Lecture notes for Math 417-517
Multivariable Calculus

J. Dimock
Dept. of Mathematics
SUNY at Buffalo
Buffalo, NY 14260

October 18, 2019

Contents

1 multivariable calculus

1.1 vectors

We start with some definitions. A real number x is positive, zero, or negative and is rational or irrational. We denote

$$\mathbb{R} = \text{set of all real numbers } x \tag{1}$$

The real numbers label the points on a line once we pick an origin and a unit of length. Real numbers are also called *scalars*

Next define

$$\mathbb{R}^2 = \text{all pairs of real numbers } x = (x_1, x_2) \tag{2}$$

The elements of $\mathbb{R}^2$ label points in the plane once we pick an origin and a pair of orthogonal axes. Elements of $\mathbb{R}^2$ are also called (2-dimensional) *vectors* and can be represented by arrows from the origin to the point represented.

Next define

$$\mathbb{R}^3 = \text{all triples of real numbers } x = (x_1, x_2, x_3) \tag{3}$$

The elements of $\mathbb{R}^3$ label points in space once we pick an origin and three orthogonal axes. Elements of $\mathbb{R}^3$ are (3-dimensional) vectors. Especially for $\mathbb{R}^3$ one might emphasize that x is a vector by writing it in bold face $\mathbf{x} = (x_1, x_2, x_3)$ or with an arrow $\vec{x} = (x_1, x_2, x_3)$ but we refrain from doing this for the time being.

Generalizing still further we define

$$\mathbb{R}^n = \text{all n-tuples of real numbers } x = (x_1, x_2, \dots, x_n) \tag{4}$$

The elements of $\mathbb{R}^n$ are the points in n-dimensional space and are also called (n-dimensional) vectors

Given a vector $x = (x_1, \dots, x_n)$ in $\mathbb{R}^n$ and a scalar $\alpha \in \mathbb{R}$ the product is the vector

$$\alpha x = (\alpha x_1, \dots, \alpha x_n) \tag{5}$$

Another vector $y = (y_1, \dots, y_n)$ can to added to x to give a vector

$$x + y = (x_1 + y_1, \dots, x_n + y_n) \tag{6}$$

Because elements of $\mathbb{R}^n$ can be multiplied by a scalar and added it is called a *vector space*. We can also subtract vectors defining $x - y = x + (-1)y$ and then

$$x - y = (x_1 - y_1, \dots, x_n - y_n) \tag{7}$$

For two or three dimensional vectors these operations have a geometric interpretation. αx is a vector in the same direction as x (opposite direction if $\alpha < 0$) with length

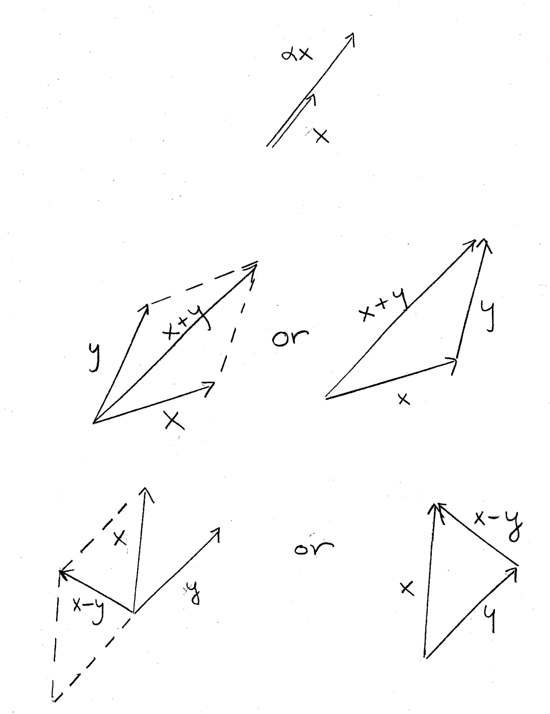


Figure 1: vector operations

increased by $|\alpha|$. The vector $x + y$ can be found by completing a parallelogram with sides x, y and taking the diagonal, or by putting the tail of y on the head of x and drawing the arrow from the tail of x to the head of y. The vector $x - y$ is found by drawing $x + (-1)y$. Alternatively if the tail of $x - y$ put a the head of y then the arrow goes from the head of y to the head of x. See figure 1.

A vector $x = (x_1, \ldots, x_n)$ has a length which is

$$|x| = \text{length of x} = \sqrt{x_1^2 + \cdots + x_n^2} \tag{8}$$

Since $x - y$ goes from the point y to the point x, the length of this vector is the distance between the points:

$$|x - y| = \text{distance between } x \text{ and } y = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} \tag{9}$$

One can also form the dot product of vectors x, y in $\mathbb{R}^n$. The result is a scalar given by

$$x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \tag{10}$$

Then we have

$$x \cdot x = |x|^2 \tag{11}$$

1.2 functions of several variables

We are interested in functions f from $\mathbb{R}^n$ to $\mathbb{R}^m$ (or more generally from a subset $D \subset \mathbb{R}^n$ to $\mathbb{R}^m$ called the *domain* of the function). A function f assigns to each $x \in \mathbb{R}^n$ a point $y \in \mathbb{R}^m$ and we write

$$y = f(x) \tag{12}$$

The set of all such points y is the *range* of the function.

Each component of $y = (y_1, \dots, y_m)$ is real-valued function of $x \in \mathbb{R}^n$ written $y_i = f_i(x)$ and the function can also be written as the collection of n functions

$$y_1 = f_1(x), \cdots, \; y_m = f_m(x) \tag{13}$$

If we also write out the components of $x = (x_1, \dots, x_n)$, then are function can be written as m functions of n variables each:

$$\begin{aligned} y_1 =& f_1(x_1, \dots, x_n) \\ y_2 =& f_2(x_1, \dots, x_n) \\ & \dots \\ y_m =& f_m(x_1, \dots, x_n) \end{aligned} \tag{14}$$

The *graph* of the function is all pairs (x, y) with $y = f(x)$. It is a subset of $\mathbb{R}^{n+m}$.

special cases:

1. $n = 1, m = 2$ (or $m = 3$). The function has the form

   $$y_1 = f_1(x) \qquad y_2 = f_2(x) \tag{15}$$

   In this case the range of the function is a curve in $\mathbb{R}^2$.

2. $n = 2, m = 1$. Then function has the form

   $$y = f(x_1, x_2) \tag{16}$$

   The graph of the function is a surface in $\mathbb{R}^3$.

3. $n = 2, m = 3$. The function has the form

   $$\begin{aligned} y_1 =& f_1(x_1, x_2) \\ y_2 =& f_2(x_1, x_2) \\ y_3 =& f_3(x_1, x_2) \end{aligned} \tag{17}$$

   The range of the function is a surface in $\mathbb{R}^3$.

4. $n = 3, m = 3$. The function has the form

   $$\begin{aligned} y_1 =& f_1(x_1, x_2, x_3) \\ y_2 =& f_2(x_1, x_2, x_3) \\ y_3 =& f_3(x_1, x_2 . x_3) \end{aligned} \tag{18}$$

   The function assigns a vector to each point in space and is called a *vector field*.

1.3 limits

Consider a function $y = f(x)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ (or possibly a subset of $\mathbb{R}^n$). Let $x^0 = (x^0_1, \dots x^0_n)$ be a point in $\mathbb{R}^n$ and let $y^0 = (y^0_1, \dots, y^0_m)$ be a point in $\mathbb{R}^m$. We say that y^0 is the *limit* of f as x goes to x^0, written

$$\lim_{x \to x^0} f(x) = y^0 \tag{19}$$

if for every $\epsilon > 0$ there exists a $\delta > 0$ so that if $|x - x^0| < \delta$ then $|f(x) - y^0| < \epsilon$. The function is *continuous at* x^0 if it is defined at and near x^0 and

$$\lim_{x \to x^0} f(x) = f(x^0) \tag{20}$$

The function is *continuous* if it is continuous at every point in its domain.

If f, g are continuous at x^0 then so are $f \pm g$. If f, g are scalars (i.e. if $m = 1$) then the the product fg is defined and continuous at x^0. If f, g are scalars and $g(x^0) \neq 0$ then f/g is defined near x^0 and and continuous at x^0.

1.4 partial derivatives

At first suppose f is a function from $\mathbb{R}^2$ to $\mathbb{R}$ written

$$z = f(x, y) \tag{21}$$

We define the *partial derivative of f with respect to x at* (x_0, y_0) to be

$$f_x(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \tag{22}$$

if the limit exists. It is the same as the ordinary derivative with y fixed at y_0, i.e

$$\left[\frac{d}{dx} f(x, y_0)\right]_{x = x_0} \tag{23}$$

We also define the *partial derivative of f with respect to y at* (x_0, y_0) to be

$$f_y(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \tag{24}$$

if the limit exists. It is the same as the ordinary derivative with x fixed at x_0, i.e

$$\left[\frac{d}{dy} f(x_0, y)\right]_{y = y_0} \tag{25}$$

We also use the notation

$$
\begin{aligned}
f_x =& \frac{\partial z}{\partial x} \qquad \left(\text{or } \frac{\partial f}{\partial x} \text{ or } z_x \right) \\
f_y =& \frac{\partial z}{\partial y} \qquad \left(\text{or } \frac{\partial f}{\partial y} \text{ or } z_y \right)
\end{aligned} \tag{26}
$$

If we let (x_0, y_0) vary the partial derivatives are also functions and we can take second partial derivatives like

$$
(f_x)_x \equiv f_{xx} \qquad \text{also written} \qquad \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^2} \tag{27}
$$

The four second partial derivatives are

$$
\begin{aligned}
f_{xx} &= \frac{\partial^2 z}{\partial x^2} & f_{xy} &= \frac{\partial^2 z}{\partial y \partial x} \\
f_{yx} &= \frac{\partial^2 z}{\partial x \partial y} & f_{yy} &= \frac{\partial^2 z}{\partial y^2}
\end{aligned} \tag{28}
$$

Usually $f_{xy} = f_{yx}$ for we have

Theorem 1 *If f_x, f_y, f_{xy}, f_{yx} exist and are continuous near (x_0, y_0) (i.e in a little disc centered on (x_0, y_0)) then*

$$
f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0) \tag{29}
$$

Example: Consider $f(x, y) = 3x^2y + 4xy^3$. Then

$$
\begin{aligned}
f_x &= 6xy + 4y^3 & f_y &= 3x^2 + 12xy^2 \\
f_{xy} &= 6x + 12y^2 & f_{yx} &= 6x + 12y^2
\end{aligned} \tag{30}
$$

We also have partial derivatives for a function f from $\mathbb{R}^n$ to $\mathbb{R}$ written $y = f(x_1, \ldots x_n)$. The partial derivative with respect to x_i at $(x_1^0, \ldots x_n^0)$ is

$$
f_{x_i}(x_1^0, \ldots x_n^0) = \lim_{h \to 0} \frac{f(x_1^0, \ldots, x_i^0 + h, \ldots, x_n^0) - f(x_1^0, \ldots x_n^0)}{h} \tag{31}
$$

It is also written

$$
f_{x_i} = \frac{\partial y}{\partial x_i} \tag{32}
$$

1.5 derivatives

A function $z = f(x,y)$ is said to be *differentiable* at (x_0, y_0) if it can be well-approximated by a linear function near that point. This means there should be constants a, b such that

$$f(x,y) = f(x_0,y_0) + a(x-x_0) + b(y-y_0) + \epsilon(x,y) \tag{33}$$

where the error term $\epsilon(x,y)$ is continuous at (x_0, y_0) and $\epsilon(x,y) \to 0$ as $(x,y) \to (x_0,y_0)$ faster than the distance between the points:

$$\lim_{(x,y)\to(x_0,y_0)} \frac{\epsilon(x,y)}{|(x,y)-(x_0,y_0)|} = 0 \tag{34}$$

Note that differentiable implies continuous.

Suppose it is true and take $(x,y) = (x_0+h, y_0)$. Then

$$f(x_0+h,y_0) = f(x_0,y_0) + ah + \epsilon(x_0+h,y_0) \tag{35}$$

and so

$$\frac{f(x_0+h,y_0) - f(x_0,y_0)}{h} = a + \frac{\epsilon(x_0+h,y_0)}{h} \tag{36}$$

Taking the limit $h \to 0$ we see that $f_x(x_0,y_0)$ exists and equals a. Similarly if we take $(x,y) = (x_0, y_0+h)$ we find that $f_y(x_0,y_0)$ exists and equals b.

Thus if f is differentiable then

$$f(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) + \epsilon(x,y) \tag{37}$$

where ϵ satisfies the above condition. The linear approximation is the function

$$z = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \tag{38}$$

The graph of the linear approximation is a plane called the *tangent plane*. See figure 2

It is possible that the partial derivatives $f_x(x_0,y_0)$ and $f_y(x_0,y_0)$ exist but still the function is not differentiable as the following example shows

example. Define a function by

$$f(x,y) = \begin{cases} 1 & x = 0 \text{ or } y = 0 \\ 0 & \text{otherwise} \end{cases} \tag{39}$$

Then

$$f_x(0,0) = 0 \qquad f_y(0,0) = 0 \tag{40}$$

But the function cannot be differentiable at $(0,0)$ since it is not continuous there. It is not continuous since for example

$$\lim_{t\to 0} f(t,t) = 0 \qquad f(0,0) = 1 \tag{41}$$

However the following is true:

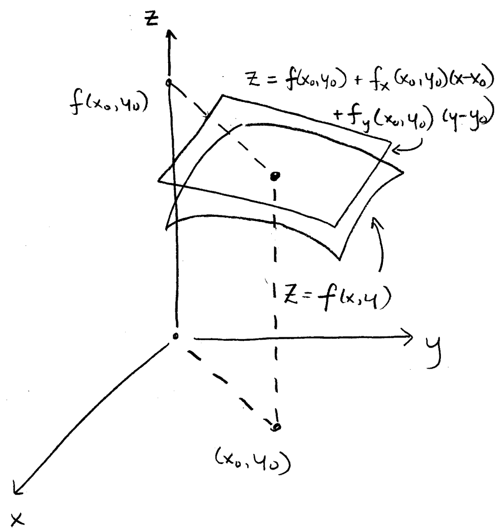


Figure 2: tangent plane

Theorem 2 *If the partial derivatives f_x, f_y exist and are continuous near (x_0, y_0) (i.e in a little disc centered on (x_0, y_0)) then f is differentiable at (x_0, y_0).*

problem: Show the function $f(x, y) = y^3 + 3x^2y^2$ is differentiable at any point and find the linear approximation (tangent plane) at $(1, 1)$.

solution: This has partial derivatives

$$f_x(x,y) = 6xy^2 \qquad f_y(x,y) = 3y^2 + 6x^2y \tag{42}$$

at any point and they are continuous. Thus the function is differentiable. The tangent plane at $(1, 1)$ is

$$\begin{aligned} z =& f(1,1) + f_x(1,1)(x-1) + f_y(1,1)(y-1) \\ =& 4 + 6(x-1) + 9(y-1) \\ =& 6x + 9y - 11 \end{aligned} \tag{43}$$

Next consider a function from $\mathbb{R}^n$ to $\mathbb{R}$ written $y = f(x) = f(x_1, \dots, x_n)$. We say f is differentiable at $x^0 = (x_1^0, \dots x_n^0)$ if there is a vector $a = (a_1, \dots, a_n)$ such that

$$f(x) = f(x^0) + a \cdot (x - x^0) + \epsilon(x) \tag{44}$$

where as before

$$\lim_{x\to x^0}\frac{\epsilon(x)}{|x-x^0|}=0 \tag{45}$$

If it is true then we find as before that the vector must be

$$a=(f_{x_1}(x^0),\dots,f_{x_n}(x^0))\equiv(\nabla f)(x^0) \tag{46}$$

also called the *gradient* of f at x^0. Thus we have

$$f(x)=f(x^0)+(\nabla f)(x^0)\cdot(x-x^0)+\epsilon(x) \tag{47}$$

Finally consider a function $y=f(x)$ from $\mathbb{R}^n$ to $\mathbb{R}^m$. We write the points and the function as column vectors:

$$y=\begin{pmatrix} y_1\\ \vdots\\ y_m\end{pmatrix}=\begin{pmatrix} f_1(x)\\ \vdots\\ f_m(x)\end{pmatrix}\qquad x=\begin{pmatrix} x_1\\ \vdots\\ x_n\end{pmatrix} \tag{48}$$

The function is differentiable at x^0 is there is an $m\times n$ matrix (m rows, n columns) A such that

$$f(x)=f(x^0)+A(x-x^0)+\epsilon(x) \tag{49}$$

where the error $\epsilon(x)\in\mathbb{R}^m$ satisfies

$$\lim_{x\to x^0}\frac{\epsilon(x)}{|x-x^0|}=0 \tag{50}$$

By considering each component separately we find that the i^{th} row of A must be the the gradient of f_i at x^0. Thus

$$A=Df(x^0)\equiv\begin{pmatrix} f_{1,x_1}(x^0) & \cdots & f_{1,x_n}(x^0)\\ \vdots & & \vdots\\ f_{m,x_1}(x^0) & \cdots & f_{m,x_n}(x^0)\end{pmatrix} \tag{51}$$

This matrix $Df(x^0)$ made up of all the partial derivatives of f at x^0 is the *derivative* of f at x^0. Thus we have

$$f(x)=f(x_0)+Df(x^0)(x-x^0)+\epsilon(x) \tag{52}$$

The derivative is also written

$$Df\equiv\begin{pmatrix} \partial y_1/\partial x_1 & \cdots & \partial y_1/\partial x_n\\ \vdots & & \vdots\\ \partial y_m/\partial x_1 & \cdots & \partial y_m/\partial x_n\end{pmatrix} \tag{53}$$

problem: Consider the function from $\mathbb{R}^2$ to $\mathbb{R}^2$ given by

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} = \begin{pmatrix} (x_1 + x_2 + 1)^2 \\ x_1(x_2 + 3) \end{pmatrix} \tag{54}$$

Find the linear approximation at $(0,0)$

solution: The derivative is

$$Df = \begin{pmatrix} \partial y_1/\partial x_1 & \partial y_1/\partial x_2 \\ \partial y_2/\partial x_1 & \partial y_2/\partial x_2 \end{pmatrix} = \begin{pmatrix} 2(x_1 + x_2 + 1) & 2(x_1 + x_2 + 1) \\ x_2 + 3 & x_1 \end{pmatrix} \tag{55}$$

The linear approximation is

$$y = f(0) + Df(0)(x - 0) \tag{56}$$

or

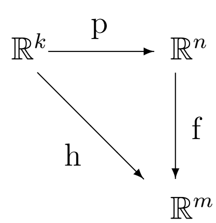
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + 2x_2 + 1 \\ 3x_1 \end{pmatrix} \tag{57}$$

1.6 the chain rule

If $f : \mathbb{R}^n \to \mathbb{R}^m$ (i.e. f is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$) and $p : \mathbb{R}^k \to \mathbb{R}^n$, then there is a composite function $h : \mathbb{R}^k \to \mathbb{R}^m$ defined by

$$h(x) = f(p(x)) \tag{58}$$

and we also write $h = f \circ p$. We can represent the situation by the diagram:



The chain gives a formula for the derivatives of the composite function h in terms of the derivatives of f and p. We start with some special cases.

k=1,n=2,m=1 In this case the functions have the form

$$\begin{aligned} u &= f(x, y) \\ x &= p_1(t) \quad y = p_2(t) \end{aligned} \tag{59}$$

and the composite is

$$u = h(t) = f(p_1(t), p_2(t)) \tag{60}$$

Theorem 3 *(chain rule) If f and p are differentiable, then so is the composite $h = f \circ p$ and the derivative is*

$$h'(t) = f_x(p_1(t), p_2(t))p_1'(t) + f_y(p_1(t), p_2(t))p_2'(t) \tag{61}$$

The idea of the proof is as follows. Since f is differentiable at $(p_1(t), p_2(t))$

$$\begin{aligned} h(t+\Delta t) - h(t) =& f((p_1(t+\Delta t), p_2(t+\Delta t)) - f(p_1(t), p_2(t)) \\ =& f_x(p_1(t), p_2(t))\Big(p_1(t+\Delta t) - p_1(t)\Big) + f_y(p_1(t), p_2(t))\Big(p_2(t+\Delta t) - p_2(t)\Big) \\ &+ \epsilon(p_1(t+\Delta t), p_2(t+\Delta t)) \end{aligned} \tag{62}$$

Now divide by Δt and let $\Delta t \to 0$. One has to show that the error term goes to zero and the result follows.

This form of the chain rule can also be written in the concise form

$$\frac{du}{dt} = \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} \tag{63}$$

But one must keep in mind that $\partial u/\partial x$ and $\partial u/\partial y$ are to be evaluated at $x = p_1(t), y = p_2(t)$. Also note that on the left u stands for the function $u = h(t)$ while on the right it stands for the function $u = f(x, y)$.

example: Suppose $u = x^2 + y^2$ and $x = \cos t$, $y = \sin t$. Find du/dt. We have

$$\begin{aligned} \frac{du}{dt} =& \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} \\ =& 2x(-\sin t) + 2y\cos t \\ =& 2\cos t(-\sin t) + 2\sin t\cos t \\ =& 0 \end{aligned} \tag{64}$$

(This is to be expected since the composite is $u = 1$).

example: Suppose $u = \sqrt{2 + x^2 + y^2}$ and $x = e^t$, $y = e^{2t}$. Find du/dt at $t = 0$. At $t = 0$ we have $x = 1, y = 1$ and so

$$\begin{aligned} \frac{du}{dt} =& \frac{\partial u}{\partial x}\frac{dx}{dt} + \frac{\partial u}{\partial y}\frac{dy}{dt} \\ =& \frac{x}{\sqrt{2 + x^2 + y^2}}\ e^t + \frac{y}{\sqrt{2 + x^2 + y^2}}\ 2e^{2t} \\ =& \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 2 = \frac{3}{2} \end{aligned} \tag{65}$$

k=1,n,m=1 In this case the functions have the form

$$\begin{aligned} &u = f(x_1, \dots x_n) \\ &x_1 = p_1(t), \dots, x_n = p_n(t) \end{aligned} \tag{66}$$

and the composite is

$$u = h(t) = f(p_1(t), \cdots, p_n(t)) \tag{67}$$

The chain rule says

$$h'(t) = f_{x_1}(p_1(t), \cdots, p_n(t))p_1'(t) + \cdots + f_{x_n}(p_1(t), \cdots, p_n(t))p_n'(t) \tag{68}$$

It can also be written

$$\frac{du}{dt} = \frac{\partial u}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial u}{\partial x_n}\frac{dx_n}{dt} \tag{69}$$

k=2,n=2,m=2 In this case the functions have the form

$$\begin{aligned} &u = f_1(x, y) \qquad v = f_2(x, y) \\ &x = p_1(s, t) \qquad y = p_2(s, t) \end{aligned} \tag{70}$$

and the composite is

$$u = h_1(s, t) = f_1(p_1(s, t), p_2(s, t)) \qquad v = h_2(s, t) = f_2(p_1(s, t), p_2(s, t)) \tag{71}$$

Taking partial derivatives with respect to s, t we can use the formula from the case k=1,n=2,m=1 to obtain

$$\begin{aligned} \frac{\partial u}{\partial s} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial u}{\partial t} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} \\ \frac{\partial v}{\partial s} &= \frac{\partial v}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial v}{\partial t} &= \frac{\partial v}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial t} \end{aligned} \tag{72}$$

Here derivatives with respect to x, y are to be evaluated at $x = p_1(s, t), y = p_2(s, t)$. This can also be written as a matrix product:

$$\begin{pmatrix} \partial u/\partial s & \partial u/\partial t \\ \partial v/\partial s & \partial v/\partial t \end{pmatrix} = \begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix} \begin{pmatrix} \partial x/\partial s & \partial x/\partial t \\ \partial y/\partial s & \partial y/\partial t \end{pmatrix} \tag{73}$$

These matrices are just the derivatives of the various functions and the last equation can be written

$$(Dh)(s,t) = (Df)(p(s,t))(Dp)(s,t) \tag{74}$$

The last form holds in the general case. If $p : \mathbb{R}^k \to \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ are differentiable then the composite function $h = f \circ p : \mathbb{R}^k \to \mathbb{R}^m$ is differentiable and

$$\underbrace{(Dh)(x)}_{m\times k \text{ matrix}} = \underbrace{(Df)(p(x))}_{m\times n \text{ matrix}} \underbrace{(Dp)(x)}_{n\times k \text{ matrix}} \tag{75}$$

In other words the derivative of a composite is the matrix product of the derivatives of the two elements. All forms of the chain rule are special cases of this equation.

