DIRECTORATE OF DISTANCE EDUCATION

# UNIVERSITY OF NORTH BENGAL

MASTERS OF SCIENCE-MATHEMATICS

## SEMESTER -I

# ANALYSIS OF SEVERAL VARIABLES

## DEMATH-1 SCORE-3

# BLOCK-2

**UNIVERSITY OF NORTH BENGAL**

Postal Address:

The Registrar,

University of North Bengal,

Raja Rammohunpur,

P.O.-N.B.U.,Dist-Darjeeling,

West Bengal, Pin-734013,

India.

Phone: ( O ) +91 0353-2776331/2699008

Fax:( 0353 ) 2776313, 2699001

Email: regnbu@sancharnet.in ; regnbu@nbu.ac.in

Wesbsite: www.nbu.ac.in

First Published in 2019

# FOREWORD

The Self-Learning Material (SLM) is written with the aim of providing simple and organized study content to all the learners. The SLMs are prepared on the framework of being mutually cohesive, internally consistent and structured as per the university’s syllabi. It is a humble attempt to give glimpses of the various approaches and dimensions to the topic of study and to kindle the learner’s interest to the subject

We have tried to put together information from various sources into this book that has been written in an engaging style with interesting and relevant examples. It introduces you to the insights of subject concepts and theories and presents them in a way that is easy to understand and comprehend.

We always believe in continuous improvement and would periodically update the content in the very interest of the learners. It may be added that despite enormous efforts and coordination, there is every possibility for some omission or inadequacy in few areas or topics, which would definitely be rectified in future.

We hope you enjoy learning from this book and the experience truly enrich your learning and help you to advance in your career and future endeavours.

# ANALYSIS OF SEVERAL VARIABLES

## BLOCK 1

Unit 1 Topology

Unit 2 The Space Of Linear Transformations From Rm To Rn

Unit 3 Local Properties Of Continuous Functions

Unit 4 The Basic Facts Of Differential Calculus Of Real-Valued Functions Of Several Variables

Unit 5 Transition To The Case Of A Relation

Unit 6 Taylor's Theorem

Unit 7 A Sufficient Condition For A Constrained Extremum

## BLOCK 2

**Unit -8: The Basic Rules Of Differentiation .......................................... 7**

**Unit - 9: The Basic Theorems Of Differential Calculus .................... 39**

**Unit - 10 : The Riemann Integral In N Variables ............................. 70**

**Unit -11 Connectedness ........................................................................ 115**

**Unit-12 : Differentiation Of Vector-Valued Functions.................... 151**

**Unit - 13 : Multivariable Differential Calculus ................................ 186**

**Unit -14 Systems Of Differential Equations And Vector Fields ..... 220**

# BLOCK-2 ANALYSIS OF SEVERAL VARIABLES

**Introduction to the Block**

In this block we will go through In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

| | |
|---|---|
| Unit 8 | The Basic Rules Of Differentiation And The Arithmetic Operations |
| Unit 9 | The Basic Theorems Of Differential Calculus |
| Unit 10 | The Riemann Integral In N Variables |
| Unit 11 | Connected Sets |
| Unit 12 | Differentiation Of Vector-Valued Functions |
| Unit 13 | Multivariable Differential Calculus |
| Unit 14 | Systems Of Differential Equations And Vector Fields |

# UNIT -8: THE BASIC RULES OF DIFFERENTIATION

**STRUCTURE**

8.0 Objectives

8.1 Introduction

8.2 Basic Rules of Differentiation & Arithmetic Operations

8.3 Differentiation Of A Composite Function (Chain Rule)

8.4 Differentiation Of An Inverse Function

8.5 Table Of Derivatives Of The Basic Elementary Functions

8.6 Let Us Sum Up

8.7 Keywords

8.8 Questions For Review

8.9 References

8.10 Answers To Check Your Progress

## 8.0 OBJECTIVES

After studying this unit,you should be able to:
Learn Understand about The Basic Rules Of Differentiation And The Arithmetic Operations

Learn Understand about Differentiation Of A Composite Function (Chain Rule)

Learn Understand about Differentiation Of An Inverse Function

Learn Understand about Table Of Derivatives Of The Basic Elementary Functions

## 8.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

The Basic Rules Of Differentiation, And The Arithmetic Operations, Differentiation Of A Composite Function (Chain Rule), Differentiation Of An Inverse Function, Table Of Derivatives Of The Basic Elementary Functions

## 8.2 THE BASIC RULES OF DIFFERENTIATION

### Differentiation And The Arithmetic Operations

Constructing the differential of a given function or, equivalently, the process of finding its derivative, is called differentiation.

Theorem . If functions f : X -A R and g : X -A R are differentiable at a point x € X,then their sum is differentiable at x

Proof. In the proof we shall rely on the definition of a differentiable function and the properties of the symbol o(-) .

(/ + g)(x + h)-(f + g)(x) = (/(x + h) + g(x + h)) -

- (f(x) + g(x)) = (/(x + h)~ f(x)) + (g(x + h) - g{x)) = = (f'{x)h + o(h)) + (g'(x)h + o(h)) = (f'(x) + g'(x))h + o(h) =

(f'+g'){x)h + o(h) .

(/ • g)(x + h) - (/ • g)(x) = /(x + h)g{x + h) - f(x)g(x) =

= (/(») + f'(x)h + o(h)) (g(x) + g'(x)h + o(h)) - f(x)g(x) =

(f'(x)g(x) + f(x)g'(x))h + o{h) .

Since a function that is differentiable at a point x G X is continuous at that point,taking account of the relation g(x) ^ 0 and the properties of continuous functions, we can guarantee that g(x + h) ^ 0 for sufficiently small values of h. In the following computations it is assumed that h is small:

(L)ix + h) _ (t\(X) = ^ + _ M =

V5/ g(x + h) g(x)

= g(x)g(x + h) (/(X + - f{x)9{x + =

= (gT^j +°(1))((f(x) + f'(x)h+°(h))9(x)-f(x)(g(x)+g'(x)h+o{h))) =

= (^y + °(1)) (if'(x)9(x) ~ f(x)g'(x))h + 0(h)) =

f'(x)g(x) - f(x)g'(x)

= h + °\h) ■

g2(x)

Here we have used the continuity of g at the point x and the relation g(x) 0 to deduce that

h->o g(x)g(x + h) g2(x) '

that is,

1 1

g(x)g(x + h) g2(x)+° ' where o(l) is infinitesimal as /i 0,a: + /i G I. □

Corollary . The derivative of a linear combination of differentiate functions equals the same linear combination of the derivatives of these functions.

Proof Since a constant function is obviously differentiate and has a derivative equal to 0 at every point,taking / = const = c in statement find (cg)'{x) = cg'{x).

Now,using statement we can write

(cif + C2g)'(x) = (cif)'(x) + (c2g)'(x) = ci/'(x) + c2g'(x) .

Taking account of what has just been proved,we verify by induction that

(Cl/l + h Cnfn)'(x) = Cif[(x) + b Cnfn(x) .

Corollary. If the functions /i, ...,fn are differentiate at x,then

(/l • • • fn)'(x) = f[(x)f2(x) ■ ■ ■ fn(x) +

+ fl(x)f2(x)f3(x) • • • fn(x) + ■•• +fl(x)--- fn-i(x)fn(x) .

Proof For n = 1 the statement is obvious.

If it holds for some n <= N,then by statement also holds for (n -f 1) e N. By the principle of induction,we conclude that the formula is valid for any n <= N.

Corollary. It follows from the relation between the derivative and the differential that Theorem can also be written in terms of differentials. To be specific:

d(/ + g)(x) = df(x) + dg(x) ;

d(/ • g)(x) = g{x)df{x) + f{x)dg(x) ;

d(J)(*) = 9^W(x)-f{x)dg(X) ifg{x) ^ 0 _

Proof Assume us verify, for example, statement a).

d(/ + g){x)h = (/ + g)'{x)h = (/' + g'){x)h =

= (f'(x) + g'{x))h = f'(x)h + g'(x)h =

= df(x)h + dg(x)h = (df(x) + dg(x))h,

and we have verified that d(/ + g)(x) and d/(x) + dg(x) are the same function.

Example. Invariance of the definition of velocity. We are now in a position to verify that the instantaneous velocity vector of a point mass defined in Subsect.is independent of the Cartesian coordinate system used to define it. In fact we shall verify this for all affine coordinate systems.

Assume (x1,x2) and (x1,x2) be the coordinates of the same point of the plane in two different coordinate systems connected by the relations

x1 = alx1 + a\x2 + b1,,.

x2 = alx1 + a2x2 + b2 .          {b'2b>

Since any vector (in affine space) is determined by a pair of points and its coordinates are the differences of the coordinates of the terminal and initial points of the vector, it follows that the coordinates of a given vector in these two coordinate systems must be connected by the relations

v1 = a\vl + a\v2, (v

v2 = a2v1+a2v2. [b'Z7)

If the law of motion of the point is given by functions xx{t) and x2(t) in one system of coordinates, it is given in the other system by functions xx(t) and x2(t) connected with the first set by relations.

Differentiating relations with respect to <=,we find by the rules for differentiation

<5-28>

x = afx + a%x .

Thus the coordinates (v1, v2) = (xx, x2) of the velocity vector in the first

• 1 • 2

system and the coordinates (y1, !)2) = (x,x) of the velocity vector in the second system are connected by relations telling us that we are dealing with two different expressions for the same vector.

Example. Assume f(x) = tanx. We shall show that f'(x) = col2 x at every point where cosx ^ 0,that is,in the domain of definition of the function tanx = COS x

It was shown that sin'(x) = cosx and cos' x = — sin x,so that by statement

we find,

when cos x^O,

/sin\' x sin'x cos x — sin x cos'x t. = I I 17!) =

cos x cos x + sin x sin x

Example. cot'x = — sin\ x wherever sinx ^ 0,

that is in the domain of definition of cotx = .

sin x

Indeed,

cos' x sin x — cos x sin' x

sin2 x

— sin x sin x — cos x cos x

sin2 x  sin2 x

Example • If P(pc) = co + c\x + • • • + cnxn is a polynomial,then P'(x) = ci + 2c2x +  1- ncnxn_1.

Indeed,since ^ = 1,

we have = nxn_1,and the statement

## 8.3 DIFFERENTIATION OF A COMPOSITE FUNCTION (CHAIN RULE)

Theorem. (Differentiation of a composite function). If the function f : X -» Y C M is differentiate at a point x G X and the function g : Y -» R differentiate at the point y = f(x) G T,then the composite function go f : X -» R differentiate at x,and the differential d(go/)(x) : TR(x) -> TR(^(/(x))) of their composition equals the composition df(y) o d/(x) o/ ^/ie2r differentials d/(x) : TR(x) -> TR(y = /(x)) and d^(y = /(x)) : TR(y) -> TR^^)) .

Proof. The conditions for differentiability of the functions / and g have the form

/(x + h) — f(x) = f'(x)h + o(h) as/i-»0,x + /iGX,

+ *) - 5(2/) = 9'{y)t + o(t) ast->0, y + teY.

We remark that in the second equality here the function o(t) can be considered to be defined for t = 0,and in the representation

o(t) = 7(<=)<=,

where 7(t) -» 0 as t -» 0,y + t G T,

we may assume 7(0) = 0.

Setting f(x) = y and f(x + ft) = y + t,

by the differentiability (and hence continuity) of / at the point x we conclude that t -» 0 as ft -» 0,and if x + ft G X,then y + t G y.

By the theorem on the limit of a composite function,

we now have 7 (/(# + ft) — /(x)) = a(ft) -» 0 as ft -» 0,x + ft G X, and thus if t = f(x + ft) — /(#), then

= 7(/(® + ft) - /(«)) (/(a; + ft) - f(x)) =

= a(h)(f'(x)h + o(h)) = a(h)f'(x)h + a(h)o(h) =

= o(ft) + o(ft) = o(ft) asft-»0,x + ftGX.

(5 ° /)(x + h)-(go f)(x) = g(f(x + ft)) - g(f(x)) =

= g(y + t)~ g(y) = g'(y)t + o(t) =

= )) (f(x + ft) - f(x)) + o(f(x + ft) - f(x)) =

= g'(f(x)) (f'(x)h + o(h)) + o(f(x + ft) - /(x» =

= g'(f(x)) (f'(x)h) + g'(f(x)) (o(ft)) + o(/(x + ft) - f(x)) .

Since we can interpret the quantity gf(f(x))(/' (x)h) as the value

dg(/(x)) o d/(x)ft of the composition ft g'(f(x)) • ff(x)h of the

mappings ft /'(#)ft,t gf{y)x at the displacement ft,to compe the proof it remains only for us to remark that the sum

(o(ft)) + o(f(x + ft) - f(x))

is infinitesimal compared with ft as ft —» 0,x + ft G X,or,as we have already established,

o(f(x + ft) — f(x)) = o(h) as ft -» 0,x + h e X .

Thus we have proved that

(0°/)(® + ft) - (g°f)(x) =

— </(/(x)) * f'(x)h + asft-»0,x + ftGX.

Corollary. The derivative (g o f)'(x) of the composition of differentiate real-valued functions equals the product gf(f(x)) • ff(x) of the derivatives of these functions computed at the corresponding points.

There is a strong temptation to give a short proof of this last statement in Leibniz' notation for the derivative,in which if z = z(y) and y = y(x), we have

dz dz dy dx dy dx '

which appears to be compely natural,

if one regards the symbol ^ or ^ not as a unit,but as the ratio of dz to dy or dy to dx.

The idea for a proof that thereby arises is to consider the difference quotient

Az _ Az Ay Ax Ay Ax

and then pass to the limit as Ax —» 0. The difficulty that arises here (which we also have had to deal with in part!) is that Ay may be 0 even if Ax ^ 0.

Corollary. If the composition (fn o • • • o f\)(x) of differentiable functions Vi = h(x),..., yn = fn{yn-i) exists,then

(fn O • • • O /l)'(x) = f'n(yn-\)f'n-l(yn-l)---f[(x) •

Proof The statement is obvious if n = 1.

If it holds for some n G N,then by Theorem 2 it also holds for n + 1,so that by the principle of induction,it holds for any n <= N.

Example. Assume us show that for a G M we have = ax01-1 in the domain x > 0,that is,dxa = ax°l~1dx and

(x + h)a — xa = ax°l~1h + o(h) as h -» 0 .

Proof We write xa = ealnx and apply the theorem,taking account Assume g(y) = ey and y = f(x) = aln(x). Then xa = (g o f)(x) and

(y ° f)'(x) = g'(y) ■ f'(x) = ev ■ - = ealnx • - = ax0"1 .

X X

Example The derivative of the logarithm of the absolute value of a differentiable function is often called its logarithmic derivative.

Since F(x) = In |/(x)| = (lno||o /)(#)

We have F'(x) = (ln|/|)'(a;) =

Thus

d (,,,i/i)(*) = <=Mdx = ig>.

Example. The absolute and relative errors in the value of a differentiable function caused by errors in the data for the argument.

If the function / is differentiable at x,then

f(x + h) - f(x) = f'(x)h + a(x; h),

where a(x; h) = o(h) as h -» 0.

Thus,if in computing the value f(x) of a function,the argument x is determined with absolute error /i,the absolute error \f(x + h) — f(x)|in the value of the function due to this error in the argument can be replaced for small values of h by the absolute value of the differential \df(x)h\ =\f'(x)h\ at displacement h.

The relative error can then be computed as the ratio or as the absolute value of the product |y^| \h\ of the logarithmic derivative of the function and the magnitude of the absolute error in the argument.

We remark by the way that if f(x) = lnx,then dlnx = and the absolute error in determining the value of a logarithm equals the relative error in the argument. This circumstance can be beautifully exploited for example,in the slide rule (and many other devices with nonuniform scales). To be specific,assume us imagine that with each point of the real line lying right of zero we connect its coordinate y and write it down above the point,while below the point we write the number x = ey. Then y = lnx. The same real half-line has now been endowed with a uniform scale y and a nonuniform scale x (called logarithmic).

Notes

To find lnx,one need only set the cursor on the umber x and read the corresponding number y written above it. Since the precision in setting the cursor on a particular point is independent of the number x or y corresponding to it and is measured by some quantity Ay (the length of the interval of possible deviation) on the uniform scale,we shall have approximately the same absolute error in determining both a number x and its logarithm y\ and in determining a number from its logarithm we shall have approximately the same relative error in all parts of the scale.

Example. Assume us differentiate a function u(x)v(x\ where u(x) and v(x) are differentiate functions and u(x) > 0. We write u(x)v^ = ev^ lnn(x)

In u(x) ,

f (\ = e^x)ln«(x)^/(a.)lnu^ =

ax \ u(x) J

= u(x)v^ - v'{x) lnw(x) + vi^uixY^-1 - v!(x) .

## 8.4 DIFFERENTIATION OF AN INVERSE FUNCTION

Theorem. (The derivative of an inverse function). Assume the functions f : X -» Y and /-1 : Y -» X be mutually inverse and continuous at points xq G X and f(xo) = yo <= Y respectively. If f is differentiable at xo and f'{xo) 7^ 0,then f~l is also differentiable at the point yo, and

= (f'(x o))-1.

Proof Since the functions f : X -$Y and /-1 :Y -» X are mutually inverse, the quantities f(x) — f(xo) and f~l(y) — f~l{yo)-> where y = f(x), are both nonzero if x ^ xo- In addition, we conclude from the continuity of / at Xo and /-1 at yo that (X 3 x -» Xo) & (Y 3 y -» yo)- Now using the theorem on the limit of a composite function and the arithmetic properties of the limit, we find

f~1(y)-f~1(y o) x-x0

lira = lim ~~<=~{ \ 77—7 =

YBy->vo 2/ - 2/o        X3x->x0 J(x) - f(x0)

XSHio //(^)-/(xo)\ f'(x0) '

\ x—xq j

Thus we have shown that the function /-1 : Y -» X has a derivative at yo and that

(/_1)'(2/o) = (f'(x o))_1

Remark. If we knew in advance that the function /-1 was differentiate at yo? we would find immediately by the identity (/-1 o /)(x) = x and the theorem on differentiation of a composite function that        — 1-

Remark . The condition f'(xo) ^ 0 is obviously equivalent to the statement that the mapping h ^ ff(xo)h realized by the differential d/(xo) : TR(xo) -» TR(yo) has the inverse mapping [d/(xo)]_1 : TR(yo) TR(xo) given by the formula r ^ (f'(xo))

Hence,in terms of differentials we can write the second statement in as follows:

If a function f is differentiate at a point xo and its differential d/(#o) : TR(xq) -» TR(yo) is invertible at that point,then the differential of the function /-1 inverse to f exists at the point yo = f(xo) and is the mapping

d/-1(2/o) = [df(xo)]-1 = ™(2/o) TR(x0), inverse to df(xo) : TR(xo) -> TR(yo)-

Example. We shall show that arcsin'y = A— for \y\ < 1. The functions

y/i-y2 sin : [—7t/2, 7t/2] -» [—1, 1] and arcsin : [—1, 1] -> [—7t/2, 7t/2] are mutually inverse and continuous (see Example 8 of Sect. 4.2) and sin'(x) = cos# ^ 0 if \x\ < 7t/2. For \x\ < ir/2 we have \y\ < 1 for the values y = sinx. Therefore,  ., 11       1          1

arcsin y = . . = =          = =       .

sin x cosx \/l — sin2 x \/l — y2

The sign in front of the radical is chosen taking account of the inequality cosx > 0 for |x| < 7r/2.

**Notes**

Example. Reasoning as in the preceding example, one can show that

arccos' y = 7 ^ for \y\ < 1 .

Indeed,

, 11 1 1

arccos y =

cos' x sin x y/1 — cos2 x

The sign in front of the radical is chosen taking account of the inequality sin x > 0 if 0 < x < it.

Example. arctan'y = y G R.

Indeed,

/II 2 1 1

arctan y = — = —r- = cos x

tan'* (^) " 1 + tan2x 1 + y* '

Example. arccot'y = —, y <= R.

Indeed

, 1 1 .o 1 1

arccot y = —r— = ;—r- = — sm x = — = —

.

y cot x 1 -f- cot x 1 + y2

\ sin^ x / *

The functions y = f(x) = ax and x = /-1(y) = logay have the derivatives /'(*) = ax In a and (/_1)'(y) =

Assume us see how this is consistent

l 111

ff(x) ax In a ylna '

fil) = (FW¥) = ut)

Example The hyperbolic and inverse hyperbolic functions and their derivatives. The functions .

sinhx = ^(ex — e-x),

cosh# = ^(ex + e-x)

are called respectively the hyperbolic sine and hyperbolic cosine8 of x.

These functions,which for the time being have been introduced purely formally,arise just as naturally in many problems as the circular functions

sin# and cosx.

We remark that

sinh(—x) = — sinhx, cosh(—x) = coshx,

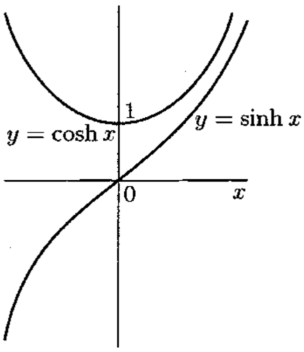
Prom the Latin phrases sinus hyperbolici and cosinus hyperbolici.

that is,the hyperbolic sine is an odd function and the hyperbolic cosine is an even function.

Moreover,the following basic identity is obvious:

cosh2 x — sinh2 x = 1 .

The graphs of the functions y = sinh# and y = cosh# are shown in




y

It follows from the definition of sinh x and the properties of the function ex that sinh# is a continuous strictly increasing function mapping 1 in a one-to-one manner onto itself. The inverse function to sinh# thus exists,is defined on R,is continuous, and is strictly increasing.

This inverse is denoted arsinh y (read "area-sine of y") . This function is easily expressed in terms of known functions. In solving the equation

I(e*-e ~*)=y

for x,we find successively

ex = y + y/l + y2

(ex > 0,and so ex ^ y — y/l.+ y2) and

x = In (y + y/l + y2) .

Thus,

arsinh y = In (y + y/l + y2), ye R .

Similarly,using the monotonicity of the function y = cosh x on the two intervals R_ = {x <= R| x < 0} and R+ = {x <= R|x > 0},we can construct functions arcosh and arcoshdefined for y > 1 and inverse to the function cosh x on R_ and R+ respectively.

They are given by the formulas

arcosh-y = In (y — \fy2 — l),

arcosh = In (y + yV2 - l) .

Prom the definitions given above,we find

sinh' x — ^ (ex + e_x) = cosh a:,

cosh' x = ^ (ex — e-x) = sinhx,

and by the theorem on the derivative of an inverse function,we find

arsinh'y = . ^^ * *

sinh x coshx \/l + sinh2 x \/l + y2

u/ 1 1 1 1 1

arcosh = ——f— = — = = = ,y > 1,

cosh x sinh x — y cosh2 x — 1 yT/2 — 1

!_/ 1 1 1 1 ,

arcosh+y = —-7— = —— =  = =      ,y > 1 .

cosh x sinh x ^/cosh2 x - 1 yV ~ 1

These last three relations can be verified by using the explicit expressions for the inverse hyperbolic functions arsinhy and arcosh y.

For example,

arsinh'y =        ) (l + hi + y2)~1/2 ■ 2y) =

y + ^l + y2K- 2           '

_ 1 \/i + y2 + y _ 1

y + y/1 + y2 \/l + y2 y/l + y2

Like tan x and cot x one can consider the functions

. sinh x . . cosh x tanh x = —-— and coth x = —, cosh x      sinh x

called the hyperbolic tangent and hyperbolic cotangent respectively,and also the functions inverse to them,the area tangent

artanhy = Ln l-tl., |y| < 1,

2 1 - y

and the area cotangent

arcoth y = Ln  \y\ > 1 .

z y l

We omit the solutions of the elementary equations that lead to these formulas.

By the rules for differentiation we have

sinh' x cosh x — sinh x cosh' x

tanh'x =

cosh2 x

cosh x cosh x — sinh x sinh x 1

Notes

cosh2 x cosh2 x '

cosh' x sinh x — cosh x sinh' x _ sinh2 x

sinh x sinh x — cosh x cosh x

sinh2 x sinh2 x

By the theorem on the derivative of an inverse function

i/ 1 1 , 2

artanh x = —j— = —=—r = cosh x =

tanh'x f—K—)

\cosh'2 x/

1 1

1 — tanh2 x 1 — y2

i/l 1 , l2

arcoth x = —i— = r = — sinh x =

coth'x (. h)

\ sinh'2 x /

l l

\y\ > i •

coth2 x — 1 y2 — 1'

The last two formulas can also be verified by direct differentiation of the explicit formulas for the functions artanh y and arcoth y.

## 8.5 DERIVATIVES OF THE BASIC ELEMENTARY FUNCTIONS

We now write out the derivatives of the basic elementary functions computed.

**Differentiation of a Very Simple Implicit Function**

Assume y = y(t) and x = x(t) be differentiate functions defined in a neighborhood U(to) of a point to G R. Assume that the function x = x(t) has an inverse t = t(x) defined in a neighborhood V(xq) of Xo = x(to)>

Then the quantity y = y(t), which depends on can also be regarde as an implicit function of x,since y(t) = y(t(x)). Assume us find the derivative of this function with respect to x at the point #o,assuming that x'(to) ^ 0. Using the theorem on the differentiation of a composite function and the theorem on differentiation of an inverse function.

If the same quantity is regarded as a function of different arguments,in order to avoid misunderstandings in differentiation,we indicate explicitly the variable with respect to which the differentiation is carried out,as we have done here.

Example. The law of addition of velocities. The motion of a point along a line is compely determined if we know the coordinate x of the point in our chosen coordinate system (the real line) at each instant t in a system we have chosen for measuring time. Thus the pair of numbers (x,t) determines the position of the point in space and time. The law of motion is written in the form of a function x = x(t). Suppose we wish to express the motion of this point in terms of a different coordinate system (x, ?). For example,the new real line may be moving uniformly with speed —v relative to the first system. (The velocity vector in this case may be identified with the single number that defines it.) For simplicity

we shall assume that the coordinates (0, 0) refer to the same point in both systems; more precisely,that at time t = 0 the point x = 0 coincided with the point x = 0 at which the clock showed t = 0.

Then one of the possible connections between the coordinate systems (x,t) and (x, ?) describing the motion of the same point observed from different oordinate systems is provided by the classical Galilean transformations:

Assume us consider a somewhat more general linear connection

x = ax + fit, . .

t = 1X + 5t.

Notes

assuming,of course,that this connection is invertible that is the determinant of the matrix [ a ^) is not zero.

V7 SJ _

Assume x = x(t) and x = x(t) be the law of motion for the point under observation,written in these coordinate systems.

We remark that,knowing the relation x = x(t), we find by formula that

x(t) = ax(t) + pt, ,u

t(t) = 7x(t) + St, (° >

and since the transformation is invertible,after writing

x = ax + (3t, (.

t = yi + Si, (5'32)

knowing x = x(<=), we find

x{t) = ax(t) + (3t, . .

t(i) = ^x(i) + St. [)

It is clear from relations that for the given point there exist mutually inverse functions t = t(t) and t = t(t).

We now consider the problem of the connection between the velocities

v^ = = and ^ = =

of the point computed in the coordinate systems (x,t) and (x,t) respectively.

Using the rule for differentiating an implicit function and formula we have

dx = f = af + /?

dt" f ^ + S

or

V(h aV(t)+0

m WWTs • <

where t and t are the coordinates of the same instant of time in the systems (x,t) and (x,t). This is always to be kept in mind in the abbreviated notation

V=%±J;

7V+6 v J

for formula

In the case of the Galilean transformations we obtain the classical law of addition of velocities from formula

V = V + v.

It has been established experimentally with a high degree of precision (and this became one of the postulates of the special theory of relativity) that in a vacuum light propagates with a certain velocity c that is independent of the state of motion of the radiating body. This means that if an explosion occurs at time t = t = 0 at the point x = x = 0,the light will reach the points x with coordinates such that x2 = (ct)2 after time t in the coordinate system (x,t), while in the system (x,?) this event will correspond to time t and coordinates x,where again x2 = (ct)2.

Thus,if x2 — c2t2 = 0,then x2 — ct2 = 0 also,and conversely. By virtue of certain additional physical considerations,one must consider that,in general

x2 — c2t2 = x2 — c2t2,

if (x, t) and (x, t) correspond to the same event in the different coordinate systems connected by relation. Conditions give the following relations on the coefficients ce,/?,7,and S of the transformation

2 22 1 or — c 7 = 1,

ol/3 - c27<5 = 0, (5.38)

f _ c252 = _c2 ?

If c = 1,we would have,instead of

2 2i

a -7 = 1,

1 = I

0 a (32 -S2 = -1,

from which it follows easily that the general solution of up to a change of sign in the pairs (ol, /3) and (7, 5)) can be given as

a = cosh cp,7 = sinh </?,(3 = sinh </?,S = cosh (p,

where <p is a parameter.

The general solution of the system then has the form

a /3\ _ (cosh (f c sinh <p \ j 5) sinh (p cosh <p J

and the transformation can be made specific:

x = cosh (p x + c sinh (p t,

t =\sinh (p x + cosh (p t .

This is the Lorentz transformation.

In order to clarify the way in which the free parameter (p is determined,we recall that the x-axis is moving with speed —v relative to the x-axis,that is,the point x = 0 of this axis,when observed in the system (x,t) has velocity —v. Setting x = 0 in we find its law of motion in the system (x, t):

x = — ctanhpt.

Therefore,

v

tanh</? = — .

c

Comparing the general law of transformation of velocities with the Lorentz transformation,we obtain

~ cosh ipV + c sinh (p 1 sinh ipV + cosh (p '

or,taking account of

y + v i +

Formula is the relativistic law of addition of velocities,which for \vV\ c2,that is,as c oo,becomes the classical law expressed by formula

The Lorentz transformation itself can be rewritten taking account of relation in the following more natural form:

^ x 4-vt x =

a/1 "(f)2'

t+%X

t =

\A-ffl

from which one can see that for |v| «c,that is,as c —>• oo,they become the classical Galilean transformations d2 fix) dx2

dx71

Also by convention,f^(x) := /(x).

The set of functions f : E R having continuous derivatives up to order n inclusive will be denoted C(n\E,R), and by the simpler symbol C(n\E), or Cn(E,R) and Cn(E) respectively wherever no confusion can arise. In particular C^°\E) = C(E) by our convention that /^(x) = /(x). Assume us now consider some examples of the computation of higher order derivatives.

Examples

| | | |
|---|---|---|
| /(*) | /'(*) | /"(*) |
| ax | ax In a | ax In2 a |
| ex | ex | ex |
| sinx | cosx | — sinx |

cosx — sinx— cosx

(1 + x)a a(l + x)°

xa axa_1

loSa WI^^"1

In |x| x-1

a(a — 1)(1 + x)a 2 a(a — l)xa"2

^x~2

(-l)rr-2

(n~1)!x-n

(-l)71"1^- l)!x"

Example Leibniz' formula. Assume u{x) and v(x) be functions having derivatives up to order n inclusive on a common set E. The following formula of Leibniz holds for the nth derivative of their product:

(««)(") = jr (™.

771=0 ^ '

Leibniz' formula bears a strong resemblance to Newton's binomial formula,and in fact the two are directly connected.

Proof. For n = 1 formula agrees with the rule already established for the derivative of a product.

If the functions u and v have derivatives up to order n +1 inclusive,then assuming that formula holds for order n,after differentiating the left and right-hand sides,we find

(uu)(n+1) = ^ ™) u^-m+l)v(m) + ^2 (™) u(n-m)v(m+l) = 771=0 ^
771=0 ^

= u("+1 V°) +S ((fc) + (jfe - 1)) u((n+1)~k)v{[k) + U(0)t;(n+1) =

= ]C (H + 1) U((n+1)_fe)^(fe) •

k=0 ^ '

Here we have combined the terms containing like products of derivatives of the functions u and v and used the binomial relation +

(":') Thus by induction we have established the validity of Leibniz' formula.

Example. If Pn(x) = Co 4- c\x H b cnxn,then

Pn(0) = C0,

P'n(x) = a + 2c2x H b ncnxn_1 and P„(0) = a,

Pn(x) — 2c2 4- 3 • 2c3x H b n(n - 1)cnxn~2 and P„(0) = 2!c2,

Pn3Hx) — 3 • 2c3 H n(n — l)(n - 2)cnxn~3 and jP^(0) = 3!c3,

PnnHx) — n(n ~ l)(n - 2) • • • 2cn and P^(0) = n\cn, P^ (x) = 0 for k > n .

Thus,the polynomial Pn(x) can be written as

Pn(x) = Pi°)(0) + 4p, (i)(0)x + IpW(0)X2 + ■ ■ ■ + LpW(0)x"

Example. Using Leibniz' formula and the fact that all the derivatives of a polynomial of order higher than the degree of the polynomial are zero,we can find the nth derivative of f(x) = x2 sinx:

/(n) (x) = sin^ (x) - x2 + (^i^j x - 2x + ^ 2) s"^n~2^ x - 2 =

= x2 sin + + 2nxsin ^x + (n — 1)7^) + ^ — n(n— 1) sin (x + n^)) = = (x2 — n(n — 1)) sin ^x + n^j — 2nx cos ^x + n^j .

Example Assume /(x) = arctanx. Assume us find the values f^n\0) (n = 1, 2, ...).

Since f'(x) =, it follows that (1 + x2)/'(x) = 1.

Applying Leibniz' formula to this last equality,we find the recursion relation

(1 + x2)/(n+1>(x) + 2nx/<n> (x) + n(n - 1)f{n~1) (x) = 0,

from which one can successively find all the derivatives of f(x).

Setting x = 0,we obtain

/(n+1)(0) = —n(n - l)f(n-V(0) .

For n = 1 we find f(2\0) = 0,and therefore f^2n\0) = 0. For derivatives of odd order we have

/(2m+1)(0) = —2m(2m - i)/(2™-i)(0)

and since /'(0) = 1,we obtain

f(2m+1)(0) =

Example: Acceleration. If x = x(t) denotes the time dependence of a point mass moving along the real line,then = x(t) is the velocity of the point,

and then = dd^ = x(t) is its acceleration at time t.

If x(t) = at + /?,then x(t) = a and x(t) = 0,that is,the acceleration in a uniform motion is zero. We shall soon verify that if the second derivative equals zero,then the function itself has the form at + (3. Thus,in uniform motions,and only in uniform motions,is the acceleration equal to zero.

But if we wish for a body moving under inertia in empty space to move uniformly in a straight line when observed in two different coordinate systems,it is necessary for the transition formulas from one inertial system to the other to be linear. That is the reason why the linear formulas were chosen for the coordinate transformations.

Example. The second derivative of a simple implicit function.
Assume y = y(t) and x = x(t) be twice-differentiable functions. Assume that the function x = x(t) has a differentiate inverse function t = t{x). Then the quantity y(t) can be regarded as an implicit function of x,since y = y{t) = y(t(x)).

Assume us find the second derivative yxx assuming that x'{t) ^ 0.

By the rule for differentiating such a function,studied in Subsect. we have

y'

Jx x'

so that

(y'xYt = Q)t = = x'ty't't - x''ty't

x't x't   x't        (x't)z

We remark that the explicit expressions for all the functions that occur here,including yxx,depend on t,but they make it possible to obtain the value of yxx at the particular point x after substituting for t the value t = t(x) corresponding to the value x.

For example,if y = e* and x = In t,then

/ Vt e* + // W-rYt et + tet, _ f

=t{t+1)e

We have deliberately chosen this simple example,in which one can explicitly express t in terms of x as t = ex and,by substituting t = ex into y(t) = e*,find the explicit dependence of y = e®* on x. Differentiating this last function,one can justify the results obtained above.

It is clear that in this way one can find the derivatives of any order by successively applying the formula

„<»,_

Uxn ~  / •

5.

and the definition of the sum of a series that

e* = l + ±x+±x2 + --- + ±xn + ...

for all xGl.

Example We obtain the expansion of the function ax for any a,

0 < a,  a / 1,similarly:

T          In a In2 a o      lnna _

a = 1 + -rrx + ^r-x2 + • • • + —rxn + • • • . 1!2!         n\

**Notes**

Example. Assume f(x) = sinx. We know (see Example 18 of Subsect. 5.2.6) that f^n\x) = sin (x + \n), n G N,and so by Lagrange's formula (5.56) with Xo = 0 and any x E R we find

rn(0;x) = ^ sin ($ + |(n + l))a;n+1,

from which it follows that rn(0;x) tends to zero for any x G M as n —> oo. Thus we have the expansion

sinI = x _ i., 3 + ^ _ ... + + ...

for every x G M.

Example 6. Similarly,for the function f(x) = cosx,we obtain

r^°;a^ = (n + l)! COS + + 1))a;n+1

and

cosx = 1 - ±x2 + ±x4 - ■ ■ ■ + ^jr*2n + • •'

for x G M.

Example. Since sinh'x = coshx and cosh'x = sinhx,formula yields the following expression for the remainder in the Taylor series of /(x) = sinh x:

r«(0;:c)= (^Ti)!/(n+1)(c)a:n+1'

where f(n+1\<=) = sinh<= if n is even and /^n+1)(<=) — cosh<= if n is odd. In any case |/^n+1)(<=)l < max {| sinhx|, |coshx|},since |<=| < |x|. Hence for any given value x G M we have rn(0;x) —> 0 as n —> oo,and we obtain the expansion

sinhrr = x + + ^x5 + ■ ■ ■ + (2w+1),s2m+1 + • • • > valid for all x G M.

Example . Similarly we obtain the expansion

cosh# = 1 + TyX2 + ^yx4 H h ^)jx2n ^ '

valid for any x G M.

Example . For the function /(x) = ln(l+x) we have f^n\x) = ^~1|1+x^~1^!, so that the Taylor series of this function at Xo = 0 is .

ln(l + x) = x — ^x2 + ^x3 h -—— xn + rn(0; x) .

Z o 72

This time we represent rn(0; a;) using Cauchy's formula

1 (—l)"n!

rn(0;^) = ^(i + g)n(a:-0 ar,

or

<x —

T

where <= lies between 0 and x.

If |x| < 1,it follows from the condition that <= lies between 0 and x that

*|-|<=I/M-|<=|, i — M /, i-|«l i,,

|TkT s th«r = 1" Tn«i s 1" ihoi = w •

!+<<=

Thus for |rr| < 1

kn(0;a:)| < |a:|ra+1,

and consequently the following expansion is valid for |x| < 1:

ln(l + x) = x — ^x2 + ^x3 h -—— xn H

Z o 72

We remark that outside the closed interval |x| < 1 the series on the right hand side of diverges at every point,since its general term does not tend to zero if Ixl > 1.

Example For the function (1 + x)a,where a G M,we have f^n\x) = a(a — 1) • • • (a —n + l)(l + x)a_n,so that Taylor's formula at x$ = 0 for this function has the form

(!+*)« = l + <=* + 5fciW..

... + a(a-l)-(» -, + l)in +

Notes

n\

Using Cauchy's formula (5.55), we find

rn(0;X) = a(tt-l)-(«-")(1 + ^a-n-l(a. _ ^nx ^ (g ?2)

where <= lies between 0 and x.

