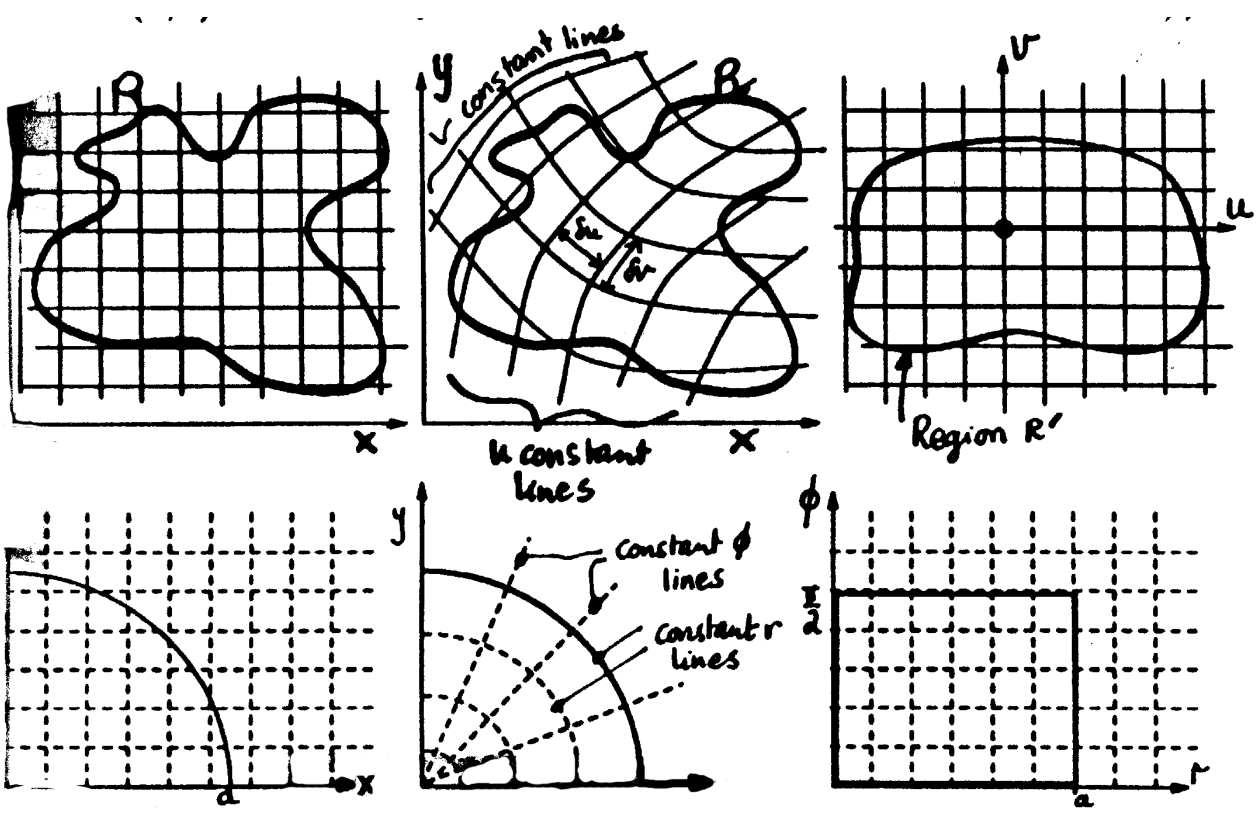

Figure 5.10: Transformation into curvilinear coordinates. Top, an arbitrary transformation: Bottom to Plane Polars. Left, the original in xy space: Middle, lines of constant uv in xy space: Right, lines of constant uv in uv space with the transformed region shape

Now to transform the integral we require three things:

- The function value in (u,v) coordinates.
- The tile area

- The region defined in (u, v) space.

We look at these in turn.

1. A function value for each (u, v). This is obtained by substituting for x and y using $x = x(u, v)$ and $y = y(u, v)$. Thus $F(u, v) = f(x(u, v), y(u, v))$.
2. The area of the tile.

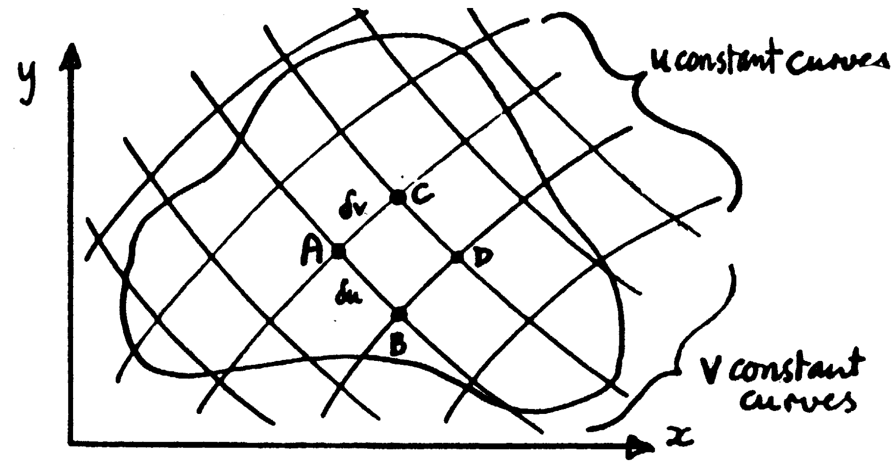

Figure 5.11: The new tile area

NOTICE THAT THIS IS NOT *dudv*. The tile is not a rectangle but rather, in the limit, a parallelogram. Consider the parallelogram ABCD. Its area is given by the modulus of the vector product:

$$\delta A \approx |\vec{AB} \times \vec{AC}|. \tag{5.41}$$

Using $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$ and $\hat{\mathbf{k}}$ as unit vectors in the x, y and z directions, we have But

$$\vec{AB} = (x_B - x_A)\hat{\mathbf{i}} + (y_B - y_A)\hat{\mathbf{j}} \tag{5.42}$$
$$\vec{AC} = (x_C - x_A)\hat{\mathbf{i}} + (y_C - y_A)\hat{\mathbf{j}} \tag{5.43}$$

But

$$(x_B - x_A) \approx \frac{\partial x}{\partial u}\delta u \tag{5.44}$$
$$(x_C - x_A) \approx \frac{\partial x}{\partial v}\delta v \tag{5.45}$$

and similarly for the y terms, so that

$$\vec{AB} = \frac{\partial x}{\partial u}\delta u\hat{\mathbf{i}} + \frac{\partial y}{\partial u}\delta u\hat{\mathbf{j}} \tag{5.46}$$
$$\vec{AC} = \frac{\partial x}{\partial v}\delta v\hat{\mathbf{i}} + \frac{\partial y}{\partial v}\delta v\hat{\mathbf{j}} \tag{5.47}$$

The vector product is thus

$$\vec{AB} \times \vec{AC} = \hat{\mathbf{k}} \left(\frac{\partial x}{\partial u} \delta u \frac{\partial y}{\partial v} \delta v - \frac{\partial x}{\partial v} \delta u \frac{\partial y}{\partial u} \delta v \right) \tag{5.48}$$

So that taking the modulus ($\hat{\mathbf{k}}$ is a unit vector!):

$$\delta A = \left| \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u} \right| \delta u \delta v \quad . \tag{5.49}$$

But the $|\dots|$ term is nothing other than the **modulus of the Jacobian**, so that in the limit

$$dA = dxdy = \left| \frac{\partial(x,y)}{\partial(u,v)} \right| dudv. \tag{5.50}$$

[This is a neater derivation, using properties of determinants and vector products. The vector product of two vectors can be written in determinant form as

$$(a_1\hat{\mathbf{i}} + b_1\hat{\mathbf{j}} + c_1\hat{\mathbf{k}}) \times (a_2\hat{\mathbf{i}} + b_2\hat{\mathbf{j}} + c_2\hat{\mathbf{k}}) = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{vmatrix} \tag{5.51}$$

Thus

$$\vec{AB} \times \vec{AC} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial x}{\partial u}\delta u & \frac{\partial y}{\partial u}\delta u & 0 \\ \frac{\partial x}{\partial v}\delta v & \frac{\partial y}{\partial v}\delta v & 0 \end{vmatrix} \tag{5.52}$$

$$= \hat{\mathbf{k}} \begin{vmatrix} \frac{\partial x}{\partial u}\delta u & \frac{\partial y}{\partial u}\delta u \\ \frac{\partial x}{\partial v}\delta v & \frac{\partial y}{\partial v}\delta v \end{vmatrix} \tag{5.53}$$

$$= \hat{\mathbf{k}}\delta u \delta v \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} \quad . \tag{5.54}$$

In agreement with what we had before.)