1.7 implicit function theorem -I

Suppose we have an equation of the form

$$f(x,y,z) = 0 \tag{76}$$

Can we solve it for z?. More precisely, is it the case that for each x, y in some domain there is a unique z so that $f(x,y,z) = 0$? If so one can define an *implicit function* by

$$z = \phi(x,y)$$

Geometrically points satisfying $f(x,y,z) = 0$ are a surface and we are asking whether the surface is the graph of a function.

Suppose there is an implicit function. Then

$$f(x,y,\phi(x,y)) = 0 \tag{77}$$

Assuming everything is differentiable we can take partial derivatives of this equation. By the chain rule we have

$$\begin{aligned} 0 = \frac{\partial}{\partial x}[f(x,y,\phi(x,y))] &= f_x(x,y,\phi(x,y)) + f_z(x,y,\phi(x,y))\phi_x(x,y) \\ 0 = \frac{\partial}{\partial y}[f(x,y,\phi(x,y))] &= f_y(x,y,\phi(x,y)) + f_z(x,y,\phi(x,y))\phi_y(x,y) \end{aligned} \tag{78}$$

Solve this for $\phi_x(x,y)$ and $\phi_y(x,y)$ and get

$$\begin{aligned} \phi_x(x,y) &= -\frac{f_x(x,y,\phi(x,y))}{f_z(x,y,\phi(x,y))} \\ \phi_y(x,y) &= -\frac{f_y(x,y,\phi(x,y))}{f_z(x,y,\phi(x,y))} \end{aligned} \tag{79}$$

This can also be written in the form

$$
\begin{aligned}
\frac{\partial z}{\partial x} &= -\frac{\partial f/\partial x}{\partial f/\partial z} \\
\frac{\partial z}{\partial y} &= -\frac{\partial f/\partial y}{\partial f/\partial z}
\end{aligned}
\tag{80}
$$

keeping in mind that the right side is evaluated at $z = \phi(x, y)$.

For this to work we need that $\partial f/\partial z \neq 0$. If this holds and if we restrict attention to a small region then there always is an implicit function. This is the content of the following

Theorem 4 *(implicit function theorem) Let $f(x, y, z)$ have continuous partial derivatives near some point (x_0, y_0, z_0). If*

$$
\begin{aligned}
f(x_0, y_0, z_0) &= 0 \\
f_z(x_0, y_0, z_0) &\neq 0
\end{aligned}
\tag{81}
$$

Then for every (x, y) near (x_0, y_0) there is a unique z near z_0 such that $f(x, y, z) = 0$. The implicit function $z = \phi(x, y)$ defined near (x_0, y_0) satisfies $z_0 = \phi(x_0, y_0)$ and has continuous partial derivatives which satisfy the equations (79).

The theorem does not tell you what ϕ is or how to find it, only that it exists. However if we take the equations (79) at the special point (x_0, y_0) we find

$$
\begin{aligned}
\phi_x(x_0, y_0) &= -\frac{f_x(x_0, y_0, z_0)}{f_z(x_0, y_0, z_0)} \\
\phi_y(x_0, y_0) &= -\frac{f_y(x_0, y_0, z_0)}{f_z(x_0, y_0, z_0)}
\end{aligned}
\tag{82}
$$

and these quantities can be computed.

example: Suppose the equation is

$$
f(x, y, z) = x^2 + y^2 + z^2 - 1 = 0 \tag{83}
$$

which describes the surface of a sphere. Is there an implicit function $z = \phi(x, y)$ near a particular point (x_0, y_0, z_0) on the sphere?

We have the derivatives

$$
f_x(x, y, z) = 2x \qquad f_y(x, y, z) = 2y \qquad f_z(x, y, z) = 2z \tag{84}
$$

Then $f_z(x_0, y_0, z_0) = 2z_0$ is not zero if $z_0 \neq 0$. So by the theorem there is an implicit function $z = \phi(x, y)$ near a point with $z_0 \neq 0$ and

$$\begin{aligned} \frac{\partial z}{\partial x} &= -\frac{f_x}{f_z} = -\frac{x}{z} \\ \frac{\partial z}{\partial y} &= -\frac{f_y}{f_z} = -\frac{y}{z} \end{aligned} \tag{85}$$

Check: this example is special in that we can find the implicit function exactly and so check these results. The implicit function is

$$z = \pm\sqrt{1 - x^2 - y^2} \tag{86}$$

with the plus sign if $z_0 > 0$ and the minus sign if $z_0 < 0$. In either case we have the expected result

$$\begin{aligned} \frac{\partial z}{\partial x} &= \pm\frac{1}{2}(1 - x^2 - y^2)^{-1/2}(-2x) = -\frac{x}{z} \\ \frac{\partial z}{\partial y} &= \pm\frac{1}{2}(1 - x^2 - y^2)^{-1/2}(-2y) = -\frac{y}{z} \end{aligned} \tag{87}$$

If $z_0 = 0$ there is no implicit function near the point as figure 3 illustrates.

problem: Consider the equation

$$f(x, y, z) = xe^z + yz + 1 = 0 \tag{88}$$

Note that $(x, y, z) = (0, 1, -1)$ is one solution. Show there is an implicit function $z = \phi(x, y)$ near this point. What are the derivatives $\partial z/\partial x, \partial z/\partial y$ at $(x, y) = (0, 1)$?

solution We have the derivatives

$$f_x = e^z \qquad f_y = z \qquad f_z = xe^z + y \tag{89}$$

Then $f_z(0, 1, -1) = 1$ is not zero so by the theorem there is an implicit function. The derivatives at (0,1) are

$$\begin{aligned} \frac{\partial z}{\partial x} &= -\frac{f_x}{f_z} = -\frac{e^z}{xe^z + y} = -e^{-1} \qquad \text{at (0,1,-1)} \\ \frac{\partial z}{\partial y} &= -\frac{f_y}{f_z} = -\frac{z}{xe^z + y} = 1 \qquad \text{at (0,1,-1)} \end{aligned} \tag{90}$$

alternate solution: Once the existence of the implicit function is established we can argue as follows. Take partial derivatives of $xe^z + yz + 1 = 0$ assuming $z = \phi(x, y)$ and obtain

$$\begin{aligned} \frac{\partial}{\partial x} &: \qquad e^z + xe^z\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial x} = 0 \\ \frac{\partial}{\partial y} &: \qquad xe^z\frac{\partial z}{\partial y} + z + y\frac{\partial z}{\partial y} = 0 \end{aligned} \tag{91}$$

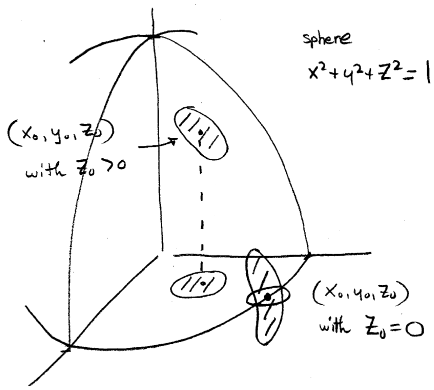


Figure 3: There is an implicit function near points with $z_0 \neq 0$, but not near points with $z_0 = 0$.

Now put in the point $(x, y, z) = (0, 1, -1)$ and get

$$\begin{aligned} e^{-1} + \frac{\partial z}{\partial x} &= 0 \\ -1 + \frac{\partial z}{\partial y} &= 0 \end{aligned} \tag{92}$$

Solving for the derivatives gives the same result.

Some remarks:

1. Why is the implicit function theorem true? Instead of solving $f(x, y, z) = 0$ for z near (x_0, y_0, z_0) one could make a linear approximation and try to solve that. Taking into account that $f(x_0, y_0, z_0) = 0$ the linear approximation is

$$f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) = 0 \tag{93}$$

If $f_z(x_0, y_0, z_0) \neq 0$ this can be solved by

$$z = z_0 - \frac{f_x(x_0, y_0, z_0)}{f_z(x_0, y_0, z_0)}(x - x_0) - \frac{f_y(x_0, y_0, z_0)}{f_z(x_0, y_0, z_0)}(y - y_0) \tag{94}$$

and this has the expected derivatives. To prove the theorem one has to argue that since the actual function is near the linear approximation there is an implicit function in that case as well.

2. Here are some variations of the implicit function theorem

   (a) If $f(x,y)=0$ one can solve for $y=\phi(x)$ near any point where $f_y\neq 0$ and $dy/dx=-f_x/f_y$.

   (b) If $f(x,y,z)=0$ one can solve for $x=\phi(y,z)$ near any point where $f_x\neq 0$ and $\partial x/\partial y=-f_y/f_x$ and $\partial x/\partial z=-f_z/f_x$

   (c) If $f(x,y,z,w)=0$ one can solve for $w=\phi(x,y,z)$ near any point where $f_w\neq 0$ and $\partial w/\partial x=-f_x/f_w$, etc.

3. One can also find higher derivatives of the implicit function. If $f(x,y,z)=0$ defines $z=\phi(x,y)$ and

$$\phi_x(x,y)=-\frac{f_x(x,y,\phi(x,y))}{f_z(x,y,\phi(x,y))} \tag{95}$$

   then one can find ϕ_{xx},ϕ_{xy} by further differentiation.

example: $f(x,y,z)=x^2+y^2+z^2-1=0$ defines $z=\phi(x,y)$ which satisfies

$$\frac{\partial z}{\partial x}=-\frac{x}{z} \tag{96}$$

Keeping in mind that z is a function of x,y further differentiation yields

$$\frac{\partial^2 z}{\partial x^2}=-\left(\frac{z-x\ \partial z/\partial x}{z^2}\right)=-\left(\frac{z-x(-x/z)}{z^2}\right)=\frac{-z^2-x^2}{z^3} \tag{97}$$

Since $x^2+y^2+z^2=1$ this can also be written as $(y^2-1)/z^3$. Similarly

$$\frac{\partial^2 z}{\partial x\partial y}=\frac{-xy}{z^3} \tag{98}$$

1.8 implicit function theorem -II

We give another version of the implicit function theorem. Suppose we have a pair of equations of the form

$$\begin{aligned}F(x,y,u,v)=&0\\G(x,y,u,v)=&0\end{aligned} \tag{99}$$

Can we solve them for u, v as functions of x, y? More precisely is it the case that for each x, y in some domain there is a unique u, v so that the equations are satisfied. If it is so we can define implicit functions

$$\begin{aligned} u =& f(x,y) \\ v =& g(x,y) \end{aligned} \tag{100}$$

If the implicit functions exist we have

$$\begin{aligned} F(x,y,f(x,y),g(x,y)) =& 0 \\ G(x,y,f(x,y),g(x,y)) =& 0 \end{aligned} \tag{101}$$

Take partial derivatives with respect to x and y and get

$$\begin{aligned} F_x + F_u f_x + F_v g_x =& 0 \\ G_x + G_u f_x + G_v g_x =& 0 \\ F_y + F_u f_y + F_v g_y =& 0 \\ G_y + G_u f_y + G_v g_y =& 0 \end{aligned} \tag{102}$$

These equations can be written as the matrix equations

$$\begin{aligned} \begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \begin{pmatrix} f_x \\ g_x \end{pmatrix} =& \begin{pmatrix} -F_x \\ -G_x \end{pmatrix} \\ \begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \begin{pmatrix} f_y \\ g_y \end{pmatrix} =& \begin{pmatrix} -F_y \\ -G_y \end{pmatrix} \end{aligned} \tag{103}$$

Note that the matrix on the left is the derivative of the function

$$(u,v) \to (F(x,y,u,v), G(x,y,u,v)) \tag{104}$$

for fixed (x, y). If the determinant of this matrix is not zero we can solve for the partial derivatives f_x, g_x, f_y, g_y. One finds for example

$$f_x = \frac{\det \begin{pmatrix} -F_x & F_v \\ -G_x & G_v \end{pmatrix}}{\det \begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix}} \tag{105}$$

The determinant of the matrix is called the *Jacobian determinant* and is given a special symbol

$$\frac{\partial(F,G)}{\partial(u,v)} \equiv \det \begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} \equiv F_u G_v - F_v G_u \tag{106}$$

With this notation the equations for the four partial derivatives can be written

$$\begin{aligned}
\frac{\partial u}{\partial x} =& -\frac{\partial(F,G)}{\partial(x,v)} \Big/ \frac{\partial(F,G)}{\partial(u,v)} \\
\frac{\partial v}{\partial x} =& -\frac{\partial(F,G)}{\partial(u,x)} \Big/ \frac{\partial(F,G)}{\partial(u,v)} \\
\frac{\partial u}{\partial y} =& -\frac{\partial(F,G)}{\partial(y,v)} \Big/ \frac{\partial(F,G)}{\partial(u,v)} \\
\frac{\partial v}{\partial y} =& -\frac{\partial(F,G)}{\partial(u,y)} \Big/ \frac{\partial(F,G)}{\partial(u,v)}
\end{aligned} \tag{107}$$

where $u = f(x,y), v = g(x,y)$ on the right. This holds provided $\partial(F,G)/\partial(u,v) \neq 0$ and this is the key condition in the following theorem which guarantees the existence of the implicit functions.

Theorem 5 *(implicit function theorem) Let F, G have continuous partial derivatives near some point (x_0, y_0, u_0, v_0). If*

$$\begin{aligned}
F(x_0, y_0, u_0, v_0) =& 0 \\
G(x_0, y_0, u_0, v_0) =& 0
\end{aligned} \tag{108}$$

and

$$\left[\frac{\partial(F,G)}{\partial(u,v)}\right](x_0, y_0, u_0, v_0) \neq 0 \tag{109}$$

Then for every (x,y) near (x_0, y_0) there is a unique (u,v) near (u_0, v_0) such that $F(x,y,u,v) = 0$ and $G(x,y,u,v) = 0$. The implicit functions $u = f(x,y), v = g(x,y)$ defined near (x_0, y_0) satisfy $u_0 = f(x_0, y_0), v_0 = g(x_0, y_0)$ and have continuous partial derivative which satisfy the equations (107).

problem: Consider the equations

$$\begin{aligned}
F(x,y,u,v) = x^2 - y^2 + 2uv - 2 =& 0 \\
G(x,y,u,v) = 3x + 2xy + u^2 - v^2 =& 0
\end{aligned} \tag{110}$$

Note that $(x,y,u,v) = (0,0,1,1)$ is a solution. Are there implicit functions $u = f(x,y), v = g(x,y)$ near $(0,0)$? What are the derivatives $\partial u/\partial x, \partial v/\partial x$ at $(x,y) = (0,0)$?

solution: First compute

$$\frac{\partial(F,G)}{\partial(u,v)} = \det\begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} = \det\begin{pmatrix} 2v & 2u \\ 2u & -2v \end{pmatrix} = -4(u^2+v^2) = -8 \tag{111}$$

Since this is not zero the implicit functions exist by the theorem. We also compute

$$\frac{\partial(F,G)}{\partial(x,v)} = \det\begin{pmatrix} F_x & F_v \\ G_x & G_v \end{pmatrix} = \det\begin{pmatrix} 2x & 2u \\ 3+2y & -2v \end{pmatrix} = -4xv - 2u(3+2y) = -6 \tag{112}$$

and

$$\frac{\partial(F,G)}{\partial(u,x)} = \det\begin{pmatrix} F_u & F_x \\ G_u & G_x \end{pmatrix} = \det\begin{pmatrix} 2v & 2x \\ 2u & 3+2y \end{pmatrix} = 2v(3+2y) - 4ux = 6 \tag{113}$$

Then

$$\frac{\partial u}{\partial x} = -\frac{-6}{-8} = -\frac{3}{4} \qquad \frac{\partial v}{\partial x} = -\frac{6}{-8} = \frac{3}{4} \tag{114}$$

alternate solution: Assuming the implicit functions exist differentiate the equations $F = 0, G = 0$ with respect to x assuming $u = f(x,y), v = g(x,y)$. This gives

$$\begin{aligned} 2x + 2v\frac{\partial u}{\partial x} + 2u\frac{\partial v}{\partial x} &= 0 \\ 3 + 2y + 2u\frac{\partial u}{\partial x} - 2v\frac{\partial v}{\partial x} &= 0 \end{aligned} \tag{115}$$

Now put in the point $(0,0,1,1)$ and get

$$\begin{aligned} 2\frac{\partial u}{\partial x} + 2\frac{\partial v}{\partial x} &= 0 \\ 3 + 2\frac{\partial u}{\partial x} - 2\frac{\partial v}{\partial x} &= 0 \end{aligned} \tag{116}$$

This again has the solution $\partial u/\partial x = -3/4$, $\partial v/\partial x = 3/4$.

1.9 inverse functions

Suppose we have a function f from $U \subset \mathbb{R}^2$ to $V \subset \mathbb{R}^2$ which we write in the form

$$\begin{aligned} x &= f_1(u,v) \\ y &= f_2(u,v) \end{aligned} \tag{117}$$

Suppose further for every (x,y) in V there is a unique (u,v) in U such that $x = f_1(u,v), y = f_2(u,v)$. (0ne says the function is *one-to-one* and *onto*.) Then there is a an inverse function g from V to U defined by

$$\begin{aligned} u &= g_1(x,y) \\ v &= g_2(x.y) \end{aligned} \tag{118}$$

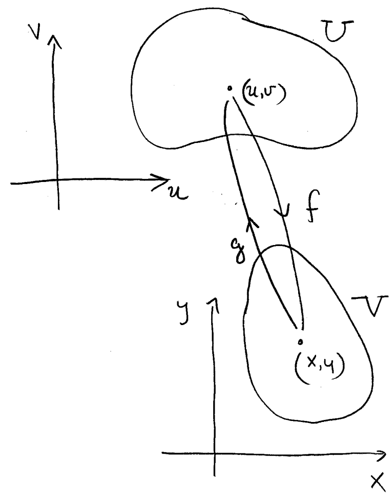


Figure 4: inverse function

So g sends every point back where it came from. See figure 4. We have

$$\begin{aligned}(g \circ f)(u, v) =&(u, v)\\(f \circ g)(x, y) =&(x, y)\end{aligned} \tag{119}$$

We write $g = f^{-1}$.

example: Suppose the function is

$$\begin{aligned}x =&au + bv\\y =&cu + dv\end{aligned} \tag{120}$$

defined on all of $\mathbb{R}^2$. This can also be written in matrix form

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{121}$$

This is invertible if we can solve the equation for (u, v) which is possible if and only if

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \neq 0 \tag{122}$$

The inverse function is given by the inverse matrix

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \tag{123}$$

where

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{124}$$

example: Consider the function

$$\begin{aligned} x =& r\cos\theta \\ y =& r\sin\theta \end{aligned} \tag{125}$$

from

$$U = \{(r,\theta) : r > 0, 0 \leq \theta < 2\pi\} \tag{126}$$

to

$$V = \{(x,y) : (x,y) \neq 0\} \tag{127}$$

This function sends (r,θ) to the point with polar coordinates (r,θ). See figure 5. The function is invertible since every point (x,y) in V has unique polar coordinates (r,θ) in U. (It would not be invertible if we took $U = \mathbb{R}^2$ since (r,θ) and $(r,\theta+2\pi)$ are sent to the same point). For (x,y) in the first quadrant the inverse is

$$\begin{aligned} r =& \sqrt{x^2+y^2} \\ \theta =& \tan^{-1}\left(\frac{y}{x}\right) \end{aligned} \tag{128}$$

1.10 inverse function theorem

Continuing the discussion of the previous section suppose that f has an inverse function g and that both are differentiable. Then differentiating $(f \circ g)(x,y) = (x,y)$ we find by the chain rule

$$(Df)(g(x,y))\ (Dg)(x,y) = I \qquad\qquad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{129}$$

and so the derivative of the inverse function is the matrix inverse

$$(Dg)(x,y) = [(Df)(g(x,y))]^{-1} \tag{130}$$

It is not always easy to tell whether an inverse function exists. The following theorem can be helpful.

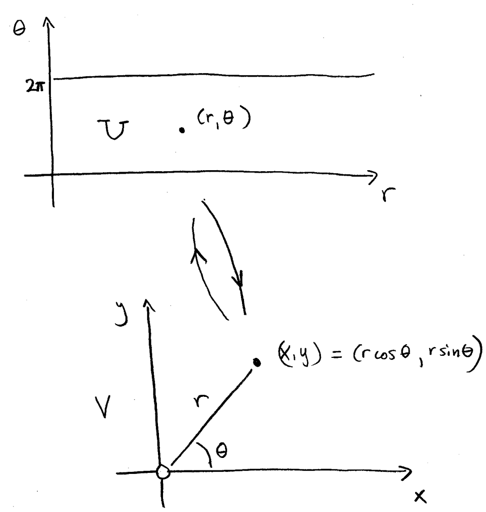


Figure 5: inverse function for polar coordinates

Theorem 6 *(inverse function theorem) Let $x = f_1(u, v), y = f_2(u, v)$ have continuous partial derivatives and suppose*

$$\begin{aligned} x_0 =& f_1(u_0, v_0) \\ y_0 =& f_2(u_0, v_0) \end{aligned} \tag{131}$$

and

$$\left(\frac{\partial(x, y)}{\partial(u, v)}\right)(u_0, v_0) \neq 0 \tag{132}$$

Then there is an inverse function $u = g_1(x, y), v = g_2(x, y)$ defined near (x_0, y_0) which satisfies

$$\begin{aligned} u_0 =& g_1(x_0, y_0) \\ v_0 =& g_2(x_0, y_0) \end{aligned} \tag{133}$$

and has a continuous derivative which satisfies (130). In particular $Dg(x_0, y_0) = [Df(u_0, v_0)]^{-1}$ or

$$\begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix}_{(x_0,y_0)} = \begin{pmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{pmatrix}^{-1}_{(u_0,v_0)} \tag{134}$$

Proof. The inverse exists if for each (x,y) near (x_0,y_0) there is a unique (u,v) near (u_0,v_0) such that

$$\begin{aligned} F(x,y,u,v) \equiv f_1(u,v) - x =& 0 \\ G(x,y,u,v) \equiv f_2(u,v) - y =& 0 \end{aligned} \tag{135}$$

This follows by the implicit functions theorem since (x_0,y_0,u_0,v_0) is one solution and at this point

$$\frac{\partial(F,G)}{\partial(u,v)} = \det\begin{pmatrix} F_u & F_v \\ G_u & G_v \end{pmatrix} = \det\begin{pmatrix} f_{1,u} & f_{1,v} \\ f_{2,u} & f_{2,v} \end{pmatrix} = \frac{\partial(x,y)}{\partial(u,v)} \neq 0$$

The differentiability of the inverse also follows from the implicit function theorem.

problem: Consider the function

$$\begin{aligned} x =& u + v^2 \\ y =& u^2 + v \end{aligned} \tag{137}$$

which sends $(u,v) = (1,2)$ to $(x,y) = (5,3)$. Show that there is an inverse function defined near $(5,3)$ and find the partial derivatives of the inverse function at $(5,3)$.

solution: We have

$$\begin{pmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{pmatrix} = \begin{pmatrix} 1 & 2v \\ 2u & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 2 & 1 \end{pmatrix} \text{ at } (1,2) \tag{138}$$

Therefore

$$\frac{\partial(x,y)}{\partial(u,v)} = \det\begin{pmatrix} 1 & 4 \\ 2 & 1 \end{pmatrix} = -7 \quad \text{at } (1,2) \tag{139}$$

This is not zero so the inverse exists by the theorem and sends $(5,3)$ to $(1,2)$. We have for the derivatives

$$\begin{aligned} \begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix}_{(5,3)} =& \begin{pmatrix} \partial x/\partial u & \partial x/\partial v \\ \partial y/\partial u & \partial y/\partial v \end{pmatrix}^{-1}_{(1,2)} \\ = \begin{pmatrix} 1 & 4 \\ 2 & 1 \end{pmatrix}^{-1} =& \begin{pmatrix} -1/7 & 4/7 \\ 2/7 & -1/7 \end{pmatrix} \end{aligned} \tag{140}$$

alternate solution: Differentiate $x = u + v^2, y = u^2 + v$ assuming u,v are functions of x,y, then put in the point and solve for the derivatives.

1.11 maxima and minima

Let $f(x) = f(x_1, \ldots, x_n)$ be a function from $\mathbb{R}^n$ to $\mathbb{R}$ and let $x^0 = (x_1^0, \ldots x_n^0)$ be a point in $\mathbb{R}^n$. We say that f has a *local maximum* at x^0 if $f(x) \leq f(x^0)$ for all x near x^0. We say that f has a *local minimum* at x^0 if $f(x) \geq f(x^0)$ for all x near x^0.

