ADVANCED CALCULUS LECTURE NOTES FOR MATHEMATICS 217-317

PART I

THIRD EDITION

JAMES S. MULDOWNEY

DEPARTMENT OF MATHEMATICS
The University of Alberta
Edmonton, Alberta, Canada

PREFACE

These notes are not intended as a textbook. It is hoped however that they will minimize the amount of notetaking activity which occupies so much of a student's classtime in most courses in mathematics. Since the material is presented in the sterile "definition, theorem, proof" form without much background colour or discussion most students will find it profitable to use the notes in conjunction with a textbook recommended by the instructor.

Probably the most important aspect of the notes is the set of exercises. You should develop the practice of attempting several of these problems every week. Many of the problems are quite difficult so please consult your instructor if you are not blessed with success initially. Do not acquire the habit of abandoning a problem if it does not yield to your first attempt; a defeatist attitude is your greatest adversary. Solution of a problem, even with some assistance from the teacher when necessary, is a fine boost for your morale. You will find that a strong effort expended on the earlier part of the course will be rewarded by growing selfconfidence and easier success later.

Thanks are due to Olwyn Buckland who typed the notes. If you find that some of the solutions given to the exercises are incorrect please complain to her about it.

<u>Notation</u>: Except when specified otherwise, upper case (capital) letters will denote sets and lower case (small) letters will denote elements of sets.

$a \in A$	means a is an element of the set A .
$A \subset B$	means A is a subset of B .
$B \supset A$	means B contains A .
$A \subsetneq B$	means A is a proper subset of B .
$P \Rightarrow Q$	means statement P implies statement Q .
$P \Leftrightarrow Q$	means P holds if and only if Q holds.
$\exists$	"there exists"
$\forall$	"for each"
$\ni$	"such that"
$\square$	"end of proof"

<u>Remark</u>: A slash through any symbol means the negation of the corresponding statement e.g. $a \notin A$ means a is not an element of A . $\{x : \ldots\}$ means the set of all things x satisfying the condition following the colon.

$A \cup B = \{x : x \in A \text{ or } x \in B\}$ (union)

$A \cap B = \{x : x \in A \text{ and } x \in B\}$ (intersection)

$A - B = \{x : x \in A \text{ and } x \notin B\}$ (difference)

$A \times B = \{(x,y) : x \in A , y \in B\}$ (Cartesian product)

ϕ : empty set

CHAPTER ONE

THE REAL NUMBER SYSTEM & FINITE DIMENSIONAL CARTESIAN SPACE

We begin by defining our basic tool, the real numbers R. The real numbers can be constructed from more primitive notions such as the natural numbers $N = \{1,2,3,\dots\}$ or even from the fundamental axioms of set theory; this topic is studied in Mathematics 417. Here we shall be content with a precise description of R.

THE REAL NUMBER SYSTEM R.

Definition: R is a complete ordered field. We explain the three words underlined.

Field.

Definition: A field is a set F together with two binary operations $+$ and $\cdot$ (addition, multiplication) which satisfy the following axioms: For all $a,b,c,\dots$ in F

(F1) $a+b \in F$ and $a\cdot b \in F$ (closure)

(F2) $a+b = b+a$ and $a\cdot b = b\cdot a$ (commutativity)

(F3) $a+(b+c) = (a+b)+c$ and $a\cdot(b\cdot c) = (a\cdot b)\cdot c$ (associativity)

(F4) $(a+b)\cdot c = (a\cdot c) + (b\cdot c)$ (distributivity)

(F5) There exist unique elements 0 and $1 \in F$ $0 \neq 1$ such that

$$a+0 = a \text{ and } a\cdot 1 = a \ \forall a \in F \quad \text{(identities)}$$

(F6) For each $a \in F$, $\exists (-a) \in F$ such that $a+(-a) = 0$ and if $a \neq 0$ $\exists a^{-1} \in F$ such that $a \cdot a^{-1} = 1$ (inverses).

Remarks:

(i) $a \cdot b$ will henceforth be written ab.

(ii) $a(b+c) = ab + ac$ is a consequence of F2 and F4 and need not be assumed (verify).

(iii) The elements $(-a)$ and a^{-1} are uniquely determined by a. Suppose, for example, that there are elements $(-a_1),(-a_2)$ such that $a+(-a_1) = 0$. Then

$$(-a_1) = (-a_1)+0 = (-a_1) + (a+(-a_2))$$
$$= ((-a_1)+a) + (-a_2) = 0 + (-a_2) = (-a_2) \quad .$$

(Justify each equality.)

(iv) It is customary to write $a-b$ for $a+(-b)$ and $\frac{a}{b}$ for ab^{-1}.

(v) aa, aaa, etc., are usually denoted as a^2, a^3, etc.

(vi) $\{1, 1+1, 1+1+1 \text{ etc.}\}$ is usually denoted $\{1,2,3, \text{etc.}\} = N$

Exercises.

1.1: Establish the following properties of a field.

(a) $a \cdot 0 = 0$

(b) $a(-1) = (-a)$

(c) $(-a)(-b) = ab$

(d) $(ab^{-1})(cd^{-1}) = ac(bd)^{-1}$

(e) If $ab = 0$ then $a = 0$ or $b = 0$.

Examples:

(i) The simplest (and least interesting) field is the set $\{\theta, e\}$ with binary operations

+	θ	e
θ	θ	e
e	e	θ

$\cdot$	θ	e
θ	θ	θ
e	θ	e

(ii) The set Q of rational numbers, i.e., numbers of the form $\frac{m}{n}$ $(n \neq 0)$ with the usual addition and multiplication is a field.

(iii) The sets R and C of real and complex numbers respectively with the usual addition and multiplication are fields.

(iv) The set $Q(t)$ of all rational functions with rational coefficients (i.e. functions of the form $\frac{p(t)}{q(t)}$ where $p(t)$ and $q(t)$ are polynomials with rational coefficients) is a field.

(v) The set $N = \{1,2,3,\ldots\}$ of natural numbers and the set Z of integers are not fields.

Ordered Fields.

Definition: A field F is ordered if there is a subset P of F (called the positive elements) such that:

(O1) $a,b \in P \Rightarrow a+b \in P$ and $ab \in P$

(O2) $0 \notin P$

(O3) $x \in F$, $x \neq 0 \Rightarrow x \in P$ or $-x \in P$ but not both.

<u>Exercise</u>.

1.2: Observe that

(a) $1 \in P$

(b) $a \neq 0 \Rightarrow a^2 \in P$

(c) If $n \in N$ then $n \in P$

(d) The field $\{\theta,e\}$ in Example (i) above cannot be ordered.

(e) The field C of complex numbers cannot be ordered.

(f) The fields Q and R are ordered by the usual notion of positivity.

(g) The field $Q(t)$ in Example (v) is ordered if $\frac{p(t)}{q(t)} \in P$ whenever the coefficient of the highest power of t in the product $p(t)\,q(t)$ is positive.

<u>Remark</u>: Every ordered field contains Q as a subfield (we do not prove this). Thus Q may be characterised as an ordered field containing no ordered proper subfield, i.e. Q is the <u>smallest</u> ordered field. (Two ordered fields are considered the same if they are isomorphic and the isomorphism preserves the order.) A discussion of this point may be found in

C. Goffman: Real Functions, Proposition 1 and 2, Chapter 3, or

E. Hewitt and K. Stromberg: Real and Abstract Analysis, Theorem 5.9.

We define a relation $>$ on an ordered field.

<u>Definition</u>. If $a,b \in F$, write $a > b$ (equivalently $b < a$) if $a-b \in P$.

Properties 01, 02, 03 yield the following:

(i) $a > b$, $b > c \Rightarrow a > c$

(ii) If $a,b \in F$ then exactly one of the following holds

$$a > b\ ,\ a = b\ ,\ a < b\ .$$

(iii) $a \geq b$, $b \geq a \Rightarrow a = b$

(iv) $a > b \Rightarrow a+c > b+c$, $\forall\ c \in F$.

(v) $a > b$, $c > d \Rightarrow a+c > b+d$

(vi) $a > b$, $c > 0 \Rightarrow ac > bc$

$a > b$, $c < 0 \Rightarrow ac < bc$

(vii) $a > 0 \Rightarrow a^{-1} > 0$

$a < 0 \Rightarrow a^{-1} < 0$

(viii) $a > b \Rightarrow a > \frac{a+b}{2} > b$

(ix) $ab > 0 \Rightarrow$ either $a > 0$ and $b > 0$

or $a < 0$ and $b < 0$.

<u>Exercise</u>.

1.3: Prove statements (i) - (ix).

<u>Complete Ordered Field</u>.

Let S be a subset of an ordered field F .

(a) $u \in F$ is an <u>upper bound</u> of S if $s \leq u$, $\forall s \in S$

(b) $w \in F$ is a <u>lower bound</u> of S if $w \leq s$, $\forall s \in S$

(c) S is bounded above (below) if it has an upper (lower) bound. S is <u>bounded</u> if it is bounded above and below.

e.g. $N = \{1,2,3,\ldots\}$ is bounded below and unbounded above;

$[0,1) = \{x : 0 \leq x < 1\}$ is bounded.

(d) u is the <u>least upper bound</u> (or <u>supremum</u>) of S if

(i) $s \leq u$, $\forall s \in S$ (i.e. u is an upper bound)

(ii) $s \leq v$, $\forall s \in S \Rightarrow u \leq v$ (i.e. u is smaller than any other upper bound).

Write $u = \sup S$ or $u = \text{lub } S$.

(e) Similarly the <u>greatest lower bound</u> (or <u>infimum</u>) of S is a number w which is a lower bound of S and exceeds all other lower bounds.

Write $w = \inf S$ or $w = \text{g}\ell\text{b } S$.

<u>Definition</u>: An ordered field F is <u>complete</u> if each nonempty subset S of F which has an upper bound has a least upper bound (supremum) in F.

<u>Remark</u>: R is the only complete ordered field (to within an isomorphism). Again we do not prove this. A discussion may be found in Hewitt and Stromberg p. 45.

Exercise.

1.4: Guess the supremum and infimum of each of the following sets (when they exist).

$(0,1) = \{x : 0 < x < 1\}$, $[0,1] = \{x : 0 \le x \le 1\}$

$\{\frac{1}{n} : n = 1,2,3,\ldots\}$, $\{1,2,3,\ldots\}$

Properties of R :

Definition: An ordered field F is Archimedean if $\forall\, a \in F$, $\exists\, n \in N = \{1,2,\ldots\}$ $\ni n > a$ (i.e. N is not bounded above).

Theorem 1.1: R is Archimedean.

Proof: Suppose not. Then $\exists\, a \ge n$, $\forall\, n \in N$, i.e. a is an upper bound of N .

$\therefore$ $b = \text{Sup}\, N$ $\exists$ since R is complete.

$\therefore$ $b \ge n$ $\forall\, n \in N$ and $b-1 < n_o$ for some $n_o \in N$ ($b-1 < b$ is not an upper bound since b is the least upper bound). But then $b < n_o+1 \in N$ contradicting $b = \sup N$. □

Corollary 1.1.1: If $a > 0$ $\exists\, n \in N \ni 0 < \frac{1}{n} < a$.

Proof: $\exists\, n > a^{-1} > 0$ (Why?)

$\therefore\, a > \frac{1}{n} > 0$ (Why?) □

Exercise.

1.5: If $a > 0 \ \exists \ n \in N \ni 0 < \frac{1}{2^n} < a$.

(Hint: Show $2^n > n$, $\forall n \in N$.)

Corollary 1.1.2: Q is an Archimedean ordered field.

Exercise.

1.6: $Q(t)$ is not Archimedean.

Corollary 1.1.3: If $a,b \in R$, $a < b$ then there is a rational r such that $a < r < b$. (This Corollary holds for any Archimedean field.)

Proof: $\exists \ n \in N \ni n(b-a) > 1$ (Why?). Let m be the least integer such that $m > na$. Hence $m-1 \leq na$

$$\Rightarrow na < m \leq na+1 < na + n(b-a) = nb$$

$$\Rightarrow a < \frac{m}{n} < b . \quad \square$$

Exercises.

1.7: Let F be an Archimedean ordered field containing an irrational element ξ . Show that if $a,b \in F$, $a < b$ then there is an irrational element η such that $a < \eta < b$.

1.8: Show that R contains an irrational element.

(Hint: Show first that no rational p satisfies $p^2 = 2$. Then show that $p = \sup\{x > 0 : x^2 < 2\}$ must satisfy $p^2 = 2$.

<u>Notation</u>: If $a,b \in R$, $a < b$

$$[a,b] = \{x \in R : a \leq x \leq b\} \quad \text{closed interval}$$

$$(a,b) = \{x \in R : a < x < b\} \quad \text{open interval}$$

$$[a,b) = \{x \in R : a \leq x < b\}$$

$$(a,b] = \{x \in R : a < x \leq b\}$$

<u>Theorem 1.2</u>: Let $I_n = [a_n,b_n]$, and $I_{n+1} \subset I_n$, $n = 1,2,\ldots$, then $\bigcap_{n=1}^{\infty} I_n \neq \phi$ (i.e. a <u>nested</u> sequence of <u>closed</u> intervals has at least one point common to all the intervals).

<u>Proof</u>: $a_n < b_m$ for all n,m.

$\therefore$ each b_m is an upper bound for $\{a_n : n \in N\}$

$\therefore$ $a = \sup \{a_n : n \in N\} \leq b_m$, $\forall\, m$

$\therefore$ $a_n \leq a \leq b_n$, $\forall\, n$; i.e. $a \in I_n$, $\forall\, n$. $\square$

<u>Exercises</u>:

1.9: Let F be an Archimedean ordered field, with the property that if $\{I_n\}$ is a nested sequence of closed intervals in F then $\bigcap_{n=1}^{\infty} I_n \neq \phi$. Show that F is complete.

Theorem 1.2 and Exercise 1.9 together show that the supremum property (completeness) and the nested interval property are equivalent.

1.10: Let $I_n = (0,\frac{1}{n})$. Show that $\cap\, I_n = \phi$.

1.11: Let $K_n = [n,\infty) = \{x : x \geq n\}$. Show $\cap K_n = \phi$.

1.12: If a set S of real numbers contains one of its upper bounds a then $a = \sup S$. Such a supremum is called a <u>maximum</u> of S .

1.13: Show that $S \subset R$ cannot have two suprema.

1.14: Show that an ordered field F is complete if and only if every non-empty subset of F which has a lower bound has an infimum.

1.15: Show that Q is not complete.

1.16: If $S \subset R$ is bounded and $S_o \subset S$ show

$$\inf S \leq \inf S_o \leq \sup S_o \leq \sup S$$

1.17: If $S = \{(-1)^n(1 - \frac{1}{n}) : n = 1,2,\ldots\}$, find $\sup S$, $\inf S$. Prove any statements you make.

<u>Definition</u>: $x \in R$

$$|x| = \begin{cases} x \ , & \text{if } x \geq 0 \\ -x \ , & \text{if } x < 0 \ , \end{cases}$$

$|x|$ is called the <u>absolute</u> <u>value</u> of x .

Proposition:

(i) $|x| = 0 \Leftrightarrow x = 0$

(ii) $|-x| = |x| \quad , \forall x$

(iii) $|xy| = |x||y| \quad , \forall x,y$

(iv) If $c \geq 0$ then $|x| \leq c \Leftrightarrow -c \leq x \leq c$

(v) $||a|-|b|| \leq |a\pm b| \leq |a| + |b|$ (triangle inequality).

Exercise.

1.18: Prove (i) - (iv).

Proof of (v):

$$(iv) \Rightarrow -|a| \leq a \leq |a| \quad , \quad -|b| \leq b \leq |b|$$

$$-(|a|+|b|) \leq a+b \leq |a| + |b|$$

$$(iv) \Rightarrow |a+b| \leq |a| + |b|$$

which is the right-hand inequality. This implies the left-hand inequality since

$$|b| = |b-a+a| \leq |b-a| + |a|$$

$$\Rightarrow |b|-|a| \leq |b-a| = |a-b| \quad ,$$

and, interchanging a,b

$$|a|-|b| \leq |a-b|$$

$$\therefore \quad ||b|-|a|| \leq |a-b| \quad .$$

The remainder of (v) follows by replacing b by $-b$.

CARTESIAN SPACES

Recall $A \times B = \{(x,y) : x \in A , y \in B\}$.

Definition: $R^n = R \times R \times \ldots \times R$ (n times)

$$= \{(x_1,\ldots,x_n) : x_i \in R , i = 1,\ldots,n\}$$

x_i , $i = 1,\ldots,n$: the components of $(x_1,\ldots,x_n)$

$p = (x_1,\ldots,x_n)$: a point in R^n or a vector in R^n

$0 = (0,\ldots,0)$: zero vector, origin.

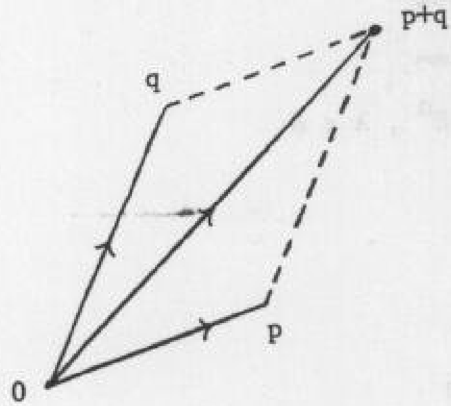
Algebra:

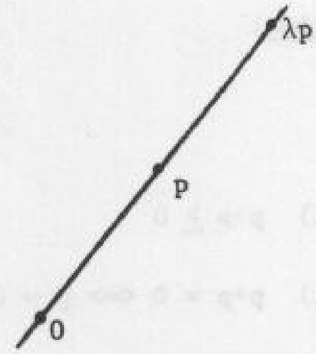
$$p = (x_1,\ldots,x_n) \quad q = (y_1,\ldots,y_n)$$

$$p+q = (x_1+y_1,\ldots,x_n+y_n) \qquad \text{(addition)}$$

$$\lambda p = (\lambda x_1,\ldots,\lambda x_n) \qquad \text{(scalar multiplication)}$$

where $\lambda \in R$ (i.e. λ a scalar).





The following properties follow immediately from those of R .

(i) $p+q = q+p$

(ii) $(p+q)+r = p+(q+r)$

(iii) $0+p = p+0 = p$

(iv) $p+(-1)p = 0$

(v) $1p = p$, $0p = 0$

(vi) $\lambda(\mu p) = (\lambda\mu)p$

(vii) $\lambda(p+q) = \lambda p + \lambda q$, $(\lambda+\mu)p = \lambda p + \mu p$.

This means R^n is a <u>vector space</u>.

<u>Inner Product</u>:

$$p = (x_1,\ldots,x_n) \quad q = (y_1,\ldots,y_n)$$

$$p\cdot q = \sum_{i=1}^{n} x_i y_i = x_1y_1 + \ldots + x_ny_n \quad .$$

Properties:

$$\forall\ p,q,r \in R^n\ ,\ \lambda \in R$$

(i) $p\cdot p \geq 0$

(ii) $p\cdot p = 0 \Leftrightarrow p = 0$

(iii) $p\cdot q = q\cdot p$

(iv) $p\cdot(q+r) = p\cdot q + p\cdot r$

(v) $(\lambda p)\cdot q = \lambda(p\cdot q) = p\cdot(\lambda q)$.

Norm:

$$|p| = \sqrt{p\cdot p} = (x_1^2 + \dots + x_n^2)^{1/2} \quad .$$

Theorem 1.3: (Cauchy-Bunyakowski-Schwarz inequality).

$$p\cdot q \leq |p|\,|q| \qquad \forall\ p,q \in R^n$$

and '=' holds $\Leftrightarrow$ $p = \lambda q\ ,\ \lambda > 0\ ,$ if $p,q \neq 0$.

Proof: Let $r = \lambda p - \mu q\ ,\ \lambda,\mu \in R$. By (i),

$$0 \leq r\cdot r$$

$$= \lambda^2\, p\cdot p - 2\lambda\mu\ p\cdot q + \mu^2\ q\cdot q \quad \text{(by (iii),(iv),(v))}$$

$$= |q|^2|p|^2 - 2|p|\,|q|\ p\cdot q + |p|^2|q|^2$$

(choosing $\lambda = |q|\ ,\ \mu = |p|$)

$$= 2|p|\,|q|\ \{|p|\,|q| - (p\cdot q)\}$$

and hence $p \cdot q \leq |p||q|$. If equality holds in this last expression then work backwards through the proof to obtain

$$0 = r = |q|p - |p|q \quad , \quad \text{i.e. } p = \frac{|p|}{|q|} q \quad . \qquad \square$$

<u>Corollary 1.3.1</u>: $|p \cdot q| \leq |p||q|$, i.e., for any n-tuples $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n)$

$$|x_1 y_1 + \ldots + x_n y_n| \leq (x_1^2 + \ldots + x_n^2)^{1/2} (y_1^2 + \ldots + y_n^2)^{1/2} \quad .$$

<u>Proof</u>: Exercise 1.19.

<u>Corollary 1.3.2</u>: (Triangle inequality).

$$||p| - |q|| \leq |p \pm q| \leq |p| + |q| \quad .$$

<u>Proof</u>:

$$\begin{aligned} |p+q|^2 &= (p+q) \cdot (p+q) = p \cdot p + 2p \cdot q + q \cdot q \\ &= |p|^2 + 2p \cdot q + |q|^2 \\ &\leq |p|^2 + 2|p||q| + |q|^2 \qquad \text{(CBS inequality)} \\ &= (|p| + |q|)^2 \end{aligned}$$

$$\therefore \quad |p+q| \leq |p| + |q| \quad .$$

The left-hand inequality follows from this as in the scalar case.

Properties of Norm:

(i) $|p| \geq 0 \quad \forall\, p \in R^n$

(ii) $|p| = 0 \Leftrightarrow p = 0$

(iii) $|\lambda p| = |\lambda||p|$

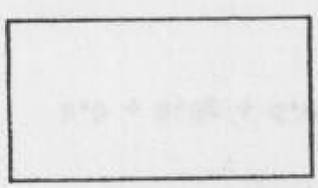
(iv) $||p|-|q|| \leq |p \pm q| \leq |p| + |q|$

An interval in R^n :

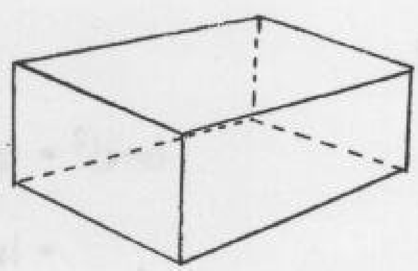
$$I = I_1 \times \ldots \times I_n$$

where I_i are intervals in R , $i = 1,\ldots,n$. I is called a closed interval if I_i are closed intervals in R; given $a_i < b_i$ $i=1,\ldots,n$:

$$I = \{(x_1,\ldots,x_n) : a_i \leq x_i \leq b_i \text{ , } i = 1,\ldots,n\} \text{ .}$$



Interval in R^2



Interval in R^3

Theorem 1.4. (Nested interval property). If $\{I_k\}$, is a sequence of closed intervals in R^n such that $I_{k+1} \subset I_k$, $k = 1,2,\ldots$, then

$$\bigcap_{k=1}^{\infty} I_k \neq \phi \text{ .}$$

<u>Proof</u>: If $I_k = I_{k1} \times \ldots \times I_{kn}$, $k = 1,2,\ldots$ where I_{ki} are closed intervals in R then $I_{k+1,i} \subset I_{k,i}$, $k = 1,2,\ldots$ and $\bigcap_k I_{ki} \neq \phi$ for each $i = 1,\ldots,n$. i.e.,

$$\exists\ x_i \in I_{ki}\ \forall k \quad , \quad i = 1,\ldots,n$$

$$(x_1,\ldots,x_n) \in I_{k1} \times \ldots \times I_{kn} = I_k\ , \forall k \quad . \qquad \square$$

FUNCTIONS

<u>Definition</u>: A subset f of $A \times B$ is a <u>function</u> from A to B if

$$(x,y_1),(x,y_2) \in f \Rightarrow y_1 = y_2 \quad .$$

Write: $f : A \to B$ ("f is a function from A to B").

Notation: If $(x,y) \in f$, $y = f(x)$

$$R_f = \{y : (x,y) \in f\} \quad \text{, the range of } f$$

$$D_f = \{x : (x,y) \in f\} \quad \text{, the domain of } f$$

If $U \subset A$, $f(U) = \{f(x) : x \in U\}$, <u>image</u> of U . If $V \subset B$, $f^{-1}(V) = \{x : f(x) \in V\}$, <u>inverse</u> <u>image</u> of V .

<u>Example</u>: Consider $f = \{(x,x^2) : -1 \leq x \leq 1\}$ (i.e. $f(x) = x^2$) $f : R \to R$

$$D_f = [-1,1] \quad , \quad R_f = [0,1] \quad ,$$

$$f([-1,\tfrac{1}{2}]) = [0,1] \quad , \quad f^{-1}([-1,\tfrac{1}{2}]) = [\tfrac{-1}{\sqrt{2}},\tfrac{1}{\sqrt{2}}]$$

<u>Definition</u>: A function $f : A \to B$ is one-to-one if it also satisfies

$$(x_1,y),(x_2,y) \in f \Rightarrow x_1 = x_2 \quad .$$

Write: $f : A \xrightarrow{(1-1)} B$.

Note: $f : A \xrightarrow{(1-1)} B$ then $\{(y,x) : (x,y) \in f\}$ is also a one-to-one function, denoted $f^{-1} : B \xrightarrow{(1-1)} A$ called the inverse function of f .

<u>Notation</u>: If $f : A \to B$, $g : B \to C$

$$g\circ f(x) = g(f(x)) \quad \underline{\text{composition}} \text{ of g with f}$$

$$D_{g\circ f} = f^{-1}(D_g) \quad .$$

<u>Example</u>:

$$f : R \to R^2 \qquad f(x) = (|x|,x^2+1)$$

$$g : R^2 \to R^2 \qquad g(u,v) = (u+v,u-v)$$

$$g\circ f : R \to R^2 \qquad g\circ f(x) = (|x|+ x^2+1,|x|-x^2-1) \quad .$$

Exercises.

1.20: $|p+q|^2 + |p-q|^2 = 2\{|p|^2+|q|^2\}$ $\forall$ $p,q \in R^n$ (Parallelogram identity).

1.21: $p = (x_1,\ldots,x_n)$

$|x_i| \leq |p| \leq \sqrt{n}\ \sup\ \{|x_1|,\ldots,|x_n|\}$.

1.22: $|p+q|^2 = |p|^2 + |q|^2 \Leftrightarrow p \cdot q = 0$.

In this case one says p and q are orthogonal.

1.23: Is it true that

$$|p+q| = |p| + |q|$$

$\Leftrightarrow$ $p = \lambda q$ or $q = \lambda p$ with $\lambda \geq 0$?

1.24: Two sets A and B have the same cardinality if there is a one-to-one function $\phi : A \to B$ such that $\phi(A) = B$ and $\phi^{-1}(B) = A$. Show that the following sets have the same cardinality:

(a) $N = \{1,2,3,\ldots\}$ and $\{2,4,6,\ldots\}$.

(b) $[0,1]$ and $[0,2]$.

(c) $(0,1)$ and $(0,\infty) = \{x : x > 0\}$.

(d) $[0,1]$ and $[0,1)$.

1.25: A set A is said to be finite if it has the same cardinality as an initial segment $\{1,\ldots,n\}$ of the natural numbers and is said to be infinite otherwise. (This means the elements can be labeled $a_1,\ldots,a_n$.) Show that a finite set of real numbers contains its sup and inf. (Hint: induction.)

1.26: A set A is countable if it has the same cardinality as N, the set of natural numbers, or is finite. Otherwise it is said to be uncountable.

Show that a countable set need not contain its sup or inf. (countability means the elements can be labeled $\{a_1, a_2, \ldots\}$).

1.27: Show that the union of a countable collection of countable sets is countable. [Hint:

$$\begin{array}{l} S_1 : a_{11}, a_{12}, a_{13}, \ldots \\ S_2 : a_{21}, a_{22}, a_{23}, \ldots \\ S_3 : a_{31}, a_{32}, a_{33}, \ldots \end{array}$$

The elements may be 'counted' by the scheme indicated omitting repetitions. Deduce that Q is countable.

1.28: $[0,1]$ is uncountable. [Sketch of proof: Suppose

$$[0,1] = \{a_1, a_2, \ldots\} \quad . \quad *$$

At least one of the intervals $[0,\frac{1}{3}]$, $[\frac{1}{3},\frac{2}{3}]$, $[\frac{2}{3},1]$ does not contain a_1; call this interval I_1. Subdivide I_1 into three closed subintervals; at least one of these does not contain a_2, call it I_2. Continuing like this we obtain a nested sequence of closed intervals $I_n \ni a_n \notin I_n$, $n = 1,2,\ldots$.

$$\therefore \quad a_n \notin \bigcap_{k=1}^{\infty} I_k \ , \quad n = 1,2,\ldots \ .$$

But $\bigcap_{k=1}^{\infty} I_k \neq \phi$ so there exists a number $x \in [0,1] \ni x \neq a_n \ \forall n$ contradicting *.]

Convexity:

Definition: $p,q \in R^n$, $p \neq q$

(i) $\{p + t(q-p) : t \in R\}$ **line** through p and q .

(ii) $\{p + t(q-p) : t \in [0,1]\}$ **line segment** between p and q .

Definition: A subset C of R^n is **convex** if

$$p,q \in C \Rightarrow p + t(q-p) \in C \quad , \forall\, t \in [0,1] \quad ,$$

i.e. if $p,q \in C$ then the line segment between them is a subset of C .

Example: $S = \{p : |p| \leq 1\}$ is convex.

Proof: $|p| \leq 1$, $|q| \leq 1$.

$$\begin{aligned} \Rightarrow |p + t(q-p)| &= |(1-t)p + t\,q| \\ &\leq |(1-t)p| + |tq| \qquad \text{(triangle inequality)} \\ &= (1-t)|p| + t|q| \qquad 0 \leq t \leq 1 \\ &\leq (1-t) + t = 1 \end{aligned}$$

$\therefore$ $p + t(q-p) \in S$, $0 \leq t \leq 1$, i.e. S is convex.

Exercises:

1.28: Prove that $\{p : |p| = 1\}$ is not convex.

1.29: Prove that $\{(x,y) \in R^2 : y > 0\}$ is convex.

1.30: Let $\mathcal{C}$ be any collection of convex sets. Show that $\cap\mathcal{C}$ is convex. Is $\cup\mathcal{C}$ necessarily convex?

1.31: The convex hull H(A) of a set A is the intersection of all convex sets containing A as a subset. Prove H(A) is convex. What is H(A) if A is a set containing two points only?

1.32: A subset C of R^n is a cone if $\{tx : x \in C\} \subset C$ for all $t \geq 0$.

(i) Prove that a cone C is a convex set if and only if

$$\{x+y : x \in C, y \in C\} \subset C .$$

(ii) Draw pictures of convex and non-convex cones in R^2.

Exercises p. 14-15 (Buck).

TOPOLOGY

Definition:

(i) If $\rho > 0$, $p_o \in R^n$ then

$$B(p_o,\rho) = \{p : |p-p_o| < \rho\}$$

is the open ball of centre p_o and radius ρ.

(ii) A <u>neighbourhood</u> of p_o is any set U which contains an open ball with centre p_o as a subset.

(iii) A set A is <u>open</u> in R^n if it is a neighbourhood of each of its points.

<u>Examples</u>:

(1) R^n is open in R^n (Why?)

(2) ϕ is open in R^n (Why?)

(3) $(0,1)$ is open in R [If $x_o \in (0,1)$ then

$$B(x_o,\delta) = (x_o-\delta, x_o+\delta) \subset (0,1) \text{ where } \delta = \min\{x_o, 1-x_o\} .]$$

(4) $[0,1)$ is <u>not</u> open in R $[B(0,\delta) = (-\delta,\delta) \not\subset [0,1), \forall\, \delta > 0 .]$

(5) $\{(x,y) : 0 < x < 1 , y = 1\}$ is <u>not</u> open in R^2 (Why?)

(6) $B(p_o,\rho)$ is open in R^n.

To see #6 let $p_1 \in B(p_o,\rho)$. We will show that $B(p_1,\rho_1) \subset B(p_o,\rho)$ where $\rho_1 = \rho - |p_1-p_o| > 0$, so that $B(p_o,\rho)$ is a neighbourhood of each of its points p_1 and hence is open. Let $p \in B(p_1,\rho_1)$

$$\begin{aligned} |p-p_o| &= |p-p_1 + p_1-p_o| \\ &\leq |p-p_1| + |p_1-p_o| && \text{(triangle inequality)} \\ &< \rho_1 + |p_1-p_o| && (p \in B(p_1,\rho_1)) \\ &= \rho - |p_1-p_o| + |p_1-p_o| && \text{(definition of } \rho_1) \\ &= \rho \end{aligned}$$

$$\therefore \quad p \in B(p_o,\rho) \quad \text{also}$$

$$\therefore \quad B(p_1,\rho_1) \subset B(p_o,\rho) \quad .$$

Open Set Properties.

(a) ϕ and R^n are open.

(b) A open, B open => $A \cap B$ open.

(c) The union of any collection of open sets is open.

Proofs:

(a) ϕ and R^n are neighbourhoods of their points (vacuously true in the case of ϕ)

(b) $p_o \in A \cap B \Rightarrow p_o \in A$ and $p_o \in B$

$\Rightarrow \exists \rho_1,\rho_2 \ni B(p_o,\rho_1) \subset A$, $B(p_o,\rho_2) \subset B$ (A,B open)

$\Rightarrow B(p_o,\rho) \subset A$ and $B(p_o,\rho) \subset B$ where $\rho = \min\{\rho_1,\rho_2\}$

$\Rightarrow B(p_o,\rho) \subset A \cap B$

$\therefore$ $A \cap B$ open.

(c) Let C be a collection of open sets.

$p_o \in \cup C \Rightarrow p_o \in A$ for some $A \in C$

$\Rightarrow B(p_o,\rho) \subset A$ for some $\rho > 0$ (A open)

$\Rightarrow B(p_o,\rho) \subset \cup C$

$\Rightarrow \cup C$ open. □

Exercise:

1.33: Property (b) implies that the intersection of any finite collection of open sets is open. Show that it is not true that this holds for all infinite collections of open sets.

Definition: A set A is closed in R^n if its complement

$$A^c = R^n - A = \{p \in R^n : p \notin A\}$$

is open.

Closed set Properties.

(a) R^n and ϕ are closed.

(b) A closed, B closed $\Rightarrow$ $A \cup B$ closed

(c) The intersection of any collection of closed sets is closed.

Exercise.

1.34: Prove (a), (b), (c).

Examples:

(1) $[0,\infty) = \{x : x \geq 0\}$ is closed in R [$(-\infty,0)$ is open]

(2) $[0,1]$ is closed in R [$(-\infty,0) \cup (1,\infty)$ is open]

(3) $[0,1)$ is not closed in R (Why?)

(4) $\{p : |p| \geq 1\}$ is closed in R^n [$B(0,1)$ is open]

(5) $\{p : |p| \leq 1\}$ is closed in R^n (Exercise).

Proposition:

$$C \text{ closed, } V \text{ open} \quad \Rightarrow \begin{cases} \text{(a)} \quad C-V \text{ closed} \\ \\ \text{(b)} \quad V-C \text{ open} \end{cases} .$$

Proof:

(a) $C-V = \{p : p \in C , p \notin V\}$

$= C \cap V^c$, closed since C and V^c are closed.

(b) Exercise.

Remark: Notice that there are sets which are neither open nor closed (e.g. $[0,1)$ in R).

The sets R^n and ϕ are both open and closed. We will see that these are the only ones in R^n with this property.