If |x| < 1,then,using the estimate,we have

|rn(0;®)| < |a(l - j) • • • (l - <=) |(1 + •

When n is increased by 1,the right-hand side of is multiplied by |(l — ^i)x\. But since |x| < 1,we shall have |(l — < Q < 1,

independently of the value of a,provided \x\ < q < 1 and n is sufficiently large.

It follows from this that rn(0;x) —> 0 as n —> oo for any and any

x in the open interval |x| < 1. Therefore the expansion obtained by Newton (Newton's binomial theorem) is valid on the open interval |x| < 1:

/., a a(a — I) 9 a(a — 1) • • • (a — n + 1) ,,

(1+s)a = l + -ra;H——-x +• • •+ — ) -x +• ■ ■

1! 2! n\

We remark that d'Alembert's test implies that for |x| > 1 the series generally diverges if a <= N. Assume us now consider separately the case when a = n E N.

In this case f(x) = (1 + x)a = (1 + x)n is a polynomial of degree n and hence all of its derivatives of order higher than n are equal to 0. Therefore Taylor's formula,together with,for example,the Lagrange form of the remainder,enables us to write the following equality:

,n n(n — 1) 9 n(n — 1) • • • 1 ,, /r,

(1 + x)n = 1 + -a; + v x + • • • + -i xn,

which is the Newton binomial theorem known from high school for a naturalnumber exponent:

(i+x)« = i+(!j)x+(;)«»+...+(^)x».

Assume us to estimate the error in computing the important elementary functions using Taylor's formula. Finally,we have obtained the power-series expansions of these functions.

Definition. If the function f(x) has derivatives of all orders n e N at a point #o,the series

f(x o) + Tjf'(xo)(x - «o) + 1- —J^n)(Xo)(x - X0)n + ■■■

1! n\

is called the Taylor series of / at the point xq.

It should not be thought that the Taylor series of an infinitely differentiable function converges in some neighborhood of #o,for given any sequence Co,ci, ...,cn, ... of numbers,one can construct (although this is not simple to do a function f(x) such that f^n\xo) = cn,n G N.

It should also not be thought that if the Taylor series converges,it necessarily converges to the function that generated it. A Taylor series converges to the function that generated it only when the generating function belongs to the class of so-called analytic functions.

Here is Cauchy's example of a nonanalytic function:

e"1/*2, if x ± 0, 0, if x = 0 .

Starting from the definition of the derivative and the fact that xke-i/x _>> o as x ->► 0,independently of the value of k one can verify that 0) = 0 for n = 0, 1, 2,

Thus the Taylor series in this case has all its terms equal to 0 and hence its sum is identically equal to 0,while f(x) ^ 0 if x ^ 0.

In conclusion,we discuss a local version of Taylor's formula.

We return once again to the problem of the local representation of a function / : E —> R by a polynomial

**Notes**

We wish to choose the polynomial Pn(xo; x) = xq-\- C\{x — xo) HK

cn(x — xo)n so as to have

f(x) = Pn(x) + o((x - <=o)n) as x Xo,x e E, or,in more detail,

f(x) = Co + Ci(x - <=o) H h cn(x - x0)n + o((x - x0)n)

as x —> xo,x e E . (5.76)

We now state explicitly a proposition that has already been proved in all its essentials.

Proposition. If there exists a polynomial Pn(xo;x) = Co + ci(x — xo) + — - + Cn(x — x0)n satisfying condition that polynomial is unique.

Proof. Indeed,from relation we obtain the coefficients of the polyno- mial successively and compely unambiguously

Co = lim<=9x_»Xo f(x),

ci = ,

_,. /(x)-[c0+---+cn_i(x-x0)n *]

cn — lim ebx^xo {x-xq)71 '

We now prove the local version of Taylor's theorem.

Proposition. (The local Taylor formula). Assume E be a closed interval having xo G M as an endpoint. If the function f : E —> R has derivatives f'(xo),. • •,f^n\xo) up to order n inclusive at the point xo,then the following representation holds:

f(x) = f(xo) + ^ ^ (x - xo) H + ^ ^°\x - xo)n +

H-o((x — xo)n) as x —> xo,x E E . (5.77)

Thus the problem of the local approximation of a differentiable function is solved by the Taylor polynomial of the appropriate order. Since the Taylor polynomial Pn(xo; x) is constructed from the requirement that its derivatives up to order n inclusive must coincide with the corresponding derivatives of the function / at xo,it follows that f^k\xo)—Pnk\xo; xo) =

0 (k = 0, 1, ...,n) and the validity of formula is established by the following lemma.

Lemma. If a function (p : E ^ R,defined on a closed interval E with endpoint xo,is such that it has derivatives up to order n inclusive at xo and ip(xo) = <pr{xo) = • • • = (p(n\xo) = 0,then (p(x) = o((x — xo)n) as x —>• xo,  x e E.

Proof. For n = 1 the assertion follows from the definition of differentiability of the function (p at xo,by virtue of which

(p{x) = (p(xo) + <p>{xo)(x — xo) + o(x — xo) as x —> xo,x E E,

and,since tp(xo) = <p'(xo) = 0,we have

(p(x) = o(x — xo) as x —> xo,x E E .

Suppose the assertion has been proved for order n = k — 1 > 1. We shall show that it is then valid for order n = k > 2.

**Check your Progress -1**

Discuss The Basic Rules Of Differentiation

______________________________________________________________

______________________________________________________________

______________________________________________________________

Discuss Differentiation Of An Inverse Function

______________________________________________________________

______________________________________________________________

## 8.6 LET US SUM UP

In this unit we have discussed the definition and example of The Basic Rules Of Differentiation And The Arithmetic Operations, Differentiation Of A Composite Function (Chain Rule), Differentiation Of An Inverse Function, Derivatives Of The Basic Elementary Functions

## 8.7 KEYWORDS

1. The Basic Rules Of Differentiation And The Arithmetic Operations Constructing the differential of a given function or, equivalently, the process of finding its derivative, is called differentiation.

2. .Differentiation Of An Inverse Function: The derivative of an inverse function). Assume the functions f : X -» Y and /-1 : Y -» X

## 8.8 QUESTIONS FOR REVIEW

Explain The Basic Rules Of Differentiation

Explain Differentiation Of An Inverse Function

## 8.9 REFERENCES


- Function of Real Variables
- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables


## 8.10 ANSWERS TO CHECK YOUR PROGRESS

The Basic Rules Of Differentiation (answer for Check your Progress - 1 Q)

Differentiation Of An Inverse Function

(answer for Check your Progress - 2 Q)

# UNIT - 9: THE BASIC THEOREMS OF DIFFERENTIAL CALCULUS

**STRUCTURE**

9.0 Objectives

9.1 Introduction

9.2 The Basic Theorems Of Differential Calculus

9.3 Nonlinear Systems Of Equations

9.4 The Inverse Function Theorem

9.5 The Implicit Function Theorem

9.6 Integral Calculus Of Several Variables

9.7 Riemann Volume In Rn

9.8 Riemann Volume In Rn Integrals Over Volumes In Rn

9.9 Basic Properties Of The Integral

9 .10 Integrals Over Rectangular Regions

9.11 Tangent And Normal Vectors

9.12 Let Us Sum Up

9.13 Keywords

9.14 Questions For Review

9.15 References

9.16 Answers To Check Your Progress

## 9.0 OBJECTIVES

After studying this unit, you should be able to:

Learn Understand about The Basic Theorems Of Differential Calculus

Learn Understand about Nonlinear Systems Of Equations

Learn Understand about The Inverse Function Theorem

Learn Understand about The Implicit Function Theorem

Learn Understand about Integral Calculus Of Several Variables

Learn Understand about Riemann Volume In Rn

Learn Understand about Riemann Volume In Rn Integrals Over Volumes In Rn

Learn Understand about Basic Properties Of The Integral

Learn Understand about Integrals Over Rectangular Regions

Learn Understand about Tangent And Normal Vectors

## 9.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

The Basic Theorems Of Differential Calculus, Nonlinear Systems Of Equations, The Inverse Function Theorem, The Implicit Function Theorem, Integral Calculus Of Several Variables, Riemann Volume In Rn, Riemann Volume In Rn Integrals Over Volumes In Rn, Basic Properties Of The Integral, Integrals Over Rectangular Regions, Tangent And Normal Vectors

## 9.2 THE BASIC THEOREMS OF DIFFERENTIAL CALCULUS

We make the preliminary remark that since

E3x—±x o      X — Xq

the existence of (p^k\xo) presumes that the function ^k~l\x) is defined on E,at least near the point xo- Shrinking the closed interval E if necessary we can assume from the outset that the functions </?(x), • • •,ip^k~l\x), where k > 2,are all defined on the whole closed interval E with endpoint xq. Since k > 2,the function ip(x) has a derivative ipf(x) on E,and by hypothesis

(¥),)/(^o) = --- = (v')(fc-1)M = o.

Therefore,by the induction assumption,

ip'(x) = o((x — xo)^"1) as x —> xo,x G E .

Then,using Lagrange's theorem,we obtain

<p(x) = tp{x) - <p(x0) = <p'(€)(x - x0) = <*(<=)(<= - ®o)(fc_1)0c - Xo),

where <= lies between xo and x,that is,|<= — xo| < \x — xo|,and a(<=) —> 0 as <= —y xo,<= G E. Hence as x —> xo,x G E,we have simultaneously <= —> xo, (GjE,and a(<=) —> 0. Since

|<p(x)| < |a(<=)| \x - a;0|fc_1|a: - a:0|,

we have verified that

ip(pc) = o((x — xo)k) as x —y xo,x G E .

Thus,the assertion of Lemma has been verified by mathematical induction. □

Relation is called the local Taylor formula since the form of the remainder term given in it (the so-called Peano form)

rn(x0; x) = o((x - x0)n),

makes it possible to draw inferences only about the asymptotic nature of the connection between the Taylor polynomial and the function as x —> xo, x e E.

Formula is therefore convenient in computing limits and describing the asymptotic behavior of a function as x —> xo,x G E,but it cannot help with the approximate computation of the values of the function until some actual estimate of the quantity rn(xo; x) = o((x — xo)n) is available.

Assume us now summarize our results. We have defined the Taylor polynomial

Pn(x0;x) = f(xo) + ^ ^ (x - x0) + h ^ ~ Xo^ '

written the Taylor formula

f(x) = f(xo) + —jj^(a: - Xo) + h ^ ~ ^o)™ + rn(x0;x),

and obtained the following very important specific form of it:

If f has a derivative of order n + 1 on the open interval with endpoints xo and x,then

/(z) = f(xo) + ^7^- JV,) H + ^ T^°1 (■>: - m)" +

+Qf('"ir', S', S|

where ^ is a point between xo and x.

If f has derivatives of orders up to n> 1 inclusive at the point xo,then

/(x) = f(xo)+ f ^ (ar-a;o) + - • •+ ^ (x-x0)n+o({x-x0)n) .

Relation called Taylor's formula with the Lagrange form of the remainder, is obviously a generalization of Lagrange's mean-value theorem to which it reduces when n = 0.

Relation called Taylor's formula with the Peano form of the remainder, is obviously a generalization of the definition of differentiability of a function at a point, to which it reduces when n = 1.

We remark that formula is nearly always the more substantive of the two. For,on the one hand,as we have seen,it enables us to estimate the absolute magnitude of the remainder term. On the other hand,when,for example,f^n+l\x) is bounded in a neighborhood of #o,it also implies the asymptotic formula

f(x) = f(xo) + ^ ^ (x — xo) H b ^ ri?°^X ~ X°^n + °^X ~ :Eo)n+1) •

Thus for infinitely differentiable functions,with which classical analysis deals in the overwhelming majority of cases,formula contains the local formula .

In particular can now write the following table of asymptotic formulas as x —> 0:

e* = 1 + I® + i*2 + • • • + -xn + 0(xn+1),

1! 2! n\

cosx = l-±x2 + ±x4----+ {-^x2n + 0(x2n+2),

sin® = « - ^ + 0(x2n+3),

cosh a; = 1 + ±x2 + ^x4 + ■ ■ ■ + J^x2n + 0{x2n+2),

sinhrr = ® + I*3 + I*3 + ... + + 0(a.2n+3) j

ln(l + x) = X - \x2 + ]-x3 + t^xn + 0(xn+1),

2 3 n

xrv, a a(a — l) o a(a — 1) • • • (a — n + 1) ,, (1 + x)a = 1 + Yf + -L^j—®
+ • • • + — —^ LXn

+ 0(xn+1) .

Assume us now consider a few more examples of the use of Taylor's formula.

Example We shall write a polynomial that makes it possible to compute the values of sin# on the interval —1 < x < 1 with absolute error at most One can take this polynomial to be a Taylor polynomial of suitable degree obtained from the expansion of sin x in a neighborhood of xo = 0. Since

sinx = x - ^a:3 + ^a:5 b (2n + + ° + r2n+2(0;a:) '

where by Lagrange'sformula

we have,for \x\ < 1,

irs"«(0;x)i <= (toTsjt'

But (2n+3)! < 10~3 for n> 2. Thus the approximation sinx « x — ^ + i^x5 has the required precision on the closed interval |x| < 1.

Example We shall show that tanx = x + |x3 + o(x3) as x —> 0. We have

tan' x = cos-2 x,

tan'' x = 2 cos-3 x sin x,

tan''' x = 6 cos-4 x sin2 x + 2 cos-2 x .

Thus,tanO = 0,tan' 0 = 1,tan'' 0 = 0,tan''' 0 = 2,and the relation now follows from the local Taylor formula.

Example. Assume a > 0. Assume us study the convergence of the series oo

E In cos For a > 0 we have —> 0 as n —> oo. Assume us estimate the

na na

n= 1

order of a term of the series:

i 1 i (, 1 1 / 1 \\ 11 / 1 \ mcos — = m 1 — — • ——h o(—— = —- • — —V o —r— .

na V 2! n2a \n2aJJ 2 n2a \n2aJ Thus we have a series of terms of constant sign whose terms are equivalent

CO

to those of the series ^2 2^' Since this last series converges only for a >

n=l

when a > 0 the original series converges only for a >|

Example 14- Assume us show that In cos x = — \x2 — j^x4 — ^x6 + 0(x8) as x —y 0.

This time,instead of computing six successive derivatives,we shall use the already-known expansions of cos# as x —> 0 and ln(l + u) as u —> 0:

In cos x = In ^1 - ^x2 + ^x4 — ^yx6 + 0(x8)) = ln(l + u) =

= u - iu2 + i»3 + 0(ul) = (- I 0(xs)) -

'' 5 ((sf*'''2' dtix°+°(x'>) + 5 ('' Wx''+0(x'>) =

Example Assume us find the values of the first six derivatives of the function In cos x at x = 0.

We have (lncos)'x = ~ s\nf,and it is therefore clear that the function

v ' COS x '

has derivatives of all orders at 0,since cosO ^=- 0. We shall not try to find functional expressions for these derivatives,but rather we shall make use of the uniqueness of the Taylor polynomial and the result of the preceding example.

If

/(x) = Co + c\x H h cnxn + o{xn) as x —> 0,

then

Cfc = ^ ^ and /(fc)(0) = k\ck ■

Thus,in the present case we obtain

(lncos)(0) = 0,(lncos)/(0) = 0,(lncos)//(0) = —^ • 2!,

(In cos) ^(0) = 0,(In cos) ^(0) = • 4!,

(lncos)(5)(0) = 0,(lncos)(6)(0) = —• 6! .

Example. Assume f(x) be an infinitely differentiable function at the point #o,and suppose we know the expansion

f'(x) = c'0 + + • • • + c'nxn + 0(xn+1)

of its derivative in a neighborhood of zero. Then,from the uniqueness of the Taylor expansion we have

(/')(fe)(0) = k\c'k,

and so /^+1^(0) = k\c'k. Thus for the function f(x) itself we have the expansion

/(*) = m + Cp + + • • • + + °(^+2) >

or,after simplification,

f{x) = /(0) + °^x + ^x2 + ■ ■ ■ + ^Y*n+1 + 0(xn+2) .

Example Assume us find the Taylor expansion of the function /(#) = arctan x at 0.

Since /'(#) = ~ (H~x2)_1 = 1 —<=2+#4 h(—l)nx2n+0(x2n+2),

by the considerations explained in the preceding example,

1-^3,1 5 i (-1)" „2n+l i ^2n+3\

fix) = /(0) + -Xx - |®3 + - • • • + + 0(x2n+3)

that is

arctan a; = x - ^a:3 + ^a:5 h ^ ^ x2n+1 + 0(x2n+3) .

3 5 2ti -f~ 1

Example. Similarly,by expanding the function arcsin'x = (1 — x2)-1/2 by Taylor's formula in a neighborhood of zero,we find successively,

(1 + u)~l/2 = 1 + —^-u H 2 ^ ^ —u2 Hh

+ -2(-2-1),"(-S-w+1)un + o(un+1), nl

(l-x2)-'/! = H-ix2 + ^|x4 + ...+

+ I'3^l2n<~1)x2n+0(X'M2)-

1 3 1-3 5

arcsin x = x + ~x + 22 . 2\ ■ 5X + '" +

j 111 x2n+1 + 0(x2n+3)

+ (2n)!!(2n+l)* + tA j'

or,after elementary transformations,

arcsinx = x + + H*5 + • • • • + [(^ ~ + 0{x2n+3) .

3! 5! (2n + l)!

Here (2n — 1)!! := 1 • 3 • • • (2n — 1) and (2n)\\ := 2 • 4 • • • (2n).

Example. Find

arctanx — sinx [x — |x3 + 0(x5)] — [x — ^x3 + 0(x5)]

tanx — arcsinx [x + ^x3 + 0(x5)] — [x + ^x3 + 0(x5)]

-±x3 + 0(x5)

— hm —gN = —1 .

ix3 + 0(x5)

## 9.3 NONLINEAR SYSTEMS OF EQUATIONS

Solving nonlinear problems is hard,no matter if they involve algebraic equations,differential equations or something more exotic. In this chapter we examine some problems in nonlinear algebraic equations. We do this not because these problems are terribly important (though they are). We do it because these problems will give us concrete and rigorous examples of the following approach to nonlinear problems that is useful in a number of different contexts.

We find a specific solution to our nonlinear problem.

We compute the linear approximation of the nonlinear problem at that solution.

We determine whether the "linearized" problem has unique solutions.

If the linearized problem has unique solutions we show that the nonlinear problem has unique solution for problems "close" to the solution found in the first step.

## 9.4 THE INVERSE FUNCTION THEOREM

How do we solve a system of n nonlinear equations in n unknowns of the form Of course, there is no truly general answer. That is unfortunate, since this is exactly the form of a coordinate transformation. We would like to be able to determine if such a transformation is invertible (both one-to-one and onto). As we will see, we won't be able to give a general answer to the problem, but we will get something reasonably close.

Before approaching the nonlinear problem, assume's consider the linear case. When can we solve a system of the form

$f(x) = Ax = p$

where A is an nx n matrix? Fortunately,this is just the situation which tells us that this problem has a unique solution for every p e Rn (that is, the function f is invertible) exactly when the n x n matrix A is invertible.

How does this help us in solving the general nonlinear problem? Well, we know that we can define a linear approximation of any nonlinear function. Since we know a lot about solving linear problems assume us see what happens when we replace f with its linear approximation. Suppose f(xo) = po. That is, we have a solution x0 for a particular p0. The linear approximation of f at x0 is defined to be

lf(xo; x) = f(xo) + Df (xo)(x - xo) = po + Df (xo)(x - xo).

So the approximate linear problem to f (x) = p is given by

lf (xo; x) = p

which reduces to

Df(xo)(x - xo) = (p - po).

This has a unique solution for every p if and only if the n x n matrix Df (xo) is invertible. In this case

x = xo + Df (xo)-1 (p - po).

Of course, one on the most common tests for invertibility of the matrix Df (xo) is to see if its determinant is nonzero. This determinant is important enough to give it a special name.

Definition Assume Q C Rn and suppose f : Q ^ Rn is differentiable at xo e Q. We define the Jacobian of f at xo to be

J f (xo) = det Df (xo).

We also use the notation The inverse function theorem says that this condition for the existence of a solution of the approximate linear problem is sufficient to guarantee that the original nonlinear problem has a unique solution for x and p close to xo and p0.

Example. Consider the system of equations

f <-> = (x2) = (U

Note that the total derivative matrix Df(xo, yo) is

Df(x, y)= (2X -2y

The Jacobian is

(x, y) = 2(x  +,/).

Thus,at every point except the origin the matrix is invertible. The inverse function theorem says that other than at the origin ,if f(x0,y0) = (u0, v0) then for (u,v) sufficiently close to (uo,vo) there exists a unique (x,y) close to (x0,y0)such that

f (x, y) =(u, v).

Note that this does not preclude the possibility that there may be more than one solution "far away" from the original solution. In fact, we have

f (x, y) = f (-x- -y)

so there is always another solution on the other side of the origin.

## 9.5 THE IMPLICIT FUNCTION THEOREM

The implicit function theorem concerns the problem of "solving" algebraic systems where there are more unknowns than equations,say n equations in n + k unknowns. The equations would have the form

f (u,v) = 0

where f G Rn,u G Rk and v G Rn. We usually refer to such systems as "underdetermined." We don't expect to be able to solve for all of the unknowns uniquely. The best we can hope for is to solve for n of the unknowns in terms of the remaining k unknowns. If we can do this, we say that our n equations define an "implicit" function v = g(u) from k unknowns u to the remaining n unknowns v.

As we did in the previous section assume us examine the linear case, both to get some ideas on reasonable conditions for the existence of a solution and to introduce some new notation. A general linear problem of n equations in n + k unknowns can be written in the form

Notes

Au + B v = b

where A is an n x k matrix,u G Rk,B is an n x n matrix,v G Rn,and b G Rn. We think of A,B and b as constant and u and v as unknown. Once again, the issue of solvability can be addressed directly by These equations can be solved uniquely by an implicit function v = g(u) if and only if B is invertible. In this case we can define

v = g(u) = -B-1Av + B-1b,

and a simple computation shows that

Au + Bg(u) = b.

As above,we wish to apply our conditions for the solution of the linear problem to the linear approximation to the general nonlinear problem. To do this we will introduce some new notation. Given an n x k matrix A and an n x n matrix B we can define an n x (n + k) partitioned matrix or block matrix

by

(an ■ ■ ■ a\k bii ... bin

C = (A B) =

and given u G Rk and v G Rn we can define a partitioned vector in Rn+k by

(ui \

Uk vi

\Vn/

Then the system of equations above can be written in the form

Cx = (A B) (^ U^) = Au +

When the independent variable of a nonlinear function is written as a partitioned vector it is natural to write the total derivative matrix of the function as a partitioned matrix. For instance, suppose that as above we write a function f : Rn+k ^ Rn as

f (u,v)

where f G Rn,u G Rk and v G Rn. We define Duf(u0,v0) to the n x k matrix and Dv f (u0,v0) to be the n x k matrix f and the linear approximation of f can be written.

lf ((uo,vo); (u,v)) = Duf(uo,vo)(u - uo) + Dvf (uo,vo)(v - vo) + f(uo,vo). Looking ahead a bit,we define the Jacobian determinant

i<=(uo, vo) ••• f^vo)

d(/l,. . ., fn) (   ^

-tt7      T (^ vo)

d(vi ..., Vn)

f (uo, vo) ••• dVn(u^vo)

We now return to the original nonlinear problem f(u,v) = 0. Suppose we know one solution f(uo,vo) = 0. Then the approximate linear problem at that solution is

lf ((uo,vo); (u,v)) = Duf (uo,vo)(u - uo) + Dvf (uo,vo)(v - vo) = 0.

Comparing to the general linear problem above,we see that if Dvf(uo,vo) is invertible,that is if

d (fl, ..., fn),

(uo,vo) = 0,

d(vi..., Vn)

then we can define

g(u) = vo - Dvf(uo,vo) 1Duf(uo,vo)(u - uo).

Simply plugging in to the approximate linear problem, we see that for any u € Rk this satisfies

lf ((uo,vo); (u,g(u))) = 0.

As in the case of the inverse function theorem, the implicit function theorem says that we can go further. If this condition for solvability of the linearized problem is satisfied, then the original nonlinear problem can be solved "close" to the initial solution.

Theorem (Implicit function theorem). Assume h C Rn+k be the domain of a Ci function f : h ^ Rn. Suppose that there is a uo G Rk and vo G Rn such that the interior point (uo,vo) G 0 satisfies

f (uo,vo) = 0. If in addition,the n x n matrix

(f (uo,vo) ••

Dv f (uo,vo)

v f(uo, vo) ••• f(uo, vo)

is invertible,i.e.

d (fl,..., fn)

(uo,vo) = 0.

d(vi . ..,V„)

then there is a ball V)(uo) C Rk about uo and a continuous function g : V)(uo) ^ Rn such that

and for every u G V)(uo)

## 9.6 INTEGRAL CALCULUS OF SEVERAL VARIABLES

Introduction to integral calculus

In this part of the book we consider integrals of vector and scalar functions defined on Rn. In doing so, we have to confront some real conceptual problems dealing with the wide variety of domains of integration possible as subsets of Rn. This issue never comes up in the calculus of a single variable. In most applications, the only "sensible" subsets of the real line over which we might wish to integrate are simple intervals. However, in Rn there are many useful types of subsets over which to integrate. We will concentrate on three:

n-dimensional volumes,

1-dimensional curves,

(n — 1)-dimensional "surfaces."

The last item might cause you to pause a bit. You should have a good idea of what a two-dimensional surface in R3 looks like. But what is a 4-dimensional surface" in R5? More generally,what do we mean by the "dimension" of a region? How do we define its "area?" Unfortunately,a compe answer to this question is beyond the scope of this book. It involves (at least) study of a subject called "measure theory" that is usually taught in more advanced analysis courses. In order to give the reader the ability to do basic integral calculations with a pretty good understanding of their theoretical basis this text contains the following elements:

Quick sketches of some rigorous definitions of concepts from measure theory, Practical formulas for computation of various integrals

Plausibility arguments (a polite way of saying "bad proofs") connecting our practical formulas with traditional notions of length,area,and volume.

References to texts in measure theory that give comp rigorous proofs of the connection between our practical formulas and the fundamental definitions of the concepts involved.

We begin our study of integral calculus by reviewing the basic results from the calculus of a single variable. Consider a real-valued function defined on a bounded interval f : [a,b] ^ R. We would like to define the definite integral

f b

/ f(x) dx J a

to the be the area between the graph of f and the x-axis.

In order to do this, we are forced to ask ourselves what we really know about the concept of area. If we go back far enough,all definitions of area can be derived from the definition of the area of a rectangle. In elementary integral calculus we define the area under a curve by approximating the area by a collection of rectangles called a Riemann sum. This is usually done in elementary calculus texts by creating a uniform partition of the interval [a,b] by defining

Xj = a + i——— ), i = 0, 1,2, . ..,N.

for N e N. This divides the interval [a,b] into N subintervals [xj_i,Xj]. From each of these subintervals we choose a sample point cj e [xj_1, xj]. Using these we define the Riemann sum

N N

(xi xj_1)

= 1 i=1

This is simply the sum of the area of N rectangles with height f (cj) and width Axj = (xj — xj_1). (We use the convention that area below the x-axis is nega- tive.)

If the limit

N

lim\' f (cj)(xj xj_1)

N

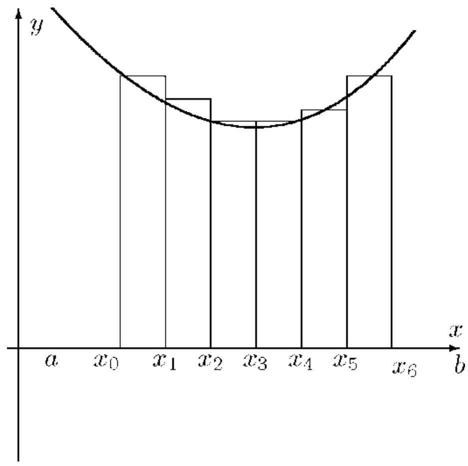
j=1

exists and is independent of the choice of sample points, we say that the function f is Riemann integrable and write

, -b N

/ f (x) dx = lim Y,f (cj)(xj — x^)

J a N




Approximating the area under a curve with a Riemann sum of the area of rectangles.

The obvious question then arises: which functions are Riemann integrable Fortunately,one can show that every continuous function is Riemann integrable. This result can be extended to functions with simple discontinuities. (These results are often stated without proof in elementary texts since a rigorous proof usually uses a concept called "uniform continuity" which is seldom covered in elementary courses.)

Once we know that the definite integral or the area under a curve is well defined for a large class of functions we are left with the problem of trying to calculate it. The fundamental theorem of calculus provides us with a relatively easy way of performing this task. While we won't be discussing vector calculus analogs of the fundamental theorem until Part,we will be using the one-dimensional version to calculate integrals in Rn,so we review it here.

Theorem. Suppose f : [a, b] ^ R is Riemann integrable. If F : [a,b] ^ R satisfies F'(x) = f (x) for all x G [a,b] then

f f (x) dx = F(b) — F(a).

Thus, we can calculate an integral over an interval by finding (guessing really) an "anti-derivative" of the function we are trying to integrate and evaluating it at the boundary points of the interval.

## 9.7 RIEMANN VOLUME IN RN

In this chapter, we define the n-dimensional Riemann volume of a set in Ru. This is a specific example of a measure - a type of function on a set designed to represent the size of a set. More advanced courses on the theory of integration consider more sophisticated measures that can evaluate the size of rather strange sets.

Assume Q C Ru be a bounded region. For a given N G N we create a uniform grid over all of Ru. We define

Xfc.ifc = Nk' ik = 0' ±1' ±2, ...,k = 1,2, ...,n.

This grid of order N divides Ru into rectangles (specifically cubes). For indices ik = 0,±1,±2, ...,±ro,k = 1,2, ...,n we label the rectangles

Rii,i2, ..., in {(x1,x2, . . ., xn) G R 1 xk, ik — 1 — xk — xk, ik,k 1,2, ..., n}. The volume of each n-dimensional rectangle is AVn = Axi, ii Ax2, i2 ... Ax,,, in

= (x1, ii — x1, ii—1)(x2, i2 — x2, i2 —1) . . . {xu, in — xu, in —1) = Nn . We now define the following subsets of the collection of grid rectangles.

We say that Rilii2i...iin is an inner rectangle of h if it lies compely in Q. That is,

Rii,i2,..., in C Q.

We use Ci(Q) to denote the union of all the inner rectangles of Q. We assume Kin(Q) be the number of inner rectangles in the grid of order N. This number must be finite since Q is bounded.

We say that Rii, i2, ..., in is an outer rectangle of Q if there is at least one point in Q inside of Riiii2i...iin. That is,

Rii, i2, ..., in C Q = 0.

We use CO(0) to denote the union of all the outer rectangles of 0. We assume KON (0) be the number of outer rectangles in the grid of order N. Again, this number must be finite since 0 is bounded.

Note the following.

Every inner rectangle is also an outer rectangle. Furthermore, Ci(0) C 0 C CO(0),

and

Ki, n(0) < Ko, n(0).

The volume of CI(0) is simply the sum of the volumes of all the rectangles that are included in the set. Since we have a uniform grid, this has a simple formula

^ AVn = AVn Kin (0).

Ri1, i2, ..., in cCj (n)

Similarly,the volume of CO (0) is given by

^ AVn = AVn Ko, n (0).

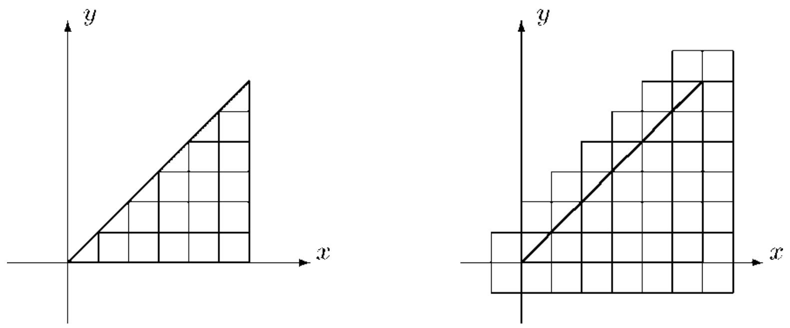
Ri1,i2, ..., in cCo (n)

We can now define the volume of 0.

Definition. If  V(0) = lim AVnKi, n(0) = lim AVNKO, N(0)

N—— ^ ' N—— '

(that is,if both limits exist and are equal) then we say V is the n-dimensional Riemann volume of 0 C Rn.




Inner and outer rectangles of a triangle.

Counting the outer rectangles (don't forget the diagonal row of rectangles that touch the triangle at one corner) gives us

Ko, n(fi) = 4(N + 1)+ N{N~ 1).

We now calculate the 2-dimensional Riemann volume (area)

V (fi) = Jim AVn K', n (fi) = Jim n <NN-11

= lim AVnKo, n(fi) = lim 8(N + ^2^ ~ 1)

n^to ' n^to 2N2

_ 1 = 2.

This is,of course,the same result as the traditional formula for the area of the triangle.

Example. Consider the line

L = {(x,0, 0) C R3|x e (0, 1)}.

To see that this one-dimensional object in R3 has zero 3-dimensional Riemann volume note that there are no inner rectangles in any grid of order N. (So the volume of CI is zero.) The line is surrounded by four rows of outer rectangles so that

Co(L) = {(x,y,z) C R3|x e [-1/N,1 + 1/N)],y, z e [-1/N,1/N]}.

The volume of CO(L) is 4(1 + 2/N)/N2. This goes to zero in the limit,so the 3-dimensional Riemann volume of L is zero.

It is pretty easy to see (if not prove) that all "lower-dimensional" objects in Rn will have zero n-dimensional Riemann volume. Thus,we will need another tool if we are to distinguish between the size of curves,surfaces,and other such objects.

Remark . In this exposition we have used a uniform grid on Rn. This makes our notation slightly easier to read and makes the grid easy to visualize. However,it isn't the most general way of setting up an appropriate grid. In addition,once the more advanced machinery necessary for rigorous proofs of our theorems is set up,the insistence on a uniform grid can make the proofs somewhat harder. All of the definitions above can be adapted to rectangular grids with variable side lengths as long as the length of the longest side goes to zero.

## 9.8 RIEMANN VOLUME IN RN

## INTEGRALS OVER VOLUMES IN Rn

In this chapter we will define the integral of a real-valued function over regions with nonzero n-dimensional volume in Rn.

**BASIC DEFINITIONS AND PROPERTIES**

We begin with the basic definition of the Riemann integral. Assume Q C Rn be a bounded region with well defined, strictly positive Riemann volume. Assume f : Q ^ R be a real-valued function. As in the previous section we define a uniform grid of order N on Rn,and we assume CI(Q) be the collection of inner rectangles contained in Q. In each rectangle Ril, i2, ..., in C Ci(Q) we choose a sample point

cii, i2, ..., in € Rii, i2, ..., in

We can now define a Riemann sum over the inner rectangles

I(Q, f, N,c) = <= f(cii, i2, ..., i„)AVn.

Ci (n)

Here the sum is over the finite collection of rectangles in Ci(Q). Our notation emphasizes that this sum depends on the domain Q,the function f,the grid of order N,and set of sample points c.

Definition. If the limit

lim I(Q,f,N,c)

N

exists and is independent of the choice of sample points, we say that the function f is Riemann integrable on Q. We write

f fdV = lim R(Q, f, N,c).

Jq n^to

Remark It is important to note that our definition of the integral is based on the fundamental notion of the volume of rectangular boxes. Thus,it is crucial that we have used a Cartesian coordinate system to describe the domain and range of the function.

Remark . There are many notations for integrals:

## Notes

We will use dV for the differential element unless we wish to emphasize the dependence of the integral on the Cartesian coordinate system,in which case we will use dV(xi,x2, ...,xn).

In the case of integrals over sets in R2 we will use the symbol dA rather than dV.

A variety of other symbols for the differential element such as

dV — dVn — dx — dxi dx 2 . .. dxn.

Some dispense with it altogether, and in the present context it doesn't really add any information that is not given by the integral sign and the specification of the domain and the function. However, as we collect a variety of types of integrals over different domains,a little redundant information can be helpful.

While we have defined the integral over an n-dimensional volume in Rn using a single integral symbol regardless of the dimension of the domain, it is very common to use two integral signs for Riemann integrals over regions Q C R2

" f dA ~ [ f dA

m jq

and three integral signs for Riemann integrals over regions Q C

fdV - fdV. q jq

While such reminders can be helpful, this can clearly become cumbersome in higher dimensions. Furthermore,the dimension of the integral is usually clear from the nature of the domain.

We will use both notations in this text, choosing the one that seems to make the exposition clearer. (Of course,when we define iterated integrals below, multiple integral signs will become a necessity.)

As was the case for functions of a single real variable, one can show that a large class of functions (continuous functions) are Riemann integrable.

# 9.9 BASIC PROPERTIES OF THE INTEGRAL

The Riemann sum definition of the integral allows us to deduce many basic properties. We will skip most of the proofs for the sake of brevity,but they are not all that difficult and we will note some of the basic ideas.

The first theorem involves the basic property of linearity.

Theorem. Assume Q C Rn have a well defined positive Riemann volume,  and suppose f : Q ^ R and g : Q ^ R are Riemann integrable. Then if a and f are any constants,af + fig is Riemann integrable and

The proof of this follows directly from the definition of a Riemann sum. Each sum has a finite number of terms and each of the terms of the sum is linear in the function f. Thus, we can simply use the distributive law to decompose the sum and then take the limit of both sides.

The next theorem involves splitting the domain of integration up into smaller subsets.

Theorem. Suppose Qi and Q2 are disjoint sets in Rn with well defined positive Riemann volume. Assume

Q = Q i U Q2,

and suppose f : Q ^ R is Riemann integrable. Then f is Riemann integrable over each of the subsets Qi and Q2 and

f fdV = f fdV + (fdV.

J Q      J Qi     J Q2

Similarly,if f is Riemann integrable over each of the subsets Qi and Q2 then f is Riemann integrable union Q and the equation above holds.

While most people consider this to be "obvious, " the proof is a bit more delicate since the inner rectangles of Q don't split neatly into the inner rectangles of Q1 and Q2. We will leave it to more advanced texts. Note,however that it implies that functions with discontinuities at the boundary of Riemann volumes are Riemann integrable if they are integrable over the relevant subsets.

The next theorem involves inequalities between integrals. It states the not too surprising result that the (generalized) volume under the graph of a big function is bigger than the volume under the graph of a small function.

Theorem. Assume Q C Rn have a well defined positive Riemann volume V(Q), and suppose f : Q ^ R and g : Q ^ R are Riemann integrable. If

f (x) < g(x)

at every x G Q then

f fdV < f g dV.

J Q      J Q

In particular,if mi and m2 are constants such that

mi < f (x) < m-2

at every x G Q then

miV (Q) < f fdV < m2V (Q).

Q

The proof of this follows directly from the formula for the Riemann sum. The following result follows immediately from the previous theorem.

Proof. Note if one wants to prove an inequality involving absolute values of the form

|a| < b

one effectively needs to prove two inequalities.

—b < a < b.

In our case this is easy, since by the basic properties of the absolute value we have

— If I < f < If I.

Thus, by the previous theorem

— f If I dV < f fdV < f |f|dV.

•J Q "J Q "J Q

This gives us our result by the observation above.