Theorem 7 *If f is differentiable at x^0 and f has a local maximum or minimum at x^0 then all partial derivatives vanish at the point, i.e.*

$$f_{x_1}(x^0) = \cdots = f_{x_n}(x^0) = 0 \tag{141}$$

Proof. For any $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ consider the function $F(t) = f(x^0 + th)$ of the single variable t. If f has a local maximum or minimum at x^0 then F has a local maximum or minimum at $t = 0$. By the chain rule $F(t)$ is differentiable and and it is a result of elementary calculus that $F'(0) = 0$. But the chain rule says

$$F'(t) = \sum_{i=1}^{n} f_{x_i}(x^0 + th) \frac{d(x_i^0 + th_i)}{dt} = \sum_{i=1}^{n} f_{x_i}(x^0 + th) h_i \tag{142}$$

Thus

$$0 = F'(0) = \sum_{i=1}^{n} f_{x_i}(x^0) h_i \tag{143}$$

Since this is true for any h it must be that $f_{x_i}(x^0) = 0$.

A point x^0 with $f_{x_i}(x^0) = 0$ is called a *critical point* for f. We have seen that if f is has a local maximum or minimum at x^0 then it is a critical point. We are interested whether the converse is true. If x^0 is a critical point is it a local maximum or minimum for f? Which is it?

To answer this question consider again $F(t) = f(x^0 + th)$ and suppose f is many times differentiable. By Taylor's theorem for one variable we have

$$F(1) = F(0) + F'(0) + \frac{1}{2} F''(0) + \frac{1}{6} F'''(s) \tag{144}$$

for some s between 0 and 1. But $F'(t)$ is computed above, and similarly we have

$$\begin{aligned} F''(t) &= \sum_{i=1}^{n} \sum_{j=1}^{n} f_{x_i x_j}(x^0 + th) h_i h_j \\ F'''(t) &= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} f_{x_i x_j x_k}(x^0 + th) h_i h_j h_k \end{aligned} \tag{145}$$

Then the Taylor's expansion becomes

$$f(x^0+h) = f(x^0) + \sum_{i=1}^{n} f_{x_i}(x^0)h_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} f_{x_ix_j}(x^0)h_ih_j + R(h) \tag{146}$$

This is an example of a multivariable Taylor's theorem with remainder. The remainder $R(h) = F'''(s)/6$ is small if h is small and one can show that there is a constant C such that for h small $|R(h)| \leq C|h|^3$.

Now suppose x^0 is a critical point so the first derivatives vanish. Also define a matrix of second derivatives by $A_{ij} = f_{x_ix_j}(x^0)$ or

$$A = \begin{pmatrix} f_{x_1,x_1}(x^0) & \cdots & f_{x_1,x_n}(x^0) \\ \vdots & & \vdots \\ f_{x_n,x_1}(x^0) & \cdots & f_{x_n,x_n}(x^0) \end{pmatrix} \tag{147}$$

called the *Hessian* of f at x^0. Then we can write our expansion as

$$f(x^0+h) = f(x^0) + \frac{1}{2}h \cdot Ah + R(h) \tag{148}$$

We are interested in whether $f(x^0+h)$ is greater than or less than $f(x^0)$ for $|h|$ small. Then idea is that since $R(h)$ is much smaller than $\frac{1}{2}h \cdot Ah$ it is the latter term which determines the answer.

A $n \times n$ matrix A is called *positive definite* if there is a constant M so $h \cdot Ah > M|h|^2$ for all $h \neq 0$. It is called *negative definite* if $h \cdot Ah < -M|h|^2$ for all $h \neq 0$.

Theorem 8 *Let x^0 be a critical point for f and suppose the Hessian A has* $\det A \neq 0$.

1. *If A is positive definite then f has a local minimum at x^0.*
2. *If A is negative definite then f has a local maximum at x^0.*
3. *Otherwise f has a saddle point at x^0, i.e f increases in some directions and decreases in other directions as you move away from x^0.*

Proof. We prove the first statement. We have $h \cdot Ah > M|h|^2$. If also $|h| \leq M/4C$. then

$$|R(h)| \leq C|h|^3 \leq \frac{1}{4}M|h|^2 \tag{149}$$

Therefore

$$f(x^0+h) \geq f(x^0) + \frac{1}{2}M|h|^2 - \frac{1}{4}M|h|^2 \tag{150}$$

or

$$f(x^0+h) \geq f(x^0) + \frac{1}{4}M|h|^2 \tag{151}$$

Thus $f(x^0+h) > f(x^0)$ for $0 < |h| \leq M/4C$ which means we have a strict local minimum at x^0.

Thus our problem is to decide whether A is positive or negative definite. First note the the Hessian is a symmetric matrix, that is if A_{ij} is the entry in the i^{th} row and j^{th} column, then $A_{ij} = f_{x_i,x_j}(x^0) = f_{x_j,x_i}(x^0) = A_{ji}$. For a symmetric matrix A one can show that A is positive definite if and only if all eigenvalues are postive and A is negative definite if and only if all eigenvalues are negative. Recall that λ is an eigenvalue if there is a vector $v \neq 0$ such that $Av = \lambda v$. One can find the eigenvalues by solving the equation

$$\det(A - \lambda I) = 0 \tag{152}$$

where I is the identity matrix.

example: Consider the function

$$f(x,y) = \exp\left(-\frac{x^2}{2} - \frac{y^2}{2} + x + y\right) \tag{153}$$

The critical points are solutions of

$$\begin{aligned} f_x =& (-x+1)\exp\left(-\frac{x^2}{2} - \frac{y^2}{2} + x + y\right) = 0 \\ f_y =& (-y+1)\exp\left(-\frac{x^2}{2} - \frac{y^2}{2} + x + y\right) = 0 \end{aligned} \tag{154}$$

Thus the only critical point is $(x,y) = (1,1)$. The second derivatives at this point are

$$\begin{aligned} f_{xx} =& (x^2 - 2x)\exp\left(-\frac{x^2}{2} - \frac{y^2}{2} + x + y\right) = -e \\ f_{xy} =& (-x+1)(-y+1)\exp\left(-\frac{x^2}{2} - \frac{y^2}{2} + x + y\right) = 0 \\ f_{yy} =& (y^2 - 2y)\exp\left(-\frac{x^2}{2} - \frac{y^2}{2} + x + y\right) = -e \end{aligned} \tag{155}$$

Thus the Hessian is

$$A = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} -e & 0 \\ 0 & -e \end{pmatrix} \tag{156}$$

It has eigenvalues $-e, -e$ which are negative. Hence A is negative definite and the function has a local maximum at $(1,1)$.

example: Consider the function

$$f(x,y) = x \sin y \tag{157}$$

The critical points are solutions of

$$\begin{aligned} f_x =& \sin y = 0 \\ f_y =& x \cos y = 0 \end{aligned} \tag{158}$$

These are points with $x = 0$ and $y = 0, \pm\pi, \pm 2\pi, \cdots$. The Hessian is

$$A = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 0 & \cos y \\ \cos y & -x \sin y \end{pmatrix} \tag{159}$$

At points $(0, \pm\pi), (0, \pm 3\pi), \ldots$ this is

$$A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \tag{160}$$

At points $(0, 0), (0, \pm 2\pi), \ldots$ this is

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{161}$$

In either case the eigenvalues are solutions of

$$\det(A - \lambda I) = \det \begin{pmatrix} -\lambda & \pm 1 \\ \pm 1 & -\lambda \end{pmatrix} = \lambda^2 - 1 = 0 \tag{162}$$

Thus the eigenvalues are $\lambda = \pm 1$. Since they have both signs every critical point is a saddle point.

1.12 differentiation under the integral sign

Theorem 9 *If $f(t, x), (\partial f / \partial t)(t, x)$ exist and are continuous*

$$\frac{d}{dt} \left[\int_a^b f(t, x) dx \right] = \int_a^b \frac{\partial f}{\partial t}(t, x)\ dx \tag{163}$$

Proof. Form the difference quotient

$$\frac{\int_a^b f(t + h, x) dx - \int_a^b f(t, x) dx}{h} = \int_a^b \frac{f(t + h, x) - f(t, x)}{h}\ dx \tag{164}$$

and take the limit $h \to 0$. The only issue is whether we can take the limit under the integral sign on the right. This can be justified under the hypotheses of the theorem.

example:

$$\frac{d}{dt}\left[\int_0^1 \log(x^2+t^2)dx\right] = \int_0^1 \frac{2t}{x^2+t^2}dx$$
$$= \left[2\tan^{-1}\left(\frac{x}{t}\right)\right]_{x=0}^{x=1} \tag{165}$$
$$= 2\tan^{-1}\left(\frac{1}{t}\right)$$

problem: Find

$$\int_0^1 \frac{\sqrt{x}-1}{\log x}dx \tag{166}$$

solution: We solve a more general problem which is to evaluate

$$\phi(t) = \int_0^1 \frac{x^t-1}{\log x}dx \tag{167}$$

Then $\phi(1/2)$ is the answer to the original problem. Differentiating under the integral sign and taking account that $\partial(x^t)/\partial t = x^t \log x$ we have

$$\phi'(t) = \int_0^1 x^t\ dx = \left[\frac{x^{t+1}}{t+1}\right]_{x=0}^{x=1} = \frac{1}{t+1} \tag{168}$$

Hence

$$\phi(t) = \log(t+1) + C \tag{169}$$

for some constant C. But we know $\phi(0) = 0$ so we must have $C = 0$. Thus $\phi(t) = \log(t+1)$ and the answer is $\phi(1/2) = \log(3/2)$.

1.13 Leibniz' rule

The following is a generalization of the previous result where we allow the endpoints to be functions of t.

Theorem 10

$$\frac{d}{dt}\left[\int_{a(t)}^{b(t)} f(t,x)dx\right] = f\Big(t,b(t)\Big)b'(t) - f\Big(t,a(t)\Big)a'(t) + \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(t,x)\ dx \tag{170}$$

Proof. Let

$$F(b,a,t) = \int_a^b f(t,x)\ dx \tag{171}$$

What we want is the derivative of $F(b(t),a(t),t)$ and by the chain rule this is

$$\begin{aligned}&\frac{d}{dt}[F(b(t),a(t),t)]\\ &= F_b(b(t),a(t),t)b'(t) + F_a(b(t),a(t),t)a'(t) + F_t(b(t),a(t),t)\end{aligned} \tag{172}$$

But

$$F_b(b,a,t) = f(t,b) \qquad F_a(b,a,t) = -f(t,a) \qquad F_t(b,a,t) = \int_a^b \frac{\partial f}{\partial t}(t,x)\ dx \tag{173}$$

which gives the result.

example:

$$\begin{aligned}&\frac{d}{dt}\left[\int_t^{t^2} \sin(t^2+x^2)dx\right]\\ =&\sin(t^2+x^2)|_{x=t^2}\frac{d(t^2)}{dt} - \sin(t^2+x^2)|_{x=t}\frac{d(t)}{dt} + \int_t^{t^2}\frac{\partial}{\partial t}\sin(t^2+x^2)dx\\ =&\sin(t^2+t^4)2t - \sin(2t^2) + 2t\int_t^{t^2}\cos(t^2+x^2)dx\end{aligned} \tag{174}$$

example: (forced harmonic oscillator). Suppose we want to find a function $x(t)$ whose derivatives $x'(t), x''(t)$ solve the ordinary differential equation

$$mx'' + kx = f(t) \tag{175}$$

with the inital conditions

$$x(0) = 0 \qquad x'(0) = 0 \tag{176}$$

Here k, m are positive constants and $f(t)$ is an arbitrary function.

We claim that a solution is

$$x(t) = \frac{1}{m\omega}\int_0^t \sin(\omega(t-\tau))f(\tau)d\tau \qquad \omega = \sqrt{\frac{k}{m}} \tag{177}$$

To check this note first that $x(0) = 0$. Then take the first derivative using the Leibniz rule and find

$$x'(t) = \frac{1}{m\omega}[\sin(\omega(t-\tau))f(\tau)]_{\tau=t} + \frac{1}{m\omega}\int_0^t \omega\cos(\omega(t-\tau))f(\tau)d\tau \tag{178}$$

The first term is zero and we also note that $x'(0) = 0$. Taking one more derivative yields

$$x''(t) = \frac{1}{m}[\cos(\omega(t-\tau))f(\tau)]_{\tau=t} + \frac{1}{m\omega}\int_0^t (-\omega^2)\sin(\omega(t-\tau))f(\tau)d\tau \tag{179}$$

which is the same as

$$x''(t) = \frac{1}{m}f(t) - \omega^2 x(t) \tag{180}$$

Now multiply by m, use $m\omega^2 = k$ and we recover the differential equation.

1.14 calculus of variations

We consider the problem of finding maximima and minima for a functional - i.e. for a function of functions.

As an example consider the problem of finding the curve between two points (x_0, y_0) and (x_1, y_1) which has the shortest length. For any such curve which is the graph of a function $y = y(x)$ with $y(x_0) = y_0$ and $y(x_1) = y_1$. The length is

$$I(y) = \int_{x_0}^{x_1} \sqrt{1 + y'(x)^2}dx \tag{181}$$

We seek to find the function $y = y(x)$ which gives the minimum value for $I(y)$,

More generally suppose we have a function $F(x, y, y')$ of three real variables (x, y, y'). For any differentiable function $y = y(x)$ satisfying $y(x_0) = y_0$ and $y(x_1) = y_1$ form the integral

$$I(y) = \int_{x_0}^{x_1} F(x, y(x), y'(x))dx \tag{182}$$

The question is which function $y = y(x)$ minimizes (or maximizes) the functional $I(y)$. We are looking for a local minimum (or maximum), that is we want to find a function y such that $I(\tilde{y}) \geq I(y)$ for all functions $\tilde{y}$ near to y in the sense that $\tilde{y}(x)$ is close to $y(x)$ for all $x_0 \leq x \leq x_1$.

Theorem 11 *If $y = y(x)$ is a local maximum or minimum for $I(y)$ with fixed endpoints then it satisfies*

$$F_y(x, y(x), y'(x)) - \frac{d}{dx}(F_{y'}(x, y(x), y'(x)) = 0 \tag{183}$$

Remarks.

1. The equation is called *Euler's equation* and is abreviated as

$$F_y - \frac{d}{dx}F_{y'} = 0 \tag{184}$$

2. Note that y, y' mean two different things in Euler's equation. First one evaluates the partial derivatives $F_y, F_{y'}$ treating y, y' as independent variables. Then in computing d/dx one treats y, y' as a function and its derivative.

3. The converse may not be true, i.e solutions of Euler's equation are not necessarily maxima or minima. They are just candidates and one should decide by other criteria.

proof: Suppose $y = y(x)$ is a local minimum. Pick any differentiable function $\eta(x)$ defined for $x_0 \leq x \leq x_1$ and satisfying $\eta(x_0) = \eta(x_1) = 0$. Then for any real number t the function

$$y_t(x) = y(x) + t\eta(x) \tag{185}$$

is differentiable with respect to x and satisfies $y_t(x_0) = y_t(x_1) = 0$. Consider the function

$$J(t) = I(y_t) = \int_{x_0}^{x_1} F(x, y_t(x), y_t'(x))dx \tag{186}$$

Since y_t is near $y_0 = y$ for t small, and since y_0 is a local minimum for I, we have

$$J(t) = I(y_t) \geq I(y_0) = J(0) \tag{187}$$

Thus $t = 0$ is a local minimum for $J(t)$ and it follows that $J'(0) = 0$.

To see what this says we differentiate under the integral sign and compute

$$\begin{aligned}
J'(t) &= \int_{x_0}^{x_1} \frac{\partial}{\partial t} F(x, y_t(x), y_t'(x))dx \\
&= \int_{x_0}^{x_1} \left(F_y(x, y_t(x), y_t'(x)) \frac{\partial}{\partial t}(y_t(x)) + F_{y'}(x, y_t(x), y_t'(x)) \frac{\partial}{\partial t}(y_t'(x)) \right) dx \\
&= \int_{x_0}^{x_1} \left(F_y(x, y_t(x), y_t'(x))\eta(x) + F_{y'}(x, y_t(x), y_t'(x))\eta'(x) \right) dx
\end{aligned} \tag{188}$$

Here in the second step we have used the chain rule. Now in the second term integrate by parts taking the derivative off $\eta'(x) = (d/dx)\eta$ and putting it on $F_{y'}(x, y_t(x), y_t'(x))$ The term involving the endpoints vanishes because $\eta(x_0) = \eta(x_1) = 0$. Then we have

$$J'(t) = \int_{x_0}^{x_1} \left(F_y(x, y_t(x), y_t'(x)) - \frac{d}{dx} F_{y'}(x, y_t(x), y_t'(x)) \right) \eta(x)dx \tag{189}$$

Now put $t = 0$ and get

$$0 = J'(0) = \int_{x_0}^{x_1} \left(F_y(x, y(x), y'(x)) - \frac{d}{dx} F_{y'}(x, y(x), y'(x)) \right) \eta(x)dx \tag{190}$$

Since this is true for an arbitrary function η it follows that the expression in parentheses must be zero which is our result. [1]

example: We return to the problem of finding the curve between two points with the shortest length. That is we seek to minimize

$$I(y) = \int_{x_0}^{x_1} \sqrt{1 + y'(x)^2} dx \tag{191}$$

This has the form (182) with

$$F(x, y, y') = \sqrt{1 + (y')^2} \tag{192}$$

The minimizer must satisfy Euler's equation. Since $F_y = 0$ and $F_{y'} = y'/\sqrt{1 + (y')^2}$ this says

$$F_y - \frac{d}{dx} F_{y'} = \frac{d}{dx} \left[\frac{y'}{\sqrt{1 + (y')^2}} \right] = 0 \tag{193}$$

Evaluating the derivatives yields

$$\frac{\sqrt{1 + (y')^2} y'' - (y')^2 y'' / \sqrt{1 + (y')^2}}{1 + (y')^2} = 0 \tag{194}$$

Now multiple by $(1 + (y')^2)^{3/2}$ and get

$$(1 + (y')^2) y'' - (y')^2 y'' = 0 \tag{195}$$

which is the same as $y'' = 0$. Thus the minimizer must have the form $y = ax + b$ for some constants a, b. So the shortest distance between two points is along a straight line as expected.

example: The problem is to find the function $y = y(x)$ with $y(0) = 0$ and $y(1) = 1$ which minimizes the integral

$$I(y) = \frac{1}{2} \int_0^1 (y(x)^2 + (y'(x))^2) dx \tag{196}$$

and again we assume there is such a minimizer. The integral has the form (182) with

$$F(x, y, y') = \frac{1}{2} (y^2 + (y')^2) \tag{197}$$

---

[1] In general if f is a continuous function $\int_a^b f(x)dx = 0$ does not imply that $f \equiv 0$. However it is true if $f(x) \geq 0$. If $\int_a^b f(x)\eta(x)dx = 0$ for any continuous function η then we can take $\eta(x) = f(x)$ and get $\int_a^b f(x)^2 dx = 0$, hence $f(x)^2 = 0$, hence $f(x) = 0$. This is not quite the situation above since we also restricted η to vanish at the endpoints. But the conclusion is still valid.

Then Euler's equation says

$$F_y - \frac{d}{dx}(F_{y'}) = y - \frac{d}{dx}y' = y - y'' = 0 \tag{198}$$

Thus we must solve the second order equation $y - y'' = 0$. Since the equation has constant coefficients one can find solutions by trying $y = e^{rx}$. One finds that $r^2 = 1$ and so $y = e^{\pm x}$ are solutions The general solution is

$$y(x) = c_1 e^x + c_2 e^{-x} \tag{199}$$

The constants c_1, c_2 are fixed by the condition $y(0) = 0$ and $y(1) = 1$ and one finds

$$y(x) = \frac{e}{e^2 - 1}(e^x - e^{-x}) = \frac{2e}{e^2 - 1}\sinh x \tag{200}$$

example : Suppose that an object moves on a line and its position at time t is given by a function $x(t)$. Suppose also we know that $x(t_0) = x_0$ and $x(t_1) = x_1$ and that it is moving in a force field $F(x) = -dV/dx$ determined by some potential function $V(x)$. What is the trajectory $x(t)$?

One way to proceed is to form a function called the Lagrangian by taking the difference of the kinetic and potential energy:

$$L(x, x') = \frac{1}{2}m(x')^2 - V(x) \tag{201}$$

For any trajectory $x = x(t)$ one forms the action

$$A(x) = \int_{t_0}^{t_1} L(x(t), x'(t))dt \tag{202}$$

According to D'Alembert's principle the actual trajectory is the one that minimizes the action. This is also called the principle of least action.

To see what it says we observe that this problem has the form (182) with new names for the variables. Euler's equation says

$$L_x - \frac{d}{dt}L_{x'} = 0 \tag{203}$$

But $L_x = -dV/dx = F$ and $L_{x'} = mx'$ so this is

$$F - mx'' = 0 \tag{204}$$

which is Newton's second law. Thus the principle of least action is an alternative to Newton's second law. This turns out to be true for many dynamical problems.

2 vector calculus in $\mathbb{R}^3$

2.1 vectors

We continue our survey of multivariable caculus but now put special emphasis on $\mathbb{R}^3$ which is a model for physical space.

Vectors in $\mathbb{R}^3$ will now be indicated by arrows or bold face type as in $\mathbf{u} = (u_1, u_2, u_3)$. Any such vector can be written

$$\begin{aligned} \mathbf{u} =& (u_1, u_2, u_3) \\ =& u_1(1,0,0) + u_2(0,1,0) + u_3(0,0,1) \\ =& u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k} \end{aligned} \tag{205}$$

where we have defined

$$\mathbf{i} = (1,0,0) \qquad \mathbf{j} = (0,1,0) \qquad \mathbf{k} = (0,0,1) \tag{206}$$

Any vector can be written as a linear combination of the independent vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$ so these form a basis for $\mathbb{R}^3$ called the *standard basis*.

We consider several products of vectors:

dot product: The dot product is defined either by

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + u_3v_3 \tag{207}$$

or by

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}|\cos\theta \tag{208}$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$. Note that $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2$. Also note that $\mathbf{u}$ is orthogonal (perpendicular) to $\mathbf{v}$ if and only if $\mathbf{u} \cdot \mathbf{v} = 0$.

The dot product has the properties

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} =& \mathbf{v} \cdot \mathbf{u} \\ (\alpha\mathbf{u}) \cdot \mathbf{v} =& \alpha(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (\alpha\mathbf{v}) \\ (\mathbf{u}_1 + \mathbf{u}_2) \cdot \mathbf{v} =& \mathbf{u}_1 \cdot \mathbf{v} + \mathbf{u}_2 \cdot \mathbf{v} \end{aligned} \tag{209}$$

Examples are

$$\begin{aligned} &\mathbf{i} \cdot \mathbf{i} = 1 \qquad \mathbf{j} \cdot \mathbf{j} = 1 \qquad \mathbf{j} \cdot \mathbf{j} = 1 \\ &\mathbf{i} \cdot \mathbf{j} = 0 \qquad \mathbf{j} \cdot \mathbf{k} = 0 \qquad \mathbf{k} \cdot \mathbf{i} = 0 \end{aligned} \tag{210}$$

This says that $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are orthogonal unit vectors. They are an example of an *orthonormal basis* for $\mathbb{R}^3$.