3. The last issue we have to consider is the region of integration. We have to express the region R' in uv space using the limits of integration of u and v to exactly cover the physical region R previously expressed in terms of x and y limits. This is shown in the top-right hand part of the figure 5.9, which is expressed in u, v space. One can imagine the whole xy plane being distorted into the uv plane until the curvilinear mesh becomes a straight "cake rack". At the same time the region R becomes distorted into the region R'. As we noted in lecture 3, the shape of the region R' often indicates whether you have made a sensible transformation or not.

♣ **Examples**

1. Evaluate, by transforming to polar coordinates, the area of that part of a circle of radius a centred at the origin of xy space which lies in the $x > 0$ half space.

   The area is $\int_{x=0}^{a} \int_{y=-\sqrt{a^2-x^2}}^{+\sqrt{a^2-x^2}} dydx$, which is $\frac{1}{2}\pi a^2$. Plane polars are $x = r\cos\phi, y = r\sin\phi$. Using the steps above we have

   - $f(x, y) = 1$, so that $F(r, \phi) = 1$.
   - Find mod Jacobian. We have done this before, so here we go ...

     $$\frac{\partial x}{\partial r} = \cos\phi, \quad \frac{\partial x}{\partial \phi} = -r\sin\phi$$

     $$\frac{\partial y}{\partial r} = \sin\phi, \quad \frac{\partial y}{\partial \phi} = r\cos\phi$$

     so

     $$\frac{\partial(x, y)}{\partial(r, \phi)} = r\cos^2\phi - (-r\sin^2\phi) = r \ .$$

     Now $r > 0$, so $|r| = r$.
   - Find region. This is $0 \le r \le a$ and $-\pi/2 \le \phi \le \pi/2$. Note that the new region is a rectangle in r, ϕ space. As we have noted before this is the sort of convenient shape we should hope for when make a "good" transformation.

   So the new integral is

   $$A = \int_{r=0}^{a} \int_{\phi=-\pi/2}^{+\pi/2} r dr d\phi \quad (5.55)$$

   $$= \frac{1}{2}\pi a^2. \quad (5.56)$$

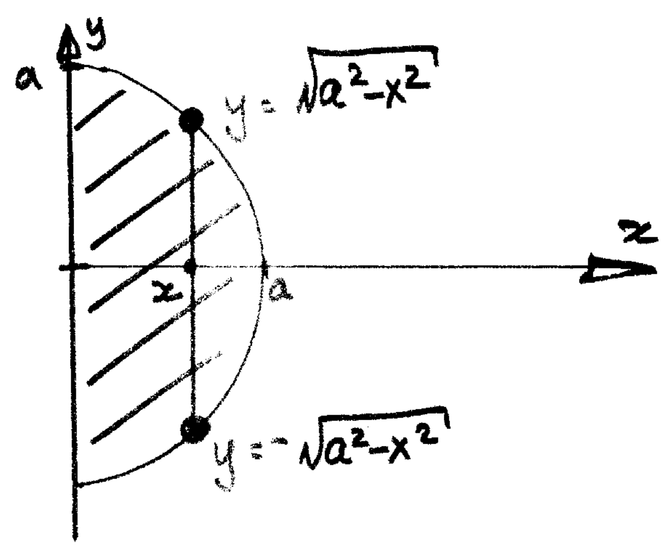

Figure 5.12: Figure for example 1

2. Evaluate

$$I \int_{y=-a}^{+a} \int_{x=-\sqrt{a^2-y^2}}^{+\sqrt{a^2-y^2}} (x^2+y^2)^{3/2} dxdy$$

using the previous transformation.

- Find function in terms of r and ϕ $f(x,y) = (x^2+y^2)^{3/2} = r^3$.
- Modulus of the Jacobian, as before, is r.
- Circle in xy space transforms to rectangle $0 \le r \le a$, $0 \le \phi \le 2\pi$.

Hence

$$I \int_{\phi=0}^{2\pi} \int_{r=0}^{a} r^3 r dr d\phi = 2\pi \frac{a^5}{5}.$$

3. Evaluate $I = \int_0^\infty e^{-x^2} dx$ by expressing I^2 as a double integral.

This is an example of double integration which is not associated a priori with a physical problem (ie like finding the mass ...).

$$I = \lim_{a \longrightarrow \infty} \int_0^a e^{-x^2} dx \quad (5.57)$$

$$\text{Hence } I^2 = \lim_{a \longrightarrow \infty} \int_0^a e^{-x^2} dx \int_0^a e^{-y^2} dy \quad (5.58)$$

$$= \lim_{a \longrightarrow \infty} \int_0^a \int_0^a e^{-x^2-y^2} dxdy \quad (5.59)$$

Transform to plane polars.

$$I^2 = \lim_{a \longrightarrow \infty} \int_{\phi=0}^{\pi/2} \int_{r=0}^{a} e^{-r^2} r dr d\phi \quad (5.60)$$

$$= \lim_{a \longrightarrow \infty} \int_{\phi=0}^{\pi/2} -\frac{1}{2} e^{-r^2} \Big|_0^a d\phi \quad (5.61)$$

$$= \lim_{a \longrightarrow \infty} \left[\frac{\pi}{4} (1 - e^{-a^2} \right] \quad (5.62)$$

$$= \frac{\pi}{4}. \quad (5.63)$$

4. Evaluate the integral

$$I = \int_R \int \left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} \right)^{1/2} dxdy$$

where R is the region enclosed by $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ and $0 < a, 0 < b$.

We shall use this example to show that transformations can be carried out in two stages.

The first transformation could be $X = x/a$, $Y = y/b$.

- The function becomes $(1 - X^2 - Y^2)^{1/2}$
- Modulus of Jacobian is Modulus$\begin{vmatrix} a & 0 \\ 0 & b \end{vmatrix} = |ab| = ab$. (This was obvious enough!)

- Region in xy space is an ellipse, centred at origin. Region in XY space is a circle, radius 1, centred at origin.

Thus

$$I = \int\int (1 - X^2 - Y^2)^{1/2} ab \; dX dY \tag{5.64}$$

Now transform to plane polars: $X = r\cos\phi,\ Y = r\sin\phi$.

$$\begin{aligned} I &= ab \int_{\phi=0}^{2\pi} \int_{r=0}^{1} (1 - r^2)^{1/2} r dr d\phi & (5.65) \\ &= -2ab\pi \frac{2}{3}\frac{1}{2} \left. (1 - r^2)^{3/2} \right|_0^1 & (5.66) \\ &= \frac{2\pi ab}{3}. & (5.67) \end{aligned}$$

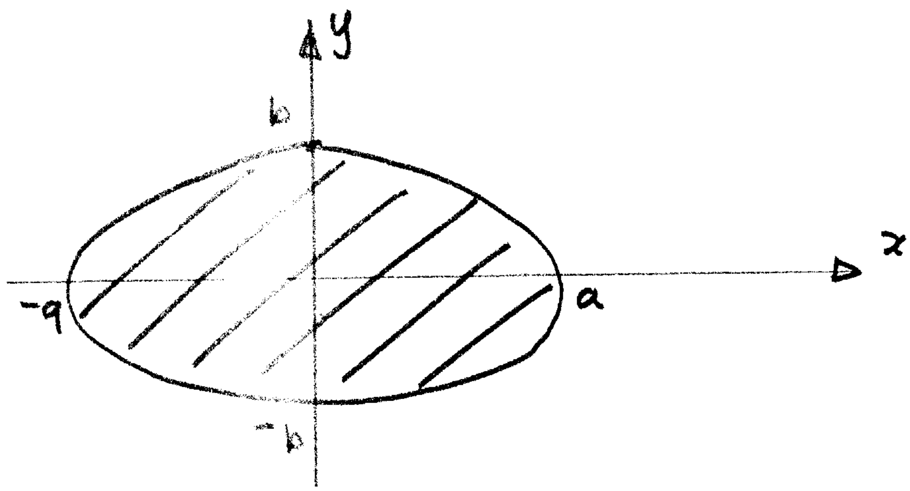

Figure 5.13: Figure for Example 4.