Definition: $S \subset R^n$. p_o is a cluster point of S if each neighbourhood of p_o contains a point $p \in S$, $p \neq p_o$.

Remark: A cluster point of S need not be an element of S , e.g. $\{\frac{1}{n} : n = 1,2,\ldots\}$ has the cluster point 0 .

<u>Definition</u>: $S \subset R^n$ is <u>bounded</u> if $\exists\ \rho > 0 \ni$

$$S \subset B(0,\rho) \quad (|p| < \rho\ ,\ \forall\ \ p \in S)\ .$$

<u>Equivalently</u> S is bounded if it is contained in some closed interval in R^n .

<u>Exercise</u>.

1.35: Prove these two definitions are equivalent.

<u>Notation</u>: $I = [a_1,b_1] \times \ldots \times [a_n,b_n]$: closed interval in R^n

$[a_i,b_i]$: <u>sides</u> of I

$$\lambda(I) = \sqrt{(b_1-a_1)^2 + \ldots + (b_n-a_n)^2} \quad : \text{ diameter of } I\ .$$

Note that if $p,q \in I$ then $|p-q| \leq \lambda(I)$.

<u>Theorem 1.5</u>. (Bolzano-Weierstrass Theorem): Every bounded infinite subset of R^n has a cluster point.

<u>Proof</u>: Let K be a bounded infinite subset of R^n .

$$\therefore \quad K \subset I_1 \quad , \quad \text{a closed interval.}$$

Bisect the sides of I_1 obtaining 2^n subintervals of I_1 . At least one of these new intervals contains infinitely many points of K – call it I_2 .

Bisect the sides of I_2. At least one of the resulting intervals contains infinitely many points of K - call it I_3.

Inductively, we thus define a sequence of closed nested intervals I_k each one containing infinitely many points of K. Furthermore

(1) $0 < \lambda(I_k) = \frac{1}{2^{k-1}} \lambda(I_1)$, $k = 1,2,\dots$

(2) $\exists\ p_o \in \bigcap_k I_k \neq \phi$ (by the Nested Interval Property, Theorem 1.4).

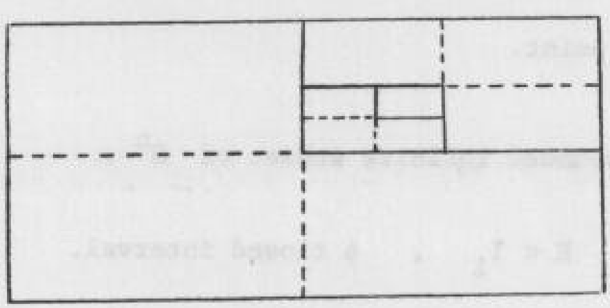
Claim: p_o is a cluster point of K.

(3) If not $\exists\ \rho > 0 \ni B(p_o,\rho)$ contains no points of K (except possibly p_o)

$$\exists\ k \ni 0 < \frac{1}{2^{k-1}} \lambda(I_1) < \rho \qquad \text{(Why?)}$$

$$0 < \lambda(I_k) < \rho \qquad \text{(by(1))}$$

If $p \in I_k$, $|p-p_o| \leq \lambda(I_k) < \rho$, i.e. $p \in B(p_o,\rho)$, i.e. $I_k \subset B(p_o,\rho)$. But I_k contains infinitely many points of K , contradicting (3). Therefore p_o is a cluster point of K.



<u>Theorem 1.6</u>: A subset K of R^n is closed

<=> K contains all of its cluster points.

<u>Proof</u>: "=>" : Let K be closed. Let p_o be a cluster point of K . If $p_o \notin K$ (i.e. $p_o \in K^c$) then since K^c is open

$$\exists\ \rho > 0 \ni B(p_o,\rho) \subset K^c$$

$$\text{i.e.} \quad B(p_o,\rho) \cap K = \phi$$

contradicting that p_o is a cluster point of K .

$$\therefore \quad p_o \in K \ .$$

"<=" : Let K contain all its cluster points. If $p_o \in K^c$ then $p_o \notin K$ and is not a cluster point of K .

$$\Rightarrow \exists \rho > 0 \ni B(p_o,\rho) \subset K^c$$

$$\Rightarrow \quad K^c \text{ open } \Rightarrow K \text{ closed .} \qquad \square$$

<u>Definition</u>:

(i) A collection G of open sets is an <u>open</u> <u>cover</u> of a set K if

$$K \subset \cup G \ .$$

(ii) A set K in R^n is <u>compact</u> if every open cover of K has a finite subcover.

Examples.

(1) A finite set is compact.

(2) $[0,\infty)$ is not compact [Let $G_n = (-1,n)$, $n = 1,2,\ldots$, $\{G_n : n = 1,2,\ldots\}$ is an open cover of $[0,\infty)$. There is no finite subcover. (Why?) $\therefore$ $[0,\infty)$ is not compact.]

(3) $(0,1)$ is not compact. (Exercise.)

Theorem 1.7: (Heine-Borel Theorem). A subset K of R^n is compact

$\Leftrightarrow$ K is closed and bounded.

Proof: "$\Rightarrow$" : Let K be compact.

(a) K is closed: Let p_o be any point in K^c ; we wish to show that $B(p_o,\rho) \subset K^c$ for some $\rho > 0$ so that K^c is open and hence K is closed. Consider

$$G_k = \{p : |p-p_o| > \tfrac{1}{k}\} \quad \text{(open set)} .$$

$G = \{G_k : k = 1,2,\ldots\}$ covers $R^n - \{p_o\}$ and, in particular, G covers K . Since K is compact there is a finite subcover $\{G_{k_i} : i = 1,\ldots,m\}$ of K . Let $k_o = \max\{k_i : i = 1,\ldots,m\}$. Since $G_i \supset G_j$ if $i \geq j$, $K \subset G_{k_o}$, and hence $G^c_{k_o} = \{p : |p-p_o| \leq \frac{1}{k_o}\} \subset K^c$.

$$\therefore \; B(p_o,\tfrac{1}{k_o}) \subset K^c$$

$$\therefore \; K^c \text{ open (i.e. K closed).}$$

(b) K is <u>bounded</u>: Consider $G_k = \{p : |p| < k\} = B(0,k)$ (open). $\{G_k : k = 1,2,\ldots\}$ covers R^n and hence covers K. Therefore, a finite subcover $\{G_{k_i} : i = 1,\ldots,m\}$ of K exists since K is compact. Since

$$G_i \subset G_j \text{ if } i \leq j \text{ , } K \subset G_{k_o} \text{ , } k_o = \max\{k_i\} \text{ .}$$

$$\therefore \ |p| \leq k_o \text{ , } \forall\, p \in K \text{ .}$$

"<=" Let K be closed and bounded.

If K is not compact then there exists an open cover by $\mathcal{G} = \{G_\alpha\}$ of K such that K is not contained in the union of finitely many G_α's. Now, K is bounded so $K \subset I_1$, some closed interval. Bisect the sides of I_1 as in the proof of Theorem 1.5 (Bolzano-Weierstrass Theorem). At least one of the resulting 2^n closed intervals intersects K and this intersection cannot be covered by finitely many G_α's. Call this I_2. Proceeding like this we obtain a nested sequence $\{I_n\}$ of closed intervals each of whose intersections with K cannot be covered by finitely many G_α's. Further

$$(1) \qquad \lambda(I_k) = \frac{\lambda(I_1)}{2^{k-1}} \text{ , } k = 1,2,\ldots \text{ .}$$

There is a point $p_o \in I_k$, $\forall\, k$ (nested intervals theorem), and as in the B-W Theorem, p_o is a cluster point of K.

$$\therefore \ p_o \in K \text{ , since } K \text{ is closed (Theorem 1.6)}$$

$$\therefore \ p_o \in G_\alpha \text{ for some } \alpha \ (\mathcal{G} \text{ covers } K)$$

$$\therefore \ p_o \in B(p_o,\rho) \subset G_\alpha \text{ , some } \rho > 0 \ (G \text{ open})$$

$$\exists\, k \ni \lambda(I_k) < \rho \quad \text{(by (1))} \text{ .}$$

If $q \in I_k$, $|q-p_o| \leq \lambda(I_k) < \rho$

$$\therefore \quad q \in B(p_o,\rho)$$

$$\therefore \quad I_k \subset B(p_o,\rho) \subset G_\alpha \quad .$$

Therefore the intersection of K with I_k is covered by <u>one</u> G_α contradicting the definition of I_k . So our assumption that K is not compact must be false. □

<u>Definition</u>: A subset D of R^n is <u>disconnected</u> if there exist open sets A and B such that

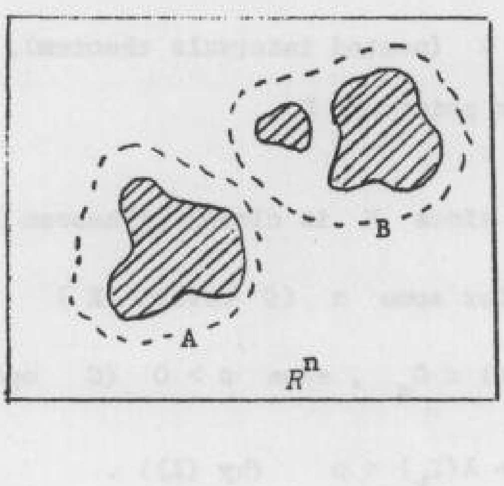
(i) $A \cap D \neq \phi$, $B \cap D \neq \phi$

(ii) $(A \cap D) \cap (B \cap D) = \phi$

(iii) $(A \cap D) \cup (B \cap D) = D$.

A and B are called a <u>disconnection</u> of D .

D is said to be <u>connected</u> if it is not disconnected.



Examples:

(1) $N = \{1,2,\ldots\}$ is disconnected $[(-\infty,\frac{3}{2}),(\frac{3}{2},\infty)$ is a disconnection of $N]$.

(2) Q is disconnected $[(-\infty,\sqrt{2}),(\sqrt{2},\infty)$ is a disconnection of Q].

Proposition: $[0,1]$ is connected.

Proof: Suppose $[0,1]$ is disconnected. $\exists$ A,B open in $R \ni$

(i) $A \cap [0,1] \neq \phi$, $B \cap [0,1] \neq \phi$

(ii) $(A\cap[0,1])\cap(B\cap[0,1]) = \phi$

(iii) $(A\cap[0,1])\cup(B\cap[0,1]) = [0,1]$.

We may suppose without loss of generality that $1 \in B$. Let

$$c = \sup(A\cap[0,1]) \quad ; \quad \therefore \quad 0 \leq c \leq 1 \quad \text{(Why?)}$$

$$\therefore \quad c \in A \cup B \quad \text{(i.e. } c \in A \text{ or } c \in B) .$$

If $c \in A$, $c \in A \cap [0,1]$. But $1 \in B$ so

(ii) $\Rightarrow$ $c \in A \cap [0,1)$; since A is open there exist points in $A \cap [0,1)$ exceeding c contradicting $c = \sup(A \cap [0,1])$. If $c \in B$, $(c-\delta,c) \subset B$ for some $\delta > 0$ (B open) so if $c \geq x$, $\forall x \in A \cap [0,1]$ every number in $(c-\delta,c)$ is also an upper bound of $A \cap [0,1]$ again contradicting the definition of c . Thus no disconnection of $[0,1]$ exists, i.e., $[0,1]$ is connected. □

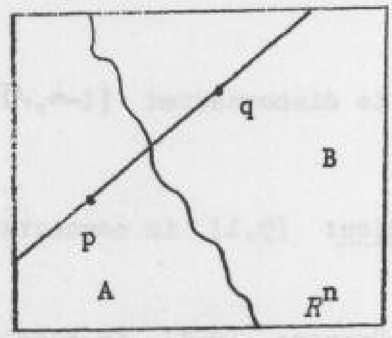
<u>Theorem 1.8</u>: R^n is connected.

<u>Proof</u>: If R^n is disconnected, there exist open sets $A,B \subset R^n$ such that

(i) $A \neq \phi$, $B \neq \phi$

(ii) $A \cap B = \phi$

(iii) $A \cup B = R^n$.



Let $p \in A$, $q \in B$ and consider

$$A_1 = \{t \in R : p + t(q-p) \in A\}$$

$$B_1 = \{t \in R : p + t(q-p) \in B\} \quad .$$

A,B open in $R^n \Rightarrow A_1$, B_1 open in R (this is not obvious so think about it). Now (i), (ii), (iii) above =>

$(i)_1$ $\quad A_1 \cap [0,1] \neq \phi$, $B_1 \cap [0,1] \neq \phi$ $\;(0 \in A_1$, $1 \in B_1)$

$(ii)_1$ $\quad (A_1 \cap [0,1]) \cap (B_1 \cap [0,1]) = \phi$

$(iii)_1$ $\quad (A_1 \cap [0,1]) \cup (B_1 \cap [0,1]) = [0,1]$.

So A_1 , B_1 is a disconnection of $[0,1]$ which, we have seen, is connected. So R^n is connected. □

<u>Corollary 1.8.1</u>: The only sets in R^n which are both open and closed are R^n and ϕ .

<u>Proof</u>: If $A \neq R^n$, ϕ, is both open and closed then so also is A^c . A and A^c would provide a disconnection of R^n . □

More restrictive notions of connectedness are sometimes used. A set C is <u>polygonally</u> <u>connected</u>, if for each pair of points p_o, $p_m \in C$, there is a finite subset $\{p_1,\ldots,p_{m-1}\}$ of C such that the polygon

$$\{p_{i-1} + t(p_i - p_{i-1}) : t \in [0,1] , i = 1,\ldots,m\}$$

is a subset of C. C is <u>arcwise</u> <u>connected</u> if, for each pair of points $p,q \in C$ there is a path joining p and q lying entirely in C (i.e., there is a continuous function $f : [0,1] \to C$ such that $f(0) = p$, $f(1) = q$). either of these types of connectedness implies connectedness in the sense adopted here. However the converse is not true; e.g. $\{(x,y) \in R^2 ; x \neq 0 , 0 < y \leq x^2\} \cup \{(0,0)\}$ is connected (in fact arcwise connected) but is not polygonally connected; $\{(x,y) \in R^2 : y = \sin \frac{1}{x} , x \neq 0\} \cup \{(0,y) : -1 \leq y \leq 1\}$ is connected but is not arcwise connected.

<u>Exercises</u>:

1.36: Prove that the intersection of any finite collection of open sets is open [use Property (b) and induction].

1.37: Prove that $\{p : |p| \leq 1\}$ is closed in R^n.

1.38: Prove that a subset U of R^n is open if and only if it is the union of a collection of open balls.

1.39: If A is a subset of R^n then $\overline{A}$, the <u>closure</u> of A, is the intersection of all closed sets which contain A as a subset.

(a) $\overline{A}$ is closed.

(b) $A \subset \overline{A}$

(c) $\overline{\overline{A}} = \overline{A}$

(d) $\overline{(A \cup B)} = \overline{A} \cup \overline{B}$

(e) $\overline{\phi} = \phi$

(f) Observe that $\overline{A}$ is the smallest closed set containing A .

(g) Prove that $\overline{B}\,(0,1) = \{p : |p| \leq 1\}$.

(h) If A and B are any subsets of R then is $\overline{A \cap B} = \overline{A} \cap \overline{B}$?

1.40: If A is a subset of R^n then A^o , the interior of A , is the union of all open sets contained in A .

(a) A^o is open

(b) $A^o \subset A$

(c) $(A^o)^o = A^o$

(d) $(A \cap B)^o = A^o \cap B^o$

(e) $(R^n)^o = R^n$

(f) Observe that A^o is the largest open set contained in A .

(g) Prove that $(\overline{B}\,(0,1))^o = B(0,1)$.

(h) Is there a subset A of R such that $A^o = \phi$ and $\overline{A} = R$?

1.41: Let $A \subset R^n$. The __derived set__ A' of A is the set of all cluster points of A .

(a) Prove that A' is closed.

(b) Prove that $\overline{A} = A \cup A'$.

(c) Theorem 1.6 says that A is closed $\Leftrightarrow A' \subset A$. A set A such that $A' = A$ is called __perfect__. Give examples of perfect and non-perfect closed sets.

1.42: If $A \subset R^n$ then ∂A , the boundary of A , is the set of all points p such that each neighbourhood of p contains a point of A and a point of A^c .

(a) Show A closed $\Leftrightarrow \partial A \subset A$.

(b) Show ∂A is closed (i.e. $\partial(\partial A) \subset \partial A$) .

(c) Show $\partial A = \overline{A} - A^o$.

1.43: For each of the following sets __state__ $\overline{A}$, A^o , A' , ∂A .

(a) $\{p \in R^n : 0 < |p| < 1\}$

(b) $\{\frac{1}{n} \in R : n = 1,2,3,\ldots\}$

(c) $\{(\frac{1}{n},\frac{1}{m}) \in R^2 : n,m = 1,2,3,\ldots\}$

(d) $\{p \in R^n : |p| < 1\}$.

1.44: Without using the Heine-Borel Theorem show that $\{(x,y) : x^2+y^2 < 1\}$ is not compact in R ?

1.45: Let A and B be open in R ; prove that $A \times B$ is open in R^2 .

1.46: Show that a finite subset of R^n is closed.

1.47: Show that a countable subset of R is not open. Show that it may or may not be closed.

1.48: Let S be an uncountable subset of R . Show that S has a cluster point. [Hint: Show that at least one interval [n, n+1] must contain uncountably many points of S .]

1.49: Show that a closed interval is a closed set.

1.50: Let Q^2 denote the set of points in R^2 with rational co-ordinates. What is the interior of Q^2 ? The boundary of Q^2 ? Show that Q^2 is not connected.

1.51: Show that Q^c is not connected. Show that $(Q^2)^c$ is connected - in fact polygonally connected.

1.52: Show that an open set which is connected is polygonally connected.

References for Chapter I

R.G. Bartle: The Elements of Real Analysis (Chapters I, II).

R.C. Buck: Advanced Calculus (Chapter I).

H. Sagan: Advanced Calculus (Chapters I, V).

CHAPTER TWO

LIMITS, CONTINUITY AND DIFFERENTIATION

SEQUENCES

Definition: A sequence in R^n is a function $p : N \to R^n$. Sequences are usually denoted $\{p_k\}$ where $p_k = p(k)$.

Examples:

(1) $p_k = \frac{1}{k}$; $\{p_k\}$ is a sequence in R.

(2) $p_k = (\frac{1}{k}, \sin(k^2))$; $\{p_k\}$ is a sequence in R^2.

Definition: A sequence $\{p_k\}$ in R^n is convergent if $\exists\ p \in R^n$ such that for each neighbourhood U of p $\exists$ a natural number $N = N(U)$ for which $k \geq N \Rightarrow$

$$p_k \in U .$$

Write: $\lim_{k\to\infty} p_k = p$, or $\lim \{p_k\} = p$. The sequence $\{p_k\}$ is said to be divergent if it is not convergent. Equivalently; $\{p_k\}$ converges if there is a point $p \in R^n$ such that for each $\varepsilon > 0\ \exists\ N = N(\varepsilon) \in N \ni$ if $k \geq N$ then

$$|p_k - p| < \varepsilon \quad ;$$

p is called the limit of the sequence.

Exercise:

2.1: Show that the two definitions are equivalent.

Proposition: A convergent sequence cannot have two limits.

Proof: Suppose $\lim \{p_k\} = p$ and $\lim \{p_k\} = q$. If $\varepsilon > 0$ then

$$\exists\ N_1 \ni k \geq N_1 \Rightarrow |p_k - p| < \varepsilon$$

and

$$\exists\ N_2 \ni k \geq N_2 \Rightarrow |p_k - q| < \varepsilon \quad .$$

Let $k = \sup \{N_1, N_2\}$; then

$$\begin{aligned} |q-p| &= |q-p_k+p_k-p| \\ &\leq |p_k-q| + |p_k-p| \\ &< 2\varepsilon \quad . \end{aligned}$$

Therefore $|q-p| < 2\varepsilon$ for each $\varepsilon > 0$; thus $|q-p| = 0$ (Archimedean Property), i.e. $q = p$.

Examples:

(1) $x_k = 1$, $k = 1,2,\ldots$ $\lim_{k\to\infty} x_k = 1$.

(2) $\lim_{k\to\infty} \frac{1}{k} = 0$ (from the Archimedean Property)

<u>Remarks:</u> Notice that if $\lim p_k = p$ then p is either

(a) a cluster point of the set $\{p_k : k = 1,2,\ldots\}$ or

(b) $\{p_k\}$ is ultimately constant and equal to p.

(i.e. $\exists\, N \ni k \geq N \Rightarrow p_k = p$). Note also $\lim\limits_{k\to\infty} p_k = p \Leftrightarrow \lim\limits_{k\to\infty} |p_k - p| = 0$.

<u>Theorem 2.1</u>: A convergent sequence is bounded.

<u>Proof</u>: If $\lim\limits_{k\to\infty} p_k = p$, $\exists$ a natural number N such that $k \geq N \Rightarrow |p_k - p| < 1$.

By the triangle inequality

$$|p_k| - |p| \leq |p_k - p| < 1 \quad , \quad \text{if } k \geq N$$

$$|p_k| \leq 1 + |p| \quad \text{if } k \geq N$$

$$\therefore \quad |p_k| \leq \sup \{|p_1|,\ldots,|p_{N-1}|, 1 + |p|\} \ , \quad k = 1,2,\ldots \qquad \square$$

<u>Example</u>: If $p_k = k$, $\{p_k\}$ is divergent. By the Archimedean Property $\{p_k\}$ is unbounded and so is divergent by Theorem 2.1.

<u>Definition</u>: If $\{k_j\}$ is a sequence of natural numbers such that $k_1 < k_2 < k_3 < \ldots$ then $\{p_{k_j}\}$ is called a <u>subsequence</u> of $\{p_k\}$.

<u>Theorem 2.2</u>: A bounded sequence has a convergent subsequence.

<u>Proof</u>: There are two cases to consider. Either $\{p_k : k = 1,2,\ldots\}$ is a finite set or it is infinite. If it is a finite set then there is at least one value p such that $p_k = p$ for infinitely many k - these terms in the sequence

form a convergent subsequence of $\{p_k\}$. In the other case $\{p_k : k = 1,2,\ldots\}$ is a bounded infinite set and so by the B.W. Theorem it has a cluster point p . Consider $B(p,1)$; $\exists\ k_1$ such that $p_{k_1} \in B(p,1)$. If k_j is such that $p_{k_j} \in B(p,\frac{1}{j})$ there exists $k_{j+1} > k_j$ such that $p_{k_{j+1}} \in B(p,\frac{1}{j+1})$ so by induction there exists a subsequence $\{p_{k_j}\}$ of $\{p_k\}$ such that $|p_{k_j} - p| < \frac{1}{j}$, $j = 1,2,\ldots$. The Archimedean Property implies $\lim\limits_{j\to\infty} p_{k_j} = p$. □

<u>Corollary 2.2.1</u>: $K \subset R^n$.

K compact <=> each sequence of points in K has a subsequence convergent to a point in K .

<u>Proof</u>:

"=>" : K compact <=> K closed and bounded (Heine-Borel). Thus any sequence of points in K is bounded and so by Theorem 2.2 contains a convergent subsequence. The limit of this sequence is either a cluster point of the sequence (and hence of K) and so is contained in K (closed) by Theorem 1.6, or else the sequence is ultimately constant so its limit is in K in this case also.

"<=" : Suppose each sequence of points in K has a subsequence convergent to a point in K .

<u>K is closed</u>: If K is not closed there exists a cluster point p of K $\ni$ $p \notin K$. Thus there exists a sequence of points $p_k \in K \ni \lim p_k = p \notin K$ contradicting our hypothesis.

<u>K is bounded</u>: If K is not bounded $\exists\ p_k \in K \ni |p_k| \geq k$, $k = 1,2,\ldots$.

$\{p_k\}$ is thus an unbounded sequence and each subsequence is unbounded and hence divergent again contradicting the hypothesis. Thus K is closed and bounded and is compact by the Heine-Borel Theorem. $\square$

<u>Theorem 2.3</u>: A sequence is convergent and has limit p <=> each subsequence is convergent and has limit p.

<u>Proof</u>:

"=>" Suppose $\lim_{k\to\infty} p_k = p$; i.e. for each $\varepsilon > 0 \ \exists\ N \ni j \geq N \Rightarrow |p_j - p| < \varepsilon$.
If $\{p_{k_j}\}$ is a subsequence of $\{p_k\}$ then $j \geq N \Rightarrow k_j \geq j \geq N \Rightarrow |p_{k_j} - p| < \varepsilon$.
Thus $\lim_{j\to\infty} p_{k_j} = p$ also.

"<=" Suppose each subsequence $\{p_{k_j}\}$ of $\{p_k\}$ satisfies $\lim_{j\to\infty} p_{k_j} = p$. $\{p_k\}$ is a subsequence of $\{p_k\}$ so this part is trivial. $\square$

<u>Example</u>: If $p_k = (-1)^k$, $\{p_k\}$ is divergent.

$$\lim_{k\to\infty} p_{2k} = 1 \quad , \quad \lim_{k\to\infty} p_{2k+1} = -1 \ .$$

If $\{p_k\}$ were convergent both of these limits would be the same number by Theorem 2.3.

<u>Theorem 2.4</u>. If $\{p_k\}$ and $\{q_k\}$ are sequences in R^n such that

$$\lim_{k\to\infty} p_k = p \ , \ \lim_{k\to\infty} q_k = q \ ,$$

then

(i) $\lim\limits_{k\to\infty} (p_k+q_k) = p+q$

(ii) $\lim\limits_{k\to\infty} p_k \cdot q_k = p \cdot q$.

Further if $\{x_k\}$ is a sequence in R such that

$$\lim_{k\to\infty} x_k = x$$

then

(iii) $\lim\limits_{k\to\infty} x_k p_k = xp$

(iv) $\lim\limits_{k\to\infty} \frac{1}{x_k} p_k = \frac{1}{x} p$ (if $x \neq 0$).

<u>Proof</u>: <u>Exercise</u> 2.2.

The following result shows that it is sufficient to consider only sequences in R .

<u>Theorem 2.5:</u> Let $p_k = (x_{1k},\ldots,x_{nk}) \in R^n$. $\{p_k\}$ is convergent <=> each of the sequences $\{x_{ik}\}$ is convergent, $i = 1,\ldots,n$. Furthermore,

$$\lim_{k\to\infty} p_k = p = (x_1,\ldots,x_n) \iff$$

$$\lim_{k\to\infty} x_{ik} = x_i \text{ , } i = 1,\ldots,n \text{ .}$$

Proof:

"=>" $|p_k - p| \geq |x_{ik} - x_i|$, $k = 1,2,\ldots$, $i = 1,\ldots,n$. (Exercise 1.21). Finish the proof.

"<=" $|p_k - p| \leq \sqrt{n}\ \sup\ \{|x_{ik} - x_i| \ : \ i = 1,\ldots,n\}$. Thus if $\lim_{k\to\infty} x_{ik} = x_i$, $i = 1,\ldots,n$, given $\varepsilon > 0\ \exists\ N_i \ni k \geq N_i \Rightarrow |x_{ik} - x_i| < \frac{\varepsilon}{\sqrt{n}}$, $i = 1,\ldots,n$. Let $N = \sup\ \{N_i : i = 1,\ldots,n\}$ so $k \geq N \Rightarrow$

$$|x_{ik} - x_i| < \frac{\varepsilon}{\sqrt{n}} \ , \ i = 1,\ldots,n \Rightarrow |p_k - p| < \varepsilon \Rightarrow \lim_{k\to\infty} p_n = p \quad . \qquad \square$$

Examples.

(1) $\lim_{k\to\infty} (\frac{1}{k}, \frac{1}{k^2} + 1) = (0,1)$. Use Theorems 2.4, 2.5

$\lim_{k\to\infty} \frac{1}{k} = 0$ (Archimedean property). Therefore,

$\lim_{k\to\infty} \frac{1}{k^2} = 0$ (Theorem 2.4) $\therefore$ $\lim_{k\to\infty} \frac{1}{k^2} + 1 = 1$ (Theorem 2.4).

(2) $\lim_{k\to\infty} \frac{2k+3}{k+2} = 2$ since $\frac{2k+3}{k+2} = \dfrac{2 + \frac{3}{k}}{1 + \frac{2}{k}}$ and $\lim_{k\to\infty} \frac{1}{k} = 0$

$$\Rightarrow \lim_{k\to\infty} \frac{3}{k} = \lim_{k\to\infty} \frac{2}{k} = 0 \ .$$

Exercises:

2.3: Let $x_k \leq z_k \leq y_k$. Prove that if $\{x_k\}$ and $\{y_k\}$ are convergent and have limit c then $\{z_k\}$ is convergent and has limit c .

2.4: Discuss the convergence or divergence of the sequences whose nth terms is given.

(a) $\frac{n}{n+1}$ (d) $\frac{2n^2+3}{3n^2+1}$

(b) $\frac{(-1)^n n}{n+1}$ (e) $(\frac{1}{n},n)$

(c) $\frac{2n}{3n^2+1}$ (f) $((-1)^n,\frac{1}{n})$.

2.5: If $\{x_n\}$ is a sequence of nonnegative real numbers which converges to x , show that $\lim_{n\to\infty} \sqrt{x_n} = \sqrt{x}$. (Hint: $\sqrt{x_n} - \sqrt{x} = \frac{x_n - x}{\sqrt{x_n} + \sqrt{x}}$ if $x \neq 0$.)

2.6: Let $x_n = \sqrt{n+1} - \sqrt{n}$. Discuss the convergence or divergence of $\{x_n\}$ and $\{\sqrt{n}\, x_n\}$.

2.7: Show that a set C in R^n is closed if and only if each convergent sequence in C has its limit in C .

2.8: Show $\lim_{k\to\infty} p_k = p \iff \lim_{k\to\infty} |p_k-p| = 0$.

2.9: If $0 < r < 1$ show $\lim\limits_{n\to\infty} r^n = 0$.

[Hint: $r = \frac{1}{1+s}$, $s > 0$. Show $(1+s)^n \geq 1 + ns$, $n = 1,2,\ldots$, so $0 < r^n = \frac{1}{(1+s)^n} \leq \frac{1}{1+ns}$.] Show

$$\lim_{n\to\infty} r^n = \begin{cases} 0 , & \text{if } |r| < 1 \\ 1 , & \text{if } r = 1 \end{cases}$$

and $\{r^n\}$ is divergent if $r = -1$ or $|r| > 1$.

2.10: (Ratio Test) Let $\{x_n\}$ be such that $\lim\limits_{n\to\infty} \left|\frac{x_{n+1}}{x_n}\right| = r$. Show

(a) If $0 \leq r < 1$, then $\lim\limits_{n\to\infty} x_n = 0$.

(b) If $r > 1$, then $\{x_n\}$ is divergent.

(c) If $r = 1$, give examples which show that $\{x_n\}$ may be convergent or divergent.

[Hint: In (a), if $r < s < 1$, show that for some constant A and all large enough n, $0 \leq |x_n| \leq A s^n$. Use 2.9.]

2.11: Show that the sequence $\{\frac{x^n}{n}\}$ has limit 0 if $-1 \leq x \leq 1$ and is divergent if $|x| > 1$.

2.12: Show that $\lim\limits_{n\to\infty} \frac{x^n}{n!} = 0$ for all real numbers x.

2.13: If $x > 0$ show $\lim_{n\to\infty} x^{1/n} = 1$.

[Hint: If $0 < x < 1$, given $\varepsilon > 0 \ \exists \ N \ni n \geq N \Rightarrow \frac{1}{1 + \varepsilon n} < x < 1$

(Why?).

$$\therefore \quad \frac{1}{(1+\varepsilon)^n} \leq \frac{1}{1+\varepsilon n} < x < 1 \qquad \text{(Why?)}$$

$$\frac{1}{1+\varepsilon} < x^{1/n} < 1 \quad \text{if} \quad n \geq N$$

so

$$|x^{1/n} - 1| < \frac{\varepsilon}{1+\varepsilon} < \varepsilon \ , \quad \text{if} \quad n \geq N \ .$$

Do the case $x > 1$.]

2.14: Show $\lim_{n\to\infty} n^{1/n} = 1$.

[Hint: Let $x_n = n^{1/n} - 1 > 0$. Show $n = (1+x_n)^n > \frac{n(n-1)}{2} x_n^2$ and hence $\lim x_n = 0$.]

2.15: (Root Test) Let $\{x_n\}$ be such that $\lim_{n\to\infty} |x_n|^{1/n} = r$. Show

(a) If $0 \leq r < 1$, then $\lim_{n\to\infty} x_n = 0$.

(b) If $r > 1$, then $\{x_n\}$ is divergent.

(c) If $r = 1$, $\{x_n\}$ may be either convergent or divergent.

2.16: If a and b are nonnegative real numbers show $\lim_{n\to\infty} (a^n + b^n)^{1/n} = \max\{a,b\}$.

Definition: A real sequence $\{x_n\}$ is

(i) increasing if $x_1 \leq x_2 \leq x_3 \leq \cdots$, and is

(ii) decreasing if $x_1 \geq x_2 \geq x_3 \geq \cdots$.

$\{x_n\}$ is monotone if it is increasing or decreasing.

Theorem 2.6: A monotone sequence is convergent if and only if it is bounded.

Proof: Let $\{x_n\}$ be an increasing sequence.

$\{x_n\}$ convergent $\Rightarrow$ $\{x_n\}$ bounded (Theorem 2.1)

$\{x_n\}$ bounded $\Rightarrow x = \sup \{x_n : n = 1,2,\ldots\}$ exists. If $\varepsilon > 0$, then $x-\varepsilon$ is not an upper bound of $\{x_n\}$ so $\exists\, N \ni x-\varepsilon < x_N \leq x$. Since $\{x_n\}$ is increasing

$$n \geq N \Rightarrow x-\varepsilon < x_N \leq x_n \leq x$$

$$\text{i.e. } n \geq N \Rightarrow |x_n - x| < \varepsilon .$$

Thus the increasing sequence converges to its supremum. □

Remark: Given any ordered fied F, each monotone bounded sequence in F is convergent (with its limit in F) if and only if F is complete.

Examples.

(1) $\lim\limits_{n\to\infty} r^n = 0$ if $0 \leq r < 1$. $0 \leq r^{n+1} = rr^n \leq r^n < 1$ so $\{r^n\}$ is decreasing and bounded, hence convergent with $0 \leq \lim r^n < 1$ $\{r^{n+1}\}$ is a sub-

sequence of r^n so $\lim r^{n+1} = \lim r^n$. But $\lim r^{n+1} = r \lim r^n < \lim r^n$ if $\lim r^n \neq 0$, a contradiction. [A different proof of this is outlined in Exercise 2.9.]

(2) If $a_n = \frac{n}{2^n}$, $\lim\limits_{n\to\infty} a_n = 0$. $a_{n+1} = \frac{n+1}{2^{n+1}} = \frac{n+1}{2n} a_n$.

$$\therefore \quad 0 < a_{n+1} \leq a_n$$

$$\therefore \{a_n\} \text{ is convergent and } \lim_{n\to\infty} a_n \geq 0 \quad .$$

Now $\lim a_n = \lim a_{n+1} = \lim \frac{n+1}{2n} a_n = \frac{1}{2} \lim a_n \Rightarrow \lim a_n = 0$.