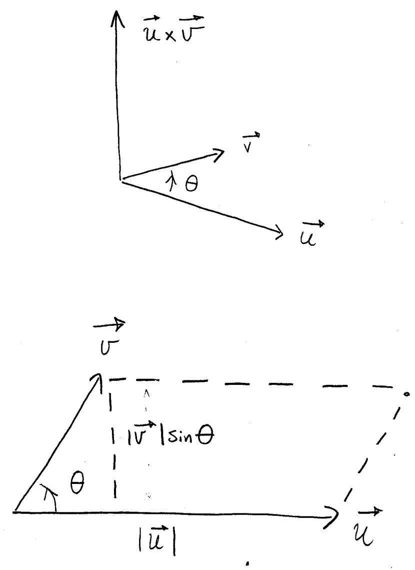


Figure 6: cross product

cross product: (only in $\mathbb{R}^3$) The *cross product* of $\mathbf{u}$ and $\mathbf{v}$ is a vector $\mathbf{u} \times \mathbf{v}$ which has length

$$|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}||\mathbf{v}| \sin \theta \tag{211}$$

Here θ is the positive angle between the vectors. The direction of $\mathbf{u} \times \mathbf{v}$ is specified by requiring that it be perpendicular to $\mathbf{u}$ and $\mathbf{v}$ in such a way that $\mathbf{u}, \mathbf{v}, \mathbf{u} \times \mathbf{v}$ form a right-handed system. (See figure 6)

The length $|\mathbf{u} \times \mathbf{v}|$ is interpreted as the area of the parallelogram spanned by $\mathbf{u}, \mathbf{v}$. This follows since the parallelogram has base $|\mathbf{u}|$ and height $|\mathbf{v}| \sin \theta$ (See figure 6) and so

$$\begin{aligned} \text{area} =& \text{ base } \times \text{height} \\ =&|\mathbf{u}| \; |\mathbf{v}| \sin \theta \\ =&|\mathbf{u} \times \mathbf{v}| \end{aligned} \tag{212}$$

An alternate definition of the cross-product uses determinants. Recall that

$$\begin{aligned} \det \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} =& a_1 \det \begin{pmatrix} b_2 & b_3 \\ c_2 & c_3 \end{pmatrix} - a_2 \det \begin{pmatrix} b_1 & b_3 \\ c_1 & c_3 \end{pmatrix} + a_3 \det \begin{pmatrix} b_1 & b_2 \\ c_1 & c_2 \end{pmatrix} \\ =& a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1) \end{aligned} \tag{213}$$

The other definition is

$$\mathbf{u}\times\mathbf{v} = \det\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} = \mathbf{i}(u_2v_3 - u_3v_2) + \mathbf{j}(u_3v_1 - u_1v_3) + \mathbf{k}(u_1v_2 - u_2v_1) \quad (214)$$

The cross product has the following properties

$$\begin{aligned} \mathbf{u}\times\mathbf{u} =& 0 \\ \mathbf{u}\times\mathbf{v} =& -\mathbf{v}\times\mathbf{u} \\ (\alpha\mathbf{u})\times\mathbf{v} =& \alpha(\mathbf{u}\times\mathbf{v}) = \mathbf{u}\times(\alpha\mathbf{v}) \\ (\mathbf{u}_1+\mathbf{u}_2)\times\mathbf{v} =& \mathbf{u}_1\times\mathbf{v} + \mathbf{u}_2\times\mathbf{v} \end{aligned} \quad (215)$$

Examples are

$$\mathbf{i}\times\mathbf{j} = \mathbf{k} \qquad \mathbf{j}\times\mathbf{k} = \mathbf{i} \qquad \mathbf{k}\times\mathbf{i} = \mathbf{j} \quad (216)$$

triple product: The triple product of vectors $\mathbf{w}, \mathbf{u}, \mathbf{v}$ is defined by

$$\begin{aligned} \mathbf{w}\cdot(\mathbf{u}\times\mathbf{v}) =& w_1(u_2v_3 - u_3v_2) + w_2(u_3v_1 - u_1v_3) + w_3(u_1v_2 - u_2v_1) \\ =& \det\begin{pmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} \end{aligned} \quad (217)$$

The absolute value $|\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})|$ is the volume of the *parallelopiped* spanned by $\mathbf{u}, \mathbf{v}, \mathbf{w}$ (see figure 16). This is so because if ϕ is the angle between $\mathbf{w}$ and $\mathbf{u}\times\mathbf{v}$ then

$$\begin{aligned} \text{volume } =& \text{ (area of base) } \times \text{ height} \\ =& \Big(|\mathbf{u}\times\mathbf{v}|\Big)\Big(|\mathbf{w}|\cos\phi\Big) \\ =& |\mathbf{w}\cdot(\mathbf{u}\times\mathbf{v})| \end{aligned} \quad (218)$$

problem: Find the volume of the *parallelopiped* spanned by $\mathbf{i}+\mathbf{j},\ \mathbf{j},\ \mathbf{i}+\mathbf{j}+\mathbf{k}$.

solution:

$$\det\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = 1 \quad (219)$$

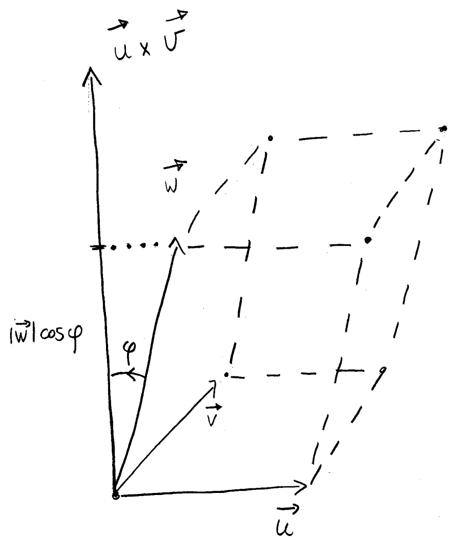


Figure 7: triple product

Planes: A plane is determined by a particular point $\mathbf{R}_0 = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k}$ in the plane and a vector $\mathbf{N} = N_1\mathbf{i} + N_2\mathbf{j} + N_3\mathbf{k}$ perpendicular to the plane, called a normal vector. If $\mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ is any other point in the plane, then $\mathbf{R} - \mathbf{R}_0$ lies in the plane, hence it is orthogonal to $\mathbf{N}$ and hence (see figure 8)

$$\mathbf{N} \cdot (\mathbf{R} - \mathbf{R}_0) = 0 \tag{220}$$

In fact this is the equation of the plane. That is a point $\mathbf{R}$ lies on the plane if and only if it satisfies the equation. Written out it says

$$N_1(x - x_0) + N_2(y - y_0) + N_3(z - z_0) = 0 \tag{221}$$

problem: Find the plane determined by the three points $\mathbf{a} = \mathbf{i}, \mathbf{b} = 2\mathbf{j}, \mathbf{c} = 3\mathbf{k}$.

solution: $\mathbf{b} - \mathbf{a} = -\mathbf{i} + 2\mathbf{j}$ and $\mathbf{c} - \mathbf{a} = -\mathbf{i} + 3\mathbf{k}$ both lie in the the plane. Their cross product is orthogonal to both, hence to the plane, and can be take as a normal vector:

$$\mathbf{N} = (\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a}) = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix} = 6\mathbf{i} + 3\mathbf{j} + 2\mathbf{k} \tag{222}$$

For the particular point take $\mathbf{R}_0 = \mathbf{a} = \mathbf{i}$. Then the equation of the plane is

$$\mathbf{N} \cdot (\mathbf{R} - \mathbf{R}_0) = 6(x - 1) + 3y + 2z = 0 \tag{223}$$

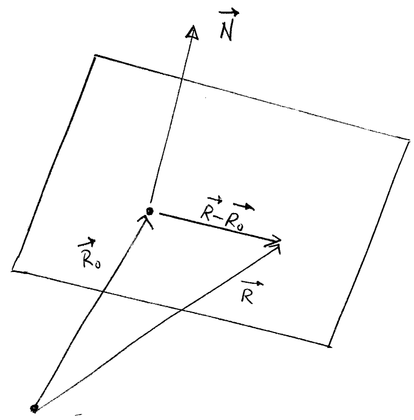


Figure 8:

2.2 vector-valued functions

A vector-valued function is a function from $\mathbb{R}$ to $\mathbb{R}^3$ (or more generally $\mathbb{R}^n$). It is written

$$\mathbf{R}(t) = (x(t), y(t), z(t)) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} \tag{224}$$

To say $\mathbf{R}(t)$ has limit

$$\lim_{t \to t_0} \mathbf{R}(t) = \mathbf{R}_0 \tag{225}$$

means that

$$\lim_{t \to t_0} |\mathbf{R}(t) - \mathbf{R}_0| = 0 \tag{226}$$

If $\mathbf{R}_0 = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{j}$ then

$$|\mathbf{R}(t) - \mathbf{R}_0| = \sqrt{(x(t) - x_0)^2 + (y(t) - y_0)^2 + (z(t) - z_0)^2} \tag{227}$$

Hence $\lim_{t \to t_0} \mathbf{R}(t) = \mathbf{R}_0$ is the same as the three limits

$$\lim_{t \to t_0} x(t) = x_0 \qquad \lim_{t \to t_0} y(t) = y_0 \qquad \lim_{t \to t_0} z(t) = z_0 \tag{228}$$

The function $\mathbf{R}(t)$ is *continuous* if

$$\lim_{t \to t_0} \mathbf{R}(t) = \mathbf{R}(t_0) \tag{229}$$

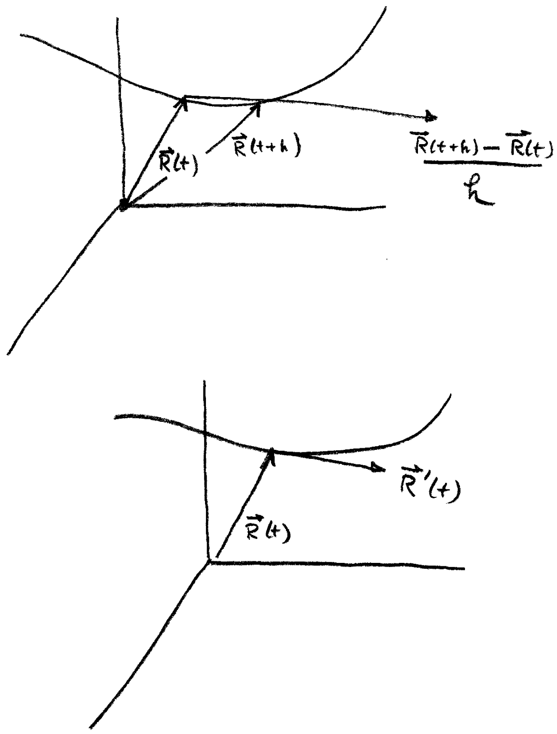


Figure 9: tangent vector to a curve

This is the same as saying that the components $x(t), y(t), z(t)$ are all continuous.

The function $\mathbf{R}(t)$ is *differentiable* if

$$\mathbf{R}'(t) = \frac{d\mathbf{R}}{dt} = \lim_{h\to 0} \frac{\mathbf{R}(t+h) - \mathbf{R}(t)}{h} \tag{230}$$

exists. Since

$$\frac{\mathbf{R}(t+h) - \mathbf{R}(t)}{h} = \frac{x(t+h) - x(t)}{h}\mathbf{i} + \frac{y(t+h) - y(t)}{h}\mathbf{j} + \frac{z(t+h) - z(t)}{h}\mathbf{k} \tag{231}$$

this is the same as saying that the components $x(t), y(t), z(t)$ are all differentiable, in which case the derivative is

$$\mathbf{R}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k} \tag{232}$$

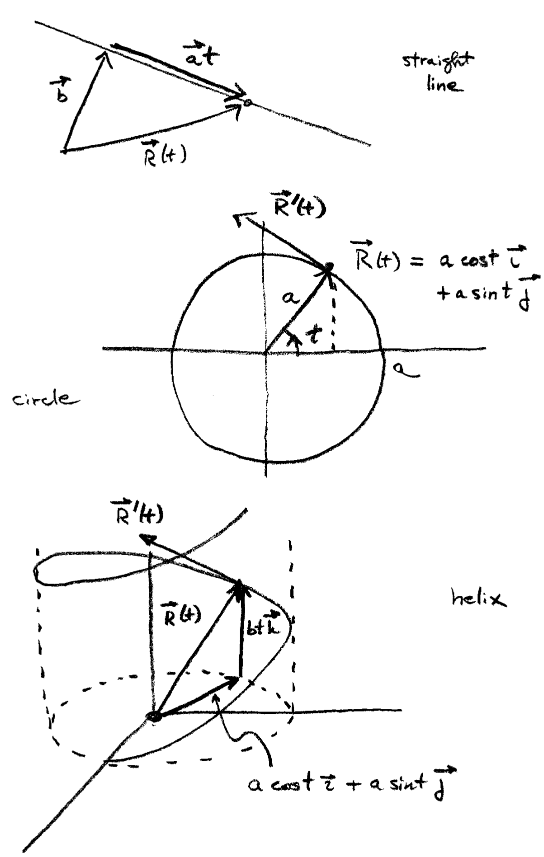


Figure 10: some examples

In other words you find the derivative by differentiating the components

The range of a continuous function $\mathbf{R}(t)$ is a curve. The derivative $\mathbf{R}'(t)$ has the interpretation of being a tangent to the curve as figure 9 shows.

A common application is that $\mathbf{R}(t)$ is the location of some object at time t. Then $\mathbf{R}'(t)$ is the *velocity* and the magnitude of the velocity $|\mathbf{R}'(t)|$ is the *speed*. The second derivative $\mathbf{R}''(t)$ is the acceleration.

Here are some examples illustrated in figure 10

example: straight line. Consider

$$\mathbf{R}(t) = \mathbf{a}t + \mathbf{b} \tag{233}$$

Then $\mathbf{R}(0) = \mathbf{b}$ and $\mathbf{R}'(t) = \mathbf{a}$ so it is a straight line through $\mathbf{b}$ in the direction $\mathbf{a}$.

example: circle. Let $\mathbf{R}(t)$ be the point on a circle of radius a with polar angle t. As t increases it travels around the circle at uniform speed. The point has Cartesian coordinates $x = a\cos t, y = a\sin t$ so

$$\mathbf{R}(t) = a\cos t\ \mathbf{i} + a\sin t\ \mathbf{j} \tag{234}$$

The velocity is

$$\mathbf{R}'(t) = -a\sin t\ \mathbf{i} + a\cos t\ \mathbf{j} \tag{235}$$

and the speed is $|\mathbf{R}'| = a$.

example: helix. To the previous example we add a constant velocity in the z-direction

$$\mathbf{R}(t) = a\cos t\ \mathbf{i} + a\sin t\ \mathbf{j} + bt\mathbf{k} \tag{236}$$

This describes a helix and we have

$$\mathbf{R}'(t) = -a\sin t\ \mathbf{i} + a\cos t\ \mathbf{j} + b\mathbf{k} \tag{237}$$

2.3 other coordinate systems

We next describe vector valued functions using other coordinate systems.

A. **Polar:** First some general remarks about vectors and polar coordinates in $\mathbb{R}^2$. Let $\mathbf{R}(r,\theta)$ be the point with polar coordinates r, θ. This has Cartesian coordinates $x = r\cos\theta, y = r\sin\theta$ so

$$\mathbf{R}(r,\theta) = r\cos\theta\mathbf{i} + r\sin\theta\mathbf{j} \tag{238}$$

If we vary r with θ fixed we get an "r-line". The tangent vector to this line is

$$\frac{\partial\mathbf{R}}{\partial r}(r,\theta) = \cos\theta\mathbf{i} + \sin\theta\mathbf{j} \tag{239}$$

If we vary θ with r fixed we get an "θ-line". The tangent vector to this line is

$$\frac{\partial\mathbf{R}}{\partial\theta}(r,\theta) = -r\sin\theta\mathbf{i} + r\cos\theta\mathbf{j} \tag{240}$$

We also consider unit tangent vectors to these coordinate lines:

$$\begin{aligned}\mathbf{e}_r(\theta) =& \frac{\partial\mathbf{R}/\partial r}{|\partial\mathbf{R}/\partial r|} = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}\\ \mathbf{e}_\theta(\theta) =& \frac{\partial\mathbf{R}/\partial\theta}{|\partial\mathbf{R}/\partial\theta|} = -\sin\theta\mathbf{i} + \cos\theta\mathbf{j}\end{aligned} \tag{241}$$

See figure 11.

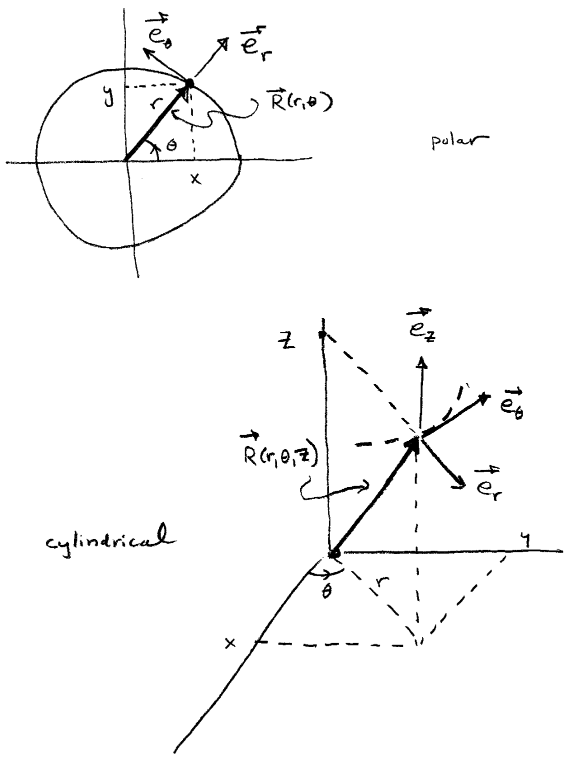


Figure 11: polar and cylindrical basis vectors

Note that

$$\mathbf{e}_r(\theta) \cdot \mathbf{e}_\theta(\theta) = \cos\theta(-\sin\theta) + \sin\theta\cos\theta = 0 \tag{242}$$

Thus for any θ the vectors $\mathbf{e}_r(\theta), \mathbf{e}_\theta(\theta)$ form an orthonormal set of vectors in $\mathbb{R}^2$ and hence an orthonormal basis. In the new basis we can write

$$\mathbf{R}(r,\theta) = r\mathbf{e}_r(\theta) \tag{243}$$

For future reference note that

$$\begin{aligned} \frac{d\mathbf{e}_r}{d\theta} &= \mathbf{e}_\theta \\ \frac{d\mathbf{e}_\theta}{d\theta} &= -\,\mathbf{e}_r \end{aligned} \tag{244}$$

Now a curve is specified in polar coordinates by a pair of functions $r(t), \theta(t)$. So the curve is

$$\mathbf{R}(t) = \mathbf{R}\Big(r(t), \theta(t)\Big) = r(t)\mathbf{e}_r(\theta(t)) \tag{245}$$

The velocity is

$$\mathbf{R}'(t) = r'(t)\mathbf{e}_r(\theta(t)) + r(t)\frac{d}{dt}\mathbf{e}_r(\theta(t)) \tag{246}$$

But

$$\frac{d}{dt}\mathbf{e}_r(\theta(t)) = \frac{d\mathbf{e}_r}{d\theta}(\theta(t))\frac{d\theta}{dt} = \theta'(t)\mathbf{e}_\theta(\theta(t)) \tag{247}$$

and so

$$\mathbf{R}'(t) = r'(t)\mathbf{e}_r(\theta(t)) + r(t)\theta'(t)\mathbf{e}_\theta(\theta(t)) \tag{248}$$

By differentiating this we get a formula for the acceleration $\mathbf{R}''(t)$:

$$\mathbf{R}''(t) = \Big(r''(t) - r(t)(\theta'(t))^2\Big)\mathbf{e}_r(\theta(t)) + \Big(r(t)\theta''(t) + 2r'(t)\theta'(t)\Big)\mathbf{e}_\theta(\theta(t)) \tag{249}$$

We summarize in an abbreviated notation

$$\begin{aligned} \mathbf{R} &= r\mathbf{e}_r \\ \mathbf{R}' &= r'\mathbf{e}_r + r\theta'\mathbf{e}_\theta \\ \mathbf{R}'' &= (r'' - r(\theta')^2)\mathbf{e}_r + (r\theta'' + 2r'\theta')\mathbf{e}_\theta \end{aligned} \tag{250}$$

example: Consider the spiral described in polar coordinates by $r = at$ and $\theta = bt$. Then $r' = a, \theta' = b$ and $r'' = 0, \theta'' = 0$ and so

$$\begin{aligned} \mathbf{R} &= at\ \mathbf{e}_r \\ \mathbf{R}' &= a\ \mathbf{e}_r + abt\ \mathbf{e}_\theta \\ \mathbf{R}'' &= -ab^2t\ \mathbf{e}_r + 2ab\ \mathbf{e}_\theta \end{aligned} \tag{251}$$

In these formulas $\mathbf{e}_r = \mathbf{e}_r(bt) = \cos(bt)\mathbf{i} + \sin(bt)\mathbf{j}$ and $\mathbf{e}_\theta = \mathbf{e}_\theta(bt) = -\sin(bt)\mathbf{i} + \cos(bt)\mathbf{j}$.

B. **cylindrical:** Cylindrical coordinates in $\mathbb{R}^3$ replace x, y by polar coordinates r, θ and leave z alone. Thus

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ z &= z \end{aligned} \tag{252}$$

A point with cylindrical coordinates r, θ, z is

$$\mathbf{R}(r, \theta, z) = r\cos\theta\ \mathbf{i} + r\sin\theta\ \mathbf{j} + z\ \mathbf{k} \tag{253}$$

Unit tangent vectors to the coordinate lines are $\mathbf{e}_r, \mathbf{e}_\theta$ as before and $\mathbf{e}_z = \mathbf{k}$. See figure 11. In the new basis we have

$$\mathbf{R}(r, \theta, z) = r\mathbf{e}_r(\theta) + z\mathbf{e}_z \tag{254}$$

A curve in cylindrical coordinate is given by $r(t), \theta(t), z(t)$ and we have

$$\mathbf{R}(t) = \mathbf{R}\Big(r(t), \theta(t), z(t)\Big) = r(t)\mathbf{e}_r(\theta(t)) + z(t)\mathbf{e}_z \tag{255}$$

As before:

$$\begin{aligned}\mathbf{R} &= r\mathbf{e}_r + z\mathbf{e}_z \\ \mathbf{R}' &= r'\mathbf{e}_r + r\theta'\mathbf{e}_\theta + z'\mathbf{e}_z \\ \mathbf{R}'' &= (r'' - r(\theta')^2)\mathbf{e}_r + (r\theta'' + 2r'\theta')\mathbf{e}_\theta + z''\mathbf{e}_z\end{aligned} \tag{256}$$

C. **spherical:** Spherical coordinates ρ, ϕ, θ label a point by its distance to the origin, the angle with the z-axis, and the polar angle when projected into the x, y plane. The corresponding Cartesian coordinates are

$$\begin{aligned}x &= \rho \sin\phi \cos\theta \\ y &= \rho \sin\phi \sin\theta \\ z &= \rho \cos\phi\end{aligned} \tag{257}$$

The point with spherical coordinates ρ, ϕ, θ is

$$\mathbf{R}(\rho, \phi, \theta) = \rho \sin\phi \cos\theta \ \mathbf{i} + \rho \sin\phi \sin\theta \ \mathbf{j} + \rho \cos\phi \ \mathbf{k} \tag{258}$$

Tangent vectors to the coordinate lines are

$$\begin{aligned}\frac{\partial \mathbf{R}}{\partial \rho} &= \sin\phi \cos\theta \ \mathbf{i} + \sin\phi \sin\theta \ \mathbf{j} + \cos\phi \ \mathbf{k} \\ \frac{\partial \mathbf{R}}{\partial \phi} &= \rho \cos\phi \cos\theta \ \mathbf{i} + \rho \cos\phi \sin\theta \ \mathbf{j} - \rho \sin\phi \ \mathbf{k} \\ \frac{\partial \mathbf{R}}{\partial \theta} &= -\rho \sin\phi \sin\theta \ \mathbf{i} + \rho \sin\phi \cos\theta \ \mathbf{j}\end{aligned} \tag{259}$$

Divide by the length and get unit tangent vectors to the coordinate lines: (see figure 12)

$$\begin{aligned}\mathbf{e}_\rho(\phi, \theta) &= \sin\phi \cos\theta \ \mathbf{i} + \sin\phi \sin\theta \ \mathbf{j} + \cos\phi \ \mathbf{k} \\ \mathbf{e}_\phi(\phi, \theta) &= \cos\phi \cos\theta \ \mathbf{i} + \cos\phi \sin\theta \ \mathbf{j} - \sin\phi \ \mathbf{k} \\ \mathbf{e}_\theta(\phi, \theta) &= -\sin\theta \ \mathbf{i} + \cos\theta \ \mathbf{j}\end{aligned} \tag{260}$$

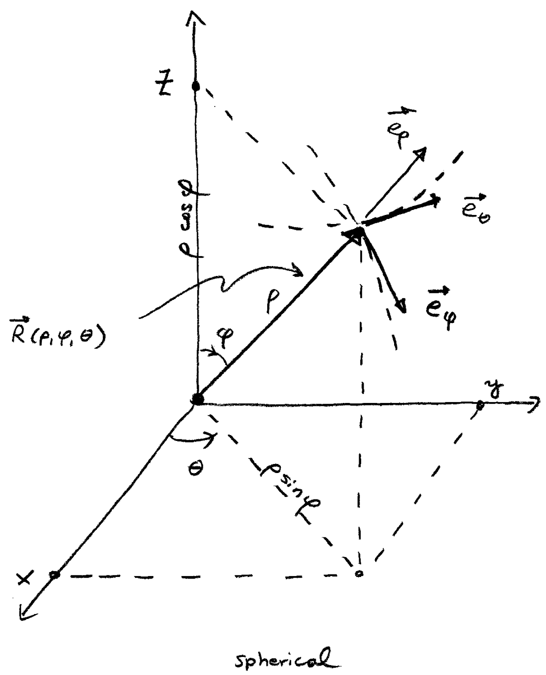


Figure 12: spherical basis vectors

In terms of the new basis vectors we have

$$\mathbf{R}(\rho,\phi,\theta) = \rho\ \mathbf{e}_\rho(\phi,\theta) \tag{261}$$

A curve in spherical coordinates is specified by three functions $\rho(t), \phi(t), \theta(t)$. So the curve is

$$\mathbf{R}(t) = \mathbf{R}\Big(\rho(t), \phi(t), \theta(t)\Big) = \rho(t)\mathbf{e}_\rho(\phi(t), \theta(t)) \tag{262}$$