You should check that we could have done all this using just one transformation: $x = ar\cos\phi$ and $y = br\sin\phi$.

5.5 Triple (and higher) integrals

The method we have discussed for double integrals is, as you might have guessed, straightforwardly extensible to triple and higher integrals.
The triple integral is defined, where definitions and caveats are carried over from the case of double integrals, as

$$\iiint_R f(x, y, z) dV = \lim_{n \to \infty} \sum_{i=1}^{n} f_i \delta V_i \tag{5.68}$$

Again we can write $dV = dxdydz$ and evaluate as a repeated integral.

Although the extension is direct, rather more care is required with triple integrals to define the region of integration R in terms of the limits of integration.

♣ **Examples** Find the mass of that part of a sphere or radius 1 which occupies the octant $x^2 + y^2 + z^2 \leq 1$ and which has volume density $\rho = \rho(x, y, z) = x^3yz$.

This is an obvious candidate for a transformation, but for now let's set up the integral in Cartesian coordinates.

$$M = \iiint x^3yz \; dxdydz$$

Let the integration be carried out first over x, with y and z constant. Variable x must go from 0 to $\sqrt{a^2 - y^2 - z^2}$, so that the little volume elements are summed into a stick of cross section δy by δz. Now sum this stick over y which varies from 0 to $\sqrt{a^2 - z^2}$ to give a slice parallel to the xy plane. Finally sum the slices along z. z varies from 0 to a. Thus the mass is

$$\begin{aligned}
M &= \int_{z=0}^{a} \int_{y=0}^{\sqrt{a^2-z^2}} \int_{x=0}^{\sqrt{a^2-y^2-z^2}} x^3yz \; dxdydz && (5.69)\\
&= \frac{1}{4} \int_{z=0}^{a} z \int_{y=0}^{\sqrt{a^2-z^2}} y(a^2 - y^2 - z^2)^2 \; dydz && (5.70)\\
&= -\frac{1}{24} \int_{z=0}^{a} \left[z(a^2 - y^2 - z^2)^3 \right]_{y=0}^{\sqrt{a^2-z^2}} && (5.71)\\
&= -\frac{1}{192} \left[(a^2 - z^2)^4 \right]_0^a && (5.72)\\
&= \frac{a^8}{192}. && (5.73)
\end{aligned}$$

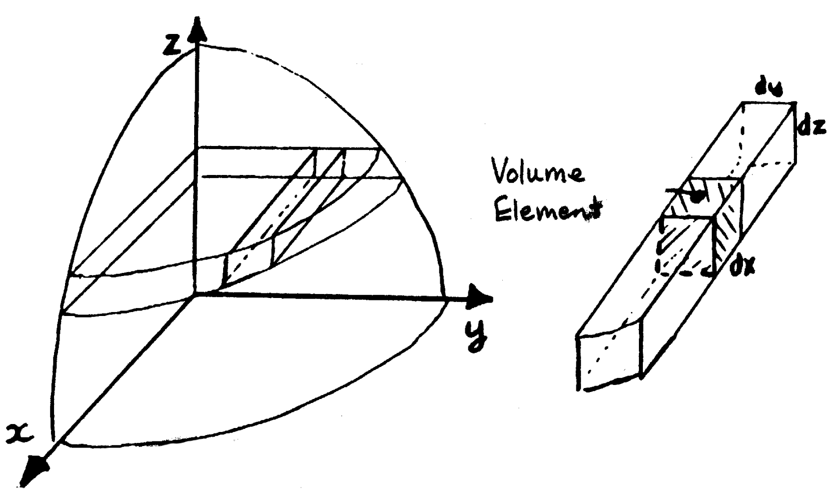

Figure 5.14:

5.6 Transformations for triple integrals

Suppose the integral is

$$I = \iiint_R f(x, y, z) \; dxdydz \qquad (5.74)$$

and we wish to change to variables u, v, w where

$$x = x(u, v, w) \quad y = y(u, v, w) \quad z = z(u, v, w)$$

The extension of the transformation for double integrals to triple (and higher) integrals is direct, namely:

$$I = \iiint_{R'} f(x(u,v,w), y(u,v,w), z(u,v,w)) \left| \frac{\partial(x,y,z)}{\partial(u,v,w)} \right| dudvdw. \tag{5.75}$$

where the Jacobian is the determinant

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{vmatrix} \tag{5.76}$$

To prove this we need to extend the previous tile construction to 3D, construct the volume between lines

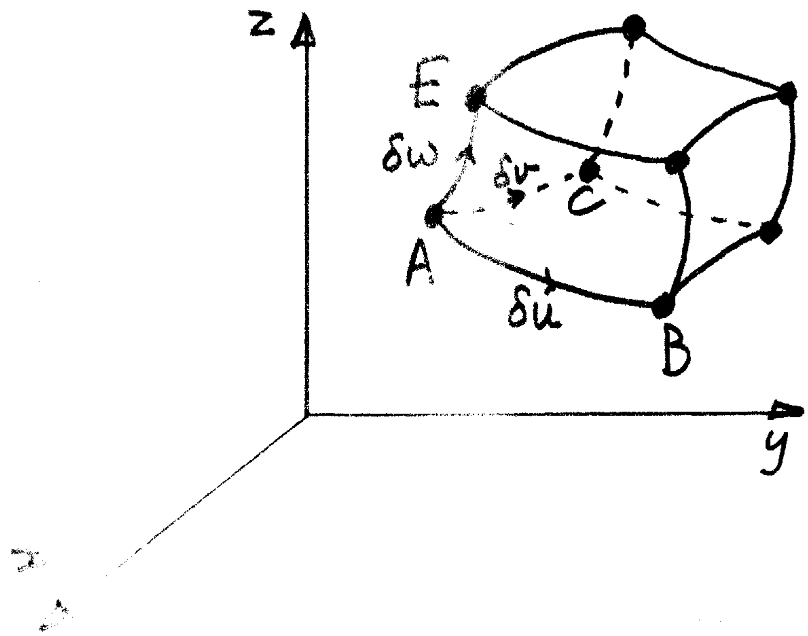

Figure 5.15: Curvilinear coordinates and the volume element

ı by the scalar triple product:

$$\delta V \approx (\vec{AB} \times \vec{AC}) \cdot \vec{AE} \quad . \tag{5.77}$$

But now

$$\vec{AB} = (x_B - x_A)\hat{\mathbf{i}} + (y_B - y_A)\hat{\mathbf{j}} + (x_C - x_A)\hat{\mathbf{k}} \tag{5.78}$$

$$= \frac{\partial x}{\partial u}\delta u\hat{\mathbf{i}} + \frac{\partial y}{\partial u}\delta u\hat{\mathbf{j}} + \frac{\partial x}{\partial w}\delta w\hat{\mathbf{k}} \tag{5.79}$$