Our final result is an integral version of the mean value theorem. It say that a continuous function must attain it average value somewhere in the domain of integration.

Theorem. Assume Q C Rn have a well defined positive Riemann volume V(Q). Suppose f : Q ^ R is continuous. Then there is a point xo G Q such that

f fdV = f (xo)V (Q).

Q

# 9.10 INTEGRALS OVER RECTANGULAR REGIONS

While Riemann sums (or more sophisticated methods of estimating integrals) are standard tools for computer calculations,they are not easy to use for hand calculations. Furthermore,they provide us only an estimate for the integral, notits exact value. The next two sections will give us a method of exact calculation of the integral using the one-dimensional version of the fundamental theorem of calculus. We begin with the simplest situation where the domain of the function is a rectangular region

R = {x G Rn|aj < Xi < bj}

and f : R ^ R is a real-valued function.

As with the one-dimensional fundamental theorem,the method here is based on finding antiderivatives. However,in this case we have to find anti-partial derivatives. We say that F : R ^ R is an antiderivative of f with respect to xi if

dF = f

**Notes**

Oxi

For example,if f (x,y,z) = x2y3z then an antiderivative with respect to x is 1 x3y3z while an antiderivative with respect to z is 2x2y3z2,and so on. If we think of the functions xi ^ f with all other variables fixed as functions of one variable, then the elementary fundamental theorem of calculus gives us

/ f (xi, x2, . ..,xi,.. .,x,,) dxi = F(xi, x2, . ..,bi, . ..,x,,)

J ai

-F(xi, x2,.. .,ai,.. .,x,,).

We refer to this calculation as the integral of f with respect to the single variable xi from bi to ai.

We can use this technique of integrating a function of several variables with respect to a single variable to calculate an integral over an n-dimensional rectangle. Our next theorem says two things.

The integral of any Riemann integrable function over an n-dimensional rectangle can be calculated by an iterated integral in which we integrate with respect to each of the n variables - one at a time.

These n integrals can be performed in any order that is convenient. Every order yields the same result.

It's worth remarking that the second part of the theorem had better be true if the first part is to be of any use. It would be pretty disquieting if the calculation of an integral depended on how we numbered the axes of our Cartesian coordinate system.

Theorem (Fubini). If f : R ^ R is Riemann integrable then

/ f (x) dV = • • • / f (xi,.. .,x,,_i, x,,) dxi dx2 ... dx,,

/ R Jan Y J a2 yJ ai J J

Further more, the integrations with respect to the n coordinates can be done in any order with the same result.. However,it is pretty easy to see the general idea of the proof. Essentially, we can group the factors in each term of our Riemann sum so that they are arranged like the appropriate iterated integrals.

For instance, for a two-dimensional example we can write the Riemann sum in the following two ways -

EE f (cil, j2)Ax2 j Axi = EE f(cii, i2)Axi j Ax2. ii\i2/ i2 V ii /

Of course,the trick is to prove rigorously that in the limit as the grid becomes infinitely fine this becomes

pbi / pb2 \ pb2 / pbi \

/ I / f (xi,X2) dx2\dx 1 = I / f (xi,X2) dxi\dx2.

Jai V Ja2 J J<%2 V Jai J

Example. Assume R = {(x,y) G R2|0 < x < 2,1 < y < 3}. We first do an iterated integral with x followed by y.

/ 6x2y dA = / (/ 6x2y d^ I dy

m J1 \J0

f' 3

= / 2x3y|!l0 dy

3

= J 16y dy = 64. Reversing the order of integration gives us the same outcome

J 6x2y dA = J (^ j 6x2y dy^ dx

= (3x2y2|y 3 dx Jo ^ ly=1

2

27x2 — 3x2 dx

0

8x3|0 = 64.

For integration of functions of a single variable, by far the most common domain of integration is an interval - the same type of domain used in the basic definition of the integral. Unfortunately, for functions of several variables, we often wish to integrate over nonrectangular volumes. This causes significant problems in calculating these integrals. In this section we give the reader the tools with which to do the job (though we only describe a few simple applications).

## 9.11 TANGENT AND NORMAL VECTORS

If M is a manifold given as the zero-level set (locally) of $: Rn — R,then we defined the normal space NaM to be the space spanned by V$1(a),...,V$^(a) If M is parametrised locally by F : Rk — Rn (where k +1 = n), then we defined the tangent space TaM to be the space spanned by dF(u),...,ddF(u), where F(u) = a.

We next give a definition of TaM which does not depend on the particular representation of M. We then show that NaM is the orthogonal complement of TaM,and so also NaM does not depend on the particular representation of M.

Definition. Assume M be a manifold in Rn and suppose a G M. Suppose ^: I M is C1 where 0 G I C R,I is an interval and -0(0) = a. Any vector h of the form

h = 0'(0) is said to be tangent to M at A. The set of all such vectors is denoted by TaM.

Theorem. The set TaM as defined above is indeed a vector space.

If M is given locally by the parametrisation F: Rfc — Rn and F(u) = a then TaM is spanned by

dF dF

dUi(u) dUk (u)1°

As in Definition these vectors are assumed to be linearly independent.

If M is given locally as the zero-level set of X:Rn ^ then TaM is the orthogonal complement of the space spanned by

VX1(a),..., VXi(a).

proof: Step 1: First suppose h = b'(0) as in the Definition. Then

¥(b(t)) = 0

for i = 1, ...,<= and for t near 0. By the chain rule

n dX dbj

gdx(a)d"(0) fori: = 1"""

i.e.

W(a) ± b'(0) for i = 1, ...,

This shows that TaM is orthogonal to the space spanned by VX1 (a),...,V$^(a), and so is a subset of a space of dimension n —

Step 2: If M is parametrised by F : Rk ^ Rn with F(u) = a,then every vector

k dF

i=i dui

is a tangent vector

To see this assume

b(t) = F (ui + tai, ..., Un + tan).

Then by the chain rule,

k dF b'(0) = X) ai Qff, (u).

Hence TaM contains the space spanned by dF (u),...,dF (u), and so contains a space of dimension k(= n — <=).

Step 3: From the last line in Steps 1 and 2,it follows that TaM is a space of dimension n — <=. It follows from Step 1 that TaM is in fact the orthogonal complement of the space spanned by VX1(a),...,VX^(a), and from Step 2 that TaM is in fact spanned by dF(u),...,dF (u).

**Check your Progress - 1**

Notes

Discuss Basic Theorems Of Differential Calculus

___

___

___

Discuss Integral Calculus Of Several Variables

___

___

___

## 9.12 LET US SUM UP

In this unit we have discussed the definition and example of The Basic Theorems Of Differential Calculus, Nonlinear Systems Of Equations, The Inverse Function Theorem, The Implicit Function Theorem, Integral Calculus Of Several Variables, Riemann Volume In Rn, Riemann Volume In Rn Integrals Over Volumes In Rn, Basic Properties Of The Integral, Integrals Over Rectangular Regions, Tangent And Normal Vectors

## 9.13 KEYWORDS

1/. The Basic Theorems Of Differential Calculus We make the preliminary remark that since E3x—±x o X — Xq

2. Nonlinear Systems Of Equations: Solving nonlinear problems is hard,no matter if they involve algebraic equations, differential equations or something more exotic

3. The Inverse Function Theorem: How do we solve a system of n nonlinear equations in n unknowns of the form Of course, there is no truly general answer

4.The Implicit Function Theorem : The implicit function theorem concerns the problem of "solving" algebraic systems where there are more unknowns than equations, say n equations in n + k unknowns

5. Riemann Volume In Rn Integrals Over Volumes In Rn In this chapter we will define the integral of a real-valued function over regions with nonzero n-dimensional volume in Rn.

6. Basic Properties Of The Integral The Riemann sum definition of the integral allows us to deduce many basic properties. We will skip most of the proofs for the sake of brevity, they are not all that difficult and we will note some of the basic ideas.

7.Integrals Over Rectangular Regions While Riemann sums (or more sophisticated methods of estimating integrals) are standard tools for computer calculations, they are not easy to use for hand calculations.

8.Tangent And Normal Vectors If M is a manifold given as the zero-level set (locally) of ∃ : Rn — R

## 9.14 QUESTIONS FOR REVIEW

Explain Basic Theorems Of Differential Calculus

Explain Integral Calculus Of Several Variables

## 9.15 REFERENCES


- Application of Several Variables
- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables


## 9.16 ANSWERS TO CHECK YOUR PROGRESS

Basic Theorems Of Differential Calculus

(answer for Check your Progress - 1 Q)

Integral Calculus Of Several Variables

(answer for Check your Progress - 1 Q)

# UNIT - 10 : THE RIEMANN INTEGRAL IN N VARIABLES

**STRUCTURE**


10.0 Objectives

10.1 Introduction

10.2 The Riemann Integral In N Variables

10.3 Integrals Over General Regions In R2

10.4 Change Of Order Of Integration In R2

10.5 Integrals Over Regions In R3

10.6 The Change Of Variables Formula

10.7 Multiple Integral

10.8 Change Of Variables

10.9 Let Us Sum Up

10.10 Keywords

10.11 Questions For Review

10.12 References

10.13 Answers To Check Your Progress


## 10.0 OBJECTIVES

The Riemann Integral In N Variables

Integrals Over General Regions In R2

Change Of Order Of Integration In R2

Integrals Over Regions In R3

The Change Of Variables Formula

# 10.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

The Riemann Integral In N Variables, Integrals Over General Regions In R2,Change Of Order Of Integration In R2,Integrals Over Regions In R3,The Change Of Variables Formula

# 10.2 THE RIEMANN INTEGRAL IN N VARIABLES

We define the Riemann integral of a bounded function f : R ^ R,where R c Rn is a cell,i.e.,a product of intervals R = I1 x ■■■ x In,where Iv = [av,bv] are intervals in R. Recall that a partition of an interval I = [a,b] is a finite collection of subintervals {Jk : 0 < k < N},disjoint except for their endpoints,whose union is I. We can take Jk = [xk, xk+i],where

a = xo < xi < ■ ■ ■ < xN < xN+1 = b.

Now,if one has a partition of each Iv into Jv1 LN ■ -U JV, N(v), then a partition P of R consists of the cells

Ra — J1ai x J2«2 x ■ ■ ■ x Jnan,

where 0 < av < N(v). For such a partition, define

maxsize (P) = max diam Ra,

a

where (diam Ra)2 = l(J1ai)2 + + l(Jnan)2. Here,l(J) denotes the length

of an interval J. Each cell has n-dimensional volume

V (Ra) = l(Jiai) ■■■ l(Jnan).

Sometimes we use Vn(Ra) for emphasis on the dimension. We also use A(R) for V2(R), and, of course,l(R) for V1(R).

We set

^ (f) = <= sup f (x) V(Ra),

a Ra

IP(f)^Rnf f (x) V(Ra).

Ra

a

Note that I_v (f) < Ip (f). These quantities should approximate the Riemann integral of f,if the partition P is sufficiently "fine."

To be more precise, if P and Q are two partitions of R, we say P refines Q and write P y Q,if each partition of each interval factor Iv of R involved in the definition of Q is further refined in order to produce the partitions of the factors Iv,used to define P. It is an exercise to show that any two partitions of R have a common refinement. Note also that

PyQ =* Ip (f) < I Q(f), and Ip (f) > IQ(f).

Consequently,if Pj are any two partitions of R and Q is a common refinement,we have

Ip! (f) < IQ (f) < I Q(f) < Ip2 (f).

Now, whenever f : R ^ R is bounded, the following quantities are well defined:

1(f)= mf 1 p(f), L(f)= sup Ip(f),

pen(R) pen(R)

where n(R) is the set of all partitions of R,as defined above.
I(f) < I(f). We then say that f is Riemann integrable (on R) provided I(f) = I(f), and in such a case, we set

J f (x) dV(x)= I(f)= I(f).

R

We will denote the set of Riemann integrable functions on R by R(R). If dim R = 2,we will often use dA(x) instead of dV(x) For general n,we might also use simply dx.

Proposition. Assume Pv be any sequence of partitions of R such that

maxsize (Pv) = bv 0,

and assume fva be any choice of one point in each cell Rva in the partition Pv. Then,whenever f € R(R),

f f(x) dV(x) = lim ^ f (fva) V(Rva).

/ v^-tt *—'

Ra

This is the multidimensional Darboux theorem. The sums that arise in

are Riemann sums.

Proposition. If fj € R(R) and Cj € R,then c\f\ + c2f2 € R(R), and

/(cifi + c%f2) dV = ci/ fi dV + c2j f2 dV.

R R R

Proposition. If f is continuous on R,then f € R(R).

Proof. As in we have that

maxsize(P) < 5 Ip(f) — I_p(f) < u(5) ■ V(R),

where w(5) is a modulus of continuity for f on R. This proves the proposition.

□

When the number of variables exceeds one,it becomes more important to identify some nice classes of discontinuous functions on R that are Riemann integrable. A useful tool for this is the following notion of size of a set S C R, called content. "upper content" cont+ and "lower content" cont" by

cont+(S) = I (xs), cont"(S) = I (xs),

where xS is the characteristic function of S. We say S has content,or "is contented, " if these quantities are equal,which happens if and only if xS € R(R), in which case the common value of cont+(S) and cont" (S) is

V(S) = y Xs(x) dV(s).

R

It is easy to see that

N

cont+(S) = inf {<= V(Rk) : S C Ri U ■ ■ ■ U Rn^,

k=1

where Rk are cells contained in R. In the formal definition, the Ra in should be part of a partition P of R,as defined above, but if {R1, ...,RN} are any cells in R,they can be chopped up into smaller cells, some perhaps thrown away, to yield a finite cover of S by cells in a partition of R,so one gets the same result.

It is an exercise to see that, for any set S C R,

cont+(S) = cont+(S), where S is the closure of S.

We note that,generally, for a bounded function f on R,

I (f)+1 (1 — f) = V (R).

In particular, given S C R,

cont_(S) + cont+(R\S) = V(R).

Using this together with S and R\S switched, we have

cont_(S) = cont_(S),

O O

where S denotes the interior of S. The difference S\S is called the boundary of S,and denoted bS.

Note that

N

cont_(S) = sup {<= V(Rk) : Ri U • • • U Rn C SI,

fc=i

where here we take {R1, ...,RN} to be cells within a partition P of R,and assume P vary over all partitions of R. Now,given a partition P of R,classify each cell in P as either being contained in R\S,or intersecting bS,or

O

contained in S. Assumeting P vary over all partitions of R,we see that

cont+(S) = cont+(bS) + cont_(S).

In particular,we have:

Proposition. If S C R,then S is contented if and only if cont+(bS) = 0.

If a set <= C R has the property that cont+(<=) = 0,we say that <= has content zero,or is a nil set. Clearly <= is nil if and only if <= is nil. If <=j are nil,1 < j < K,then (jj=i <=j is nil. If S1,S2 C R and S = S1US2,then S = S1US2 and S D S1US2. Hence bS C b(S1) U b(S2). If S1 and S2 are contented,so is S1 U S2.

Clearly,if Sj are contented,so are Sj = R\ Sj.

It follows that,if S1 and S2 are contented,so is S1 n S2 = (Sj U Sj)°. A family F of subsets of R is called an algebra of subsets of R provided the following conditions hold:

R eF,

Sj eF^ S1 U S2 e F,and

S eF^R\S eF.

Algebras of sets are automatically closed under finite intersections also.

Proposition. The family of contented subsets of R is an algebra of sets.The following result specifies a useful class of Riemann integrable func- tions.

Proposition. If f : R ^ R is bounded and the set S of points of discontinuity of f is a nil set,then f € R(R).

Proof. Suppose \f\ < M on R,and take e > 0. Take a partition P of R, and write P = P' U Pwhere cells in P' do not meet S,and cells in P" do intersect S. Since cont+(S) = 0,we can pick P so that the cells in P" have total volume < e. Now f is continuous on each cell in P . Further refining the partition if necessary,we can assume that f varies by < e on each cell in P . Thus

Ip(f) - Ip(f) < [V(R) +2M]e.

This proves the proposition.

To give an example,suppose K C R is a closed set such that bK is nil. Assume f : K ^ R be continuous. Define f : R ^ R by

f(x) = f (x) for x € K,

0 for x € R\K.

Then the set of points of discontinuity of f is contained in bK. Hence f € R(R). We set

JfdV = j f dV.

K R

In connection with this, we note the following fact,whose proof is an exercise. Suppose R and R are cells, with R C R. Suppose that g € R(R) and that g is defined on R,to be equal to g on R and to be 0 on R\R. Then

g €R(R), and J g dV = J g dV.

R R

This can be shown by an argument involving refining any given pair of partitions of R and R,respectively,to a pair of partitions PR and Pr with the property that each cell in PR is a cell in Pr.

The following describes an important class of sets S C Rra that have content zero.

Proposition. Assume <= C Rra_1 be a closed bounded set and assume g : <= ^ R be continuous. Then the graph of g,

G = {(x, g(x)) : x € <=}

is a nil subset of Rra.

Proof. Put S in a cell R0 C Rn-1. Suppose \f\< M on E. Take N € Z+ and set e = M/N. Pick a partition P0 of R0,sufficiently fine that g varies by at most e on each set E n Ra,for any cell Ra €P0. Partition the interval I = [-M, M] into 2N equal intervals Jv,of length e. Then {Ra x Jv} = {Qav} forms a partition of R0 x I. Now,over each cell Ra € P0,there lie at most 2 cells Qav which intersect G,so cont+(©) < 2e • V(R0). Assumeting N ^ tx>,we have the proposition.

Similarly,for any j € {1, ...,n},the graph of Xj as a continuous function of the complementary variables is a nil set in Rn. So are finite unions of such graphs. Such sets arise as boundaries of many ordinary-looking regions in Rn Here is a further class of nil sets.

Proposition. Assume O C Rra be open and assume S C O be a compact nil subset. Assume f : O ^ Rra is a Lipschitz map. Then f (S) is a nil subset of Rra.

Proof. The Lipschitz hypothesis on f is that there exists L < to such that, for p, q € O,

\f (p) - f (q)\ < L\p - q\.

If we cover S with k cells (in a partition), of total volume < a,each cubical with edgesize S,then f (S) is covered by k sets of diameter < L^Jnb,hence it can be covered by k cubical cells of edgesize L^JnS,having total volume < (L^fn)aa. From this we have the (not very sharp) general bound

cont+(f(S)) < (L^n)n cont+(S),

which proves the proposition. □

In evaluating n-dimensional integrals,it is usually convenient to reduce them to iterated integrals. The following is a special case of a result known as Fubini's Theorem.

Theorem. Assume E C Rn_1 be a closed,bounded contented set and assume gj : E ^ R be continuous,with g0(x) < g1(x) on E. Take

Q = {(x, y) € Rn : x € E,go(x) < y < gi(x)}.

Then Q is a contented set in Rn. If f : Q ^ R is continuous,then rsi(x)

V(x) = f (x, y) dy

Jga(x)

is continuous on E,and

J fdVn = J <pdVn-i,

n s

i.e.,

f f (rgi(x) \

fdVn = / / f (x, y) dy\ dVn-\(x).

n I \Jg°(x) )

Proof. The continuity is an exercise in one-variable integration Assume w(5) be a modulus of continuity for g0,g1,and f,and also <p.

We also can assume that w(5) > 5.

Put E in a cell R0 and assume P0 be a partition of R0. If A < g0 < g1 < B, partition the interval [A, B],and from this and P0 construct a partition P of R = R0 x [A,B]. We denote a cell in P0 by Ra and a cell in P by Ral = Ra x J I. Pick points iai € Ral.

O

Write Po = P'0 U P0 U PQ',consisting respectively of cells inside E, meeting bE,and inside R0\E. Similarly write P = P' UP" UP consisting

O

respectively of cells inside Q,meeting bQ,and inside R\Q,as illustrated

For fixed a,assume

z'(a) = : Ral € P and assume z"(a) and zm (a) be similarly defined. Note that

Z (a) = ^ Ra € K

provided we assume maxsize (P) < 5 and 25 < min[gi(x) — g0(x)],as we will from here on.

that

| fB(a)

| V f(iai)l(Ji) — f(x, y) dy < (B — A)u(5), V x € Ra,

JA(a)

l<=z'(a)

where (Jlez, (a) Jl = [A(a), B(a)]. Note that A(a) and B(a) are within 2u(5) of g0(x) and g1(x), respectively, for all x € Ra,if Ra €P'0. Hence,if \f\< M,

B(a)

|| B(a) ||

/ f (x, y) dy — p(x) < 4Mw(5), Vx € Ra.

A(a)

Thus, with C = B — A + 4M,

| Y f (iai)l(Ji) — <p(x) < Cu(5), V x € Ra €P0.

l^z! (a)

Multiplying by Vn-1(Ra) and summing over Ra €P'0,we have

| Y,f (ial)Vn(Ral) — Y ^a)Vn-l(Ra)|< CV(Ro)u(5),

Rai&V RaeP'0

where xa is an arbitrary point in Ra.

<pt ptf *pnt

## 10.3 INTEGRALS OVER GENERAL REGIONS IN R2

We begin with R2. We describe two types of regions over which the calculation of the integral is relatively easy.

Definition. Suppose 0 C R2 has a well defined positive Riemann volume.

If there exist constants a < b and functions yi : [a, b] ^ R and y2 : [a,b] ^ R such that

0 = {(x, y) e R2|a < x < b,yi(x) < y < y2(x)} we say that 0 is simple in the y-direction,or y-simple.

If there exist constants c < d and functions x1 : [c,d] ^ R and x2 : [c,d] ^ R such that

0 = {(x,y) e R2|c < y < d,xi(y) < x < X2(y)} we say that 0 is simple in the x-direction,or x-simple.

While this is the most useful form of the definition,it can be summarized as follows.

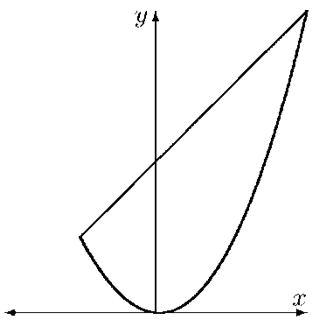
A region is y-simple if The region lie between two vertical lines,Every vertical line between those two lines touches the boundary at either one or two points. A region is x simple if The region lies between two horizontal lines,Every horizontal line between those two lines touches the boundary at either one or two points. Graph a the y-simple domain

0i = {(x,y)|- 1 < x < 2,x2 < y < x + 2}.

Note that this is also an x-simple domain. However,it is much easier to describe as a y-simple domain since the function bounding the domain on the left would have to be "defined piecewise, " using different formulas for different values of y. That is

0i = {(x,y)|0 < y < 4,f (y) < x < ^y},

There is nothing wrong with this,but it can make calculation more difficult.

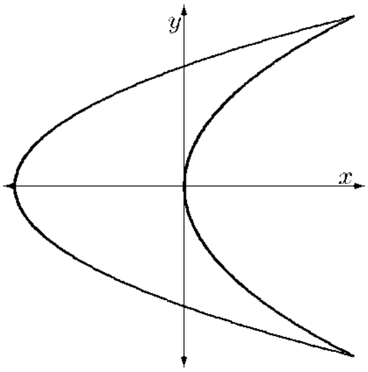


The y-simple region ^ = {(x, y)|— 1 < x < 2,x2 < y < x + 2}.

displays the graph of the x-simple region

^2 = {(x,y)|— 1 < y < 1,2y2 — 1 < x < y2}.

Note that this is not a y-simple region since vertical lines can cross the boundary at up to four places.




The x-simple region ^2 = {(x, y)|— 1 < y < 1,2y2 — 1 < x < y2}.

Our basic theorem is a version of Fubini's theorem given above for rectangular regions.

Theorem. Suppose Q C R has a well defined positive Riemann volume and f : Q ^ R is Riemann integrable.

If Q is y-simple then

r r rb f ry2(x) \

f dA = / / f (x, y) dy dx.

JJq Ja \jyi(x) )

If Q is x-simple then

r r rd (rx2(y) \

f dA = / (/ f (x,y) dx) dy.

JJq Jc \Jx 1 (y))

/2 n x+2

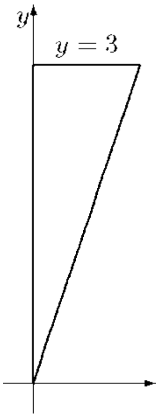
/ 2(x +-lJ x

y=x

^3 i o^2

## 10.4 CHANGE OF ORDER OF INTEGRATION IN R2

Of course,there are lots of situations where a region is simple in both directions. In that case we can compute an iterated integral in either order and get the same




x = 0

y = 3x

x = y x =3

Triangular region of integration Q.

For instance,suppose Q is the triangle

Q = {(x,y)|0 < x < 1,3x < y < 3}.

Of course we can also describe Q as an x-simple region

Q = {(x,y)|0 <y< 3,0 <x< y/3}.

Assume's integrate the function f (x,y) = 12xy2 using the two possible iterated integrals. We start by integrating y before x.

1 r3

12xy dy dx

' 0 J 3x 1

31 y=3 7 4xy I dx

° ly=3x

0

As expected,doing the integration in the other order gives the same result.

As you might expect,sometimes there are advantages to choosing one order of integration over the other. For instance,suppose we wish to integrate the function g(x, y) = 54xcos(y3) over the triangle U given above. One of the iterated integrals

0 3x

cannot be integrated in closed form. However,the other order of integration is tractable.3 rv/3

n3 r3 rv/'3

54x cos(y3) dy dx = 54x cos(y3) dx dy

x J 0 J 0 00 3

7x2 cos(y3)

0

3

= / 3y2 cos(y3) dy

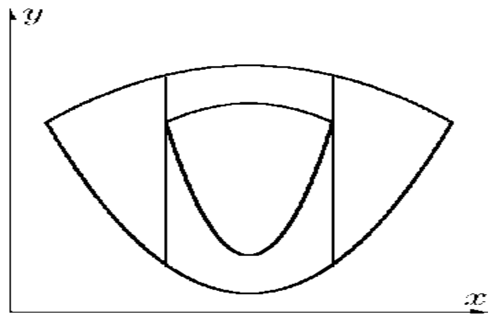
0

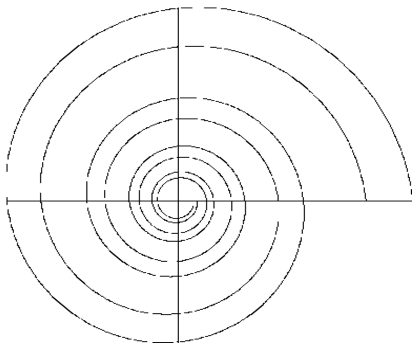
= sin(y3)|0 = sin(27).

Remark At the end of this chapter there are several problems in which you will be asked to do iterated integrals like where you must change the order of integration to do the computation. My best advice to you is always draw a picture of the region of integration. It is always worth the time no matter how obvious you think the change in the limits.

Of course,not all regions in the plane are simple. For such regions, our strategy is to express the domain of integration as the union of a collection of simple regions as illustrated.

Unfortunately,one can easily construct examples of domains that cannot be broken up into a finite collection of simple domains displays a pair of exponentially decaying spiral curves. The region between them cannot be broken up into a finite collection of simple domains since the curves cross both axes infinitely often. Of course,this is rarely a problem in practice.



A non-simple region broken up into four y-simple regions.

The area between the two curves is meant to suggest an infinite spiraling domain that cannot be written as the union of a finite collection of simple domains.

## 10.5 INTEGRALS OVER REGIONS IN R3

In R2 we describe a region as simple if it lies between the graphs of two functions of one variable defined on a common interval. In R3 we describe a region as simple if it lies between the graphs of two functions of two variables with a common domain in a plane and the common domain is a simple region in the plane. Since there are three possible coordinate planes and two possible directions for the planar domain to be simple,there would be six combinations of coordinates for which we could describe a version of Fubini's theorem for a simple region in R3. We will give one version and leave the rest to the reader.

Theorem. Suppose that a region Q G R3 can be described in the following way. There are constants

a < b,

and continuous functions yi : [a,b] ^ R and y2 : [a,b] ^ R with

yi(x) < y2(x)

for all x G [a,b]. These define a domain

Q' = {(x,y) G R2|a < x < b,yi(x) < y < y2(x)}.

On the domain Q' there are two continuous functions zi : Q' ^ R and Z2 : Q' ^ R with

zi(x, y) < Z2(x, y) for all (x,y) G Q. Finally,we can describe

Q = {(x,y,z) G R3|a < x < b,yi(x) < y < y2(x), zi(x, y) < z < Z2(x,y)}
Then if f : Q ^ R is Riemann integrable on Q we have

f dV

Remark. Again,there is nothing special about the order of the coordinates. The same result is obtained for as long as the domain can be described in the way indicated.

Remark. We can think of the common domain Q' as the "shadow" of the volume Q in the xy-plane caused by a light shining down the z-axis. The important thing is that Q have a well defined "top" and "bottom" perpendicular to this axis.

Remark. Note that as we integrate each successive variable,the variable is eliminated from the calculation. Once we integrate with respect to z,the remaining calculation depends only on x and y. Once we integrate with respect to y the remaining calculation depends only on x.

Example. Assume us consider the three-dimensional region inside the cylinder x2 + y2 = 1,below the plane z = 4 + y and above the plane z =

2 + x. Suppose we wish to integrate the function f (x,y,z) = 1 — x over this region. Since we are inside the cylinder,it is easy to identify the "shadow" domain the unit disk in the xy-plane. We can describe our domain as

Q = {(x,y,z)|— 1 < x < 1,— \j 1 — x2 < y < \j 1 — x2,2 + x < z < 4 + y}.

Our integral becomes

/

1 /■ Vl-x2 p4+y

I I 1 — x dz dy dx

-lj-Vl-x2 J2+x /* 1 /* Vl-x2 = 11 (2 + y — x)(1 — x) dy dx

J—1 J-Vl-x2

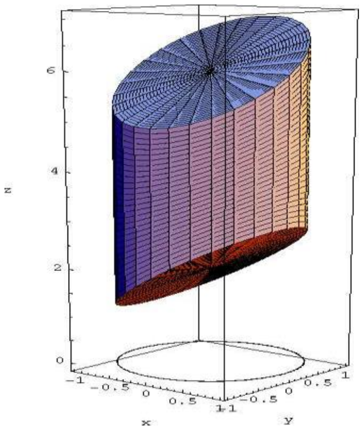
(x2 — 3x + 2)y + (1 — x)y2/2|y=-%/I-x2 dx

J 2(x2 — 3x + 2)^1 — x2 dx 1 l

-(^1 — x2 (8 + 7x — 8x2 + 2x3) + 9 arcsin(x))) 9n

T

.6:



The region inside the cylinder x2 + y2 = 1 above the plane z = 2 + x and below the plane z = 4 + y. Its "shadow" is the unit disk in the yz-plane.

Example. Suppose we wish to find the volume of the region in the first octant bounded by the planes z = y,x = y,and y = 1. We can think of the "shadow" domain as the region q' = {(x,y)|0 < x < 1,x < y < 1}.

Over this domain in the xy-plane,the three-dimensional region is bounded above

by the plane z = y and below by z = 0. Thus,

q = {(x, y, z)|0 < x < 1,x < y < 1,0 < z < y}. We can set up our volume integral as

»i r 1 ry

dz dy dx

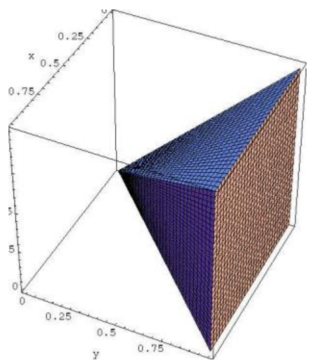
'0 Jx J 0 -1 f 1

y dy dx

0

1 1 1

2(1 - x2) dx = 3




.

The region in the first octant bounded above by the plane z = y and by the planes y = 1 and x = y. Not shown in this figure are the sides x = 0 and y = 0.

Example. As a simple example of using a different order of coordinates consider the problem of trying to find the volume of the sphere

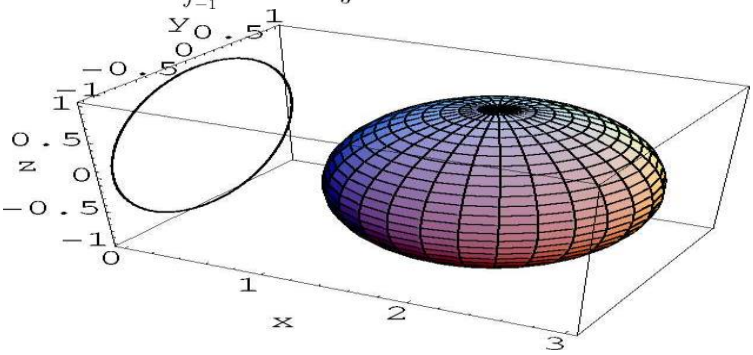
(x - 2)2 + y2 + z2 < 1.

Of course,we could describe this in the same way as above,but instead assume's look at the shadow in the yz-plane and describe the region as

Notes

-1 < y < 1,-^1 - y2 < z < \/1 - y2, 2 - v21 - y2 - z2 < x < 2+ ^1 - y2 -.Our volume integral becomes

$$\begin{aligned}\iiint_\Omega 1\,dV &= \int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\int_{2-\sqrt{1-y^2-z^2}}^{2+\sqrt{1-y^2-z^2}} dx\,dz\,dy\\ &= \int_{-1}^{1}\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} 2\sqrt{1-y^2-z^2}\,dz\,dy\\ &= 2\int_{-1}^{1} \frac{z}{2}\sqrt{1-y^2-z^2}+\frac{1-y^2}{2}\arcsin\left(\frac{z}{\sqrt{1-y^2}}\right)\Bigg|_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dy\\ &= \pi\int_{-1}^{1} 1-y^2\,dy = \frac{4}{3}\pi.\end{aligned}$$




The sphere (x — 2)2 + y2 + z)2 = 1 and its "shadow" the in the yz-plane. The "shadow" is the common domain of the functions describing the hemispheres: x = 2+^/1 — y2 — z2 and x = 2 — \J 1 — y2 — z2 respectively.

## 10.6 THE CHANGE OF VARIABLES FORMULA

One of the most important integration formulas in elementary calculus is the change of variables or "u-substitution" formula.

Theorem. Suppose f : R ^ R is continuous and g : R ^ R is differentiable. Then

rb ra(b)

/ f (g(x))g'(x) dx = / f (u) du. Ja Jg(a)

We make the change of variables by making the formal substitution

u = g(x), du = g'(x) dx.

For example, we can simplify the integral

/■2

J sin(nx3)x2 dx

by making the change

3

u = nx, du = 3nx2 dx,

so that the integral becomes

f2 1 /"8n —1

sin(nx3)x2 dx = — sinu du = — cosu

Ji K ' 3n Jn 3n

In one dimension,it is pretty easy to think of the proof of this theorem with- out worrying about the geometry. In higher dimensions,the geometry is more crucial. The geometric key to the one-dimensional version of the formula is the "fudge factor" g'(u) that relates the length of the grid dx on the x-axis to the grid du on the u-axis.

What is the correct analog for this fudge factor in higher dimensions? For example,suppose we have an invertible transformation

,. (x\ (X(u, v)

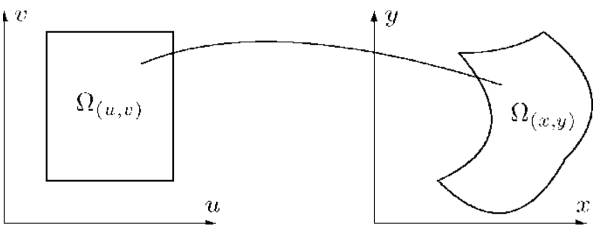
x(u)=l >) = 1«».»)

that maps a region <=l(u, v) in the uv-plane into a region ^(x, y) in the xy-plane. Can we derive a formula analogous to the change of variables formula in one dimension? That is,a formula of the form

f (x,y) dA(x,y) = f (x(u, v),y(u, v)) Fudge Factor dA(u, v).

J J Q/,,,

l(x, y) 'J'J^(u, v)



Transformation from the uv-plane to the xy-plane. This is a typical situation in changing variables in multiple dimensions. The reason for the change is that the domain in the xy-plane is complicated. We have a much simpler rectangular domain in the uv-plane. This is never a consideration in one dimension,where domains are almost always intervals.

In order to make a guess at the fudge factor consider the linear transformation

X = x(u) = Au :

where

ab cd

is a nonsingular matrix. Assume us consider what the transformation does to the domain

^(u, v) = {(x, y)|0 < u < 1,0 < v < 1}.

It is pretty easy to see the following.

• The sides of 0(u, v) transform into lines connecting the respective vertices. For example a point on the line segment (1, t) t G [0, 1] transforms to

(c)+t (d), t G [0, 1]-

The interior of <=l(u, v) transforms into the interior of the parallelogram formed by the vectors (a,c) and (b,d).

Problem asks you to show that the area of a parallelogram formed by the vectors (a,c) and (b,d) was given by the absolute value of the determinant of the matrix A with those vectors as columns. Thus,the square region <=l(u, v) of area one was mapped to a parallelogram ^(x, y) of area|det A|. In fact,one can prove something much more general.

The details above give the basic idea of a proof. We place a uniform grid on &(u, v). The cubic cells of the grid get mapped to similar parallelograms as above,and the ratio between the areas of the transformed parallelograms and the original cubes is the absolute value of the determinant. This can be factored out of sum of the areas of the interior cubes and the the interior parallelograms. In the limit we get the desired relationship.

Remark. Note that the matrix A is the total derivative matrix of the linear transformation X at every point. Thus the determinant of A is the Jacobian of the transformation.

Remark. In fact,the same theorem as above is true for nonsingular linear (affine) transformations from Rn to Rn for any n. While we really haven't studied the tools necessary to prove this in general,it should be fairly obvious in R3 from the relationship between the determinant of a 3 x 3 matrix and the scalar triple product.

Of course,the next step in deriving a general change of variables formula is to go from linear transformations to a general nonlinear transformation X(u). Not to give away the punch line,but here is our basic theorem.

Theorem. Suppose C Rn and C Rn and X : hu ^ is a smooth,invertible transformation. Then if f : ^ R is integrable,the composite function u ^ f (X(u)) is integrable over qu and

$$\iint_{\Omega_{\mathbf{x}}} f(\mathbf{x})\, dV(\mathbf{x}) \quad \iint_{\Omega_{\mathbf{u}}} f(\hat{\mathbf{x}}(\mathbf{u})) \left| \frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)} \right| \, dV(\mathbf{u}).$$

The proof of this follows many of the basic ideas from the previous proof. Again,we break up the domain hu into a regular cubic grid in u-space in order to approximate the integral of the composite f (X) over Qu. However,instead of using a regular cubic grid for the domain we use the curves formed by the transformed coordinate line in u-space. Thus,small cubes in u-space are transformed into small curved regions in x-space. While we can't compute the ratio of the volumes of the corresponding regions exactly,we can use the fact that the nonlinear transformation can be approximated by an affine transformation.

As above,the affine transformation would transform the cube in u-space to a (generalized) parallelogram in x-space. Here the ratio of the volumes is known: the absolute value of the determinant of the total derivative matrix defining the best affine approximation. That is,the ratio of the

Notes

volumes (and hence the fudge factor we have been seeking) is the absolute value of the Jacobian of the transformation.