Now we can take derivatives and we find

$$\begin{aligned}
\mathbf{R} =&\rho \mathbf{e}_\rho \\
\mathbf{R}' =&\rho' \mathbf{e}_\rho + \rho\phi' \ \mathbf{e}_\phi + \rho\theta' \sin\phi \ \mathbf{e}_\theta \\
\mathbf{R}'' =&\Big(\rho'' - \rho(\phi')^2 - \rho(\theta')^2 \sin^2\phi\Big)\mathbf{e}_\rho \\
&+\Big(\rho\phi'' + 2\rho'\phi' - \rho(\theta')^2 \sin\phi\cos\phi\Big)\mathbf{e}_\phi \\
&+\Big(\rho\theta''\sin\phi + 2\rho'\phi'\sin\phi + 2\rho\theta'\phi'\cos\phi\Big)\mathbf{e}_\theta
\end{aligned} \tag{263}$$

example: Suppose that $\rho = 1, \phi = at, \theta = bt$ with b much greater than a. This represents a point on a sphere sprialing down from the north pole. Then $\rho' = 0, \phi' = a, \theta' = b$ and $\rho'' = 0, \phi'' = 0, \theta'' = 0$ and we have with $\mathbf{e}_\rho = \mathbf{e}_\rho(at, bt)$, etc

$$\begin{aligned}
\mathbf{R} =&\mathbf{e}_\rho \\
\mathbf{R}' =&a \ \mathbf{e}_\phi + b\sin at \ \mathbf{e}_\theta \\
\mathbf{R}'' =&\Big(-a^2 - b^2\sin^2 at\Big)\mathbf{e}_\rho + \Big(-b^2 \sin at\cos at\Big)\mathbf{e}_\phi + \Big(2ab\cos at\Big)\mathbf{e}_\theta
\end{aligned} \tag{264}$$

2.4 line integrals

We want to define the length of a curve $\mathcal{C}$ in $\mathbb{R}^3$. Suppose the curve is the range of a vector valued function $\mathbf{R}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$, $a \leq t \leq b$. We say that $\mathbf{R}(t)$ is a *parametrization* of $\mathcal{C}$. There will be many parametrizations, but we pick one. We divide up the interval $[a, b]$ by picking points

$$a = t_0 < t_1 < t_2 < \cdots < t_n = b \tag{265}$$

This gives a sequence of points on the curve $\mathbf{R}(t_0), \mathbf{R}(t_1), \ldots \mathbf{R}(t_n)$. (see figure 13). If $\Delta t_i = t_{i+1} - t_i$ is small then for any t_i^* in the interval $[t_i, t_{i+1}]$

$$\begin{aligned}
\mathbf{R}(t_{i+1}) - \mathbf{R}(t_i) &= (x(t_{i+1}) - x(t_i))\mathbf{i} + (y(t_{i+1}) - y(t_i))\mathbf{j} + (z(t_{i+1}) - z(t_i))\mathbf{k} \\
&\approx x'(t_i^*)\Delta t_i\mathbf{i} + y'(t_i^*)\Delta t_i\mathbf{j} + z'(t_i^*)\Delta t_i\mathbf{k} \\
&= \mathbf{R}'(t_i^*)\Delta t_i
\end{aligned} \tag{266}$$

(The mean value theorem says there is a point t_i^* so that $(x(t_{i+1}) - x(t_i)) = x'(t_i^*)\Delta t_i$. Changing to an arbitrary point in the interval is second order small and negligible). Let Δs_i be the length of the straight line from $\mathbf{R}(t_i)$ to $\mathbf{R}(t_{i+1})$. Then

$$\Delta s_i = |\mathbf{R}(t_{i+1}) - \mathbf{R}(t_i)| \approx |\mathbf{R}'(t_i^*)|\Delta t_i \tag{267}$$

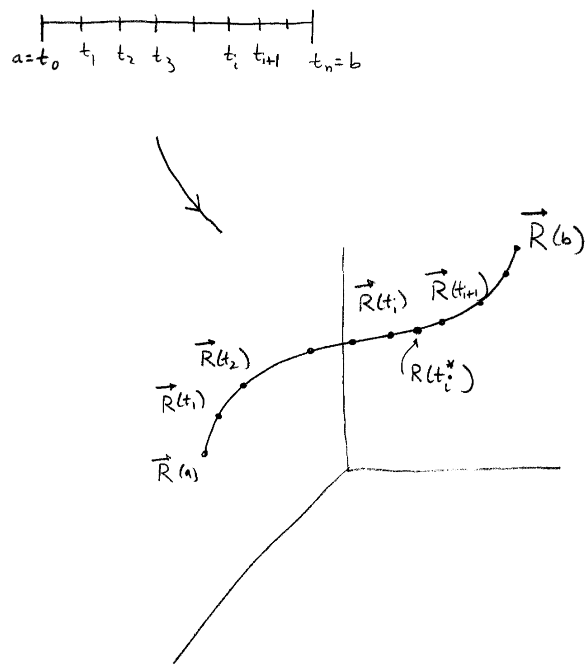


Figure 13: line integral

Then we have

$$\text{length of } \mathcal{C} \approx \sum_{i=0}^{n-1} \Delta s_i \approx \sum_{i=0}^{n-1} |\mathbf{R}'(t_i^*)| \Delta t_i \tag{268}$$

This is a Riemann sum and as the division becomes increasingly fine, i.e as $\max_i \Delta t_i$ tends to 0, this converges a Riemann integral which we take as the definition

$$\text{length of } \mathcal{C} = \int_a^b |\mathbf{R}'(t)| dt \tag{269}$$

One can show that this depends only on $\mathcal{C}$ and not on the particular parametrization.

More generally we want to define the integral of a function $f(\mathbf{R}) = f(x, y, z)$ over the curve $\mathcal{C}$. An approximation to what we want is

$$\sum_{i=0}^{n-1} f(\mathbf{R}(t_i^*))\Delta s_i \approx \sum_{i=0}^{n-1} f(\mathbf{R}(t_i^*))|\mathbf{R}'(t_i^*)|\Delta t_i \tag{270}$$

As the division becomes fine this converges to a Riemann integral which we take as the definition of the integral of f over $\mathcal{C}$. It is denoted $\int_{\mathcal{C}} f(\mathbf{R})ds$ and is given by

$$\int_{\mathcal{C}} f(\mathbf{R})ds = \int_a^b f(\mathbf{R}(t))|\mathbf{R}'(t)|dt \tag{271}$$

This is also independent of parametrization. A short way to remember it is to replace $\mathbf{R}$ by its parametrization $\mathbf{R}(t)$, replace $\mathcal{C}$ by the interval $[a, b]$ and replace the formal symbol ds by

$$ds = \left|\frac{d\mathbf{R}}{dt}\right| dt \tag{272}$$

Note also that if $f(\mathbf{R}) = 1$ we have

$$\int_{\mathcal{C}} ds = \text{length of } \mathcal{C} \tag{273}$$

The same formulas hold in $\mathbb{R}^2$ but now $\mathbf{R}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$.

example: Consider the helix parametrized by

$$\mathbf{R}(t) = a\cos t\ \mathbf{i} + a\sin t\ \mathbf{j} + bt\ \mathbf{k} \tag{274}$$

with $0 \le t \le 2\pi$. Then

$$\frac{d\mathbf{R}}{dt} = -a\sin t\ \mathbf{i} + a\cos t\ \mathbf{j} + b\ \mathbf{k} \tag{275}$$

and

$$ds = \left|\frac{d\mathbf{R}}{dt}\right| dt = \sqrt{a^2 + b^2}\ dt \tag{276}$$

The length is then

$$\int_{\mathcal{C}} ds = \int_0^{2\pi} \sqrt{a^2 + b^2}\ dt = 2\pi\sqrt{a^2 + b^2} \tag{277}$$

example: Suppose we have a thin semi-circular wire of radius a with uniform linear density ρ (mass per unit of length). We want to find the coordinates $(\bar{x}, \bar{y})$ of the center of mass. These are defined by dividing the wire up into segments of length Δs_i and mass $\Delta m_i = \rho\Delta s_i$ and computing

$$\bar{x} = \frac{\sum_i x_i\Delta m_i}{\sum_i \Delta m_i} = \frac{\sum_i x_i\Delta s_i}{\sum_i \Delta s_i} \qquad \bar{y} = \frac{\sum_i y_i\Delta m_i}{\sum_i \Delta m_i} = \frac{\sum_i y_i\Delta s_i}{\sum_i \Delta s_i} \tag{278}$$

where (x_i, y_i) are the coordinates of the i^{th} segment. As the division becomes fine this is expressed as a ratio of line integrals over the semi-circle $\mathcal{C}$

$$\bar{x} = \frac{\int_{\mathcal{C}} xds}{\int_{\mathcal{C}} ds} \qquad \bar{y} = \frac{\int_{\mathcal{C}} yds}{\int_{\mathcal{C}} ds} \tag{279}$$

To compute it parametrize the semi-circle in polar coordinates. In fact it is a theta line given by

$$\mathbf{R}(\theta) = \mathbf{R}(a, \theta) = a\cos\theta\mathbf{i} + a\sin\theta\mathbf{j} \tag{280}$$

with $0 \leq \theta \leq \pi$. Then

$$\frac{d\mathbf{R}}{d\theta} = -a\sin\theta\ \mathbf{i} + a\cos\theta\ \mathbf{j} \tag{281}$$

and

$$ds = \left|\frac{d\mathbf{R}}{d\theta}\right| d\theta = ad\theta \tag{282}$$

Then since $x = a\cos\theta$,y = a $\sin\theta$

$$\begin{aligned} \int_{\mathcal{C}} x\ ds &= \int_0^{\pi} a\cos\theta\ ad\theta = 0 \\ \int_{\mathcal{C}} y\ ds &= \int_0^{\pi} a\sin\theta\ ad\theta = 2a^2 \end{aligned} \tag{283}$$

and

$$\int_{\mathcal{C}} ds = \int_0^{\pi} ad\theta = a\pi \tag{284}$$

Thus

$$\bar{x} = \frac{0}{a\pi} = 0 \qquad \bar{y} = \frac{2a^2}{a\pi} = \frac{2a}{\pi} \tag{285}$$

2.5 double integrals

Let $\mathcal{R}$ be a region in $\mathbb{R}^2$ and let $f(x, y)$ be a function defined on $\mathcal{R}$. We want to define the integral of f over $\mathcal{R}$ denoted by

$$\int_{\mathcal{R}} f(x,y)dA \quad \text{or} \quad \int_{\mathcal{R}} f(x,y)dxdy \quad \text{or} \quad \int\int_{\mathcal{R}} f(x,y)dxdy \tag{286}$$

It is supposed to be the sum of the values of the function weighted by area.

To define it put a rectangular grid over the region (see figure 14) and suppose the rectangles are enumerated by some index i. The i^{th} rectangle will have some dimensions $\Delta x_i, \Delta y_i$. Let $\Delta A_i = \Delta x_i \Delta y_i$ be the area of the i^{th} rectangle. Also let (x_i^*, y_i^*) be any point in the i^{th} rectangle. An approximation to what we want is the Riemann sum

$$\sum_i f(x_i^*, y_i^*)\Delta A_i \tag{287}$$

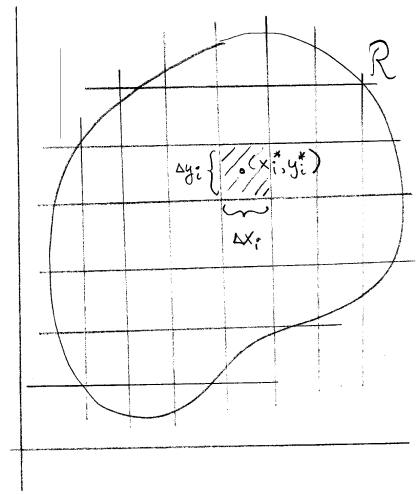


Figure 14: double integral

If these expressions approach a definite number as the grid becomes fine then this is the integral we want. The fineness of the grid can be measured by

$$h = \max_i \{\Delta x_i, \Delta y_i\} \tag{288}$$

Here is an exact definition of the integral.

definition: If there is a number I so that for any $\epsilon > 0$ there is a $\delta > 0$ such that for any grid over $\mathcal{R}$ with $h < \delta$ and any choice of points (x_i^*, y_i^*) in the grid we have

$$|\sum_i f(x_i^*, y_i^*)\Delta A_i - I| < \epsilon \tag{289}$$

then f is integrable over $\mathcal{R}$ and we define

$$\int_{\mathcal{R}} f(x,y)dA = I \tag{290}$$

For short one can write this as

$$\int_{\mathcal{R}} f(x,y)dA = \lim_{h\to 0} \sum_i f(x_i^*, y_i^*)\Delta A_i \tag{291}$$

although it is not an ordinary limit since the right side is not a function of h.

One can show:

Theorem 12 *Continuous functions are integrable.*

Here are some applications of double integrals:

1. With $f = 1$
$$\int_{\mathcal{R}} dA = \text{ area of } \mathcal{R} \tag{292}$$

2. If $\mathcal{R}$ represents a thin plate and $f(x, y)$ is the density of the plate (mass per unit area) then $f(x_i^*, y_i^*)\Delta A_i$ is the approximate mass of the i^{th} rectangle and so
$$\int_{\mathcal{R}} f(x, y)dA = \text{ total mass of plate} \tag{293}$$

3. If $f(x, y) \geq 0$ then $f(x_i^*, y_i^*)\Delta A_i$ is the approximate volume of the column above the i^{th} rectangle and under the graph and so
$$\int_{\mathcal{R}} f(x, y)dA = \text{ volume under the graph of } z = f(x, y) \text{ above } \mathcal{R} \tag{294}$$

Here are some properties of double integrals:

1. For any two functions f$_1, f_2$ on $\mathcal{R}$
$$\int_{\mathcal{R}} (f_1 + f_2)dA = \int_{\mathcal{R}} f_1 dA + \int_{\mathcal{R}} f_2 dA$$

2. If α is a constant
$$\int_{\mathcal{R}} \alpha f dA = \alpha \int_{\mathcal{R}} f dA$$

3. If $\mathcal{R}$ can be written as a disjoint union $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$ then
$$\int_{\mathcal{R}} f dA = \int_{\mathcal{R}_1} f dA + \int_{\mathcal{R}_2} f dA$$

To compute double integrals one writes them as iterated integrals in one variable and then uses the fundamental theorem of calculus.

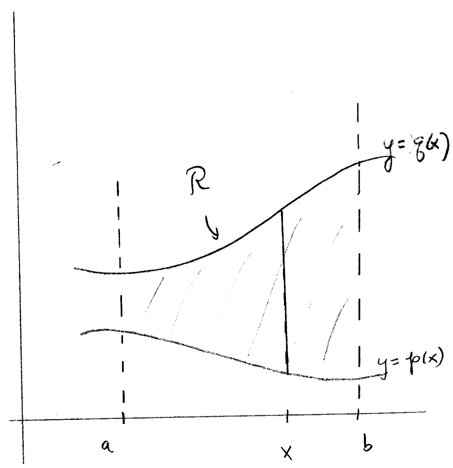


Figure 15: iterated integral

Theorem 13 *Suppose the region* $\mathcal{R}$ *has the form*

$$\mathcal{R} = \{(x,y) : a \leq x \leq b,\ p(x) \leq y \leq q(x)\} \tag{295}$$

for some functions p, q. *(See figure 15). Then*

$$\int_{\mathcal{R}} f(x,y)dA = \int_a^b \Big(\int_{p(x)}^{q(x)} f(x,y)dy \Big) dx \tag{296}$$

This says fix x and integrate over the y values in the region for this value of x. This gives you a function of x which you integrate over the x values for the region.

example: Suppose the region $\mathcal{R}$ below the graph of $y = -x^2 + 1$ in the first quadrant. Thus $\mathcal{R}$ is defined by $0 \leq x \leq 1$ and $0 \leq y \leq -x^2 + 1$. The problem is to find $\int_{\mathcal{R}} xdA$.

We compute

$$\begin{aligned} \int_{\mathcal{R}} xdA = \int_0^1 \left(\int_0^{-x^2+1} xdy \right) dx = \int_0^1 [xy]_{y=0}^{y=-x^2+1} dx \\ = \int_0^1 (-x^3 + x) dx = -\frac{1}{4} + \frac{1}{2} = \frac{1}{4} \end{aligned} \tag{297}$$

Alternatively $\mathcal{R}$ can be regarded as the region $0 \leq y \leq 1$ and $0 \leq x \leq \sqrt{1-y}$ (draw a picture). Then we can do the x integral first:

$$\begin{aligned}\int_{\mathcal{R}} x dA &= \int_0^1 \left(\int_0^{\sqrt{1-y}} x dx \right) dy = \int_0^1 \left[\frac{x^2}{2} \right]_{x=0}^{x=\sqrt{1-y}} dy \\ &= \int_0^1 \frac{1-y}{2} dy = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}\end{aligned} \tag{298}$$

2.6 triple integrals

Now let $\mathcal{R}$ be a region in $\mathbb{R}^3$ and let f be a function defined on $\mathcal{R}$. We want to define the integral of f over $\mathcal{R}$ which will be denoted

$$\int_{\mathcal{R}} f(x,y,z) dV \qquad \text{or} \qquad \int_{\mathcal{R}} f(x,y,z) dxdydz \tag{299}$$

To define it divide the region $\mathcal{R}$ up into many small rectangular boxes. Suppose the i^{th} box has dimensions $\Delta x_i, \Delta y_i, \Delta z_i$ and volume $\Delta V_i = \Delta x_i \Delta y_i \Delta z_i$ and let (x_i^*, y_i^*, z_i^*) be an arbitrary point in the i^{th} box. Also let

$$h = \max_i \{\Delta x_i, \Delta y_i, \Delta z_i\} \tag{300}$$

be the largest dimension in the grid. Then we define

$$\int_{\mathcal{R}} f(x,y,z) dV = \lim_{h \to 0} \sum_i f(x_i^*, y_i^*, z_i^*) \Delta V_i \tag{301}$$

If $f = 1$ then $\int_{\mathcal{R}} dV$ is interpreted as the volume of $\mathcal{R}$. Another application is that $\mathcal{R}$ could represent a solid object. If $f(x,y,z)$ is the density at the point (x,y,z) (mass per unit volume) then $\int_{\mathcal{R}} f(x,y,z) dV$ is the total mass of the object.

Suppose the region $\mathcal{R}$ is the region between the graphs of $z = \phi(x,y)$ and $z = \psi(x,y)$ with (x,y) restricted to some plane region E (see figure 16). Then we can write the triple integral as a single integral followed by a double integral:

$$\int_{\mathcal{R}} f(x,y,z) dV = \int_E \left(\int_{\phi(x,y)}^{\psi(x,y)} f(x,y,z) dz \right) dA \tag{302}$$

If in addition the plane region E is the region between two curves $y = p(x)$ and $y = q(x)$ with $a \leq x \leq b$ then the double integral can be written as an iterated in integral and we have

$$\int_{\mathcal{R}} f(x,y,z) dV = \int_a^b \left(\int_{p(x)}^{q(x)} \left(\int_{\phi(x,y)}^{\psi(x,y)} f(x,y,z) dz \right) dy \right) dx \tag{303}$$

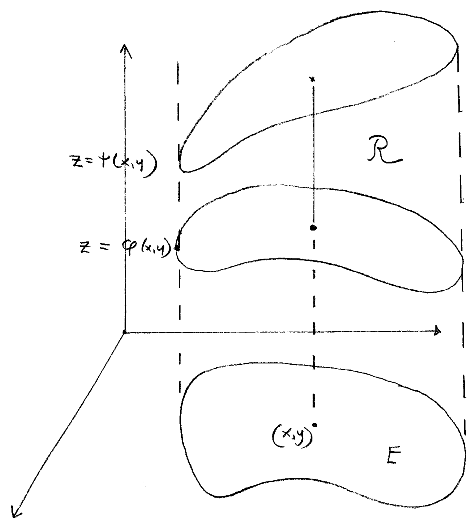


Figure 16:

example: Suppose we are given the problem of finding the volume between the paraboloid $z = 2 - x^2 - y^2$ and the plane $z = 1$.

These surfaces intersect when $x^2 + y^2 = 1$. The problem must be refering to the region with $x^2 + y^2 \leq 1$ since the region with x$^2 + y^2 \geq 1$ is infinite. Thus we want to find the volume of the region $\mathcal{R}$ below $z = 2-x^2-y^2$ and above $z = 1$ with $x^2+y^2 \leq 1$. It is

$$\begin{aligned}
\int_{\mathcal{R}} dV &= \int_{x^2+y^2\leq 1} \left(\int_1^{2-x^2-y^2} dz \right) dA \\
&= \int_{x^2+y^2\leq 1} (1 - x^2 - y^2) dA \\
&= \int_{-1}^{1} \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (1 - x^2 - y^2) dy \right) dx \\
&= \int_{-1}^{1} \frac{4}{3} (1 - x^2)^{3/2} dx \\
&= \frac{\pi}{2}
\end{aligned} \tag{304}$$

The last integral is left as an exercise. (An alternative is to evaluate the integral $\int_{x^2+y^2\leq 1}(1-x^2-y^2)dA$ in polar coordinates, a topic we take up later.)

example: Let $\mathcal{R}$ be the region bounded by the planes $x=0, y=0, z=0$ and $x+y+z=1$ and suppose we want to write $\int_{\mathcal{R}} x\ dV$ as an interated integral.

The intersection of $\mathcal{R}$ with the plane $z=0$ is the region E bounded by the lines $x=0, y=0, x+y=1$. In fact the region $\mathcal{R}$ lies between $z=1-x-y$ and $z=0$ and above E. (draw a picture). Thus we have

$$\int_{\mathcal{R}} x\ dV = \int_E \left(\int_0^{1-x-y} xdz \right) dA \tag{305}$$

But since E lies between $y=0$ and $y=1-x$ with $0\leq x\leq 1$ this can be expressed as

$$\int_{\mathcal{R}} x\ dV = \int_0^1 \left(\int_0^{1-x} \left(\int_0^{1-x-y} xdz \right) dy \right) dx \tag{306}$$

The evaluation is left as an exercise.

2.7 parametrized surfaces

Consider a function from $\mathcal{R}\subset\mathbb{R}^2$ to $\mathbb{R}^3$ which we write as

$$x = x(u,v) \qquad y = y(u,v) \qquad z = z(u,v) \tag{307}$$

The range of this function is a surface $\mathcal{S}$ and the function is called a *parametrization* of the surface. (A surface has many possible parametrizations, but we pick one). The function can also be written

$$\mathbf{R}(u,v) = x(u,v)\mathbf{i} + y(u,v)\mathbf{j} + z(u,v)\mathbf{k} \tag{308}$$

example: Consider the function

$$\begin{aligned} x =& a\sin\phi\cos\theta \\ y =& a\sin\phi\sin\theta \\ z =& a\cos\phi \end{aligned} \tag{309}$$

with $0<\phi<\pi$ and $0<\theta<2\pi$. Then $\mathcal{S}$ is the surface of a sphere of radius a, and it is parametrized by spherical coordinates.

example: Suppose $\mathcal{S}$ is the graph of a function $z=\phi(x,y)$ with $(x,y)\in\mathcal{R}$. Then $\mathcal{S}$ can be parametrized by

$$x = u \qquad y = v \qquad z = \phi(u,v) \tag{310}$$

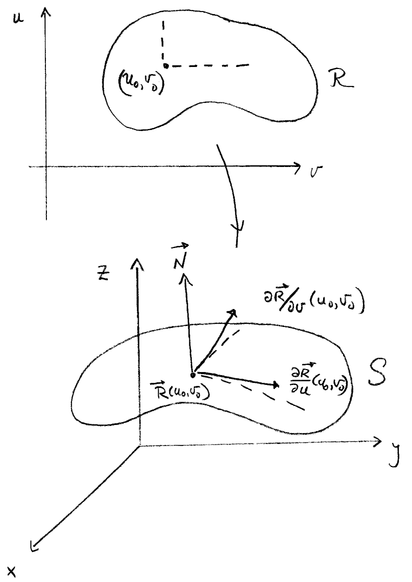


Figure 17:

with $(u, v) \in \mathcal{R}$. This can also be written

$$x = x \qquad y = y \qquad z = \phi(x, y) \tag{311}$$

with $(x, y) \in \mathcal{R}$.