(5.80)

with similar expressions for $\vec{AC}$ and $\vec{AE}$. Using the expression for cross products we have

$$\vec{AB} \times \vec{AC} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \frac{\partial x}{\partial u}\delta u & \frac{\partial y}{\partial u}\delta u & \frac{\partial z}{\partial u}\delta u \\ \frac{\partial x}{\partial v}\delta v & \frac{\partial y}{\partial v}\delta v & \frac{\partial z}{\partial v}\delta v \end{vmatrix} \tag{5.81}$$

Taking the scalar product with $\vec{AE}$ gives

$$\vec{AB} \times \vec{AC} \cdot \vec{AE} = \begin{vmatrix} \frac{\partial x}{\partial w}\delta w & \frac{\partial y}{\partial w}\delta w & \frac{\partial z}{\partial w}\delta w \\ \frac{\partial x}{\partial u}\delta u & \frac{\partial y}{\partial u}\delta u & \frac{\partial z}{\partial u}\delta u \\ \frac{\partial x}{\partial v}\delta v & \frac{\partial y}{\partial v}\delta v & \frac{\partial z}{\partial v}\delta v \end{vmatrix} \tag{5.82}$$

Taking δw, δu and δv out as factors, and doing two row swaps (one swap changes the sign, another swap changes it back again)

$$dV = \vec{AB} \times \vec{AC} \cdot \vec{AE} = \delta u \delta v \delta w \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & \frac{\partial z}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & \frac{\partial z}{\partial v} \\ \frac{\partial x}{\partial w} & \frac{\partial y}{\partial w} & \frac{\partial z}{\partial w} \end{vmatrix} \tag{5.83}$$

as required.

♣ **Examples**

1. Cylindrical polar coordinates are related to Cartesian coordinates by

$$x = r\cos\phi \quad y = r\sin\phi \quad z = z$$

Compute the moment of inertia I_x about the x-axis of the solid cylinder of $x^2 + y^2 \leq a^2$ bounded by $z = 0$ and $z = b$. Assume uniform density.

The moment of inertial of a little element $dxdydz$ about the x-axis is $\rho(y^2 + z^2)dxdydz$. So

$$I_x = \rho \textstyle\iiint_R (y^2 + z^2)dxdydz \tag{5.84}$$

There are again three tasks in making the transformation: (i) rewrite the function, (ii) derive the Mod Jacobian, (iii) rewrite the region.

- The function $(y^2 + z^2)$ is $(r^2 \sin^2 \phi + z^2)$.
- The Jacobian is

$$\begin{vmatrix} \cos\phi & r\sin\phi & 0 \\ \sin\phi & -r\cos\phi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r$$

  Now as $r > 0$ always, $|r| = r$.
- The region is $0 \leq r \leq a$, $0 \leq \phi \leq 2\pi$ and $0 \leq z \leq b$.

So

$$I_x = \rho \int_{r=0}^{a} r dr \int_{\phi=0}^{2\pi} d\phi \int_{z=0}^{b} (r^2 \sin^2 \phi + z^2) dz \quad (5.85)$$

$$= \rho \int_{r=0}^{a} r dr \int_{\phi=0}^{2\pi} (br^2 \sin^2 \phi + b^3/3) d\phi \quad (5.86)$$

Now

$$\int_0^{2\pi} \sin^2 \phi d\phi = \frac{1}{2} \int_0^{2\pi} (1 - \cos 2\phi) d\phi = \pi$$

so that

$$I_x = \rho \int_{r=0}^{a} r dr (r^2 \pi b + \frac{b^3}{3} 2\pi) \quad (5.87)$$

$$= \rho \left[\frac{a^4 \pi b}{4} + \frac{b^3 2\pi a^2}{6} \right] \quad (5.88)$$

$$= \rho \frac{a^2 b \pi}{12} (3a^2 + 4b^2). \quad (5.89)$$

5.7 Standard transformations revisited

Here we specialize the general theory to look at the shapes of the area and volume elements generated by the standard transformations.

1. **2D Cartesian to plane polars.** The Jacobian is r (see Lecture 3) which is always positive so that the modulus is r. Thus the element of area is $dA = rdrd\phi$. This is readily seen from the diagram.

2. **3D Cartesian to spherical polars.** The Jacobian is $r^2 \sin\theta$ (see Lecture 3). Now theta ranges from 0 to π, so that the Jacobian is always positive. Thus the element of area is $dA = r^2 \sin\theta drd\theta d\phi$. Again this is readily seen from the diagram.

3. **3D Cartesian to cylindrical polars.** The Jacobian is r (see Lecture 3) so that the Jacobian is always positive. Thus the element of area is $dA = rdrd\phi dz$. Again this is readily seen from the diagram.

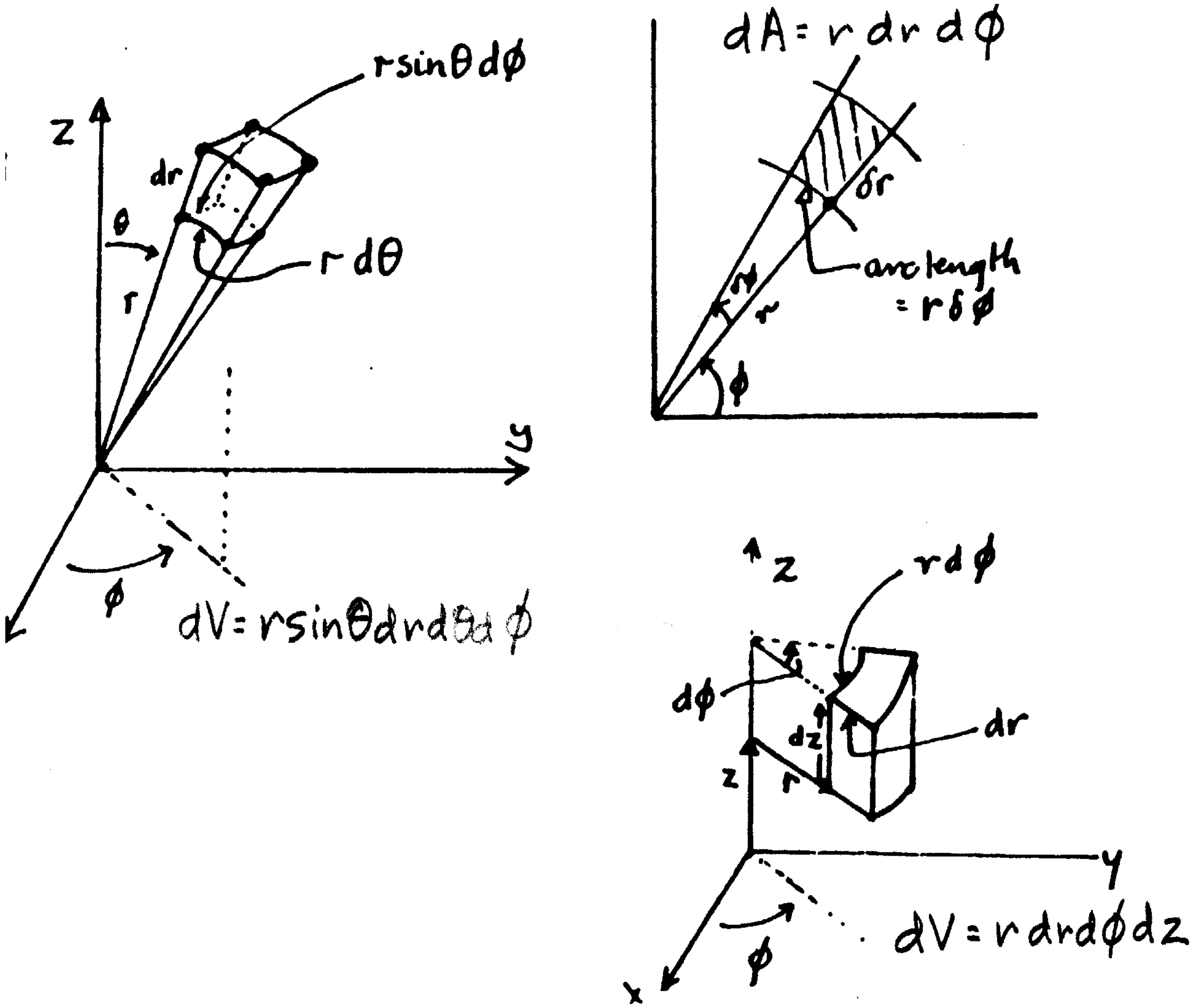

Figure 5.16: Area and volume elements for the standard transformations

Chapter 6

Volumes of revolution, line and surface integrals

6.1 Volumes of Revolution: a double integral

Suppose a closed curve C enclosing a region R lying in the postive quadrant of the xy plane is rotated about the x-axis. It sweeps out a volume known as a *volume of revolution*.