(3) $a_n = (1 + \frac{1}{n})^n$, $\{a_n\}$ is convergent and $2 \leq \lim a_n \leq 3$.

$$a_n = (1 + \tfrac{1}{n})^n$$

$$= 1 + \frac{n}{1!}\frac{1}{n} + \frac{n(n-1)}{2!}\frac{1}{n^2} + \dots + \frac{n(n-1)\dots 3\cdot 2}{(n-1)!}\frac{1}{n^{n-1}} + \frac{n!}{n!}\frac{1}{n^n}$$

$$= 1 + 1 + \frac{1}{2!}(1 - \tfrac{1}{n}) + \frac{1}{3!}(1 - \tfrac{1}{n})(1 - \tfrac{2}{n}) + \dots + \frac{1}{(n-1)!}(1 - \tfrac{1}{n}) \dots (\tfrac{2}{n})$$

$$+ \frac{1}{n!}(1 - \tfrac{1}{n}) \dots (\tfrac{2}{n})\,\tfrac{1}{n} .$$

$$a_{n+1} = 1 + 1 + \frac{1}{2!}(1 - \tfrac{1}{n+1}) + \frac{1}{3!}(1 - \tfrac{1}{n+1})(1 - \tfrac{2}{n+1}) + \dots \quad .$$

Each term in the expression for a_n is no greater than the corresponding term in a_{n+1} (and there is also one more term in a_{n+1} than in a_n). Thus $a_{n+1} > a_n$ so $\{a_n\}$ is <u>increasing</u>; it remains to show it is bounded.

$$2 \le a_n \le 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}$$

$$\le 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^n} = 1 + \frac{1 - \frac{1}{2^{n+1}}}{1 - \frac{1}{2}} \le 3$$

Thus $2 \le \lim_{n\to\infty} (1 + \frac{1}{n})^n \le 3$. This is sometimes taken as the <u>definition</u> of e .

<u>Note</u>: We used the fact that $1 + r + \dots + r^n = \frac{1 - r^{n+1}}{1 - r}$ if $r \ne 1$. You will recall that this is easily proved by denoting the left-hand side by s_n and showing $s_n - r\,s_n = 1 - r^{n+1}$.

(4) If $A > 0$, $x_1 > 0$ and $x_{n+1} = \frac{1}{2}(x_n + \frac{A}{x_n})$, $n = 1,2,\dots$ then $\lim_{n\to\infty} x_n = \sqrt{A}$. This is an algorithm for computing square roots.

$$x_1 > 0 \Rightarrow x_n > 0 \quad \text{(by induction)}$$

$${x_{n+1}}^2 = \frac{1}{4}({x_n}^2 + 2A + \frac{A^2}{{x_n}^2})$$

$${x_{n+1}}^2 - A = \frac{1}{4}({x_n}^2 - 2A + \frac{A^2}{{x_n}^2}) = \frac{1}{4}(x_n - \frac{A}{x_n})^2 \ge 0$$

$${x_{n+1}}^2 \ge A \quad \text{i.e.} \quad x_{n+1} \ge \sqrt{A} \ , \quad n = 1,2,\dots$$

$$x_n \ge \sqrt{A} \ , \quad n = 2,3,\dots \quad (*)$$

$$x_n - x_{n+1} = x_n - \frac{1}{2}(x_n + \frac{A}{x_n}) = \frac{1}{2}(x_n - \frac{A}{x_n})$$

$$= \frac{1}{2}\,\frac{{x_n}^2 - A}{x_n} \ge 0 \quad \text{by } (*) \quad n = 2,3,\dots .$$

Thus $\{x_n\}$ is <u>decreasing</u> and bounded below by $\sqrt{A} > 0$ (for $n \geq 2$)

$$\therefore \quad \lim x_n = L \geq \sqrt{A} \ .$$

But $x_{n+1} = \frac{1}{2}(x_n + \frac{A}{x_n}) \Rightarrow L = \frac{1}{2}(L + \frac{A}{L}) \Rightarrow L^2 = A \Rightarrow L = \sqrt{A}$.

The results on monotone sequences are interesting in that, unlike the preceeding examples, it was not necessary to first guess the limit of a sequence in order to show that it converged. Fortunately we are able to do this in general.

<u>Definition</u>: A sequence $\{p_n\}$ in R^k is a Cauchy Sequence if, for each $\varepsilon > 0$, there exists a natural number $N = N(\varepsilon)$ $\ni$ if $n,m \geq N$ then

$$|p_n - p_m| < \varepsilon \ .$$

<u>Theorem 2.7</u>: (Cauchy Criterion). $\{p_n\}$ is convergent $\Leftrightarrow$ $\{p_n\}$ is a Cauchy sequence.

<u>Proof</u>:

"$\Rightarrow$" : Suppose $\lim\limits_{n\to\infty} p_n = p$. Thus, for each $\varepsilon > 0$, $\exists$ $N \ni n \geq N \Rightarrow$ $|p_n - p| < \frac{\varepsilon}{2}$. Hence, if $n,m \geq N$

$$\begin{aligned} |p_n - p_m| &= |p_n - p + p - p_m| \\ &\leq |p_n - p| + |p_m - p| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

i.e. $\{p_n\}$ is a Cauchy sequence.

"<=" :

(a) $\{p_n\}$ <u>is bounded and so by Theorem 2.2 (or the B.W. Theorem) has a convergent subsequence.</u>

<u>Proof</u>: $\exists N \ni n,m \geq N \Rightarrow |p_n - p_m| < 1$. In particular,

$$n \geq N \Rightarrow |p_n - p_N| < 1$$

$$\Rightarrow |p_n| < |p_N| + 1 .$$

$$\therefore \quad |p_n| < \sup\{|p_1|,\ldots,|p_{N-1}|,|p_N|+1\} .$$

(b) <u>If</u> $\lim\{p_{n_k}\} = p$ <u>then</u> $\lim\{p_n\} = p$ <u>also</u>.

<u>Proof</u>: $\lim\{p_{n_k}\} = p$ so, for each $\varepsilon > 0$, $\exists N_1$ and $k \geq N_1 \Rightarrow |p_{n_k} - p| < \frac{\varepsilon}{2}$. But $\{p_n\}$ is a Cauchy sequence, so $\exists N_2$ $m,n \geq N_2 \Rightarrow |p_n - p_m| < \frac{\varepsilon}{2}$. Choose $k = \sup\{N_1,N_2\}$. Then if $n \geq \sup\{N_1,N_2\}$

$$|p_n - p| = |p_n - p_{n_k} + p_{n_k} - p|$$

$$\leq |p_n - p_{n_k}| + |p_{n_k} - p|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon .$$

Thus $\lim\{p_n\} = p$, i.e. $\{p_n\}$ is convergent.

Remark: An ordered field F is complete if and only if each Cauchy sequence in F is convergent (with its limit in F).

Examples:

(1) $\{(-1)^n\}$ is divergent. $x_n = (-1)^n$, $|x_{n+1}-x_n| = 2$ for all n so $\{x_n\}$ is not a Cauchy sequence.

(2) $x_n = 1 + \frac{1}{2} + \dots + \frac{1}{n}$, $\{x_n\}$ is divergent.

$$\begin{aligned} |x_{2n}-x_n| &= \frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n} \\ &\geq \frac{1}{2n} + \frac{1}{2n} + \dots + \frac{1}{2n} \\ &= \frac{n}{2n} = \frac{1}{2} . \end{aligned}$$

Therefore $|x_{2n}-x_n| \geq \frac{1}{2}$ for all n so $\{x_n\}$ is not a Cauchy sequence.

Exercises:

2.17: Show that the following sequences are divergent by proving directly they are not Cauchy sequences.

(a) $\{n\}$ (b) $\{(-1)^n(1 - \frac{1}{n})\}$.

2.18: Show directly that the following are Cauchy sequences and hence convergent.

(a) $\{\frac{n+1}{n}\}$ (b) $\{1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!}\}$.

2.19: Determine whether each of the following sequences is convergent or divergent. In the case of convergence find the limit.

(a) $\{\frac{4n^2 + n + 4}{(1+n)(2+\frac{3}{n})}\}$ (b) $\frac{n^3 + n^2 + 1}{(n+1)^4}$

(c) $\{\frac{(n^2+1)^2}{(n+1)(n+2)(n+3)}\}$ (d) $\{n^2+n\}$

(e) $\{1-n^3\}$ (f) $\{\frac{\sin(\frac{1}{3} n\pi) + 3n}{n}\}$

(g) $\{\frac{\sin(\frac{1}{3} n\pi) + 2n^2}{n}\}$ (h) $\{\frac{n^2+1}{n^3+1} \cos^2 (\frac{1}{4} n\pi)\}$

(i) $\{\frac{1}{3} n - [\frac{1}{3} n]\}$ where $[x]$ denotes the greatest integer not exceeding x .

2.20: For what values of x are the following convergent, divergent? Wherever you can, give the limit.

(a) $\{\frac{x^n}{(n+1)(n+2)}\}$ (b) $\{(n+1)(n+2)\ x^n\}$

(c) $\{\frac{x^n + n}{x^{n+1} + n+1}\}$ (d) $\{\frac{nx^n + x^{n-1} + 1}{nx^{n-1} + 1}\}$

(e) $\{\frac{x^n}{n!}\}$ (f) $\{n!\ x^n\}$

2.21: If $a_1 = 2$, $a_{n+1} = \sqrt{6+a_n}$, show that $\{a_n\}$ is increasing and $\lim \{a_n\} = 3$.

2.22: Show that the sequence $1, 1.4, \ldots, a_n, \ldots$, in which $(2a_n+3)a_{n+1} = 4 + 3a_n$, is monotonic and that $\lim \{a_n\} = \sqrt{2}$.

2.23: If $a_1 = 1$ and $a_{n+1}(1+a_n) = 12$ show that $\lim \{a_n\} = 3$. [Hint: show $\{a_{2n+1}\}$ is increasing and $\{a_{2n}\}$ is decreasing.]

2.24: (a) If $\lim\limits_{n\to\infty} a_n = 0$, show $\lim\limits_{n\to\infty} \sigma_n = 0$ where $\sigma_n = \dfrac{a_1 + \ldots + a_n}{n}$.

(b) If $\lim\limits_{n\to\infty} a_n = a$, show $\lim\limits_{n\to\infty} \sigma_n = a$. [Hint: $\lim\limits_{n\to\infty} (a_n - a) = 0$, use (a) with a_n replaced by $a_n - a$.]

(c) Give an example to show that $\{\sigma_n\}$ may be convergent even though $\{a_n\}$ is not.

2.25: We have seen that $\lim\limits_{n\to\infty} x^{1/n} = 1$ if $x > 0$ (Exercise 2.14). An easier proof is now available to us from our results on monotone sequences. Show $\{x^{1/n}\}$ is monotone and bounded, hence convergent. Consider the subsequence $\{x^{1/2n}\}$ and deduce the result.

2.26: Let $\{x_n\}$ be a sequence of positive real numbers. Show

$$\lim_{n\to\infty} \frac{x_{n+1}}{x_n} = r \Rightarrow \lim_{n\to\infty} x_n^{1/n} = r .$$

[Hint: If $r > 0$ and $\varepsilon > 0$ then there exist positive real numbers A, B and a natural number N such that $A(r-\varepsilon)^n \le x_n \le B(r+\varepsilon)^n$ for $n \ge N$.]

2.27: By applying Exercise 2.26 to the sequence $\{\frac{n^n}{n!}\}$ show that

$$\lim \{\frac{n}{(n!)^{1/n}}\} = e \quad (e = \lim (1+\frac{1}{n})^n) \quad .$$

2.28: Let $s_n = (1+\frac{1}{n})^n$, $t_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!}$. We have seen that both sequences are convergent. Show that they have the same limit. [Hint: First use the Binomial Theorem to show $s_n = 1 + \frac{1}{1!} + \frac{1}{2!}(1-\frac{1}{n}) + \ldots + \frac{1}{n!}(1-\frac{1}{n}) \ldots (1-\frac{n-1}{n}) \leq t_n$. Next, with m fixed, and $n \geq m$, show

$$s_n \geq 1 + \frac{1}{1!} + \frac{1}{2!}(1-\frac{1}{n}) + \ldots + \frac{1}{m!}(1-\frac{1}{n}) \ldots (1-\frac{m-1}{n})$$

and deduce $\lim\limits_{n\to\infty} s_n \geq t_m$ for each m .]

2.29: Show that every sequence of real numbers has a monotone subsequence (be careful).

2.30: Let F be an ordered field. Show that the following statements are equivalent.

(a) F is complete (sets bounded above have supremums).

(b) F has the nested interval property.

(c) Each bounded monotone sequence in F converges to an element of F .

(d) Each Cauchy sequence in F converges to an element of F .

<u>Discussion</u>: All four statements in Problem 2.30 are equivalent. However (a), (b), (c) are inapplicable if one wishes to consider completeness for sets which are not ordered whereas (d) is applicable to any set X for which distance between points, i.e., a metric, has been defined. A function $\rho : X \times X \to R$ is a <u>metric</u> on X if

(i) $\rho(x,y) \geq 0$, $\forall\ x,y \in X$

(ii) $\rho(x,y) = 0 \iff x = y$

(iii) $\rho(x,y) = \rho(y,x)$, $\forall\ x,y \in X$

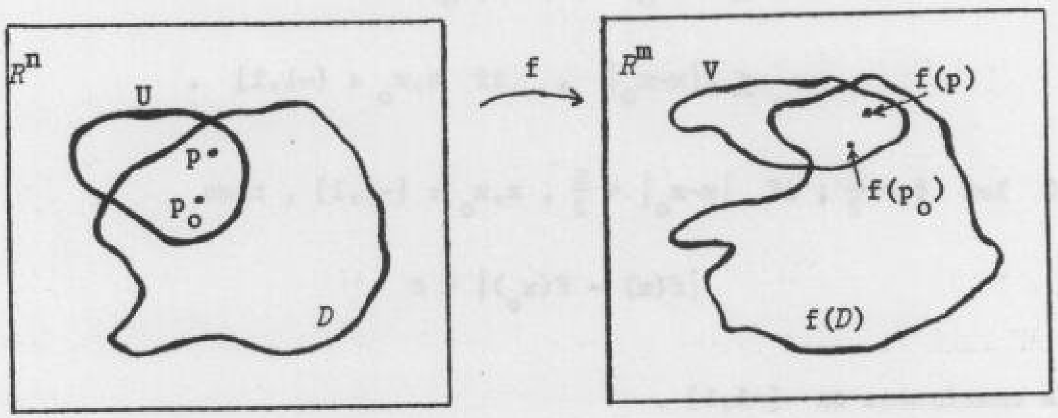
(iv) $\rho(x,y) \leq \rho(x,z) + \rho(z,y)$, $\forall\ x,y,z \in X$. (triangle inequality)

The pair $\{X,\rho\}$ is called a <u>metric space</u>. A sequence $\{x_n\}$ in X is a <u>Cauchy sequence</u> (or fundamental sequence) if, for each $\varepsilon > 0$ there exists a natural number N such that $m,n \geq N$ implies $\rho(x_n,x_m) < \varepsilon$. A metric space $\{X,\rho\}$ is called <u>complete</u> if each Cauchy sequence in $\{X,\rho\}$ is convergent (i.e. $\exists x \in X \ni \lim \rho(x_n,x) = 0$ if $\{x_n\}$ is a Cauchy sequence). For example, the space $\{R^n,\rho\}$ is a complete metric space if $\rho(p,q) = |p-q|$. It is easily seen that ρ is a metric from the norm properties (i), (ii), (iii), (iv) above; Th. 2.7 implies completeness of R^n .

CONTINUITY

Let $f : R^n \to R^m$, and let $D \subset R^n$ be the domain of f .

__Definition__: Suppose $p_o \in D$; f is continuous at p_o if, for each neighbourhood V of $f(p_o)$, there exist a neighbourhood U of p_o such that if $p \in U \cap D$ then $f(p) \in V$ (i.e. $f(U) \subset V$).



Equivalently f is continuous at p_o if, for each $\varepsilon > 0$, $\exists$ a $\delta > 0 \ni$ $p \in D$, $|p-p_o| < \delta$

$$\Rightarrow |f(p) - f(p_o)| < \varepsilon \quad .$$

Notice: In general $\delta = \delta(\varepsilon, p_o)$.

__Exercise__:

2.31: Prove that the two definitions are equivalent.

<u>Definition</u>: f is continuous on D if f is continuous at each point p_o of D.

<u>Example</u>: Let $f(x) = x^2$, $-1 \leq x \leq 1$

$$D = [-1,1] \quad , \quad f(D) = [0,1]$$

$$|f(x)-f(x_o)| = |x^2-x_o^2| = |(x-x_o)(x+x_o)| = |x-x_o||x+x_o|$$

$$\leq |x-x_o| \; (|x| + |x_o|)$$

$$\leq 2|x-x_o| \quad , \quad \text{if} \quad x,x_o \in [-1,1] \quad .$$

If $\varepsilon > 0$ let $\delta = \frac{\varepsilon}{2}$; if $|x-x_o| < \frac{\varepsilon}{2}$, $x,x_o \in [-1,1]$, then

$$|f(x) - f(x_o)| < \varepsilon$$

so f is continuous on $[-1,1]$.

<u>Exercise</u>:

2.32: Let $f(x) = \sqrt{x}$, $x \geq 0$. Show f is continuous on $[0,\infty)$.

<u>Theorem 2.8</u>: Let $f : D \to R^m$, $D \subset R^k$; f is continuous at $p_o \in D$ <=> for <u>each</u> sequence $\{p_n\}$ in D such that

$$\lim \{p_n\} = p_o$$

then

$$\lim \{f(p_n)\} = f(p_o) \quad .$$

Proof:

"=>" : f continuous at $p_o \Rightarrow \forall\, \varepsilon > 0$, $\exists\, \delta > 0 \ni$ if $p \in D$, $|p-p_o| < \delta$ then $|f(p)-f(p_o)| < \varepsilon$. Now $p_n \in D$, $\lim_{n\to\infty} p_n = p_o$.

$$\Rightarrow \exists\, N \ni \text{if } n \geq N \text{ then } |p_n - p_o| < \delta \ .$$

Hence if $n \geq N$, $|f(p_n)-f(p_o)| < \varepsilon$, i.e.

$$\lim_{n\to\infty} f(p_n) = f(p_o) \ .$$

"<=" : Suppose $\lim_{n\to\infty} f(p_n) = f(p_o)$ for each sequence $\{p_n\}$ in D such that $\lim_{n\to\infty} p_n = p_o$. Assume f is not continuous at p_o. Then, negating the definition of continuity, $\exists\, \varepsilon_o > 0 \ni$ each neighbourhood U of p_o contains a point p for which

$$|f(p) - f(p_o)| \geq \varepsilon_o \ .$$

Consider $B(p_o,\frac{1}{n})$, $n = 1,2,\ldots$; $\exists\, p_n \in B(p_o,\frac{1}{n}) \ni |f(p_n)-f(p_o)| \geq \varepsilon_o$. Hence $\lim_{n\to\infty} p_n = p_o$, but $\{f(p_n)\}$ does not converge to $f(p_o)$ contrary to our hypothesis. Hence the assumption that f is not continuous at p_o is false. □

Corollary 2.8.1:

(i) If $f,g : R^n \to R^m$ are continuous at p_o then $f+g$ and $f\cdot g$ are continuous at p_o.

(ii) If $f ; R^n \to R^m$ and $\lambda : R^n \to R$ are continuous at p_o then λf is continuous at p_o and $\frac{1}{\lambda} f$ is continuous at p_o if $\lambda(p_o) \neq 0$.

Proof: This follows immediately from the corresponding theorem for sequences (Theorem 2.4).

Corollary 2.8.2: $f : R^n \to R^m$. f is continuous at p_o if and only if each component of f is continuous at p_o .

Proof: If $f(p) = (f_1(p),\ldots,f_m(p))$ and $\lim\limits_{n\to\infty} p_n = p_o$ then $\lim\limits_{n\to\infty} f(p_n) = f(p_o) \Leftrightarrow \lim\limits_{n\to\infty} f_i(p_n) = f_i(p_o)$, $i = 1,\ldots,m$, by Theorem 2.5. □

Examples.

(1) $D = [0,1]$, $f(x) = 1$, $0 < x \leq 1$

$f(0) = 0$.

f is discontinuous at 0 ; if x_n is any sequence in $(0,1]$ such that $\lim\limits_{n\to\infty} x_n = 0$ then $f(x_n) = 1$ so $\lim\limits_{n\to\infty} f(x_n) = 1 \neq f(0)$.

(2) $D = R$, $f(x) = \sin \frac{1}{x}$, $x \neq 0$

$f(0) = 0$.

f is discontinuous at 0 since $\lim\limits_{n\to\infty} \frac{2}{(2n+1)\pi} = 0$ but $f(\frac{2}{(2n+1)\pi}) = (-1)^n$ is not convergent.

The discontinuity in Example (1) is removable, i.e. the discontinuity at 0 can be removed by changing the value of the function at 0. The discontinuity in Example (2) is essential; no matter what value is assigned to the function at 0 the discontinuity cannot be removed.

(3) $D = R^2$, $f(x,y) = \frac{xy}{x^2+y^2}$, $(x,y) \neq (0,0)$

$f(0,0) = 0$.

f is discontinuous at $(0,0)$. Consider

$$p_n = (\frac{1}{n}, \frac{1}{n}) \;;$$

$\lim_{n\to\infty} p_n = (0,0)$ and $f(p_n) = \frac{1}{n^2} / 2\frac{1}{n^2} = \frac{1}{2}$ so $\lim_{n\to\infty} f(p_n) = \frac{1}{2} \neq f(0,0)$.

The discontinuity at $(0,0)$ is essential since if $q_n = (\frac{1}{n}, \frac{1}{2n})$,

$\lim_{n\to\infty} q_n = (0,0)$ and $f(q_n) = \frac{1}{2n^2} / (\frac{1}{n^2} + \frac{1}{4n^2}) = \frac{2}{5}$ so

$$\lim_{n\to\infty} f(q_n) = \frac{2}{5} \neq \frac{1}{2} = \lim_{n\to\infty} f(p_n) .$$

(4) $D = R^2$, $f(x,y) = \frac{x^2y^2}{x^2+y^2}$, $(x,y) \neq (0,0)$

$f(0,0) = 0$.

f is continuous on R^2 .

$\underline{(x_o,y_o) \neq (0,0)}$: If $\lim_{n\to\infty} (x_n,y_n) = (x_o,y_o)$ then $\lim_{n\to\infty} x_n^2y_n^2 = x_o^2y_o^2$ and $\lim_{n\to\infty} (x_n^2 + y_n^2) = (x_o^2 + y_o^2) \neq 0$ so

$$\lim_{n\to\infty} f(x_n,y_n) = \lim_{n\to\infty} \frac{x_n^2\ y_n^2}{x_n^2+y_n^2} = \frac{x_o^2\ y_o^2}{x_o^2+y_o^2} = f(x_o,y_o) \qquad \text{(Corollary 2.8.2).}$$

$\underline{(x_o,y_o) = (0,0)}$: Corollary 2.8.2 cannot be used here since $x_o^2 + y_o^2 = 0$.

However

$$|f(x,y) - f(0,0)| = \frac{x^2y^2}{x^2+y^2} \ , \ \text{if} \quad (x,y) \neq (0,0)$$

$$\leq \frac{\frac{1}{2}(x^4 + 2x^2y^2 + y^4)}{x^2+y^2} = \frac{1}{2}(x^2+y^2)$$

$$= \frac{1}{2}|(x,y)|^2 = \frac{1}{2}|(x,y)-(0,0)|^2 \ .$$

Hence, if $|(x,y) - (0,0)| < \sqrt{2\varepsilon}$, $|f(x,y) - f(0,0)| < \varepsilon$.

Recall that if $f : R^k \to R^\ell$, $g : R^\ell \to R^m$, then $g \circ f$ is defined by $g \circ f(p) = g(f(p))$ the domain of $g \circ f$ being $\{p \in D_f : f(p) \in D_g\} = f^{-1}(D_g) \subset D_f$.

<u>Theorem 2.9</u>: Suppose $p_o \in D_{g \circ f}$ then $g \circ f$ is continuous at p_o if

(i) f is continuous at p_o and

(ii) g is continuous at $f(p_o)$.

<u>Proof</u>: If $\{p_n\}$ is a sequence in $D_{g \circ f} \subset D_f$ such that

$$\lim_{n\to\infty} p_n = p_o$$

then $\{f(p_n)\}$ is a sequence in D_g such that

$$\lim_{n\to\infty} f(p_n) = f(p_o) \quad \text{(Theorem 2.8)}$$

since f is continuous at p_o , and since g is continuous at $f(p_o)$

$$\lim_{n\to\infty} g(f(p_n)) = g(f(p_o)) \quad \text{(Theorem 2.8)}$$

i.e.

$$\lim_{n\to\infty} g\circ f(p_n) = g\circ f(p_o) \quad .$$

Thus, by Theorem 2.8, $g\circ f$ is continuous at p_o .

Corollary 2.9.1: If $f : R^n \to R^m$ is continuous at p_o then $|f| : R^n \to R$ is continuous at p_o .

Proof: If $g(q) = |q|$, $q \in R^m$ then g is continuous on R^m since

$$||q| - |q_o|| \leq |q - q_o| \quad .$$

Thus if $\varepsilon > 0$, $|q-q_o| < \varepsilon \Rightarrow ||q| - |q_o|| < \varepsilon$. Therefore $g\circ f = |f|$ is continuous at p_o since g is continuous at $f(p_o)$. □

GLOBAL PROPERTIES OF CONTINUOUS FUNCTIONS

Again recall that if $f : R^n \to R^m$ with domain $D \subset R^n$ and range $f(D) \subset R^m$ then if $A \subset R^n$, $f(A) = \{f(p) : p \in A \cap D\}$, and if $B \subset R^m$, $f^{-1}(B) = \{p : f(p) \in B\}$.

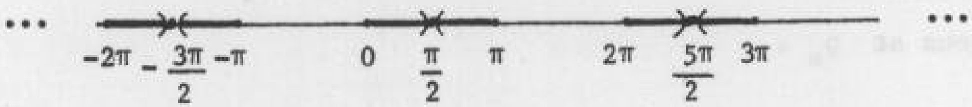
Examples:

(1) $f(x) = x^2$, $-1 \leq x \leq 1$.

$$f([0,\tfrac{1}{2}]) = [0,\tfrac{1}{4}] \ , \ f([\tfrac{1}{2},2]) = [\tfrac{1}{4},1] \ , \ f^{-1}([\tfrac{1}{4},3]) = [-1,-\tfrac{1}{2}]\cup[\tfrac{1}{2},1] \quad .$$

(2) $f(x) = \sin x \ , \ -\infty < x < \infty$.

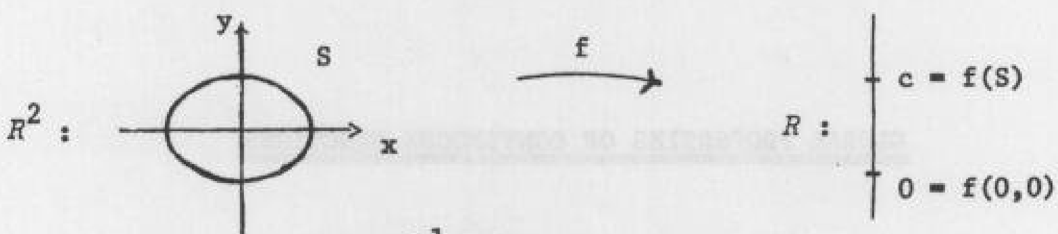
$$f^{-1}([0,1)) = \bigcup_{n=-\infty}^{\infty} \left([2n\pi,(2n+1)\pi] - \{(4n+1)\tfrac{\pi}{2}\}\right)$$



(3) $f(x,y) = \dfrac{x^2}{4} + y^2 \ , \ (x,y) \in R^2$,

$f(R^2) = \{x \in R : x \geq 0\}$,

$S = \{(x,y) : \dfrac{x^2}{4} + y^2 = c\} \ , \ c > 0$.



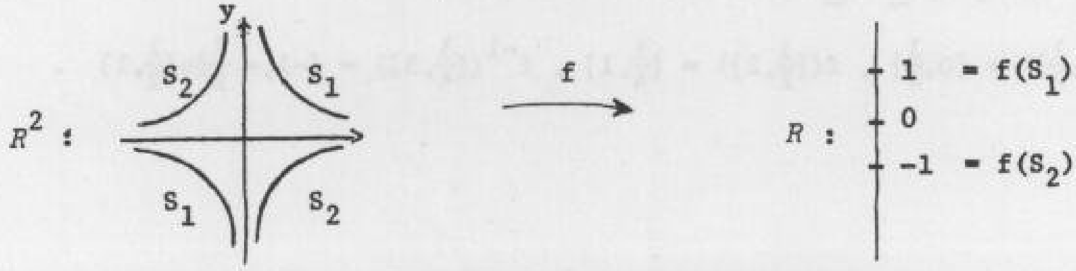
$S = f^{-1}(\{c\})$

(4) $f(x,y) = x\,y \ , \quad (x,y) \in R^2$

$f(R^2) = R$

$S_1 = \{(x,y) : xy = 1\}$

$S_2 = \{(x,y) : xy = -1\}$



$\{\text{x-axis}\} \cup \{\text{y-axis}\} = f^{-1}(\{0\})$

<u>Theorem 2.10</u>. (Global Continuity Theorem). $f : R^n \to R^m$. f is continuous on its domain <=> for each open set $V \subset R^m$ $\exists$ an open set $U \subset R^n$ such that

$$f^{-1}(V) = U \cap D .$$

<u>Proof</u>:

"=>" : Let f be continuous on D and V an open set in R^m. If $p_o \in f^{-1}(V) \subset D$ then f is continuous at p_o.

$$f(p_o) \in V .$$

Since V is open and f is continuous at p_o, $\exists\ \delta(p_o) > 0 \ni$

$$p \in B(p_o,\delta(p_o)) \cap D \Rightarrow f(p) \in V \quad ,$$

(1) $$\text{i.e. } B(p_o,\delta(p_o)) \cap D \subset f^{-1}(V) \ , \ \forall\ p_o \in f^{-1}(V) \ .$$

Hence, if $U = \bigcup_{p_o \in f^{-1}(V)} B(p_o,\delta(p_o))$, U is <u>open</u> and (1) implies

(2) $$U \cap D \subset f^{-1}(V) \ .$$

However, from the definition of U, $p_o \in f^{-1}(V) \Rightarrow p_o \in U \cap D$.

(3) $$\text{i.e. } U \cap D \supset f^{-1}(V) \ .$$

So U is an open subset of R^n and, from (2), (3),

$$U \cap D = f^{-1}(V) \ .$$

"<=" : Suppose that for each open $V \subset R^m$ $\exists$ open $U \subset R^n$ such that

$$U \cap D = f^{-1}(V) \quad .$$

Let $p_o \in D$, $\varepsilon > 0$; consider $V = B(f(p_o),\varepsilon)$, $\exists\ U \subset R^n \ni U$ open and

$$U \cap D = f^{-1}(V) \quad .$$

In particular $p_o \in U$ so U is a neighbourhood of p_o and

$$p \in U \cap D \Rightarrow f(p) \in V \quad .$$

Thus f is continuous at p_o . □

<u>Theorem 2.11</u>. (Preservation of Connectedness.): $f : R^n \to R^m$. If D is connected and f is continuous on D then $f(D)$ is connected.

<u>Proof</u>: Suppose $f(D)$ is not connected; then $\exists$ open sets V_1 and V_2 in $R^m \ni$

(i) $V_1 \cap f(D) \neq \phi$, $V_2 \cap f(D) \neq \phi$.

(ii) $[V_1 \cap f(D)] \cap [V_2 \cap f(D)] = \phi$.

(iii) $[V_1 \cap f(D)] \cup [V_2 \cap f(D)] = f(D)$.

By the Global Continuity Theorem (Theorem 2.10) there exist open sets U_1 and U_2 in R^n such that

$$U_1 \cap D = f^{-1}(V_1) \quad , \quad U_2 \cap D = f^{-1}(V_2) \quad ;$$

then

(i)' $U_1 \cap D \neq \phi$, $U_2 \cap D \neq \phi$ (from (i))

(ii)' $(U_1 \cap D) \cap (U_2 \cap D) = \phi$ (from (ii))

(iii)' $(U_1 \cap D) \cup (U_2 \cap D) = D$ (from (iii))

Thus D is disconnected contradicting our hypothesis that D is connected. Therefore our assumption is false, $f(D)$ is connected. □

Corollary 2.11.1. (Intermediate Value Theorem): $f : R^n \to R$. Let D be a connected subset of R^n and f a real valued continuous function on D . If $p, q \in D$ and $f(p) < k < f(q)$ then $\exists\ p_o \in D \ni f(p_o) = k$.

Proof: If not then $V_1 = (-\infty, k)$, $V_2 = (k, \infty)$ provide a disconnection of $f(D)$ contradicting Theorem 2.11. □

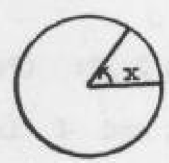
Example: At any time there are two antipodal points on the equator at the same temperature.

$T(x)$: temperature

$T(x+2\pi) = T(x)$

We suppose T is continuous

Consider $f(x) = T(x+\pi) - T(x)$



$$f(0) = T(\pi) - T(0)$$

$$f(\pi) = T(2\pi) - T(\pi) = T(0) - T(\pi) = -f(0) .$$

Case (i) : $f(0) = 0 \Rightarrow T(\pi) = T(0)$.

Case (ii) : $f(0) \neq 0$ e.g. $f(0) > 0 \Rightarrow f(\pi) = -f(0) < 0$

$$\Rightarrow \exists c \in (0,\pi) \ni f(c) = 0$$

$$\text{i.e. } T(c+\pi) = T(c) .$$

<u>Exercises</u>.

2.31: Show that every real polynomial of odd degree has at least one real root.

2.32: Show that $x^4 + 7x^3 - 9$ has at least two real roots.

2.33: Suppose f and g are continuous real valued functions on $[0,1]$ such that $f(0) < g(0)$ and $f(1) > g(1)$. Prove that there exists $x \in (0,1)$ such that $f(x) = g(x)$. Draw a picture.

<u>Theorem 2.12</u>: (Preservation of compactness.) $f : R^n \to R^m$. If D is compact and f is continuous on D then $f(D)$ is compact.

<u>Proof</u>: Let G be an open covering of $f(D)$. For each $V \in G$ there is an open set $U \subset R^n \ni$

(1) $$U \cap D = f^{-1}(V) \qquad \text{(Theorem 2.10)}$$

Since $\{V : V \in G\}$ covers $f(D)$, $\{U : U \cap D = f^{-1}(V)\ , V \in G\}$ covers D - a compact set so a finite subcover $\{U_1,\ldots,U_k\}$ exists; from (1) the corresponding V's satisfy $f(U_i) \subset V_i$ so $\{V_1,\ldots,V_k\}$ covers $f(D)$ i.e. G contains a finite subcover of $f(D)$ so $f(D)$ is compact. $\square$

<u>Exercise</u>.

2.34: Give an alternative proof by showing directly that $f(D)$ is closed and bounded if f is continuous on D and D is closed and bounded.

<u>Hint</u>: If $f(D)$ is not bounded $\exists\, p_n \in D \ni |f(p_n)| > n$, $n = 1,2,\ldots$. If $f(D)$ is not closed there is at least one cluster point q of $f(D)$, $q \notin f(D)$ i.e. $\exists\, p_n \in D \ni \lim f(p_n) = q \notin f(D)$. Show that each case leads to a contradiction of the hypothesis.

<u>Corollary 2.12.1</u>: A continuous function on a compact set is bounded.

<u>Proof</u>: $f(D)$ closed and <u>bounded</u> by the Heine-Borel Theorem. $\square$

<u>Corollary 2.12.2</u>: A continuous real-valued function on a nonempty compact set achieves its supremum and infimum, i.e. it has a maximum and a minimum value.