Now suppose $\mathcal{S}$ is a surface parametrized by $\mathbf{R}(u, v)$. At (u_0, v_0) if we vary u we get a u-line in $\mathcal{S}$. Then

$$\frac{\partial \mathbf{R}}{\partial u}(u_0, v_0) = \text{ tangent vector to } u\text{-line through } \mathbf{R}(u_0, v_0)$$
$$\frac{\partial \mathbf{R}}{\partial v}(u_0, v_0) = \text{ tangent vector to } v\text{-line through } \mathbf{R}(u_0, v_0)$$

Together these two tangent vectors determine the tangent plane to the surface at the point

$$\mathbf{R}_0 = \mathbf{R}(u_0, v_0) \tag{312}$$

A normal to this tangent plane is (see figure 17)

$$\mathbf{N} = \frac{\partial \mathbf{R}}{\partial u}(u_0, v_0) \times \frac{\partial \mathbf{R}}{\partial v}(u_0, v_0) \tag{313}$$

With this $\mathbf{N}$ the equation of the tangent plane to the surface $\mathcal{S}$ at $\mathbf{R}_0$ is

$$\mathbf{N} \cdot (\mathbf{R} - \mathbf{R}_0) = 0 \tag{314}$$

example: Suppose we want to find the tangent plane to the surface

$$x = u + v \qquad y = u - v \qquad z = 2uv \tag{315}$$

at the point $u = 1, v = 1$. In this case

$$\mathbf{R}(u, v) = (u + v)\mathbf{i} + (u - v)\mathbf{j} + 2uv\mathbf{k} \tag{316}$$

and the point is

$$\mathbf{R}_0 = \mathbf{R}(1, 1) = 2\mathbf{i} + 2\mathbf{k} \tag{317}$$

At this point

$$\begin{aligned} \frac{\partial \mathbf{R}}{\partial u} &= \mathbf{i} + \mathbf{j} + 2v\mathbf{k} = \mathbf{i} + \mathbf{j} + 2\mathbf{k} \\ \frac{\partial \mathbf{R}}{\partial v} &= \mathbf{i} - \mathbf{j} + 2u\mathbf{k} = \mathbf{i} - \mathbf{j} + 2\mathbf{k} \end{aligned} \tag{318}$$

so the normal is

$$\mathbf{N} = \frac{\partial \mathbf{R}}{\partial u} \times \frac{\partial \mathbf{R}}{\partial v} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 2 \\ 1 & -1 & 2 \end{pmatrix} = 4\mathbf{i} - 2\mathbf{k} \tag{319}$$

Also we have

$$\mathbf{R} - \mathbf{R}_0 = (x - 2)\mathbf{i} + y\mathbf{j} + (z - 2)\mathbf{k} \tag{320}$$

The equation of the tangent plane at this point is $\mathbf{N} \cdot (\mathbf{R} - \mathbf{R}_0) = 0$ which says

$$4(x - 2) - 2(z - 2) = 0 \tag{321}$$

which can also be written $z = 2x - 2$.

example: Suppose our surface is the graph of a function $z = \phi(x, y)$. Then it can be parametrized by

$$\mathbf{R}(x, y) = x\mathbf{i} + y\mathbf{j} + \phi(x, y)\mathbf{k} \tag{322}$$

We want to find the normal and the tangent plane to the graph at (x_0, y_0) This is the point

$$\mathbf{R}_0 = \mathbf{R}(x_0, y_0) = x_0\mathbf{i} + y_0\mathbf{j} + \phi(x_0, y_0)\mathbf{k} \tag{323}$$

The derivatives at (x_0, y_0) are

$$\begin{aligned} \frac{\partial \mathbf{R}}{\partial x} &= \mathbf{i} + \phi_x(x_0, y_0)\mathbf{k} \\ \frac{\partial \mathbf{R}}{\partial y} &= \mathbf{j} + \phi_y(x_0, y_0)\mathbf{k} \end{aligned} \tag{324}$$

and so the normal is

$$\mathbf{N} = \frac{\partial \mathbf{R}}{\partial x} \times \frac{\partial \mathbf{R}}{\partial y} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \phi_x(x_0, y_0) \\ 0 & 1 & \phi_y(x_0, y_0) \end{pmatrix} \tag{325}$$

which says

$$\mathbf{N} = -\phi_x(x_0, y_0)\mathbf{i} - \phi_y(x_0, y_0)\mathbf{j} + \mathbf{k} \tag{326}$$

The equation of the tangent plane $\mathbf{N} \cdot (\mathbf{R} - \mathbf{R}_0) = 0$ is then

$$-\phi_x(x_0, y_0)(x - x_0) - \phi_y(x_0, y_0)(y - y_0) + (z - \phi(x_0, y_0)) = 0 \tag{327}$$

This can also be written

$$z = \phi(x_0, y_0) + \phi_x(x_0, y_0)(x - x_0) + \phi_y(x_0, y_0)(y - y_0) \tag{328}$$

which agrees with our earlier definition of the tangent plane.

2.8 surface area

Let $\mathcal{S}$ be a surface parametrized by a function $\mathbf{R}(u, v)$ with $(u, v) \in \mathcal{R}$. We assume the function is one-to-one so it only covers $\mathcal{S}$ once. We want to define the area of $\mathcal{S}$.

To do so we divide up $\mathcal{R}$ into a fine rectangular grid (see figure 18). The lines of the grid are mapped to lines in the surface and this divides up the surface into little pieces (no longer rectangles). Suppose the i^{th} rectangle has lower left corner (u_i, v_i) and dimensions $\Delta u_i, \Delta v_i$. The image of this rectangle is a patch with corners

$$\mathbf{R}(u_i, v_i), \quad \mathbf{R}(u_i + \Delta u_i, v_i), \quad \mathbf{R}(u_i, v_i + \Delta v_i), \quad \mathbf{R}(u_i + \Delta u_i, v_i + \Delta v_i)$$

The area of this patch is approximated as the area of the parallelogram spanned by

$$\begin{aligned} \mathbf{a}_i =& \mathbf{R}(u_i + \Delta u_i, v_i) - \mathbf{R}(u_i, v_i) \\ \mathbf{b}_i =& \mathbf{R}(u_i, v_i + \Delta v_i) - \mathbf{R}(u_i, v_i) \end{aligned} \tag{329}$$

This area is

$$\Delta \sigma_i = |\mathbf{a}_i \times \mathbf{b}_i| \tag{330}$$

However since Δu_i and Δv_i are assumed small we have the approximations

$$\begin{aligned} \mathbf{a}_i \approx& \frac{\partial \mathbf{R}}{\partial u}(u_i, v_i)\Delta u_i \\ \mathbf{b}_i \approx& \frac{\partial \mathbf{R}}{\partial v}(u_i, v_i)\Delta v_i \end{aligned} \tag{331}$$

Hence

$$\Delta \sigma_i \approx \left| \frac{\partial \mathbf{R}}{\partial u}(u_i, v_i) \times \frac{\partial \mathbf{R}}{\partial v}(u_i, v_i) \right| \Delta u_i \Delta v_i \tag{332}$$

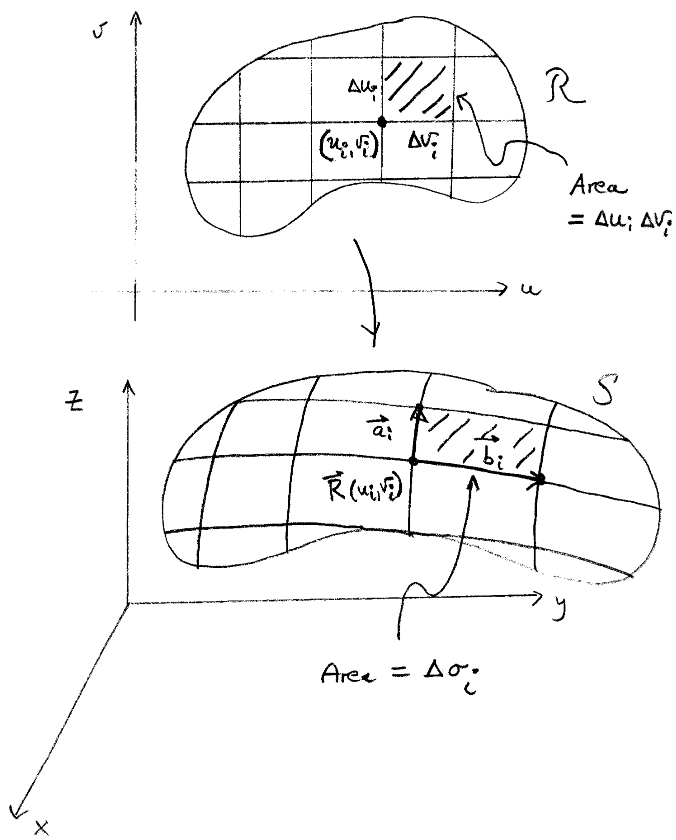


Figure 18:

Now if $h = \max_i\{\Delta u_i, \Delta v_i\}$ is the maximum dimension in the grid we define

$$\begin{aligned}\text{Area of } \mathcal{S} &= \lim_{h\to 0}\sum_i \Delta\sigma_i \\ &= \lim_{h\to 0}\sum_i \left|\frac{\partial \mathbf{R}}{\partial u}(u_i, v_i) \times \frac{\partial \mathbf{R}}{\partial v}(u_i, v_i)\right| \Delta u_i \Delta v_i \qquad (333) \\ &= \int_{\mathcal{R}} \left|\frac{\partial \mathbf{R}}{\partial u} \times \frac{\partial \mathbf{R}}{\partial v}\right| du dv\end{aligned}$$

This definition turns out to be independent of the particular parametrization we have chosen.

example: Find the area of a spherical cap of radius a and angle α. In spherical coordinates this is the surface described

$$r = a \qquad 0 \leq \phi \leq \alpha \qquad 0 \leq \theta \leq 2\pi$$

We parametrize with spherical coordinates and take

$$\mathbf{R}(\phi, \theta) = a \sin\phi \cos\theta \mathbf{i} + a \sin\phi \sin\theta \mathbf{j} + a \cos\phi \mathbf{k} \tag{334}$$

with $0 \leq \phi \leq \alpha, 0 \leq \theta \leq 2\pi$. Then we compute

$$\begin{aligned} \frac{\partial \mathbf{R}}{\partial \phi} \times \frac{\partial \mathbf{R}}{\partial \theta} =& \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a\cos\phi\cos\theta & a\cos\phi\sin\theta & -a\sin\phi \\ -a\sin\phi\sin\theta & a\sin\phi\cos\theta & 0 \end{pmatrix} \\ =& a^2 \sin^2\phi \cos\theta \mathbf{i} + a^2 \sin^2\phi \sin\theta \mathbf{j} + a^2 \cos\phi \sin\phi \mathbf{k} \end{aligned} \tag{335}$$

Then

$$\begin{aligned} \left| \frac{\partial \mathbf{R}}{\partial \phi} \times \frac{\partial \mathbf{R}}{\partial \theta} \right| =& \sqrt{a^4 \sin^4\phi(\cos^2\theta + \sin^2\theta) + a^4 \cos^2\phi \sin^2\phi} \\ =& a^2 \sin\phi \sqrt{\sin^2\phi + \cos^2\phi} \\ =& a^2 \sin\phi \end{aligned} \tag{336}$$

and the area is

$$\begin{aligned} \text{Area } =& \int_0^{2\pi} \int_0^{\alpha} \left| \frac{\partial \mathbf{R}}{\partial \phi} \times \frac{\partial \mathbf{R}}{\partial \theta} \right| d\phi \; d\theta \\ =& \int_0^{2\pi} \int_0^{\alpha} a^2 \sin\phi \; d\phi \; d\theta \\ =& a^2 \Big(\int_0^{2\pi} d\theta \Big) \Big(\int_0^{\alpha} \sin\phi \; d\phi \; \Big) \\ =& 2\pi a^2 (1 - \cos\alpha) \end{aligned} \tag{337}$$

Note that if $\alpha = \pi$ the area is $4\pi a^2$ which is what we expect for the whole sphere.

2.9 surface integrals

We continue to consider a surface $\mathcal{S}$ parametrized by a function $\mathbf{R}(u, v)$ with $(u, v) \in \mathcal{R}$. Also let $f(\mathbf{R}) = f(x, y, z)$ be a function defined on $\mathcal{S}$ (and possibly elsewhere in $\mathbb{R}^3$). We want to define the integral of f over $\mathcal{S}$ which will be written $\int_{\mathcal{S}} f(\mathbf{R}) d\sigma$.

To define it we again divide up the parameter space into a rectangular grid. We also let (u_i, v_i) be the corner point in the i^{th} rectangle. (see figure 18 again). Then we define

$$\begin{aligned}\int_{\mathcal{S}} f(\mathbf{R})d\sigma &= \lim_{h\to 0}\sum_i f\big(\mathbf{R}(u_i,v_i)\big)\Delta\sigma_i \\ &= \lim_{h\to 0}\sum_i f\big(\mathbf{R}(u_i,v_i)\big)\left|\frac{\partial \mathbf{R}}{\partial u}(u_i,v_i)\times\frac{\partial \mathbf{R}}{\partial v}(u_i,v_i)\right|\Delta u_i\Delta v_i \\ &= \int_{\mathcal{R}} f\big(\mathbf{R}(u,v)\big)\left|\frac{\partial \mathbf{R}}{\partial u}\times\frac{\partial \mathbf{R}}{\partial v}\right| du\ dv\end{aligned} \tag{338}$$

For short one just has to remember

$$d\sigma = \left|\frac{\partial \mathbf{R}}{\partial u}\times\frac{\partial \mathbf{R}}{\partial v}\right| du\ dv \tag{339}$$

A special case is with $f(\mathbf{R}) = 1$ which gives

$$\int_{\mathcal{S}} d\sigma = \int_{\mathcal{R}}\left|\frac{\partial \mathbf{R}}{\partial u}\times\frac{\partial \mathbf{R}}{\partial v}\right| du\ dv = \text{Area of } \mathcal{S} \tag{340}$$

example: Let $\mathcal{S}$ represent a thin hemispherical shell of uniform density which has radius a and is centered on the orgin. We want to find the z-component of the center of mass. Since it is a thin shell it is reasonable to represent in terms of surface integrals and we take the definition

$$\bar{z} = \frac{\int_{\mathcal{S}} z\ d\sigma}{\int_{\mathcal{S}} d\sigma} \tag{341}$$

The hemisphere is parametrized as before by $x = a\sin\phi\cos\theta$, $y = a\sin\phi\sin\theta$, and $z = a\cos\phi$ with $0 \le \phi \le \pi/2$ and $0 \le \theta \le 2\pi$. We also have as before

$$d\sigma = \left|\frac{\partial \mathbf{R}}{\partial \phi}\times\frac{\partial \mathbf{R}}{\partial \theta}\right| d\phi\ d\theta = a^2\sin\phi\ d\phi\ d\theta \tag{342}$$

Then we can compute

$$\begin{aligned}\int_{\mathcal{S}} z\ d\sigma &= \int_0^{2\pi}\int_0^{\pi/2}(a\cos\phi)a^2\sin\phi\ d\phi\ d\theta \\ &= a^3\Big(\int_0^{2\pi} d\theta\Big)\Big(\int_0^{\pi/2}\cos\phi\sin\phi\ d\phi\ \Big) \\ &= a^3\cdot 2\pi\cdot\frac{1}{2} \\ &= \pi a^3\end{aligned} \tag{343}$$

From our earlier calculation of area we have

$$\int_{\mathcal{S}} d\sigma = 2\pi a^2 \tag{344}$$

Thus

$$\bar{z} = \frac{\pi a^3}{2\pi a^2} = \frac{a}{2} \tag{345}$$

example: Suppose that the surface $\mathcal{S}$ is the graph of a function $z = \phi(x, y)$ with $(x, y) \in \mathcal{R}$. As noted previously we can parametrize $\mathcal{S}$ by

$$\mathbf{R}(x, y) = x\mathbf{i} + y\mathbf{j} + \phi(x, y)\mathbf{k} \qquad (x, y) \in \mathcal{R} \tag{346}$$

We also computed earlier

$$\frac{\partial \mathbf{R}}{\partial x} \times \frac{\partial \mathbf{R}}{\partial y} = -\phi_x \mathbf{i} - \phi_y \mathbf{j} + \mathbf{k} \tag{347}$$

Therefore

$$d\sigma = \left| \frac{\partial \mathbf{R}}{\partial x} \times \frac{\partial \mathbf{R}}{\partial y} \right| \, dxdy = \sqrt{1 + \phi_x^2 + \phi_y^2} \, dxdy \tag{348}$$

and so

$$\int_{\mathcal{S}} f(x, y, z) d\sigma = \int_{\mathcal{R}} f(x, y, \phi(x, y)) \sqrt{1 + \phi_x^2 + \phi_y^2} \, dxdy \tag{349}$$

In particular

$$\text{Area of } \mathcal{S} = \int_{\mathcal{S}} d\sigma = \int_{\mathcal{R}} \sqrt{1 + \phi_x^2 + \phi_y^2} \, dxdy \tag{350}$$

2.10 change of variables in $\mathbb{R}^2$

Consider a special case of the surface integral in which the surface $\mathcal{S}$ lies in the xy plane. Then the parametrization has the form $x = x(u, v), y = y(u, v), z = 0$ for $(u, v) \in \mathcal{R}$. In vector form

$$\mathbf{R}(u, v) = x(u, v)\mathbf{i} + y(u, v)\mathbf{j} \tag{351}$$

In this case

$$\begin{aligned} \frac{\partial \mathbf{R}}{\partial u} \times \frac{\partial \mathbf{R}}{\partial v} &= \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_u & y_u & 0 \\ x_v & y_v & 0 \end{pmatrix} = \det \begin{pmatrix} x_u & y_u \\ x_v & y_v \end{pmatrix} \mathbf{k} \\ &= \det \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \mathbf{k} = \frac{\partial(x, y)}{\partial(u, v)} \mathbf{k} \end{aligned} \tag{352}$$

and so

$$d\sigma = \left| \frac{\partial \mathbf{R}}{\partial u} \times \frac{\partial \mathbf{R}}{\partial v} \right| dudv = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv \tag{353}$$

Then our surface integral is evaluated as

$$\int_{\mathcal{S}} f(x, y) d\sigma = \int_{\mathcal{R}} f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv \tag{354}$$

However parametrizing $\mathcal{S}$ as the graph of the function $z = 0$ we have

$$d\sigma = \sqrt{1 + (\partial z/\partial x)^2 + (\partial z/\partial y)^2} dxdy = dxdy = dA \tag{355}$$

Thus we have demonstrated the following change of variables formula:

Theorem 14 *Let a region $\mathcal{S} \subset \mathbb{R}^2$ be the image of a region $\mathcal{R} \subset \mathbb{R}^2$ under a differentiable invertible function $x = x(u, v), y = y(u, v)$. Then*

$$\int_{\mathcal{S}} f(x, y) dA = \int_{\mathcal{R}} f(x(u, v), y(u, v)) \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv \tag{356}$$

One can think of (u, v) as new coordinates for the region $\mathcal{S}$. Then a short version of the theorem is

$$dA = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| dudv \tag{357}$$

expressing the area element in the new coordinates.

As a special case consider polar coordinates $x = r\cos\theta, y = r\sin\theta$. Then

$$\frac{\partial(x, y)}{\partial(r, \theta)} = \det \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix} = \det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r \tag{358}$$

and so $dA = rdrd\theta$. The change of variables formula is

$$\int_{\mathcal{S}} f(x, y)\ dA = \int_{\mathcal{R}} f(r\cos\theta, r\sin\theta)\ rdrd\theta \tag{359}$$

Here $\mathcal{R}$ is all points (r, θ) such that $(x, y) = (r\cos\theta, r\sin\theta)$ is in $\mathcal{S}$. Thus $\mathcal{R}$ is just $\mathcal{S}$ described in polar coordinates.

example: Suppose $\mathcal{S}$ is the half-disc $x^2 + y^2 \leq 4, y > 0$ and we want to evaluate

$$\int_{\mathcal{S}} (3 - x^2 - y^2) dA \tag{360}$$

In polar coordinates $\mathcal{S}$ becomes the region $\mathcal{R}$ defined by $0 < r < 2, 0 < \theta < \pi$. Thus

$$\begin{aligned} \int_{\mathcal{S}} (3 - x^2 - y^2) dA &= \int_{\mathcal{R}} (3 - r^2) rdrd\theta \\ &= \int_0^\pi \int_0^2 (3r - r^3) drd\theta \\ &= \pi \left[\frac{3r^2}{2} - \frac{r^4}{4} \right]_0^2 \\ &= 2\pi \end{aligned} \tag{361}$$

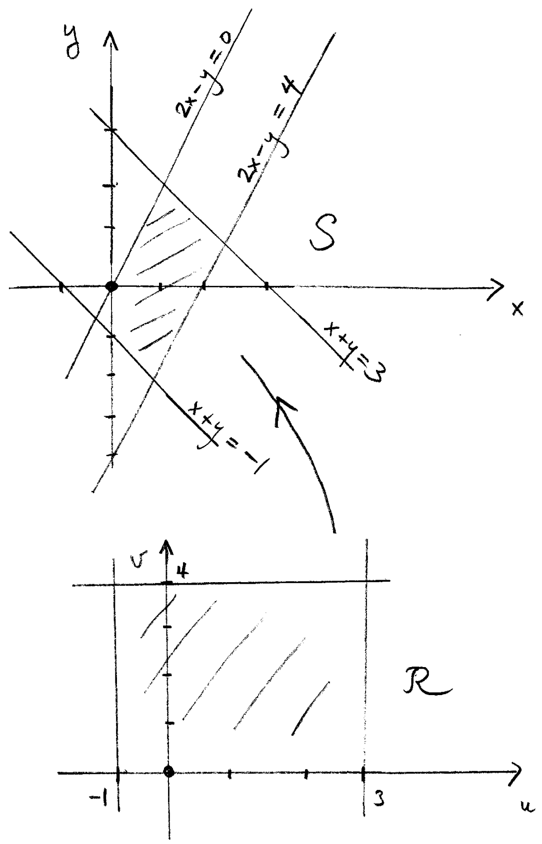


Figure 19:

example: Let $\mathcal{S}$ be the region bounded by the lines $x + y = -1, x + y = 3, 2x - y = 0, 2x - y = 4$. (see figure 19) We want to evaluate the integral

$$\int_{\mathcal{S}} (x + y) dA \tag{362}$$

We make a change of variables suggested by the boundary lines and set

$$\begin{aligned} u =& x + y \\ v =& 2x - y \end{aligned} \tag{363}$$

The lines $x + y = c$ are sent to the line $u = c$ and the lines $2x - y = c$ are sent to the lines $v = c$. Thus the region $\mathcal{S}$ is sent to the region $\mathcal{R}$ bounded by the lines $u = -1, u = 3, v = 0, v = 4$.

For the change of variables formula we need the inverse function

$$\begin{aligned} x =& \frac{u+v}{3} \\ y =& \frac{2u-v}{3} \end{aligned} \tag{364}$$

which sends $\mathcal{R}$ back to $\mathcal{S}$.

We compute

$$\frac{\partial(x,y)}{\partial(u,v)} = \det \begin{pmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{pmatrix} = -\frac{1}{3} \tag{365}$$

and then

$$\begin{aligned} \int_{\mathcal{S}} (x+y) dA =& \int_{\mathcal{R}} u \left| \frac{\partial(x,y)}{\partial(u,v)} \right| \, dudv \\ =& \int_0^4 \int_{-1}^3 u \cdot \frac{1}{3} \, dudv \\ =& 4 \left[\frac{u^2}{6} \right]_{-1}^3 = \frac{16}{3} \end{aligned} \tag{366}$$

2.11 change of variables in $\mathbb{R}^3$

In $\mathbb{R}^3$ the change of variables formula is the following:

Theorem 15 *Let a region $\mathcal{V} \subset \mathbb{R}^3$ be the image of a region $\mathcal{R} \subset \mathbb{R}^3$ under a differentiable invertible function $x = x(u, v, w), y = y(u, v, w), z = z(u, v, w)$. Then*

$$\int_{\mathcal{V}} f(x, y, z)dV = \int_{\mathcal{R}} f(x(u, v, w), y(u, v, w), z(u, v, w)) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| dudvdw \tag{367}$$

For short we can write

$$dV = \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| dudvdw \tag{368}$$

Also note that in the special case $f(x, y, z) = 1$ we have

$$\text{Volume of } \mathcal{V} = \int_{\mathcal{V}} dV = \int_{\mathcal{R}} \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| dudvdw \tag{369}$$

Proof. Divide up the region $\mathcal{R}$ into a grid of small boxes. The i^{th} box will have a corner (u_i, v_i, w_i) and dimensions $(\Delta u_i, \Delta v_i, \Delta w_i)$. The volume of this box is $\Delta V_i = \Delta u_i \Delta v_i \Delta w_i$.