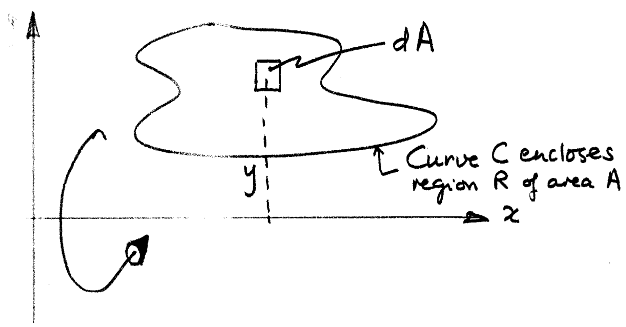

Figure 6.1: Volume of revolution

To calculate the volume of the v.o.r. first calculate the element of volume created by rotation an area element dA about the x-axis. This is obviously:

$$dV = 2\pi y dA \quad . \tag{6.1}$$

The total volume of revolution is thus:

$$V = \iint_R 2\pi y dA \tag{6.2}$$

$$= \iint_R 2\pi y dx dy \quad . \tag{6.3}$$

But the y coordinate of the centroid of the region R is given as

$$\bar{y} = \frac{\iint_R y dx dy}{\iint_R dx dy} \tag{6.4}$$

$$= \frac{\iint_R y dx dy}{A} \tag{6.5}$$

where A is the area of the region. In other words we arrive at the intuitive result that

$$V = 2\pi A \bar{y} \tag{6.6}$$

6.2 Line integrals

6.2.1 Defining a curve

A *plane curve* is a one-parameter set of points $P(x, y)$ where $x = x(t)$ and $y = y(t)$, where t (the "one-parameter") takes all values in some range $a \leq t \leq b$. If the points given by $t = a$ and $t = b$ coincide then the curve is said to be *closed*, and is otherwise *open*. If the curve does not cross itself, it is said to be *simple*. If it does cross itself, the curve is *non-simple*, and the points of intersection are called *multiple points*. If the only multiple point is the $t = a$ and $t = b$ point, then the curve is said to be a *simple closed curve*. A curve consisting of several simple curves joined together is called *piecewise smooth*.

6.2.2 Defining the line integral

Suppose we have a function $f(x, y)$ defined in the xy plane. Thus for each value of the parameter t, we have an associated value of the function. Now suppose we form the sum

$$I_n = \sum_{i=1}^{n} f(\bar{x}_i, \bar{y}_i)\delta s_i \tag{6.7}$$

Here δs_i is the length of the curve between points P_{i-1} and P_i and $\bar{x}_i, \bar{y}_i$ are chosen somewhere between P_{i-1} and P_i.

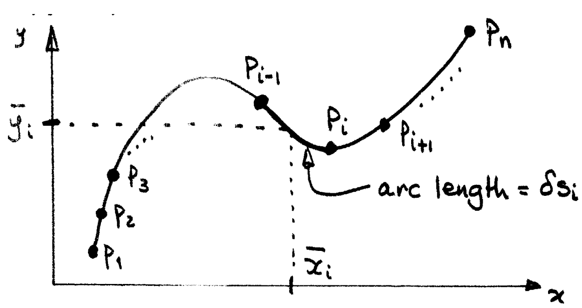

Figure 6.2: The line integral

If as $\delta s_i \to 0$ (nomatter how exactly) the limit $\lim_{n\to\infty} I_n$ exists and has the same value, then

$$\int_C f(x, y)ds = \lim_{n\to\infty} I_n = \lim_{n\to\infty} \sum_{i=1}^{n} f(\bar{x}_i, \bar{y}_i)\delta s_i \tag{6.8}$$

is called the line integral of $f(x, y)$ along the curve C defined by $x = x(t)$ and $y = y(t)$.

6.2.3 Evaluating line integrals

There are several ways of evaluating these integrals, depending on how x and y are related.

Method 1A. Suppose $y = y(x)$. We can turn $f(x, y)$ into a function of x alone. Now the question is how to express ds? Using Pythagoras' theorem we see that in the limit

$$(ds)^2 = (dx)^2 + (dy)^2 \tag{6.9}$$

Thus

$$\frac{ds}{dx} = \left(1 + \left[\frac{dy}{dx}\right]^2\right)^{1/2} \tag{6.10}$$

and so the line integral is

$$\int_C f(x,y)ds = \int_{x=a}^{x=b} f(x,y(x)) \left(1 + \left[\frac{dy}{dx}\right]^2\right)^{1/2} dx \tag{6.11}$$

Method 1B. A similar expression can be found when $x = x(y)$.

$$\int_C f(x,y)ds = \int_{y=a}^{y=b} f(x(y),y) \left(1 + \left[\frac{dx}{dy}\right]^2\right)^{1/2} dy \tag{6.12}$$

Method 2. Often, curves are expressed as functions of arc-length, s. In other words the parameter t equals s. Then $y = y(s)$ and $x = x(s)$ and the integrand can be written as a function of s alone. Thus the line integral is

$$\int_C f(x,y)ds = \int_{s=s_1}^{s=s_2} f(x(s),y(s))ds. \tag{6.13}$$

Method 3. Finally we consider the case when the parameter t is an arbitrary parameter. We have $x = x(t)$ and $y = y(t)$. But t must be dependent on s alone, so that ds/dt must exist. Indeed in the limit

$$\frac{ds}{dt} = \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2\right]^{1/2} \tag{6.14}$$

and the line integral is

$$\int_C f(x,y)ds = \int_{t=t_1}^{t=t_2} f(x(t),y(t)) \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2\right]^{1/2} dt. \tag{6.15}$$

♣ Examples

1. Show that $\int_C (x^2 + y^2)ds$ where C is the arc of a circle $x = a\cos(s/a)$ and $y = a\sin(s/a)$ $(0 \le s \le \pi a)$ is πa^3.

   This is an example of method 2. $(x^2 + y^2) = a^2$ so that the integral is

   $$\int_0^{\pi a} a^2 ds = a^2 s\Big|_0^{\pi a} = \pi a^3 \quad . \tag{6.16}$$

2. Derive $I = \int_C (x - y^2)ds$ (i) where C is a segment of the straight line $y = 2x$ such that $0 \le x \le 1$; and (ii) where curve C is the arc of a circle $x = a\cos t$, $y = a\sin t$, $0 \le t \le \pi/2$.