Example Suppose D(xy) is the parallelogram bounded by the lines

x

2 + 3

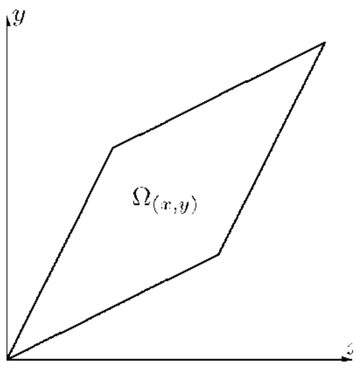
y y

If we wished to compute

I 3y — 6x dA

L(x, y)

directly it would be possible, but difficult. We would have to split the parallel- ogram into simple regions and do more than one double integral.




Domain in the xy-plane with|< y < X +3 and|< x < y +3.

It will be much easier to create a transformation that will represent the domain as the image of a rectangle. There are many transformations that will do this. For instance,if we assume

u = y — 2x, v = 2y — x.

Then the sides of our domain transform as follows.

y= x 2 ~ v = 0,

y= x ~ v = 6,

y= 2x ~ u = 0,

y= 2x — 6 ~ u = —

Thus the equivalent domain in the uv-plane is the square

q(u, v) = {(u, v)|0 < u < 6,0 < v < 6}.

In order to use the change of variables formula we need to compute the

. While there is more than one way to compute this,assume's invert

our transformation to give x and y as functions of u and v. A little linear algebra on equations gives us

1

x = u + v,

3'

12

u +— v.

3 3

This give us the Jacobian

d(x, y) _

d(u,v)

If we note that 3y - 6x = 3u we get Consider the integral(y4 — x4)exy dA(x,y)

D,

(x, y)

where D(x, y) is the region in the xy-plane bounded by the hyperbolic curves xy =1,xy = 2,x2 — y2 = 1,and x2 — y2 = 2.

Here both the integrand and the domain are problematic,but we concentrate on the domain first. It's pretty easy to see that we can define a one-to-one,onto (and hence invertible) map from D(x y) to the square

$D(u, v) = \{(u,v)|1 < u < 2,1 < v < 2\}$

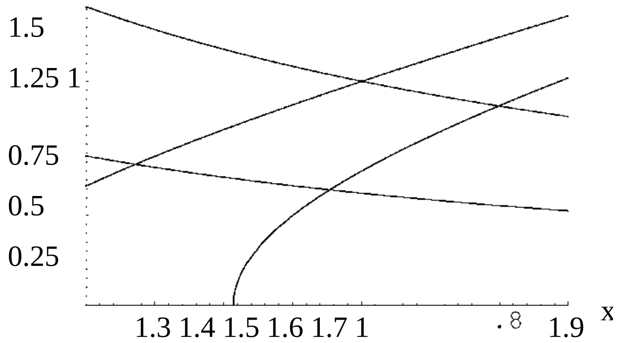
using

$u = xy,$

22

$v = x — y .$

While an inverse map (x(u,v), y(u,v)) exists,it is not necessary for us to find y




The region bounded by the hyperbolic curves $xy = 1, xy = 2$, $x2 — y2 = 1$ and $x2 — y2 = 2$. it explicitly. Instead we note that

d (x, y)

d(u,v)

Example. Consider the integral

/ / sin(x2 + y2) dV(x,y)

3 J D(x, y)

where $D(x, y) = \{(x, y)|x2 + y2 < 1\}$ is the unit disk. Given the circular symmetry of the domain and the integrand it seems sensible to convert to polar coordinates. We use the transformation

Under this transformation,the disk (in the xy-plane) is the image of rectangle

Example . Suppose we wish to find the volume of the region Q(x, y, z) above the cone

z = \j x2 + y2

and below the parabola

z = 2 — x2 — y2.

These two surfaces intersect at the circle x2 + y2 = 1 when z = 1. The common domain Q' of the functions describing the surfaces is the unit circle in the xy-plane. This can be described as a y-simple two-dimensional domain

Q' = {(x, y)|— 1 < x < 1,— \f1 — x2 < y < \f1 — x2}. We can describe

Q(x, y, z) = {(x,y,z) |(x,y) e Q',Jx2 + y2 < z < 2 — x2 — y2}. So we have

V (Q(x, y, z)) =

(x, y, z)

r1 r-vT—x2 /»2 —x2—y2

dz dy dx

- 1 -/ — y/1 — x2 J x2 + y2

This is a rather nasty integral to compute in Cartesian coordinates. However,in cylindrical coordinates it is rather easy. Recall that the cylindrical coordinate transformation is given by x

y

= Pc(r,d,z) =

Under this transformation the cone is given by

z=r

while the parabola is

z = 2 — r

Thus,domain can be described by

q.(r, g, z) = {(r,9,z)|0 < 9 < 2n,0 < r < 1,r < z < 2 — r2}. We compute the Jacobian of the transformation

dx dx dx

d(x,y,z) d(r,9,z)

r sin 9 r cos 9 0

1 = r(cos2 9 + sin2 9) = r. The volume can be computed using the change of variables formula

V (^(x, y, z}) —

(x, y, z)

r dV(r,9,z)

(r, 9, z) c2n r 1 /»2 —

r dz dr d9

/0 ^0 r2n /• 1

p2n pi

(2 — r2 — r) r dr d9

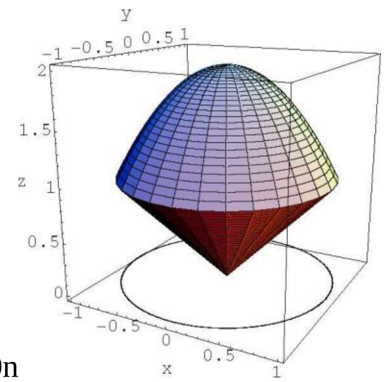
00 f2n

1

r2 _ 1 r4 _ i r3

4 3

0




10n

The region between the cone z = \jx2 + y2 and the parabola z = 2 — x2 — y2. The common domain of the two functions is indicated in the xy plane.

Example. Consider the three-dimensional region U(x, y, z) bounded by the spheres of radius one and two and the cones

z = V3^ x2 + y2

and

z = \jx2 + y2

In we show this region and its cross section in the xz-plane.

While computing the volume of this region as an integral would be a mess to even describe in Cartesian coordinates,it is rather easy in spherical coordinates. Recall that the spherical coordinate transformation is given by

x \ (x(p,0,^>)\(p cos 0 sin ^ \

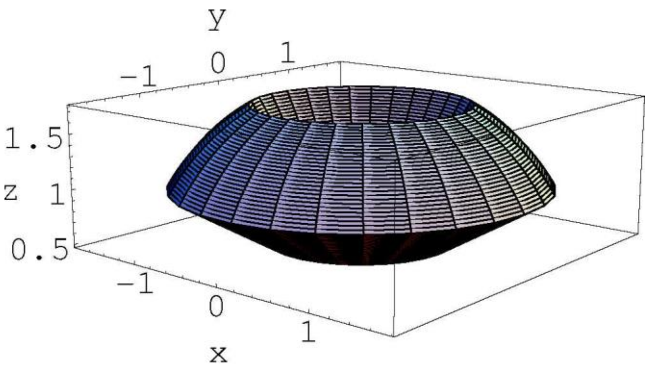
y I = ps(p, 0, ^)= I y(p,0,I = I psin0sin^ I

z) \ z(p, 0, ^) /\pcos $ J

Of course,in this system the spheres of radius one and two are described by the equations p =1 and p = 2 respectively. The cones are described by the equations ^ = n/6 and ^ = n/3 respectively. This can be seen by symmetry or we can determine this analytically as follows. We transform the equation z = a/3^/x2 + y2 into spherical coordinates to get

p cos ^ = V^y p2 cos2 0 sin2 ^ + p2 sin2 0 sin2 ^. Using the fact that p > 0 and sin ^ > 0,this can be reduced to

cos ^ = v3 sin ^




z

x

-2 -1

1 2

Region between the spheres p =1 and p = 2 and the cones 6 = n/6 and 6 = n/3. Both a perspective plot and the cross section of the region in the xz-plane are displayed

or

tan 6 = —=.

V3

Which gives us 6 = n/6. The cone 6 = n/3 can be determined in a similar way. Thus we have

n n ^

6,6)|1 < P < 2,0 < 6 < 2n,— < 6 < _ r .

63

We now compute the Jacobian of the transformation cos e sin ^ —p sin e sin ^ p cos e cos ^ sin e sin ^ p cos e sin ^ p sin e cos ^ cos ^ 0

—p sin ^

—p sin e sin ^ p cos e cos ^ p cos e sin ^ p sin e cos ^

cos e sin ^ —p sin e sin ^ sin e sin ^ p cos e sin ^

= cos ^(—p2 sin2 e sin ^ cos ^ — p2 cos2 e sin ^ cos ^)

—p sin ^(p cos2 e sin2 ^ + p sin2 e sin2 ^)

= —p2 sin ^.

We use this (after taking its absolute value) in the change of variables formula to compute the volume

v (q(x, y, z)) =

(x, y, z)

p sin ^ dV(p,e,^>)

p2 sin ^ dp d^ de

. n n.,23 13. 7n a-

= 2n(c«s 6 — coe 6)("2 — yl-y— 1).

Problem. Consider the system

u = x — y,v = 2x + y.

Solve the system for x and y in terms of u and v. Compute the Jacobian

d (x, y)

d (u,v)

Using this transformation,find the region <=l(u, v) in the uv-plane correspond- ing to the triangular region U(x y) with vertices (0, 0), (1, 1), and (1,—2) in the xy-plane. Sketch the region in the uv-plane.

Use the calculations above to write the integral.

,1,x

/ / 3x dy dx Jo J-2x

as an integral in the uv-plane.

Compute both integrals and show they are the same.

Use the same transformation to evaluate the integral

(2x2 — xy — y2) dx dy

(x, y)

where U(x, y) is the region in the first quadrant bounded by the lines y = —2x+4, y = —2x + 7,y = x — 2,and y = x + 1.

Problem. Consider the change of variables

(x,y) = X(u,v) = (4u,2u + 3v).

Assume U(x, y) = {(x,y)|0 < x < 1,1 < y < 2}.

Find U(u, v) such that X(U(u, v)) = U(x, y),

Use the change of variables formula to calculate

xy dx dy

^(x, y)

as an integral over D(u, v).

Problem. Consider the change of variables

(x,y) = X(u,v) = (u,v(1 + u)).

Assume U(x, y) = {(x,y)|0 < x < 1,1 < y < 2}.

Find U(u v) such that X(U(u v)) = U(x, y),

Use the change of variables formula to calculate

(x — y) dx dy

^ (x, y)

as an integral over U(u v).

Problem. Consider the change of variables

(x,y) = X(u,v) = (u2 — v2,uv).

Assume U(u v) = {(u,v)|u2 + v2 < 1,0 < u}.

Find U(x, y) = X(U(u, v)),

Evaluate

dx dy.

(x, y)

Problem. Convert the following double integrals in Cartesian coordinates to integrals in polar coordinates and evaluate the polar integral.

(a) f^/1 — X2

dy dx.

—~Z 1—X2(b) r1 r\/1—y2

I I xX + y2 dx dy.

1—1J—X1~

y(c) 1 r Z1—

e (x +y) dy dx.

10 J 0

Problem. Convert the integral below to an equivalent integral in cylindrical coordinates and evaluate the integral.

/

1 ryj~y2 rx

/ / (x2 + y2) dz dx dy.

1 0 0

y2 rx

## 10.7 MULTIPLE INTEGRAL

The multiple integral is a definite integral of a function of more than one real variable,for example,f(x,y) or f(x,y,z). Integrals of a function of two variables over a region in $R^2$ are called double integrals,and integrals of a function of three variables over a region of $R^3$ are called triple integralsIntroduction

Just as the definite integral of a positive function of one variable represents the area of the region between the graph of the function and the x-axis,the double integral of a positive function of two variables represents the volume of the region between the surface defined by the function (on the three-dimensional Cartesian plane where z = f(x,y) and the plane which contains its domain. If there are more variables,a multiple integral will yield hypervolumes of multidimensional functions.

Multiple integration of a function in n variables: $f(x_1,x_2,...,x_n)$ over a domain D is most commonly represented by nested integral signs in the reverse order of execution (the leftmost integral sign is computed last), followed by the function and integrand arguments in proper order (the integral with respect to the rightmost argument is computed last). The domain of integration is either represented symbolically for every argument over each integral sign,or is abbreviated by a variable at the rightmost integral sign:

Since the concept of an antiderivative is only defined for functions of a single real variable,the usual definition of the indefinite integral does not immediately extend to the multiple integral.

### MATHEMATICAL DEFINITION

For n > 1,consider a so-called "half-open" n-dimensional hyperrectangular domain T,defined as: Partition each interval $[a_j,b_j)$ into a finite family $I_j$ of non-overlapping subintervals $i_{j\alpha}$,with each subinterval closed at the left end,and open at the right end.

Then the finite family of subrectangles C given by

is a partition of T; that is,the subrectangles $C_k$ are non-overlapping and their union is T.

Assume f : T → R be a function defined on T. Consider a partition C of T as defined above,such that C is a family of m subrectangles $C_m$ and

We can approximate the total (n + 1)th-dimensional volume bounded below by the n-dimensional hyperrectangle T and above by the n-dimensional graph of f with the following Riemann sum:

where $P_k$ is a point in $C_k$ and $m(C_k)$ is the product of the lengths of the intervals whose Cartesian product is $C_k$,also known as the measure of $C_k$.

The diameter of a subrectangle $C_k$ is the largest of the lengths of the intervals whose Cartesian product is $C_k$. The diameter of a given partition of T is defined as the largest of the diameters of the subrectangles in the partition. Intuitively,as the diameter of the partition C is restricted smaller and smaller,the number of subrectangles m gets larger,and the measure $m(C_k)$ of each subrectangle grows smaller. The function f is said to be Riemann integrable if the limit

exists,where the limit is taken over all possible partitions of T of diameter at most δ

If f is Riemann integrable,S is called the Riemann integral of f over T and is denoted

Frequently this notation is abbreviated as

where x represents the n-tuple $(x_1,... x_n)$ and $d^nx$ is the n-dimensional volume differential.

The Riemann integral of a function defined over an arbitrary bounded n-dimensional set can be defined by extending that function to a function defined over a half-open rectangle whose values are zero outside the domain of the original function. Then the integral of the original function over the original domain is defined to be the integral of the extended function over its rectangular domain,if it exists.

In what follows the Riemann integral in n dimensions will be called the multiple integral.

Properties

Multiple integrals have many properties common to those of integrals of functions of one variable (linearity,commutativity,monotonicity,and so on). One important property of multiple integrals is that the value of an integral is independent of the order of integrands under certain conditions. This property is popularly known as Fubini's theorem.[4]

Particular cases

In the case of $T \subseteq R^2$,the integral

is the double integral of f on T,and if $T \subseteq R^3$ the integral

is the triple integral of f on T.

Notice that,by convention,the double integral has two integral signs,and the triple integral has three; this is a notational convention which is convenient when computing a multiple integral as an iterated integral,as shown later in this article**.**

**METHODS OF INTEGRATION**

The resolution of problems with multiple integrals consists,in most cases,of finding a way to reduce the multiple integral to an iterated integral,a series of integrals of one variable,each being directly solvable. For continuous functions,this is justified by Fubini's theorem. Sometimes,it is possible to obtain the result of the integration by direct examination without any calculations.

Integrating constant functions

When the integrand is a constant function c,the integral is equal to the product of c and the measure of the domain of integration. If c = 1 and the domain is a subregion of $R^2$,the integral gives the area of the region,while if the domain is a subregion of $R^3$,the integral gives the volume of the region.

Use of symmetry

When the domain of integration is symmetric about the origin with respect to at least one of the variables of integration and the integrand is odd with respect to this variable,the integral is equal to zero,as the integrals over the two halves of the domain have the same absolute value but opposite signs. When the integrand is even with respect to this variable,the integral is equal to twice the integral over one half of the domain,as the integrals over the two halves of the domain are equal.

**Normal domains on "R"$^2$**

the projection of D onto either the x-axis or the y-axis is bounded by the two values,a and b

any line perpendicular to this axis that passes between these two values intersects the domain in an interval whose endpoints are given by the graphs of two functions,α and β.

Such a domain will be here called a normal domain. Elsewhere in the literature,normal domains are sometimes called type I or type II domains,depending on which axis the domain is fibred over. In all cases,the function to be integrated must be Riemann integrable on the domain,which is true (for instance) if the function is continuous.

**Normal domains on $R^3$**

If T is a domain that is normal with respect to the xy-plane and determined by the functions α(x,y) and β(x,y), then

This definition is the same for the other five normality cases on $R^3$. It can be generalized in a straightforward way to domains in $R^n$.

# 10.8 CHANGE OF VARIABLES

The limits of integration are often not easily interchangeable (without normality or with complex formulae to integrate). One makes a change of variables to rewrite the integral in a more "comfortable" region,which can be described in simpler formulae. To do so,the function must be adapted to the new coordinates.

Example 1a. The function is $f(x,y) = (x - 1)^2 + \sqrt{y}$; if one adopts the substitution $x' = x - 1, y' = y$ therefore $x = x' + 1, y = y'$ one obtains the new function $f_2(x,y) = (x')^2 + \sqrt{y}$.

the differentials dx and dy transform via the absolute value of the determinant of the Jacobian matrix containing the partial derivatives of the transformations regarding the new variable (consider,as an example,the differential transformation in polar coordinates).

There exist three main "kinds" of changes of variable (one in $R^2$,two in $R^3$); however,more general substitutions can be made using the same principle.

**Polar coordinates**

Transformation from cartesian to polar coordinates.

In $R^2$ if the domain has a circular symmetry and the function has some particular characteristics one can apply the transformation to polar coordinates (see the example in the picture) which means that the generic points $P(x,y)$ in Cartesian coordinates switch to their respective points in polar coordinates. That allows one to change the shape of the domain and simplify the operations.

**The Jacobian determinant of that transformation**

which has been obtained by inserting the partial derivatives of $x = \rho \cos(\varphi)$, $y = \rho \sin(\varphi)$ in the first column respect to $\rho$ and in the second respect to $\varphi$,so the dx dy differentials in this transformation become $\rho \, d\rho \, d\varphi$.

Once the function is transformed and the domain evaluated,it is possible to define the formula for the change of variables in polar coordinates:

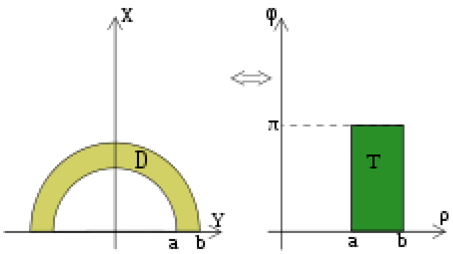
φ is valid in the [0,2π] interval while ρ,which is a measure of a length,can only have positive values.

Example. The function is f(x,y) = x + y and applying the transformation one obtains

Example. The function is $f(x,y) = x^2 + y^2$,in this case one has:

using the Pythagorean trigonometric identity (very useful to simplify this operation).

The transformation of the domain is made by defining the radius' crown length and the amplitude of the described angle to define the ρ,φ intervals starting from x,y.



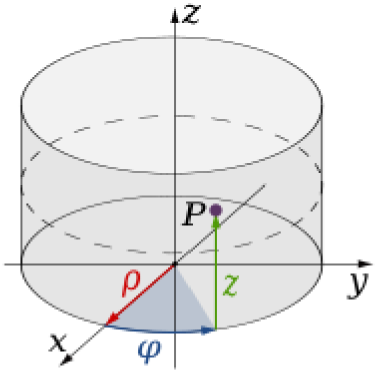

**Example of a domain transformation from cartesian to polar**.

Example. The domain is $D = \{x^2 + y^2 \leq 4\}$,that is a circumference of radius 2; it's evident that the covered angle is the circle angle,so φ varies from 0 to 2π,while the crown radius varies from 0 to 2 (the crown with the inside radius null is just a circle).

Example. The domain is $D = \{x^2 + y^2 \leq 9, x^2 + y^2 \geq 4, y \geq 0\}$,that is the circular crown in the positive y half-plane (please see the picture in the example); φ describes a plane angle while ρ varies from 2 to 3. Therefore the transformed domain will be the following rectangle:

Example  The function is f(x,y) = x and the domain is the same as in Example 2d. From the previous analysis of D we know the intervals of ρ (from 2 to 3) and of φ (from 0 to π).

**Cylindrical coordinates**




Cylindrical coordinates.

In $R^3$ the integration on domains with a circular base can be made by the passage to cylindrical coordinates; the transformation of the function is made by the following relation:

The domain transformation can be graphically attained,because only the shape of the base varies,while the height follows the shape of the starting region.

Example. The region is $D = \{x^2 + y^2 \leq 9, x^2 + y^2 \geq 4, 0 \leq z \leq 5\}$ (that is the "tube" whose base is the circular crown of Example 2d and whose height is 5); if the transformation is applied,this region is obtained:

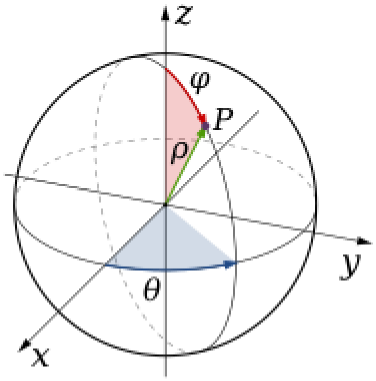
Because the z component is unvaried during the transformation,the dx dy dz differentials vary as in the passage to polar coordinates: therefore,they become ρ dρ dφ dz.

Finally,it is possible to apply the final formula to cylindrical coordinates:

This method is convenient in case of cylindrical or conical domains or in regions where it is easy to individuate the z interval and even transform the circular base and the function.

Example. The function is $f(x,y,z) = x^2 + y^2 + z$ and as integration domain this cylinder: $D = \{x^2 + y^2 \leq 9, -5 \leq z \leq 5\}$. The transformation of D in cylindrical coordinates

**Spherical coordinates**




Spherical coordinates.

In $R^3$ some domains have a spherical symmetry,so it's possible to specify the coordinates of every point of the integration region by two angles and one distance. It's possible to use therefore the passage to spherical coordinates; the function is transformed by this relation:

Points on the z-axis do not have a precise characterization in spherical coordinates,so $\theta$ can vary between 0 and $2\pi$.

The better integration domain for this passage is the sphere.

Example. The domain is $D = x^2 + y^2 + z^2 \leq 16$ (sphere with radius 4 and center at the origin); applying the transformation you get the region

The Jacobian determinant of this transformation is the following:

The dx dy dz differentials therefore are transformed to $\rho^2 \sin(\varphi)\, d\rho\, d\theta\, d\varphi$.

This yields the final integration formula:

It is better to use this method in case of spherical domains and in case of functions that can be easily simplified by the first fundamental relation of trigonometry extended to $R^3$

The extra $\rho^2$ and $\sin \varphi$ come from the Jacobian.

In the following examples the roles of $\varphi$ and $\theta$ have been reversed.

Example. D is the same region as in Example 4a and $f(x,y,z) = x^2 + y^2 + z^2$ is the function to integrate

**DOUBLE INTEGRAL OVER A RECTANGLE**

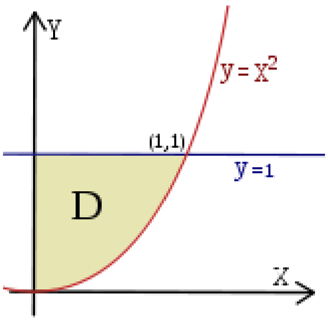
Assume us assume that we wish to integrate a multivariable function f over a region A:

The iterated integral

The inner integral is performed first,integrating with respect to x and taking y as a constant,as it is not the variable of integration. The result of this integral,which is a function depending only on y,is then integrated with respect to y then integrate the result with respect to y.

In cases where the double integral of the absolute value of the function is finite,the order of integration is interchangeable,that is,integrating with respect to x first and integrating with respect to y first produce the same result. That is Fubini's theorem. For example,doing the previous calculation with order reversed gives the same result:

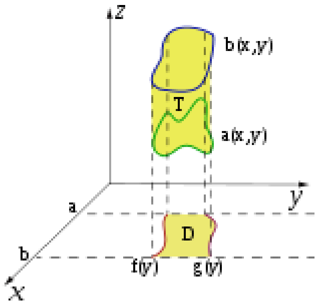
Double integral over a normal domain[edit]




Example: double integral over the normal region D

This domain is normal with respect to both the x- and y-axes. To apply the formulae it is required to find the functions that determine D and the intervals over which these functions are defined. In this case the two functions are:

while the interval is given by the intersections of the functions with x = 0,so the interval is [a,b] = [0,1] (normality has been chosen with respect to the x-axis for a better visual understanding).




Example of domain in $R^3$ that is normal with respect to the xy-plane.

Calculating volume

Using the methods previously described,it is possible to calculate the volumes of some common solids. Cylinder: The volume of a cylinder with height h and circular base of radius R can be calculated by integrating the constant function h over the circular base,using polar coordinates.

Sphere: The volume of a sphere with radius R can be calculated by integrating the constant function 1 over the sphere,using spherical coordinates.

Tetrahedron (triangular pyramid or 3-simplex): The volume of a tetrahedron with its apex at the origin and edges of length ℓ along the x - , y- and z-axes can be calculated by integrating the constant function 1 over the tetrahedron.

**MULTIPLE IMPROPER INTEGRAL**

In case of unbounded domains or functions not bounded near the boundary of the domain,we have to introduce the double improper integral or the triple improper integral.

Multiple integrals and iterated integral

that is,if the integral is absolutely convergent,then the multiple integral

will give the same result as either of the two iterated integrals:

In particular this will occur if |f(x,y)| is a bounded function and A and B are bounded sets.

If the integral is not absolutely convergent,care is needed not to confuse the concepts of multiple integral and iterated integral,especially since the same notation is often used for either concept. The notation means,in some cases,an iterated integral rather than a true double integral. In an iterated integral,the outer integral is the integral with respect to x of the following function of x:

A double integral,on the other hand,is defined with respect to area in the xy-plane. If the double integral exists,then it is equal to each of the two iterated integrals (either "dy dx" or "dx dy") and one often computes it by computing either of the iterated integrals. But sometimes the two iterated integrals exist when the double integral does not,and in some such cases the two iterated integrals are different numbers This is an instance of rearrangement of a conditionally convergent integral. on [0,1] × [0,1] and both iterated integrals exist,then they are equal. Moreover,existence of the inner integrals ensures existence of the outer integrals

**Check your Progress - 1**

Discuss The Riemann Integral In N Variables

________________________________________________________

________________________________________________________

________________________________________________________

Discuss **Multiple Integral_**

_______________________________________________________

________________________________________________________

________________________________________________________

## 10.9 LET US SUM UP

In this unit we have discussed the definition and example of The Riemann Integral In N Variables, Integrals Over General Regions In R2, Change Of Order Of Integration In R2, Integrals Over Regions In R3, The Change Of Variables Formula, **Multiple Integral, Change Of Variables**

## 10.10 KEYWORDS

1. The Riemann Integral In N Variables : We define the Riemann integral of a bounded function f : R ^ R

2.Integrals Over General Regions In R2 We begin with R2. We describe two types of regions over which the calculation of the integral is relatively easy

Change Of Order Of Integration In R2 There are lots of situations where a region is simple in both directions.

3.Integrals Over Regions In R3 In R2 we describe a region as simple if it lies between the graphs of two functions of one variable defined on a common interval.

4.The Change Of Variables Formula: One of the most important integration formulas in elementary calculus is the change of variables or "u-substitution" formula

**5.Multiple Integral :** The multiple integral is a definite integral of a function of more than one real variable,for example,f(x,y) or f(x,y,z).

**6.Change Of Variables:** The limits of integration are often not easily interchangeable (without normality or with complex formulae to integrate).

## 10.11 QUESTIONS FOR REVIEW

Explain The Riemann Integral In N Variables

Explain **Multiple Integral**

## 10.12 REFERENCES

- Analysis of Several Variables
- Application of Several Variables
- Function of Several Variables
- Several Variables
- Function of Variables
- System of Equation
- Function of Real Variables
- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variabless

## 10.13 ANSWERS TO CHECK YOUR PROGRESS

The Riemann Integral In N Variables

(answer for Check your Progress - 1 Q)

**Multiple Integra**

(answer for Check your Progress - 1 Q)

# UNIT -11 CONNECTEDNESS

**STRUCTURE**

11.0 Objectives

11.1 Introduction

11.2 Connected Sets

11.3 Connectedness In R

11.4 Path Connected Sets

11.5 Differentiation Of Real-Valued Functions

11.6 Algebraic Preliminaries

11.7 Partial Derivatives

11.8 Directional Derivatives

11.9 The Differential (Or Derivative)

11.10 The Gradient

11.11 Geometric Interpretation Of The Gradient

11.12 Level Sets And The Gradient

11.13 Mean Value Theorem And Consequences

11.14 Continuously Differentiable Functions

11.15 Higher-Order Partial Derivatives

11.16 Taylor's Theorem

11.17 The General Linear Group GLn(R)

11.18 Let Us Sum Up

11.19 Keywords

11.20 Questions For Review

11.21 References

11.22 Answers To Check Your Progress

## 11.0 OBJECTIVES

After studying this unit you should be able to:

- Understand about Connected Sets

- Understand about Connectedness In R
- Understand about Path Connected Sets
- Understand about Differentiation Of Real-Valued Functions
- Understand about Algebraic Preliminaries
- Understand about Partial Derivatives
- Understand about Directional Derivatives
- Understand about The Differential (Or Derivative)
- Understand about The Gradient
- Understand about Geometric Interpretation Of The Gradient
- Understand about Level Sets And The Gradient
- Understand about Mean Value Theorem And Consequences
- Understand about Continuously Differentiable Functions
- Understand about Higher-Order Partial Derivatives
- Understand about Taylor's Theorem
- Understand about The General Linear Group GLn(R)

## 11.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables Connected Sets,Connectedness In R,Path Connected Sets,Differentiation Of Real-Valued Functions,Algebraic Preliminaries,Partial Derivatives,Directional Derivatives,The Differential (Or Derivative), The Gradient,Geometric Interpretation Of The Gradient,Level Sets And The Gradient,Mean Value Theorem And Consequences,Continuously Differentiable Functions,Higher-Order Partial Derivatives,Taylor's Theorem,The General Linear Group GLn(R)

## 11.2 CONNECTED SETS

## CONNECTEDNESS

Introduction

One intuitive idea of what it means for a set S to be "connected" is that S cannot be written as the union of two sets which do not "touch" one another. Another informal idea is that any two points in S can be connected by a "path" which joins the points and which lies entirely in S.

These two notions are distinct, though they agree on open subsets of Rra

Definition. A metric space (X,d) is connected if there do not exist two non-empty disjoint open sets U and V such that X = U U V.

The metric space is disconnected if it is not connected,i.e. if there exist two non-empty disjoint open sets U and V such that X = U U V.

A set S C X is connected (disconnected) if the metric subspace (S,d) is connected (disconnected).

T2 is y = sin(1 /x) curve

The sets U and V in the previous definition are required to be open in X.

For example,assume

A = [0, 1] U (2,3],

We claim that A is disconnected.

Assume U = [0, 1] and V = (2,3]. Then both these sets are open in the metric subspace (A,d) (where d is the standard metric induced from R). To see this,note that both U and V are the intersection of A with sets which are open in R It follows from the definition that X is disconnected.

In the definition,the sets U and V cannot be arbitrary disjoint sets. For example,R is connected.

But R = U U V where U and V are the disjoint sets (—to,0] and (0,to) respectively.

Q is disconnected. To see this write

Q = (q n (—to,v^)) u (q n (v^

,to)),

The following proposition gives two other definitions of connectedness. Proposition. A metric space (X, d) is connected

iff there do not exist two non-empty disjoint closed sets U and V such that X = U U V;

iff the only non-empty subset of X which is both open and closed is X itself.

proof: Suppose X = U U V where U n V = 0. Then U = X\V and V = X\U. Thus U and V are both open iff they are both closed

In order to show the second condition is also equivalent to connected-ness,first suppose that X is not connected and assume U and V be the open sets.

Then U = X\V and so U is also closed. Since U = 0, X,(2) in the statement of the theorem is not true.

Conversely,if the statement of the theorem is not true assume E C X be both open and closed and E = 0, X. Assume U = E,V = X\E. Then U and V are non-empty disjoint open sets whose union is X,and so X is not connected.

Example We saw before that if A = [0, 1] U (2,3] (c R), then A is not connected. The sets [0, 1] and (2,3] are both open and both closed in A.

## 11.3 CONNECTEDNESS IN R

Not surprisingly,the connected sets in R are precisely the intervals in R. We first need a precise definition of interval.

Definition. A set S C R is an interval if

a, b E S and a < x < b ^ x E S.

Theorem . S C R is connected iff S is an interval.

proof: (a) Suppose S is not an interval. Then there exist a, b E S and there exists x E (a,b) such that x E S.

Then

S = (s n (—to,x) u (s n (x,to)) .

Both sets on the right side are open in S,are disjoint,and are non-empty (the first contains a,the second contains b). Hence S is not connected.

(b) Suppose S is an interval.

Assume that S is not connected. Then there exist nonempty sets U and V which are open in S such that

S = u u V,u n V = %.

Choose a E U and b E V. Without loss of generality we may assume a < b. Since S is an interval,[a,b] C S.

Assume

c = sup([a,b] n U).

Since c E [a, b] C S it follows c E S,and so either c E U or c E V.

Suppose c E U. Then c = b and so a < c < b. Since c E U and U is open, there exists c' E (c,b) such that c1 E U. This contradicts the definition of c as sup([a,b] n U).

Suppose c E V. Then c = a and so a < c < b. Since c E V and V is open,there exists c" E (a,c) such that [c",c] C V. But this implies that c is again not the sup. Thus we again have a contradiction.

Hence S is connected. IRemark There is no such simple characterization in Rra for n > 1.

## 11.4 PATH CONNECTED SETS

Definition. A path connecting two points x and y in a metric space (X, d) is a continuous function f: [0, 1] (C R) ^ X such that f (0) = x and

f (1) = y.

Notes

Definition. A metric space (X,d) is path connected if any two points in X can be connected by a path in X.

A set S C X is path connected if the metric subspace (S,d) is path connected.

The notion of path connected may seem more intuitive than that of connected. However,the latter is usually mathematically easier to work with.

Every path connected set is connected. A connected set need not be path connected,but for open subsets of Rra the two notions of connectedness are equivalent.

Theorem . If a metric space (X,d) is path connected then it is connected.

proof: Assume X is not connected

Thus there exist non-empty disjoint open sets U and V such that X = U U V.

Choose x E U,y E V and suppose there is a path from x to y,i.e. suppose there is a continuous function f: [0, 1] (C R) ^ X such that f (0) = x and f(1) = y.

Consider f-1[U], f-1[V] C [0, 1]. They are open (continuous inverse images of open sets), disjoint (since U and V are), non-empty (since 0 E f-1[U], 1 E f-1[V]), and [0, 1] = f-1 [U] U f-1[V] (since X = U U V). But this contradicts the connectedness of [0, 1]. Hence there is no such path and so X is not path connected.

Br(x) C R2 is path connected and hence connected. Since for u,v e Br (x) the path f: [0, 1] ^ R2 given by f (t) = (1 — t)u + tv is a path in R2 connecting u and v. The fact that the path does lie in R2 is clear, and can be checked from the triangle inequality (exercise).

The same argument shows that in any normed space the open balls Br (x) are path connected,and hence connected. The closed balls {y : d(y, x) < r} are similarly path connected and hence connected.

A = R2\{(0,0), (1,0), (2,0) (3,0),...,(~,0),...} is path connected (take a semicircle joining points in A) and hence connected.

Assume

A = {(x,y) : x > 0 and y = sin X,or x = 0 and y E [0, 1]}.

Then A is connected but not path connected (*exercise).

Theorem . Assume U C Rn be an open set. Then U is connected iff it is path connected.

proof: From Theorem it is sufficient to prove that if U is connected then it is path connected.

Assume then that U is connected.

The result is trivial if U = 0 (why?). So assume U = 0 and choose some a E U. Assume

E = {x E U : there is a path in U from a to x}.

We want to show E = U. Clearly E = 0 since a e E . If we can show that E is both open and closed,it will follow from Proposition that

E = U .

To show that E is open,suppose x e E and choose r > 0 such that Br (x) C U. for each y e Br (x) there is a path in Br (x) from x to y. If we "join" this to the path from a to x,it is not difficult to obtain a path from a to y. Thus y e E and so E is open.

To show that E is closed in U,suppose (xra)^c=1 C E and xn ^ x e U. We want to show x e E. Choose r > 0 so Br (x) C U. Choose n so xn e Br (x). There is a path in U joining a to xn (since xn e E) and a path joining xn to x (as Br(x) is path connected). As before,it follows there is a path in U from a to x. Hence x e E and so E is closed.

Since E is open and closed,it follows as remarked before that E = U, and so we are done.

Basic Results

Theorem . The continuous image of a connected set is connected.

proof: Assume f: X ^ Y,where X is connected.

Suppose f [X] is not connected (we intend to obtain a contradiction).

Then there exists E C f [X],E = f [X],and E both open and closed in f [X]. It follows there exists an open E' C Y and a closed E" C Y such that

E = f [X ] n E' = f [X ] n E'.

In particular,

f-1[E] = f-1 [E ' = f-1 [E"],

and so f-1[E] is both open and closed in X. Since E = 0, f[X] it follows that f-1[E] = 0, X. Hence X is not connected,contradiction.

Thus f [X] is connected.

The next result generalises the usual Intermediate Value Theorem.

Corollary. Suppose f : X ^ R is continuous,X is connected,and f takes the values a and b where a < b. Then f takes all values between a and b.

proof: By the previous theorem,f [X] is a connected subset of R. Then, by Theorem, f [X] is an interval.

Since a, b E f [X] it then follows c E f [X] for any c E [a, b].

## 11.5 DIFFERENTIATION OF REAL-VALUED FUNCTIONS

Introduction

In this Chapter we discuss the notion of derivative (i.e. differential) for func- tions f : D (c Rra) ^ R. In the next chapter we consider the case for functions f: D (c Rra) ^ Rra.

We can represent such a function (m = 1) by drawing its graph,as is done in the first diagrams in Section 10.1 in case n = 1 or n = 2,or as is done "schematically" in the second last diagram in Section 10.1 for arbitrary n. In case n = 2 (or perhaps n = 3) we can draw the level sets,as is done in

Section. Convention Unless stated otherwise,we will always consider functions f: D(c Rra) ^ R where the domain D is open. This implies that for any x e D there exists r > 0 such that Br (x) c D.

## 11.6 ALGEBRAIC PRELIMINARIES

The inner product in Rn is represented by

y • x = ylxl + ... + ynxn

where y = (y1, ...,yn) and x = (x1, ...,xn).

For each fixed y E Rn the inner product enables us to define a linear function

Ly = L : Rn ^ R

given by

L(x) = y •x.

Conversely, we have the following.

Proposition. For any linear function

L: Rn ^ R there exists a unique y e Rn such that

L(x) = y •x Vx E Rn.