<u>Proof</u>: Let f be continuous on D (compact). Then $f(D)$ is a compact subset of R (Theorem 2.12). Let $M = \text{Sup}\,\{f(p) : p \in D\} = \text{Sup}\ f(D)$. We wish to show $M \in f(D)$ i.e. $M = f(p_o)$ for some p_o . Suppose $M \notin f(D)$ i.e.

$$M > f(p) \ , \ \forall\, p \in D \ .$$

Consider

$$g(p) = \frac{1}{M-f(p)} > 0 \ ;$$

g is continuous on D (Why?), D is compact so g is bounded (Corollary 2.12.1).

$$\therefore \ \exists\, A \ni 0 < g(p) \le A \ , \ \forall\, p \in D$$

$$\text{i.e.} \quad 0 < \frac{1}{M-f(p)} \le A \quad , \ \forall\, p \in D$$

$$\text{i.e.} \quad \frac{1}{A} \le M - f(p) \quad , \ \forall\, p \in D$$

$$\text{i.e.} \quad f(p) \le M - \frac{1}{A} < M \ , \ \forall\, p \in D$$

contradicting the fact that M is the least upper bound of $f(D)$. Hence we must have $M \in f(D)$. □

UNIFORM CONTINUITY

Definition: f is uniformly continuous on D if, for each $\varepsilon > 0$, $\exists\, \delta > 0 \ni$ if

$$p,q \in D \quad \text{and} \quad |p-q| < \delta$$

then

$$|f(p) - f(q)| < \varepsilon \ .$$

N.B.: δ depends on ε only.

Examples:

(1) $f(x) = x$, $-\infty < x < \infty$; $D = R$, f uniformly continuous on D . If $x,y \in R$, $|x-y| < \varepsilon$ then $|f(x) - f(y)| < \varepsilon$. f is uniformly continuous on R .

(2) $f(x) = x^2$, $0 \leq x \leq 1$; $D = [0,1]$, f uniformly continuous on D . $|f(x) - f(y)| = |x^2-y^2| = |x-y|\ |x+y| \leq 2|x-y|$ if $x,y \in [0,1]$. If $x,y \in D$, $|x-y| < \frac{\varepsilon}{2}$ then $|f(x) - f(y)| < \varepsilon$ so f is uniformly continuous on $[0,1]$.

(3) $f(x) = x^2$, $0 \leq x < \infty$; $D = [0,\infty)$, f is <u>not</u> uniformly continuous on D .

$$|f(x) - f(y)| = |x-y|\ |x+y| \quad .$$

We will show that the condition in the definition is not satisfied, for $\varepsilon = 1$, by any $\delta > 0$. Let $\delta > 0$; if $x = \frac{1}{\delta}$, $y = \frac{1}{\delta} + \frac{\delta}{2}$ then $|x-y| = \frac{\delta}{2} < \delta$ and $|f(x)-f(y)| = \frac{\delta}{2}\left(\frac{2}{\delta} + \frac{\delta}{2}\right) > 1$. Thus f is not uniformly continuous on $[0,\infty)$.

(4) $f(x) = \sqrt{x}$, $1 \leq x < \infty$; $D = [1,\infty)$. f is uniformly continuous on D . $|f(x)-f(y)| = |\sqrt{x} - \sqrt{y}| = \dfrac{|x-y|}{\sqrt{x} + \sqrt{y}} \leq \frac{1}{2}|x-y|$. If $|x-y| < 2\varepsilon$, $x,y \in [1,\infty)$ then $|f(x)-f(y)| < \varepsilon$.

(5) $f(x) = \sqrt{x}$, $0 \leq x < \infty$; $D = [0,\infty)$. f is uniformly continuous on D .

Again $|f(x) - f(y)| = \dfrac{x-y}{\sqrt{x} + \sqrt{y}}$ (we may assume without loss of generality that $x > y \geq 0$).

Suppose $|x-y| < \delta$.

<u>Case (a)</u>: $x > y > \delta$: In this case $|f(x)-f(y)| < \dfrac{\delta}{\sqrt{\delta}} = \sqrt{\delta} < 3\sqrt{\delta}$.

<u>Case (b)</u>: $0 \leq y < \delta$: Since $x > y$ and $|x-y| < \delta$ we have $0 \leq x < 2\delta$. Hence

$$|f(x)-f(y)| = \sqrt{x} - \sqrt{y} \leq \sqrt{x} + \sqrt{y} \leq \sqrt{2\delta} + \sqrt{\delta} < 3\sqrt{\delta} \quad .$$

In any case if $x \geq 0$, $y \geq 0$ and $|x-y| < \delta$ then

$$|f(x)-f(y)| < 3\sqrt{\delta} \quad .$$

If $\varepsilon > 0$, let $\delta = \dfrac{\varepsilon^2}{9}$ to see that f is uniformly continuous on $[0,\infty)$.

<u>Exercise</u>.

2.34: Show that $f(x) = \dfrac{1}{x}$ is not uniformly continuous on $(0,1]$.

<u>Theorem 2.13</u>: $f : R^n \to R^m$. If f is continuous on D and D is compact then f is uniformly continuous on D .

<u>Proof</u>: Let $\varepsilon > 0$. For each $p_o \in D$, $\exists\, \delta(\varepsilon,p_o)$

$$\ni \; p \in B(p_o,\delta(\varepsilon,p_o)) \cap D \Rightarrow |f(p) - f(p_o)| < \frac{\varepsilon}{2}$$

$$\bigcup_{p_o \in D} B(p_o\, , \tfrac{1}{2}\,\delta(\varepsilon,p_o)) \supset D$$

$$D \text{ compact} \Rightarrow \exists\, p_1,\ldots,p_k \in D \ni$$

(1) $$\bigcup_{j=1}^{k} B(p_j\,,\ \tfrac{1}{2}\,\delta\,(\varepsilon,p_j)) \supset D\ .$$

Let $\delta(\varepsilon) = \frac{1}{2}\min\{\delta(\varepsilon,p_j) : j = 1,\ldots,k\}$; if $p,q \in D$, $|p-q| < \delta\,(\varepsilon)$ then $p \in B(p_j\,,\ \frac{1}{2}\,\delta(\varepsilon,p_j))$ for some j (from (1)). Thus

(2) $$|p-p_j| < \tfrac{1}{2}\,\delta\,(\varepsilon,p_j)$$

But $|p-q| < \delta(\varepsilon) \leq \frac{1}{2}\,\delta\,(\varepsilon,p_j)$ and hence (from (2))

(3) $$|p_j-q| \leq |p-p_j| + |p-q| < \delta(\varepsilon,p_j)$$

$$(2) \Rightarrow |f(p)-f(p_j)| < \frac{\varepsilon}{2} \quad \text{and} \quad (3) \Rightarrow |f(q)-f(p_j)| < \frac{\varepsilon}{2}\ .$$

Therefore, from the triangle inequality, if $p,q \in D$, $|p-q| < \delta(\varepsilon)$, then

$$|f(p)-f(q)| \leq \varepsilon\ ,$$

so f is uniformly continuous on D . $\square$

<u>Alternate Proof</u>: Suppose f is not uniformly continuous on D . Negating the definition of uniform continuity we find $\exists\, \varepsilon_o > 0 \ni \forall\, \delta > 0\ \exists$ $p,q \in D$ with $|p-q| < \delta$ and $|f(p)-f(q)| \geq \varepsilon_o > 0$. In particular $\exists p_n,q_n \in D$, $|p_n-q_n| < \frac{1}{n}$

$$\ni |f(p_n)-f(q_n)| \geq \varepsilon_o > 0$$

$$p_n \in D \text{ (compact)} \Rightarrow \lim_{k\to\infty} p_{n_k} = p_o \in D$$

for some subsequence $\{p_{n_k}\}$ of $\{p_n\}$ (Corollary 2.2.1). Now

$$|q_{n_k}-p_o| \leq |q_{n_k}-p_{n_k}| + |p_{n_k}-p_o|$$

$$\leq \frac{1}{n_k} + |p_{n_k}-p_o| \xrightarrow[k\to\infty]{} 0$$

so that $\lim\limits_{k\to\infty} q_{n_k} = p_o \in D$ also. But since f is continuous on D (in particular at $p_o \in D$) $\lim\limits_{k\to\infty} \{f(p_{n_k})-f(q_{n_k})\} = f(p_o) - f(p_o) = 0$ (Theorem 2.8), contradicting $|f(p_n)-f(q_n)| \geq \varepsilon_o > 0$, for all n . □

<u>Exercises</u>.

2.35: If $f(x) = \frac{1}{x}$, $x \neq 0$, then f is continuous on its domain.

2.36: Show that a polynomial

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_o \quad (a_i \text{ constants})$$

is continuous on R [show that $f(x) = \text{constant}$ and $g(x) = x$ are continuous on R and deduce the general result from Corollary 2.8.1].

2.37: Let $f : R \to R$ be defined by

$$f(x) = 1-x \quad , \quad x \in Q \, ,$$
$$f(x) = x \quad , \quad x \notin Q \, .$$

Show that f is continuous at $\frac{1}{2}$ and discontinuous elsewhere.

2.38: Let $f : R \to R$ be continuous on R. Show that if $f(x) = 0$ for all $x \in Q$ then $f(x) = 0$ for all $x \in R$.

2.39: Use the inequalities $|\sin x| \leq x$, $|\cos x| \leq 1$ and the formula

$$\sin x - \sin u = 2 \sin \left(\frac{x-u}{2}\right) \cos \left(\frac{x+u}{2}\right)$$

to show that the sine function is uniformly continuous on R.

2.40: Using the results of 2.35 and 2.39 show that if

$$g(x) = x \sin \frac{1}{x} \quad , \quad x \neq 0$$

$$g(0) = 0$$

then g is continuous on R.

2.41: Is it possible for f and g to be discontinuous and yet for $g \cdot f$ to be continuous? How about $g \circ f$?

2.42: (a) Is it true that f continuous on D and D open => $f(D)$ open? [This one is easy.]

(b) Is it true that f continuous on D and D closed => $f(D)$ closed? [Consider $f(x) = \dfrac{1}{1 + x^2}$, $x \in R$.]

2.43: If $f(x) = \dfrac{1}{1 + x^2}$ then f is uniformly continuous on R.

2.44: Let f be a real valued function with domain D, an open subset of R^n. Prove that f is continuous on D if and only if the sets

$$\{p : f(p) > \alpha\} \quad , \quad \{p : f(p) < \alpha\}$$

are open for each $\alpha \in R$.

2.45: If $f(x,y) = xy + x^4 - y^4$ then the equation $f(x,y) = 0$ has at least four solutions on any circle $x^2 + y^2 = R^2 > 0$.

2.46: (a) Let f be uniformly continuous on $(0,1]$ and $\{x_n\}$ a sequence of real numbers such that $0 < x_n \leq 1$ and $\lim_{n\to\infty} x_n = 0$. Prove that $\lim_{n\to\infty} f(x_n)$ exists. [What was Cauchy's first name?]

(b) If f is uniformly continuous on $(0,1]$ then f may be defined at 0 so that f is continuous on $[0,1]$.

(c) If $f : R^n \to R^m$, domain $D \subset R^n$, is uniformly continuous on D then f may also be defined on $\overline{D}-D$ so that f is continuous on $\overline{D}$ ($\overline{D}$ denotes the closure of D, see Exercises 1.39 and 1.41).

2.47: If $f : R^n \to R^m$, domain $D \subset R^n$; the set

$$G = \{(p,f(p)) : p \in D\} \subset R^{n+m}$$

is called the <u>graph</u> of f. Suppose D is <u>connected</u>; is it true that f is continuous on $D \iff G$ is <u>connected</u>? What if "connected" is replaced by "closed"? "Compact"? Prove statements you believe to be true and find a counterexample for any false statement.

2.48: If f is real valued and f is continuous at p_o with $f(p_o) > 0$ then $f(p) > 0$ for each p in a neighbourhood of p_o [Hint: $\varepsilon = \frac{1}{2} f(p_o)$.]

2.49: Prove that f is continuous at a point p_o in its domain D if and only if, for each sequence $\{p_n\}$ of points in D such that $\lim\limits_{n\to\infty} p_n = p_o$ we have $\{f(p_n)\}$ convergent. (We say nothing about $\lim\limits_{n\to\infty} f(p_n)$.)

LIMITS

The notion of a limit which was introduced for sequences can, as you recall from first year Calculus, be extended to any function.

Definition: $f : R^n \to R^m$, domain $D \subset R^n$. Let p_o be a cluster point of D (it is not assumed that $p_o \in D$). $L \in R^m$ is the limit at p_o of f if, for each $\varepsilon > 0$, there exists a $\delta > 0$ such that $p \in D$ and $0 < |p-p_o| < \delta$

$$\Rightarrow \quad |f(p) - L| < \varepsilon \quad .$$

Write

$$\lim_{p\to p_o} f(p) = L \quad \text{or} \quad \lim_{p_o} f = L \; .$$

If no such L exists we say the limit at p_o of f does not exist.

Examples:

(1) $f(x,y) = \dfrac{xy}{\sqrt{x^2+y^2}}$, $(x,y) \neq (0,0)$; $D = R^2 - \{(0,0)\}$, $\lim\limits_{(0,0)} f = 0$

$$\left|\frac{xy}{\sqrt{x^2+y^2}}\right| \leq \frac{1}{2}\,\frac{x^2+y^2}{\sqrt{x^2+y^2}} \qquad \left\{ \begin{array}{l} \text{Recall } 2ab \leq a^2+b^2 \\ \text{since } a^2-2ab+b^2 = (a-b)^2 \geq 0 \end{array}\right.$$

$$= \frac{1}{2}\sqrt{x^2+y^2}$$

$$= \frac{1}{2}\,|(x,y)| \ .$$

If $\varepsilon > 0$ then $0 < |(x,y)| < 2\varepsilon$

$$\Rightarrow \ |f(x,y)| < \varepsilon \qquad (\text{i.e. } \delta = 2\varepsilon)$$

(2) $f(x,y) = 0$ if $y \neq x^2$

$f(x,y) = 1$ if $y = x^2$.

$\lim\limits_{(0,0)} f$ does not exist since each neighbourhood of $(0,0)$ contains points $(x,y) \neq (0,0)$ at which $f(x,y) = 0$ and $f(x,y) = 1$.

Exercise:

2.50: If $f(x) = \dfrac{x^2-a^2}{x-a}$, $x \neq a$, show $\lim\limits_{x \to a} f(x) = 2a$.

Just as in the case of sequences there is a Cauchy Criterion for the existence of $\lim_{p \to p_o} f(p)$.

<u>Theorem 2.14</u> (Cauchy Criterion).

$$\lim_{p \to p_o} f(p) \ \exists \iff \forall\ \varepsilon > 0 \ , \exists\ \delta > 0 \ni$$

if $p,q \in D$ and $0 < |p-p_o| < \delta$, $0 < |q-p_o| < \delta$ then $|f(p)-f(q)| < \varepsilon$.

<u>Proof:</u>

"=>" Suppose $\lim_{p \to p_o} f(p) = L$ i.e. if $\varepsilon > 0\ \exists\ \delta$

$$\ni p \in D \ , \ 0 < |p-p_o| < \delta \Rightarrow |f(p)-L| < \frac{\varepsilon}{2} \ .$$

Then if $p,q \in D$ and $0 < |p-p_o| < \delta$, $0 < |q-p_o| < \delta$ then

$$|f(p)-f(q)| \leq |f(p)-L| + |f(q)-L|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \ ,$$

so Cauchy's Criterion is satisfied.

"<=" Let Cauchy's Criterion be satified i.e. if $\varepsilon > 0$

$$\exists\ \delta > 0 \ni p,q \in D \ , \ 0 < |p-p_o| < \delta \ , \ 0 < |q-p_o| < \delta = \delta(\varepsilon)$$

$$\Rightarrow |f(p) - f(q)| < \varepsilon \ .$$

Let $\{p_n\}$ be any sequence in D for which $\lim_{n \to \infty} p_n = p_o$, $p_n \neq p_o$, $n = 1,2,\ldots$

$$\exists\ N \ni n \geq N \Rightarrow 0 < |p_n - p_o| < \delta(\varepsilon)$$

$$\therefore\ m,n \geq N \Rightarrow p_n, p_m \in D\ ,\ 0 < |p_n - p_o| < \delta \text{ and } 0 < |p_m - p_o| < \delta$$

$$\Rightarrow |f(p_n) - f(p_m)| < \varepsilon$$

$$\Rightarrow \{f(p_n)\} \text{ is a Cauchy Sequence}$$

$$\Rightarrow \lim_{n\to\infty} f(p_n) = L \text{ (say) exists.}$$

We now show that $\lim\limits_{p\to p_o} f(p) = L$. Since $\lim\limits_{n\to\infty} p_n = p_o$ and $\lim\limits_{n\to\infty} f(p_n) = L$ it follows that $\exists\ n_1 \ni 0 < |p_{n_1} - p_o| < \delta(\frac{\varepsilon}{2})$ and $|f(p_{n_1}) - L| < \frac{\varepsilon}{2}$. Therefore if $p \in D$ and $0 < |p - p_o| < \delta(\frac{\varepsilon}{2})$

$$|f(p) - L| \leq |f(p) - f(p_{n_1})| + |f(p_{n_1}) - L|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

$$\text{i.e. } \lim_{p\to p_o} f(p) = L\ . \qquad \square$$

In fact general limits are reduced to consideration of limits of sequences by the following result

<u>Theorem 2.15</u>: $\lim\limits_{p\to p_o} f(p)$ exists and equals $L \iff \lim\limits_{n\to\infty} f(p_n)$ exists and equals L for each sequence $\{p_n\}$ in D such that $\lim\limits_{n\to\infty} p_n = p_o$, $p_n \neq p_o$.

<u>Proof</u>: <u>Exercise</u> 2.51:

A point $p_o \in D$ is called <u>isolated</u> if it is not a cluster point of D. Check back to the definition of continuity and observe that a function is always continuous at an isolated point of its domain.

<u>Corollary 2.15.1</u>: f is continuous at a cluster point of its domain

$$\Leftrightarrow \lim_{p \to p_o} f(p) = f(p_o) \ .$$

<u>Proof</u>: Theorems 2.8, 2.15.

The following Corollaries follow immediately from the corresponding statements for sequences.

<u>Corollary 2.15.2</u>: $f : R^n \to R^m$, $f = (f_1, \ldots, f_m)$

$$L = (L_1, \ldots, L_m)$$

$$\lim_{p_o} f = L \ \Leftrightarrow \ \lim_{p_o} f_i = L_i \quad , \quad i = 1, \ldots, m \ .$$

<u>Corollary 2.15.3:</u> $f, g : R^n \to R$

$$\lim_{p_o} f = L \quad , \quad \lim_{p_o} g = M$$

(i) $\lim\limits_{p_o} (f+g) = L + M$

(ii) $\lim\limits_{p_o} fg = LM$

(iii) $\lim\limits_{p_o} \frac{f}{g} = \frac{L}{M}$ (if $M \neq 0$) .

If p_o is a cluster point of $D_o \subset D$ we may talk about the limit with respect to D_o at p_o, i.e. $(D_o) \lim\limits_{p \to p_o} f(p) = L$ if, for each $\varepsilon > 0$ $\exists\, \delta > 0 \ni$ if $p \in D_o$ and $0 < |p-p_o| < \delta$ then $|f(p)-L| < \varepsilon$.

Special Notation:

$$D_o = (x_o,\infty) : (D_o) \lim_{x \to x_o} f(x) = \lim_{x \to x_o^+} f(x)$$

$$D_o = (-\infty,x_o) : (D_o) \lim_{x \to x_o} f(x) = \lim_{x \to x_o^-} f(x) \ .$$

Examples.

(1) $f(x) = 1 \ , \ x > 0$

$f(x) = x \ , \ x \leq 0$

$$\lim_{x \to 0^+} f(x) = 1 \quad , \quad \lim_{x \to 0^-} f(x) = 0 \ .$$

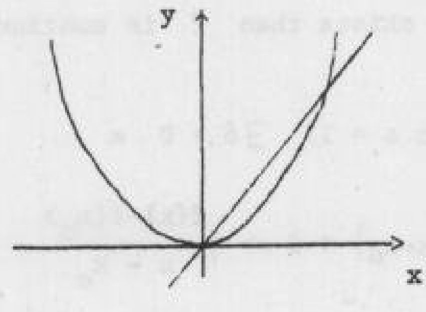
(2) $f(x,y) = 0 \ , \ y \neq x^2$

$f(x,y) = 1 \ , \ y = x^2$.

Evidently the limit at $(0,0)$ with respect to points on the parabola $y = x^2$ is 1 while the limit with respect to the complement of the parabola is 0. Notice that in this case even though the limit with respect to any ray through $(0,0)$ exists and is 0,

$$\text{i.e.} \quad \lim_{t \to 0} f(tx_o,ty_o) = 0 \quad \text{if} \quad (x_o,y_o) \neq (0,0) ,$$

$$\lim_{(x,y)\to(0,0)} f(x,y) \text{ does not exist.}$$



The moral of this is that when limit properties of functions of more than one variable are being considered it is not enough to look at their behaviour on straight lines.

Exercise:

2.52: Show that $\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4}$ does not exist even though $0 =$

$(D_o) \lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4}$ for any straight line D_o through $(0,0)$.

DIFFERENTIATION OF REAL VALUED FUNCTIONS OF A REAL VARIABLE

We review the most important results on differentiation from first year Calculus. Throughout this section $f : R \to R$.

Definition: If f is defined in a neighbourhood of x_o and

$$\lim_{x\to x_o} \frac{f(x) - f(x_o)}{x - x_o}$$

exists this limit is called the <u>derivative</u> of f at x_o and is denoted $f'(x_o)$.

<u>Proposition</u>: If $f'(x_o)$ exists then f is continuous at x_o .

<u>Proof</u>: $f'(x_o)$ $\exists$ so (with $\varepsilon = 1$) $\exists \delta > 0$ $\ni$

$$0 < |x-x_o| < \delta \Rightarrow \left|\frac{f(x)-f(x_o)}{x - x_o} - f'(x_o)\right| < 1$$

$$\Rightarrow |f(x) - f(x_o)| < \{|f'(x_o)| + 1\} |x-x_o|$$

$$\Rightarrow \lim_{x \to x_o} f(x) = f(x_o) \quad .$$

<u>Exercise</u>:

2.53: Recall the rules for differentiation of constants, sums, products, composites (chain rule) which you have learned. Prove one or two of them.

<u>Definition</u>: f has an <u>interior</u> <u>relative</u> <u>maximum</u> (<u>minimum</u>) at c if there is a neighbourhood U of c such that

$$f(x) \leq f(c) \qquad (f(x) \geq f(c))$$

for each $x \in U$.

<u>Theorem 2.16</u>: If f has an interior relative <u>maximum</u> (<u>minimum</u>) at c and $f'(c)$ exists then $f'(c) = 0$.

Proof:

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \quad \exists \quad .$$

Suppose f has an interior relative maximum at c . $\exists\, \delta > 0 \ni$ if $|x-c| < \delta$ then $f(x) \leq f(c)$

$$\therefore \quad c < x < c+\delta \Rightarrow \frac{f(x) - f(c)}{x - c} \leq 0$$

(1) $$\Rightarrow \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c) \leq 0 \quad \text{(Why?)} \quad .$$

Similarly, $c-\delta < x < c \Rightarrow \dfrac{f(x) - f(c)}{x - c} \geq 0$

(2) $$\Rightarrow \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'(c) \geq 0$$

(1) and (2) $\Rightarrow$ $f'(c) = 0$. □

Exercise.

2.54: Draw the graph of the function $f(x) = |\sin x|$, $x \in R$. Check that the derivative exists and is 0 at each relative maximum but does not exist at each relative minimum.

Corollary 2.16.1. (Rolle's Theorem): Suppose

(i) f is continuous on $[a,b]$,

(ii) f' exists on (a,b) ,

(iii) $f(a) = f(b)$.

Then $\exists\ c \in (a,b) \ni f'(c) = 0$.

Proof:

(a) If $f(x) = f(a) = f(b)$, $\forall\ x \in (a,b)$ then $f'(x) = 0$, $\forall\ x \in (a,b)$.

(b) If $f(x_1) > f(a) = f(b)$ for some $x_1 \in (a,b)$ then since f is continuous on $[a,b]$ (compact) f achieves the value $m = \sup\{f(x) : x \in [a,b]\}$ at some point $c \in (a,b)$ (Corollary 2.12.2); thus f has an interior relative maximum at c so $f'(c) = 0$.

(c) If $f(x_1) < f(a) = f(b)$ for some $x_1 \in (a,b)$ then f has an interior relative minimum at some point c and $f'(c) = 0$. □

Corollary 2.16.2: (Mean Value Theorem): Suppose

(i) f is continuous on $[a,b]$,

(ii) f' exists on (a,b) .

Then $\exists\ c \in (a,b) \ni f'(c)(b-a) = f(b) - f(a)$.

Proof: Consider $\phi(x) = [f(x)-f(a)](b-a) - [f(b)-f(a)](x-a)$; ϕ is continuous on $[a,b]$, ϕ' exists on (a,b) and $\phi(a) = \phi(b) = 0$. By Rolle's Theorem $\exists\ c \in (a,b) \ni$

$$0 = \phi'(c) = f'(c)(b-a) - [f(b)-f(a)] \ . \qquad \square$$

Corollary 2.16.3: (Cauchy Mean Value Theorem): Suppose

(i) f,g are continuous on $[a,b]$,

(ii) f',g' exist on a,b .

Then $\exists\, c \in (a,b) \ni f'(c)\,[g(b)-g(a)] = g'(c)\,[f(b)-f(a)]$.

Proof: Consider $\phi(x) = [f(x)-f(a)][g(b)-g(a)] - [f(b)-f(a)][g(x)-g(a)]$. □

Application (L'Hospital's Rule): If f,g are differentiable on (a,b) and $g'(x) \neq 0 \;\forall\, x \in (a,b)$ and either

(i) $\lim\limits_{x\to b^-} f(x) = 0$, $\lim\limits_{x\to b^-} g(x) = 0$,

or

(ii) $\lim\limits_{x\to b^-} f(x) = \infty$, $\lim\limits_{x\to b^-} g(x) = \infty$,

then

$$\lim_{x\to b^-} \frac{f'(x)}{g'(x)} = L \Rightarrow \lim_{x\to b^-} \frac{f(x)}{g(x)} = L \quad .$$

[Note: $\lim\limits_{x\to b^-} f(x) = \infty$ means that, for each real number N , $\exists\, \delta > 0$ and $x \in (b-\delta,b) \Rightarrow f(x) > N$.]

Proof:

Case (i). Suppose $\lim\limits_{x\to b^-} f(x) = \lim\limits_{x\to b^-} g(x) = 0$ and $\lim\limits_{x\to b^-} \frac{f'(x)}{g'(x)} = L$. If $\varepsilon > 0$, $\exists\, \delta > 0 \ni$ if $x \in (b-\delta,b)$ then

$$\left|\frac{f'(x)}{g'(x)} - L\right| < \varepsilon \quad . \tag{1}$$

Let f,g be defined at b by $f(b) = g(b) = 0$; then f and g are continuous on $[x,b]$ if $x > a$. From the Cauchy Mean Value Theorem

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(b)}{g(x) - g(b)} = \frac{f'(c)}{g'(c)}\ , \quad b-\delta < x < c < b\ .$$

Therefore

$$\left|\frac{f(x)}{g(x)} - L\right| = \left|\frac{f'(c)}{g'(c)} - L\right| < \varepsilon\ , \quad \text{if} \quad x \in (b-\delta,b)\ ;$$

thus

$$\lim_{x\to b^-} \frac{f(x)}{g(x)} = L\ .$$

<u>Case (ii)</u>. Suppose $\lim\limits_{x\to b^-} f(x) = \lim\limits_{x\to b^-} g(x) = \infty$ and $\lim\limits_{x\to b^-} \frac{f'(x)}{g'(x)} = L$.

As before, (1) holds if $x \in (b-\delta,b)$ for some $\delta > 0$. However we cannot define $f(b)$, $g(b)$ in this case so that f is continuous at b . Let $x_o = b-\delta$, $x \in (b-\delta,b)$;

$$\frac{f(x) - f(x_o)}{g(x) - g(x_o)} = \frac{f'(c)}{g'(c)}\ , \quad \text{some } c\ ,\ b-\delta < c < x < b\ .$$

Thus, from (1),

$$\left|\frac{f(x) - f(x_o)}{g(x) - g(x_o)} - L\right| < \varepsilon$$

$$\text{i.e.} \quad \left|\frac{f(x)}{g(x)}\ \frac{1 - f(x_o)/f(x)}{1 - g(x_o)/g(x)} - L\right| < \varepsilon$$

(2) i.e. $\left|\frac{f(x)}{g(x)}\, h(x) - L\right| < \varepsilon$, where $h(x) = \dfrac{1 - f(x_o)/f(x)}{1 - g(x_o)/g(x)}$.

Notice that $\lim\limits_{x \to b^-} h(x) = 1$ so $\exists\ \delta_1 > \delta$ such that

$$(3) \qquad x \in (b-\delta_1, b) \Rightarrow |h(x)-1| < \varepsilon \text{ and } h(x) > \frac{1}{2} .$$

Thus, if $x \in (b-\delta_1, b)$,

$$\begin{aligned} \frac{1}{2} \left|\frac{f(x)}{g(x)} - L\right| &< \left|h(x)\left\{\frac{f(x)}{g(x)} - L\right\}\right| \quad , \text{ from (3)} \\ &= \left|h(x)\frac{f(x)}{g(x)} - L + L - h(x)L\right| \\ &\le \left|h(x)\frac{f(x)}{g(x)} - L\right| + |L|\ |1 - h(x)| \\ &< \varepsilon + |L|\ \varepsilon \qquad , \text{ from (2) and (3).} \end{aligned}$$

Thus, if $\varepsilon > 0$, $\exists\ \delta_1 \ni$ if $x \in (b-\delta_1, b)$ then

$$\left|\frac{f(x)}{g(x)} - L\right| < 2(1 + |L|)\varepsilon \quad .$$

Therefore

$$\lim_{x \to b^-} \frac{f(x)}{g(x)} = L \quad . \qquad \Box$$

<u>Remark</u>: Define $\lim\limits_{x\to\infty} f(x) = L$ if , for each $\varepsilon > 0$, $\exists\ N \ni x \ge N \Rightarrow |f(x) - L| < \varepsilon$. L'Hospital's Rule is also valid if "$\lim\limits_{x\to b^-}$" is replaced by "$\lim\limits_{x\to\infty}$". Minor changes are required in the proofs however.

<u>Exercises</u>: [In these exercises f' denotes the derivative of f when it exists, $f^{(2)}$ or f'' denotes the derivative of f' and, inductively, $f^{(n)}$ the derivative of $f^{(n-1)}$. We will also write $f^{(n)}(x) = \frac{d^n}{dx^n} f(x)$.]

2.55: If $f'(x) = 0$ for each $x \in (a,b)$ then f is constant on (a,b).

2.56: If $|f'(x)| \leq M$ for each $x \in (a,b)$ then

$$|f(x) - f(y)| \leq M\ |x-y|$$

for all $x,y \in (a,b)$ and f is uniformly continuous on (a,b).

2.57: Let f be Lipschitzian of type α on $[a,b]$ i.e. there are positive constants α and M such that

$$|f(x)-f(y)| \leq M\ |x-y|^{\alpha} \quad \text{for each} \quad x,y \in [a,b]\ .$$

(i) Prove that f is uniformly continuous on $[a,b]$.

(ii) Prove that f is a constant if $\alpha > 1$.

2.58: If $h = fg$ prove

(i) $h' = f'g + fg'$,

(ii) $h^{(n)} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} f^{(n-k)}\ g^{(k)}$ (Leibniz's Rule).

Assume all the differentiablity you need $(0! = 1)$.

2.59: Find a point $c \in (a,b)$ such that $f'(c) = 0$ where

$$f(x) = (x-a)(b-x) \quad , \quad x \in [a,b]\ .$$

2.60: Prove that there is no real number k such that the equation

$$x^3 - 3x + k = 0$$

has two distinct solutions in $[0,1]$.

2.61: If $c_o,\ldots,c_n$ are real numbers such that

$$c_o + \frac{c_1}{2} + \ldots + \frac{c_{n-1}}{n} + \frac{c_n}{n+1} = 0$$

prove that

$$c_o + c_1 x + \ldots + c_{n-1} x^{n-1} + c_n x^n = 0$$

has at least one solution in $(0,1)$.

2.62: If $f(x) = \sqrt{|x|}$ then $f'(x)$ exists if $x \neq 0$ and does not exist if $x = 0$.

2.63: Prove that $10.243 < \sqrt{105} < 10.250$.

[Hint: By the Mean Value Theorem $\sqrt{105} = \sqrt{100} + \frac{1}{2\sqrt{c}} 5$ where $100 < c < 105$.]

2.64: Prove l'Hospital's Rule with "$\lim\limits_{x \to b^-}$" replaced by "$\lim\limits_{x \to \infty}$". Prove that $\lim\limits_{x \to \infty} x e^{-x} = 0$.

2.65: Prove

(i) $\lim_{x \to 0} \dfrac{\text{arc tan } x}{x} = 1$

(ii) $\lim_{x \to 0} \dfrac{e^x - e^{-x} - 2 \tan x}{\sin^3 x} = -\dfrac{1}{3}$

(iii) $\lim_{x \to 2} \dfrac{x^4 - 5x^3 + 6x^2 + 4x - 8}{x^4 - 7x^3 + 18x^2 - 20x + 8} = 3$

(iv) $\lim_{x \to 0} (\cot x - \frac{1}{x}) = 0$

(v) $\lim_{x \to 0} (1 + \tan x)^{\operatorname{cosec} x} = e$

(vi) $\lim_{x \to \frac{\pi}{2}+} \dfrac{\log (x - \frac{\pi}{2})}{\tan x} = 0$.

2.66: If $f'(x)$ exists for all x near c, f is continuous at c and

$$\lim_{x \to c} f'(x) = A$$

then $f'(c)$ exists and equals A .

2.67: (Darboux Property of the Derivative) Let $f'(x)$ exist for each $x \in [a,b]$ and $f'(a) = \alpha$, $f'(b) = \beta$. Suppose $\alpha < \gamma < \beta$ show that $\exists\, c \in (a,b) \ni f'(c) = \gamma$.

[Hint: Show that $g(x) = f(x) - \gamma x$ must achieve its minimum in (a,b).] This exercise shows that derivatives, like continuous functions, have the Intermediate Value Property! However, a derivative need not be continuous on its domain, see Exercise 2.71 (i) and (iv).

2.68: Let $f(x) = 0$, $-1 \leq x \leq 0$

$f(x) = 1$, $0 < x \leq 1$.

Is there a function g such that $g'(x) = f(x)$, $-1 \leq x \leq 1$?

2.69: The function defined by

$$g(x) = x^2 \ , \quad x \in Q$$
$$g(x) = 0 \ , \quad x \notin Q$$

is continuous at exactly one point. Is it differentiable there?

2.70: (Taylor's Theorem) Suppose that

(i) $f^{(n-1)}$ exists and is continuous on $[\alpha, \beta]$,

(ii) $f^{(n)}$ exists on (α, β) .