   Using method 1A,

   $$ds = \sqrt{1 + 2^2}dx \quad . \tag{6.17}$$

Thus

$$I = \int_C (x-y^2)ds \tag{6.18}$$
$$= \int_0^1 (x-4x^2)\sqrt{5}dx \tag{6.19}$$
$$= \sqrt{5}\left[\frac{x^2}{2} - \frac{4}{3}x^3\right|_0^1 = -\frac{5\sqrt{5}}{6} \ . \tag{6.20}$$

The second curve us and example of Method 3.

$$ds = [(-a\sin t)^2 + (a\cos t)^2]^{1/2}dt = adt \tag{6.21}$$

Thus

$$I = \int_C (x-y^2)ds \tag{6.22}$$
$$= \int_0^{\pi/2} (a\cos t - a^2\sin^2 t)adt \tag{6.23}$$
$$= \left[a^2\sin t - a^3(\frac{t}{2} - \frac{\sin 2t}{4})\right|_0^{\pi/2} \tag{6.24}$$
$$= a^2 - \pi a^3/4 \ . \tag{6.25}$$

6.2.4 Line integrals independent of path

A region R is said to be connected if every two points in R can be joined by a piecewise smooth curve that lies wholly in R. If **every** curve can be shrunk by an arbitrary continuous deformation and still be contained in the region, then the region is simply connected. If this is not the case, then the region is multiply connected.
Let $\phi(x,y)$ be some differentiable (and hence single valued) function defined in the connected region R. Its total differential is

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy \tag{6.26}$$

Now suppose we formed the line integral

$$I = \int_C d\phi = \int_C \left[\frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy\right] \tag{6.27}$$

over some arbitrary C joining points P_1 and P_2 in R.
This integral is

$$I = \int_C d\phi = \phi(P_2) - \phi(P_1) \tag{6.28}$$

which is independent of the precise path C taken between P_1 and P_2.
So, if the integrand in the line integral

$$\int_C [M(x,y)dx + N(x,y)dy]$$

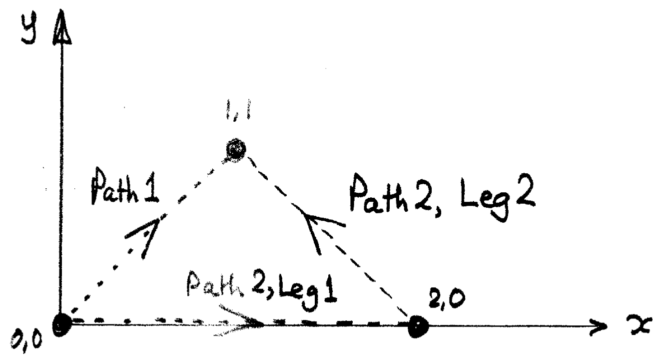

Figure 6.3: Line integral indepent of path

is the total differential $d\phi$ of some function $\phi(x, y)$ with continuous partial derivatives, then the line integral between any two points is independent of path. Equivalently, the line integral around a piecewise smooth *closed* path in R is identically zero.

♣ **Example**

[Let us first choose some function from which to create the total differential, and then using this to form the line integral. The function is $f(x, y) = (x^2 + y^2 x)$. Thus the differential is

$$df = (2x + y^2)dx + (2xy)dy \quad . \tag{6.29}$$

So the question might be posed as:]

Show that the line integral

$$\int_C [(2x + y^2)dx + (2xy)dy]$$

from $(0, 0)$ to $(1, 1)$ is the same for the paths

1. a straight line from $(0, 0)$ to $(1, 1)$.
2. a path in two stages from $(0, 0)$ to $(2, 0)$, then from $(2, 0)$ to $(1, 1)$.

Path 1. The line integral for the first path is

$$\int_C [(2x + y^2)dx + (2xy)dy] = \int_0^1 (2x + x^2)dx + \int_0^1 (2y^2)dy \tag{6.30}$$

$$= 4/3 \quad + \quad 2/3 \quad = 2 \quad . \tag{6.31}$$

Path 2. The first leg of the second path has $y = 0$, thus

$$\int_C [(2x + y^2)dx + (2xy)dy] = \int_0^2 2x dx \tag{6.32}$$

$$= 4 \tag{6.33}$$

The second leg is $y = 2 - x$ so that

$$\int_C \quad [\quad (2x + y^2)dx + (2xy)dy] \tag{6.34}$$

$$= \int_2^1 (2x + (2 - x)^2)dx + \int_0^1 2(2 - y)y dy \tag{6.35}$$

$$= \left[-x^2 + x^3/3 + 4x\right|_2^1 + \left[2(-y^3/3 + y^2)\right|_0^1 \tag{6.36}$$

$$= -10/3 \quad + \quad 4/3 = -2 \tag{6.37}$$

So the total for this path is $4 - 2 = 2$, the same as the first path.

6.3 Surface Integrals: a double integral

Just as line integrals were concerned with integrating alonge a curve some function $f(s)$ which is defined at each point s on that curve, so surface integrals are concernec with integrating over a surface some function $f(S)$ defined at every point S on the surface.
(It is important to realize that the surface is some real surface, *not* the notional surface of the function f.)
For example, we may have beaten a sheet of metal (of uniform surface density 1) into a curved object. The thickess T of the metal varies from point to point on the surface. What is the mass of some region P of the surface? It is

$$M = \iint_P T(S) dS \quad . \tag{6.38}$$

The key to understanding this problem is to notice how a small area patch dS on the surface projects onto a patch dA in the xy plane. Consider the surface shown in the figure. The surface is described by $z = f(x, y)$. On the flat part of the surface parallel to the xy plane it is obvious that $dS = dA$, but as the surface becomes steeper then $dS > dA$. Indeed if the surface were like a cliff face, then a finite dA would back project onto an infinite dS.

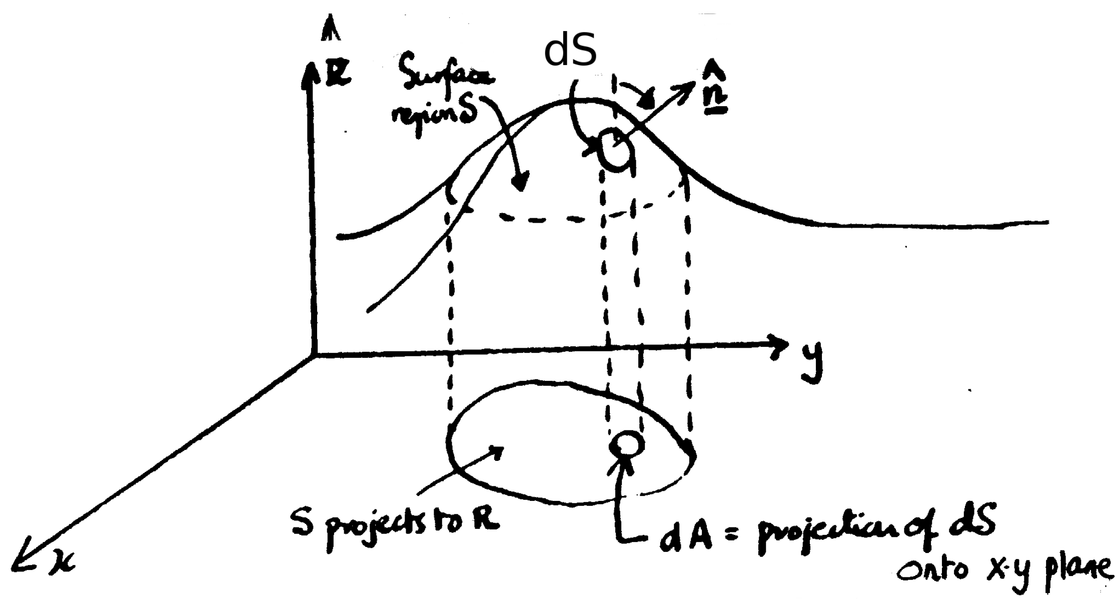

Figure 6.4: Surface integrals

Let us construct $\hat{\mathbf{n}}$, a unit normal to the surface patch dS. It is clear from the projection operation that

$$dS = \frac{1}{\cos\theta} dA \tag{6.39}$$

where θ is the angle between the z-axis and the surface normal.
So, returning to $M = \iint_P T(S) dS$, what can we say. Let us suppose for now that each point on the xy-plane maps onto a unique point on the surface. Thus we can replace $T(S)$ by $T(x, y)$. We also know that the angle θ depends on position on the surface, so that $\theta = \theta(x, y)$. Thus we can rewrite the integral as

$$M = \iint_R T(x, y) \frac{1}{\cos\theta(x, y)} dA \tag{6.40}$$

where the region of integration in the xy plane is R, the *projection* of P onto the xy plane.
The question now is how to find $\cos\theta(x,y)$? The scalar product $\mathbf{a}\cdot\mathbf{b}$ between two vectors is $\mathbf{a}\cdot\mathbf{b} = ab\cos\alpha$ where α is the angle between the two vectors. Now θ is the angle between the surface normal and the z-axis. If we attached unit vectors $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ to the x, y, z axes, we see that

$$\cos\theta = \hat{\mathbf{n}}\cdot\hat{\mathbf{k}} \quad . \tag{6.41}$$

So what is the vector $\hat{\mathbf{n}}$?
Suppose we moved some small distance *on the surface*. The change in z is related to those in x and y by the total differential