The components of y are given by y% = L(ei).

proof: Suppose L: Rn ^ R is linear. Define y = (y1, ...,yn) by

y = L(ei) i =1, ..., n.

Then

L(x) = L(x1 e1 + ••• + xnen)

= x1L(ei) + ••• + xnL(en)

= xlyl + •••+xnyn = y •x

This proves the existence of y satisfying

Notes

The uniqueness of y follows from the fact that if is true for some y,then on choosing x = ei it follows we must have

L(ei)= y i =1, ..., n.

Note that if L is the zero operator,i.e. if L(x) = 0 for all x e Rn,then the vector y corresponding to L is the zero vector.

## 11.7 PARTIAL DERIVATIVES

Definition The zth Partial derivative of f at x is defined by

f (x + tei) - f (x)

df () 8X(x)

t

f (x1 x^ I t xn) f (x1 x^ xn)

-

lim

t0

provided the limit exists. The notation Af (x) is also used.

df

Thus -r-^(x) is just the usual derivative at t = 0 of the real-valued func-

dxiK J J

tion g defined by g(t) = f (x1,..., x% + t, ...,xn). Think of g as being defined along the line L,with t = 0 corresponding to the point x.

## 11.8 DIRECTIONAL DERIVATIVES

Definition. The directional derivative of f at x in the direction v = 0 is defined by

Dvf (x) = lim f (x + tv) - f (x),

t0

provided the limit exists.

It follows immediately from the definitions that

f (x) = Det f(x).

Note that Dvf (x) is just the usual derivative at t = 0 of the real-valued function g defined by g(t) = f (x + tv). As before, think of the function g as being defined along the lin

e L in the previous diagram.

Thus we interpret Dv f (x) as the rate of change of f at x in the direction v; at least in the case v is a unit vector.

Exercise: Show that Dav f (x) = aDv f (x) for any real number a.

## 11.9 THE DIFFERENTIAL (OR DERIVATIVE)

Motivation Suppose f : I (c R) ^ R is differentiable at a E I. Then f'(a) can be used to define the best linear approximation to f (x) for x near a. Namely:

f (x) ~ f (a) + f'(a)(x — a).

f(x)

f(a)+f(a)(x-a)

graph of x |-> f(x)

graph of x |-> f(a)+f(a)(x-a)

Note that the right-hand side of is linear in x.

The error,or difference between the two sides of approaches zero as x ^ a,faster than \x — a\ ^ 0. More precisely

f (x) — (7(a) + f'(a)(x — a)

\x — a\

f (x) — f (a)

We make this the basis for the next definition in the case n > 1

Definition. Suppose f : D (c Rra) a e D if there is a linear function L: Rra —

R. Then f is differentiable at R such that

f (x) - f (a) + L(x - a))

x — a

The linear function L is denoted by f '(a) or df (a) and is called the derivative or differential of f at a. that if L exists.

The idea is that the graph of x ^ f (a) + L(x — a) is "tangent" to the graph of f (x) at the point (a,f (a).

Notation: We write (df (a),x — a) for L(x — a), and read this as "df at a applied to x — a". We think of df (a) as a linear transformation (or function) which operates on vectors x — a whose "base" is at a.

The next proposition gives the connection between the differential operating on a vector v,and the directional derivative in the direction corresponding to v. In particular,it shows that the differential is uniquely defined Temporarily,we assume df (a) be any linear map satisfying the definition for the differential of f at a.

Proposition. Assume v e Rra and suppose f is differentiable at a.

Then Dv f (a) exists and

(df (a),v) = D v f (a)•

In particular,the differential is unique.proof: Assume x = a + tv in  Then f(a + tv) - j(a) + (df (a),tv)

t

lim f (a + tv) - f (a) - a), v) = 0. t^o t

Dv f (a) = (d/ (a),v)Thus (df (a), v) is just the directional derivative at a in the direction v.

The next result shows df (a) is the linear map given by the row vector of partial derivatives of f at a.

Corollary. Suppose f is differentiable at a. Then for any vector v,

n df

(df (a),v) = <= v%—(a).

proof: (df (a), v) = {df (a), vlei + + vnen)

= vl(df (a), ei) + + vn(df (a), en)

= v1 De! f (a) + ... + vnDeJ (a)

Df

Example Assume f (x,y,z) = x2 + 3xy2 + y3z + z.

Then

(df (a), v) = v 1 dx (a)+ v2 dy (a)+ v3 j (a)

= vi(2ai + 3a22) + v2 (6aia2 + 3a22«3) + v3 (a23 + 1).

Thus df (a) is the linear map corresponding to the row vector (2ai+3a22,6aia2+ 3a22a3,a23 + 1).

If a = (1,0, 1) then (df (a), v) = 2vi + v3. Thus df (a) is the linear map corresponding to the row vector (2,0,1).

df

If a = (1,0, 1) and v = ei then (df (1,0, 1), ei) = —(1,0, 1) = 2.

Rates of Convergence If a function ^(x) has the property that (x)|

7 —^ 0 as x —^ a,

xa

then we say "|-0(x)| ^ 0 as x ^ a,faster than |x — a| ^ 0". We write o(|x — a|) for ^(x), and read this as "little oh of |x — a|".

If

W'(x)| < M V|x — a| < e,

xa

U(x) |

for some M and some e > 0,i.e. if is bounded as x ^ a,then we say

| x — a|

"|-0(x)| ^ 0 as x ^ a,at least as fast as |x — a|^ 0". We write O(|x — a|) for ^(x), and read this as "big oh of |x — a|".

For example,we can write

o(|x — a|) for ^ — a^/2,

and

O(|x — a|) for sin(x — a).

Clearly,if ^(x) can be written as o(|x — a|) then it can also be written as O(|x — a|), but the converse may not be true as the above example shows.

The next proposition gives an equivalent definition for the differential of a function.

Proposition If f is differentiable at a then

f (x) = f (a) + <df (a) >x — a) + ^ (x)>

where ^(x) = o(|x — a|).

Conversely,suppose

f (x) = f (a) + L(x — a) + ^(x),

vjhere L : Rra ^ R is linear and ^(x) = o(|x — a|). Then f is differentiable at a and df (a) = L.

proof: Suppose f is differentiable at a. Assume

^(x) = f (x) — (f (a) + <df (a),x — a

Then

f (x) = f (a) + <f (a),x — a) + ^ (x),

and ^(x) = o(|x — a|).

Conversely,suppose

f (x) = f (a) + L(x - a) + ^(x), where L: Rra - R is linear and ^(x) = o(|x — a|). Then

f (x) — (f (a) + L(x — a)) (x)

: = -—-—— — 0 as x — a,

|x — a| |x — a|

and so f is differentiable at a and df (a) = L. I

Remark The word "differential" is used in [Sw] in an imprecise,and different,way from here.

Finally we have:

Proposition. If f, g : D (c Rra) — R are differentiable at a e D, then so are af and f + g. Moreover,

d(af)(a) = adf (a),

d(f + g)(a) = df (a) + dg(a).

proof: This is straightforward The previous proposition corresponds to the fact that the partial derivatives for f + g are the sum of the partial derivatives corresponding to f and g respectively. Similarly for af.

## 11. 10 THE GRADIENT

Notes

Strictly speaking,df (a) is a linear operator on vectors in Rra (where,for convenience,we think of these vectors as having their "base at a").

We saw in Section that every linear operator from Rra to R corresponds to a unique vector in Rra. In particular,the vector corresponding to the differential at a is called the gradient at a.

Definition. Suppose f is differentiable at a. The vector V f (a) e Rra (uniquely) determined by

Vf (a) • v = idf (a),v) Vv e RC

is called the gradient of f at a.Proposition. If f is differentiable at a,then

Vf (a)= (f (a))-

proof: It follows from that the components of Vf (a) df

are (df (a), ei), i.e. —(a).

Example For the example in Section we have

Vf(a) = (2ai + 3a22,6aia2 + 3a22a^,+ 1),

Vf (1,0, 1) = (2, 0, 1).

## 11.11 GEOMETRIC INTERPRETATION OF THE GRADIENT

Proposition Suppose f is differentiable at x. Then the directional derivatives at x are given by

Dvf (x) = V • Vf (x).

The unit vector v for which this is a maximum is v = Vf (x)/\Vf (x)| (assuming \Vf (x)\ = 0), and the directional derivative in this direction is \Vf (x)\.

proof: It follows that

Vf (x) • v = (df (x),v) = Dv f (x)

This proves the first claim.

Now suppose v is a unit vector. From the Cauchy-Schwartz Inequality we have

Vf (x) • v <|Vf (x)\

By the condition for equality in,equality holds in iff v is a positive multiple of Vf (x). Since v is a unit vector,this is equivalent to v = Vf (x)/\Vf (x)\. The left side of is then \Vf (x)\.

## 11.12 LEVEL SETS AND THE GRADIENT

Definition. If f: Rn ^ R then the level set through x is {y: f(y) = f(x) }.

For example,the contour lines on a map are the level sets of the height function.

2 Xi2 - X22 = -.5

x12 - x22 = 0

Xi2 - X22 = 2

Xi2 - X22 = 2

level sets of f(x) = Xi2 - X22 (the graph of f looks like a saddle)

Definition A vector v is tangent at x to the level set S through x if

Dvf (x) = 0.

This is a reasonable definition,since f is constant on S,and so the rate of change of f in any direction tangent to S should be zero.

x12+x22 = 25

Proposition. Suppose f is differentiable at x. Then Vf (x) is orthogonal to all vectors vjhich are tangent at x to the level set through x.

Notes

proof: In the previous proposition,we say Vf (x) is orthogonal to the level set through x.

Some Interesting Examples

An example where the partial derivatives exist but the other directional derivatives do not exist.

Assume

f (x, y) = (xy)1/3.

Then

df

— (0,0) = 0 since f = 0 on the x-axis; df

~q~(0,0) = 0 since f = 0 on the y-axis;

Assume v be any vector. Then

Dvf (0,0) = lim f (tV) - f i0' 0)

J K ' t^0 t

t2/3(ViV2)1/3

= lim

t^o t

v (VlV2)1/3

= im-ti/^ ■

This limit does not exist,unless v1 = 0 or v2 = 0.

An example where the directional derivatives at some point all exist,but the function is not differentiable at the point.

Assume

f(x, y) = (xZ+p <x, y' = (°, 0)

[ 0 (x, y) = (0, 0)

Assume v = (v1,v2) be any non-zero vector. Then

f (tv) - f (0,0)

Dvf (0, 0) = lim

t^0 t

t3v1v22

0

lim t2V12 + t4V24

i;^0 t

2

v1v22

lim

Thus the directional derivatives Dv f (0,0) exist for all v

In particular

f (0, 0) = f (0, 0) = 0.

dx dy

But if f were differentiate at (0,0), then we could compute any directional derivative from the partial drivatives. Thus for any vector v we would have

Dv f (0,0) = {df (0,0) v)

= - If (^ 0) +

= 0 from

This contradicts .

An Example where the directional derivatives at a point all exist,but the function is not continuous at the point

Notes

Take the same example as in. Approach the origin along the curve x = A2,y = A. Then

A4 1

lim f (V, A) = lim 2A4 =5•

But if we approach the origin along any straight line of the form (Avi,Av2), then we can check that the corresponding limit is 0.

Thus it is impossible to define f at (0, 0) in order to make f continuous there.

Differentiability Implies Continuity

If f is differentiable at a,then it is continuous at a.

proof: Suppose f is differentiable at a. Then

n If

f (x) = f (a) + ox (a)(xi" a%) + o(lx - a|)-

Since x% — a1 ^ 0 and o(|x — a|) ^ 0 as x ^ a,it follows that f (x) ^ f (a) as x ^ a. That is,f is continuous at a.

## 11.13 MEAN VALUE THEOREM AND CONSEQUENCES

Theorem Suppose f is continuous at all points on the line segment L joining a and a + h; and is differentiable at all points on L,except possibly at the end points

f (a + h) - f (a) = {df (x), h)

t f (x) ft-

i=l UXfor some x E L,x not an endpoint of L.

proof: Define the one variable function g by g(t) = f (a + th).

Then g is continuous on [0, 1] (being the composition of the continuous functions

t ^ a + th and x ^ f (x)). Moreover,

g(0) = f(a) g(1) = f(a + h).

We next show that g is differentiable and compute its derivative.

If 0 < t < 1,then f is differentiable at a + th,and so

f (a + th + w) — f (a + th) — {df (a + t h), w)

0 = lim

|w|^0

Assume w = sh where s is a small real number,positive or negative. Since |w| = ±s|h|,and since we may assume h = 0

f ((a + (t + s)h) — f (a + th) — {df (a + t h),sh)

0 = lim

s—* 0

s

= lim 1g(t + s> - -f(a + th). h)'

s^o y    s

using the linearity of df (a + th).

Hence g'(t) exists for 0 < t < 1,and moreover

g'(i) = {df (a +1h).h)

By the usual Mean Value Theorem for a function of one variable, applied to g,we have

g(1) - g(°) = g'CO

for some t E (°,1).

If the norm of the gradient vector of f is bounded by M, then it is not surprising that the difference in value between f (a) and f (a + h) is bounded by M|h|. More precisely.

Corollary Assume the hypotheses of the previous theorem and sup- pose IVf (x)| < M for all x e L. Then

If (a + h) - f (a)|< M |h|

proof: From the previous theorem

f (a + h) — f (a)| = Kdf (x), h)| for some x e L

= Vf (x) • h|

Vf (x)||h|

M|h|.

Corollary Suppose H C Rra is open and connected and f :H ^ R.

Suppose f is differentiable in H and df (x) = 0 for all x e H .

Then f is constant on H.

proof: Choose any a E H and suppose f (a) = a. Assume

E = {x E H : f (x) = a}.

Then E is non-empty (as a e E). We will prove E is both open and closed in H. Since H is connected,this will imply that E is all of H . This establishes the result.

To see E is open ,suppose x e E and choose r > 0 so that Br (x) C H. If y E Br (x), then from some u between x and y,

f (y) — f (x) = idf (u), y — x)

= 0,by hypothesis.

Thus f (y) = f (x) (= a), and so y E E.

Hence Br (x) C E and so E is open.

To show that E is closed in Q,it is sufficient to show that Ec ={y : f (x) = a} is open in Q.

From Proposition we know that f is continuous. Since we have Ec = f _1[R\{a}] and R\{a} is open,it follows that Ec is open in Q. Hence E is closed in Q,as required.

Since E = 0,and E is both open and closed in Q,it follows E = Q (as Q is connected). In other words,f is constant (= a) on Q.

# 11.14 CONTINUOUSLY DIFFERENTIABLE FUNCTIONS

The partial derivatives (and even all the directional derivatives) of a function can exist without the function being differentiable.

However,we do have the following important theorem:

Theorem  Suppose f : Q (C Rra) ^ R where Q is open. If the partial derivatives of f exist and are continuous at every point in Q; then f is differentiable everywhere in Q.

Remark: If the partial derivatives of f exist in some neighbourhood of, and are continuous at,a single point,it does not necessarily follow that f is differentiable at that point. The hypotheses of the theorem need to hold at all points in some open set Q.

proof: We prove the theorem in case n = 2 (the proof for n > 2 is only notationally more complicated).

Suppose that the partial derivatives of f exist and are continuous in Q. Then if a e Q and a + h is sufficiently close to a,

f (a1 + hl, a2 + h2) = f (a1, a2)

+f (a1 + h1,a2) — f(a1,a2)

+f (a1 + h1,a2 + h2) — f (a1 + h1, a2)

= f(a1, a2) + f (^, a2)h1 + f (a1 + h1, <=2)h2' for some f1 between a1 and a1 + h1,and some f2 between a2 and a2 + h2. The first partial derivative comes from applying the usual Mean Value Theorem, for a function of one variable,to the function f (x1, a2) obtained by fixing a2 and taking x1 as a variable. The second partial derivative is similarly obtained by considering the function f (a1 + hl, x2), where a1 + h1 is fixed and x2 is variable.

Hence

dx1 dx2

— (a1 ~2^ u

dx1

f (a1 + h1, a2 + h2) = f (al, a2) + (a1, a2)h1 + ^2. (a1, a2)h2

+ (d— ie-'a) - dx— (a"a2))h

+ (dx(a' +h, 2) - dx (a"a2))h

= f(a1, a2) + L(h) + ^(h), say.

Here L is the linear map defined by

L is represented by the previous 1 x 2 matrix. We claim that the error term

Thus

(dx K'-a2) - dx <a1•a2)) h1+(dx(a1+h, 2) - dx h2

can be written as o(|h|)

This follows from the facts:

df df

(<=\ a2) ^ 7— (a1,a2) as h ^ 0 (by continuity of the partial deriva-

dx dx

tives),

df df

-7— (a1 + h, 2) ^ j—2(a1, a2) as h ^ 0 (again by continuity of the

dx dx

partial derivatives),

Ih1l < |h|,|h21 < |h|.

It now follows from that f is differentiable at a,and the differential of f is given by the previous 1 x 2 matrix of partial derivatives.

Since a e H is arbitrary,this compes the proof

Definition If the partial derivatives of f exist and are continuous in the open set Q,we say f is a C1 (or continuously differentiable) function on Q. One writes f E C 1(Q).

It follows from the previous Theorem that if f E C 1(Q) then f is indeed differentiable in Q. Exercise: The converse may not be true,give a simple counterexample in R.

## 11.15 HIGHER-ORDER PARTIAL DERIVATIVES

df df

Suppose f : Q (c Rn) ^ R. The partial derivatives 7-—,v—,if they

dx1 dxn

exist,are also functions from Q to R,and may themselves have partial deriva- tives.

df

The jth partial derivative of is denoted by

d 2f

or fij or Dij f.

dxjdxi

If all first and second partial derivatives of f exist and are continuous

Notes

we write

f E C2(Q).

Similar remarks apply to higher order derivatives,and we similarly define Cq(Q) for any integer q > 0.

Note that

C0(Q) D C 1(Q) D C2(Q) d ...

The usual rules for differentiating a sum,product or quotient of functions of a single variable apply to partial derivatives. It follows that Ck(Q) is closed under addition,products and quotients (if the denominator is non-zero).

The next theorem shows that for higher order derivatives,the actual order of differentiation does not matter,only the number of derivatives with respect to each variable is important. Thus

d2f d2f

dxidxj dxjdxi'

and so

d3f d3f d3f

etc.

Theorem If f E C 1(H) and both fj and fj exist and are contin- uous (for some i = j) in Q,then fj = fj in Q.

In particular,if f E C2(Q) then fj = fj for all i = j.

proof: For notational simplicity we take n = 2. The proof for n > 2 is very similar.

Suppose a e Q and suppose h > 0 is some sufficiently small real number. Consider the second difference quotient defined by

A(h) = h"2 ^(f (al + h, a2 + h) - f (a1, a2 + h)

—(f(a1 + h, a2) — f (al, ci2))^ (1T.1T)

= (#(a2 + h) — 9(a2)), (17.18)

where

9(x'2) = f (a1 + h, x2) — f (a\x2).

(ai,a2+h) (ai+h,a2+h)

a = (ai,a2) (ai+h,a2)

A(h) = ((f(B) - f(A)) - (f(D) - f(C))) / h2 = ((f(B) - f(D)) - (f(A) - f(C))) / h2

From the definition of partial differentiation,g'(x2) exists and

g'(x2) = d-L (a1 + h, x2) — dX (a1, x2) (17.19)

for a2 < x < a2 + h.

Applying the mean value theorem for a function of a single variable to (17.18), we see from (17.19) that

A(h) = — g'(f2) some f2 E (a2, a2 + h) h

= h(dx <a'+h, f2) — f <aA2>) ■ <17-20'

df

Applying the mean value theorem again to the function (x\<=2), with <=2 fixed,we see

d2 f

A(h) = dXcVhF (<=l'<=2) some <= 1 E (a1, 0jl + h).

If we now rewrite as

A(h) = — ^ f (a1 + h,a2 + h) — f (a1 + h, a2)j

— f (a1, a2 + h) — f(a1 + a2))

and interchange the roles of x1 and x2 in the previous argument,we obtain

A(h) = sxdx1(n1, n2)

Notes

for some n1 E (a1,a1 + h), n2 E (a2,a2 + h).

If we assume h ^ 0 then (<=1, <=2) and (n1, n2) ^ (a1, a2), and so from and the continuity of f12 and f21 at a,it follows that

f12(a) = f21 (a) •

This compes the proof.

## 11.16 TAYLOR'S THEOREM

If g E C1 [a,6],then we know

f b

g(b) = g(a) + / g'(t) dt

J a

This is the case k = 1 of the following version of Taylor's Theorem for a function of one variable.

Theorem (Taylor's Formula; Single Variable,First Version)

Suppose g E Ck [a,b]. Then

1 2!

1 jk-D^,k-1, rb (b — t)k-1

proof: An elegant (but not obvious) proof is to begin by computing:

d (g<p(k-1) - g'f(k-2) + g<p(k-3) + (-1)fc-Vfc-1V)

= {gp(k) + g'f(k-1)) - {g'p(k-1) + g"f(k-2)) + (g"p(k-2) + g'"f(k-3))

+ (-1)k-1 (g(k-1)v' + g'k)v)

= gf(k) + (-1}k-kg(k)f.

Now choose

f(t)

Then

Dividing by (- 1)k 1 and integrating both sides from a to b,we get

g(b) -g(a) + g'(a)(b - a

rb • (b - t)k-1

dt.

(k - 1)!

This gives formula

(b - a)2 (k 1) (b - a) ap— H g{) (a J

2!

Theorem . (Taylor's Formula; Single Variable,Second Version)

Suppose g E Ck [a, b]. Then

g(b) = g(a)+ g'(a)(b - aH^y g"(a)(b - a)2 + ••g(k-1)(a)(b - a)k-1 + - g(k)(f)(b - a)k

(k - 1)!

for some f E (a,b).

proof: Since g(k is continuous in [a,b],it has a minimum value m,and a maximum value M,say.

By elementary properties of integrals,it follows that

rb (b - t)k—1. rb (k), Ab - t)k—1. f\ Ab - t)k—1,

m— dt < g('(tW —— dt < M—rvrdt,

(k - 1)! ~Ja (k - 1)! J a (k - 1)!

(b - t)k-1

Ja* k^Ak-1)! ^ M

< —— .. < M.

fb (b - t)k

dt

(k - 1)!By the Intermediate Value Theorem,g(k takes all values in the range [m,M],and so the middle term in the previous inequality must equal g(k'(<=) for some <= E (a, b). Since

rb (b - t)k-1 (b - a)k

Ja (k - 1)! = k!,

it follows

r g^^ dt = g(k'(<=).

Taylor's Theorem generalises easily to functions of more than one variable.

Theorem . (Taylor's Formula; Several Variables)

Suppose f E Ck(Q) where Q C Rn; and the line segment joining a and a + h is a subset of Q.

Then

n 1 n

f (a + h) = f (a) + <= A f (a) hi + - <= D,, f (a) h'h' + •••

i—1 2! i, j=1

1 n

+7^ E Di,,...ik-t f (a) hi1 • ... • hik-1 + Rk(a,h) A ii, --ik-1 = 1

where

1 n r 1

Rk(a,h) = (k _ 1)! ^ J (1 - t)k—1Dii...ikf (a + th) dt

()! ii a=1 0

il, ..., ifc = 1 '

1n

k! Di1, ..., ik f (a + sh) hi1 • ...• hik for some s E (0, 1).

k! ii, ..., ik = 1

proof: First note that for any differentiable function F : D (c Rn) ^ R we have

d "

— F(a + th) = y DtF(a + th) hi.

dt

This is just a particular case of the chain rule,which we will discuss later. Assume

g(t) = f (a + th).

Then g: [0, 1] ^ R. We will apply Taylor's Theorem for a function of one variable to g.

From we have

g'(t) = <= D<f (a + th) hi.

i— 1

Differentiating again,and applying to DiF,we obtain

n / n \

g''(t) = <= (<= d,, f (a + th) h') hi

i=i \'=i J

n

= ^2 D' f (a + th) hih'.

i, '=1

Similarly

g'''(t)= Y,Di'kf (a + th) hih'hk,

i, ', k=1

etc. In this way,we see g E Ck [0, 1] and obtain formulae for the derivatives of g.

g(1) = g(0) + g'(0)

+ ^ g"(0) + ••• + g(k—1>(0)

1 f1(1 - t)k—

1g(k>(t) dt

(k — 1)! Jo

or

—g(k)(s) some s E (0, 1).

Remark The first two terms of Taylor's Formula give the best first order approximation in h to f (a + h) for h near 0. The first three terms give the best second order approximation in h,the first four terms give the best third order approximation,etc.

Note that the remainder term Rk(a,h) in Theorem can be written as O(|h|k) (see theRemarks on rates of convergence in Section), i.e.

is bounded as h —► 0.

This follows from the second version for the remainder in Theorem and the facts:

Dil...ik f (x) is continuous,and hence bounded on compact sets,

Ihh • ...•hikl< |h|k.

Example Assume

f (x, y) = (1 + y2)1/2 cos x.

One finds the best second order approximation to f for (x, y) near (0, 1) as follows.

First note that Moreover,

2-3/2 at (0, 1f (x, y) = 21/2 + 2-1/2(y — 1) — 21/2x2 + 2-3/2(y — 1)2 + R^(0, 1), (x, y)), where

r3^(0, 1), (x, y)) = O [\(x, y) — (0, 1)|3) = O (7;r2 + (y — 1)2)

# 11.17 THE GENERAL LINEAR GROUP,GLN(R)

Recall from the Groups page that a group a set G with a binary operation $\cdot : G \times G \to G$ where:

**1)** For all a, b, $c \in G$ we have that $(a\cdot b)\cdot c = a\cdot(b\cdot c)$ (Associativity of $\cdot$).

**2)** There exists an element $e \in G$ such that $a\cdot e = a$ and $e\cdot a = a$ (The existence of an identity for $\cdot$).

**3)** For all $a \in G$ there exists a $a^{-1} \in G$ such that $a\cdot a^{-1} = e$ and $a^{-1}\cdot a = e$ (The existence of inverses for each element in G).

We will now look at the group of invertible $n\times n$ matrices with real entries under matrix multiplication $\cdot$ which is often called the **General Linear Group $GL_n(\mathbb{R})$. Assume A, B, $C \in GL_n(\mathbb{R})$.
Consider the product $A\cdot B$. Since A and B are invertible $n\times n$ matrices, we know from linear algebra that $A\cdot B$ will be an invertible $n\times n$ matrix (whose inverse is $(A\cdot B)^{-1} = B^{-1}\cdot A^{-1}$) and so $(A\cdot B) \in GL_n(\mathbb{R})$ so $GL_n(\mathbb{R})$ is closed under $\cdot$
We also already know that matrix multiplication is associative from linear algebra, and so $A\cdot(B\cdot C) = (A\cdot B)\cdot C$.
The identity for $\cdot$ is the $n\times n$ identity matrix $I_n$ whose main diagonal entries are all 1s and every other entry is a 0,i.e.:

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{n\times n}$$

We know from linear algebra that for any $n\times n$ square diagonal matrix A that $\det A = \prod_{i=1}^{n} a_{ii}$. So $\det I_n = \prod_{i=1}^{n} a_{ii} = \prod_{i=1}^{n} 1 = 1 \neq 0$ and so indeed $I_n$ is invertible,so $I_n \in GL_n(\mathbb{R})$.

Notes

For each element $A \in GL_n(R)$ we denote the inverse under · to be matrix inverse of A which we denote $A^{-1} \in GL_n(R)$ such that $A \cdot A^{-1} = I_n$ and $A^{-1} \cdot A = I_n$.

Hence, $(GL_n(R), \cdot)$ is a group.

Check your Progress - 1

Discuss Differentiation Of Real-Valued Functions & The Gradient

___________________________________________________________

___________________________________________________________

___________________________________________________________

Discuss The Gradient

___________________________________________________________

___________________________________________________________

___________________________________________________________

## 11.18 LET US SUM UP

In this unit we have discussed the definition and example of Connected Sets, Connectedness In R, Path Connected Sets, Differentiation Of Real-Valued Functions, Algebraic Preliminaries, Partial Derivatives, Directional Derivatives, The Differential (Or Derivative),The Gradient, Geometric Interpretation Of The Gradient, Level Sets And The Gradient, Mean Value Theorem And Consequences, Continuously Differentiable Functions, Higher-Order Partial Derivatives, Taylor's Theorem, The General Linear Group $GL_n(R)$

## 11.19 KEYWORDS

1. Connected Sets: One intuitive idea of what it means for a set S to be "connected" is that S cannot be written as the union.

2. Connectedness In R: The connected sets in R are precisely the intervals in R. We first need a precise definition of interval.

3. Path Connected Sets: A path connecting two points x and y in a metric space (X, d) is a continuous function

4. Differentiation Of Real-Valued Functions: In this Chapter we discuss the notion of derivative The inner product in Rn is represented by y • x = ylxl + ... + ynxn

5. Algebraic Preliminaries The zth Partial derivative of f at x is defined by f (x + tei) - f (x)

6.Partial Derivatives The directional derivative of f at x in the direction v = 0 is defined by Dvf (x) = lim f (x + tv) - f (x),

7. Directional Derivatives Motivation Suppose f : I (c R) ^ R is differentiable at a E I.

8.The Gradient Suppose f is differentiable at x. Then the directional derivatives at x are given by Dvf (x) = V • Vf (x).

9.Geometric Interpretation Of The Gradient If f: Rn ^ R then the level set through x is {y: f(y) = f(x) }

10.Level Sets And The Gradient Suppose f is continuous at all points on the line segment L joining a and a + h

11. Mean Value Theorem And Consequences, Continuously Differentiable Functions, Higher-Order Partial Derivatives

12. Taylor's Theorem If g E C1 [a,6],then f b g(b) = g(a) + / g'(t) dt

13. The General Linear Group GLn(R) a group a set G with a binary operation ·:G×G→G

# 11.20 QUESTIONS FOR REVIEW

Explain Differentiation Of Real-Valued Functions

Explain The Gradient

# 11.21 REFERENCES


- Analysis of Several Variables
- Application of Several Variables

- Function of Several Variables
- Several Variables
- Function of Variables
- System of Equation
- Function of Real Variables
- Real Several Variables
- Elementary Variables

## 11.22 ANSWERS TO CHECK YOUR PROGRESS

Differentiation Of Real-Valued Functions

(answer for Check your Progress - 1 Q)

The Gradient (answer for Check your Progress - 1 Q)

# UNIT-12 : DIFFERENTIATION OF VECTOR-VALUED FUNCTIONS

**STRUCTURE**

12.0 Objectives

12.1 Introduction

12.2 Differentiation Of Vector-Valued Functions

12.3 Partial And Directional Derivatives

12.4 The Chain Rule

12.5 The Inverse Function Theorem And Its Applications

12.6 Implicit Function Theorem

12.7 Maximum,Minimum,And Critical Points

12.8 Lagrange Multipliers

12.9 Let Us Sum Up

12.10 Keywords

12.11 Questions For Review

12.12 References

12.13 Answers To Check Your Progress

## 12.0 OBJECTIVES

After studying this unit you should be able to:

Learn Understand about Differentiation Of Vector-Valued Functions

Learn Understand about Partial And Directional Derivatives

Learn Understand about The Chain Rule

Learn Understand about The Inverse Function Theorem And Its Applications

Learn Understand about Implicit Function Theorem

Learn Understand about Maximum,Minimum,And Critical Points

Learn Understand about Lagrange Multipliers

Learn Understand about Partial And Directional Derivatives

Learn Understand about The Chain Rule

Learn Understand about The Inverse Function Theorem And Its Applications

Learn Understand about Implicit Function Theorem

Learn Understand about Maximum,Minimum,And Critical Points

Learn Understand about Lagrange Multipliers

## 12.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

Differentiation Of Vector-Valued Functions,Partial And Directional Derivatives,The Chain Rule,The Inverse Function Theorem And Its Applications,Implicit Function Theorem,Maximum,Minimum And Critical Points,Lagrange Multipliers,Partial And Directional Derivatives,The Chain Rule,The Inverse Function Theorem And Its Applications,Implicit Function Theorem,Maximum,Minimum,And Critical Points,Lagrange Multipliers

## 12.2 DIFFERENTIATION OF VECTOR-VALUED FUNCTIONS

Introduction

In this chapter we consider functions

f: D (C R") — R",

with m > 1.

We write

f (;r1...., r")= (/1 (x1, ..., x"),..., /'"(x1, ..., x"))

where

/1: D — R i = 1, ..., m

are real-valued functions.

Example Assume

f (x,y,z) = (x2 — y2,2xz + 1).

Then / 1(x,y,z) = x2 — y2 and /2(x,y,z) = 2xz + 1.

Reduction to Component Functions For many purposes we can reduce the study of functions f,as above,to the study of the corresponding real valued functions /1, ...,/m. However,this is not always a good idea,since studying the /1 involves a choice of coordinates in R",and this can obscure the geometry involved.

In Definitions  we define the notion of partial derivative, directional derivative, and differential of f without reference to the component functions. definitions are equivalent to definitions in terms of the component functions.

Paths in Rm

In this section we consider the case corresponding to n = 1 in the notation of the previous section. This is an important case in its own right and also helps motivates the case n > 1.

Definition. Assume I be an interval in R. If f : I — Rra then the derivative or tangent vector at t is the vector

f'(t) = lim f (t + s) ~ f (t),

•^0      s

provided the limit exists . In this case we say f is differentiable at t. If, moreover, f'(t) = 0 then f'(t)/|f'(t)| is called the unit tangent at t.

Remark Although we say $f'(t)$ is the tangent vector at t, we should really think of $f'(t)$ as a vector with its "base" at f (t). See the next diagram.

Proposition. Assume $f(t) = f1 (t),..., fm(t)$. Then f is differentiable at t iff f1, . .., fm are differentiable at t. In this case

$f'm = (f1 '(t),..., r'(t)$

Proof: Since

f (t + s) - f (t) = / f 1(t + s) - f 1(t) fm(t + s) - fm(t) \ s s s

Definition. If $f(t) = f1 (t),...,fm(t)$ then f is C1 if each f is C1.

We have the usual rules for differentiating the sum of two functions from I to ^m, and the product of such a function with a real valued function (exer- cise : formulate and prove such a result). The following rule for differentiating the inner product of two functions is useful.

Proposition. If f 1,f2: I — Rra are differentiable at t then d

— (fi(t) f2(t)) = (f/l(t), f2(t)) + (f1(t) f2(t^ .

Proof: Since

m

(f 1(t),f2(t^ ^f1(t)f2(t),

i=1

the result follows from the usual rule for differentiation sums and products.

If f : I Rn, we can think of f as tracing out a "curve" in Rn (we will make this precise later). The terminology tangent vector is reasonable, as we see from the following diagram. Sometimes we speak of the tangent vector at f (t) rather than at t, but we need to be careful if f is not one-one, as in the second figure.

Examples

Assume

f (t) = (cos t,sin t) t E [0,2n). This traces out a circle in R2 and

f'(t) = (— sint,cos t).

Assume

f (t) = (t, t2). This traces out a parabola in R2 and

f'(t) = (1,2t).

f'(t) = (-sin t,cos t)

f(t)=(cos t,sin t)l

Example Consider the functions

fi(t) = (t, t3) t E R,

f2(t) = (t3, t9) t E R,

f3(t) = (tft,t) t E R.

Then each function fj traces out the same "cubic" curve in R ,(i.e.,the image is the same set of points), and

fi(0) = f2(0) = f3(0) = (0,0).

However,

f'i(0) = (1,0), f2(0) = (0,0), f3(0) is undefined.

Intuitively,we will think of a path in Rra as a function f which neither stops nor reverses direction. It is often convenient to consider the variable t as representing "time". We will think of the corresponding curve as the set of points traced out by f. Many different paths (i.e. functions) will give the same curve; they correspond to tracing out the curve at different times and velocities. We make this precise as follows:

Definition. We say f: I — Rra is a path in Rra if f is C1 and f'(t) = 0 for t E I. We say the two paths f 1: I1 — Rra and f2 : I2 — Rra are equivalent if there exists a function 0: I1 — I2 such that f 1 = f2 o 0,where 0 is C1 and 0'(t) > 0 for t E I1.

A curve is an equivalence class of paths. Any path in the equivalence class is called a parametrisation of the curve.

We can think of 0 as giving another way of measuring "time".

We expect that the unit tangent vector to a curve should depend only on the curve itself,and not on the particular parametrisation.

fi'(t) / |fi'(t)| = f2'(9 (t)) / |f2'(9 (t))l

2

Proposition Suppose f 1: I1 — Rra and f2 : I2 — Rra are equivalent parametrisations; and in particular f 1 = f2 o 0 where 0: I1 — I2,0 is C1 and 0'(t) > 0 for t E I1. Then f 1 and f2 have the same unit tangent vector at t and 0(t) respectively.

proof: From the chain rule for a function of one variable,we have

no = (/: '(t),..., /r(t)}

/(m ^'(t),..., frm)) m

f 2(0(t)) 0'(t)-

f l(t) f 2(t)

\m\ \f2(t)i

Definition. If f is a path in Rn, then the acceleration at t is f"(t).

Example If \f'(t)\ is constant (i.e. the "speed" is constant) then the velocity and the acceleration are orthogonal.

proof: Since \f(t)\2 = f'(t), f'(y)) is constant, we have from Proposition that

d

d (f'(t),f

if"(t), f'{y}).

This gives the result.

Arc length

Suppose f \[a,b] ^ Rn is a path in Rn. Assume a = ti <t2 < ... < tn = b be a partition of [a,b],where ti — ti-i = St for all i.

We think of the length of the curve corresponding to f as being

<=\f (ti) — / (ti_i)! = <= '/(fi) st« t \f (t) dt.

i=2 i=2

Definition. Assume f: [a, b] ^ Rra be a path in Rra. Then the length of the curve corresponding to f is given by

f b

/ \f\t)\ dt.

J a

The next result shows that this definition is independent of the particular parametrisation chosen for the curve.

Proposition. Suppose fi : [ai, bi] ^ Rra and f2 : [a2, b2] ^ Rra are equivalent parametrisations; and in particular f i = f2 o 0 where 0 :[ai,bi] ^ [a2, b2] 0 is C1 and 0'(t) > 0 for t E Ii. Then

f \fi(t)\dt =/ \f2(s)\ds.

Jai Ja2

proof: From the chain rule and then the rule for change of variable of integration,

\fi(t)\ dt = I \f2(0(t))\ 0'(t)dt

\f 2(s)\ds.

a2

# 12.3 PARTIAL AND DIRECTIONAL DERIVATIVES

Notes

Definition. The ith partial derivative of f at x is defined by df n _
uf(x + tei) — f(x)

7-— (x) for Dif(x)) = lim BXiK ' V 1\>)

dxi ^ ' t^o t

provided the limit exists. More generally,the directional derivative of f at x in the direction v is defined by

, , f(x + tv) — f(x)

Dv f (x) = lim ————^,

v K ' t^o t

provided the limit exists.

It follows immediately from the Definitions that

dX W = De f (*)•

The partial and directional derivatives are vectors in Rra. In the terminology of the previous section,(x) is tangent to the path t ^f (x + tej) and Dvf (x) is tangent to the path t ^ f (x + tv). Note that the curves corresponding to these paths are subsets of the image of f.