(a) Show that, if $m > 0$,

$$f(\beta) = f(\alpha) + \frac{(\beta-\alpha)}{1!} f'(\alpha) + \dots + \frac{(\beta-\alpha)^{n-1}}{(n-1)!} f^{(n-1)}(\alpha) + R_n(f;\alpha,\beta)$$

where

$$R_n(f;\alpha,\beta) = \frac{(\beta-\alpha)^m (\beta-\gamma)^{n-m}}{m(n-1)!} f^{(n)}(\gamma)$$

for some point $\gamma \in (\alpha,\beta)$. This is called Schlömilch's form of the remainder. Special cases are given by taking

$$m = n \ : \ R_n(f;\alpha,\beta) = \frac{(\beta-\alpha)^n}{n!} f^{(n)}(\gamma) \qquad \text{Lagrange form}$$

$$m = 1 \ : \ R_n(f;\alpha,\beta) = \frac{(\beta-\alpha)(\beta-\gamma)^{n-1}}{(n-1)!} f^{(n)}(\gamma) \qquad \text{Cauchy form}$$

The Lagrange form is the easiest to remember and is adequate for most purposes. [Hint: Let the constant C be defined by

$$f(\beta) - f(\alpha) - \frac{(\beta-\alpha)}{1!} f'(\alpha) - \dots - \frac{(\beta-\alpha)^{n-1}}{(n-1)!} f^{(n-1)}(\alpha) - (\beta-\alpha)^m C = 0$$

and apply Rolle's Theorem to

$$\phi(x) = f(\beta) - f(x) - \frac{(\beta-x)}{1!} f'(x) - \dots - \frac{(\beta-x)^{n-1}}{(n-1)!} f^{(n-1)}(x) - (\beta-x)^m C .]$$

(b) (i) If $f(x) = e^x$ show $\lim_{n\to\infty} R_n(f; 0,x) = 0$ for all x.

(i.e. $\lim_{n\to\infty} (1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^{n-1}}{(n-1)!}) = e^x$ for all x.)

(ii) If $f(x) = \sin x$ then $\lim_{n\to\infty} R_n(f;0,x) = 0$ for all x.

(iii) If $f(x) = \frac{1}{1-x}$, $x \neq 1$, then $\lim_{n\to\infty} R_n(f;0,x) = 0$, $-1 < x < 1$.

(c) (i) How large must I take n to approximate e to four decimal places by the expression

$$1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{(n-1)!} \,?$$

(ii) Approximate $\sqrt[3]{e}$ to five decimal places.

(iii) Prove that the closest integer to $\frac{n!}{e}$ is divisible by $(n-1)$. [Hint: Use the Taylor expansion for e^{-1}.]

(iv) Compute $\sqrt{97}$ to four decimals.

(d) Suppose f'' exists and is non-negative (non-positive) in a neighbourhood of c and $f'(c) = 0$ then f has a relative minimum (maximum) at c.

(e) Suppose f'' exists in a neighbourhood of c and is continuous at c then f has a relative minimum (maximum) at c if

$$f'(c) = 0 \quad \text{and} \quad f''(c) > 0 \quad (< 0) .$$

2.71: Let $f_o(x) = \sin\frac{1}{x}$, $f_1(x) = x \sin\frac{1}{x}$, $f_2(x) = x^2 \sin\frac{1}{x}$, $f_3(x) = x^3 \sin\frac{1}{x}$, $x \neq 0$ and $f_i(0) = 0$, $i = 0,1,2,3$.

(i) f_i are differentiable at any point $x \neq 0$ (Chain Rule)

(ii) f_o is discontinuous at 0

(iii) f_1 is continuous at 0 but is not differentiable at 0

(iv) f_2 is differentiable at 0, f_2' is discontinuous at 0

(v) f_3 is differentiable at 0, f_3' is continuous at 0.

2.72: Let p_n be defined by

$$p_n(x) = \frac{d^n}{dx^n}(x^2-1)^n \quad , \quad n = 1,2,\ldots$$

(i) Show that p_n is a polynomial of degree n.

(ii) The equation $p_n(x) = 0$ has exactly n roots in $(-1,1)$.

2.73: Doesn't time fly when you're having fun?

<u>References for Chapter II</u>

Bartle: Chapters II, III, IV, V

Buck: Chapters 2, 3.

CHAPTER THREE

RIEMANN INTEGRATION

The definition of the Riemann integral is essentially the same in higher dimensions as in one dimension. You may have learned Riemann integration of functions of one variable by the lower and upper Riemann sums approach; the treatment adopted here is slightly different but equivalent to that approach (see Exercise 3.17).

Closed Interval in R^n :

$$I = [a_1,b_1] \times \ldots \times [a_n,b_n] = \{(x_1,\ldots,x_n) : a_i \leq x_i \leq b_i \ ,\ i = 1,\ldots,n\} \ .$$

Diameter of I :

$$\lambda(I) = [(b_1-a_1)^2 + \ldots + (b_n-a_n)^2]^{1/2} = \sqrt{\sum_{k=1}^{n} (b_k-a_k)^2} \ .$$

If $p,q \in I$ then $|p-q| \leq \lambda(I)$.

Content of I. (Jordan measure of I):

$$\mu(I) = (b_1-a_1)(b_2-a_2)\ldots(b_n-a_n) = \prod_{k=1}^{n} (b_k-a_k) \ .$$

In R^1 , R^2 , R^3 μ is length, area and volume respectively. If you wish to emphasize the dimension of the space in which you are working write μ_n instead of μ .

<u>Content Zero</u>: A set $S \subset R^n$ has <u>content zero</u> (Jordan measure zero) if, for each $\varepsilon > 0$, there exists a <u>finite</u> collection of intervals $\{I_i : i=1,\dots,k\}$ such that

$$S \subset \bigcup_{i=1}^{k} I_i$$

and

$$\sum_{i=1}^{k} \mu(I_i) \le \varepsilon \quad .$$

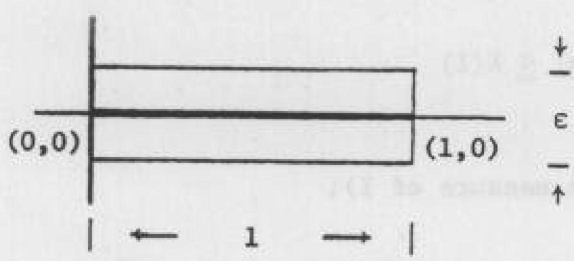
For such a set S we will simply write $\mu(S) = 0$. $\mu(S)$ is now defined if S is an interval or if S has zero content.

<u>Examples</u>.

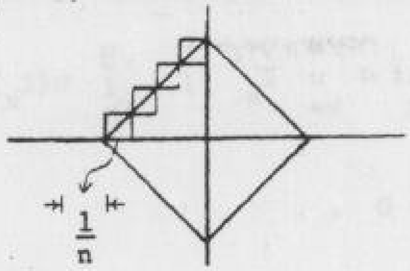
(1) A set containing a single point only has content zero in R^n.

(2) A set containing a finite number of points has content zero in R^n.

(3) $\{(x,0) : 0 \le x \le 1\}$ has content zero in R^2.



(4) $\{(x,y) : |x| + |y| = 1\}$ has content zero in R^2 .



$\sum \mu(I_i) = \frac{4}{n} < \varepsilon$, if n sufficiently large. [(4) also follows from (5) and (6).]

(5) If f is a continuous real valued function on $[0,1]$ then $\{(x,f(x)) : 0 \leq x \leq 1\}$, the <u>graph</u> of f , has content zero in R^2 . Let $\varepsilon > 0$; f is uniformly continuous on $[0,1]$ (Theorem 2.13) so $\exists \delta < 1 \ni x,y \in [0,1]$, $|x-y| \leq \delta \Rightarrow |f(x)-f(y)| \leq \frac{\varepsilon}{4}$. Now let m be that natural number such that

$$m\,\delta < 1 \quad , \quad (m+1)\delta \geq 1 \quad .$$

Consider the following intervals in R^2 :

$$I_k = [k\delta,(k+1)\delta]\times[f(k\delta) - \frac{\varepsilon}{4} , f(k\delta) + \frac{\varepsilon}{4}] \text{ , } k = 0,\ldots,m-1 \text{ .}$$

$$I_m = [m\delta,1] \times [f(m\delta)- \frac{\varepsilon}{4}, f(m\delta) \ \frac{\varepsilon}{4}]$$

If $x \in [k\delta,(k+1)\delta] \cap [0,1]$ then $|x-k\delta| \leq \delta$ so

$$|f(x)-f(k\delta)| \leq \frac{\varepsilon}{4}$$

$$\text{i.e.} \quad f(x) \in [f(k\delta) - \frac{\varepsilon}{4} , f(k\ \delta) + \frac{\varepsilon}{4}] \quad ;$$

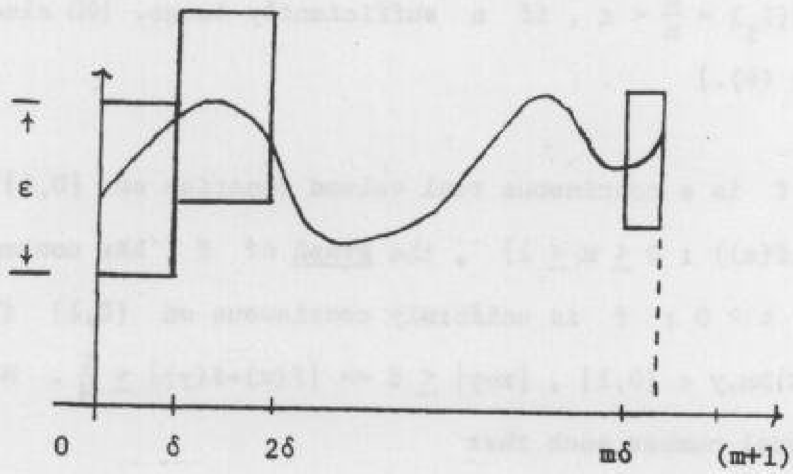
thus

$$(x,f(x)) \in I_k \quad \text{if} \quad x \in [k\delta,(k+1)\delta] \cap [0,1] \quad .$$

Hence

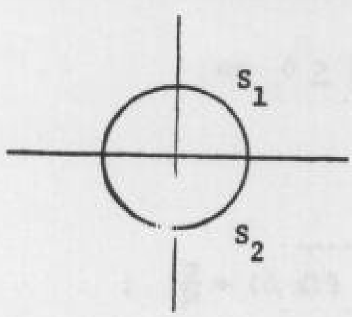
$$\{(x,f(x)) : 0 \le x \le 1\} \subset \bigcup_{k=}^{m} I_k \quad ; \quad \sum_{k=}^{m} \mu(I_k) \simeq \frac{\varepsilon}{2}(m+1)\delta < \frac{\varepsilon}{2}(1+\delta) < \varepsilon$$

so $\mu(\{(x,f(x) : 0 \le x \le 1\} = 0$.



(6) The union of any finite collection of sets with zero content has zero content.

(7) A circle has zero content in R^2



$C = S_1 \cup S_2$

S_i each has zero content by (5).

<u>Partition of I</u>:

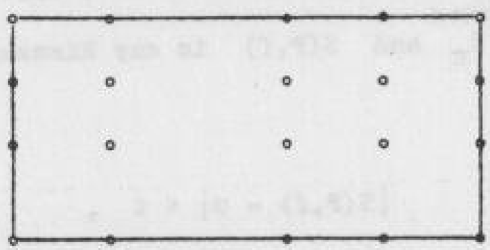
$$I = [a_1,b_1] \times [a_2,b_2] \times \ldots \times [a_n,b_n] \quad .$$

Let P_i be finite sets of real numbers

$$\{t_{ij} : j = 0,\ldots,m_i\} \quad , \quad i = 1,\ldots,n \quad ,$$

such that $a_i = t_{i0} < t_{i1} < \ldots < t_{im_i} = b_i$; then $P = P_1 \times P_2 \times \ldots \times P_n$ is said to be a <u>partition</u> of I .

I :

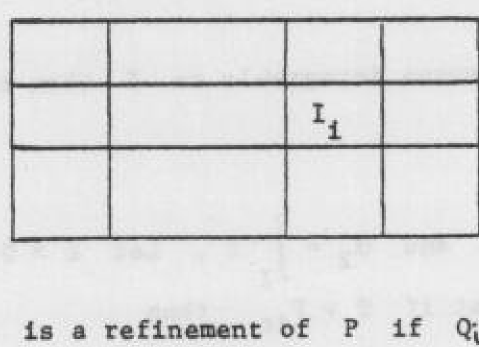


Notice that a partition of I generates a subdivision of I into a finite collection of closed nonoverlapping subintervals $\{I_i\}$ of the form

$$[t_{j1},t_{j+1,1}] \times [t_{k2},t_{k+1,2}] \times \ldots \times [t_{\ell n},t_{\ell+1,n}] \quad .$$

There are $m_1\, m_2 \ldots m_n$ such subintervals generated by P .

I :



A partition Q of I is a refinement of P if $Q_i \supset P_i$. Notice that a refinement of P further subdivides the intervals I_i generated by P .

Riemann Sums: Let I be a closed interval in R^n and $f : I \to R^m$. If P is a partition of I which generates a subdivision of I into subintervals I_i then

$$S(P,f) = \sum_i f(p_i) \; \mu(I_i) \quad , \quad p_i \in I_i$$

is a Riemann sum of f corresponding to the partition P.

Definition (Riemann Integral). Let $f : I \to R^m$ where I is a closed interval in R^n. Let $\alpha \in R^m$. If, for each $\varepsilon > 0$, there exists a partition P_ε of I such that if a partition $P \supset P_\varepsilon$ and $S(P,f)$ is any Riemann sum corresponding to P then

$$|S(P,f) - \alpha| < \varepsilon \quad ,$$

f is said to be Riemann integrable on I and to have Riemann integral α.

Write:

$$\alpha = \int_I f \quad ,$$

or

$$\alpha = \int_I f \, d\mu \quad .$$

Proposition: If f is Riemann integrable on I then the integral of f on I is unique.

Proof: Suppose $\alpha_1 = \int_I f$ and $\alpha_2 = \int_I f$. Let $\varepsilon > 0$. $\exists$ partitions $P_{1\varepsilon}$, $P_{2\varepsilon}$ of I such that if $P \supset P_{i\varepsilon}$ then

$$|S(P,f) - \alpha_i| < \varepsilon \quad , \; i = 1,2 \; .$$

But $P_1 \cup P_2 \supset P_1$ and $P_1 \cup P_2 \supset P_2$ so

$$|S(P_1 \cup P_2, f) - \alpha_i| < \varepsilon \;, \quad i = 1,2 \; .$$

Therefore

$$|\alpha_1 - \alpha_2| \leq |S(P_1 \cup P_2, f) - \alpha_1| + |S(P_1 \cup P_2, f) - \alpha_2|$$

$$< 2\varepsilon \; .$$

We have shown $|\alpha_1 - \alpha_2| < 2\varepsilon$ for each $\varepsilon > 0$. Hence $|\alpha_1 - \alpha_2| = 0$, i.e. $\alpha_1 = \alpha_2$. □

__Exercises:__

3.1: If f is integrable on I then f is bounded.

3.2: $f : I \to R^m$ is integrable on I <=> each component f_i, $i = 1,\dots,m$, of f is integrable on I and

$$\int_I f = \left(\int_I f_1, \dots, \int_I f_m\right) \; .$$

It therefore suffices to consider only real-valued functions just as in the case of sequences.

<u>Theorem 3.1</u> (Cauchy Criterion): $f : I \to R^m$. $\int_I f$ exists <=> for each $\varepsilon > 0$ there exists a partition P_ε of I such that if partitions $P,Q \supset P_\varepsilon$ then

$$|S(P,f) - S(Q,f)| < \varepsilon$$

for all Riemann sums $S(P,f)$, $S(Q,f)$ corresponding to P,Q .

<u>Proof</u>:

"=>" $\alpha = \int_I f \ \exists$ => for each $\varepsilon > 0$, $\exists$ a partition P_ε of $I \ni$ if a partition $P \supset P_\varepsilon$ then

(1) $$|S(P,f) - \alpha| < \frac{\varepsilon}{2} \quad .$$

Let $P,Q \supset P_\varepsilon$ be partitions ; (1) =>

$$|S(P,f) - S(Q,f)| \le |S(P,f) - \alpha| + |S(Q,f) - \alpha|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad .$$

i.e. the Cauchy Criterion is satisfied.

"<=" Suppose the Cauchy Criterion is satisfied. There exist partitions P_n of I such that partitions $P,Q \supset P_n$ =>

(2) $$|S(P,f) - S(Q,f)| < \frac{1}{n} \ , \quad n = 1,2,... \quad .$$

Let $Q_n = \bigcup_{k=1}^{n} P_k$; for each $n = 1,2,...$ consider a fixed Riemann sum $S_o(Q_n,f)$. If $m \ge n$, then $Q_m \supset P_n$ and $Q_n \supset P_n$ so, from (2),

$$|S_o(Q_n,f) - S_o(Q_m,f)| < \frac{1}{n} .$$

Thus $\{S_o(Q_n,f)\}$ is a Cauchy sequence of real numbers so

(3) $$\lim_{n\to\infty} S_o(Q_n,f) = \alpha \quad \text{(say) exists.}$$ (Theorem 2.7)

It remains to show that $\alpha = \int_I f$. From (3) it follows that, if $\varepsilon > 0$, there exists a natural number N such that

(4) $$\frac{1}{N} < \frac{\varepsilon}{2} \quad \text{and} \quad |S_o(Q_N,f) - \alpha| < \frac{\varepsilon}{2} .$$

If a partition P is any refinement of Q_N then both P and Q_N are refinements of P_N (from the definition of Q_N) so each Riemann sum $S(P,f)$ satisfies

$$|S(P,f)-\alpha| \le |S(P,f) - S_o(Q_N,f)| + |S_o(Q_N,f)-\alpha|$$

$$< \frac{1}{N} + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

by (2) and (4). □

<u>Lemma</u>: Let I be an interval in R^n . If P is a partition of I subdividing I into a finite collection of nonoverlapping subintervals $\{I_i\}$; then

$$\mu(I) = \sum_i \mu(I_i) .$$

This can be proved by a straightforward induction on the number of subintervals into which I is partitioned. Do it.

Corollary 3.1.1 gives an equivalent but more readily applicable version of the Cauchy Criterion.

Corollary 3.1.1 (Cauchy Criterion). $\int_I f$ exists <=> for each $\varepsilon > 0$ there exists a partition P_ε of I such that if $S_1(P_\varepsilon,f)$ and $S_2(P_\varepsilon,f)$ are any two Riemann sums corresponding to P_ε then

$$|S_1(P_\varepsilon,f) - S_2(P_\varepsilon,f)| < \varepsilon \quad .$$

Proof: It is evident from Theorem 3.1 that the condition is necessary. To see that it is sufficient let P_ε satisfy the requirement of Corollary 3.1.1 and let P and Q be refinements of P_ε, generating subintervals $\{A_k\}$ and $\{B_j\}$ respectively.

(5) $$|S(P,f) - S(Q,f)| = |\sum_k f(p_k)\,\mu(A_k) - \sum_j f(q_j)\,\mu(B_j)|$$

$$= \Big|\sum_i \Big[\sum_{A_k \subset I_i} f(p_k)\,\mu(A_k) - \sum_{B_j \subset I_i} f(q_j)\,\mu(B_j)\Big]\Big|$$

where I_i are the subintervals generated by P_ε. Now there exists points r_i^*, $r_{i*} \in I_i$ such that

(6) $$|\sum_{A_k \subset I_i} f(p_k)\,\mu(A_k) - \sum_{B_j \subset I_i} f(q_j)\,\mu(B_j)|$$

$$\leq [f(r_i^*) - f(r_{i*})]\,\mu(I_i) \quad . \quad (\dagger)$$

(†) To see that (6) is true let $f(r_i^*) = \sup\,\{f(p_k), f(q_j)\}$, $f(r_{i*}) = \inf\,\{f(p_k), f(q_j)\}$ where the sup and inf are taken over all k and j for which $A_k \subset I_i$ and $B_j \subset I_i$. There is no problem about the sup and inf existing since the sets are finite.

$$f(r_{i*}) \sum_{A_k \subset I_i} \mu(A_k) - f(r_i^*) \sum_{B_j \subset I_i} \mu(B_j)$$

$$\leq \sum_{A_k \subset I_i} f(p_k)\mu(A_k) - \sum_{B_j \subset I_i} f(q_j)\,\mu(B_j)$$

$$\leq f(r_i^*) \sum_{A_k \subset I_i} \mu(A_k) - f(r_{i*}) \sum_{B_j \subset I_i} \mu(B_j)$$

and (from the Lemma)

$$\sum_{A_k \subset I_i} \mu(A_k) = \sum_{B_j \subset I_i} \mu(B_j) = \mu(I_i)$$

so (6) follows.

From (5) and (6)

$$|S(P,f) - S(Q,f)| \leq \sum_i [f(r_i^*) - f(r_{i*})]\,\mu(I_i)$$

$$= |S_1(P_\varepsilon, f) - S_2(P_\varepsilon, f)| < \varepsilon .$$

Thus the Cauchy Criterion for integrability of f (as proved in Theorem 3.1) is satisfied, i.e. $\int_I f$ exists. □

<u>Theorem 3.2</u>: If f is a <u>continuous</u> real-valued function on I then $\int_I f$ exists.

<u>Proof</u>: Since I is compact (why?) f is uniformly continuous on I (Theorem 2.13). If $\varepsilon > 0$, $\exists\, \delta > 0 \ni |p-q| < \delta \Rightarrow |f(p)-f(q)| < \varepsilon/\mu(I)$. We may choose a partition P_ε sufficiently fine to ensure that the subintervals I_i which it generates have diameter

$$\lambda(I_i) < \delta \quad .$$

Hence if $p, q \in I_i$ then $|p - q| \leq \lambda(I_i) < \delta$ and $|f(p) - f(q)| < \varepsilon/\mu(I)$ so that if $S_1(P_\varepsilon, f)$ and $S_2(P_\varepsilon, f)$ are any two Riemann sums corresponding to P_ε

$$\begin{aligned}
|S_1(P_\varepsilon,f) - S_2(P_\varepsilon,f)| &= \left|\sum_i [f(p_i) - f(q_i)]\, \mu(I_i)\right| \\
&\leq \sum_i |f(p_i) - f(q_i)|\ \mu(I_i) \\
&< \frac{\varepsilon}{\mu(I)} \sum_i \mu(I_i) = \frac{\varepsilon}{\mu(I)}\, \mu(I) \quad \text{(by the lemma)} \\
&= \varepsilon \quad .
\end{aligned}$$

Thus by the Cauchy Criterion (Corollary 3.1.1) $\int_I f$ exists. □

<u>Theorem 3.3</u>: Suppose

(i) f is bounded on I,

(ii) the set of points of discontinuity of f has content zero.

Then $\int_I f$ exists.

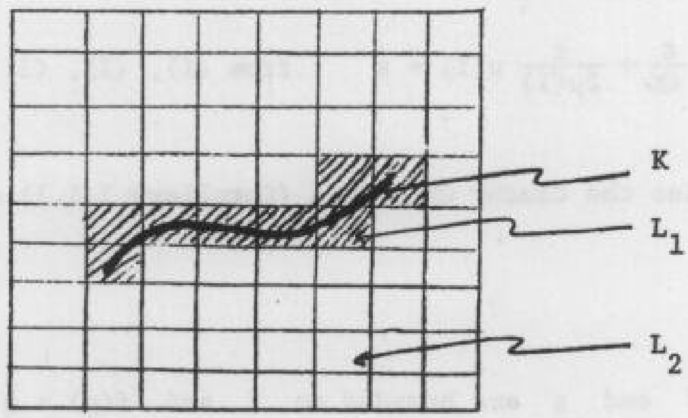
<u>Proof</u>: Suppose

$$(1) \qquad |f(p)| \leq N \quad , \quad \forall\, p \in I$$

and K is the set of points of discontinuity of f in I. Let $\varepsilon > 0$. There exists a partition $P_{0\varepsilon}$ of I such that, if I_i are the subintervals generated by $P_{0\varepsilon}$,

(2) $$\sum_{K\cap I_i\neq\phi} \mu(I_i) < \frac{\varepsilon}{4N}$$

since K has content zero.



Let $L_1 = \bigcup_{K\cap I_i\neq\phi} I_i$ (shaded), $L_2 = \bigcup_{K\cap I_i=\phi} I_i$; f is continuous on L_2 which is compact and hence f is uniformly continuous on L_2. There exists $\delta > 0 \ni p,q \in L_2$, $|p-q| < \delta \Rightarrow |f(p)-f(q)| < \frac{\varepsilon}{2\mu(I)}$. Let P_ε be a partition of I generating subintervals $\tilde{I}_i$, such that

$$P_\varepsilon \supset P_{o\varepsilon} \quad \text{and} \quad \lambda(\tilde{I}_i) < \delta \;;$$

thus if $p_i, q_i \in \tilde{I}_i \subset L_2$ then $|p_i - q_i| \leq \lambda(\tilde{I}_i) < \delta$ so

(3) $$|f(p_i) - f(q_i)| < \frac{\varepsilon}{2\mu(I)} \;.$$

Thus

$$|S_1(P_\varepsilon,f) - S_2(P_\varepsilon,f)| = \left|\sum_i [f(p_i)-f(q_i)]\, \mu(\tilde{I}_i)\right|$$
$$\leq \sum_i |f(p_i)-f(q_i)|\, \mu(\tilde{I}_i)$$

$$\leq \sum_{I_i \subset L_1} |f(p_i)-f(q_i)|\,\mu(\tilde{I}_i) + \sum_{I_i \subset L_2} |f(p_i)-f(q_i)|\,\mu(\tilde{I}_i)$$

$$< 2N \sum_{I_i \subset L_1} \mu(\tilde{I}_i) + \frac{\varepsilon}{2\mu(I)} \sum_{I_i \subset L_2} \mu(\tilde{I}_i)$$

$$\leq 2N\,\frac{\varepsilon}{4N} + \frac{\varepsilon}{2\mu(I)}\,\mu(I) = \varepsilon \qquad \text{from (1), (2), (3)} \quad .$$

Thus f satisfies the Cauchy Criterion (Corollary 3.1.1) and $\int_I f$ exists.

<u>Exercise</u>.

3.3: Suppose f and g are bounded on I and $f(p) = g(p)\ \forall\ p \in I-K$ where K has content zero; prove that

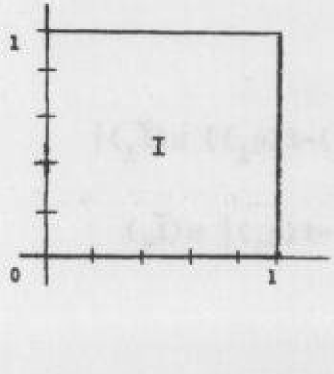
$$\int_I f\ \exists \Rightarrow \int_I g\ \exists = \int_I f \quad .$$

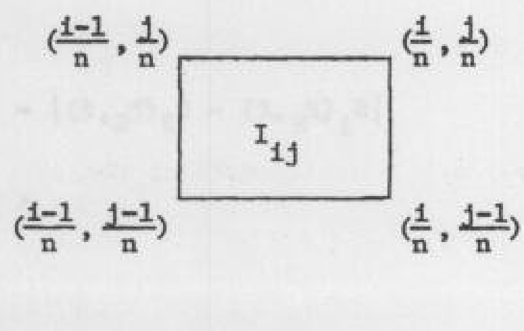
<u>Example</u>:

$$f(x,y) = xy^2\ ,\ 0 \leq x \leq 1\ ,\ 0 \leq y \leq 1$$

$I = [0,1] \times [0,1]$; f is continuous on I so $\int_I f$ exists. Partition I into n^2 subintervals of side $\frac{1}{n}$

(Partition P_n) : $I_{ij} = [\frac{i-1}{n}, \frac{i}{n}] \times [\frac{j-1}{n}, \frac{j}{n}]$, $i = 1,\ldots,n$, $j = 1,\ldots,n$.





If $p_{ij} \in I_{ij}$, $(\frac{i-1}{n})(\frac{j-1}{n})^2 \leq f(p_{ij}) \leq \frac{i}{n}(\frac{j}{n})^2$, so

$$\sum_{i,j=1}^{n} (\frac{i-1}{n})(\frac{j-1}{n})^2 \frac{1}{n^2} \leq S(P_n,f) \leq \sum_{i,j=1}^{n} \frac{i}{n}(\frac{j}{n})^2 \frac{1}{n^2}$$

and the same estimate holds for $S(P,f)$ if $P \supset P_n$.

$$\sum_{i,j=1}^{n} (\frac{i-1}{n})(\frac{j-1}{n})^2 \frac{1}{n^2} = \frac{1}{n^5} \sum_{i=0}^{n-1} i \sum_{j=0}^{n-1} j^2$$

$$= \frac{1}{n^5} [\frac{(n-1)n}{2}][\frac{(n-1)n(2n-1)}{6}] \xrightarrow[n\to\infty]{} \frac{1}{6}$$

$$\sum_{i,j=1}^{n} \frac{i}{n}(\frac{j}{n})^2 \frac{1}{n^2} = \frac{1}{n^5} \sum_{i=1}^{n} i \sum_{j=1}^{n} j^2$$

$$= \frac{1}{n^5} [\frac{n(n+1)}{2}][\frac{n(n+1)(2n+1)}{6}] \xrightarrow[n\to\infty]{} \frac{1}{6}$$

Thus $\int_I f = \frac{1}{6}$.

<u>Definition</u>: Suppose

(i) D is a bounded subset of R^n (so $D \subset I$, some closed interval in R^n)

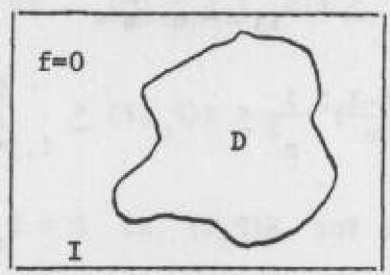
(ii) $f : D \to R$.

Extend the domain of f by defining

$$f(p) = 0 \quad \text{if} \quad p \notin D \quad .$$

Then we say $\int_D f$ exists if and only if $\int_I f$ exists and define

$$\int_D f = \int_I f \quad .$$



<u>Theorem 3.4</u>: Suppose

(i) D is <u>bounded</u> and ∂D has content zero,

(ii) f is <u>continuous</u> and <u>bounded</u> on D .

Then $\int_D f$ exists [∂D denotes the <u>boundary</u> of D (Exercise 1.42)].

<u>Proof</u>: f , with its domain extended to an interval $I \supset D$ as in the preceding definition, is continuous at each point in I except possibly at points of ∂D . But $\mu(\partial D) = 0$, so by Theorem 3.3 $\int_I f$ exists. Therefore $\int_D f$ exists and, by definition, is equal to $\int_I f$. □

<u>Definition</u>: A bounded set $D \subset R^n$ has content if $\int_D 1$ exists and

$$\mu(D) \overset{\text{def}}{=} \int_D 1 \quad ,$$

equivalently.

Define the <u>characteristic</u> <u>function</u> χ_D of D by

$$\chi_D (p) = 1 \quad , \quad \text{if} \quad p \in D$$

$$\chi_D (p) = 0 \quad , \quad \text{if} \quad p \notin D$$

and let I be a closed interval containing D as a subset. D has content if $\int_I \chi_D$ exists and then

$$\mu(D) \overset{\text{def}}{=} \int_I \chi_D \ .$$

<u>Theorem 3.5</u>: A bounded set $D \subset R^n$ has content $\iff$ $\mu(\partial D) = 0$.

<u>Proof</u>: <u>Exercise</u> 3.4.

<u>Exercise</u>.

3.5: Give an example of a countable set which has content and one which does not have content (contented and discontented sets?). Is there a countable set with positive content?

<u>Theorem 3.6</u> (Properties of the Integral):

(a) If $\int_D f$ and $\int_D g$ exist and $\alpha, \beta \in R$ then $\int_D (\alpha f + \beta g)$ exists,

and $$\int_D (\alpha f + \beta g) = \alpha \int_D f + \beta \int_D g \ .$$

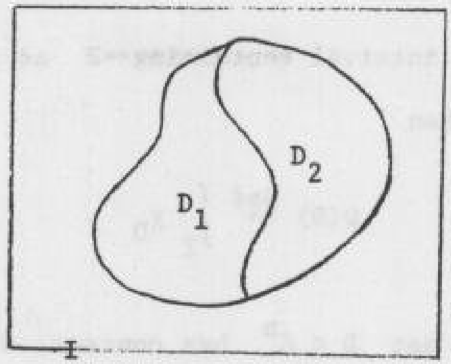
(b) If $f(p) \geq 0$, $\forall\, p \in D$ and $\int_D f$ exists then

$$\int_D f \geq 0 \ .$$

(c) If $\int_D f$ exists then $\int_D |f|$ exists, and $\left|\int_D f\right| \leq \int_D |f|$.

(d) If $\int_{D_1} f$ and $\int_{D_2} f$ exist and $\mu(D_1 \cap D_2) = 0$ then

$$\int_{D_1 \cup D_2} f = \int_{D_1} f + \int_{D_2} f \ .$$



(e) If (i) D has content,

(ii) $\int_D f$ exists,

(iii) $m \leq f(p) \leq M$, $\forall\, p \in D$, then

$$m\,\mu(D) \leq \int_D f \leq M\,\mu(D) \quad .$$

(f) (Mean Value Theorem for integrals). If D is a <u>compact</u> connected set with <u>content</u> (i.e. $\mu(\partial D) = 0$) and f is <u>continuous</u> on D then

$$\exists\; p_o \in D \ni \int_D f = f(p_o)\,\mu(D) \quad .$$

<u>Proof of (a)</u>: There is no loss of generality in assuming that D is an interval in this case. Let $\varepsilon > 0$.

$\int_D f \;\exists \Rightarrow \exists$ a partition $P_{1\varepsilon}$ of $D \ni$ if $P \supset P_{1\varepsilon}$ then $\left|S(P,f) - \int_D f\right| < \varepsilon$

$\int_D f \;\exists \Rightarrow \exists$ a partition $P_{2\varepsilon}$ of $D \ni$ if $P \supset P_{2\varepsilon}$ then $\left|S(P,g) - \int_D f\right| < \varepsilon$

Let $P_\varepsilon = P_{1\varepsilon} \cup P_{2\varepsilon}$; if $P \supset P_\varepsilon$ then

$$|S(P,\alpha f+\beta g) - \alpha\int_D f - \beta\int_D g| = |\alpha\, S(P,f) + \beta\, S(P,g) - \alpha\int_D f - \beta\int_D g|$$

$$\leq |\alpha|\ |S(P,f) - \int_D f| + |\beta|\ |S(P,g) - \int_D g|$$

$$\leq (|\alpha| + |\beta|)\varepsilon \quad .$$

Thus $\int_D (\alpha f + \beta g)$ exists and equals $\alpha\int_D f + \beta\int_D g$.

<u>Proof of (b)</u>: <u>Exercise</u> 3.6.

<u>Proof of (c)</u>: <u>Exercise</u> 3.7.

<u>Proof of (d)</u>: Let $\chi_{D_i}(p) = 1 \ , \ p \in D_i \ , \quad i = 1,2$

$$\chi_{D_i}(p) = 0 \ , \ p \notin D_i \ ,$$

and let I be an interval $I \supset D_1 \cup D_2$. By definition

$$\int_{D_1} f = \int_I f\,\chi_{D_1} = \int_{D_1 \cup D_2} f\,\chi_{D_1}$$

$$\int_{D_2} f = \int_I f\,\chi_{D_2} = \int_{D_1 \cup D_2} f\,\chi_{D_2} \quad .$$

By part (a) of the present theorem

$$\int_{D_1 \cup D_2} f\,(\chi_{D_1} + \chi_{D_2}) \ \exists = \int_{D_1 \cup D_2} f\,\chi_{D_1} + \int_{D_1 \cup D_2} f\,\chi_{D_2}$$

$$= \int_{D_1} f + \int_{D_2} f \quad .$$

But, by definition

$$\int_{D_1 \cup D_2} f\,(\chi_{D_1} + \chi_{D_2}) = \int_I f\,(\chi_{D_1} + \chi_{D_2})\,\chi_{D_1 \cup D_2}$$

$$= \int_I f\,\chi_{D_1 \cup D_2} \qquad (\dagger)$$

$$= \int_{D_1 \cup D_2} f \;.$$

Thus $\int_{D_1 \cup D_2} f$ exists and equals $\int_{D_1} f + \int_{D_2} f$.