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \quad , \tag{6.42}$$

or

$$\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy - dz = 0 \quad . \tag{6.43}$$

But the movement (dx, dy, dz) can be written as a vector as

$$\hat{\mathbf{i}}dx + \hat{\mathbf{j}}dy + \hat{\mathbf{k}}dz \quad . \tag{6.44}$$

Now if we remain on the surface, so that eq 6.43 is satisfied, the vector $\hat{\mathbf{i}}dx + \hat{\mathbf{j}}dy + \hat{\mathbf{k}}dz$ must be a tangent to the surface.
But using the dot product we could write equation 6.43 as

$$(\hat{\mathbf{i}}\frac{\partial f}{\partial x} + \hat{\mathbf{j}}\frac{\partial f}{\partial y} - \hat{\mathbf{k}})\cdot(\hat{\mathbf{i}}dx + \hat{\mathbf{j}}dy + \hat{\mathbf{k}}dz) = 0 \quad . \tag{6.45}$$

But the equation $\mathbf{a}\cdot\mathbf{b} = 0$ means that $\mathbf{a}$ and $\mathbf{b}$ are orthogonal, so $(\hat{\mathbf{i}}\frac{\partial f}{\partial x} + \hat{\mathbf{j}}\frac{\partial f}{\partial y} - \hat{\mathbf{k}})$ is orthogonal to the tangent, so it must be the normal to the surface.
Now in fact the normal could be $\pm(\hat{\mathbf{i}}\frac{\partial f}{\partial x} + \hat{\mathbf{j}}\frac{\partial f}{\partial y} - \hat{\mathbf{k}})$, but as we have taken the surface to be above the xy plane we would like the $\hat{\mathbf{k}}$ component to be positive. Hence we choose the normal as

$$(-\hat{\mathbf{i}}\frac{\partial f}{\partial x} - \hat{\mathbf{j}}\frac{\partial f}{\partial y} + \hat{\mathbf{k}}) \quad . \tag{6.46}$$

The last step is to find the unit normal as:

$$\hat{\mathbf{n}} = (-\hat{\mathbf{i}}\frac{\partial f}{\partial x} - \hat{\mathbf{j}}\frac{\partial f}{\partial y} + \hat{\mathbf{k}})\left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1\right]^{-1/2} \tag{6.47}$$

Taking the scalar product with $\hat{\mathbf{k}}$ we have

$$\cos\theta = \hat{\mathbf{n}}\cdot\hat{\mathbf{k}} = \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1\right]^{-1/2} \tag{6.48}$$

and

$$\frac{1}{\cos\theta} = \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1\right]^{1/2} \tag{6.49}$$

So, returning to the original problem, we find for a surface $z = f(x, y)$:

$$M = \iint_R T(x,y) \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y}) \right)^2 + 1 \right]^{1/2} dxdy \tag{6.50}$$

where, note again, the region of integration is the projection of the actual surface onto the xy plane.

♣ **Example**

Find the surface area of the sphere $x^2+y^2+z^2 = a^2$ which is cut off by the cylinder $(x-a/2)^2 + y^2 = (a/2)^2$. Using symmetry, we need to consider only the positive octant, then quadruple the

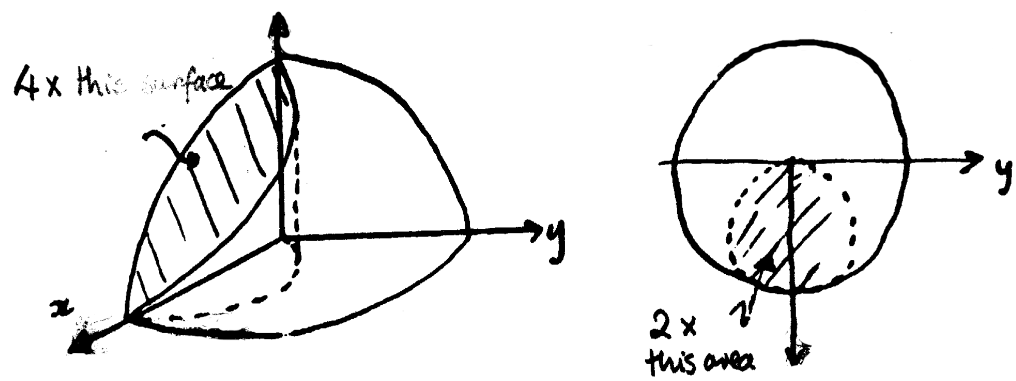

Figure 6.5:

answer. The surface area is given by

$$S = 4\iint_R \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y}) \right)^2 + 1 \right]^{1/2} dxdy \quad . \tag{6.51}$$

The surface is that of the sphere, so that

$$z = (a^2 - x^2 - y^2)^{1/2} \tag{6.52}$$

$$z_x = \frac{-x}{(a^2 - x^2 - y^2)^{1/2}} \tag{6.53}$$

$$z_y = \frac{-y}{(a^2 - x^2 - y^2)^{1/2}} \tag{6.54}$$

The region is given by integrating y from 0 to $\sqrt{ax - x^2}$ followed by integration over x from 0 to a. Thus the surface area is

$$S = 4 \int_{x=0}^{a} \int_{y=0}^{\sqrt{ax-x^2}} \left[\frac{x^2+y^2}{a^2-x^2-y^2} + 1 \right]^{1/2} dydx \tag{6.55}$$

$$= 4 \int_{x=0}^{a} \int_{y=0}^{\sqrt{ax-x^2}} \frac{a}{\sqrt{a^2-x^2-y^2}} dydx \tag{6.56}$$

To evaluate this transform to cylindrical polars. The equations of the cylinder and sphere are:

$$r = a\cos\phi \tag{6.57}$$

$$z = \sqrt{a^2 - x^2 - y^2} = \sqrt{a^2 - r^2} \tag{6.58}$$

Thus

$$S = 4\int_{\phi=0}^{\pi/2}\int_{r=0}^{a\cos\phi}\frac{ardrd\phi}{\sqrt{a^2-r^2}} \quad (6.59)$$

$$= 4a^2\left(\frac{\pi}{2}-1\right) \quad (6.60)$$

6.3.1 Standard tranformations revisited – again

Spherical polar coordinates.
By considering the volume element for spherical polars, it is evident that the surface element in spherical polar coordinates (r,θ,ϕ) is $dS = r^2\sin\theta d\theta d\phi$. But when considering the element, we must also specify the direction of the normal to the surface. In spherical coordinates, the normal is obviously just $\hat{\mathbf{r}}$, a unit vector in the r direction. In cartesian coordinates it is $\hat{\mathbf{i}}\sin\theta\cos\phi + \hat{\mathbf{j}}\sin\theta\sin\phi + \hat{\mathbf{k}}\cos\theta$.