Let $\Delta V_i'$ be the volume of image of the i^{th} box. The image has corners $x_i = x(u_i, v_i, w_i), y_i = y(u_i, v_i, w_i), z_i = z(u_i, v_i, w_i)$ and curved sides. The volume is approximately the volume of a parallelopiped spanned by vectors $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i$ joining the corners. (See figure 20). Thus

$$\Delta V_i' \approx |\mathbf{a}_i \cdot (\mathbf{b}_i \times \mathbf{c}_i)| \tag{370}$$

If we write the the funtion as

$$\mathbf{R}(u, v, w) = x(u, v, w)\mathbf{i} + y(u, v, w)\mathbf{j} + z(u, v, w)\mathbf{k} \tag{371}$$

then

$$\mathbf{a}_i = \mathbf{R}(u_i + \Delta u_i, v_i w_i) - \mathbf{R}(u_i, v_i w_i) \approx \frac{\partial \mathbf{R}}{\partial u}(u_i, v_i, w_i)\Delta u_i \tag{372}$$

and similarly

$$\mathbf{b}_i \approx \frac{\partial \mathbf{R}}{\partial v}(u_i, v_i, w_i)\Delta v_i \qquad \mathbf{c}_i \approx \frac{\partial \mathbf{R}}{\partial w}(u_i, v_i, w_i)\Delta w_i$$

Therefore

$$\begin{aligned} \Delta V_i' &\approx \left| \frac{\partial \mathbf{R}}{\partial u} \cdot \left(\frac{\partial \mathbf{R}}{\partial v} \times \frac{\partial \mathbf{R}}{\partial w} \right) \right| \Delta u_i \Delta v_i \Delta w_i \\ &= \left| \det \begin{pmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \\ x_w & y_w & z_w \end{pmatrix} \right| \Delta V_i = \left| \det \begin{pmatrix} x_u & x_v & x_w \\ y_u & y_v & y_w \\ z_u & z_v & z_w \end{pmatrix} \right| \Delta V_i \\ &= \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| \Delta V_i \end{aligned} \tag{374}$$

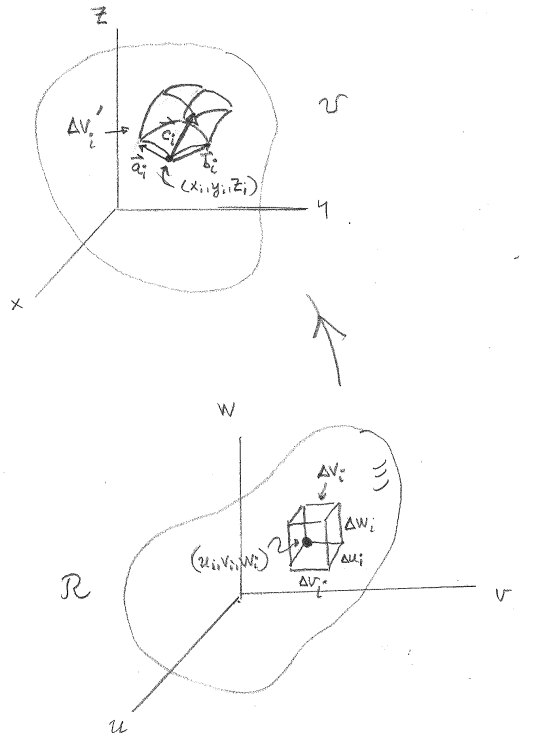


Figure 20:

So the Jacobian determinant tells how volumes increase.

Then we have

$$\begin{aligned}&\sum_i f(x_i, y_i, z_i)\Delta V_i' \\ =&\sum_i f(x(u_i, v_i, w_i), y(u_i, v_i, w_i), z(u_i, v_i, w_i)) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)}(u_i, v_i, w_i) \right| \Delta V_i\end{aligned} \tag{375}$$

Now taking the limit as the grid size goes to zero we obtain the result (although the expression on the left is not the standard Riemann integral).

special cases:

1. cylindrical coordinates:

$$\begin{aligned} x =& r\cos\theta \\ y =& r\sin\theta \\ z =& z \end{aligned} \tag{376}$$

In this case

$$dV = \left|\frac{\partial(x,y,z)}{\partial(r,\theta,z)}\right| drd\theta dz = rdrd\theta dz \tag{377}$$

and

$$\int_{\mathcal{V}} f(x,y,z)dV = \int_{\mathcal{R}} f(r\cos\theta, r\sin\theta, z)rdrd\theta dz \tag{378}$$

where $\mathcal{R}$ is the region $\mathcal{V}$ described in cylindrical coordinates.

2. spherical coordinates:

$$\begin{aligned} x =& \rho\sin\phi\cos\theta \\ y =& \rho\sin\phi\sin\theta \\ z =& \rho\cos\phi \end{aligned} \tag{379}$$

In this case (check it!)

$$dV = \left|\frac{\partial(x,y,z)}{\partial(\rho,\phi,\theta)}\right| d\rho d\phi d\theta = \rho^2 \sin\phi d\rho d\phi d\theta \tag{380}$$

and

$$\int_{\mathcal{V}} f(x,y,z)dV = \int_{\mathcal{R}} f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\rho^2\sin\phi d\rho d\phi d\theta \tag{381}$$

where $\mathcal{R}$ is the region $\mathcal{V}$ described in spherical coordinates.

example: Suppose we want to find the volume of the quarter-cone in a sphere $\mathcal{V}$ which is described in spherical coordinates by $0 < \rho < 1$, $0 < \phi < \pi/4$, $0 < \theta < \pi/2$. We compute

$$\begin{aligned} \text{Volume} =& \int_{\mathcal{V}} dV \\ =& \int_0^{\pi/2}\int_0^{\pi/4}\int_0^1 \rho^2 \sin\phi \; d\rho \; d\phi \; d\theta \\ =& \Big(\int_0^1 \rho^2 d\rho\Big)\Big(\int_0^{\pi/4} \sin\phi \; d\phi \;\Big)\Big(\int_0^{\pi/2} d\theta\Big) \\ =& \frac{1}{3}\cdot\big[-\cos\phi\big]_0^{\pi/4}\cdot\pi/2 \\ =& \frac{\pi}{6}\left(1-\frac{1}{\sqrt{2}}\right) \end{aligned} \tag{382}$$

2.12 derivatives in $\mathbb{R}^3$

In $\mathbb{R}^3$ we continue to use the notation

$$\mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = (x, y, z) \tag{383}$$

A *scalar* is a function from $\mathbb{R}^3$ to $\mathbb{R}$ and has the form

$$u = u(\mathbf{R}) = u(x, y, z) \tag{384}$$

A scalar can be drawn (not very well) by shading points in $\mathbb{R}^3$ proportional to the value of u at that point. Examples of quantites that can be represented by scalars are density and temperature.

A *vector field* is function from $\mathbb{R}^3$ to $\mathbb{R}^3$ and has the form

$$\begin{aligned} \mathbf{v} = \mathbf{v}(\mathbf{R}) =& v_1(\mathbf{R})\mathbf{i} + v_2(\mathbf{R})\mathbf{j} + v_3(\mathbf{R})\mathbf{k} \\ =& v_1(x,y,z)\mathbf{i} + v_2(x,y,z)\mathbf{j} + v_3(x,y,z)\mathbf{k} \end{aligned} \tag{385}$$

A vector field can be repesented by drawing a vector $\mathbf{v}(\mathbf{R})$ at the point $\mathbf{R}$ for some representative points $\mathbf{R}$. Examples of quantities that can be represented by vector fields are forces and the velocity of a fluid.

We want to define various derivatives of scalars and vector fields. These are specified with the operator

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{386}$$

which is called *del* or *nabla*. If u is a scalar we define a vector field

$$\textit{gradient} \text{ of } u = \nabla u = \mathbf{i}\frac{\partial u}{\partial x} + \mathbf{j}\frac{\partial u}{\partial y} + \mathbf{k}\frac{\partial u}{\partial z} \tag{387}$$

If $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ is a vector field we define a scalar

$$\textit{divergence} \text{ of } \mathbf{v} = \nabla \cdot \mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z} \tag{388}$$

and also a vector field

$$\begin{aligned} \textit{curl} \text{ of } \mathbf{v} =& \nabla \times \mathbf{v} \\ =& \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ v_1 & v_2 & v_3 \end{pmatrix} \\ =& \mathbf{i}\left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right) + \mathbf{j}\left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right) + \mathbf{k}\left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right) \end{aligned} \tag{389}$$

Finally if u is a scalar we define

$$\textit{Laplacian} \text{ of } u = \Delta u = \nabla \cdot \nabla u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \tag{390}$$

example: If $u = x^2 + xy + y^2 + yz + z^2 + zx$ then

$$\begin{aligned} \nabla u =&(2x + y + z)\mathbf{i} + (2y + x + z)\mathbf{j} + (2z + z + y)\mathbf{k} \\ \Delta u =&\nabla \cdot \nabla u = 2 + 2 + 2 = 6 \end{aligned} \tag{391}$$

example: If $\mathbf{v} = yz^2\mathbf{i} + xz^2\mathbf{j} + (2xyz + z)\mathbf{k}$ then

$$\begin{aligned} \nabla \cdot \mathbf{v} =&2xy + 1 \\ \nabla \times \mathbf{v} =&\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ yz^2 & xz^2 & (2xyz + z) \end{pmatrix} \\ =&(2xz - 2xz)\mathbf{i} + (2zy - 2zy)\mathbf{j} + (z^2 - z^2)\mathbf{k} = 0 \end{aligned} \tag{392}$$

The derivatives satisfy various identities some of which we list. These hold for any scalar u and any vector field $\mathbf{v}$, assuming only they are twice continuously differentiable.

1. $\nabla \times \nabla u = 0$
2. $\nabla \cdot (\nabla \times \mathbf{v}) = 0$
3. $\nabla \cdot (u\mathbf{v}) = \nabla u \cdot \mathbf{v} + u(\nabla \cdot \mathbf{v})$
4. $\nabla \cdot (\mathbf{v} \times \mathbf{w}) = (\nabla \times \mathbf{v}) \cdot \mathbf{w} - \mathbf{v} \cdot (\nabla \times \mathbf{w})$
5. $\nabla \times (u\mathbf{v}) = u(\nabla \times \mathbf{v}) + \nabla u \times \mathbf{v}$
6. $\nabla \times (\nabla \times \mathbf{v}) = \nabla(\nabla \cdot \mathbf{v}) - \Delta \mathbf{v}$ where $\Delta \mathbf{v} = \Delta v_1 \mathbf{i} + \Delta v_2 \mathbf{j} + \Delta v_3 \mathbf{k}$

proof of (1.). We compute

$$\begin{aligned} \nabla \times \nabla u =&\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ \partial u/\partial x & \partial u/\partial y & \partial u/\partial z \end{pmatrix} \\ =&\Big(\frac{\partial^2 u}{\partial y \partial z} - \frac{\partial^2 u}{\partial z \partial y}\Big)\mathbf{i} + \Big(\frac{\partial^2 u}{\partial z \partial x} - \frac{\partial^2 u}{\partial x \partial z}\Big)\mathbf{j} + \Big(\frac{\partial^2 u}{\partial x \partial y} - \frac{\partial^2 u}{\partial y \partial x}\Big)\mathbf{k} \\ =&0 \end{aligned} \tag{393}$$

We also note that the chain rule

$$\begin{aligned} &\frac{d}{dt} u\Big(x(t), y(t), z(t)\Big) \\ =&\frac{\partial u}{\partial x}\Big(x(t), y(t), z(t)\Big)\frac{dx}{dt} + \frac{\partial u}{\partial y}\Big(x(t), y(t), z(t)\Big)\frac{dy}{dt} + \frac{\partial u}{\partial z}\Big(x(t), y(t), z(t)\Big)\frac{dz}{dt} \end{aligned} \tag{394}$$

can be written in a vector notation as

$$\frac{d}{dt}u(\mathbf{R}(t)) = \nabla u(\mathbf{R}(t)) \cdot \frac{d\mathbf{R}}{dt} \tag{395}$$

2.13 gradient

One of our goals is to interpret the gradient, divergence, and curl. Here we give three interpretations of the gradient.

1. $(\nabla u)(\mathbf{R}_0)$ is normal to the level surface u = constant through $\mathbf{R}_0$, (The level surface is all points $\mathbf{R}$ such that $u(\mathbf{R}) = u(\mathbf{R}_0)$).

   To see this let $\mathbf{R}(t)$ be any curve in the surface with $\mathbf{R}(0) = \mathbf{R}_0$. Thus

   $$u(\mathbf{R}(t)) = u(\mathbf{R}_0) \tag{396}$$

   Taking the derivative with respect to t and using the chain rule gives

   $$(\nabla u)(\mathbf{R}(t)) \cdot \mathbf{R}'(t) = 0 \tag{397}$$

   At $t = 0$ this says

   $$(\nabla u)(\mathbf{R}_0) \cdot \mathbf{R}'(0) = 0 \tag{398}$$

   Any tangent vector to the surface at $\mathbf{R}_0$ has the form $\mathbf{R}'(0)$ for some curve through $\mathbf{R}_0$. Thus $(\nabla u)(\mathbf{R}_0)$ is normal to any tangent vector at $\mathbf{R}_0$ and hence is normal to the surface. See figure 21.

2. If $\mathbf{n}$ is a unit vector, then $(\nabla u)(\mathbf{R}_0) \cdot \mathbf{n}$ is the rate of change of u at $\mathbf{R}_0$ in the direction $\mathbf{n}$, also called the directional derivative.

   To see this use the chain rule to calculate the rate of change as

   $$\frac{d}{dt}u(\mathbf{R}_0 + t\mathbf{n})\Big|_{t=0} = (\nabla u)(\mathbf{R}_0 + t\mathbf{n}) \cdot \frac{d}{dt}(\mathbf{R}_0 + t\mathbf{n})\Big|_{t=0} = (\nabla u)(\mathbf{R}_0) \cdot \mathbf{n} \tag{399}$$

3. $(\nabla u)(\mathbf{R}_0)$ is the direction of greatest increase for u at $\mathbf{R}_0$.

   To see this consider that the direction of greatest increase is the unit vector $\mathbf{n}$ which maximizes the directional derivative $(\nabla u)(\mathbf{R}_0) \cdot \mathbf{n}$. This occurs when $\mathbf{n}$ is parallel to $(\nabla u)(\mathbf{R}_0)$.

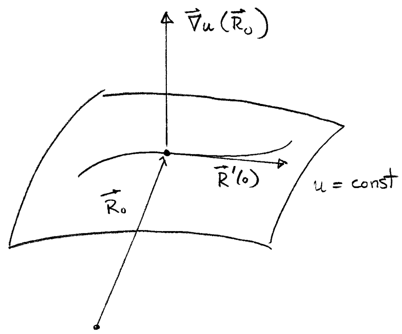


Figure 21:

problem Find the normal to the surface which is the graph of $z = f(x, y)$

solution The surface is

$$u(x, y, z) = -f(x, y) + z = 0 \tag{400}$$

A normal is

$$\nabla u = u_x \mathbf{i} + u_y \mathbf{j} + u_z \mathbf{k} = -f_x \mathbf{i} - f_y \mathbf{j} + \mathbf{k} \tag{401}$$

as before.

problem Find the direction of greatest change for $u = x^2 + 2xy + y^2 + 3z^2$ at the point $(1, 1, 1)$, i.e. find a unit vector.

solution The gradient at this point is

$$\nabla u = (2x + 2y)\mathbf{i} + (2x + 2y)\mathbf{j} + 6z\mathbf{k} = 4\mathbf{i} + 4\mathbf{j} + 6\mathbf{k} \tag{402}$$

A unit vector in this direction is

$$\mathbf{n} = \frac{\nabla u}{|\nabla u|} = \frac{4\mathbf{i} + 4\mathbf{j} + 6\mathbf{k}}{\sqrt{68}} \tag{403}$$

2.14 divergence theorem

The divergence is important because it appears in the following theorem.

Theorem 16 *(Divergence Theorem). Let $\mathcal{R}$ be a solid region in $\mathbb{R}^3$ with boundary surface $\mathcal{S}$. Let $\mathbf{n}$ be the unit outward normal on $\mathcal{S}$. Then for any continuously differentiable vector field $\mathbf{v}$ on $\mathcal{R}$*

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{v} \; dV = \int_{\mathcal{S}} \mathbf{v} \cdot \mathbf{n} \; d\sigma \tag{404}$$

Proof. Suppose $\mathbf{n} = n_1\mathbf{i} + n_2\mathbf{j} + n_3\mathbf{k}$. If suffices to show that

$$\begin{aligned} \int_{\mathcal{R}} \frac{\partial v_1}{\partial x} \; dV &= \int_{\mathcal{S}} v_1 n_1 d\sigma \\ \int_{\mathcal{R}} \frac{\partial v_2}{\partial y} \; dV &= \int_{\mathcal{S}} v_2 n_2 \; d\sigma \\ \int_{\mathcal{R}} \frac{\partial v_3}{\partial z} \; dV &= \int_{\mathcal{S}} v_3 n_3 \; d\sigma \end{aligned} \tag{405}$$

Then adding them together gives the result.

We prove the last, the others are similar. To prove it suppose that $\mathcal{R}$ if the region between the graphs of two functions. It is defined by $\phi(x,y) \leq z \leq \psi(x,y)$ with with (x,y) in some region E in the plane. Then we have

$$\begin{aligned} \int_{\mathcal{R}} \frac{\partial v_3}{\partial z} \; dV &= \int_E \left(\int_{\phi(x,y)}^{\psi(x,y)} \frac{\partial v_3}{\partial z} dz \right) dxdy \\ &= \int_E \left(v_3(x,y,\psi(x,y)) - v_3(x,y,\phi(x,y) \right) dxdy \end{aligned} \tag{406}$$

We need to show that $\int_{\mathcal{S}} v_3 n_3 \; d\sigma$ has the same expression. Now $\mathcal{S}$ has three parts $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$, see figure 22, and

$$\int_{\mathcal{S}} v_3 n_3 \; d\sigma = \int_{\mathcal{S}_1} v_3 n_3 \; d\sigma + \int_{\mathcal{S}_2} v_3 n_3 \; d\sigma + \int_{\mathcal{S}_3} v_3 n_3 \; d\sigma \tag{407}$$

The surface $\mathcal{S}_1$ is the graph of $z = \psi(x,y)$ and a normal vector is $\mathbf{N} = -\psi_x \mathbf{i} - \psi_y \mathbf{j} + \mathbf{k}$. This is an upward normal since the third component is positive. On this surface upward is outward and so the unit outward normal is

$$\mathbf{n} = \frac{\mathbf{N}}{|\mathbf{N}|} = \frac{-\psi_x \mathbf{i} - \psi_y \mathbf{j} + \mathbf{k}}{\sqrt{1 + \psi_x^2 + \psi_y^2}} \tag{408}$$

Also on $\mathcal{S}_1$ we have

$$d\sigma = \sqrt{1 + \psi_x^2 + \psi_y^2} \; dxdy \tag{409}$$

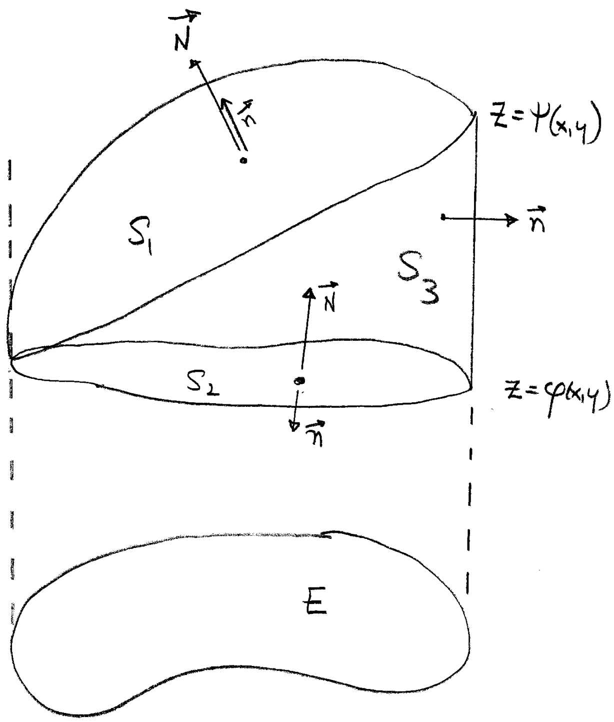


Figure 22:

Therefore

$$\begin{aligned} \int_{\mathcal{S}_1} v_3 n_3 \, d\sigma &= \int_E v_3(x,y,\psi(x,y)) \frac{1}{\sqrt{1+\psi_x^2+\psi_y^2}} \sqrt{1+\psi_x^2+\psi_y^2} \, dxdy \\ &= \int_E v_3(x,y,\psi(x,y)) \, dxdy \end{aligned} \tag{410}$$

The surface $\mathcal{S}_2$ is the graph of $z = \phi(x,y)$ and an upward normal vector is $\mathbf{N} = -\phi_x \mathbf{i} - \phi_y \mathbf{j} + \mathbf{k}$. On this surface upward is inward and so the unit outward normal is

$$\mathbf{n} = -\frac{\mathbf{N}}{|\mathbf{N}|} = \frac{\phi_x \mathbf{i} + \phi_y \mathbf{j} - \mathbf{k}}{\sqrt{1+\phi_x^2+\phi_y^2}} \tag{411}$$

Also on $\mathcal{S}_2$ we have

$$d\sigma = \sqrt{1+\phi_x^2+\phi_y^2} \, dxdy \tag{412}$$

Therefore

$$\begin{aligned} \int_{\mathcal{S}_2} v_3 n_3 \, d\sigma &= \int_E v_3(x,y,\phi(x,y)) \frac{-1}{\sqrt{1+\phi_x^2+\phi_y^2}} \sqrt{1+\phi_x^2+\phi_y^2} \, dxdy \\ &= -\int_E v_3(x,y,\phi(x,y)) \, dxdy \end{aligned} \tag{413}$$

On the surface $\mathcal{S}_3$ we have $n_3 = 0$ and so $\int_{\mathcal{S}_3} v_3 n_3 \, d\sigma = 0$

Adding the contributions from the three surfaces gives the desired result

$$\int_{\mathcal{S}} v_3 n_3 \, d\sigma = \int_E (v_3(x,y,\psi(x,y)) - v_3(x,y,\phi(x,y)) \, dxdy \tag{414}$$

The divergence theorem can be used in various ways. Here we just offer a couple of examples illustrating what it says.

example Let $\mathcal{R}$ be a sphere of radius a with surface $\mathcal{S}$. Check the divergence theorem for this region and the vector field $\mathbf{v}(\mathbf{R}) = \mathbf{R} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

For the volume integral note that $\nabla \cdot \mathbf{R} = 3$ so

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{v} \, dV = 3 \int_{\mathcal{R}} dV = 3 \times \text{ volume of } \mathcal{R} = 3\left(\frac{4}{3}\pi a^3\right) = 4\pi a^3 \tag{415}$$

On the other hand for a sphere the unit normal to the surface $\mathcal{S}$ at $\mathbf{R}$ is $\mathbf{n} = \mathbf{R}/|\mathbf{R}|$. Hence on $\mathcal{S}$ we have $\mathbf{v} \cdot \mathbf{n} = \mathbf{R} \cdot \mathbf{R}/|\mathbf{R}| = |\mathbf{R}| = a$. Therefore

$$\int_{\mathcal{S}} \mathbf{v} \cdot \mathbf{n} d\sigma = a \int_{\mathcal{S}} d\sigma = a \text{ area of } \mathcal{S} = a\left(4\pi a^2\right) = 4\pi a^3 \tag{416}$$

which agrees with the volume integral.

example Let $\mathcal{R}$ be the region defined by $0 \leq z \leq 1 - x^2 - y^2$ with $x^2 + y^2 \leq 1$. Let $\mathcal{S}$ be the boundary of $\mathcal{R}$. Check the divergence theorem for this region and the vector field $\mathbf{v} = \frac{1}{2}(x^2 + y^2)\mathbf{k}$.