As we will discuss later,we may regard the partial derivatives at x as a basis for the tangent space to the image of f at f (x) .

Proposition. If f1,..., fm are the component functions of f then

df (df1 dfm \

sx(a) = (f(a)-"f(a)] for' = 1--n

Dvf(a) = (dv f 1(a),..., Dvfm(a})

in the sense that if one side of either equality exists,then so does the other, and both sides are then equal.

Example Assume f: R2 ^ R3 be given by

f(x, y) = (x2 — 2xy,X + y3,sin x).

Then

df, , (df1 df2 df3\,

dX (x'y) = [jx-lx'ix\= (2x — 2y, 2x'cos x),

df () (df1 df2 df3 \ 3 2 0)

oy,(x'y) = = ("2x-3y-0)'

are vectors in R3.

The linear transformation L is denoted by f'(a) or df (a) and is called the derivative or differential of f at a .

A vector-valued function is differentiable iff the corresponding component functions are differentiable. More precisely:

Proposition. f is differentiable at a iff f1,..., fm are differentiable at a. In this case the differential is given by

(df (a) v) = (y(df 1 (a), v), ..., (df m (a), v ^.

In particular,the differential is unique.

proof: For any linear map L : Rn — Rn,and for each i = 1, ..., m,assume Li: Rn — R be the linear map defined by L% (v) = ^L(v)j .

it follows

f (x) - (f (a) + L(x - a)

| x — a|

iff

f(x) — (f-(a) + L'(x — a))

— 0 as x — a for i = 1,..., m.

I x — a |

Thus f is differentiable at a iff f1,..., fm are differentiable at a.

**Notes**

In this case we must have

L- = df- (a) i = 1, ..., m

(by uniqueness of the differential for real-valued functions), and so

L(v) = ((df 1 (a),v),..., (d^f m (a),v))-

But this says that the differential df (a) is unique

Corollary. If f is differentiable at a then the linear transformation df (a) is represented by the matrix

proof: The ith column of the matrix corresponding to df (a) is the vector (df(a), ej) this is the column vector corresponding to

\df1 (a), e i),..., (dfm (a), e i)N

i.e. to

(f (a)).

This proves the result.

Remark The jth column is the vector in Rn corresponding to the partial

d f derivative (a). The ith rovj represents dfj(a).

The following proposition is immediate.

Proposition. If f is differentiable at a then

f(x) = f(a) + (df(a), x - a) + ^(x),

vjhere ^(x) = o(|x — a|).

Conversely,suppose

f (x) = f (a) + L(x — a) + ^ (x),

vjhere L: Rn ^ Rn is linear and ^(x) = o(|x — a|). Then f is differentiable at a and df (a) = L.

proof: Thus as is the case for real-valued functions,the previous proposition implies f (a) + (df (a), x — a) gives the best first order approximation to f (x) for x near a.

Example Assume f: R2 ^ R2 be given by

f(x, y) = (x — 2xy, x + y).

Find the best first order approximation to f (x) for x near (1,2).
Solution:So the best first order approximation near (1,2) is

3 — 2(x - 1) - 4(y - 2) 9 + 2(x — 1) + 12(y — 2)

7 — 2x — 4y

17 + 2x + 12y

Alternatively,working with each component separately,the best first order approximation is

f1(1-2) + f (1, 2)(x -f(1-2)(y—2)-

f 2(1, 2) + f (1, 2)(x — 1) + % (y — 2))

= (—3 — 2(x — 1) — 4(y — 2), 9 + 2(x — 1) + 12(y — 2)) = (7 — 2x — 4y,—17 + 2x + 12y).

Remark One similarly obtains second and higher order approximations by using Taylor's formula for each component function.

Proposition. If f,g : D (c Rra) ^ Rra are differentiable at a e D, then so are af and f + g. Moreover

proof: The previous proposition corresponds to the fact that the partial deriva- tives for f + g are the sum of the partial derivatives corresponding to f and g respectively. Similarly for af.

Higher Derivatives We say f E Ck(D) iff f1,..., fm E Ck(D).

It follows from the corresponding results for the component functions that

f E C 1(D) ^ f is differentiable in D;

C0(D) D C 1(D) D C2(D) D ....

## 12.4 THE CHAIN RULE

Motivation The chain rule for the composition of functions of one variable says that

d

dx^f (x)) = 9'(f (x)) f'(x).

Or to use a more informal notation,if g = g(f) and f = f (x), then

dg dg df dx df dx

This is generalised in the following theorem. The theorem says that the linear approximation to g o f (computed at x) is the composition of the linear approximation to f (computed at x) followed by the linear approximation to g (computed at f (x)).

A Little Linear Algebra Suppose L: Rra ^ Rra is a linear map. Then we define the norm of L by

= max{|L(x)| : |x| < 1} .

A simple result (exercise) is that

for any x E Rra.

It is also easy to check (exercise) that || • || does define a norm on the vector space of linear maps from Rra into Rra.

Theorem (Chain Rule) Suppose f : D (c Rra) ^ Q(c Rra) and g :0(c Rra) ^ Kr. Suppose f is differentiable at x and g is differentiable at f (x). Then g o f is differentiable at x and

d(g o f)(x) = dg(f(x)) o df(x). D (C Rn) -U Q(c Rn) -

d(gof)(x) = dg(f (x)) o df (x)

>n df(x) T-n dg(f(x))

Rn

Example To see how all this corresponds to other formulations of the chain rule,suppose we have the following:

Thus coordinates in R3 are denoted by (x,y,z), coordinates in the first copy of R2 are denoted by (u, v) and coordinates in the second copy of R2 are denoted by (p,q).

The functions f and g can be written as follows:

f : u = u(x, y, z), v = v(x, y, z), g : p = p^^ q = q(u, v).

Thus we think of u and v as functions of x,y and z; and p and q as functions of u and v.

We can also represent p and q as functions of x,y and z via p = p(u(x, y, z), v(x, y, zf), q = q(Ku(x, y, z), v(x, y, zf).

The usual version of the chain rule in terms of partial derivatives is:

dp dp du dp dv

dx du dx dv dx dp dp du dp dv

dx du dx dv dx

dq dq du dq dv dz du dz + dv dz

In the first equality,|| is evaluated at (x, y, z), djp and dp are evaluated at (u(x, y,z),v(x, y,z)), and dp and dx are evaluated at (x, y, z). Similarly for the other equalities.

In terms of the matrices of partial derivatives:

du du du dx dy dz

dv dy

df (x)

where x = (x,y,z).

Proof of Chain Rule: We want to show

(f O g)(a + h) = (f O g)(a) + L(h) + o(\h\),

where L = df (g(a)) o dg(a).

Now

(f o g)(a + h)

f)g(a + h)

f fg(a) + g(a + h) - g(a)

f(g(a)) + (df(g(a)), g(a + h) - g(a)

+o(^|g(a + h) - g(a

.. .by' the differentiability of f

f(g(a)) + (df(g(a)), (dg(a), h) + o(|h|)

+o(\g(a + h) - g(a)\)

.. .by the differentiability of g

f(g(a)) + (df(g(a)), (dg(a), h))

+ (df(g(a)), o(|h|)) + o(|g(a + h) -g(a A + B + C + D

But B = ^df (^g(a)J o dg(a), hy,by definition of the "composition" of two maps. Also C = o(|h|). Finally,for D we have

(dg(a), h) + o(|h|) ...by differentiability of g <||dg(a)|| |h| + o(|h|) ... = O(|h|) ... why?

Substituting the above expressions into A + B + C + D,we get (f o g)(a + h) = f(g(a)) + (df(g(a))) o dg(a), h) + o(|h|).

If follows that f o g is differentiable at a, and moreover the differential equals df (g(a)) o dg(a). This proves the theorem.

## 12.5 THE INVERSE FUNCTION THEOREM AND ITS APPLICATIONS

Inverse Function Theorem

Motivation

Suppose

f :Q(C Rn) ^ Rn

and f is C1. Note that the dimension of the domain and the range are the same. Suppose f (x0) = y0. Then a good approximation to f (x) for x near x0 is gven by

x ^ f (xo) + (f'(xo),x-xo).

first order map x |-> f(xo) + <f(xo), x-xg>

We expect that if f'(xo) is a one-one and onto linear map,(which is the same as det f'(x0) = 0 and which implies the map in (19.1) is one-one and onto), then f should be one-one and onto near x0. This is true, and is called the Inverse Function Theorem.

Consider the set of equations

f 1(x1, ..., xn) = y1 f 2(x1, ..., xn) = y2

fn(x1, ..., xn) = yn,

where f , .. ., fn are certain real-valued functions. Suppose that these equations are satisfied if (x1, ...,xn) = (x0, ...,x%) and (y1, ...,yn) = (y0, ..., y0), and that det f '(x0) = 0. Then it follows from the In- verse Function Theorem that for all (y1, ...,yn) in some ball centred at (y^, ..., y0) the equations have a unique solution (x1, ...,xn) in some ball centred at (x^, . .., xn).

Theorem. (Inverse Function Theorem) Suppose f :Q(c Rn) ^ Rn is C1 and Q is open1. Suppose f'(x0) is invertible for some x0 E Q.

Then there exists an open set U 3 x0 and an open set V 3 f (x0) such that

f '(x) is invertible at every x E U,

f: U ^ V is one-one and onto,and hence has an inverse g: V ^ U,

g is C1 and g'(f (x)) = [f'(x)]_1 for every x E U.

Proof: Step 1 Suppose

y* E Bs(f (xo)).

(We will take the set V in the theorem to be the open set Bs(f(xo)))

For each such y,we want to prove the existence of x (= x*,say) such that

f (x) = y*.

We write f (x) as a first order function plus an error term. Thus we want to solve (for x)

f (xo) + {f'(xo),x - xo) + R(x) = y*,

where

R(x) := f (x) - f (xo) - {f'(xo),x xo). In other words,we want to find x such that

{f'(xo), x - xo) = y* - f (xo) - R(x),

i.e. such that

x = xo + ([/M] 1, y* - f(xo^ -([f'(xo)] 1, R(x))

(w% ?).

The right side is the sum of two terms. The first term,that is xo + {[f'(xo)]-1,y* - f(xo)),is the solution of the linear equation y* = f (xo) + {f'(xo),x - xo). The second term is the error term - {[f'(xo)]-1, R(x)), which is o(|x - xo|) because R(x) is o(|x - xo|) and [f'(xo)]-1 is a fixed

<[f(Xo)]-1,R(x)>

graph of f.

R(X)

* V x

Xo X

A *

Xo + <[f'(Xo)]-1,y -f(Xo)>

Step 2 Because of define

Ay* (x) := x + ([/(xo)}-1, y* — f(xo)) — ([f'(x0)}-i, R(x)) .

Note that x is a fixed point of Ay* iff x satisfies and hence solves. We claim that

Ay* : Be(xo) ^ Be(xo)

and that Ay* is a contraction map,provided e > 0 is sufficiently small (e will depend only on x0 and f) and provided y* E B$ (y0) (where 5 > 0 also depends only on x0 and f).

To prove the claim,we compute

R(x2) — R(xi) = f (x2) — f (xi) — (f'(x0),x2 — xi) .

We apply the mean value theorem to each of the components of this equation to obtain

Ri(x2) — R'(xi)

fi (Ci),x2 — xi) — f (x0),x2 — xi for i = 1,..., n and some ^ E Rra between xi and x2

fi (Ci) — fi (x0),x2 — xi f (Ci) — f (x0) \x2 — xi1,

by Cauchy-Schwartz,treating fi as a "row vector".

By the continuity of the derivatives of f,it follows

\R(x2) — R(xi)\ < 2K\x2 — xi\,

provided x\, x2 E Be(x0) for some t> 0 depending only on f and x0.

\Ay*(xi) - Ay*(x2)\ < |xi - x2\.

2

This proves

Ay* : Bn (x0) ^

is a contraction map.

For this we compute

\Ay*(x) - xo\ < |([f'(x0)]-1, V* -f(x0^| + ^[f/(xo)]-1, ^(x))| < K\V* - f (xo)\ + K\R(x)\

= K\v* - f (x0)\ + K\R(x) - R(x0)\ as R(x0) = 0

K\V* -f (x0)\ + 2\x -x0\

e/2 + e/2 = e,

provided x E Be(x0) and y* E B$(f (x0)) (if K5 < e). This establishes and compes the proof of the claim.

Step 3 We now know that for each y E B$ (f (x0)) there is a unique x E Be(x0) such that f (x) = y. Denote this x by g(y). Thus

g: Bs (f (x0)) ^ Be(x0).

We claim that this inverse function g is continuous.

To see this assume xi = g(vi) for i = 1,2. That is,f (xi) = yi,or equivalently xi = Ayi(xi) Then

\g(yi) - g(V2)\ = \xi - x2\

K \yi - V2\ + K \R(xi) - R(x2)\ from (19.8)

K \yi - V2\ + K2K\xi - x2\ from (19.10)

= K \yi - V2\ + 2\g(yi) - g(V2)\.Thus

2\g(vi) - g(V2)\ <K Ivi- V2 \,

and so

\ g(yi) - g(V2)\< 2KI yi- V2I

In particular,g is Lipschitz and hence continuous.

Step 4 Assume

V = Bs(f (xo)), U = g [B&(f (x0))]

Since U = Be(x0) H f 1[V] (why?), it follows U is open. We have thus proved the second part of the theorem.

The first part of the theorem is easy. All we need do is first replace Q by a smaller open set containing xo in which f '(x) is invertible for all x.

This is possible as det f'(x0) = 0 and the entries in the matrix f '(x) are continuous.

Step 5 We claim g is C1 on V and

g'(f (x)) = [f'(x)]-1.

To see that g is differentiable at y E V is true,suppose y, y E V,and assume f (x) = y,f (x) = y where x, x E U. Then

|g(y) -g(y) - ([f'(x)]-1, y- y)\

\y - y\

= \x - x - ([f'(x)]-\ f(x) - f(x))\

\y - y\

[f'(x)]-1,(f'(x),x - x) - f (x) + f (x)

If we fix y and assume y ^ y,then x is fixed and x ^ x. Hence the last line in the previous series of inequalities ^ 0,since f is differentiable at x and \x - x\/\y - y\< K/2  Hence g is differentiable at y and the derivative .

The fact that g is C1 follows from the expression for the inverse of a matrix.

Remark We have

1

g' (y) = [f'(g(y))Y Ad [f '(g(y))]

det[f '(g(y))]'

where Ad [f'(g(y))] is the matrix of cofactors of the matrix [f (g(y))].

If f is C2,then since we already know g is C1,it follows that the terms in the matrix are algebraic combinations of C1 functions and so are C1. Hence the terms in the matrix g' are C1 and so g is C2.

Similarly,if f is C3 then since g is C2 it follows the terms in the matrix are C2 and so g is C3.

By induction we have the following Corollary.

Corollary.If in the Inverse Function Theorem the function f is Ck then the local inverse function g is also Ck.

Summary of Proof of Theorem

Write the equation f (x*) = y as a perturbation of the first order equation obtained by linearising around x0..

Write the solution x as the solution T (y*) of the linear equation plus an error term E(x),

x = T(y*) + E(x) =: Ay*(x)

Show Ay* (x) is a contraction map on Be(x0) (for e sufficiently small and y* near y0) and hence has a fixed point. It follows that for all y* near y0 there exists a unique x* near x0 such that f (x*) = y*. Write g(y*) = x*.

The local inverse function g is close to the inverse T(y*) of the linear function. Use this to prove that g is Lipschitz continuous.

Wrap up the proof of parts 1 and 2 of the theorem.

Write out the difference quotient for the derivative of g and use this and the differentiability of f to show g is differentiable.

## 12.6 IMPLICIT FUNCTION THEOREM

Motivation We can write the equations in the previous "Motivation" section as

f (x) = y,

where x = (x1xn) and y = (y1,..., yn).

More generally we may have n equations

f (x, u) = y,

i.e.

f 1(x1, ..., xn, u1, ..., um) f 2(x1,..., x'n, ul, ..., um) f n(x1,..., xn, u\..., um) = yn,

where we regard the u = (u1, ...,um) as parameters. Write

' df

det

dx

Thus det[df/dx] is the determinant of the derivative of the map f (x1, ...,xn), where x1, ...,xm are taken as the variables and the u1,..., um are taken to be fixed.

Now suppose that f (xo, uo) — yo, det From the Inverse Function Theorem (still thinking of u1, ...,um as fixed), for y near y0 there exists a unique x near x0 such that

f (x,uo) — y.

The Implicit Function Theorem says more generally that for y near y0 and for u near u0,there exists a unique x near x0 such that

f (x, u) = y.

In applications we will usually take y — y0 — c(say) to be fixed. Thus we consider an equation

f (x,u) — c

where

f (x0,u0) c,

Hence for u near u0 there exists a unique x — x(u) near x0 such that

f (x(u),u) — c.

In words,suppose we have n equations involving n unknowns x and certain parameters u. Suppose the equations are satisfied at (x0, u0) and suppose that the determinant of the matrix of derivatives with respect to the x variables is non-zero at (x0, u0). Then the equations can be solved for x — x(u) if u is near u0.

Notes

Moreover,differentiating the ith equation in with respect to uj we obtain

,, df df + df — 0

— 0.

dxk duj duj

That iswhere the first three matrices are n x n,n x m,and n x m
+
respectively,and the last matrix is the n x m zero matrix. Since det [df /dx]^XQ Uo) — 0,it follows

(Xo, Uo)

Example. Consider the circle in R2 described by

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

Write

F (x,y) = 1.

Thus u is replaced by y and c is replaced by 1.

y

(xo, yo)

Suppose F(xo, yo) = 1 and dF/dxo\(Xa, yo) = 0 (i.e. xo = 0). Then for y near y0 there is a unique x near x0 satisfying In fact x = ±V 1 — y2 according as x0 > 0 or x0 < 0. See the diagram for two examples of such points (x0, y0).

Similarly,if dF/dy0\(Xoyo) = 0,i.e. y0 = 0,Then for x near x0 there is a unique y near y0 satisfying

Example Suppose a "surface" in R  is described by

F(x,y,z) = 0.

Suppose F(x0, y0, z0) = 0 and dF/dz (x0, y0, z0) = 0.

(x, y, z) = 0

(xo, yo)

Then by the Implicit Function Theorem,for (x, y) near (x0, y0) there is a unique z near z0 such that T(x, y, z) = 0. Thus the "surface" can locally be written as a graph over the x-y plane

More generally,if VT(x0, y0, z0) = 0 then at least one of the derivatives dF/dx (x0, y0, z0), dF/dy (x0, y0, z0) or dF/dz (x0, y0, z0) does not equal 0. The corresponding variable x,y or z can then be solved for in terms of

the other two variables and the surface is locally a graph over the plane corresponding to these two other variables.

Example Suppose a "curve" in R3 is described by

®(x, y, z)

Suppose (x0, yo, z0) lies on the curve,i.e. $(x0, y0, z0) = ^(x0, yo, z0) = 0. Suppose moreover that the matrix

dx dy dz di &W di dx dy dz _

(xo, yo, zo)

has rank 2. In other words,two of the three columns must be linearly independent. Suppose it is the first two. Then By the Implicit Function Theorem,we can solve for (x, y) near (x0, y0) in terms of z near z0. In other words we can locally write the curve as a graph over the z axis.

Example Consider the equations

Zi(xi, x2, yi, y2, y3) = 2exi + x2yi - 4y2 + 3 f2(xi, x2, yi, y2, y3) = x2 cos xi - 6xi + 2yi - y-3.

Consider the "three dimensional surface in R " given by fi(xi, x2, yi, y2, y3) = 0,f2(xi, x2, yi, y2, y3) = 0 . We easily check that

f (0, 1,3,2,7) = 0

231 -4 0 -612 0-1

The first two columns are linearly independent and so we can solve for x1,x2 in terms of y1, y2, y3 near (3,2,7).

Moreover

**Notes**

dxi dxi dxi

dyi dy2 dy3

dx2 dx2 dx2

dyi dy2 dy3

It folows that for $(y_1, y_2, y_3)$ near (3,2,7) we have

K ^ K ^ 3 . xi ~ 0 + 4(yi -3) + 5(y2- 2) - ^(ys- 7)

K ^ 6, 1,

x2 ~ 1 - 2(yi - 3) + 5(y2- 2) + ^(ys- 7).

We now give a precise statement and proof of the Implicit Function Theorem.

Theorem. (Implicit function Theorem) Suppose f : D (c Rra x Rk) ^ Rra is

C1 and D is open. Suppose f (x0, u0) = y0 where x0 E Rra and u0 E Rm. Suppose det [df /dx] |(xo, uo) = 0.

Then there exist e, 8 > 0 such that for all y E B$(y0) and all u E B$(u0) there is a unique x E Be(x0) such that

f (x, u) = y.

If we denote this x by g(u, y) then g is C1. Moreover,

(xo, uo)

proof: Define

F: D Rra x Rm

by

F (x, u) = f (x, u),u). Then clearly5 F is C1 anddet F'l(xouo) = det

Also

F (x0, u0) = (yo,Uq).

From the Inverse Function Theorem,for all (y,u) near (y0,u0) there exists a unique (x, w) near (x0, u0) such that

F (x, w) = (y, u).

Moreover,x and w are C1 functions of (y, u). But from the definition of F it follows that holds iff w = u and f (x, u) = y. Hence for all (y, u) near (y0, u0) there exists a unique x = g(u, y) near x0 such that

f (x, u) = y.

Moreover,g is a C1 function of (u, y).

The expression for dug follows from differentiating precisely

as in the derivation

Manifolds

Discussion Loosely speaking,M is a fc-dimensional manifold in Rn if M locally looks like the graph of a function of k variables. Thus a 2-dimensional manifold is a surface and a 1-dimensional manifold is a curve.

We will give three different ways to define a manifold and show that they are equivalent.

We begin by considering manifolds of dimension n — 1 in Rn (e.g. a curve in R2 or a surface in R3). Such a manifold is said to have codimension one.

Suppose

$:Rn ^ R

is C1 . Assume

M = {x : $(x) = 0}.

If VT(a) = 0 for some a G M,then M locally as the graph of a function of one of the variables xi in terms of the remaining n — 1 variables.

This leads to the following definition.

**Notes**

Definition [Manifolds as Level Sets] Suppose M C Rn and for

each a G M there exists r > 0 and a C1 function T: Br (a) ^ R such that

M n Br(a) = {x : $(x) = 0}.

Suppose also that V$(x) = 0 for each x E Br(a).

Then M is an n — 1 dimensional manifold in Rn. We say M has codimension one.

The one dimensional space spanned by V$(a) is called the normal space to M at a and is denoted by NaM

Remarks

Usually M is described by a single function $ defined on Rn

for a discussion of V$(a) which motivates the defini- tion of Na M.

locally write M as the graph of a function

xi f (x11 ■ ■ ■ j xi-1 j xi+11 ■ ■ ■ j xn)

for some 1 < i < n.

Higher Codimension Manifolds Suppose more generally that

$: Rn -»■ Re

is C1 and I > 1. Now

M = M1 n — n m<=,

where

Mi = {x : $i(x) = 0}.

Note that each Ti is real-valued. Thus we expect that,under reasonable conditions,M should have dimension n — t in some sense. In fact,if

VT1(x),..., VT^(;r)

are linearly independent for each x E M, then the same argument as for Example3: in the previous section shows that M is locally the graph of a function of t of the variables x1,..., xn in terms of the other n — t variables.

This leads to the following definition which generalises the previous one.

Definition [Manifolds as Level Sets] Suppose M C Rn and for each a E M

there exists r > 0 and a C1 function T: Br (a) — R^ such that

M n Br(a) = {x : $(x) = 0}.

Suppose also that VT1(x),...,VT^(x) are linearly independent for each x E Br (a).

Then M is an n — t dimensional manifold in Rn. We say M has co dimension t.

The t dimensional space spanned by VT1(a),..., VT^(a) is called the normal space to M at a and is denoted by NaM .

Remarks locally write M as the graph of a function of t of the variables in terms of the remaining n — t variables.

Equivalent Definitions There are two other ways to define a manifold. For simplicity of notation we consider the case M has codimension one,but the more general case is compely analogous.

Definition [Manifolds as Graphs] Suppose M C Rn and that for each a E M there exist r > 0 and a C1 function f: Q(c Rn -1) - R such that for some 1 < i < n

M n Br (a) = {x E Br (a) : xi = f (x1, .. ., x—1, xi+1, ...,xn)}.

Then M is an n — 1 dimensional manifold in Rn.

The space NaM does not depend on the particular $ used to describe M.

Equivalence of the Level-Set and Graph Definitions Suppose M is a manifold as in the Graph Definition. Assume

T(x) = Xi - f (xi,. . ., Xi-x, Xi+l,.. ., Xn).

Then

^ ^ (df df df 6f\

V (x)=(9xi ^.^ dX-i, , 3x1+i,.. dXn)

In particular,VT(x) = 0 and so M is a manifold in the level-set sense.

Conversely,

If M is a manifold in the level-set sense then it is also a manifold in the graphical sense.

Definition [Manifolds as Parametrised Sets] Suppose M C Rn and that for each a E M there exists r > 0 and a C1 function

F :Q(C Rn-i) ^ Rn

such that

M n Br(a) = F[Q] n Br(a).

Suppose moreover that the vectors

dF () ()

(u),...,(u)

UUi UUn—1

are linearly independent for each u E Q.

Then M is an n — 1 dimensional manifold in Rn. We say that (F,Q) is a parametrisation of (part of) M.

The n — 1 dimensional space spanned by dF (u),...,JF (u) is called the tangent space to M at a = F(u) and is denoted by TaM.

Equivalence of the Graph and Parametrisation Definitions Suppose M is a manifold as in the Parametrisation Definition. We want to show that M is locally the graph of a C1 function.

First note that the n x (n — 1) matrix dF (p)

has rank n - 1 and so n — 1 of the rows are linearly independent. Suppose the first n — 1 rows are linearly independent.

= (x1, ..., xn.1)

It follows that the (n — 1) x (n — 1) matrix jp- (p) is invert-

lauj ll<i, j<n-l

ible and hence by the Inverse Function Theorem there is locally a one-one correspondence between u = (u, ...,un-i) and points of the form

(xi, .. ., Xn-i) = (F1 (u),...,Fn-l(u)) e Rn-1 - Rn-1 x {0} (c Rn),

with C1 inverse G (so u = G(xl, ...,xn-l)).

Thus points in M can be written in the form

(F1 (u),...,Fn-l(u),Fn(u)) = (xi, .. ., Xn-i,(Fn o G)(xi, .. ., Xn-l)) . Hence M is locally the graph of the Cl function Fn o G.

Conversely,suppose M is a manifold in the graph sense. Then locally, after perhaps relabelling coordinates,for some Cl function f: Q (c Rn" -l) - R,

M = {(xi,..., xn): xn = f (xi,. . ., xn-l)}.

It follows that M is also locally the image of the Cl function F : Q (c Rn-i) — Rn defined by

F (xi,. . ., xn-l) = (xi, . . .,xn-l, f(xi,. . .,x„- i)) .

Moreover,

dF df

rv ei + rv en

dxi dxi

for i = 1, . .., n — 1,and so these vectors are linearly independent.

In conclusion, we have established the following theorem.

Theorem. The level-set,graph and parametrisation definitions of a manifold are equivalent.

Notes

Remark If M is parametrised locally by a function F: Q (c Rk) — Rn and also given locally as the zero-level set of $: Rn — R^ then it follows that

k + I = n.

To see this, note that previous arguments show that M is locally the graph of a function from Rk — Rra-fc and also locally the graph of a function from R«.-r — r^_ This makes it very plausible that k = n — I. A strict proof requires a little topology or measure theory.

## 12.7 MAXIMUM,MINIMUM,AND CRITICAL POINTS

In this section suppose F:Q (c Rn) ^ R,where Q is open.

Definition. The point a G Q is a local minimum point for F if for some r > 0

F(a) < F(x)

for all x G Br(a).

A similar definition applies for local maximum points.

Theorem. If F is C1 and a is a local minimum or maximum point

for F,then

## 12.8 LAGRANGE MULTIPLIERS

We are often interested in the problem of investigating the maximum and minimum points of a real-valued function F restricted to some manifold M in Rn.

Definition. Suppose M is a manifold in Rn. The function F: Rn ^ R has a local minimum (maximum) at a E M vjhen F is constrained to M if for some r > 0,

F(a) < (>) F(x)

for all x E Br(a).

If F has a local (constrained) minimum at a E M then it is intuitively reasonable that the rate of change of F in any direction h in TaM should be zero. Since

DhF(a) = VF(a) • h,

this means VF (a) is orthogonal to any vector in TaM and hence belongs to NaM. We make this precise in the following Theorem.

Theorem. (Method of Lagrange Multipliers) Assume M be a man- ifold in Rn given locally as the zero-level set of T:Rn ^ R^ 11.

Thus $ is C1 and for each x E M the vectors V$1(x),...,VF<=(x) are linearly independent.

Suppose

F: Rn R

is C1 and F has a constrained minimum (maximum) at a E M. Then

VF (a) = <= A,(a)

}=i

for some Ai, . .., Ae E R called Lagrange Multipliers.

Equivalently,assume H: Rn+ ^ R be defined by

H (xi,..., Xn, ai, ..., ae) = F (xi,. . .,Xn)-(Ji^1(Xi,. . .,Xn)-. . .-cri¥(xi,. . .,Xn). Then H has a critical point at ai, ..., an, Ai, ..., Ae for some Ai, ..., Ae

proof: Suppose 0: I ^ M where I is an open interval containing 0,-0(0) = a and 0 is Ci.

Then F(0(t)) has a local minimum at t = 0 and so by the chain rule

n dF d0-

0 = § dx-(a) (0)

i.e.

Notes

VF(a) ± 0'(0).

Since 0'(0) can be any vector in TaM,it follows VF(a) E NaM. Hence

VF (a) = <= A,V<I>J (a)

J=i

for some Ai, ..., A<=. This proves the first claim.

For the second claim just note that

dH = 0^ = _$J

dx- dx- jaj dx-, 8aj

J J

Since (a) = 0 it follows that H has a critical point at ai, ..., an, Ai, ..., Ai iff

dF <,8& ()

-—(a) = > A, —— (a)

dx-y) = j dx- K)

for i = 1, .. ., n. That is,

VF (a) = <= A,VSJ (a).

j =i

Example Find the maximum and minimum points of

F(x,y, z) = x + y + 2z

on the ellipsoid

M = {(x,y,z) : x2 + y2 + 2z2 = 2}.

Solution: Assume

$(x,y,z) = x2 + y2 + 2z2 — 2. At a critical point there exists A such that

VF = AVT.

That is

1 = A(2x)

= A(2y)

= A(4z).Moreover

These four equations give

1

2A' * 2A' ~ 2A' AHence

(x, y, z) = ±^(1, 1, 1).

Since F is continuous and M is compact,F must have a minimum and a maximum point. Thus one of ±(1, 1, 1)/v^2 must be the minimum point and the other the maximum point. A calculation gives

F (t5(1' 1-1}) = 2V2 F (—X5(1' = —2V5-

Thus the minimum and maximum points are — (1,1,1) /\[2 and +(1,1,1) /\[2 respectively.

.

Check your Progress - 1

Discuss Differentiation Of Vector-Valued Functions

__________________________________________________________

__________________________________________________________

__________________________________________________________

Discuss Partial And Directional Derivatives

_________________________________________________________

__________________________________________________________

__________________________________________________________

## 12.9 LET US SUM UP

In this unit we have discussed the definition and example of Differentiation Of Vector-Valued Functions, Partial And Directional Derivatives, The Chain Rule, The Inverse Function Theorem And Its Applications, Implicit Function Theorem, Maximum,Minimum,And Critical Points, Lagrange Multipliers, Partial And Directional Derivatives, The Chain Rule, The Inverse Function Theorem And Its Applications, Implicit Function Theorem, Maximum,Minimum,And Critical Points, Lagrange Multipliers

## 12.10 KEYWORDS

1.Differentiation Of Vector-Valued Functions: In this chapter we consider functions f: D (C R") — R"

2/.Partial And Directional Derivatives: The ith partial derivative of f at x is defined by df n _ uf(x + tei) — f(x)

3.The Chain Rule: Motivation The chain rule for the composition of functions of one variable

4. Implicit Function Theorem f (x) = y, where x = (x1xn) and y = (y1,..., yn).

5. Maximum, Minimum, And Critical Points: In this section suppose F:Q (c Rn) ^ R,where Q is open.

6. Lagrange Multipliers: We are often interested in the problem of investigating the maximum and minimum points of a real-valued function F restricted to some manifold M in Rn.

## 12.11 QUESTIONS FOR REVIEW

Explain Differentiation Of Vector-Valued Functions

Explain Partial And Directional Derivative

## 12.12 REFERENCES


- Analysis of Several Variables
- Application of Several Variables
- Function of Variables
- System of Equation
- Function of Real Variables
- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables


## 12.13 ANSWERS TO CHECK YOUR PROGRESS

Differentiation Of Vector-Valued Functions

(answer for Check your Progress - 1 Q)

Partial And Directional Derivative

(answer for Check your Progress - 1 Q)

# UNIT - 13 : MULTIVARIABLE DIFFERENTIAL CALCULUS

STRUCTURE

13.0 Objectives

13.1 Introduction

13.2 Multivariable Differential Calculus

13.3 The Derivative

13.4 Inverse Function And Implicit Function Theorem

13.5 Let Us Sum Up

13.6 Keywords

13.7 Questions For Review

13.8 References

13.9 Answers To Check Your Progress

## 13.0 OBJECTIVES

After studying this unit you should be able to:
Learn Understand about Multivariable Differential Calculus

Learn Understand about The Derivative

Learn Understand about Inverse Function And Implicit Function Theorem

## 13.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

# 13.2 MULTIVARIABLE DIFFERENTIAL CALCULUS

Introduction

This chapter develops differential calculus on domains in n-dimensional Euclidean space Rra.

In the derivative of a function F : O ^ Rm,where O is an open subset of Rra,as a linear map from Rra to Rm. We establish some basic properties,such as the chain rule. We use the one-dimensional integral as a tool to show that,if the matrix of first order partial derivatives of F is continuous on O,then F is differentiable on O. We also discuss two convenient multi-index notations for higher derivatives,and derive the Taylor formula with remainder for a smooth function F on O C Rra.

we establish the Inverse Function Theorem,stating that a smooth map F : O ^ Rra with an invertible derivative DF(p) has a smooth inverse defined near q = F(p). We derive the Implicit Function Theorem as a consequence of this. As a tool in proving the Inverse Function Theorem,we use a fixed point theorem known as the Contraction Mapping Principle.

In systems of differential equations. We establish a basic existence and uniqueness theorem and also study the smooth dependence of a solution on initial data. We interpret the solution operator as a flow generated by a vector field and introduce the concept of the Lie bracket of vector fields. We also consider the linearization of a system of ODEs about a solution. Within the setting of linear systems,we introduce the matrix exponential as a tool and derive a number of its basic properties.

# 13.3 THE DERIVATIVE

**Notes**

Assume O be an open subset of Rn,and F : O ^ Rm a continuous function. We say F is differentiable at a point x € O,with derivative L,if L : Rn ^ Rm is a linear transformation such that,for y € Rn,small,

F(x + y) = F(x) + Ly + R(x,y)

RXiyfW 0 as y ^ 0.

We denote the derivative at x by DF(x) = L. With respect to the standard bases of Rn and Rm,DF(x) is simply the matrix of partial derivatives,

(dFi/dxi

(dF)

\dxk)

\dFm/dxi

dFm/dxn) so that,if v = (v\, ...,vn)f,(regarded as a column vector) then

fJ2(dF\/dxk)vk\

k

DF (x)v =

T, (dFm/dxk)vk

\ k /Recall the definition of the partial derivative dfj/dxk. It will be shown below that F is differentiable whenever all the partial derivatives exist and are continuous on O. In such a case we say F is a C1 function on O. More generally,F is said to be Ck if all its partial derivatives of order < k exist and are continuous. If F is Ck for all k,we say F is C™.

we can use the Euclidean norm on Rn and Rm. As this norm is defined bfor x = (x1, ...,xn) € Rn.

An application of the Fundamental Theorem of Calculus,to functions of each variable xj separately,yields the following. If we assume F : O ^ Rm is differentiable in each variable separately,and that each dF/dxj is continuous

on O,then

n

F(x + y) = F(x) + ^2 [F(x + Zj) - F(x + zj-i)] j=i

= F(x) + Y^ Aj(x,y)yj

j=i

,/ x f1 dF (       ) 1

aj(x,y) = Jo dx,(x + zj-i + tyj j ^

where zo = 0,zj = (yi, ...,yj,0, ...,0), and {ej} is the standard basis of Rn. Consequently,

n dF

F(x + y) = F(x) + <= — (x) yj + R(x,y),

j=i

j=1

Now  implies F is differentiable on O,as we stated below . Thus we have established the following.

Proposition. If O is an open subset of Rn and F : O ^ Rm is of class C1,then F is differentiable at each point x € O.

As is shown in many calculus texts,one can use the Mean Value Theorem instead of the Fundamental Theorem of Calculus,and obtain a slightly more general result.

Assume us give some examples of derivatives. First,take n = 2,m = 1,and

set

F (x) = (sin xi)(sin x2).

Then

DF(x) = ((cosxi)(sinx2), (sinxi)(cosx2)).

Next, take n = m = 2 and

F(x) = (

x2 - x22

Then

<2-L11>          df<*> = (<=i .

We can replace Rn and Rm by more general finite-dimensional real vector spaces, isomorphic to Euclidean space. For example, the space M(n,R) of real n x n matrices is isomorphic to Rra . Consider the function

5 : M(n, R) —> M(n, R), 5(X) = X2. We have

(X + Y)2 = X2 + XY + YX + Y2

2

= X2 + DS (X)Y + R(X, Y),

with R(X, Y) = Y2,and hence

DS(X)Y = XY + YX.

For our next example,we take

O = Gl(n,R) = {X e M(n,R) : det X = 0},  which,as shown below,is open in M(n,R). We consider

$ : Gl(n,R) —> M(n,R), $(X) = X-1,

and compute D$(I). We use the following. If,for A e M(n,R),

= sup{||Av|| : v e Rra,\\v\\ < 1},

A,B e M(n,R) ^ || A + B\\ < ||A| + ||B||

and

||AB||<||A||■||B||,  so Y e M(n,R) ^ WYk|< ||Y f.

Sk = I - Y + Y2          + (-1)kYk

^ YSk = SkY = Y - Y2 + Y3  + (-1)kYk+1

^ (I + Y)Sk = Sk (I + Y) = I + (-1)k Yk+1

$(I + Y)-1 = J2(-1)kYk = I - YD\$(I)Y = -Y.$

Related calculations show that Gl(n,R) is open in M(n,R). In fact,given X e Gl(n,R), Y e M(n,R),

$X + Y = X (I + X-1Y)$, is invertible as long as

$\|X-1Y\| < 1.$

One can proceed from here to compute D$(X). See the exercises.