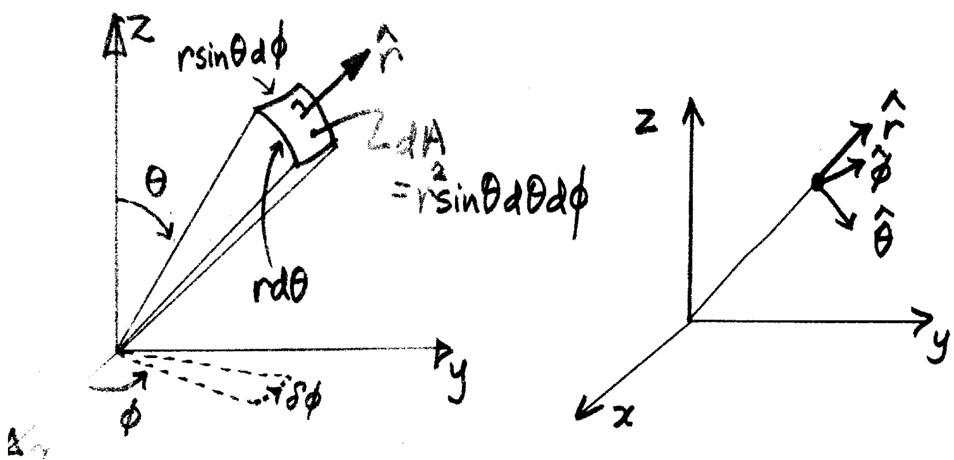

Figure 6.6: Surface patch in spherical polar coordinates

♣ **Example** Calculate the surface area of a sphere of radius R.
In spherical polars this is just

$$S = \iint_{sphere} dS \quad (6.61)$$

$$= \int_{\theta=0}^{\pi}\int_{\phi=0}^{\phi=2\pi} R^2\sin\theta d\phi d\theta \quad (6.62)$$

$$= R^2 2\pi[-\cos\theta|_0^\pi \quad (6.63)$$

$$= 4\pi R^2 \quad . \quad (6.64)$$

An alternative which you might try is to proceed as in the earlier example involving the sphere and cylinder. As there, you will need to switch to cylindrical polars to complete the integral.

Cylindrical polar coordinates.
The surface area element is $dS = rdzd\phi$, and the normal direction is $\hat{\mathbf{r}}$, or in Cartesians $\hat{\mathbf{i}}\cos\phi + \hat{\mathbf{j}}\sin\phi$.

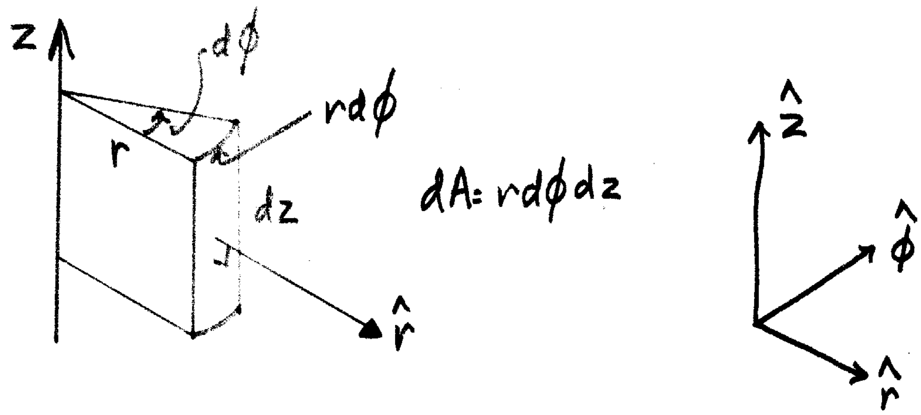

Figure 6.7: Surface patch in cylindrical polar coordinates

6.4 Line and surface integrals involving vectors

Line and surface integrals occur rather frequently, but most often as a result of vector operations. We cannot get too involved in *vector field theory* but here are a couple of examples.

6.4.1 Line integrals

Consider a body with a force $\mathbf{F}$ acting on it in the xy plane and let its components be functions of position:

$$\mathbf{F} = f(x,y)\hat{\mathbf{i}} + g(x,y)\hat{\mathbf{j}} \quad . \tag{6.65}$$

Suppose the body moves a small amount $d\mathbf{s} = \hat{\mathbf{i}}dx + \hat{\mathbf{j}}dy$ in the xy plane. Then the work done on the body is

$$dW = \mathbf{F} \cdot d\mathbf{s} \quad . \tag{6.66}$$

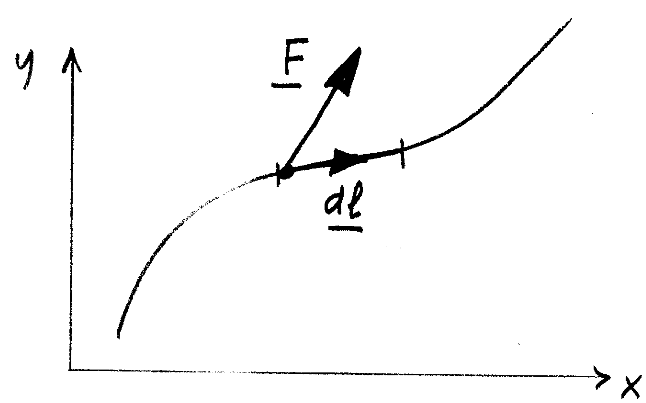

Figure 6.8: Work done on some body by a variable force over some path

Now let the body follows some path C. Then the total work done on the body is just:

$$W = \int_C \mathbf{F} \cdot d\mathbf{s} = \int_C [f(x,y)dx + g(x,y)dy] \quad , \tag{6.67}$$

that is, a line integral!
Now suppose that

$$\frac{\partial}{\partial y} f(x,y) = \frac{\partial}{\partial x} g(x,y) \ . \tag{6.68}$$

The integrand would we a total different, and the work done would only depend on the start and end points of the path, not on the actual path taken. Equivalently, the work done traversing a loop is zero.
Such a force field is said to be *conservative.*

6.4.2 Surface integrals

Consider the fountain sketched below. A particular volume of water is pumped per second into the fountain head, and (using the principal of what goes up must come down) we know that the volume of water leaving any of the surfaces drawn around the fountain head must be the same as that entering the fountain head.

How would we begin to analyze this situation using surface integrals? Suppose we are able to define, at any position in space, the water *flux* $\mathbf{v}$. This is a vector whose magnitude is the volume of water per second crossing a unit area, and whose direction indicates the direction of flow.
The total flow per second through a surface S is then

$$\int_S \mathbf{v} \cdot d\mathbf{S} \tag{6.69}$$

where $d\mathbf{S} = \hat{\mathbf{n}} dS$ and $\hat{\mathbf{n}}$ is the unit outward normal to the surface. Given an expression for $\mathbf{v}$, ie

$$\mathbf{v} = f(x,y,z)\hat{\mathbf{i}} + g(x,y,z)\hat{\mathbf{j}} + h(x,y,z)\hat{\mathbf{k}}$$

one can take the dot product and perform the integral using the methods described earlier.
This type of integral occurs in optics, electricity, magnetism, hydrodynamics, etc etc.

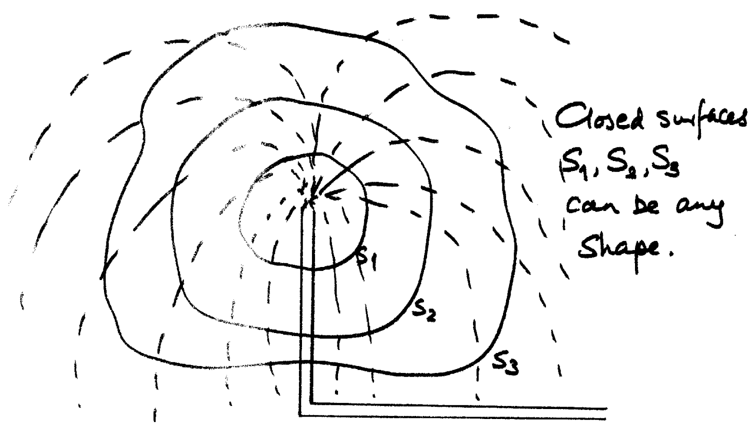

Figure 6.9: Water flux from a fountain