The surface has two pieces. The top piece called $\mathcal{S}_1$ is the graph of the function $z = 1 - x^2 - y^2$ above the disc $x^2 + y^2 \leq 1$. A normal on this surface is

$$\mathbf{N} = -\frac{\partial z}{\partial x}\mathbf{i} - \frac{\partial z}{\partial y}\mathbf{j} + \mathbf{k} = 2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k} \tag{417}$$

This points upward which is outward for this surface. Thus the unit outward normal is

$$\mathbf{n} = \frac{\mathbf{N}}{|\mathbf{N}|} = \frac{2x\mathbf{i} + 2y\mathbf{j} + \mathbf{k}}{\sqrt{1 + 4x^2 + 4y^2}} \tag{418}$$

and on the surface we have

$$\mathbf{v} \cdot \mathbf{n} = \frac{1}{2} \frac{x^2 + y^2}{\sqrt{1 + 4x^2 + 4y^2}} \tag{419}$$

Futhermore we have for this surface

$$d\sigma = \sqrt{1 + 4x^2 + 4y^2} dxdy \tag{420}$$

Combining the above we have

$$\begin{aligned}
\int_{\mathcal{S}_1} \mathbf{v} \cdot \mathbf{n} \, d\sigma &= \int_{x^2+y^2 \leq 1} \frac{1}{2} \frac{x^2 + y^2}{\sqrt{1 + 4x^2 + 4y^2}} \sqrt{1 + 4x^2 + 4y^2} dxdy \\
&= \int_{x^2+y^2 \leq 1} \frac{1}{2} (x^2 + y^2) dxdy \\
&= \int_{r \leq 1} \frac{1}{2} r^2 \cdot r dr d\theta \\
&= \frac{1}{2} \int_0^1 r^3 dr \int_0^{2\pi} d\theta \\
&= \frac{1}{2} \cdot \frac{1}{4} \cdot (2\pi) = \frac{\pi}{4}
\end{aligned} \tag{421}$$

The bottom piece is called $\mathcal{S}_2$. It is the disc $x^2 + y^2 \leq 1$, $z = 0$. The unit outward normal is $\mathbf{n} = -\mathbf{k}$, so we have

$$\mathbf{v} \cdot \mathbf{n} = -\frac{1}{2}(x^2 + y^2) \tag{422}$$

The surface is flat so $d\sigma = dxdy$ and we have

$$\int_{\mathcal{S}_2} \mathbf{v} \cdot \mathbf{n} \, d\sigma = -\int_{x^2+y^2 \leq 1} \frac{1}{2}(x^2 + y^2) dxdy = -\frac{\pi}{4} \tag{423}$$

Combining the two pieces we have for the surface integral

$$\int_{\mathcal{S}} \mathbf{v} \cdot \mathbf{n} \, d\sigma = \int_{\mathcal{S}_1} \mathbf{v} \cdot \mathbf{n} \, d\sigma + \int_{\mathcal{S}_2} \mathbf{v} \cdot \mathbf{n} \, d\sigma = \frac{\pi}{4} - \frac{\pi}{4} = 0 \tag{424}$$

For the volume integral we note that $\nabla \cdot \mathbf{v} = \partial/\partial z\Big((x^2 + y^2)/2\Big) = 0$ and so

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{v} \, dV = 0 \tag{425}$$

which agrees with the surface integral.

remark: In evaluating surface integrals we are frequently canceling awkward square roots. We can avoid this as follows. If $\mathcal{S}$ is a surface with unit normal vector $\mathbf{n}$ define a formal symbol $d\vec{\sigma} = \mathbf{n} d\sigma$ and then

$$\int_{\mathcal{S}} \mathbf{v} \cdot \mathbf{n}\ d\sigma = \int_{\mathcal{S}} \mathbf{v} \cdot d\vec{\sigma} \tag{426}$$

If $\mathcal{S}$ is parametrized by a function $\mathbf{R}(u,v)$ then

$$\begin{aligned} \mathbf{n} &= \pm \frac{\mathbf{R}_u \times \mathbf{R}_v}{|\mathbf{R}_u \times \mathbf{R}_v|} \\ d\sigma &= |\mathbf{R}_u \times \mathbf{R}_v|\ dudv \\ d\vec{\sigma} &= \pm\left(\mathbf{R}_u \times \mathbf{R}_v\right)\ dudv \end{aligned} \tag{427}$$

If $\mathcal{S}$ is the graph of a function $z = f(x,y)$ then

$$\begin{aligned} \mathbf{n} &= \pm \frac{-f_x\mathbf{i} - f_y\mathbf{j} + \mathbf{k}}{\sqrt{1 + f_x^2 + f_y^2}} \\ d\sigma &= \sqrt{1 + f_x^2 + f_y^2}\ dxdy \\ d\vec{\sigma} &= \pm\left(-f_x\mathbf{i} - f_y\mathbf{j} + \mathbf{k}\right)\ dxdy \end{aligned} \tag{428}$$

In either case the square roots are gone from $d\vec{\sigma}$. The only difficulty is that one must still think about which normal one wants to determine whether to take the plus sign or the minus sign.

2.15 applications

We give some applications of the divergence theorem. In the first we answer the question "what is divergence?". The others are derivations of some basic partial differential equations.

A. steady fluid fow The steady (i.e. time independent) flow of a fluid is discribed by the following quantities:

$$\begin{aligned} \mathbf{v}(\mathbf{R}) &= \text{ velocity of the fluid at } \mathbf{R} \qquad (cm/sec) \\ \rho(\mathbf{R}) &= \text{ density of the fluid at } \mathbf{R} \qquad (gr/cm^3) \\ \mathbf{u}(\mathbf{R}) = \rho(\mathbf{R})\mathbf{v}(\mathbf{R}) &= \text{ mass flow density} \qquad (gr/cm^2 \cdot sec) \end{aligned}$$

Also if $\mathcal{S}$ is a surface with unit normal $\mathbf{n}$ we define

$$\int_{\mathcal{S}} \mathbf{u} \cdot \mathbf{n}\ d\sigma = \text{ flux of } \mathbf{u} \text{ over } \mathcal{S} \text{ in direction } \mathbf{n}$$

We want to interpret this flux and also the divergence of $\mathbf{u}$. (They are related).

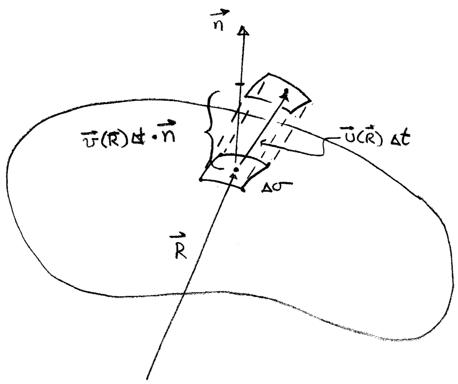


Figure 23:

Pick a point $\mathbf{R}$ on the the surface $\mathcal{S}$, let $\Delta\sigma$ be a small piece of surface around $\mathbf{R}$. Then (see figure 23)

$$\begin{aligned} &\text{mass through } \Delta\sigma \text{ in time } \Delta t \qquad\qquad (gr) \\ \approx\, &\text{density at } \mathbf{R} \times \text{ volume through } \Delta\sigma \text{ in time } \Delta t \\ \approx\, &\rho(\mathbf{R}) \times \Delta\sigma \times (\mathbf{v}(\mathbf{R})\Delta t) \cdot \mathbf{n} \\ =\, &\mathbf{u}(\mathbf{R}) \cdot \mathbf{n}\ \Delta\sigma\Delta t \end{aligned} \tag{429}$$

If we divide by Δt we get

$$\text{rate of mass flow thru } \Delta\sigma = \mathbf{u}(\mathbf{R}) \cdot \mathbf{n}\ \Delta\sigma \qquad (gr/sec)$$

Now sum over pieces $\Delta\sigma_i$ covering the surface $\mathcal{S}$ and take the limit as the partition becomes fine. This gives

$$\text{rate of mass flow thru } \mathcal{S} = \lim \sum_i \mathbf{u}(\mathbf{R}_i) \cdot \mathbf{n}_i\ \Delta\sigma_i = \int_{\mathcal{S}} \mathbf{u}(\mathbf{R}) \cdot \mathbf{n}\ d\sigma \qquad (gr/sec)$$

Thus we have an interpretation of the flux of $\mathbf{u}$ over $\mathcal{S}$. It is the rate of mass flow through $\mathcal{S}$.

Now for the divergence of $\mathbf{u}$ at any point $\mathbf{R}$ let D_ϵ be a sphere of radius ϵ around $\mathbf{R}$. let S_ϵ be the surface of that sphere with outward normal $\mathbf{n}$. Then by the divergence theorem

$$\begin{aligned}(\nabla \cdot \mathbf{u})(\mathbf{R}) &= \lim_{\epsilon \to 0} \frac{1}{\text{Vol } D_\epsilon} \int_{D_\epsilon} \nabla \cdot \mathbf{u}\, dV \\ &= \lim_{\epsilon \to 0} \frac{1}{\text{Vol } D_\epsilon} \left(\int_{S_\epsilon} \mathbf{u} \cdot \mathbf{n}\, d\sigma \right)\end{aligned} \tag{430}$$

This is the flux divided by the volume. Thus the divergence is the rate of outward mass flow per unit volume (gr/ $\text{cm}^3 \cdot$ sec). A large divergence at a point means a lot of fluid is entering the system at that point.

B. fluid dynamics

Again we consider fluid flow, but now the velocity $\mathbf{v}(\mathbf{R}, t)$, the density $\rho(\mathbf{R}, t)$, and the mass flow density $\mathbf{u}(\mathbf{R}, t) = \rho(\mathbf{R}, t)\mathbf{v}(\mathbf{R}, t)$ all depend on the time t.

In our fluid consider any region $\mathcal{R}$ with surface $\mathcal{S}$ and outward normal $\mathbf{n}$. Conservation of mass says that

$$\text{rate of mass flow out of } \mathcal{S} = \text{ rate of decrease of mass in } \mathcal{R} \tag{431}$$

which means that

$$\int_{\mathcal{S}} \mathbf{u} \cdot \mathbf{n}\, d\sigma = -\frac{d}{dt} \left(\int_{\mathcal{R}} \rho(\mathbf{R}, t)\, dV \right) \tag{432}$$

Use the divergence theorem on the left, and differentiate under the integral sign on the right to obtain

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{u}\, dV = -\int_{\mathcal{R}} \frac{\partial \rho}{\partial t}\, dV \tag{433}$$

which is the same as

$$\int_{\mathcal{R}} \left(\nabla \cdot \mathbf{u} \ + \frac{\partial \rho}{\partial t} \right) dV = 0 \tag{434}$$

Since this holds for an arbitrary region $\mathcal{R}$ it must be that

$$\nabla \cdot \mathbf{u} \ + \frac{\partial \rho}{\partial t} = 0 \tag{435}$$

which is also written

$$\nabla \cdot (\rho \mathbf{v}) \ + \frac{\partial \rho}{\partial t} = 0 \tag{436}$$

This is the *continuity equation* which must be satisfied by any flow.

We can rewrite this as

$$\rho \nabla \cdot \mathbf{v} \ + \nabla \rho \cdot \mathbf{v} + \frac{\partial \rho}{\partial t} = 0 \tag{437}$$

Now define the total derivative of ρ to be

$$\frac{D\rho}{Dt} = \frac{\partial \rho}{\partial t} + \nabla \rho \cdot \mathbf{v} \tag{438}$$

Then the continuity equation can be written as

$$\frac{D\rho}{Dt} + \rho \; \nabla \cdot \mathbf{v} \; = 0 \tag{439}$$

The total derivative has the interpretation

$$\frac{D\rho}{Dt} = \text{ rate of change of the density at a test particle moving in the fluid} \tag{440}$$

To see that this is true let $\mathbf{R}(t)$ be the trajectory of the test particle. Moving with the fluid means that $d\mathbf{R}/dt = \mathbf{v}(\mathbf{R}(t), t)$. The rate of change of the density at the test particle is by the chain rule

$$\begin{aligned} \frac{d}{dt}\rho(\mathbf{R}(t), t) =& \nabla \rho(\mathbf{R}(t), t) \cdot \frac{d\mathbf{R}}{dt} + \frac{\partial \rho}{\partial t}(\mathbf{R}(t), t) \\ =& \nabla \rho(\mathbf{R}(t), t) \cdot \mathbf{v}(\mathbf{R}(t), t) + \frac{\partial \rho}{\partial t}(\mathbf{R}(t), t) \\ =& \frac{D\rho}{Dt}(\mathbf{R}(t), t) \end{aligned} \tag{441}$$

as claimed.

The field is *incompressible* if the test particle sees no change in the density, that is

$$D\rho/Dt = 0 \tag{442}$$

By the second form of the continuity equation this is the same as

$$\nabla \cdot \mathbf{v} = 0 \tag{443}$$

C. heat equation For any object let $T(\mathbf{R}, t)$ be the temperature of the object at position $\mathbf{R}$ and time t. We want to derive an equation which describes how the temperature evolves in time. For this we also need to consider the heat (energy) in the system. The heat density (calories/ cm^3) at $\mathbf{R}, t$ is proportional to the temperature there and has the form $c\rho T(\mathbf{R}, t)$ where the specific heat c is a constant depending on the material and ρ is the mass density, assumed constant. We also need the heat flow $u(\mathbf{R}, t)$ at $\mathbf{R}, t$. This is analagous to the mass flow in the previous examples but is now the flow of energy (calories/ $\text{cm}^2 \cdot$ sec).

In our object consider any region $\mathcal{R}$ with surface $\mathcal{S}$ and outward normal $\mathbf{n}$. Conservation of energy says that

$$\text{rate of heat flow out of } \mathcal{S} = \text{ rate of decrease of heat in } \mathcal{R} \tag{444}$$

which means that

$$\int_{\mathcal{S}} \mathbf{u} \cdot \mathbf{n} \; d\sigma = -\frac{d}{dt}\left(\int_{\mathcal{R}} c\rho T(\mathbf{R}, t) \; dV \right) \tag{445}$$

Again use the divergence theorem on the left, and differentiate under the integral sign on the right to obtain

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{u} \; dV = -\int_{\mathcal{R}} c\rho \frac{\partial T}{\partial t} \; dV \tag{446}$$

which is the same as

$$\int_{\mathcal{R}} \left(\nabla \cdot \mathbf{u} \; + c\rho \frac{\partial T}{\partial t} \right) \; dV = 0 \tag{447}$$

Since this holds for any region $\mathcal{R}$ it must be that

$$\nabla \cdot \mathbf{u} \; + c\rho \frac{\partial T}{\partial t} = 0 \tag{448}$$

Now we need another fact. This is the thermal conduction law which says that the heat flow is proportional to the negative gradient of the temperature:

$$\mathbf{u} = -k \nabla T \tag{449}$$

Here k is a positive constant called the *thermal conductivity* . The law says that heat flows in the direction of greatest temperature decrease. Inserting this in the above equation gives

$$c\rho \frac{\partial T}{\partial t} - k \Delta T = 0 \tag{450}$$

This is called the *heat equation* . If the temperature is independent of time then this becomes

$$\Delta T = 0 \tag{451}$$

which is known as *Laplace's equation*.

2.16 more line integrals

We define a line integral of vector fields. Let $\mathcal{C}$ be a directed curve parametrized by $\mathbf{R}(t)$, $a \leq t \leq b$ and let $\mathbf{v}(\mathbf{R})$ be a vector field defined on $\mathcal{C}$. We define

$$\int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{R} = \int_a^b \mathbf{v}(\mathbf{R}(t)) \cdot \frac{d\mathbf{R}}{dt} \; dt \tag{452}$$

Thus we replace the curve by the parameter and interpret $d\mathbf{R} = (d\mathbf{R}/dt)dt$. This definition turns out to be independent of parametrization as long as we respect the direction of the curve.

Let

$$\mathbf{T} = \frac{d\mathbf{R}}{dt} / |\frac{d\mathbf{R}}{dt}| \tag{453}$$

be a unit tangent vector to the curve. Then the vector line integral is related to a scalar line integral by

$$\int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{R} = \int_{\mathcal{C}} \mathbf{v} \cdot \mathbf{T}\ ds \tag{454}$$

Thus we are integrating the tangential component of $\mathbf{v}$ along the curve. To see that it is true compute

$$\begin{aligned} \int_{\mathcal{C}} \mathbf{v} \cdot \mathbf{T}\ ds &= \int_a^b \mathbf{v}(\mathbf{R}(t)) \cdot \frac{d\mathbf{R}/dt}{|d\mathbf{R}/dt|} |d\mathbf{R}/dt|\ dt \\ &= \int_a^b \mathbf{v}(\mathbf{R}(t)) \cdot d\mathbf{R}/dt\ dt \\ &= \int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{R} \end{aligned} \tag{455}$$

problem Let $\mathcal{C}$ be a straight line from $\mathbf{a} = \mathbf{i}+\mathbf{k}$ to $\mathbf{b} = 2\mathbf{i}+\mathbf{j}+3\mathbf{k}$. Let $\mathbf{v}(\mathbf{R}) = x\mathbf{i}+yz\mathbf{k}$. Evlauate $\int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{R}$.

solution: $\mathcal{C}$ can be parametrized by $\mathbf{R}(t) = (1-t)\mathbf{a} + t\mathbf{b}$ with $0 \leq t \leq 1$ which is the same as

$$\mathbf{R}(t) = (t+1)\mathbf{i} + t\mathbf{j} + (2t+1)\mathbf{k} \tag{456}$$

Then

$$d\mathbf{R} = \frac{d\mathbf{R}}{dt}\ dt = (\mathbf{i} + \mathbf{j} + 2\mathbf{k})dt \tag{457}$$

Also

$$\mathbf{v}(\mathbf{R}(t)) = (t+1)\mathbf{i} + t(2t+1)\mathbf{k} \tag{458}$$

Then we compute

$$\int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{R} = \int_0^1 \Big((t+1) + 2t(2t+1)\Big)dt = \int_0^1 (1+3t+4t^2)dt = \frac{23}{6} \tag{459}$$

We list some properties of line integrals

- For any vector fields $\mathbf{v}, \mathbf{w}$

$$\int_{\mathcal{C}} (\mathbf{v}+\mathbf{w}) \cdot d\mathbf{R} = \int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{R} + \int_{\mathcal{C}} \mathbf{w} \cdot d\mathbf{R} \tag{460}$$

- For a constant α

$$\int_{\mathcal{C}} \alpha\mathbf{v} \cdot d\mathbf{R} = \alpha \int_{\mathcal{C}} \mathbf{v} \cdot d\mathbf{R} \tag{461}$$

- Let $\mathcal{C}_2$ start where $\mathcal{C}_1$ finishes and let $\mathcal{C}_1+\mathcal{C}_2$ be the curve which first traverses $\mathcal{C}_1$ and then traverses $\mathcal{C}_2$. Then

$$\int_{\mathcal{C}_1+\mathcal{C}_2} \mathbf{v}\cdot\mathbf{R} = \int_{\mathcal{C}_1} \mathbf{v}\cdot\mathbf{R} + \int_{\mathcal{C}_2} \mathbf{v}\cdot\mathbf{R} \tag{462}$$

- Let $-\mathcal{C}$ be the curve $\mathcal{C}$ traversed in the opposite direction. Then

$$\int_{-\mathcal{C}} \mathbf{v}\cdot\mathbf{R} = -\int_{\mathcal{C}} \mathbf{v}\cdot\mathbf{R} \tag{463}$$

(For scalar integrals on the other hand we have $\int_{-\mathcal{C}} f ds = \int_{\mathcal{C}} f ds$)

application: Let $\mathbf{F}(\mathbf{R})$ be the force applied to an object at position $\mathbf{R}$. Then the line integral $\int_{\mathcal{C}} \mathbf{F}(\mathbf{R})\cdot d\mathbf{R}$ is interpreted as the work done (energy expended) in moving the object along $\mathcal{C}$.

another notation If $\mathbf{v} = v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}$ and $d\mathbf{R} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}$ then

$$\mathbf{v}\cdot d\mathbf{R} = v_1 dx + v_2 dy + v_3 dz \tag{464}$$

This is called a *differential form*. For us it is just a formal symbol whose integral has a meaning, but it can be given a separate precise meaning in higher mathematics. In this notation our definition of the line integral of $\mathbf{v}$ along a curve $\mathcal{C}$ parametrized by $\mathbf{R}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ with $a \leq t \leq b$ is

$$\begin{aligned}&\int_{\mathcal{C}} v_1 dx + v_2 dy + v_3 dz\\ &= \int_a^b \left(v_1(x(t),y(t),z(t))\frac{dx}{dt} + v_2(x(t),y(t),z(t))\frac{dy}{dt} + v_3(x(t),y(t),z(t))\frac{dz}{dt}\right)\, dt\end{aligned} \tag{465}$$

In other words we interpret dx as $(dx/dt)dt$ and so forth.

This notation is also used for $\mathbb{R}^2$. If the curve $\mathcal{C}$ is parametrized by $x = x(t), y = y(t)$ with $a \leq t \leq b$ and $M(x,y)$ and $N(x,y)$ are functions on $\mathcal{C}$, then

$$\int_{\mathcal{C}} M dx + N dy = \int_a^b \left(M(x(t),y(t))\frac{dx}{dt} + N(x(t),y(t))\frac{dy}{dt}\right)\, dt \tag{466}$$

definitions A curve $\mathcal{C}$ is *simple* if it does not intersect itself (except possibly at the endpoints). A curve is *closed* if two endpoints are the same point. See figure 24 for examples.

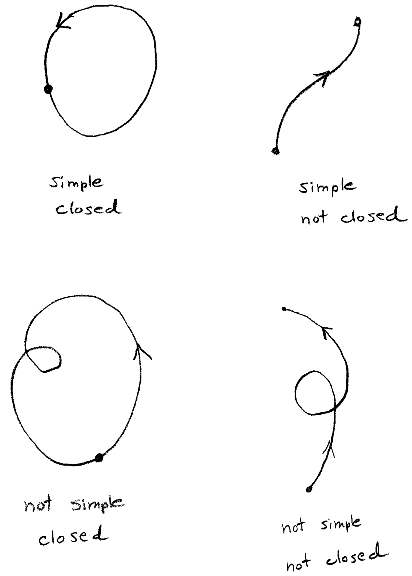


Figure 24:

Theorem 17 *(Green's Theorem). Let $\mathcal{C}$ be a simple closed curve in the plane traversed counterclockwise with interior $\mathcal{R}$. If M, N are continuously differentiable everywhere in $\mathcal{R}$ then*

$$\int_{\mathcal{C}} M dx + N dy = \int_{\mathcal{R}} \left(\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} \right) \, dx dy \tag{467}$$

Proof. Suppose the region $\mathcal{R}$ lies between the graphs of two functions $y = p(x)$ and $y = q(x)$ with $a \leq x \leq b$ and $p(x) \leq q(x)$. Call these two curves $\mathcal{C}_1$ and $\mathcal{C}_2$ both traversed in the direction of increasing x. Then $\mathcal{C} = \mathcal{C}_1 - \mathcal{C}_2$. (see figure 25). We compute

$$\begin{aligned} \int_{\mathcal{R}} -\frac{\partial M}{\partial y} \, dx dy = & - \int_a^b \left[\int_{p(x)}^{q(x)} \frac{\partial M}{\partial y}(x, y) dy \right] dx \\ & = - \int_a^b (M(x, q(x)) - M(x, p(x)) dx \\ & = - \int_{\mathcal{C}_2} M dx + \int_{\mathcal{C}_1} M dx \\ & = \int_{\mathcal{C}_1 - \mathcal{C}_2} M dx = \int_{\mathcal{C}} M dx \end{aligned} \tag{468}$$

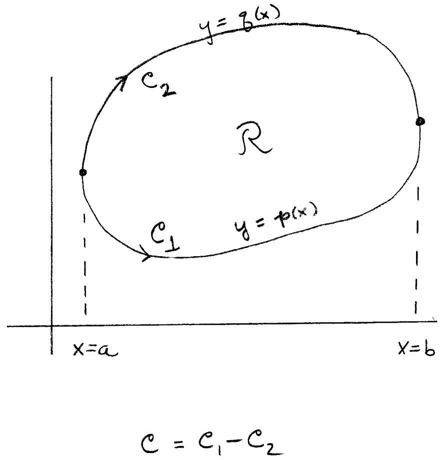


Figure 25:

Here we have used that $\mathcal{C}_2$ can be parametrized by $x = x, y = q(x), a \leq x \leq b$ and that $\mathcal{C}_1$ can be parametrized by $x = x, y = p(x), a \leq x \leq b$.

Similarly one shows that

$$\int_{\mathcal{R}} \frac{\partial N}{\partial x}\, dxdy = \int_{\mathcal{C}} Ndy \tag{469}$$

Adding the two equations gives the result.

Corollary: If $\partial N/\partial x = \partial M/\partial y$ everywhere inside a simple closed curve $\mathcal{C}$ then

$$\int_{\mathcal{C}} Mdx + Ndy = 0 \tag{470}$$

Corollary: With $M = -y/2$ and $N = x/2$

$$\frac{1}{2}\int_{\mathcal{C}} -ydx + xdy = \int_{\mathcal{R}} dxdy = \text{ area of } \mathcal{R} \tag{471}$$

problem Find the area inside the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{472}$$

solution: The ellipse can be parametrized by $x = a\cos t$, $y = b\sin t$ with $0 \leq t \leq 2\pi$. Then we have

$$dx = -a\sin t dt \qquad dy = b\cos t dt \tag{473}$$

Hence

$$\begin{aligned} \text{Area} =& \frac{1}{2}\int_{\mathcal{C}} -ydx + xdy \\ =& \frac{1}{2}\int_0^{2\pi} (ab\sin^2 t + ab\cos^2 t)dt \\ =& \frac{1}{2}\int_0^{2\pi} ab\ dt = \pi ab \end{aligned} \tag{474}$$