Proof of (e): Exercise 3.8. [Hint: Use parts (a) and (b).]

Proof of (f):

Case (i): $\mu(D) = 0$:

$$\int_D f = 0 \text{ , by (e) since } m\,\mu(D) = M\,\mu(D) = 0$$

$$\therefore \quad \int_D f = f(p_o)\,\mu(D) \text{ , } \forall\, p_o \in D \;.$$

Case (ii): $\mu(D) \neq 0$.

By (e) $m\,\mu(D) \leq \int_D f \leq M\,\mu(D)$ where

$(\dagger)$ follows from Exercise 3.3 since

$$f(\chi_{D_1} + \chi_{D_2})\chi_{D_1 \cup D_2} = f\,\chi_{D_1 \cup D_2}$$

except on the set $D_1 \cap D_2$ and $\mu(D_1 \cap D_2) = 0$.

$$m = \inf \{f(p) : p \in D\} , \quad M = \sup \{f(p) : p \in D\} .$$

Thus

$$m \leq \frac{\int_D f}{\mu(D)} \leq M .$$

By Corollary 2.12.2, D compact and f continuous on $D \Rightarrow \exists\ p_1, p_2 \in D$ $\ni f(p_1) = m$, $f(p_2) = M$; therefore by Corollary 2.11.1 (the Intermediate Value Theorem) there exists $p_o \in D$ such that

$$f(p_o) = \frac{\int_D f}{\mu(D)} ,$$

$$\text{i.e. } f(p_o)\, \mu(D) = \int_D f . \qquad \square$$

<u>Exercises</u>.

3.9: Prove that $\{(\frac{1}{m}, \frac{1}{n}) : m,n = 1,2,\ldots\}$ has content zero in R^2 .

3.10: Prove that the set of points in $[0,1]\times[0,1]$ with rational coordinates does not have content in R^2 .

3.11: Let D be a subset of R^n which has content zero and f any <u>bounded</u> function on D . Prove that $\int_D f$ exists and equals 0.

3.12: Let D be any bounded subset of R^n and $f(p) = 0$, $\forall\ p \in D$. Prove that $\int_D f$ exists and equals 0. Note that D does not necessarily have content.

3.13: Show that $\int_0^1 \sin\frac{1}{x}\,dx$ exists and is independent of the value assigned to the integrand at $x = 0$.

3.14: If $\int_D f$ exists and D_1 is any subset of D with content then $\int_{D_1} f$ exists (D itself does not necessarily have content). Note that this exercise proves in particular that $\int_a^b f \ \exists \Rightarrow \int_a^c f \ \exists$, $\forall\, c \in (a,b)$.

3.15: If $f(x) = 0$, $x \notin Q$, $f(x) = 1$, $x \in Q$, then f is not Riemann integrable on $[0,1]$.

3.16: If $f(x) = 0$, $x \notin Q$ and $f(\frac{m}{n}) = \frac{1}{n}$ (where m and n have no common divisors) then $\int_0^1 f = 0$.

3.17: If P is a partition of the interval I and $f : I \to R$, f bounded, define $\underline{S}(P,f)$ and $\overline{S}(P,f)$, the <u>lower</u> and <u>upper</u> Riemann sums corresponding to the partition P to be $\inf\{S(P,f)\}$ and $\sup\{S(P,f)\}$ respectively, the inf and sup being taken over all Riemann sums corresponding to the partition P. Let $\underline{\int}_I f = \sup_P \{\underline{S}(P,f)\}$ and $\overline{\int}_I f = \inf_P \{\overline{S}(P,f)\}$. The sup and inf in this case being taken over all partitions P of I.

(i) Show that $Q \supset P \Rightarrow \underline{S}(P,f) \leq \underline{S}(Q,f) \leq \overline{S}(Q,f) \leq \overline{S}(P,f)$

(ii) Show that $\underline{\int}_I f \leq \overline{\int}_I f$.

(iii) (Definition). f is integrable on I if $\underline{\int}_I f = \overline{\int}_I f$ and then we define $\int_I f = \underline{\int}_I f = \overline{\int}_I f$.

(iv) Show that f is integrable in this sense if and only if there is exactly one number α such that

$$\underline{S}(P,f) \leq \alpha \leq \overline{S}(P,f)$$

for every partition P of I and then $\int_I f = \alpha$.

(v) (Cauchy Criterion). Show that $\int_I f$ exists in this sense if and only if, for each $\varepsilon > 0$, there exists a partition P_ε of I such that

$$|\underline{S}(P_\varepsilon,f) - \overline{S}(P_\varepsilon,f)| < \varepsilon \quad .$$

(vi) Show that, if f is bounded, f is integrable on I in the sense (iii) <=> f is integrable in the sense adopted in these notes and both definitions give the same value for $\int_I f$.

Discussion: A subset K of R^n has Jordan measure zero if, for each $\varepsilon > 0$, there exists a finite collection of intervals $\{I_k\}$ such that

(1) $$K \supset \bigcup_k I_k \quad \text{and} \quad \sum_k \mu(I_k) \leq \varepsilon \quad .$$

K has Lebesgue measure zero if, for each $\varepsilon > 0$, there exists a countable collection of intervals I_k satisfying (1) ($\sum_{k=1}^{\infty} c_k < \varepsilon$ means $\lim_{n\to\infty} \sum_{k=1}^{n} c_k$ exists and is less than ε). This simple extension of the concept of measure zero has profound consequences. We know that the set of rationals in $[0,1]$ does not have Jordan content. However if this set is denoted by $\{r_k : k = 1,2,\ldots\}$ then $r_k \in I_k$, $\mu(I_k) = \frac{\varepsilon}{2^k}$ so that

$$\{r_k ; k = 1,2,\ldots\} \subset \cup I_k \text{ and } \sum \mu(I_k) = \varepsilon \sum_{k=1}^{\infty} \frac{1}{2^k} = \varepsilon \quad .$$

Thus the set of rationals in $[0,1]$ has Lebesgue measure zero. In fact the same argument shows that Q or any countable set has Lebesgue measure zero. Countable sets are not the only sets of Lebesgue measure zero however; in fact they can be quite complicated. The remarkable Cantor set which you will study in Math 417 has Lebesgue measure zero but nevertheless has the same cardinality as $[0,1]$ i.e. it has zero "length" but has the same "number" of points as $[0,1]$. A simple discussion of the Cantor set may be found in Bartle's book (page 51).

An interesting theorem of Lebesgue states that, if f is bounded, the Riemann integral $\int_I f$ exists if and only if the set of points in I at which f is discontinuous has Lebesgue measure zero. For example consider the functions in Exercises 3.15, 3.16; show that the function in 3.15 is discontinuous at every point in $[0,1]$; on the other hand show that the function in 3.16 is discontinuous only at rational points in $[0,1]$.

EVALUATION OF INTEGRALS

Real valued functions of a real variable:

Theorem 3.7 (Fundamental Theorem of Calculus): If f is continuous on $[a,b]$ then there is a function F_o on $[a,b]$ such that

$$F_o'(x) = f(x) \quad , \quad a \leq x \leq b \quad .$$

Proof: Consider $F_o(x) = \int_a^x f$, $a \leq x \leq b$ (exists by Exercise 3.14).

$$F_o(x+h) - F_o(x) = \int_x^{x+h} f \quad \text{(Theorem 3.6(d))}$$

$$= f(c_h)h \quad \text{(MVTh for integrals, Theorem 3.6(f))}$$

where $c_h \in [x,x+h]$ if $h > 0$, and $c_h \in [x+h,x]$ if $h < 0$. Since f is continuous on $[a,b]$

$$\lim_{h\to 0} \frac{F_o(x+h) - F_o(x)}{h} = \lim_{h\to 0} f(c_h) = f(x) \quad . \qquad \Box$$

Any function F such that $F'(x) = f(x)$, $a \leq x \leq b$ is an anti-derivative of f on $[a,b]$.

Proposition: If F_1 and F_2 are antiderivatives of f then F_1-F_2 is a constant function.

Proof: $F_1'(x) - F_2'(x) = f(x) - f(x) = 0$, all x .

Therefore $F_1 - F_2$ is a constant (cf. Exercise 2.55). □

Theorem 3.8: If f is continuous on $[a,b]$ and F is any antiderivative of f on $[a,b]$ then

$$\int_a^b f = F(b) - F(a) .$$

Proof:

$$\int_a^x f = F_o(x) = F(x) - C \qquad \text{(preceding Proposition)}$$

$$x = b : \int_a^b f = F_o(b) = F(b) - C$$

$$x = a : \int_a^a f = F_o(a) = 0 = F(a) - C \qquad \text{(Exercise 3.11)}$$

$$\therefore \quad F(a) = C$$

$$\text{so} \qquad \int_a^b f = F(b) - F(a) .$$ □

Corollary 3.8.1 (Change of Variable Formula): Suppose

(i) ϕ' exists and is continuous on $[a,b]$,

(ii) f is continuous on $\phi([a,b])$.

Then

$$\int_{\phi(a)}^{\phi(b)} f = \int_a^b (f\circ\phi)\,\phi'$$

$$\text{i.e.} \quad \int_{\phi(a)}^{\phi(b)} f(x)\, dx = \int_a^b f(\phi(u))\, \phi'(u)\, du \,.$$

<u>Proof</u>: If F is an antiderivative of f (i.e. $F'(x) = f(x)$) then $\frac{d}{du} F(\phi(u)) = f(\phi(u))\, \phi'(u)$ so that $F \circ \phi$ is an antiderivative of $(f \circ \phi)\, \phi'$. Therefore

$$\int_{\phi(a)}^{\phi(b)} f = F(\phi(b)) - F(\phi(a)) = \int_a^b (f \circ \phi)\, \phi' \,. \qquad \Box$$

You will notice that we have used the Chain Rule here although it has not yet been proved in these notes; a proof will appear in a more general context in the next chapter.

<u>Example</u>:

$$\int_0^{\pi/2} \sin^2 u \cos u\, du \qquad\qquad f(x) = x^2$$

$$\phi(u) = \sin u$$

$$= \int_{\phi(0)}^{\phi(\pi/2)} x^2\, dx \qquad\qquad \phi'(u) = \cos u$$

$$= \int_0^1 x^2\, dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3} \,.$$

<u>Corollary 3.8.2</u> (Integration by Parts Formula): If f' and g' are continuous on $[a,b]$ then

$$\int_a^b fg' + \int_a^b f'g = f(b)g(b) - f(a)g(a) \,.$$

<u>Proof</u>: fg is an antiderivative of $fg' + f'g$. Therefore

$$\int_a^b (fg' + f'g) = f(x)\ g(x)\Big|_a^b \quad . \qquad \square$$

<u>Example</u>:

$$\int_0^1 x\ e^x\ dx = x\ e^x\Big|_0^1 - \int_0^1 e^x\ dx = e-(e^x\Big|_0^1)$$

$$= e-(e-1) = 1 \ .$$

<u>Real valued functions on R^2</u> :

The important theorem of Fubini allows us to extend the use of the Fundamental Theorem of Calculus to functions of more than one variable.

<u>Theorem 3.9</u> (Fubini's Theorem): $f : I \to R$,

$$I = [a,b]\times[c,d] = \{(x,y) : a \leq x \leq b \ , \ c \leq y \leq d\} \quad .$$

Suppose:

(i) $\int_I f$ exists

(ii) $\int_c^d f(x,y)dy = F(x)$ exists for each $x \in [a,b]$. Then $\int_a^b F = \int_a^b \left\{ \int_c^d f(x,y)dy \right\} dx$ exists and equals $\int_I f$.

<u>Proof</u>: The definition of $\int_I f$ states that, if $\varepsilon > 0$, there is a partition P_ε of I such that if $P \supset P_\varepsilon$ then

$$|S(P,f) - \int_I f| < \varepsilon$$

for any Riemann sum $S(P,f)$. Thus, if

$$P = \{x_o,x_1,\ldots,x_m\}\times\{y_o,y_1,\ldots,y_n\} \supset P_\varepsilon ,$$

$$|\sum_{i=1}^{m} \sum_{j=1}^{n} f(\overline{x}_i,\overline{y}_j)(x_i-x_{i-1})(y_j-y_{j-1}) - \int_I f| < \varepsilon$$

if $x_{i-1} \leq \overline{x}_i \leq x_i$, $y_{j-1} \leq \overline{y}_j \leq y_j$. This may be written

$$(1) \qquad |\sum_{i=1}^{m} (x_i-x_{i-1}) \{\sum_{j=1}^{n} f(\overline{x}_i,\overline{y}_j)(y_j-y_{j-1})\} - \int_I f| < \varepsilon .$$

Condition (ii) implies that for any fixed set of numbers $\overline{x}_i$, $i = 1,\ldots,m$, the partition $\{y_j\}$ of $[c,d]$ may be chosen so fine that

$$|\sum_{j=1}^{n} f(\overline{x}_i,\overline{y}_j)(y_j-y_{j-1}) - \int_c^d f(\overline{x}_i,y)dy| < \frac{\varepsilon}{b-a}$$

$i = 1,\ldots,m$, since each of the integrals $\int_c^d f(\overline{x}_i,y)dy$ exists. Therefore

$$(2) \qquad |\sum_{j=1}^{n} f(\overline{x}_i,\overline{y}_j)(y_j-y_{j-1}) - F(\overline{x}_i)| < \frac{\varepsilon}{b-a} , \quad i = 1,\ldots,m .$$

(2) implies

$$(3) \qquad |\sum_{i=1}^{m} (x_i-x_{i-1}) \{\sum_{j=1}^{n} f(\overline{x}_i,\overline{y}_j)(y_j-y_{j-1})\} - \sum_{i=1}^{m} (x_i-x_{i-1}) F(\overline{x}_i)|$$

$$< \sum_{i=1}^{m} \frac{\varepsilon}{b-a} (x_i-x_{i-1}) = \frac{\varepsilon}{b-a} (b-a) = \varepsilon .$$

The triangle inequality with (1) and (3) gives

$$\left| \sum_{i=1}^{m} F(\overline{x}_i)(x_i - x_{i-1}) - \int_I f \right| < 2\varepsilon \quad .$$

Thus, from the definition of the integral, $\int_a^b F$ exists and equals $\int_I f$. □

<u>Corollary 3.9.1</u> (Interchanging the order of integration): $I = [a,b]\times[c,d]$.

If

(i) $\int_I f \ \exists$,

(ii) $\int_c^d f(x,y)\ dy \ \exists \ \forall\ x \in [a,b]$,

(iii) $\int_a^b f(x,y)\ dx \ \exists \ \forall\ y \in [c,d]$,

then

$$\int_a^b \{\int_c^d f(x,y)dy\}\ dx = \int_I f = \int_c^d \{\int_a^b f(x,y)dx\}\ dy \quad .$$

<u>Corollary 3.9.2</u>: $I = [a,b]\times[c,d]$. Suppose

(i) f is <u>bounded</u> on I ,

(ii) f is <u>continuous</u> on $I-K$, $\mu_2(K) = 0$,

(iii) $\mu_1(K\cap\{(x,y) : c \le y \le d\}) = 0$ for each $x \in [a,b]$.

Then

$$\int_I f = \int_a^b \{\int_c^d f(x,y)\ dy\}\ dx \quad .$$

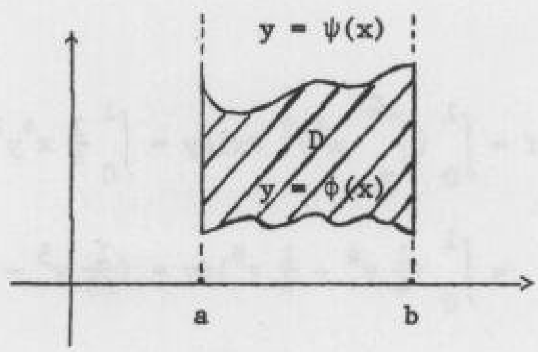
<u>Proof</u>: (i), (ii) => $\int_I f \ \exists$, by Theorem 3.3. Condition (iii) says that the intersection of K with each vertical line (considered as a set in R)

has content zero so $\int_c^d f(x,y)\,dy$ exists for each $x \in [a,b]$ again by Theorem 3.3. Therefore all the conditions of Fubini's Theorem are satisfied. □

<u>Corollary 3.9.3</u>: If $D = \{(x,y) : a \leq x \leq b\ ,\ \phi(x) \leq y \leq \psi(x)\}$ where ϕ and ψ are continuous on $[a,b]$ and $f : R^2 \to R$ is continuous on D then

$$\int_D f = \int_a^b \{ \int_{\phi(x)}^{\psi(x)} f(x,y)\,dy\}\ dx \quad .$$

<u>Proof</u>: Use Corollary 3.9.2. The graphs of ϕ and ψ have zero content in R^2 . Each vertical line intersects each graph once.



□

<u>Examples</u>.

(1) $f(x,y) = xy^2$, $I = [0,1]\times[0,1]$

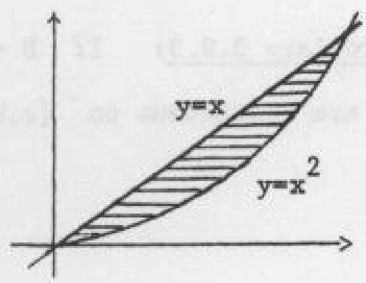
$$\int_I f = \int_0^1 \{\int_0^1 xy^2\ dy\}dx = \int_0^1 \tfrac{1}{3}\,x\ dx = \tfrac{1}{6} \quad .$$

Also

$$\int_I f = \int_0^1 \{\int_0^1 xy^2\ dx\}dy = \int_0^1 \tfrac{1}{2}\,y^2\ dy = \tfrac{1}{6} \ .$$

We have already seen from first principles that this is the value of the integral.

(2) $f(x,y) = x^3y^2$, $D = \{(x,y) : 0 \le x \le 1 , x^2 \le y \le x\}$



$$\int_D f = \int_0^1 \{\int_{x^2}^{x} x^3y^2 dy\} \; dx$$

$$= \int_0^1 \frac{1}{3} x^3y^3 \Big|_{y=x^2}^{y=x} dx$$

$$= \int_0^1 (\frac{1}{3} x^6 - \frac{1}{3} x^9) \; dx$$

$$= (\frac{1}{21} x^7 - \frac{1}{30} x^{10})\Big|_0^1 = \frac{1}{70} .$$

Also

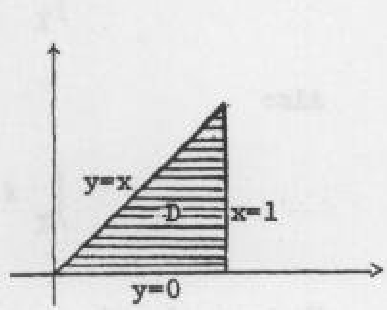
$$\int_D f = \int_0^1 \{\int_y^{\sqrt{y}} x^3y^2 \; dx\}dy = \int_0^1 \frac{1}{4} x^4y^2\Big|_{x=y}^{x=\sqrt{y}} \; dy$$

$$= \int_0^1 (\frac{1}{4} y^4 - \frac{1}{4} y^6)dy = (\frac{1}{20} y^5 - \frac{1}{28} y^7)\Big|_0^1 = \frac{1}{70} .$$

In Examples (1), (2) the iterated integrals are all easily evaluated. However one sometimes encounters iterated integrals where antiderivatives cannot be found in terms of elementary functions. A simplification is sometimes achieved by using Fubini's Theorem to reverse the order of integration.

(3) $\int_0^1 \{\int_y^1 e^{y/x} \; dx\} \; dy$

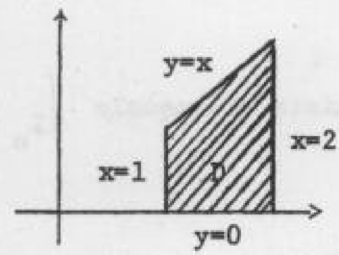
$= \int_D f$, by Fubini's Theorem



$$= \int_0^1 \{\int_0^x e^{y/x}\, dy\} dx \quad , \quad \text{again by Fubini's Theorem}$$

$$= \int_0^1 x\, e^{y/x} \Big|_{y=0}^{y=x} dx = \int_0^1 x(e-1)\, dx = \frac{1}{2}\,(e-1) \quad .$$

(4) $$\int_1^2 \{\int_0^x f(x,y)\, dy\}\, dx = \int_D f$$



$$= \int_0^2 \{\int_{\phi(y)}^2 f(x,y)\,dx\}\, dy$$

$$\text{where} \quad \phi(y) = \begin{cases} 1 \ , & 0 \le y \le 1 \\ y \ , & 1 \le y \le 2 \end{cases}$$

$$= \int_0^1 \{\int_1^2 f(x,y)\,dx\}dy + \int_1^2 \{\int_y^2 f(x,y)\,dx\}\, dy \quad .$$

Further worked examples may be found in Buck pp. 115-119.

<u>Real valued functions on</u> R^n :

Fubini's Theorem may be stated for n-dimensional intervals as follows:

<u>Theorem 3.10</u> (Fubini's Theorem): Let I_ℓ , I_m be closed intervals in R^ℓ , R^m , $\ell+m = n$, so

$$I_n = I_\ell \times I_m = \{(p,q) : p \in I_\ell \ , \ q \in I_m\}$$

is a closed interval in R^n . Let $f : I_n \to R$ and suppose

(i) $\int_{I_n} f$ exists,

(ii) $F(p) = \int_{I_m} f(p,q)\,dq$, exists for each $p \in I_\ell$. Then

$$\int_{I_\ell} F = \int_{I_\ell} \{\int_{I_m} f(p,q)\,dq\}\,dp$$

exists and equals $\int_{I_n} f$.

Notes:

(1) The proof of Theorem 3.10 is exactly the same as that given before when $n=2$.

(2) The symbols "dp" , "dq" above are simply used as devices to indicate the spaces on which we are integrating.

(3) For a more general formulation of Fubini's Theorem where the condition (ii) is dropped see "Calculus on Manifolds" by M. Spivak (p. 58). However the above statement of the theorem is sufficient for our needs.

(4) We will prove a change of variables formula for integrals in higher dimensions in a later chapter.

Exercises.

3.18: Let f be a real-valued function on $[a,b]$ such that $f'(x)$ exists for each $x \in [a,b]$- and $\int_a^b f'$ exists. Prove

$$f(b) - f(a) = \int_a^b f' \quad .$$

(You may not use the Fundamental Theorem of Calculus. Why?)

3.19: Let f be a non-negative continuous function on $[a,b]$. Prove that $\int_a^b f$ is the content in R^2 of the set

$$D = \{(x,y) : a \leq x \leq b \ , \ 0 \leq y \leq f(x)\} \quad .$$

(i.e. show $\int_D 1 = \int_a^b f$).

3.20: Let $D = \{(x,y) : 1 \leq x \leq 3 \ , \ x^2 \leq y \leq x^2+1\}$. Show that

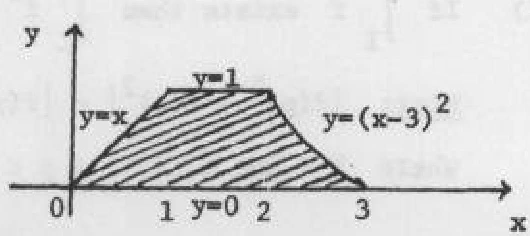
$$\mu(D) = \int_1^3 \{\int_{x^2}^{x^2+1} dy\} \, dx = 2 \quad .$$

3.21: Let $f(x,y) = x^2+y^2$, $D = \{(x,y) : 0 \leq x \leq a \ , \ 0 \leq y \leq b\}$. Show

$$\int_D f = \frac{1}{3} ab \, (a^2+b^2) \quad .$$

3.22: Let $f(x,y) = xy$, D the region in the diagram. Show $\int_D f = \frac{131}{120}$.

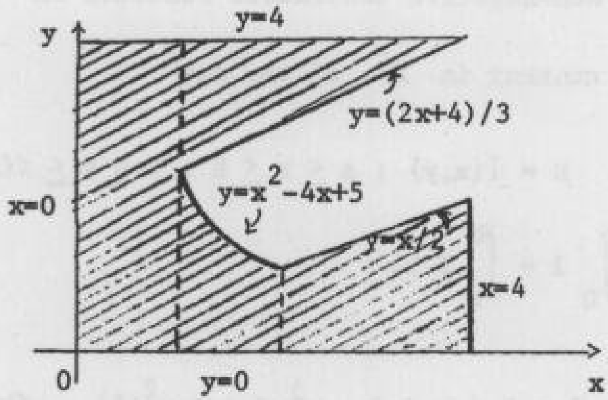
(Do this in two ways.)



3.23: Find $\int_D f$; $f(x,y) = \sin(\frac{x}{a} + \frac{y}{b})$, $D = \{(x,y) : 0 \le x \le \frac{\pi a}{2} , 0 \le y \le \frac{\pi b}{2}\}$.

3.24: Let D be the region bounded by the curves $x^2-y^2 = 1$, $x^2+y^2 = 4$ which contains $(0,0)$. Find $\int_D f$ where $f(x,y) = x^2$.

3.25: Let $f(x,y) = x$; prove that $\int_D f = \frac{77}{4}$ where D is the region in R^2 illustrated below.



3.26: (i) Let $f(x,y) = g(x)$, $(x,y) \in [a,b]\times[c,d] = I$. Prove that if $\int_a^b g$ exists then $\int_I f$ exists. Deduce from this that $\int_D f$ exists where D is any subset of I which has content.

(ii) Suppose g is defined on $[0,1]$ and $\int_0^1 g$ exists. Prove that

$$\int_0^1 [\int_x^1 g(t)dt]\, dx = \int_0^1 t\, g(t)\, dt \quad .$$

3.27: (i) If $\int_I f$ exists then $\int_I f^2$ exists.

<u>Hint</u>: $|f(p)^2 - f(q)^2| = |f(p) - f(q)|\ |f(p) + f(q)| \le 2M|f(p)-f(q)|$ where $M = \sup\{|f(p)| : p \in I\}$.

(ii) Deduce from (i) that if $\int_I f$, $\int_I g$ exist then $\int_I fg$ exists. (Hint: $(f-g)^2 = f^2 - 2fg + g^2$.)

3.28: The mean-value Theorem for integrals (Theorem 3.6(f)) implies the following: "If f is <u>continuous</u> on $[a,b]$ then $\exists\ c \in [a,b] \ni$

$$\int_a^b f = f(c)\ (b-a) \text{ ."}$$

Can you replace "continuous" by a less restrictive condition which still implies this result? What was Darboux's first name?

3.29: (Cavalieri's Principle) Let A and B be subsets of R^2 with content. If $x \in R$ define

$$A_x = \{y : (x,y) \in A\} \quad , \quad B_x = \{y : (x,y) \in B\}$$

(sections of A and B). Suppose that, for each x, A_x and B_x have content in R and $\mu_1(A_x) = \mu_1(B_x)$. Prove that $\mu_2(A) = \mu_2(B)$. Spell Fubini.

3.30: Let f be a real-valued function on $[0,1]$ such that $\int_0^1 f$ exists; define a_n by

$$a_n = \frac{1}{n} \sum_{k=1}^{n} f\left(\frac{k}{n}\right) \quad , \quad n = 1,2,\ldots \quad .$$

Show that $\{a_n\}$ is convergent and

$$\lim_{n\to\infty} a_n = \int_0^1 f \quad .$$

3.31: Let f,g be integrable on $[a,b]$ and $g(x) \geq 0$, $a \leq x \leq b$, f continuous on $[a,b]$. Prove that $\exists\ c \in [a,b] \ni$

$$\int_a^b fg = f(c) \int_a^b g \ .$$

3.32: Let D be a compact subset of R^n and $f : R^n \to R^m$ be continuous on D . Show that $\{(p,f(p)) : p \in D\}$ has Jordan content zero in R^{n+m} .

3.33: (i) Let f be a continuous real-valued function on $[0,1]$. Prove

$$\int_0^\pi x\, f(\sin x)dx = \pi \int_0^{\pi/2} f(\sin x)\ dx \ .$$

(ii) Deduce $\int_0^\pi \frac{x \sin x}{2 - \sin^2 x} dx = \frac{1}{4}\pi^2$ from (i).

3.34: Use Fubini's Theorem to show that

$$\int_0^1 \{\int_{\sqrt{y}}^{2-\sqrt{y}} f(x,y)dx\}dy = \int_0^1 \{\int_0^{x^2} f(x,y)dy\}dx + \int_1^2 \{\int_0^{(x-2)^2} f(x,y)dy\}dx \ .$$

3.35: Show that $\int_D 1 = \frac{1}{6}$ where $D = \{(x,y,z) : x \geq 0 , y \geq 0 , z \geq 0 , 0 \leq x+y+z \leq 1\}$.

3.36: Let D be a subset of R^2 with content and f a positive continuous function on D . Use Fubini's Theorem to show, if

$$K = \{(x,y,z) : (x,y) \in D , 0 \leq z \leq f(x,y)\} \ ,$$

then

$$\mu_3(K) = \int_K 1 = \int_D f \quad .$$

Deduce $m\,\mu_2(D) \leq \mu_3(K) \leq M\,\mu_2(D)$ if m,M are lower and upper bounds on f in D .

3.37: Evaluate $\int_0^2 \int_y^2 e^{x^2}\,dx\,dy$ and $\int_0^\pi \int_x^\pi \frac{\sin y}{y}\,dy\,dx$.

3.38: Show

$$\int_0^1 \{ \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \sin \frac{\pi y}{\sqrt{1-x^2}}\,dx\}\,dy = 1 \quad .$$

Explain carefully why each integral you consider exists.

See also Exercises pp. 122-124, Buck.

References for Chapter III

R.G. Bartle: Chapter VI

R.C. Buck: Chapter III.

CHAPTER FOUR

DIFFERENTIATION OF FUNCTIONS OF SEVERAL VARIABLES

PRELIMINARIES

Linear Functions.

Definition: $L : R^n \to R^m$, L is linear if, for all $p,q \in R^n$ and $\lambda \in R$,

(i) $L(p+q) = L(p) + L(q)$ (additive)

(ii) $L(\lambda p) = \lambda L(p)$ (homogeneous)

If $v_o \in R^m$ and L is linear than $M : R^n \to R^m$, $M(p) = v_o + L(p)$, is an affine function.

Theorem 4.1: $L : R^n \to R^m$ is linear <=> there is a matrix $[C_{ij}]$, $i = 1,\ldots,m$, $j = 1,\ldots,n$ such that if $p = (x_1,\ldots,x_n) \in R^n$ and $L(p) = q = (y_1,\ldots,y_m) \in R^m$ then

$$\begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ y_m \end{bmatrix} = \begin{bmatrix} C_{11} , \ldots , C_{1n} \\ \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \\ C_{m1} , \ldots , C_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} , \text{ or } q^T = C\, p^T \quad (T = \text{transpose})$$

i.e. $y_i = \sum_{j=1}^{n} C_{ij}\, x_j$, $i = 1,\ldots,m$.

Proof:

"<=" : If L is representable by a matrix then L is clearly linear.

"=>" : Conversely let L be linear.

$$L(p) = (L_1(p),\ldots,L_m(p)) = (y_1,\ldots,y_m) \quad .$$

if $e_j = (0,\ldots,\underset{\substack{\uparrow \\ j\text{th entry}}}{1},\ldots,0)$, $j = 1,\ldots,n$, then since

$$p = (x_1,\ldots,x_n) = x_1e_1 + \ldots + x_ne_n$$

$$\begin{aligned} L(p) &= L(x_1e_1 + \ldots + x_ne_n) \\ &= x_1\, L(e_1) + \ldots + x_n\, L(e_n) \quad . \end{aligned}$$

Thus

$$L_i(p) = x_1\, L_i(e_1) + \ldots + x_n\, L_i(e_n) \quad , \quad i = 1,\ldots,m$$

i.e.,

$$y_i = \sum_{j=1}^{n} L_i(e_j)x_j \quad , \quad C_{ij} = L_i(e_j) \quad ,$$

$$C = \begin{bmatrix} | & & | \\ L(e_1)^T & ,\ \ldots\ , & L(e_n)^T \\ | & & | \end{bmatrix} \quad . \qquad \square$$

Recall from linear algebra that the <u>rank</u> of a matrix C is the number of vectors in the largest linearly independent set which can be chosen from the columns of C . Also, if $L : R^n \to R^m$ is linear with matrix representation C then

$$\text{rank } C = \text{dimension } L(R^n)$$

i.e. $L(R^n) = \{x_1 L(e_1) + \ldots + x_n L(e_n) : x_i \in R\}$, from Theorem 4.1, so the minimum number of vectors which it takes to span $L(R^n)$ in R^m is precisely rank C. The range $L(R^n)$ is a linear subspace of R^m (e.g. if $m = 3$ it is either the origin, or a line or a plane containing the origin or R^3 itself). If $v_o \in R^m$ then

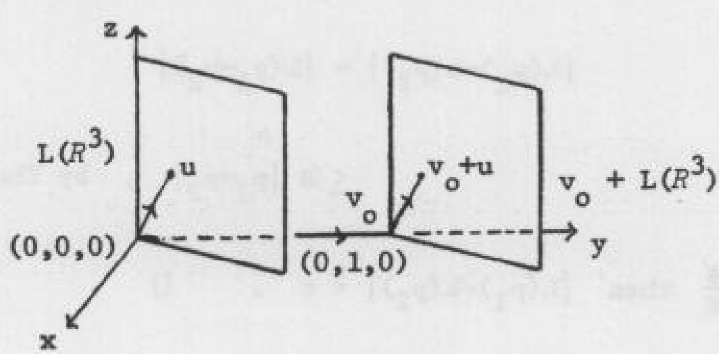
$$v_o + L(R^n) = \{v_o + v : v \in L(R^n)\} = \{v_o + L(u) : u \in R^n\}$$

is called an <u>affine</u> <u>space</u> and is evidently $L(R^n)$ translated by the vector v_o .

<u>Example</u>:

$$\text{rank} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} = 2$$

Thus the range $L(R^3)$ of the corresponding linear function $L : R^3 \to R^3$ has dimension 2. In fact it consists of all vectors of the form (s,s,t) i.e., the plane $y = x$. An example of an affine space in R^3 is $(0,1,0) + L(R^3)$ which is the parallel plane through $(0,1,0)$, i.e. $y = x+1$.



Theorem 4.2. If L is linear there is a constant M such that

$$|L(p)| \leq M|p| \quad , \quad \forall\, p \in R^n .$$

Proof:

$$|y_i| = |L_i(p)| = \left| \sum_{j=1}^{n} c_{ij}\, x_j \right|$$

$$\leq \sqrt{\sum_{j=1}^{n} c_{ij}^{\;2}}\; \sqrt{\sum_{j=1}^{n} x_j^2} \qquad \text{(CBS)}$$

$$|L(p)| = \sqrt{\sum_{i=1}^{m} y_i^2}$$

$$\leq \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}^{\;2}}\; \sqrt{\sum_{j=1}^{n} x_j^2}$$

$$= M\,|p| \quad \text{where} \quad M = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij}^{\;2}} \qquad \square$$

Corollary 4.2.1: A linear function $L : R^n \to R^m$ is uniformly continuous on R^n .

Proof:

$$|L(p_1)-L(p_2)| = |L(p_1-p_2)|$$

$$\leq M\,|p_1-p_2| \quad , \quad \text{by Theorem 4.2.}$$

Thus if $|p_1-p_2| < \frac{\varepsilon}{M}$ then $|L(p_1)-L(p_2)| < \varepsilon$. $\square$

<u>Definition</u>: $L : R^n \to R^m$ is one-to-one (1-1) if $L(p_1) = L(p_2) \Rightarrow p_1 = p_2$ equivalently $p_1 \neq p_2 \Rightarrow L(p_1) \neq L(p_2))$.