We return to general considerations,and derive the chain rule for the derivative. Assume F : O ^ Rm be differentiable at x GO,as above,assume U be a neighborhood of z = F (x) in Rm,and assume G : U ^ Rk be differentiable at z. Consider H = G o F. We have

$H (x + y) = G(F (x + y))$

$= G(F(x) + DF(x)y + R(x,y))$

$= G(z) + DG(z) (DF(x)y + R(x,y)) + Ri(x,y)$

$= G(z) + DG(z)DF (x)y + R2(x,y)$

with

Wy'11 ^ 0 as y ^ 0.

yyy

Thus G o F is differentiable at x,and

$D(G o F)(x) = DG(F(x)) \cdot DF(x).$

Another useful remark is that,by the Fundamental Theorem of Calculus, applied to p(t) = F(x + ty),

$F(x + y) = F(x) + / DF(x + ty)y dt,$

J 0

provided F is C1. For a typical application

For the study of higher order derivatives of a function,the following result is fundamental.

Proposition. Assume F : O ^ Rm is of class C2,with O open in Rra. Then,for each x GO,1 < j, k < n,

^ d dF d dF . .

dx 0x~k (x) = dxk dxj(x).

Proof. It suffices to take m = 1. We label our function f : O ^ R. For

1 < j < n,we set

(TP28) Ajhf (x) = -1 (f (x + hej) - f (x)),

where {e1, ...,en} is the standard basis of Rn. The mean value theorem (for functions of xj alone) implies that if dj f = df/dxj exists on O,then,for x G O,h > 0 sufficiently small,

Aj, hf (x) = dj f (x + aj hej),

for some aj <= (0, 1), depending on x and h. Iterating this,if dj (dk f) exists on O,then,for x <= O,h > 0 sufficiently small,

^k, h^j, hf (x) = dk (Ajf)(x + ak hek)

= Aj, h(dk f)(x + ak hek)

= dj dk f (x + ak hek + aj hej), with aj,ak <= (0, 1). Here we have used the elementary result

dk Aj, h f = Aj, h(dk f).

We deduce the following.

Proposition. If dk f and djdk f exist on O and djdk f is continuous at x0 <= O,then

dj dk f (xo) = lim Ak, hAj, hf (xo).

h^0

Clearly

(2W33) Ak, hAj, hf = Aj, hAk, hf,

so we have the following,which easily implies Propo

Corollary. If dj f and dkdj f exist on O and dkdj f is continuous at x0,then

dj dk f (xo) = dk dj f (xo).

We now describe two convenient notations to express higher order deriva- tives of a Ck function f : Q ^ R,where Q C Rn is open. In one,assume J be a fc-tuple of integers between 1 and n; J = (jl, ...,jk). We set

d

(2.^35) f (J)(x) = djk • •• djif (x), dj = d^r..

We set\J\ = k,the total order of differentiation. As we have seen in Proposi- tion 2.1.2,didjf = djdif provided f <= C2(Q). It follows that,if f <= Ck(Q), then djk • • • dj1 f = dlk • • • dll f whenever [ll, ...,lk} is a permutation of [ji, ...,jk}. Thus, another convenient notation to use is the following. Assume a be an n-tuple of non-negative integers, a = (al, ...,an). Then we set

f (a)(x) = df1 ••• dan f (x), \a\ = ai + + an.

Note that,if\J\ = \a\ = k and f <= Ck(Q),

f(J)(x) = f(a)(x), with ai = #[l : ji = i}.

Correspondingly, there are two expressions for monomials in x = (x-\^, ...,xn): xJ = x • . . . x ■ xa = xa1 ■ ■ ■ xan

^ • 14 CP J j 1 j k J 1 'n. J

and xJ = xa provided J and a are related. Both these notations are called "multi-index" notations.

We now derive Taylor's formula with remainder for a smooth function F : Q ^ R,making use of these multi-index notations.

2 f"(0)i2+■■■+-

with

1 r t

rk(t) = - J (t — s)kf(k+1)(s) ds,

given f <= Ck+1 (I), I = (-a,a).

Assume us assume 0 <= Q, and that the line segment from 0 to x is contained in Q. We set f (t) = F(tx), with t = 1. Applying the chain rule,we have

n

f'(t) = ^ dj F (tx)xj = 22 F (J)(tx)xJ.

j=1 |J|=1

Differentiating again,we have

f"(t) = ^ F(J+K)(tx)xJ+K = ^ F(J)(tx)xJ,

|J| = 1, |K| = 1 |J|=2

where,if\J\ = k,\K\ = I,we take J + K = (j1, ...,jk, k1, ...,kg). Inductively,we have

f(k)(t) = ^ F(J)(tx)

|J|=k

Hence,with t = 1,

F(x) = F(0) + J2 F(J)(0)xJ + ■ ■ ■ + -1 ^ F(J)(0)xJ + Rk(x),

J |=1 " J=

F(x) = ^2 JF(J)(0)xJ + Rk(x),

J ^k

where

Rk(x) = 1 22 (yj (1 — s)kF(J)(sx) d2jxJ.

|J|=k+1 Jo

This gives Taylor's formula with remainder for F <= Ck+1(Q), in the J-multi- index notation.

We also want to write the formula in the a-multi-index notation. We have

^ F(J)(tx)xJ = ^ v(a)F(a)(tx)xa,

\J\=k \a\=k

where

v(a) = #{J : a = a(J)},

and we define the relation a = a(J) to hold provided the condition holds,or equivalently provided xJ = xa. Thus v(a) is uniquely defined by

^ v(a)xa = ^ xJ = (xi ++ xn)k.

\a\=k \J\=k

One sees that,if |a| = k,then v(a) is equal to the product of the number of combinations of k objects,taken a1 at a time,times the number of combinations of k — a1 objects,taken a2 at a time,• • • times the number of combinations of k — (a1 + • • • + an-1) objects,taken an at a time. Thus

v(a) = (k ^ (k — a^ (k — a1 — • • • — an-1^ Va1/ V a2 / V an J

k!

a1la2l • • • an!

In other words,for|a| = k,

k!

v(a) = —,where a! = a1! • • • an!

a!

Thus the Taylor formula can be rewritten

F (x)= V 1F (a)(0)xa + Rk (x),

a!

\a\<k

where

Rk(x) = ^ k + 1 ^ / (1 — s)kF(a\sx) ds^Jxa.

**Notes**

|a|=k+1 a!

holds for F € Ck. In fact,for such F,  with k replaced by k — 1,to get

F (x)= V 1F (a)(0)xa + Rk- 1(x),

■, a!

\a\<k—1

with

Rk-1 (x) = ^ ~"r(/ (1 — s)k-1 F(a)(sx) ds^Jxa.

\a\=k a!

We can add and subtract F(a)(0) to F(a)(sx) in the integrand above,and obtain the following.

Proposition. If F € Ck on a ball Br(0), the formula holds for x € Br (0), with

Rk(x)= ^ (1 - s)k"1[F(a)(sx) - F(a)(0)] ds^xa.

\a\=k (X'

\Rk(x)| < V ^ sup \F(a)(sx) - F(a)(0)\.

a! o<s<i

\a\=k — —

The term corresponding to\J\ = 2,or \a\ = 2

It is

i i n dp,F

J2f<j\°) j = 2 (») xjxk.

\J\=2 j, k=1 k j

We define the Hessian of a C2 function F : O ^ R as an n x n matrix:

d2f y =(dSj .

Then the power series expansion of second order about 0 for F takes the form

F(x) = F(0) + DF(0)x + 1 x ■ D2F(0)x + R2(x),

\ 2 sup \F(a) (sx) F(a 0<s<1, \a\=2

In all these formulas we can translate coordinates and expand about y € O.

F(x) = F(y) + DF(y)(x -y) + 1(x -y) ■ D2F(y)(x -y) + R2(x, y), with

|2 sup |F(a o<s<1, \a\=2

Example. If we take F(x) as so DF(x) is as in then

2 - sin x1 sin x2 cos x1 cos x2

D2F (x) = ...

cos x1 cos x2 - sin x1 sin x2

The results are useful for extremal problems,i.e.,deter- mining where a sufficiently smooth function F : O ^ R has local maxima and local minima. Clearly if F € Cl(O) and F has a local maximum or minimum at x0 € O,then DF(x0) = 0. In such a case,we say x0 is a critical point of F. To check what kind of critical point xo is,we look at the n x n matrix A = D2F(x0), assuming F € C2(O). A is a symmetric matrix.

A basic result in linear algebra is that if A is a real,symmetric n x n matrix, then Rn has an orthonormal basis of eigenvectors,{vl, ...,vn},satisfying Avj = XjVj; the real numbers Xj are the eigenvalues of A.

We say A is positive definite if all Xj > 0,and we say A is negative definite if all Xj < 0. We say A is strongly indefinite if some Xj > 0 and another Xk < 0. Equivalently,given a real,symmetric matrix A,

A positive definite v ■ Av > C\v\2, A negative definite v ■ Av < —C\v\,

for some C > 0,all v € Rn,and

A strongly indefinite 3 v, w € Rn,nonzero,such that v ■ Av > C\v\2,w ■ Aw < —C\w\2,

for some C > 0.

Proposition. Assume F € C2(O) is real valued,O open in Rn. Assume x0 € O be a critical point for F. Then

D2F(x0) positive definite ^ F has a local minimum at x0,

D2F(x0) negative definite ^ F has a local maximum at x0,

D2F (x0) strongly indefinite ^ F has neither a local maximum nor a local minimum at x0.

In case (iii), we say x0 is a saddle point for F.

The following is a test for positive definiteness.

Proposition. Assume A = (aij) be a real,symmetric,n x n matrix. For 1 < 1 < n,form the 1 x 1 matrix Ai = (aj)i<i, j<<=. Then

A positive definite det Ai > 0,V I € {1, ...,n}.

Regarding the implication ^,note that if A is positive definite,then det A = det An is the product of its eigenvalues,all > 0,hence is > 0. Also in this case,it follows from the hypothesis on that each Ai must be positive definite,hence have positive determinant,so we have ^.

The implication ^ is easy enough for 2 x 2 matrices. If A is symmetric and det A > 0,then either both its eigenvalues are positive (so A is positive definite) or both are negative (so A is negative definite). In the latter case, Ai = (an) must be negative,so we have ^ in this case.

We prove ^ for n > 3,using induction. The inductive hypothesis implies that if det Al > 0 for each 1 < n,then An-\ is positive definite. The next lemma then guarantees that A = An has at least n — 1 positive eigenvalues. The hypothesis that det A > 0 does not allow that the remaining eigenvalue be < 0,so all the eigenvalues of A must be positive.

Lemma.If An-i is positive definite,then A = An has at least n — 1 positive eigenvalues.

Proof. Since A is symmetric,Rn has an orthonormal basis v\, ..., vn of eigenvectors of A; Avj = XjVj. If the conclusion of the lemma is false,at least two of the eigenvalues,say Xi, X2,are < 0. Assume W = Span(vi,v2), so

w <= W =^ w • Aw < 0.

Since W has dimension 2,Rn-i c Rn satisfies Rn-i n W = 0,so there exists a nonzero w <= Rn-i n W,and then

w • An-iw = w • Aw < 0,

contradicting the hypothesis that An-i is positive definite. □

Remark. Given we see by taking A ^ —A that if A is a real, symmetric n x n matrix,

A negative definite (—1)1 det A<= > 0,V I e{1, ..., n}.

We return to higher order power series formulas with remainder and complement 1/(k + 1) times a weighted average of ^(k+i1 (s) over s <= [0,t]. Hence we can write

rk(t) = + v(k+l)(dt), for some d <= [0, 1],

if ^ is of class Ck+i. This is the Lagrange form of the remainder; see Appendix A.4 for more on this,and for a comparison with the Cauchy form of the remainder. If y is of class Ck,we can replace k + 1 by k in and write

y(t) = y(0) + <p'(0)t + • • • + k-^y(k-i)(0)tk-i + 1 <p(k)(0t)tk,

for some 0 € [0, 1].

for <^(t) = F(tx) gives

F(x) = <= JFJ(0)xJ + 1 <= F(J\dx)xJ,

\J\<k-1 1 ' \J\=k

for some 0 € [0, 1] (depending on x and on k, but not on J), when F is of class Ck on a neighborhood Br(0) of 0 € Rn. Similarly, using the a-multi-index notation

Notes

F (x)= <= -1 F(a)(0)xa 1 F(a)(0x)xa,

\a\<k-1 \a\=k

for some d € [0, 1] (depending on x and on |a|, but not on a), if F € Ck(Br (0)). Note also that

i i n d 2 F

- ^ F(J)(0x)xJ = - ——-— (0x)xjxk ,x 2 ^ v ! 2 ^ dxkOxi ! j

\J\=2 j, k=1 k j

= 1 x ■ D2F(0x)x,

with D2F(y) as in,so if F € C2(Br(0)), we have, as an alternative F(x) = F(0) + DF(0)x + 1 x ■ D2F(0x)x,

for some 0 € [0, 1].

We next complemant the multi-index notations for higher derivatives of a function F by a multi-linear notation,defined as follows. If k € N,F € Ck(U), and y € U C Rn,set

DkF(y)(ui,..., Uk) = dq ■ ■■ dtkF(y + tiui + + tkUk)

for u1, ...,uk € Rn. For k = 1,this formula is equivalent to the definition of DF given at the beginning of this section. For k = 2,we have

D2F(y)(u,v) = u ■ D2F(y)v,

with D2F(y) defines DkF(y) as a symmetric,k-linear form in u1, ...,uk € Rn.

J-multi-index notation as follows.

We start with

dti F (y + tu + + tk uk) = ^ F (J)(y + Etj uj)uJ

\J\=1

and inductively obtain

9t! ••• dtkF(y + TtjUj) = ^ F(Jl+-+Jk)(y + EjUj)uJl ••• uJkk,

\Ji\=~=\Jk\=i

hence

Dk F (y)(ui, ..., uk) = ^ F (Jl+"+Jk)(y)uJ1 ••• uJk.

\Ji\=• • •=\Jk\=i In particular,if u1 = • • • = uk = u,

Dk F (y)(u,..., u)= ^ F (J)(y)uJ.

\J\=k

Hence the multi-linear Taylor formula with remainder

F (x) = F (0) + DF (0)x + ••• + * Dk-1F (0)(x,..., x)

(k 1)!

+A"DkF(9x)(x, ...,x), k!

for some 9 € [0, 1],if F € Ck(Br(0)). In fact,rather than appealing to we can note that

<^(t) = F(tx) =^ ^(k)(t) = dti ••• dtkf(t + ti +  + tk)

ti=• • •=tk=0

= DkF(tx)(x, ...,x), and also use the notation

Dj F (y)x®j = Dj F (y)(x,..., x),

with j copies of x within the last set of parentheses,and as

F (x) = F (0) + DF (0)x + • • • + 1 Dk-1F (0)x®(k-1)

(k " 1)!

+-1 Dk F (9x)x®k. k!

Convergent power series and their derivatives

Here we consider functions given by convergent power series, of the form

F(x) = J2 bo

vaJ

a> 0

We allow ba € C,and take x = (x1, ...,xn) € Rn,with xa given. Here is our first result.

Proposition. Assume there exist y € Rn and C0 < o such that

\yk\ = ak > 0,Vk,\baya\ < Co,Va.

Then,for each 5 € (0, 1), the series  converges absolutely and uniformly on each set

Rs = {x € Rn : \xk\ < (1 — 5)ak,Vk}.

The sum F(x) is continuous on R = {x € Rn : \xk\ < ak,Vk}.

Proof. We have

x € Rs =^\baxa\< Co(1 — 5)H,Va,  hence

Y\b»xa\< Co <=(1 — 5)lal < o.

a>0      a>0

Thus the power series is absolutely convergent whenever x € Rs. We also have,for each N € N,

F(x)= <= baxa + RN(x),

\a\<N

and,for x € Rs,

\Rn(x)\ < Y \b»xa\

\a\>N

< Co Y (1 — 5)H

\a\>N

= <=n —— 0 as N —y co.

This shows that RN(x) — 0 uniformly for x € Rs

We next discuss differentiability.

Proposition. F is differentiable on R and,for each j € {1, ...,n},

dF

d—j(x) = Yj ajbaxa-<=j,Vx € R.

3 a><=j

Here,we set <=j = (0, ...,1, ...,0), with the 1 in the jth slot. It is convenient to begin with the following.

Lemma. In the setting for each j € {1, ...,n}, Gj(x) = ^ ajbaxa-<=j

is absoassumeely convergent for x € R,uniformly on Rs for each 5 € (0, 1), therefore defining Gj as a continuous function on R.

Proof. Take a = (ai, ...,an), with aj as Given x € Rs,we have ^ aj\baxa-<=j\< ^ aj(1 - 5)\a\-l\baaa-<=j |

a><=j a><=j

C

< S aj(1 - 5>a.

and this is

< Ms < &>,V5 € (0, 1).

This gives the asserted convergence on Rs and hence defines the function Gj,continuous on R. □

we need to show that

dF

gj = Gj on R,

for each j. Assume us use the notation

xj = (xi, ...,xj-i,0,xj+i, ...,xn) = x — xjej,

where ej is the jth standard basis vector of Rn. Now,given x € Rs,5 € (0, 1), the uniform convergence of on Rs implies X

Gj (xj + tej) dt = ajba (xj + tej)a <=j dt

Jo

a><=j

V,ajbaa-ixa

a><=j = ^ baxa

= F (x) — F (Xj).

Applying d/dxj to the left side and using the fundamental the- orem of calculus then yields  as desired. This gives the identity. Since each Gj is continuous on R,this implies F is differentiable on R.

We can iterate obtaining dkdjF(x) = dkGj(x) as a convergent power series on R,etc. In particular,we have the following.

Corollary. F € C ^(R). Consider the following function f : R2 R:

f (x, y) = (cos x)(cos y).

Find all its critical points,and determine which of these are local maxima, local minima,and saddle points.

Assume M(n,R) denote the space of real n x n matrices. Assume F, G : M(n,R) ^ M(n,R) are of class C1. Show that H(X) = F(X)G(X) defines a C1 map H : M(n,R) M(n,R), and

DH (X)Y = DF (X)YG(X) + F (X)DG(X)Y.

Assume Gl(n,R) C M(n,R) denote the set of invertible matrices. Show that

$ : Gl(n,R) —> M(n,R), $(X) = X_1 is of class C1 and that

D$(X)Y = -X-1YX-1.

Identify R2 and C via z = x + iy. Then multiplication by i on C corresponds to applying

' - C -')

Assume O C R2 be open,f : O R2 be C1. Say f = (u, v). Regard Df (x, y) as a 2 x 2 real matrix. One says f is holomorphie,or complex-analytic, provided the Cauchy-Riemann equations hold:

du dv du dv dx dy dy dx'

Show that this is equivalent to the condition

Df (x, y) J = J Df (x, y).

Generalize to O open in Cm,f : O ^ Cn.

Assume f be C1 on a region in R2 containing [a, b] x {y}. Show that,as h ^ 0,

1 df

h f (x,y + h) - f (x,y)] —> dy (x,y), uniformly on [a,b] x {y}.

Hint. Show that the left side is equal to

1 fh df h j o

and use the uniform continuity of df/dy on [a,b] x [y-5,y+5]

Show that

d [ f (x, y) dx = [ f (x, y) dx. dy J a J a dy

Considering the power series

f(j) (y)

f (x) = f (y) + f (y)(x - y) + ••• + ■,(x - y)j + Rj(x,y),

show that

dR . 1

j f (j+1)(y)(x - y)j,Rj(x, x) = 0.

dy j

Use this to re-derive and hence

We define "big oh" and "little oh" notation:

f (x)

f (x) = O(x) (as x ^ 0) ^

x

f (x)

f (x) = o(x) (as x ^ 0) ^ > 0 as x ^ 0.

Assume O C Rra be open and y GO. Show that

f G Ck+1 (O) ^ f (x) = <= 1 f (a)(y)(x - y)a + 0(|x - y\k+1),

\Kk a!

f G Ck(O) ^ f (x) = <= 1 f (a)(y)(x - y)a + o(|x - y\k).

a!

Ia<k

Assume G : U -G O,F : O -G Q. Show that F, G G C1 F o G G C1.More generally,show that,for k € N,

F, G € Ck F o G € Ck.

Hint. Write H = F o G,with hi(x) = fl(g1(x),...,gn(x)), and use to get

n

djhi(x) = Y^ dk fe(gi, ...,gn)djgk■

k=1

Show that this yields. To proceed,deduce from that

jjhi(x) = Y dkidk2fi(gi, ..., gn)(djigk!Xjgk2)

ki, k2 = 1 n

+ y ^ dk fl(g1,. . .,gn)dji dj2gk.

k=1

Use this to get for k = 2. Proceeding inductively,show that there exist constants C(g,J#,k#) = C(g,J1, ...,J^, k1, ...,kM) such that if F, G € Ck and\J\ < k,

h{eJ)(x) = Y C(g,J#,k#)gJ • •• gl;)fHk1-k")(g1, ...,gn),

where the sum is over

g < \J\,J1 + ••• + J ~ J7 \JV \F 1,

and J1 + • • • + J^ ~ J means J is a rearrangement of J1 + • • • + J^. Show that follows from this.

Show that the map T : Gl(n,R) ^ Gl(n,R) given by T(X) = X-1 is Ck for each k,i.e.,T € C^.

Hint. Start with the material of Exercise 3. Write DT(X)Y = —X-1YX-1

as

d

dimT(X) = - T(X) = DT(X)Eim = -$(X)E*m$(X),

dx<=m

where X = (xlm) and Elm has just one nonzero entry,at position (I,m). Iterate this to get

dl2m2 dlimi

T(X) = -(d^T(X))EiimiT(X) - T(X)Elimi(dhm2T(X)), and continue & deal with properties of the determinant,as a differentiable function on spaces of matrices.

Assume M(n,R) be the space of n x n matrices with real coefficients,det : M(n,R) ^ R the determinant. Show that,if I is the identity matrix,

D det(I)B = Tr B,

i.e.,

d

— det(I + tB) |t=o = Tr B.

If A(t) = (ajk(t)) is a smooth curve in M(n,R), use the expansion of (d/dt)det A(t) as a sum of n determinants,in which the rows of A(t) are successively differentiated,to show that

d dt

and deduce that,for A, B € M(n,R),

D det(A)B = Tr(Cof(A)4 • B).

Suppose A € M(n,R) is invertible. Using

det(A + tB) = (det A) det(I + tA-1B),

show that

D det(A)B = (det A)Tr(A-1B).

Comparing this result with that of Exercise 12,deduce a second proof of Cramer's formula:

(det A)A-1 = Cof(A)4.

## 13.4 INVERSE FUNCTION AND IMPLICIT FUNCTION THEOREM

The Inverse Function Theorem gives a condition under which a function can be locally inverted. This theorem and its corollary the Implicit Function Theorem are fundamental results in multivariable calculus. First we state the Inverse Function Theorem. Here,we assume k > 1.

Theorem. Assume F be a Ck map from an open neighborhood Q of p0 € Rra to Rra,with q0 = F(p0). Suppose the derivative DF(p0) is invertible. Then there is a neighborhood U of p0 and a neighborhood V of q0 such that F : U V is one-to-one and onto,and F-1 : V ^ U is a Ck map. (One says F : U ^ V is a diffeomorphism.)

First we show that F is one-to-one on a neighborhood of p0,under these hypotheses. In fact,we establish the following result,of interest in its own right.

Proposition. Assume Q C Rn is open and convex,and assume f : Q ^ Rn be C1. Assume that the symmetric part of Df (u) is positive-definite,for each u € Q. Then f is one-to-one on Q.

Proof. Take distinct points u1,u2 € Q,and set u2 — u1 = w. Consider p : [0, 1] ^ R,given by

p(t) = w ■ f (u1 + tw).

Then p'(t) = w ■ Df (u1 + tw)w > 0 for t € [0, 1],so p(0) = p(1). But p(0) = w ■ f (U1) and p(1) = w ■ f (u2), so f (U1) = f (u2). □

assume us set

f (u) = A(F (po + u) — q0), A = DF (p0)_1.

Then f (0) = 0 and Df (0) = I,the identity matrix. We will show that f maps a neighborhood of 0 one-to-one and onto some neighborhood of 0. We can write

f (u) = u + R(u), R(0)=0,DR(0) = 0, and R is C1. Pick b > 0 such that

||u||< 2b =^||DR(u)||< 2.

Then Df = I + DR has positive definite symmetric part on

B2b(0) = {u € Rn : ||u|| < 2b},

f : B2b(0) —> Rn is one-to-one.

We will show that the range f (B2b(0)) contains Bb(0), that is to say,we can solve

f (u) = v,

given v € Bb(0), for some (unique) u € B2b(0). This is equivalent to u + R(u) = v.

To get the solution,we set

Tv (u) = v — R(u).

Then solving is equivalent to solving

Tv (u) = u.

We look for a fixed point

$u = K (v) = f-1(v)$.

Also,we want to show that $DK(0) = I$,i.e.,that

$K (v) = v + r(v), r(v) = o(||v||)$.

It follows that,for y close to qo,$G(y) = F—1(y)$ is defined. Also taking

$x = po + u,y = F(x), v = f (u) = A(y - qo), G(y) = po + u = po + K (v)$

$= po + K(A(y - qo))$

$= po + A(y - qo) + o(||y - qo||)$.

Hence G is differentiable at qo and

$DG(qo) = A = DF (po)-1$.

A parallel argument, with po replaced by a nearby x and $y = F(x)$, gives

$DG(y) = DF (G(y))-1$.

Thus our task is  To do this, we use the following general result,known as the Contraction Mapping Theorem.

Theorem. Assume X be a compe metric space, and assume T : X ^ X satisfy

$dist(Tx, Ty) < r dist(x, y)$,

for some $r < 1$. (We say T is a contraction.) Then T has a unique fixed point x. For any yo € X,Tkyo x as k ^ <x>.

Proof. Pick yo € X and assume $yk = Tkyo$. Then $dist(yk, yk+i) < rk dist(yo, yi)$, so

$dist(yk, yk+m) < dist(yk, yk+i)+ \quad + dist(yk+ m— 1,yk+m)$

$< \{rk ++ rk+m—1) dist(yo, yi)$

$< rk(1 - r)— 1 dist(yo, y1)$.

It follows that (yk) is a Cauchy sequence,so it converges; yk ^ x. Since $Tyk = yk+1$ and T is continuous,it follows that $Tx = x$,i.e.,x is a fixed point. Uniqueness of the fixed point is clear from the estimate dist(Tx,Tx') < r dist(x, x'), which implies dist(x, x') = 0 if x and x' are fixed points.

||v|| < b =^ Tv : Xv Xv,

Tv : Xv ^ Xv

where

Xv = {u e B2b(0) : \\u — v\\ < Av},

Av = sup \\R(w)\\.

IMI<2|MI

||w||< 26 ^||R(w)||< 1 ||w||,and ||R(w)|| = o(||w||).

Hence

\\v\\ < b Av < ||vB,and Av = o(||v||).

Thus \\u — v|| < Av ^ u e Xv. Also

u e Xv ||u|| < 2\vB

=^||R(u)||< Av

BTv(u) — vB < Av,

As for the contraction property,given Uj € Xb,||v|| < b,

\\TV(ui) - Tv(U2)| = ||E(U2) - R(ui)ll (2.2.17) i

< 2llui - U2||,

there is a unique fixed point,u = K(v) € Xv. Also,since u € Xv,

\\K(v) - v\\ < Av = o(|v|).

This establishes the existence of the inverse function G = F-i : V U,and for the derivative DG. Since G is differentiable on V,it is certainly continuous implies DG is continuous,given F € Ci(U).

To finish the proof of the Inverse Function Theorem and show that G is Ck if F is Ck,for k > 2,Thus if DF is invertible on the domain of F,F is a local diffeomorphism. Stronger hypotheses are needed to guarantee that F is a global diffeomorphism onto its range Here is a slight strengthening.

Corollary. Assume Q C Rra is open and convex,and that F : Q ^ Rra is Ci. Assume there exist n x n matrices A and B such that the symmetric part of A DF (u) B is positive definite for each u € Q. Then F maps Q diffeomorphically onto its image,an open set in Rra.

We make a comment about solving the equation F(x) = y,under the hypotheses of is close to q0. The fact that finding the fixed point for Tv in is accomplished by taking the limit of Tk(v) implies that,when y is sufficiently close to q0,the sequence (xk), defined by

xo = po,xk+i = xk + DF(po)-i(y - F(xk)),

converges to the solution x. An analysis of the rate at which xk ^ x,and F(xk) ^ y,can be made by applying F to yielding

F(xk+i) = F(xk + DF(po)-i(y - F(xk))

= F (xk) + DF (xk)DF (po)-i(y - F (xk)) + R(xk,DF(po)-i(y - F(xk))),

and hence

y - F(xfc+i) = (I - DF(xk)DF(po)-1) (y - F(xk)) + R(xk, y - F(xk)), +

with \\R(xk, y - F(xk))|| = o(||y - F(xk)||).

It turns out that replacing p0 by xk in yields a faster approximation. This method,known as Newton's method,is described in the exercises.

We consider some examples of maps to which applies. First,we look at

F:(0, to) x R _ R2,F(r.6) = ^ = (*$).

Then

n F(r 6) = (9rx 9°^ fcos 6 -r sin ^

DF(r, 6) = y Qey) ^sin 6 rcos 6),

so

det DF(r,6) = r cos2 6 + r sin2 6 = r.

Hence DF(r, 6) is invertible for all (r,6) € (0,to) x R. implies that each (r0, 60) € (0,to) x R has a neighborhood U and (x0, y0) =(r0 cos 60, r0 sin 60) has a neighborhood V such that F is a smooth diffeomorphism of U onto V. In this simple situation,it can be verified directly that

F : (0,to) x (-n,n) —> R2\{(x,0) : x < 0} is a smooth diffeomorphism.

Note that DF(1, 0) = I. Assume us check the domain of applicability. The symmetric part of DF(r, 6) is

SM)=(1 (1 -16 * Vcri r 0-

By this is positive definite if and only if

cos 6> 0, and

det S(r,6) = r cos2 6 - ^(1 - r)2 sin2 6 > 0.

Now holds for 6 € (-n/2, n/2), but not on all of (-n, n). Furthermore for (r, 6) in a neighborhood of (r0, 60) = (1, 0), but it does not hold on all of (0,to) x (-/2, n/2). It is not capture the full force of the diffomorphism property

We move on to another example replacing Rra by a finite dimensional real vector space,isometric to a Euclidean space,such as M(n,R) & Rra . As an example,consider

^ 1

Exp : M(n,R) —> M(n,R), Exp(X) = eX = ^ — Xk.

k=0

Since

Exp(Y) = I + Y + 1Y2 + •••, we have

DExp(0)Y = Y,V Y e M(n,R),

so DExp(0) is invertible implies that there exist a neighborhod U of 0 e M(n,R) and a neighborhood V of I e M(n,R) such that Exp : U V is a smooth diffeomorphism.

To motivate the next result,we consider the following example. Take a> 0 and consider the equation

x2 + y2 = a2,F (x, y) = x2 + y2.

Note that

DF(x, y) = (2x 2y), DxF(x, y) = 2x,DyF(x, y) = 2y.

The equation defines y "implicitly" as a smooth function of x if \x\ < a. Explicitly,

\x\ < a =^ y = Va2 - x2,

Similarly,defines x implicitly as a smooth function of y if \y\ < a; explicitly

\y\ < a x = ^a2 - y2.

Now,given xo e R,a > 0,there exists y0 e R such that F(x0, yo) = a2 if and only if \xo\ < a. Furthermore,

given F(xo, yo) = a2,DyF(xo, yo) = 0 ^ \xo\ < a.

Similarly,given yo e R,there exists xo such that F(xo, yo) = a2 if and only if \yo\ < a,and

given F(xo, yo) = a2,DxF(xo, yo)=0 ^ \xo\ < a.

Note also that,whenever (x,y) e R2 and F(x,y) = a2 > 0,

DF(x,y) = 0,

so either DxF(x, y) = 0 or DyF(x, y) = 0,and,as seen above whenever (xo, yo) e R2 and F(xo, yo) = a2 > 0,we can solve F(x, y) = a2 for either y

as a smooth function of x for x near x0 or for x as a smooth function of y for y near y0.

We move from these observations to the next result,the Implicit Function Theorem.

Theorem. Suppose U is a neighborhood of x0 € Rm,V a neighborhood of y0 € R*,and we have a Ck map

$F : U \times V$ -^ R*,F(x0, y0)=

Assume DyF(xo, yo) is invertible. Then the equation F(x, y) = u0 defines y = g(x, u0) for x near x0 (satisfying g(x0, u0) = y0) with g a Ck map.

To prove this,consider H : U x V ^ Rm x R^ defined by

H(x, y) = (x, F(x,y)).

(Actually,regard (x, y) and (x, F(x, y)) as column vectors.) We have

Thus DH(x0, y0) is invertible,so J = H-1 exists,on a neighborhood of (x0,u0), and is Ck,by the Inverse Function Theorem. It is clear that J(x,u0) has the form and g is the desired map.

Setis invertible,so (with (u, v) in place of y and (x, y) in place of x) implies that the equation

F (u, v, x, y)=Q) defines smooth functions

u = u(x, y), v = v(x, y),

for (x, y) near (xo, yo) = (1, 1),

satisfying with (u(1, 1),v(1, 1)) =(2,0).

Assume us next focus on the case 1 = 1 of Theorem,so

2: = (x,y) € Rn,x € Rn-\ y € R,F(z) € R.

Then DyF = dyF. If F(x0, y0) = u0,Theorem says that if

dy F (xo, yo) = 0,then one can solve

F(x, y) = uo for y = g(x, uo),

for x near xo (satisfying g(xo, uo) = yo), with g a Ck function.

the following. Set (x, y) = z = (z1, ..., z„), zo = (xo, yo).

The condition is that dZnF(zo) = 0. Now a simple permutation of variables allows us to assume

dZj F (zo) = 0,F (zo) = uo,  and deduce that one can solve

F(z) = uo,for zj = g(zi, ...,zy-i,zy+i, ...,z.n).

Assume us record this result,changing notation and replacing z by x.

Proposition. Assume Q be a neighborhood of xo € Rn. Asume we have a Ck function

F : Q —> R,F(xo) = uo,  and assume

DF(xo) = 0,i.e.,(diF(xo),..., d,,F(xo)) = 0.

Then there exists j € {1, ...,n} such that one can solve F(x) = uo for

xj = g(xi, .. ., xj-i, xj+i, .. ., x,,), with (xio, ...,xjo, ...,xno) = xo,for a Ck function g.

Remark. For F : Q ^ R,it is common to denote DF(x) by VF(x),

VF (x) = (diF (x),c)nF (x)).

Using the notation (x, y) = (x1, x2), set

F : R2 —> R,F(x, y) = x2 + y2 — x.

Then

VF(x, y) = (2x — 1,2y),

which vanishes if and only if x = 1/2,y = 0. Hence Proposition applies if and only if (x0, y0) = (1/2,0).

Assume us give an example involving a real valued function on M(n,R), namely

det : M(n,R) —> R.

if det X = 0,

D det(X) Y = (det X) Tr(X-1Y), so

det X = 0 =^ Ddet(X) = 0.

We deduce that,if

Xo € M(n,R), det Xo = a = 0,  then,writing

X = (xjk ^iKjfeKni there exist /i, v € {1, ...,n} such that the equation det X = a has a smooth solution of the form

x^v = g(xay : (a,ft) = (g,v)),

such that,if the argument of g consists of the matrix entries of X0 other than the g,v entry is the g,v entry of X0.

Assume us return to the setting of Theorem with 1 not necessarily equal to 1. In notation parallel to that of we assume F is a Ck map,

F : Q —> Rl,F(zo) = uo,

where Q is a neighborhood of zo in Rn. We assume

DF(zo) : Rn —> Rl is surjective.

Then,upon reordering the variables z = (zi, ..., zn), we can write z = (x, y), x = (x1, ...,xn-l), y = (y1, ...,yl), such that DyF(z0) is invertible, and Theorem applies. Thus (for this reordering of variables), we have a Ck solution to

F (x, y) = uo,y = g(x, uo),

satisfying yo = g(xo, uo), zo = (xo, yo).

To give one example to which this result applies,we take another look at F : R4 R2. We have

DF(u,v, x,y) = (2™ 2x "2 + v*)) .

The reader is invited to determine for which (u, v, x, y) € R4 the matrix on rank 2.

Here is another example,involving a map defined on M(n,R). Set

F : M(n,R) R2,F(X) = .

Parallel to if det X = 0,Y € M(n,R),

DF (X)Y = (<det X >»<*- Y >).

Hence,given det X = 0,DF(X) : M(n,R) ^ R2 is surjective if and only if

L : M(n,R) ^ R2,LY = (^TT^)

is surjective. This is seen to be the case if and only if X is not a scalar multiple of the identity I € M(n,R).

**Check your Progress - 1**

Discuss Multivariable Differential Calculus

____________________________________________________________

____________________________________________________________

____________________________________________________________

Discuss Derivative

____________________________________________________________

____________________________________________________________

____________________________________________________________

## 13.5 LET US SUM UP

In this unit we have discussed the definition and example of Multivariable Differential Calculus, The Derivative, Inverse Function And Implicit Function Theorem

## 13.6 KEYWORDS

1. Multivariable Differential Calculus: This chapter develops differential calculus on domains in n-dimensional Euclidean space Rra.

2. The Derivative: Assume O be an open subset of Rn,and F : O ^ Rm a continuous function

3.:Inverse Function And Implicit Function Theorem: The Inverse Function Theorem gives a condition under which a function can be locally inverted.

## 13.7 QUESTIONS FOR REVIEW

Explain Multivariable Differential Calculus

Explain Derivative

## 13.8 REFERENCES


- Function of Several Variables
- Several Variables
- Function of Variables
- System of Equation
- Function of Real Variables
- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables


## 13.9 ANSWERS TO CHECK YOUR PROGRESS

Multivariable Differential Calculus

(answer for Check your Progress - 1 Q)

Derivative

(answer for Check your Progress - 1 Q)

# UNIT -14 SYSTEMS OF DIFFERENTIAL EQUATIONS AND VECTOR FIELDS

**STRUCTURE**

14.0 Objectives

14.1 Introduction

14.2 Systems Of Differential Equations And Vector Fields

14.3 Linear Systems

14.4 Sard's Theorem

14.5 Morse Functions

14.6 Differential Forms And The Gauss-Green-Stokes Theorem

14.7 Differential Forms

14.8 The General Stokes Theorem

14.9 Applications Of The Gauss-Green-Stokes Theorem

14.10 Let Us Sum Up

14.11 Keywords

14.12 Questions For Review

14.13 References

14.14 Answers To Check Your Progress

## 14.0 OBJECTIVES

After studying this unit, you should be able to:

Learn Understand about Systems Of Differential Equations And Vector Fields

Learn Understand about Linear Systems

Learn Understand about Sard's Theorem

Learn Understand about Morse Functions

Learn Understand about Differential Forms And The Gauss-Green-Stokes Theorem

Learn Understand about Differential Forms

Learn Understand about The General Stokes Theorem

Learn Understand about Applications Of The Gauss-Green-Stokes Theorem

## 14.1 INTRODUCTION

In mathematics advanced calculus whose aim is to provide a firm logical foundation of analysis of calculus and a course in linear algebra treats analysis in one variable & analysis in several variables

Systems Of Differential Equations And Vector Fields, Linear Systems, Sard's Theorem, Morse Functions, Differential Forms And The Gauss-Green-Stokes Theorem,Differential Forms,The General Stokes Theorem, Applications Of The Gauss-Green-Stokes Theorem

## 14.2 SYSTEMS OF DIFFERENTIAL EQUATIONS AND VECTOR FIELDS

In this section we study n x n systems of ODE,

= F(t, y), y(to) = yo-

To begin,we prove the following fundamental existence and uniqueness result.