<u>Theorem 4.3</u>: $L : R^n \to R^m$, L linear. L is one-to-one $\Leftrightarrow$ there exists a constant $k > 0$ such that $|L(p)| \geq k\,|p|$ for all $p \in R^n$.

<u>Proof</u>:

"$\Leftarrow$" : If such a constant k exists then

$$|L(p_1)-L(p_2)| = |L(p_1-p_2)| \geq k|p_1-p_2|$$

$$\therefore\ p_1 \neq p_2 \Rightarrow L(p_1) \neq L(p_2) \quad \text{so } L \text{ is } (1\text{-}1) \ .$$

"$\Rightarrow$" : L is continuous on R^n (Corollary 4.2.1)

$\Rightarrow |L|$ is continuous on R^n (Corollary 2.9.1)

$\Rightarrow |L|$ is continuous on $S = \{p : |p| = 1\}$.

Now S is a compact set so, since $|L|$ is continuous on S , $|L|$ achieves a minimum value on S (Theorem 2.12.2), at p_o say. Thus

$$k \overset{\text{def}}{=} \min\{|L(p)| : p \in S\} = |L(p_o)| \ , \ |p_o| = 1 \ .$$

Now $k = |L(p_o)| > 0$ since $L(0) = 0$ (from linearity) and thus $L(p) \neq 0$ if $p \neq 0$ because L is (1-1) ; in particular $L(p_o) \neq 0$ ($|p_o| = 1$) . If $p \in R^n$, $p \neq 0$ then $\frac{1}{|p|}\,p \in S$ so

$$k \leq |L(\tfrac{1}{|p|}\,p)| = \frac{1}{|p|}\,|L(p)| \ . \qquad \square$$

<u>Exercises</u>.

4.1: Let $L : R^2 \to R^3$ be linear with $L(e_1) = (2,1,0)$, $L(e_2) = (1,0,-1)$, $e_1 = (1,0)$, $e_2 = (0,1)$. Find $L(2,0)$, $L(1,1)$, $L(1,3)$. Draw pictures.

4.2: Show that $L(R^2) \neq R^3$ for the function in Exercise 4.1.

4.3: Show that if $L : R^2 \to R^3$ is linear then

$$L(R^2) \neq R^3 .$$

4.4: Let $L : R^3 \to R^2$ be linear. Show that there are non-zero vectors $p \in R^3$ such that $L(p) = 0$.

4.5: If $L : R^2 \to R^2$ has matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, show that

(i) $L(R^2)$ = single point $\Leftrightarrow a = b = c = d = 0$.

(ii) $L(R^2)$ = line $\Leftrightarrow \Delta = ad-bc = 0$, $a^2 + b^2 + c^2 + d^2 > 0$.

(iii) $L(R^2) = R^2 \Leftrightarrow \Delta \neq 0$.

4.6: Show that in case (iii) L is one-to-one (and only in that case) and L^{-1} is linear with matrix

$$\begin{pmatrix} d/\Delta & -b/\Delta \\ -c/\Delta & a/\Delta \end{pmatrix} .$$

4.7: Show that the sum and composition of two linear functions are linear. What are the matrix representations of these?

4.8: If $L : R^n \to R^m$ is linear and (1-1) then L^{-1} is linear on its domain.

4.9: Let $f : R^n \to R^m$ be such that

(i) $f(p+q) = f(p) + f(q)$, $\forall\ p,q \in R^n$ (additive)

(ii) f is continuous at 0 .

Prove that f is linear.

4.10: Let $f : R \to R^m$ be such that

$$f(\lambda x) = \lambda\, f(x) \qquad \text{(homogeneous)}$$

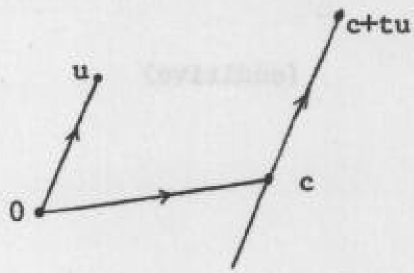
Notice that f is linear, i.e. $f(x) = f(1)\, x$. Show that homogeneity does not imply linearity for functions of more than one variable.

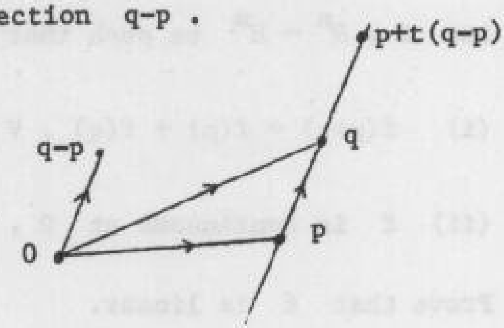
(Hint: Consider $f(x,y) = \dfrac{x^3+y^3}{x^2+y^2}$, $(x,y) \neq (0,0)$, $f(0,0) = 0$.)

See also Buck: Exercises pp. 227-229, 237-238.

Straight Lines and Curves.

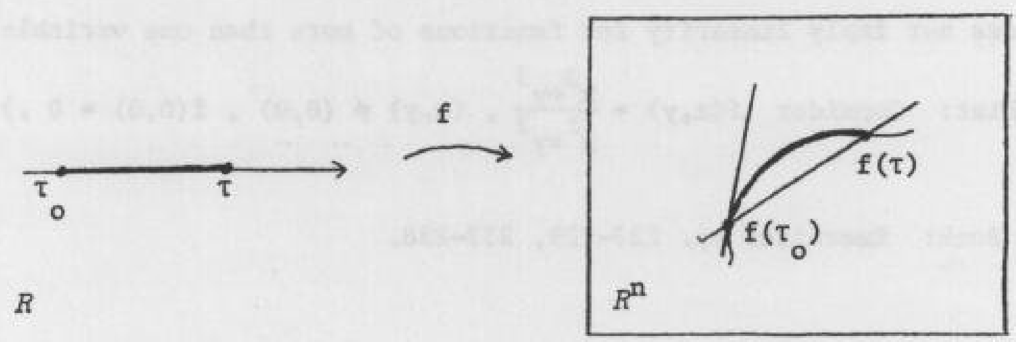
If c and u are points in R^n then $\{c+tu : t \in R\}$ is the straight line through c in the direction u. Recall that $\{p + t(q-p) : t \in R\}$ is the straight line through p and q ; it may also be considered the straight line through p in the direction q-p .





Notice that this representation of a line is not unique; you may replace u by any multiple λu $(\lambda \neq 0)$ and still get the same line.

If $f : R \to R^n$ then we say $\{f(t) : t \in R\}$ is a curve in R^n (a continuous curve if f is continuous).



The line through $f(\tau_o)$ and $f(\tau)$ $(f(\tau) \neq f(\tau_o))$ is $\{f(\tau_o) + t(f(\tau_o) - f(\tau)) : t \in R\}$ or equivalently

$$\{f(\tau_o) + t\,\frac{f(\tau) - f(\tau_o)}{\tau - \tau_o} : t \in R\} \quad ,$$

i.e. it is the line through $f(\tau_o)$ in the direction $\dfrac{f(\tau) - f(\tau_o)}{\tau - \tau_o}$. We therefore define the <u>tangent</u> to the curve at $f(\tau_o)$ to be $\{f(\tau_o) + t\, f'(\tau_o) : t \in R\}$ if $f'(\tau_o) = \lim\limits_{\tau\to\tau_o} \dfrac{f(\tau) - f(\tau_o)}{\tau - \tau_o}$ exists (this is the line through $f(\tau_o)$ in the direction $f'(\tau_o)$). Evidently

$$f'(\tau_o) = \lim_{\tau\to\tau_o} \frac{f(\tau)-f(\tau_o)}{\tau - \tau_o} = \lim_{\tau\to\tau_o} \left(\frac{f_1(\tau)-f_1(\tau_o)}{\tau - \tau_o}, \ldots, \frac{f_n(\tau)-f_n(\tau_o)}{\tau - \tau_o}\right)$$

$$= (f_1'(\tau_o), \ldots, f_n'(\tau_o)) .$$

It is also convenient sometimes to think of $f(t)$ as being the <u>position</u> of a particle at time t; then $f'(t) = (f_1'(t), \ldots, f_n'(t))$ is the <u>velocity vector</u> of the particle at time t.

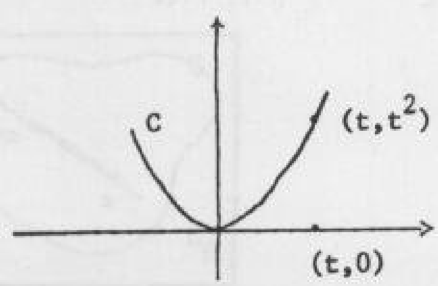
Notice also that the linear function $L(t) = t\, f'(\tau_o)$ is the best linear approximation to $f(\tau_o+t) - f(\tau_o)$ in a neighbourhood of $t = 0$ in the sense that

$$\lim_{t\to 0} \left|\frac{f(\tau_o+t)-f(\tau_o) - t\, f'(\tau_o)}{t}\right| = 0 .$$

<u>Example</u>:

The direction of the tangent to $\{(t,t^2) : t \in R\} = C$ in R^2 at the point $(0,0)$ is $(1,2t)\big|_{t=0} = (1,0)$. So the tangent line to C at $(0,0)$ is

$\{(0,0) + t(1,0) : t \in R\}$ or $\{(t,0) : t \in R\}$.



THE DIRECTIONAL DERIVATIVE AND THE DIFFERENTIAL

Definition: $f : R^n \to R^m$, domain $D \subset R^n$.

(i) If c is an interior point of D, $u \in R^n$, and

$$\lim_{t\to 0} \{f(c+tu) - f(c)\} \frac{1}{t} , \quad (t \in R)$$

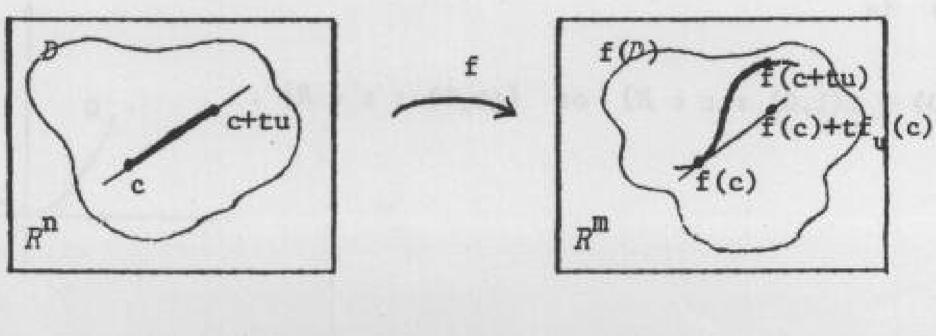
exists then this limit is called the directional derivative of f at c in the direction u and is denoted $f_u(c)$.

(ii) If $e_i = (0,\ldots,\underset{\substack{\uparrow \\ i\text{th entry}}}{1},\ldots,0)$, $f_{e_i}(c)$ is usually denoted $\frac{\partial f}{\partial x_i}(c)$, $i = 1,\ldots,n$ - the partial derivatives of f at c.

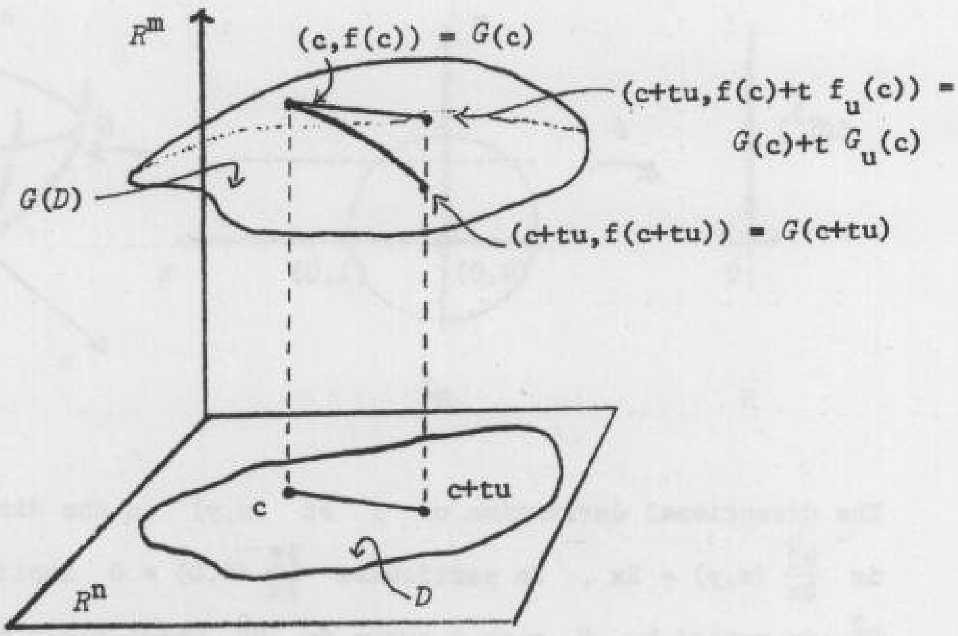
When we compute the directional derivative

$$f_u(c) = \lim_{t\to 0} \{f(c+tu) - f(c)\} \frac{1}{t}$$

we are in fact restricting our attention to the behaviour of the function f at c with respect to a straight line through c in the direction u. This straight line in R^n (at least the portion of it that lies in D) is mapped by f into a curve in $f(D) \subset R^m$. The vector $f_u(c)$ is the direction of the tangent to this curve in R^m.



It is also instructive to consider the graph $G(D) \subset R^{n+m}$ of f where $G : R^n \to R^{n+m}$ is defined by $G(p) = (p, f(p))$, $p \in D$.



$$G_u(c) = \lim_{t\to 0} [G(c+tu)-G(c)]\,\frac{1}{t} = \lim_{t\to 0} [(c+tu, f(c+tu))-(c,f(c))]\,\frac{1}{t}$$

$$= \lim_{t\to 0} (tu, f(c+tu)-f(c))\,\frac{1}{t} = (u, f_u(c)) \quad .$$

$G_u(c) = (u, f_u(c))$ is the direction (in R^{n+m}) of the tangent to the curve

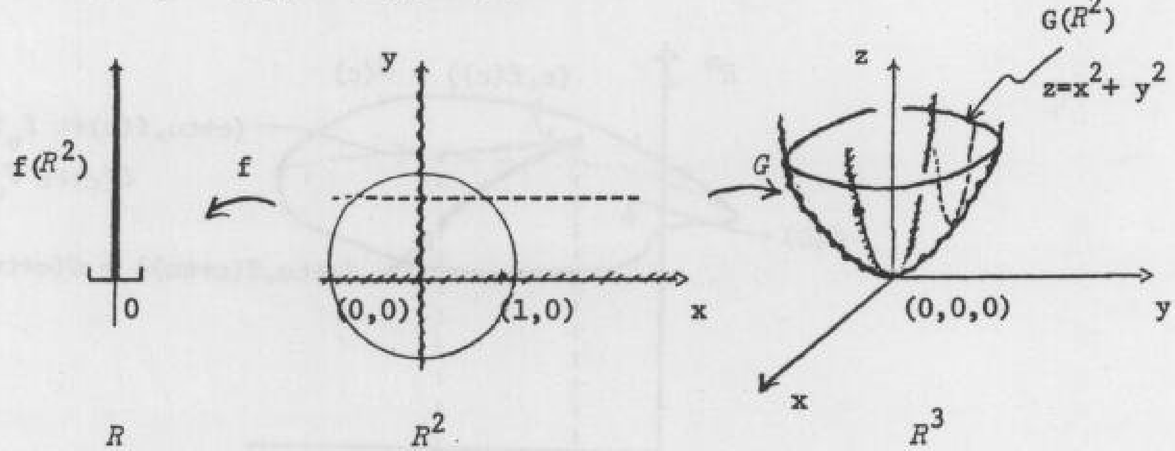
$$\{G(c+tu) : t \in R\} = \{(c+tu, f(c+tu) : t \in R\} \quad ,$$

at the point $(c, f(c))$.

Examples.

(1) $f : R^2 \to R$, $f(x,y) = x^2 + y^2$

$G : R^2 \to R^3$, $G(x,y) = (x,y,x^2+y^2)$



The directional derivative of f at (x,y) in the direction $e_1 = (1,0)$ is $\frac{\partial f}{\partial x}(x,y) = 2x$. In particular $\frac{\partial f}{\partial x}(0,0) = 0$ implies the x-axis in R^2 is mapped by G onto a curve in R^3 which has a horizontal tangent at $G(0,0) = (0,0,0)$.

(2) $f(x_1,x_2) = (x_1,x_2,x_1^2 + x_2^2)$

$c = (c_1,c_2)$, $u = (u_1,u_2)$

$$[f(c+tu)-f(c)]\frac{1}{t}$$

$$= [(c_1+tu_1,c_2+tu_2,(c_1+tu_1)^2+(c_2+tu_2)^2)-(c_1,c_2,c_1^2+c_2^2)]\frac{1}{t}$$

$$= (tu_1,tu_2,2tu_1c_1 + 2tu_2c_2 + t^2u_1^2 + t^2u_2^2)\frac{1}{t}$$

$$= (u_1,u_2,2u_1c_1 + 2u_2c_2 + tu_1^2 + tu_2^2)$$

$$\therefore \quad f_u(c) = \lim_{t\to 0} [f(c+tu) - f(c)] \frac{1}{t}$$

$$= (u_1, u_2, 2u_1c_1 + 2u_2c_2) \quad .$$

Notice that

$$\begin{bmatrix} u_1 \\ u_2 \\ 2u_1c_1 + 2u_2c_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2c_1 & 2c_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

so that $f_u(c)$ is a <u>linear function</u> of u .

(3) $f : R \to R$ such that $f'(c)$ exists.

$$[f(c+tu) - f(c)]/t = [f(c+tu) - f(c)] \frac{1}{tu} \; u$$

$$\lim_{t\to 0} \frac{f(c+tu)-f(c)}{t} = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h} \; u = f'(c) \; u$$

$$\text{i.e.} \quad f_u(c) = f'(c) \; u \quad .$$

Again $f_u(c)$ is linear in u (with matrix $[f'(c)]$). Now for the bad news:

(4) $f : R^2 \to R$, $f(x,y) = 0$, $y \neq x^2$

$f(x,y) = 1$, $y = x^2$, $(x,y) \neq (0,0)$

$f(0,0) = 0$.

Then $f_{(\alpha,\beta)}(0,0) = 0$, $\forall$ $(\alpha,\beta) \in R^2$. Thus $f_u(0,0)$ exists for every $u \in R^2$ and is linear in u but f is <u>not</u> continuous at $(0,0)$.

The following example shows that $f_u(c)$ does not have to be linear in u.

(5) $f : R^2 \to R$, $f(x,y) = \dfrac{x^2y}{x^3-y^2}$, $x^3 \neq y^2$

$f(x,y) = 0$, $x^3 = y^2$.

f is not continuous at $(0,0)$ but $f_{(\mu,\nu)}(0,0)$ exists for each $(\mu,\nu) \in R^2$ and

$$f_{(\mu,\nu)}(0,0) = -\frac{\mu^2}{\nu} \ , \ \text{if} \ \nu \neq 0$$

$$f_{(\mu,\nu)}(0,0) = 0 \ , \ \text{if} \ \nu = 0 \ .$$

Notice that $f_{(\mu,\nu)}(0,0)$ is not linear in (μ,ν) .

Remark: Notice from the preceding examples that, when f is a "nice" function at c , $f_u(c)$ is linear in u but in general need not be linear even if it exists for all u .

Observe however that all the $f_u(c)$ above, while some are not additive, are homogeneous in u , i,e. $f_{\lambda u}(c) = \lambda\, f_u(c)$ for all $\lambda \in R$. This is always true when $f_u(c)$ exists.

Exercises:

4.11: Prove that if $f_u(c)$ exists then $f_{\lambda u}(c)$ exists for all $\lambda \in R$ and $f_{\lambda u}(c) = \lambda\, f_u(c)$.

4.12: Prove that the expressions given for the directional derivatives in Examples (4), (5) above are correct.

4.13: Check that the following partial derivatives are correct:

(a) $f(x,y) = x^2 + y^3$,

$$\frac{\partial f}{\partial x}(x,y) = 2x \quad , \quad \frac{\partial f}{\partial y}(x,y) = 3y^2 \ .$$

(b) $f(x,y) = (\sin(xy) \ , \ e^x \ , \ \cos y)$

$$\frac{\partial f}{\partial x} = (y \cos(xy) \ , \ e^x \ , \ 0) \quad , \quad \frac{\partial f}{\partial y} = (x \cos(xy) \ , \ 0 \ , \ -\sin y) \ .$$

4.14: $f : R^2 \to R^3$, $f(x,y) = (x^2+y^3 \ , \ \sin y \ , \ e^x)$.

Prove that $f_{(\alpha,\beta)}(x,y)$ exists for each $(x,y) \in R^2$ and each direction (α,β) . Show also that $f_{(\alpha,\beta)}(x,y)$ is linear in (α,β) with matrix

$$\begin{bmatrix} 2x & 3y^2 \\ 0 & \cos y \\ e^x & 0 \end{bmatrix} .$$

<u>Motivation</u>.

Although the directional derivative performs some of the tasks that the derivative did in determining the properties of functions we have seen that it has embarassing shortcomings. For example a function may be discontinuous at a point where all directional derivatives exist. Therefore, to be useful, a more restrictive type of differentiation must be introduced in order that the role of differentiation may be extended to higher dimensions. To this end we reformulate the definition of the derivative for functions of one variable.

$$f : R \to R \ , \quad f \text{ is differentiable at } c \text{ if } \exists \ f'(c) \ni$$

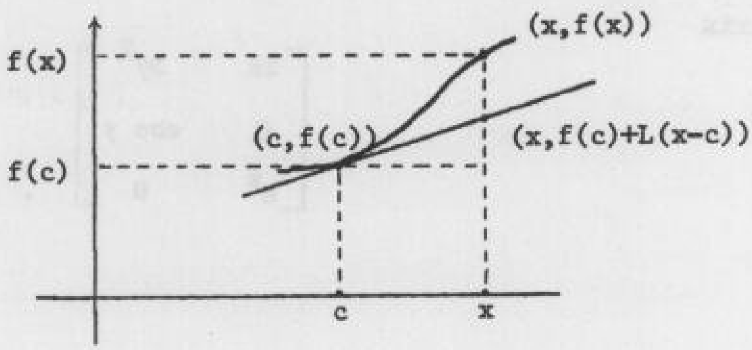
$$\lim_{x \to c} \frac{f(x)-f(c)}{x-c} = f'(c)$$

$$\text{i.e.} \quad \lim_{x \to c} \left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| = \lim_{x \to c} \left| \frac{f(x)-f(c)-f'(c)(x-c)}{|x-c|} \right| = 0 \ .$$

Thus $f : R \to R$ is differentiable at c if and only if there exists a linear function $L : R \to R$ such that

$$(*) \qquad \lim_{x \to c} \frac{|f(x)-f(c)-L(x-c)|}{|x-c|} = 0$$

and L is given by $L(u) = f'(c)u$, $\forall u \in R$. We can consider L to be the best linear approximation in the sense of (*) to $f(x) - f(c)$ in a neighbourhood of c .



Definition: Let $f : R^n \to R^m$ with c an interior point of D , the domain of f . Suppose that $L : R^n \to R^m$ is <u>linear</u> and

$$\lim_{p \to c} \frac{|f(p)-f(c) - L(p-c)|}{|p-c|} = 0 \ .$$

Then we say f is <u>differentiable</u> at c and L is called the differential of f at c, and denoted $\mathcal{D}\, f(c)$; i.e. $L(u) = \mathcal{D}\, f(c)(u)$, $\forall\, u \in \mathcal{R}^n$.

<u>Example</u>. $f : \mathcal{R} \to \mathcal{R}$, $f(x) = x^3$, $f'(c) = 3c^2$

$$L(u) = 3c^2 u \quad , \quad \forall\, u \in \mathcal{R}^n$$

$$\begin{aligned} |f(p)-f(c)-L(p-c)| &= |p^3 - c^3 - 3c^2(p-c)| \\ &= |p-c||p^2 + pc + c^2 - 3c^2| \\ &= |p-c||p^2 + pc - 2c^2| \end{aligned}$$

and

$$\lim_{p \to c} |p^2 + pc - 2c^2| = 0 \quad .$$

<u>Theorem 4.4</u>: $f : \mathcal{R}^n \to \mathcal{R}^m$

$$\mathcal{D}\, f(c)(u) = L(u) \iff \mathcal{D}\, f_i(c)(u) = L_i(u) \, , \; \forall\, u \in \mathcal{R}^n \; , \; i = 1,\ldots,m$$

where

$$f(p) = (f_1(p),\ldots,f_m(p)) \; , \; L(u) = (L_1(u),\ldots,L_m(n)) \quad .$$

<u>Proof</u>:

$$\lim_{p \to c} \frac{|f(p)-f(c) - L(p-c)|}{|p-c|} = 0$$

$$\iff \lim_{p \to c} \frac{|f_i(p) - f_i(c) - L_i(p-c)|}{|p - c|} = 0 \quad , \; i = 1,\ldots,m \quad .$$

□

<u>Theorem 4.5</u>: $f : R^n \to R^m$. If f is differentiable at c then:

(i) $f_u(c)$ exists for each $u \in R^n$ and $D\, f(c)(u) = f_u(c)$.

(ii) The matrix of the linear function $D\, f(c)$ is

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(c) & , & \cdots & , & \frac{\partial f_1}{\partial x_n}(c) \\ \vdots & & & & \vdots \\ \frac{\partial f_m}{\partial x_1}(c) & , & \cdots & , & \frac{\partial f_m}{\partial x_n}(c) \end{bmatrix}$$

This is the <u>Jacobian</u> <u>Matrix</u> of f at c, which is usually denoted $[\frac{\partial f}{\partial x}(c)]$ or $f'(c)$.

<u>Notes</u>:

(1) We are assuming $f(p) = (f_1(p),\ldots,f_m(p))$, $p = (x_1,\ldots,x_n)$.

(2) In the case $m = n$ the determinant $\det f'(c)$ is called the Jacobian of f at c and is often denoted

$$J_f(c) \quad \text{or} \quad \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(c) .$$

(3) Part (ii) of the Theorem means that if $u = (u_1,\ldots,u_n) \in R^n$ then, $i = 1,\ldots,m$,

$$L_i(u) = D\, f_i(c)(u) = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(c)\, u_j \quad \text{(cf. Theorem 4.1)}$$

Proof of Theorem 4.5:

(i) $\mathcal{D}\, f(c)\ \exists \Rightarrow$ If $u \in R^n$, $u \neq 0$ then

$$\lim_{t\to 0} \frac{|f(c+tu) - f(c) - \mathcal{D}\, f(c)(tu)|}{|tu|} = 0$$

$$\therefore \quad \lim_{t\to 0} \left|\frac{f(c+tu) - f(c)}{t} - \mathcal{D}\, f(c)(u)\right| = 0$$

since, by linearity, $\mathcal{D}\, f(c)(tu) = t\, \mathcal{D}\, f(c)(u)$

$$\therefore \quad \lim_{t\to 0} \frac{f(c+tu)-f(c)}{t} = \mathcal{D}\, f(c)(u) \ .$$

i.e. $f_u(c)$ exists and equals $\mathcal{D}\, f(c)(u)$.

(ii) Recall from Theorem 4.1 that the matrix of L is $[c_{ij}] = [L_i(e_j)]$. Here

$$\begin{aligned} L_i(e_j) &= \mathcal{D}\, f_i(c)(e_j) \\ &= f_{i_{e_j}}(c) \text{ , by part (i)} \\ &= \frac{\partial f_i}{\partial x_j}(c) \quad , \quad i = 1,\dots,m \ , \ j = 1,\dots,n \ . \end{aligned}$$

Corollary 4.5.1: If the differential exists it is unique.

Proof: $f_u(c)$ is unique (uniqueness of limits) for each $u \in R$. Alternatively the partials $\frac{\partial f_i}{\partial x_j}(c)$ are unique and hence the matrix $f'(c) = [\frac{\partial f_i}{\partial x_j}(c)]$

is unique.

<u>Theorem 4.6</u>: If f is differentiable at c then f is continuous at c.

<u>Proof</u>: Let $L = \mathcal{D} f(c)$. From the definition of $\mathcal{D} f(c)$, if $\varepsilon > 0$ there is a $\delta(\varepsilon) > 0$ such that if $0 < |p-c| < \delta(\varepsilon)$ then $\dfrac{|f(p)-f(c) - L(p-c)|}{|p-c|} < \varepsilon$, i.e. $|f(p) - f(c) - L(p-c)| < \varepsilon\, |p-c|$. In particular, with $\varepsilon = 1$, if $0 < |p-c| < \delta(1)$ then

$$
\begin{aligned}
|f(p)-f(c)| &\le |p-c| + |L(p-c)| \\
&\le (1+M)\ |p-c| \qquad \text{(L is linear, Theorem 4.2)}
\end{aligned}
$$

$$\therefore \quad \lim_{p\to c} f(p) = f(c) \ . \qquad \square$$

<u>Theorem 4.7</u> (Important): $f : R^n \to R^m$. If the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$, $i = 1,\ldots,m$, $j = 1,\ldots,n$, <u>exist</u> in a neighbourhood of c and are continuous at c then f is differentiable at c.

<u>Proof</u>: By Theorem 4.4 it is sufficient to prove the case $m = 1$ (i.e. f is a real-valued function). By the hypothesis, for each $\varepsilon > 0$, there exists a $\delta > 0$ such that if $|p-c| < \delta$ then

$$\frac{\partial f}{\partial x_i}(p) \ \text{ exists and } \ \left|\frac{\partial f}{\partial x_i}(p) - \frac{\partial f}{\partial x_i}(c)\right| < \varepsilon \ .$$

Let

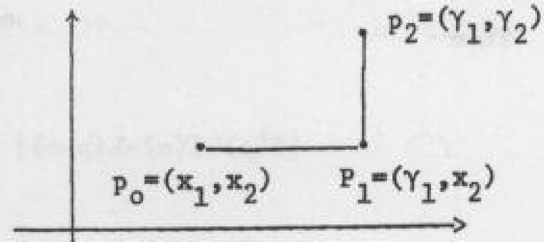
$$p = (x_1,\ldots,x_n) \ , \ c = (\gamma_1,\ldots,\gamma_n) \ , \ |p-c| < \delta \ .$$

Define $p = (x_1,\ldots,x_n) = p_0$

$(\gamma_1,x_2,\ldots,x_n) = p_1$

$(\gamma_1,\gamma_2,x_3,\ldots,x_n) = p_2$

$c = (\gamma_1,\gamma_2,\ldots,\gamma_n) = p_n$



e.g. in R^2

(1) $$|p_k-c| \leq |p-c| < \delta$$

(2) $$f(p)-f(c) = f(p_o)-f(p_n) = \sum_{k=1}^{n} [f(p_{k-1}) - f(p_k)] \quad .$$

Now on each line segment between p_{k-1} and p_k, $k = 1,\ldots,n$ we are really considering a real-valued differentiable function of a real variable so we may use the Mean Value Theorem.

(3) $$f(p_{k-1}) - f(p_k) = \frac{\partial f}{\partial x_k}(\overline{p}_k)(x_k-\gamma_k)$$

where $\overline{p}_k$ is in the line-segment from p_{k-1} to p_k. Clearly $\overline{p}_k \in \{p : |p-c| < \delta\}$ by (1), since this set is convex. From (2), (3)

$$f(p) - f(c) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\overline{p}_k)(x_k-\gamma_k) \quad .$$

If $u = (\mu_1,\ldots,\mu_n)$ let $L(u)$ be defined by

$$L(u) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(c)\, \mu_k \quad .$$

Then

$$|f(p)-f(c)-L(p-c)| = \left|\sum_{k=1}^{n} \left[\frac{\partial f}{\partial x_k}(\bar{p}_k) - \frac{\partial f}{\partial x_k}(c)\right](x_k-\gamma_k)\right|$$

$$\leq \left[\sum_{k=1}^{n}\left(\frac{\partial f}{\partial x_k}(\bar{p}_k) - \frac{\partial f}{\partial x_k}(c)\right)^2\right]^{1/2}\left[\sum_{k=1}^{n}(x_k-\gamma_k)^2\right]^{1/2} \quad \text{(CBS)}$$

$$\leq \sqrt{n}\,\varepsilon\,|p-c| \quad \text{since} \quad |\bar{p}_k-c| < \delta \quad \text{if } |p-c| < \delta$$

$$\text{i.e.} \quad \frac{|f(p)-f(c)-L(p-c)|}{|p-c|} < \sqrt{n}\,\varepsilon\ , \quad \text{if } |p-c| < \delta\ .$$

Therefore $\mathcal{D}\, f(c)$ exists and $\mathcal{D}\, f(c) = L$. □

Please note that the Theorem gives a sufficient condition only for a function to be differentiable at a point.

Examples.

(1) $f : R^2 \to R^3$, $f(x_1,x_2) = (x_1,x_2,x_1^2+x_2^2)$.

The partial derivatives are continuous on R^2 so f is differentiable and $\mathcal{D}\, f(x_1,x_2)$ has matrix

$$f'(x_1,x_2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2x_1 & 2x_2 \end{bmatrix} .$$

For example if $L = \mathcal{D}\, f(0,0)$, $L : R^2 \to R^3$ and

$$\begin{bmatrix} L_1\ u \\ L_2\ u \\ L_3\ u \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ 0 \end{bmatrix} \quad \text{where } u = (u_1,u_2)$$

i.e. $\mathcal{D}\,f(0,0)(u_1,u_2) = (u_1,u_2,0)$ for each $(u_1,u_2) \in R^2$. Check for yourself that

$$\mathcal{D}\,f(0,1)(u_1,u_2) = (u_1,u_2,2u_2)$$

and

$$\mathcal{D}\,f(1,1)(u_1,u_2) = (u_1,u_2,2u_1+2u_2) \quad .$$

(2) $f : R \to R^3 \quad f(t) = (\cos t, \sin t, t)$

$$f'(t) = \begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix} \quad ; \qquad \begin{array}{l} \mathcal{D}\,f(0)(u) = (0,u,u) \\ \mathcal{D}\,f(\frac{\pi}{2})(u) = (-u,0,u) \\ \text{for each } \; u \in R \; . \end{array}$$

(3) $f : R^3 \to R$, $f(x,y,z) = x^2y + z$

$f'(x,y,z) = [2xy, x^2, 1]$

$$\mathcal{D}\,f(0,0,0)(u_1,u_2,u_3) = u_3$$

$$\mathcal{D}\,f(1,1,0)(u_1,u_2,u_3) = 2u_1 + u_2 + u_3 \quad .$$

(4) $f : R^2 \to R^2 \quad f(x,y) = (x+y,(x+y)^2)$

$$f'(x,y) = \begin{bmatrix} 1 & 1 \\ 2(x+y) & 2(x+y) \end{bmatrix}$$

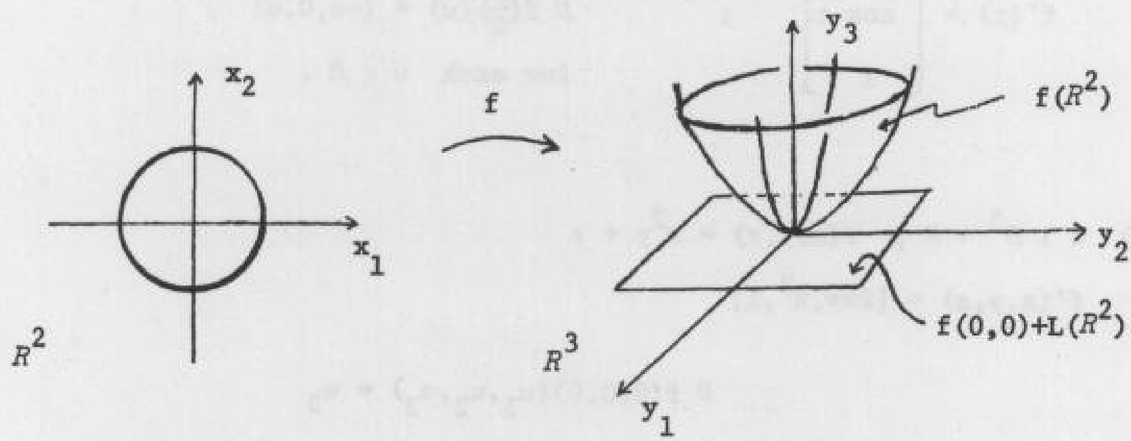
$$\mathcal{D}\,f(x_o,y_o)(u,v) = (u+v, 2(x_o+y_o)(u+v)) \quad .$$

<u>Interpretation</u>: $f : R^n \to R^m$, $L : R^n \to R^m$ (linear).

$$\lim_{n\to 0} \frac{|f(c+u)-f(c)-L(u)|}{|u|} = 0 \quad \text{if} \quad L = D\,f(c) \quad ,$$

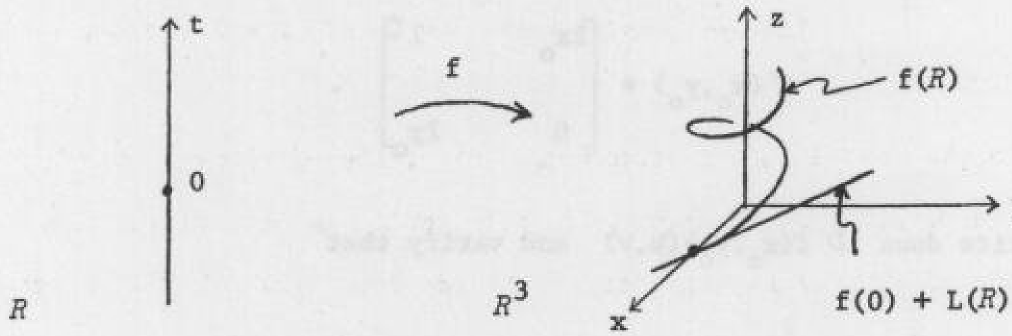
so $L(u)$ is the best linear approximation to $f(c+u)-f(c)$ near $u = 0$, equivalently $f(c) + L(u)$ is the best affine approximation to $f(c+u)$ near $u = 0$. You can think of this as saying that the affine set $f(c) + L(R^n)$ is <u>tangent</u> to $f(R^n)$ at $f(c)$.