Theorem. Assume y0 € Q,an open subset of Rra,I c R an interval containing t0. Suppose F is continuous on I x Q and satisfies the following Lipschitz estimate in y :

\\F(t,yi) - F(t, y2)\\ < L\\yi — y21|

for t € I,yj € Q. Then the equation has a unique solution on some t-interval containing t0.

To begin the proof,we note that the equation is equivalent to the integral equation

y(t) = yo +/ F(s, y(s)) ds.

J to

Existence will be established via the Picard iteration method,which is the following. Guess y0(t), e.g.,y0(t) = y0. Then set

yk(t) = yo + [ F(s, yk-i(s)) ds.

to

We aim to show that,as k ^ x>,yk(t) converges to a (unique) solution of at least for t close enough to t0.

We look for a fixed point of T,defined by

(Ty)(t) = y0 +/ F(s, y(s)) ds.

to

Assume

X = {u € C(J,Rra) : u(t0) = y0,sup ||u(t) - y0\ < R}.

t&J

Here J = [t0 — T,t0 + T],where T will be chosen,sufficiently small,below. The quantity R is picked so that

Bn(y0) = {y : \\y — y0\ < R}

is contained in Q,and we also suppose J c I. Then there exists M such that

sup \\F(s,y)|| < M.

se.J, \\y-yo\\<R

Then,provided

R

T < m,

we have

T : X ^ X.

Now,using the Lipschitz hypothesis,we have,for t € J,

||($y)(t) - ($z)(t)\\< f L\\y(s) - z(s)\| ds

to

< TL sup ||y(s) - z(s)\| seJ

assuming y and z belong to X. It follows that $ is a contraction on X provided one has

T< 1

L

in addition to the hypotheses

Note that the bound M and the Lipschitz hypothesis on F were needed only on BR(y0). The following setting:

For each compact K C Q,there exists MK < such that

\\f(t, x)\\ < mk,Vx € K,t € I,  and

For each K as above,there exists LK < to such that

\\F(t, x) - F(t, y)\\ < Lk\\x - y\\,Vx, y € K,t € I.

Note that,if K C Q is compact,there exists RK > 0 such that

K = {x € Mra : dist(x,K) < RK} C Q,

and K is compact. It follows that for each y0 € K,the solution to exists on the interval

{t € I : \t - t01 < min(RKMk,1/2Lk)}•

**Notes**

Now that we have local solutions to it is of interest to investigate when global solutions exist. Here is an example where breakdown occurs:

dy = y2,y(0) = 1.

The solution blows up in finite time.

It is useful to know that "blowing up" is the only way a solution can fail to exist globally. Wehave the following result.

Proposition. Assume F be as in but with the boundedness and Lipschitz hypotheses replaced by. Assume [a, b] is contained in the open interval I,and assume y(t) solves for t € (a, b). Assume there exists a compact K C Q such that y(t) € K for all t € (a, b). Then there exist a\ < a and b\ > b such that y(t) solves for t € (ai, bi).

Proof. We deduce that there exists 5 > 0 such that for each y1 <= K,11 <= [a,b],the solution to

'dl = F (t, y), y(1i) = yi

exists on the interval [11 — 5, 11 + 5]. Now,under the current hypotheses,take 11 <= (b — 5/2,b) and y1 = y(11), with y(1) continues y(1) past 1 = b. Similarly one can continue y(1) past 1 = a. □

Here is an example of a global existence result that can be deduced from Proposition Consider the 2 x 2 system for y = (x,v):

dx

= v,

f

dv 3

= —x. d1

Here we take Q = R2,F(1, y) = F(1, x, v) = (u,— x3). If holds for 1 <= (a,b), we have

d (v2 x4\dv 3 dx

—(— + —)= ^ + x —— = 0,

v ! d1\ 2 4) d1 d1

so each y(1) = (x(1),v(1)) solving lies on a level curve x4/4 + v2/2 = C,hence is confined to a compact subset of R2,yielding global existence of solutions.

The discussion above dealt with first order systems. Often one wants to deal with a higher-order ODE. There is a standard method of reducing an nth-order ODE

y(n)(1) = f (1, y, y', ..., y(n-1])

to a first-order system. One sets u = (u0, ...,un-1) with

uo = y,uj = y^, and then

du

— = (u1, ..., un-1, f (1, uo,. . ., un-1)) = g(1, u).

If y takes values in Rk,then u takes values in Rkn.

If the system is non-autonomous,i.e.,if F explicitly depends on 1,it can be converted to an autonomous system (one with no explicit 1-dependence) as follows. Set z = (1, y). We then have

dz (dy ^ = (l») = (1, F <z»= G(z).

Sometimes this process destroys important features of the original system

# 14.3 LINEAR SYSTEMS

Here we consider linear systems of the form dx

— = A(t)x,x(0) = x0,

given A(t) continuous in t € I (an interval about 0), with values in M(n,R). to establish global existence of solutions. It suffices to establish the following.

Proposition. If ||A(t)|| < K for t € I,then the solution satisfies

||x(i)|| < eK|t|||xo||.

**Notes**

Proof. It suffices to prove for $t > 0$. Then $y(t) = e-Ktx(t)$ satisfies

dy = c(t)y,y(0) = xo,C(t) = A(t) - KI.

We claim that,for $t > 0$,

$||y(t)|| < ||y(0)||$, which then implies (2.3.24), for $t > 0$. In fact,

d

dt|y(i)"2 = y'(t) ' y(t)+ y(t) ' y'(t)

= 2y(t) •(A(t) - K)y(t)■

Now

y(t) • A(t)y(t) < ||y(t)| • ||A(t)y(t)| < ||A(t)| • ||y(t)|2, so the hypothesis ||A(t)|| < K implies

d

- ||y(t)||2 < 0.

for s, t € I,the solution operator S(t, s) € M(n,R), S(t, s)x(s)= x(t).

We have

d

—S(t, s) = A(t)S (t, s), S (s, s) = I.

Note that S(t, s)S(s, r) = S(t, r). In particular,S(t, s) = S(s, t) 1.

We can use the solution operator S(t,s) to solve the inhomogeneous system

dx

— = A(t)x + f (t), x(to) = xo.

Namely,we can take

x(t) = S(t, t0)x0 + / S(t, s)f (s) ds.

Jt 0

We study how the solution to a system of differential equations dx

— = F (x), x(0) = y

depends on the initial condition y. We will assume F : Q ^ Rra is smooth,Q c Rra open and convex,and denote the solution to

by x = x(t, y). We want to examine smoothness in y. Assume DF(x) denote the n x n matrix valued function of partial derivatives of F.

To start,we assume F is of class C1,i.e.,DF is continuous on Q,and we want to show x(t,y) is differentiable in y. Assume us recall what this means. Take y € Q and pick R > 0 such that BR(y) is contained in Q. We seek an n x n matrix W(t, y) such that,for w0 € Rra,\\w0\\ < R,

x(t, y + W0) = x(t, y) + W(t, y)w0 + r(t, y,W0), where

r(t, y, W0) = o(||w0||), which means

lim r(t;y'W0> =0.

W0^0

When this holds,x(t,y) is differentiable in y,and Dy x(t, y) = W (t, y).

In other words,

x(t,y + W0) = x(t,y) + Dyx(t,y)w0 + o(\w0\).

In the course of proving this differentiability,we also want to produce an equation for W(t, y) = Dyx(t, y). This can be done as follows. Suppose x(t,y) were differentiable in y. (We do not yet know that it is,but that is okay.) Then F(x(t,y)) is differentiable in y,so we can apply Dy to (2.3.32). Using the chain rule,we get the following equation, dW

— = DF (x)W,W (0, y) = I,

called the linearization of I is the n x n identity matrix.
Equivalently,given w0 € Rra,

w(t,y) = W(t,y)wo is expected to solve

dw

Notes

— = DF (x)w,w(0) = wo.

Now,we do not yet know that x(t, y) is differentiable,

It remains to show that,with such a choice of W(t, y),

To rephrase the task,set

x(t) = x(t, y), x1(t) = x(t, y + w0), z(t) = xi(t) — x(t),

and assume w(t)

iiz(t)—w(t)y = o(HwoN)-

implies

dz

— = F (xi) — F (x), z(0) = wo.

Now the fundamental theorem of calculus gives

F (x1) — F (x) = G(x1,x)(x1 — x), with

G(x1, x) = / DF(tx1 + (1 — t)x) dr.

o

If F is C1,then G is continuous.

dz

— = G(x1,x)z,z(0) = w0.

Given that

IIDF(u)| < L,Vu € Q,

which we have by continuity of DF,after possibly shrinking Q slightly

||z(t)|| < eltlL\ that is,

\x(t, y) — x(t, y + w0)\\ < eltL\

This establishes that x(t,y) is Lipschitz in y.

To proceed,since G is continuous and G(x, x) = DF(x), dz

— = G(x + z, x)z = DF (x)z + R(x, z), z(0) = wo, dt

where

F e C:(Q) =^ \lR(x, z)\l = o(||z||) = o(||wo||). d

— (z — w) = DF (x)(z — w) + R(x, z), (z — w)(0) = 0.

Then Duhamel's formula gives

z(t) — w(t) = / S(t, s)R(x(s),z(s)) ds,

o

where S(t,s) is the solution operator for d/dt—B(t), with B(t) = G(x\(t),x(t)), satisfies

||S(t, s)||< e\t-SL.

(TT56) |z(t) — w(t)|| = o(|wo|).

This is precisely what is required to show that x(t,y) is differentiable with respect to y,with derivative W = Dyx(t, y) satisfying. Hence we have:

Proposition. If F e C:(Q) and if solutions exist for t e (—T0, T\),

then,for each such t,x(t, y) is C1 in y,with derivative Dyx(t, y)

We have shown that x(t, y) is both Lipschitz and differentiable in y. The continuity of W(t,y) in y follows easily by comparing the differential equations of the form (2.3.39) for W(t, y) and W(t, y + w0), in the spirit of the analysis of z(t).

If F possesses further smoothness,we can establish higher differentia-bility of x(t, y) in y by the following trick. get a system of differential equations for (x,W):

dx

dt = F (x)

dW

~dT = DF (x)W

with initial conditions

x(0) = y,W (0) = I.

We can reiterate the preceding argument,getting results on Dy(x,W), hence on D^x(t, y), and continue,proving:

Proposition. If F € Ck(Q), then x(t, y) is Ck in y.

Similarly,we can consider dependence of the solution to

dx

dt = f (t, x), x(0)= y

on a parameter t,assuming F smooth jointly in (t,x). This result can be deduced from the previous one by the following trick. Consider the system dx dz

— = F(z,y), - = 0,x(0) = y,z(0) = t.

Then we get smoothness of x(t, T, y) jointly in (t,y). As a special case,assume F(t, x) = tF(x). In this case x(t0, T, y) = x(Tt0,y), so we can improve the conclusion in Proposition to the following:

F € Ck(Q) =^ x € Ck jointly in (t, y).

Vector fields and flows

Assume U C Rn be open. A vector field on U is a smooth map

X : U —> Rn.

Consider the corresponding ODE

dt = X(y^ y(0) = x,

with x € U. A curve y(t) solving is called an integral curve of the vector field X. It is also called an orbit. For fixed t,write

y = y(t, x)= Fix (x).

The locally defined Flx,mapping (a subdomain of) U to U,is called the flow generated by the vector field X. As a consequence of the results on smooth

dependence of solutions to ODE,y is a smooth function of (t,x).

The vector field X defines a differential operator on scalar functions,as follows:

d

Cxf(x) = hm h-1 [f (FXx) — f (x)] = (FXx) \t=0.

We also use the common notation

Cx f (x) = Xf,

that is,we apply X to f as a first order differential operator.

Note that, if we apply the chain rule

Cxf (x) = X(x) • Xf (x) = ^ aj(x)d-L,

if X = Y^ aj(x)ej,with {ej} the standard basis of Rra. In particular, using the notation, we have

aj (x) = Xxj.

d

X = S aj(x) Qj.

We note that X is a derivation, that is,a map on C^(U), linear over R,satisfying

X (fg) = (Xf)g + f (Xg).

Conversely, any derivation on C^(U) defines a vector field

Proposition. If X is a derivation on C™(U), then X has the form Proof. Set aj(x) = Xxj,X# = ^ aj(x)d/dxj,and Y = X — X#. Then Y is a derivation satisfying Yxj = 0 for each j. We aim to show that Yf = 0 for all f. Note that whenever Y is a derivation

1 • 1 = 1 ^ Y • 1 = 2Y • 1 ^ Y • 1 = 0.

**Notes**

Thus Y annihilates constants. Thus in this case Y annihilates all polynomials of degree < 1.

Now we show that Yf (p) = 0 for all p € U. Without loss of generality, we can suppose p = 0. Then,with bj (x) = /^(dj f)(tx) dt,we can write

f (x) = f (0) + ^ bj (x)xj.

It immediately follows that Yf vanishes at 0.

A fundamental fact about vector fields is that they can be "straightened out" near points where they do not vanish. To see this,assume X be a smooth vector field on U,and suppose X(p) = 0. Then near p there is a hyperplane H that is not tangent to X near p,say on a portion we denote M. We can choose coordinates near p so that p is the origin and M is given by {xn = 0}. Thus we can identify a point X € Rn-1 near the origin with X € M. We can define a map

F : M x (-to, to) —> U by

F(X,t) = FX(X).

This is C™ and has surjective derivative at (0, 0), and so by the inverse function theorem is a local diffeomorphism. This defines a new coordinate system near p,in which the flow generated by X has the form

FX (X, t) = (X, t + s).

If we denote the new coordinates by (u1, ...,un), we see that the following result is established.

Theorem. If X is a smooth vector field on U and X(p) = 0,then there exists a coordinate system (u1, ..., un), centered at p (so Uj(p) = 0) with respect to which

d

X =

dun

We consider further mapping properties of vector fields. If F : V ^ W is a diffeomorphism between two open domains in Rn,and Y is a vector field on W,we define a vector field F#Y on V so that

F1f#y = F-1 oFly ◦ F,  or equivalently,by the chain rule,

F#Y (x) = (DF-1)(F (x))y (F (x)).

In particular,if U C Rn is open and X is a vector field on U,defining a flow Ff,then for a vector field Y,F#Y is defined on most of U,for \t\ small,and we can define the Lie derivative:

lim h-1

h^0 V #  dt

as a vector field on U.

Another natural construction is the operator-theoretic bracket: [X, Y]= XY - YX,

where the vector fields X and Y are regarded as first order differential op- erators on C^(U). One verifies that (2.3.78) defines a vector field on U. In fact,if X = ^ aj(x)d/dxj,Y = ^ bj(x)d/dxj,then

(2^79) X.Y1 = E (ak <= - b'k <=) Wt-

The basic fact about the Lie bracket is the following.

Theorem. If X and Y are smooth vector fields,then

LxY = [X,Y].

Proof. We examine LXY = (d/ds)FX#Y|s=0,using which implies that

Ys(x) = FSX#Y (x) = DF-s(Fsx (x))Y (Fsx (x)).

Assume us set Gs = DF-s. Note that Gs : U ^ End(Rra). Hence,for x € U,DGs(x) is an element of Hom(Rra,End(Rra)). Taking the s-derivative of

we have

d

-Ys(x)= - DX(Fsx(x))Y{Fxx(x))

+ DG s(Fsx (x))X (Fsx(x))Y (Fx(x))

+ DF-s(Fsx (x))DY (Fsx (x))X (Fsx (x)).

Note that G0(x) = I € End(Rra) for all x € U,so DG° = 0. Thus

d|

Lx Y = — Ys(x) |s=0 = -DX (x)Y (x) + DY (x)X (x), for [X, Y]. □

Corollary. If X and Y are smooth vector fields on U,then

d

-Fx#Y = Fx#[X,Y] for all t.

Proof. Since locally Ft+s = FxFy,we have the same identity for Ft+#. Hence

d d |

dt F'x #Y = Y #fx #Y l.=0 = Ft #LX Y

## 14.4 SARD'S THEOREM

Assume $F$ : $O$ ^ R^ra be a $C^1$ map,with $O$ open in R^ra. If $p$ € $O$ and $DF(p)$ : R^ra ^ R^ra is not surjective,then p is said to be a critical point,and F(p) a critical value. The set C of critical points can be a large subset of O, even all of it,but the set of critical values F(C) must be small in R^ra,as the following result implies.

***Proposition.*** If F : O ^ $R^n$ is a $C^1$ map,C C O its set of critical points,and K C O compact,then F(C n K) is a nil subset of $R^n$.

**Proof.** Without loss of generality,we can assume $K$ is a cubical cell. Assume $P$ be a partition of $K$ into cubical cells $R_a$,all of diameter 5. Write $P = P' UP''$, where cells in $P'$ are disjoint from C,and cells in $P''$ intersect C. Pick $x_a$ <= $R_a$ n $C$,for $R_a$ <= $P''$.

Fix e > 0. Now we have

F (Xa + y) = F (Xa) + DF (Xa) + ra(y),

and,if $5 > 0$ is small enough,then $|r_a(y)| < e\backslash y\backslash < e5$,for $x_a + y <= R_a$. Thus $F(R_a)$ is contained in an e5-neighborhood of the set $H_a = F(x_a) + DF(x_a)(R_a — x_a)$, which is a parallelipiped of dimension < n — 1,and diameter < M5,if \DF\ < M. Hence

$cont+ F(Ra) < Ce5^n < C'eV(Ra)$, *for* Ra <= P".

Thus

cont+ $F(C$ n $K) <$ ^ cont+ $F(R_a) < C''e$.

$R_a$eP "

Taking e 0,

This is the easy case of a result known as Sard's Theorem,which also treats the case F : *O* ^ $R^n$ when *O* is an open set in $R^m$,m > n. Then a more elaborate argument is needed,and one requires more differentiability, namely that *F* is class $C^k$,with $k = m — n + 1$.

# 14.5 MORSE FUNCTIONS

If Q C Rn is open,a C2 function f : Q ^ R is said to be a Morse function if each critical point of f is nondegenerate,i.e.,

Vp <= Q,Vf (p) = 0 ^ D2 f (p) is invertible,

where D2f (p) is the symmetric n x n matrix of second order partial deriva- tives defined in More generally,if M is an n-dimensional surface,a C2 function f : M ^ R is said to be a Morse function if f o p is a Morse function on Q for each coordinate patch p : Q U C M.

Our goal here is to establish the existence of lots of Morse functions on an n-dimensional surface M. For simplicity,we restrict attention to the case where M is compact. Here is our main result.

Theorem. Assume M c RN be a compact,smooth,n-dimensional surface. For a e RN,set pa : M —> R,pa(x) = a ■ x,x e M.

Take f e C2(M). Then the set Of of a e RN such that

f + pa : M —> R is a Morse function is a dense open subset of RN.

It is easy to verify that Of is open,since when holds,a small C2 perturbation g of f has the property that D2g(x) is invertible for x near p. What is not so easy is to show that Of is dense (or even nonempty!). Our proof of such denseness will make use of Sard's theorem,from. We begin with an easy special case.

Proposition. Assume N = n + 1 and M = dQ,with Q c Rn+1 open. Then

{a e Sn : a e O0} is a nil set hence has empty interior in the unit sphere Sn.

Proof. Here we are examining when pa is a Morse function on M. Assume N : M Sn be the exterior unit normal. Then xo e M is a critical point of pa if and only if N(x0) = ±a. Such a point x0 is a nondegenerate critical point of pa if and only if it is not a critical point of N. Hence,if ±a e Sn are regular values of N,then pa is a Morse function,i.e.,a e O0. By Sard's theorem,the set of points in Sn that are critical values of N is a nil set.

Lemma. Assume Q c Rn be open,and take g e C2(Q). Assume U c Q be the closure of a smoothly bounded open U. Set ga(x) = g(x) + a ■ x. Assume Og denote the set of a e Rn such that ga\u has only nondegenerate critical points. Then Rn \Og is a nil set.

Proof. Consider

F(x) = -Vg(x), F : Q Rn.

A point x e Q is a critical point of ga if and only if F(x) = a,and this critical point is degenerate only if,in addition,a is a critical value of F. Hence the desired conclusion holds for all a Rn that are not critical values of Flu. Again Sard's theorem applies.

Proof. Each p € M has a neighborhood Up in M such that Up C Hp C M and some n of the coordinates Xj on RN produce coordinates on Qp. Say x1, ..., xn do it. Assume (an+1, ...,aN) be fixed,but arbitrary. Then can be applied to g = f + J2N+1 ajXj,treated as a function of (x1, . . ., xn). It follows that,for all (a1, ...,an) but a nil set,f + ta has only nondegenerate critical points in Up. Thus

{a € RN : f + <pa has only nondegenerate critical points in Up}

is dense in RN. We also know this set is open. Now M can be covered by a finite collection of such sets Up,so Of,defined in Theorem is a finite intersection of open dense subsets of RN,hence it is open and dense, as asserted.

## 14.6 DIFFERENTIAL FORMS AND THE GAUSS-GREEN-STOKES THEOREM

The calculus of differential forms,one of E. Cartan's fundamental contributions to analysis,provides a superb set of tools for calculus on surfaces and other manifolds. A 1-form a on an open set Q C Rn can be written a = a\(x) dx\ + • • • + an(x) dxn. One can integrate such a 1-form over a smooth curve 7 : I Q,via

Y I

where 7*a is the pull-back of a,given by ^j aj (7(t))yj (t) dt. More generally, a fc-form is a finite sum of terms

aj (x) dxj! /•••/ dxjk,j = (ji, ..., jk),

where the "wedge product" satisfies the anticommutativity relation dxi / dxm = -dxm / dx1.

If p : O Q is a smooth map,and a is a fc-form on Q,one has the pull-back <p*a to a fc-form on O,satisfying

P*(a / (3) = p*a / p*@,(<p ∘ ^)*a = ^*(p*a),

if also ^ : U ^ O is a smooth map.

Another fundamental ingredient is the exterior derivative, d :Ak(Q) —^ Afc+1(Q),

Notes

where Ak (Q) denotes the space of smooth fc-forms on Q. One has the crucial identities

dda = 0,d(p*a) = p*(da).

The action of p* on n-forms (for Q,O open in Rn) is given by p*(F(x) dx1 A - ■ ■ A dxn) = F(p(x))(det Dp(x)) dx1 A - ■ ■ A dxn. J a = J p*a,

n o

provided p : O Q is a diffeomorphism such that det Dp(x) > 0 on O. (One says p preserves orientation.) Given this,one can define

J f

M

whenever f is a k form and M c Rn is a fc-dimensional surface,assuming M possesses an "orientation."

Complementing the important identitiesone has the following,which could be called the "fundamental theorem of the calculus of differential forms, "

da = a,

M dM

when M is a fc-dimensional oriented sirface (or manifold) with smooth boundary dM,and a is a smooth (fc — 1)-form on M. This identity general- izes classical identities of Gauss,Green,and Stokes,and is called the general Stokes formula.

The calculus of differential forms,particularly of the Gauss-Green-Stokes formula.

## 14.7 DIFFERENTIAL FORMS

It is very desirable to be able to make constructions that depend as little as possible on a particular choice of coordinate system. The calculus of differential forms,whose study we now take up,is one convenient set of tools for this purpose.

We start with the notion of a 1-form. It is an object that gets integrated over a curve; formally,a 1-form on Q C Mra is written

a = ^ aj(x) dxj.

j

If y : [a, b] ^ Q is a smooth curve,we set

Ja = J ^ aj (y (t))j'j(t) dt.

i

In other words,

J a = J y*a

y I

where I = [a,b] and

Y*a = ^ aj(Y(t))Yj(t) dt j

is the pull-back of a under the map y. More generally,if F : O ^ Q is a smooth map (O C open), the pull-back F*a is a 1-form on O defined by

d F

F*a = <= aj (F(y)) dyk.

j, k

The usual change of variable formula for integrals gives

J a = j F*a

Y a

if y is the curve F ◦ a.

If F : O Q is a diffeomorphism ,and

d

## Notes

(4L6) X = s b (x) dXj

is a vector field on Q, recall from that we have the vector field on O:

F#X (y) = (DF-1(p))X (p), p = F (y).

If we define a pairing between 1-forms and vector fields on Q by

{X, a) = ^ bj(x)aj(x) = b ■ a,

j

a simple calculation gives

{F#X,F*a) = {X,a) o F.

Thus,a 1-form on Q is characterized at each point p € Q as a linear transformation of the space of vectors at p to R.

More generally,we can regard a fc-form a on Q as a fc-multilinear map on vector fields:

a(Xi, ..., Xk) € C~(Q);

we impose the further condition of anti-symmetry when fc > 2:

a(Xi,..., Xj,..., Xe, ..., Xk) = -a(Xi,..., Xe, ..., Xj, .. ., Xk).

Assume us note that a 0-form is simply a function.

There is a special notation we use for fc-forms. If 1 < ji < ■■■ < jk < n,j = (ji, ..., jk), we set

(LL12) a = aj(x) dxji *■■■* dxjk

j

where

aj(x) = a(Dj1, ...,Djk), Dj = d/dxj.

More generally,we assign meaning to summed over all fc-indices (ji, ...,jk), where we identify

(LL14) dxji A ■ ■ ■ A dxjk = (sgn a) dxjam A ■ ■ ■ A dxja(k), a being a permutation of {1, ...,fc}. If any jm = ji (m = I), then vanishes.

A common notation for the statement that a is a fc-form on Q is

a € Ak(Q).

In particular,we can write a 2-form fi as

f = 1 ^ bjk(x) dxj A dxk

and pick coefficients satisfying bjk(x) = -bkj(x).

If we set U = Y,uj(x)d/dxj and V = J2 vj(x)d/dxj,then

f(U, V) = ^ bjk (x)uj (x)vk (x).

If bjk is not required to be antisymmetric,one gets f (U,V) = (1/2) J2(bjk — bkj)uj vk.

If F : O ^ Q is a smooth map as above,we define the pull-back F*a of a fc-form a,given by to be

F*a = <= a (F(y))(F*dxn) A^A (F*dxJk)

j

where

dF-

F*dxj = Y] -F dye,

v °yi

the algebraic computation in being performed using the rule

If F is a diffeomorphism,we have

(F *a)(F#Xl, ..., F#Xk) = a(Xl, ..., Xk) o F. If B = (bjk) is an n x n matrix,then

bik dxk) A b"2k dx2j A-^ A ^2 bnk dx2j

k k k

= biki • • • bnkn dx^ A • • • A dxkn

D, ..., kn

= ((sgn &)bla(1)b2a(2) • • • bna(n)) dxi A • • • A dxn

= (det B) dx1 A • • • A dxn.

Here Sn denotes the set of permutations of {1, ...,n},and the last identity is the formula for the determinant presented. It follows that if F : O ^ Q is a C1 map between two domains of dimension n and

a = A(x) dx1 A^ • • A dxn is an n-form on Q,then

F*a = det DF(y) A(F(y)) dy1 A • • • A dyn.

Comparison with the change of variable formula for multiple integrals suggests that one has an intrinsic definition of fQ a when a is an n-form on Q,n = dim Q. To implement this,we need to take into account that det DF(y) rather than|det DF(y)|. We say a smooth map F : O ^ Q between two open subsets of Rn preserves orientation if det DF(y) is everywhere positive. The object called an "orientation" on Q can be identified as an equivalence class of nowhere vanishing n-forms on Q,two such forms being equivalent if one is a multiple of another by a positive function in C^ (Q); the standard orientation on Rn is determined by dx1 A • • •A dxn. If S is an n-dimensional surface in Rn+k,an orientation on S can also be specified by a nowhere vanishing form u € An(S).

If such a form exists,S is said to be orientable.

The equivalence class of positive multiples a(x)w is said to consist of "positive" forms. A smooth map ^ : S ^ M between oriented n-dimensional surfaces preserves orientation provided ^*a is positive on S whenever a € An(M) is positive. If S is oriented,one can choose coordinate charts which are all orientation preserving. We mention that there exist surfaces that cannot be oriented,such as the famous "Mobius strip, " and also the projective space P2.

We define the integral of an n-form over an oriented n-dimensional surface as follows. First,if a is an n-form supported on an open set Q C Rn, then we set

J a = j A(x) dV(x),

n n

the right side defined If O is also open in Rn and F : O Q is an orientation preserving diffeomorphism,we have

J F*a = j a,

o n

as a consequence and the change of variable formula More generally,if S is an n-dimensional surface with an orientation,say the image of an open set O C Rn by p : O S,carrying the natural orientation of O,we can set

J a = J p*a

S O

for an n-form a on S. If it takes several coordinate patches to cover S,define fs a by writing a as a sum of forms,each supported on one patch.

We need to show that this definition of S a is independent of the choice of coordinate system on S (as long as the orientation of S is respected). Thus,suppose p : O ^ U C S and ^ : Q ^ U C S are both coordinate patches,so that F = ^-1 op : O Q is an orientation-preserving diffeomorphism.

We need to check that,if a is an n-form on S,supported on U,then

J p*a = J ^*a.

o n

To establish this,we first show that,for any form a of any degree,

^ o F = p =^ p*a = F*^*a.

dxj = J2(d^j/dxi) dxg,so

F*dxj = W d^j dxm,p* dxj = W dxm;

dxm dxi m dxm

i, m m

but the identity of these forms follows from the chain rule:

Notes

(,oo) ^ = W)(DF)^ M

Now

J F*(^*a),

O

which is equal to the right side of. Thus the integral of an n-form over an oriented n-dimensional surface is well defined.

If F : U0 ^ U\ and G : U\ U2 are smooth maps and a € Ak(U2), then implies

(G o F)*a = F*(G*a) in Ak(Uo).

In the special case that Uj = Rra and F and G are linear maps,and k = n, show that this identity implies

det(GF) = (det F)(det G).

Compare this with the derivation .

Assume AkRra denote the space of k-forms with constant coefficients. Show that

dimR Ak Rra = Q.

If T : Rm ^ Rra is linear,then T* preserves this class of spaces; we denote the map

AkT* : AkRra —> AkRm.

Similarly,replacing T by T* yields

AkT : AkRm —> AkRra

Show that AkT is uniquely characterized as a linear map from AkRm to AkRn which satisfies

(AkT)(vi A • • • A Vk) = (Tvi) A • • • A (Tvk), Vj € Rm

Show that if S, T : Rn ^ Rn are linear maps,then (Ak(ST) = (AkS) o (AkT).

Relate this to

If {e1, ...,en} is the standard orthonormal basis of Rn,

define an inner product on AkRn by declaring an orthonormal basis to be

(4.R38) {ej! A^^A j : 1 < ji < ••• <jk < n}.

If A : AkRn ^ AkRn is a linear map,define At : AkRn ^ AkRn by

(Aa,0) = (a,At0), a, P € AkRn, where (,) is the inner product on AkRn defined above.

Show that,if T : Rn ^ Rn is linear,with transpose Tt,then

(Ak T)t = Ak (T t).

Hint. Check the identity ((AkT)a, 0) = (a,(AkTt)0) when a and 0 run over the orthonormal basis that is,show that if a = ej1 A^ • • A ejk, 0 = eil A ••• A eik,then (TT41) (Tejl A^^A Tejk,eh /\-A eik) = (ej1 A^A e-k,Tteii A^A Tteik)■ Hint. Say T = (tij). In the spirit expand Te-1 A^ • • A Te-k,and equal to

E (sgn a)tia(1)ji • • • tia(k)jk,

(7^Sk

where Sk denotes the set of permutations of {1, ...,k}.

E (sgn T)tii jT(i) • • • tikjT(k).

T €Sk

Show that if {u1, ...,un} is any orthonormal basis of Rn,then the set {uj1 A • • • A Ujk : 1 < j1 < • • • < jk < n} is an orthonormal basis of AkRn. show that if T : Rn Rn is an orthogonal

transformation on Rn (i.e.,preserves the inner product) then AkT is an orthogonal transformation on AkRn.

Assume Vj, Wj € Rn,1 < j < k (k < n). Form the matrices V,

whose k columns are the column vectors vi, ...,vk,and W,

whose k columns are the column vectors wi, ...,wk. Show that (vi A • • • A vk, w1 A • • • A wk) = det Wf'V

= det VtW.

If vj,wj € Rn,then

(vi A • • • A vk, wi A • • • A wk) = ^(sgnn)(vi, wn(i)) • • • (vk, wn(k)),

n

where n ranges over the set of permutations of {1, ...,k}.

Assume A, B : Rk ^ Rn be linear maps and set w = ei A ••• A ek € AkRk. We have AkAw,AkBw € AkRn.

Deduce from that

(AkAw,AkBw) = det BlA.

Assume p : O ^ Rn be smooth,with O c Rm open. Deduce from Exercise 10 that,for each x € O,

||AmDp(x)w||2 = det Dp(x)tDp(x),

where w = ei A^ • •A em.

Deduce that if p : O ^ U c M is a coordinate patch on a smooth m-dimensional surface M c Rn and f € C(M) is supported on U,then

J fdS = J f (p(x))||AmDp(x)w|| dx.

M O

# 14.8 THE GENERAL STOKES THEOREM

The Stokes formula involves integrating a fc-form over a fc-dimensional sur- face with boundary. We first define that concept. Assume S be a smooth fc- dimensional surface (say in RN), and assume M be an open subset of S,such that its closure M (in RN) is contained in S. Its boundary is dM = M\M. We say M is a smooth surface with boundary if also dM is a smooth (fc — 1)- dimensional surface. In such a case,any p € dM has a neighborhood U C S with a coordinate chart p : O U,where O is an open neighborhood of 0 in Rk,such that p(0) = p and p maps {x € O : x\ = 0} onto U n dM.

If S is oriented,then M is oriented,and dM inherits an orientation, uniquely determined by the following requirement: if

M = R— = {x € Rk : xi < 0},

then dM = {(x2, ...,xk)} has the orientation determined by dx2 A • • - A dxk. We can now state the Stokes formula.

Proposition. Given a compactly supported (fc — 1)-form 4 of class C1 on an oriented fc-dimensional surface M (of class C2) with boundary dM, with its natural orientation,

J dp = J p.

M dM

Proof. Using a partition of unity and invariance of the integral and the exterior derivative under coordinate transformations,it suffices to prove this when M has the form In that case,we will be able to deduce from the

Fundamental Theorem of Calculus. Indeed,if

4 = bj(x) dx1 A • • • A dxj A • • • A dxk, with bj(x) of bounded support,we have

db

dp = (—1)j_1 —b^ dx1 A • • • A dxk.

dxj

If j > 1,we have

dbL

—^ ^ '^j

and also k* p = 0,where k : dM M is the inclusion. On the other hand, for j = 1,we have

J' dp = J dX dx2 ■ ■ ■ dxk

M

=/61(0, x/) dx'

= / p

dM

This proves Stokes' formula

## 14.9 APPLICATIONS OF THE GAUSS-GREEN-STOKES THEOREM

The first set of applications,given in deals with complex function theory. If Q c C is an open set,a C1 function f : Q ^ C is said to be holomorphic if it is complex differentiable,of equivalently if it satisfies a set of equations called the Cauchy-Riemann equations. We deduce from Green's theorem that if Q is a smoothly bounded domain and f € C 1(Q) is holomorphic on Q,then we have the Cauchy integral theorem,

dn

and the Cauchy integral formula,

These key results lead to further results on holomorphic functions,such as power series developments.

In we also consider functions on domains Q c Rra that are harmonic, and use Gauss-Green formulas to establish results about such functions, such as mean value properties,and Liouville's theorem,which states that a bounded harmonic on all of Rra must be constant. These results specialize to holomorphic functions on C. One consequence is the fundamental theorem of algebra,which states that if p(z) is a nonconstant polynomial on C,it must have a complex root.

The second set of applications,given in yields important results on the topological behavior of smooth maps on regions in Rra,and on surfaces and more generally on manifolds. A central notion here is that of degree theory. If X is a smooth,compact,oriented,n-dimensional surface,and F : X Y is a smooth map to a compact,connected,oriented, n-dimensional surface Y,then the degree of F is given by where w is an n-form on Y such that fY w = 1. That this is well-defined, independent of the choice of such w,is a consequence of the fundamental exactness criterion,given in Proposition that says a smooth n-form a on Y is exact,i.e.,has the form a = d@,if and only if fY a = 0. With this, we are able to develop degree theory as a powerful tool. Applications range from the Brouwer fixed-point theorem and the Jordan-Brouwer separation theorem (in the smooth case) to a degree-theory proof of the fundamental theorem of algebra.

We also consider on a compact surface M a vector field X with non-degenerate critical points,define the index of such a vector field,and show that Index X = x(M) is independent of the choice of such a vector field. This defines an invariant x(M), called the Euler characteristic. Investigations of x(M)

**Check your Progress - 1**

Discuss Systems Of Differential Equations And Vector Fields

__________________________________________________________

__________________________________________________________

__________________________________________________________

Discuss Differential Forms And The Gauss-Green-Stokes Theorem

_________________________________________________________

__________________________________________________________

__________________________________________________________

## 14.10 LET US SUM UP

In this unit we have discussed the definition and example of Systems Of Differential Equations And Vector Fields, Linear Systems, Sard's Theorem, Morse Functions, Differential Forms And The Gauss-Green-Stokes Theorem, Differential Forms, The General Stokes Theorem, Applications Of The Gauss-Green-Stokes Theorem, Systems Of Differential Equations And Vector Fields

## 14.11 KEYWORDS

1. Systems Of Differential Equations And Vector Fields In this section we study n x n systems of ODE F(t, y), y(to) = yo-

2. Linear Systems: Here we consider linear systems of the form dx — = A(t)x,x(0) = x0

3. Sard's Theorem: Assume $F : O \wedge R^{ra}$ be a $C^1$ map, with $O$ open in $R^{ra}$. If $p \in O$ and $DF(p)$

4. Morse Functions: If Q C Rn is open, a C2 function f : Q ^ R is said to be a Morse function if each critical point of f is non degenerate.

5. Differential Forms And The Gauss-Green-Stokes Theorem: The calculus of differential forms, one of E.

6. Applications Of The Gauss-Green-Stokes Theorem: The Stokes formula involves integrating a fc-form over a fc-dimensional sur- face with boundary.

7.Systems Of Differential Equations And Vector Fields The first set of applications, given in deals with complex function theory.

## 14.12 QUESTIONS FOR REVIEW

Explain Systems Of Differential Equations And Vector Fields

Explain Differential Forms And The Gauss-Green-Stokes Theorem

## 14.13 REFERENCES


- Analysis of Several Variables
- Application of Several Variables
- Function of Several Variables
- System of Equations

- Function of Real Variables
- Real Several Variables
- Elementary Variables
- Calculus of Several Variables
- Advance Calculus of Several Variables

## 14.14 ANSWERS TO CHECK YOUR PROGRESS

Systems Of Differential Equations And Vector Fields

(answer for Check your Progress - 1 Q)

Differential Forms And The Gauss-Green-Stokes Theorem

(answer for Check your Progress - 1 Q)