In Example 1 above $f(x_1,x_2) = (x_1,x_2,x_1^2+x_2^2)$.



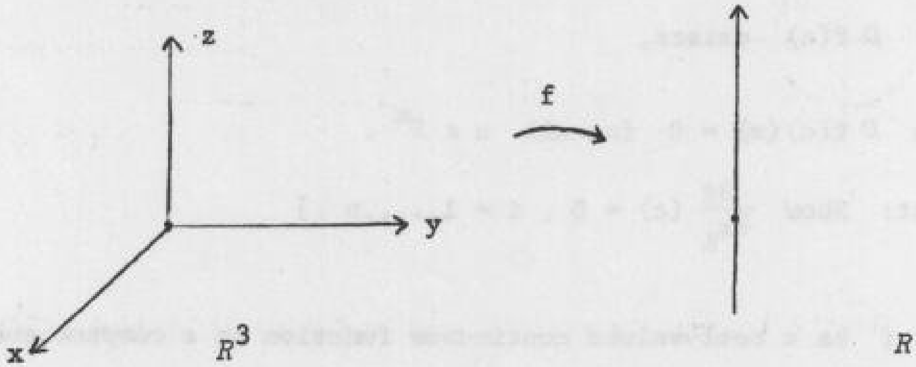
If $L = D\,f(0,0)$ the tangent at $f(0,0) = (0,0,0)$ is $f(0,0) + L(R^2)$, the set of all vectors of the form $(0,0,0) + (u_1,u_2,0)$, i.e. $\{(u_1,u_2,0) ; (u_1,u_2) \in R^2\}$ which is the plane $y_3 = 0$. Similarly the tangent at $f(0,1) = (0,1,1)$ is $f(0,1) + L(R^2) = \{(0,1,1) + (u_1,u_2,2u_2) : (u_1,u_2) \in R^2\} = \{(u_1,u_2+1,2u_2+1) : (u_1,u_2) \in R^2\}$ (here $L = D\,f(0,1)$). Thus the tangent is the plane $y_3-1 = 2(y_2-1)$. Notice that the tangent at every point on $f(R^2)$ is a plane since rank $f'(x_1,x_2) = 2$ for all $(x_1,x_2) \in R^2$.

In Example 2, $f(t) = (\cos t, \sin t, t)$



The tangent at $f(0) = (1,0,0)$ is $f(0) + L(R)$ $(L = D\, f(0))$ i.e., $\{(1,0,0) + (0,u,u) : u \in R\} = \{(1,u,u) : u \in R\}$ a straight line through $(1,0,0)$ and $(1,1,1)$.

In Example 3, $f(x,y,z) = x^2y + z$.



The tangent at $f(0,0,0) = 0$, if $L = D\, f(0,0,0)$, is $f(0,0,0) + L(R^3) = \{u_3 : (u_1,u_2,u_3) \in R^3\}$ i.e. the whole set R. The picture is not very informative in this case.

Exercises:

4.15: In Example 4 above sketch the range $f(R^2)$ (it is a curve in R^2). Find the tangent at one or two points of the range.

4.16: $f : R^2 \to R^2$, $f(x,y) = (x^2+y,y^2)$. Check that

$$f'(x_o,y_o) = \begin{bmatrix} 2x_o & 1 \\ 0 & 2y_o \end{bmatrix} .$$

Write down $D\, f(x_o,y_o)(\mu,\nu)$ and verify that

$$\lim_{(\mu,\nu) \to (0,0)} \frac{|f(\mu,\nu)-f(0,0) - D\, f(0,0)(\mu,\nu)|}{|(\mu,\nu)|} = 0 .$$

4.17: If $f : R^n \to R$ is such that

(i) f has an interior relative minimum (maximum) at c ,

(ii) $D\, f(c)$ exists,

then $D\, f(c)(u) = 0$ for all $u \in R^n$.

[Hint: Show $\frac{\partial f}{\partial x_i}(c) = 0$, $i = 1,\ldots,n$.]

4.18: Let f be a real-valued continuous function on a compact subset K of R^n such that

(i) $f(p) = 0$ if $p \in \partial K$, the boundary of K ,

(ii) $D\, f(p)$ exists if $p \in K^o \neq \phi$ (the interior of K).

Show that there is a point $p_o \in K^o$ such that

$$D\, f(p_o)(u) = 0 , \quad \text{for all } u \in R^n .$$

[This is a generalization of Rolle's Theorem.]

4.19: For each of the following functions write down the Jacobian matrix at the point indicated:

(a) $f(x,y) = 3x^2y - xy^3 + 2$; $(1,2)$.

(b) $h(u,v) = (u \sin uv\ ,\ v \cos uv)$; $(\frac{\pi}{4}, \frac{\pi}{2})$.

(c) $g(x,y,z) = (x^2yz, 3xz^2)$; $(1,2,-1)$.

4.20: Let f be a real-valued function on an open set U in R^2 such that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist and are bounded on U. Prove that f is continuous on U.

[Hint: $f(x,y)-f(x_o,y_o) = f(x,y) - f(x_o,y) + f(x_o,y) - f(x_o,y_o)$, the old polygon-in-a-convex-set trick.]

4.21: Let f be a real-valued function on an open connected set U in R^2 such that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist and are zero at each point of U. Prove that f is constant.

4.22: If f is a real-valued differentiable function on R^n and $c \in R^n$ then the vector $\text{grad}\, f(c) = \nabla f(c) = (\frac{\partial f}{\partial x_1}(c), \ldots, \frac{\partial f}{\partial x_n}(c))$ is called the gradient of f at c. Show that

$$f_u(c) = D\, f(c)(u) = \nabla f(c) \cdot u \text{ , for each } u \in R^n .$$

Deduce that the largest value of $f_u(c)$, if $|u| = 1$, is $|\nabla f(c)|$ and, if $\nabla f \neq 0$, this value is attained when $u = \nabla f(c)/|\nabla f(c)|$. This means that the direction of maximum rate of increase of f at c is the direction of the gradient vector. What was Bunyakowski-Schwarz's first name?

4.23: Sketch the surface $\{(x,y,xy) : (x,y) \in R^2\}$ in R^3 (i.e. $z = xy$) and show that the tangent to this surface at the point $(1,1,1)$ is $(1,1,1) + \{(u,v,u+v) : (u,v) \in R^2\}$ (i.e. $z+1 = x+y$).

4.24: If $L : R^n \to R^m$ is linear then the differential of L exists and equals L at each point in R^n.

Differentiation Rules:

Theorem 4.8: $\phi,\psi : R^n \to R^m$, both differentiable at $c \in R^n$.

(i) If $h = \alpha\phi + \beta\psi$ $(\alpha,\beta \in R)$ then h is differentiable at c and

$$D\,h(c)(u) = \alpha\, D\phi(c)(u) + \beta\, D\psi(c)(u)\ , \text{ for each } u \in R^n .$$

(ii) If $k = \phi\cdot\psi$ (so $k : R^n \to R$) then k is differentiable at c and

$$D\,k(c)(u) = \phi(c) \cdot D\psi(c)(u) + \psi(c) \cdot D\phi(c)(u),, \text{ for each } u \in R^n .$$

Proof:

(i) Let $L_\phi = D\,\phi(c)$, $L_\psi = D\,\psi(c)$ and consider the function $L : R^n \to R^m$, $L(u) = \alpha\, L_\phi(u) + \beta\, L_\psi(u)$. Clearly L is linear and

$$\left|\frac{h(p)-h(c)-L(p-c)}{|p-c|}\right|$$

$$= |\alpha\phi(p) + \beta\psi(p) - \alpha\phi(c) - \beta\psi(c) - \alpha\, L_\phi(p-c) - \beta\, L_\psi(p-c)| / |p-c|$$

$$\leq |\alpha| \frac{|\phi(p)-\phi(c)-L_\phi(p-c)|}{|p-c|} + |\beta| \frac{|\psi(p)-\psi(c)-L_\psi(p-c)|}{|p-c|} ,$$

by the triangle inequality. Both terms in this last expression have limit 0 as $p \to c$; therefore $D\,h(c)$ exists $= L$.

(ii) Let $L : R^n \to R$ be defined by

$$L(u) = \phi(c) \cdot L_\psi(u) + \psi(c) \cdot L_\phi(u) \quad , \quad u \in R^n .$$

Then

$$\frac{|k(p)-k(c)-L(p-c)|}{|p-c|}$$

$$= |\phi(p)\cdot\psi(p) - \phi(c)\,\psi(c) - \phi(c)\cdot L_\psi(p-c) - \psi(c)\cdot L_\phi(p-c)|/|p-c|$$

$$= |\phi(p)\cdot\{\psi(p) - \psi(c) - L_\psi(p-c)\} + \psi(c)\cdot\{\phi(p) - \phi(c) - L_\phi(p-c)\}$$

$$+ \phi(p)\cdot L_\psi(p-c) - \phi(c)\cdot L_\psi(p-c)|/|p-c|$$

$$\leq |\phi(p)| \frac{|\psi(p)-\psi(c)-L_\psi(p-c)|}{|p-c|} + |\psi(c)| \frac{|\phi(p)-\phi(c)-L_\phi(p-c)|}{|p-c|} + |\phi(p)-\phi(c)| \frac{|L_\psi(p-c)|}{|p-c|}$$

$\to 0$ as $p \to c$. In the last step we have used the CBS inequality several times, the fact that $\frac{|L_\psi(p-c)|}{|p-c|}$ is bounded (Theorem 4.2, since L_ψ is linear) and the fact that ϕ is continuous at c (Theorem 4.6).

Theorem 4.9. (The Chain Rule; very important): Suppose $\phi : R^n \to R^m$, $\psi : R^m \to R^\ell$ and

(i) ϕ is differentiable at c $(\in R^n)$,

(ii) ψ is differentiable at $b = \phi(c)$ $(\in R^m)$.

Then $f = \psi \circ \phi$ [i.e. $f(p) = \psi(\phi(p))$] is differentiable at c and

$$D\,f(c) = D\,\psi(b) \circ D\,\phi(c) \quad .$$

In particular the Jacobian matrices f' , ψ' , ϕ' satisfy

$$f'(c) = \psi'(b)\,\phi'(c) \quad .$$

i.e.,

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} , & \cdots , & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_\ell}{\partial x_1} , & \cdots , & \frac{\partial f_\ell}{\partial x_n} \end{bmatrix}_c = \begin{bmatrix} \frac{\partial \psi_1}{\partial y_1} , & \cdots , & \frac{\partial \psi_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial \psi_\ell}{\partial y_1} , & \cdots , & \frac{\partial \psi_\ell}{\partial y_m} \end{bmatrix}_b \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} , & \cdots , & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial \phi_m}{\partial x_1} , & \cdots , & \frac{\partial \phi_m}{\partial x_n} \end{bmatrix}_c$$

i.e.,

$$\frac{\partial f_i}{\partial x_j}(c) = \sum_{k=1}^{m} \frac{\partial \psi_i}{\partial y_k}(b)\,\frac{\partial \phi_k}{\partial x_j}(c) \quad , \quad i=1,\ldots,\ell \ ; \ j=1,\ldots,m \ .$$

This may be stated more conveniently but less precisely as follows: If $f = f(y_1,\ldots,y_m)$, $y_k = y_k(x_1,\ldots,x_n)$, $k = 1,\ldots,m$ then

$$\frac{\partial f}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial f}{\partial y_k}\,\frac{\partial y_k}{\partial x_j}$$

provided all the functions involved are "smooth".

Examples:

(1) $n = m = \ell = 1$.

$$f(t) = \psi(\phi(t)) \quad .$$

If $\phi'(t_o)$ and $\psi'(\phi(t_o)$ exist then $f'(t_o)$ exists and

$$f'(t_o) = \psi'(\phi(t_o))\ \phi'(t_o) \ ,$$

less precisely,

$$\frac{df}{dt} = \frac{d\psi}{d\phi}\ \frac{d\phi}{dt} \quad .$$

(2) $n = m = 2$, $\ell = 1$.

$$f(x,y) = \psi(u,v) \quad \text{where} \quad u = u(x,y) \ , \ v = v(x,y)$$

i.e. $f = \psi \circ \phi$ where $\phi(x,y) = (u(x,y),v(x,y))$.

If $D\ \phi(x_o,y_o)$ and $D\ \psi(u_o,v_o)$ exist, where $u_o = u(x_o,y_o)$, $v_o = v(x_o,y_o)$, then $D\ f(x_o,y_o)$ exists and

$$f'(x_o,y_o) = \psi'(u_o,v_o)\ \phi'(x_o,y_o) \ .$$

i.e.

$$[\frac{\partial f}{\partial x} \ , \ \frac{\partial f}{\partial y}]_{(x_o,y_o)} = [\frac{\partial \psi}{\partial u} \ , \ \frac{\partial \psi}{\partial v}]_{(u_o,v_o)} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}_{(x_o,y_o)}$$

i.e.,

$$\frac{\partial f}{\partial x} = \frac{\partial \psi}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial \psi}{\partial v}\frac{\partial v}{\partial x}$$

$$\frac{\partial f}{\partial y} = \frac{\partial \psi}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial \psi}{\partial v}\frac{\partial v}{\partial y}$$

e.g. if $h(r,\theta) = g(u,v)$, where $u = r\cos\theta$, $v = r\sin\theta$

$$\frac{\partial h}{\partial r} = \frac{\partial g}{\partial u}\cos\theta + \frac{\partial g}{\partial v}\sin\theta ,$$

$$\frac{\partial h}{\partial \theta} = \frac{\partial g}{\partial u}(-r\sin\theta) + \frac{\partial g}{\partial v} r\cos\theta .$$

(3) $n = 1$, $m = 2$, $\ell = 1$.

$$F(t) = h(x,y) \text{ where } x = r(t) , y = s(t)$$

$$[\frac{dF}{dt}] = [\frac{\partial h}{\partial x} , \frac{\partial h}{\partial y}] \begin{bmatrix} \frac{dx}{dt} \\ \\ \frac{dy}{dt} \end{bmatrix} ,$$

i.e.,

$$\frac{dF}{dt} = \frac{\partial h}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{dy}{dt} .$$

(4) Suppose a particle's position (x,y,z) in space is given at time t by $x = \cos t$, $y = \sin t$, $z = t$ (it is moving on a helix), and the temperature at any point (x,y,z) is given by $T(x,y,z) = x^2 + y^2 + z^2$. If $H(t)$ is the temperature of the particle at time t find $\frac{dH}{dt}$.

In this case $H = T \circ f$ where

$$T(x,y,z) = x^2 + y^2 + z^2 , \quad f(t) = (\cos t, \sin t, t) .$$

From the Chain Rule $H'(t) = T'(x,y,z) f'(t)$

$$[\frac{dH}{dt}] = [2x, 2y, 2z] \begin{bmatrix} -\sin t \\ \cos t \\ 1 \end{bmatrix}$$

$$\frac{dH}{dt} = -2x \sin t + 2y \cos t + 2z = 2t \quad .$$

In this case it is in fact easier to find $\frac{dH}{dt}$ directly

$$H(t) = T(x,y,z) = x^2 + y^2 + z^2 = \cos^2 t + \sin^2 t + t^2 = 1 + t^2$$

so $\frac{dH}{dt} = 2t$.

<u>Proof of Theorem 4.9</u>: Let $L_\phi = D\,\phi(c)$, $L_\psi = D\psi(b)$, $b = \phi(c)$. We make two assertions

(1) $\lim\limits_{p \to c} |\psi(\phi(p)) - \psi(\phi(c)) - L_\psi(\phi(p)-\phi(c))| / |p-c| = 0$

(2) $\lim\limits_{p \to c} |L_\psi(\phi(p)-\phi(c)-L_\phi(p-c))| / |p-c| = 0$.

Now, if $f = \psi \circ \phi$,

$$\frac{|f(p)-f(c)-L_\psi \circ L_\phi (p-c)|}{|p-c|}$$

$$= |\psi(\phi(p))-\psi(\phi(c))-L_\psi(\phi(p)-\phi(c)) + L_\psi(\phi(p)-\phi(c)-L_\phi(p-c))| / |p-c|$$

$$\leq \frac{|\psi(\phi(p))-\psi(\phi(c))-L_\psi(\phi(p)-\phi(c))|}{|p-c|} + \frac{|L_\psi(\phi(p)-\phi(c)-L_\phi(p-c))|}{|p-c|}$$

$\to 0$ as $p \to c$ by (1) and (2). Since $L_\psi \circ L_\phi$ is linear $D\,f(c)$ exists and equals $L_\psi \circ L_\phi$.

Proof of assertion (1): Let $\varepsilon > 0$. Since ϕ is differentiable at c there is a neighbourhood U of c and a constant $K > 0$ such that if $p \in U$ then

$$|\phi(p)-\phi(c)| \leqslant K|p-c| \tag{3}$$

(see the proof of Theorem 4.6). L_ψ is the differential of ψ at $b = \phi(c)$ so $\exists\ \delta > 0 \ni |q-b| < \delta \Rightarrow$

$$|\psi(q)-\psi(b) - L_\psi(q-b)| \leq \frac{\varepsilon}{K} |q-b| \quad . \tag{4}$$

Thus, from (3) and (4), if p is sufficiently close to c,

$$|\psi(\phi(p))-\psi(\phi(c)) - L_\psi(\phi(p)-\phi(c))| \leq \frac{\varepsilon}{K} |\phi(p)-\phi(c)|$$

$$\leqslant \varepsilon|p-c| \quad .$$

Therefore assertion (1) is true.

Proof of assertion (2): Since L_ψ is linear there is a constant $M > 0$ such that $|L_\psi(u)| \leq M|u|$, for all $u \in R^m$. Thus

$$\frac{|L_\psi(\phi(p)-\phi(c) - L_\phi(p-c))|}{|p-c|} \leq \frac{M|\phi(p)-\phi(c) - L_\phi(p-c)|}{|p-c|}$$

$\to 0$, as $p \to c$, by the definition of L_ϕ. □

The Chain Rule in higher dimensions is even more interesting than in one dimension; for example it includes the rules for differentiating sums and products as special cases. Thus Theorem 4.8 is a corollary to Theorem 4.9.

This can be seen by considering the functions

$$F : R^n \to R^{2m} \quad , \quad F(p) = (\phi(p), \psi(p))$$

$$G : R^{2m} \to R^m \quad , \quad G(q_1, q_2) = \alpha q_1 + \beta q_2 \quad , \quad q_1, q_2 \in R^m$$

$$H : R^{2m} \to R \quad , \quad H(q_1, q_2) = q_1 \cdot q_2$$

so $G \circ F = \alpha\phi + \beta\psi$ and $H \circ F = \phi \cdot \psi$.

Theorem 4.10. (Mean-Value Theorem): $f : R^n \to R$. If $a,b \in R^n$ and f is differentiable at each point of the line segment S between a and b then there is a point $c \in S$, $c \neq a,b$ such that

$$f(b)-f(a) = \mathcal{D}\, f(c)(b-a) \quad ,$$

i.e. $f(b)-f(a) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(c)(\beta_j - \alpha_j)$, where $a = (\alpha_1, \ldots, \alpha_n)$, $b = (\beta_1, \ldots, \beta_n)$.

In the notation of Exercise 4.22 this may be written

$$f(b)-f(a) = \nabla f(c) \cdot (b-a) \ .$$

Proof: Consider the real-valued function F on $[0,1]$

$$\begin{aligned} F(t) &= f(a+t(b-a)) \ , & 0 \le t \le 1 \\ &= f(\lambda(t)) \ , & \lambda(t) = a+t(b-a) \ . \end{aligned}$$

By the Chain Rule $F'(t)$ exists $0 \le t \le 1$ and

$$\begin{aligned} F'(t) &= \mathcal{D}\, F(t)(1) = \mathcal{D}\, f(\lambda(t)) \circ \mathcal{D}(\lambda(t))(1) \\ &= \mathcal{D}\, f(\lambda(t))(\lambda'(t)) = \mathcal{D}\, f(\lambda(t))(b-a) \ . \end{aligned}$$

By the Mean-Value Theorem for F

$$F(1)-F(0) = F'(t_o)(1-0) = F'(t_o) \;,$$

for some $t_o \in (0,1)$, i.e.,

$$f(b)-f(a) = \mathcal{D}\, f(\lambda(t_o))(b-a) = \mathcal{D}\, f(c)(b-a) \quad , \quad c = \lambda(t_o) \;. \quad \square$$

This theorem does not hold for functions $f : R^n \to R^m$ if $m > 1$. Can you tell why? If you can't, try to carry out the proof with $m > 1$. However see Exercise 4.29 on page 174.

__Exercises:__

4.25: Let $f : R \to R$ be differentiable. If $F : R^2 \to R$ is defined by

(a) $F(x,y) = f(xy)$ then $x \dfrac{\partial F}{\partial x} = y \dfrac{\partial F}{\partial y}$,

(b) $F(x,y) = f(ax+by)$ then $b \dfrac{\partial F}{\partial x} = a \dfrac{\partial F}{\partial y}$,

(c) $F(x,y) = f(x^2+y^2)$ then $y \dfrac{\partial F}{\partial x} = x \dfrac{\partial F}{\partial y}$.

4.26: $f : R^2 \to R$, $f(x,y) = (x^2+y^2) \sin \dfrac{1}{x^2+y^2}$, if $(x,y) \neq (0,0)$, $f(0,0) = 0$. Show that f is differentiable at $(0,0)$ but $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$ are not continuous at $(0,0)$.

4.27: Let $f : R^n \to R$ be a differentiable function and C a smooth curve in R^n on which the function f is constant. Prove that for any point $c \in C$ the tangent to C at c is perpendicular to $\nabla f(c) = (\dfrac{\partial f}{\partial x_1}(c) , \ldots , \dfrac{\partial f}{\partial x_n}(c))$.

4.28: $f : R^n \to R^n$, f differentiable. If for all collections of n colinear points $\{p_1,\ldots,p_n\}$ in R^n the determinant

$$\begin{vmatrix} \frac{\partial f_1}{\partial x_1}(p_1) & , & \cdots & , & \frac{\partial f_1}{\partial x_n}(p_1) \\ \vdots & & & & \vdots \\ \frac{\partial f_n}{\partial x_1}(p_n) & , & \cdots & , & \frac{\partial f_n}{\partial x_n}(p_n) \end{vmatrix}$$

is non-zero show that f is (1-1) on R^n . Thus if $n = 1$ and $f'(x) \neq 0$ for each $x \in R$ then f is (1-1) on R . Show that if $f(x,y) = (x^3-y, e^{x+y})$ then f is (1-1) on R^2 .

4.29: If $f : R^n \to R^m$ is differentiable at each point c in the line segment between a and b and satisfies $|D\, f(c)(u)| \leq M|u|$, for each $u \in R^n$ then $|f(b)-f(a)| \leq M|b-a|$.

4.30: (Euler's Theorem) $f(x_1,\ldots,x_n)$ is homogeneous of degree m if

$$(*) \qquad f(t\,x_1,\ldots,t\,x_n) = t^m f(x_1,\ldots,x_n) \;, \quad \forall\, t \in R \,.$$

For example, $x^3 + y^3 + 3x^2y$, $\frac{x^5 + y^5 + z^5}{(x+y+z)^5}$, $x^2y^7 + 2z^4x^5$ are homogeneous of degree 3 , 0 , 9 respectively. Prove that if f is differentiable and homogeneous of degree m then

$$x_1 \frac{\partial f}{\partial x_1} + \ldots + x_n \frac{\partial f}{\partial x_n} = m\, f \,.$$

[Hint: Differentiate (*) with respect to t and set $t = 1$.]

Was Euler an Edmonton hockey player???

4.30: (Project)

(i) Suppose f is continuous on $[a,b]\times[c,d] = I$ to R ; prove that if

$$F(x) = \int_c^d f(x,y)\, dy$$

then F is continuous on $[a,b]$. [f is uniformly continuous on I.]

(ii) Suppose f , $\frac{\partial f}{\partial x}$ are continuous on I . Let F be as in (i). Prove that F' exists and is continuous on $[a,b]$ and

$$F'(x) = \int_c^d \frac{\partial f}{\partial x}(x,y)\, dy \ .$$

[Let $\phi(x) = \int_c^d \frac{\partial f}{\partial x}(x,y)\, dy$; ϕ is continuous by (i), use the Fubini Theorem to show that $\int_a^x \phi = F(x)-F(a)$. Hence F is an antiderivative of ϕ so F' exists and equals ϕ .]

(iii) Suppose $\alpha(x)$ and $\beta(x)$ have continuous derivatives on $[a,b]$ and f , $\frac{\partial f}{\partial x}$ are continuous on I . Show

$$\frac{d}{dx}\int_{\alpha(x)}^{\beta(x)} f(x,t)dt = f(x,\beta(x))\beta'(x) - f(x,\alpha(x))\alpha'(x) + \int_{\alpha(x)}^{\beta(x)} \frac{\partial f}{\partial x}(x,t)dt.$$

[Apply the Chain Rule and (ii) to

$$F(x,y,z) = \int_y^z f(x,t)dt \ , \quad x = x \ , \ y = \alpha(x) \ , \ z = \beta(x) \ .]$$

(iv) From the formula

$$\int_0^\pi \frac{dx}{a + b\cos x} = \frac{\pi}{(a^2-b^2)^{1/2}} \ , \quad a > 0 \ , \quad |b| < a \ ,$$

establish the results

$$\int_0^{\pi} \frac{dx}{(a + b \cos x)^2} = \frac{\pi a}{(a^2-b^2)^{3/2}}$$

$$\int_0^{\pi} \frac{\cos x \ dx}{(a + b \cos x)^2} = \frac{-\pi b}{(a^2-b^2)^{3/2}} \ .$$

(v) Evaluate $\int_0^b \frac{dx}{x^2+a^2}$, $a > 0$, and from your result deduce

$$\int_0^b \frac{dx}{(x^2+a^2)^2} = \frac{1}{2a^3} \tan^{-1} (\frac{b}{a}) + \frac{b}{2a^2(b^2+a^2)} \ , \quad a > 0 \ .$$

Show also that

$$\int_0^{\infty} \frac{dx}{(x^2+a^2)^2} = \frac{\pi}{4a^3} \ , \quad a > 0 \ .$$

[$\int_0^{\infty} f \overset{\text{def}}{=} \lim_{T\to\infty} \int_0^T f$ if this limit exists.]

PARTIAL DERIVATIVES OF HIGHER ORDER

Definition: Let $f : R^n \to R$ be such that $\frac{\partial f}{\partial x_i}$ exists at some point. If $\frac{\partial}{\partial x_j} (\frac{\partial f}{\partial x_i})$ also exists at this point then we write

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} (\frac{\partial f}{\partial x_i})$$

and, in particular,

$$\frac{\partial^2 f}{\partial x_i{}^2} = \frac{\partial}{\partial x_i} (\frac{\partial f}{\partial x_i}) \ .$$

<u>Example:</u> $f(x,y) = x^3 + 3y^2 + 2xy$.

$$\frac{\partial f}{\partial x} = 3x^2 + 2y \qquad \frac{\partial f}{\partial y} = 6y + 2x$$

$$\frac{\partial^2 f}{\partial x^2} = 6x \quad , \quad \frac{\partial^2 f}{\partial x \partial y} = 2 = \frac{\partial^2 f}{\partial y \partial x} \quad , \quad \frac{\partial^2 f}{\partial y^2} = 6 \quad .$$

Partial derivatives of third and higher order are similarly defined.

It is usually (but not always, cf. Exercise 13, page 249, Buck) the case that successive partial differentiations may be taken in any order we please, e.g., in the above example

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

The following two theorems give sufficient conditions for this. There is no loss of generality in the fact that these theorems are proved in R^2 only; we are only concerned with the behaviour of f with respect to two variables anyway.

<u>Theorem 4.11</u>. $f : R^2 \to R$. If $\frac{\partial^2 f}{\partial x \partial y}, \frac{\partial^2 f}{\partial y \partial x}$ exist and are continuous on an open set $U \subset R^2$ then

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \quad \text{on} \quad U \ .$$

<u>Proof</u>: Let $I = [a,b]\times[c,d] \subset U$. By the Fubini Theorem

$$\mathcal{J}_1 = \int_I \frac{\partial^2 f}{\partial x \partial y} = \int_c^d \{\int_a^b \frac{\partial^2 f}{\partial x \partial y} (x,y)dx\} \ dy$$

$$= \int_c^d \{\frac{\partial f}{\partial y}(b,y) - \frac{\partial f}{\partial y}(a,y)\}\, dy$$

$$= f(b,d) - f(b,c) - f(a,d) + f(a,c) \quad ,$$

$$J_2 = \int_I \frac{\partial^2 f}{\partial y \partial x} = \int_a^b \{\int_c^d \frac{\partial^2 f}{\partial y \partial x}\, dy\}\, dx$$

$$= \int_a^b \{\frac{\partial f}{\partial x}(x,d) - \frac{\partial f}{\partial x}(x,c)\}\, dx$$

$$= f(b,d) - f(a,d) - f(b,c) + f(a,c) \quad ;$$

therefore $J_1 = J_2$ for each interval $I \subset U$. Now suppose $\frac{\partial^2 f}{\partial x \partial y}(x_o,y_o) \neq \frac{\partial^2 f}{\partial y \partial x}(x_o,y_o)$ for some $(x_o,y_o) \in U$. We may suppose further that $\frac{\partial^2 f}{\partial x \partial y}(x_o,y_o) - \frac{\partial^2 f}{\partial y \partial x}(x_o,y_o) > 0$; by the continuity, $\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} > 0$ on some interval I, $(x_o,y_o) \in I$. But then $J_1 - J_2 = \int_I (\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x}) > 0$, by Theorem 3.6(e), contradicting $J_1 - J_2 = 0$. Therefore $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ throughout U.

Theorem 4.12 is more general than Theorem 4.11 but also more difficult to prove.

<u>Theorem 4.12</u>: $f : R^2 \to R$. If

(i) $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial^2 f}{\partial x \partial y}$ exist in a neighbourhood U of (x_o,y_o),

(ii) $\frac{\partial^2 f}{\partial x \partial y}$ is continuous at (x_o,y_o) ,

then $\frac{\partial^2 f}{\partial y \partial x}(x_o,y_o)$ exists and equals $\frac{\partial^2 f}{\partial x \partial y}(x_o,y_o)$.

<u>Proof</u>: (Optional). We must show that

$$\frac{\partial^2 f}{\partial y \partial x}(x_o,y_o) = \lim_{k\to 0}\left[\frac{\partial f}{\partial x}(x_o,y_o+k) - \frac{\partial f}{\partial x}(x_o,y_o)\right]/k \ \exists = \frac{\partial^2 f}{\partial x \partial y}(x_o,y_o) \quad .$$

Let $\Delta_h f(x,y) \overset{\text{def}}{=} [f(x+h,y) - f(x,y)]/h$. Then

(1) $\lim_{h\to 0} \Delta_h f(x,y) \ \exists = \frac{\partial f}{\partial x}(x,y)$, by (i), if $(x,y) \in U$,

(2) $\frac{\partial}{\partial y} \Delta_h f(x,y) \ \exists$, by (i) if $(x,y) \in U$ and h small.

Now consider

$$(\#) \quad \left[\frac{\partial f}{\partial x}(x_o,y_o+k) - \frac{\partial f}{\partial x}(x_o,y_o)\right]/k \ , \quad k \text{ small,}$$

$$= \lim_{h\to 0}[\Delta_h f(x_o,y_o+k) - \Delta_h f(x_o,y_o)]/k \ , \quad \text{by (1)}$$

$$= \lim_{h\to 0}\frac{\partial}{\partial y}\Delta_h f(x_o,y_o+\theta k) \ , \ 0 < \theta = \theta(k) < 1 \ , \quad \text{by M-V Theorem.}$$

But (*) $\frac{\partial}{\partial y}\Delta_h f(x_o,y_o+\theta k)$

$$= \left[\frac{\partial f}{\partial y}(x_o+h,y_o+\theta k) - \frac{\partial f}{\partial y}(x_o,y_o+\theta k)\right]/h \quad (\text{definition of } \Delta_h f)$$

$$= \frac{\partial^2 f}{\partial x \partial y}(x_o+\phi h,y_o+\theta k) \ , \quad 0 < \phi = \phi(h,k) < 1 \ , \text{ by M-V Theorem.}$$

By (ii), if $\varepsilon > 0$, $\exists\ \delta(\varepsilon) > 0 \ni |h| < \delta(\varepsilon)$, $|k| < \delta(\varepsilon) \Rightarrow$

$$\left|\frac{\partial^2 f}{\partial x \partial y}(x_o+h,y_o+k) - \frac{\partial^2 f}{\partial x \partial y}(x_o,y_o)\right| < \varepsilon \ .$$

Therefore, from (*) $\left|\frac{\partial}{\partial y}\Delta_h f(x_o,y_o+\theta k) - \frac{\partial^2 f}{\partial x \partial y}(x_o,y_o)\right| < \varepsilon$. Now from (#) we may take $\lim_{h\to 0}$ in this inequality to obtain

$$\left|\left[\frac{\partial f}{\partial x}(x_o,y_o+k) - \frac{\partial f}{\partial x}(x_o,y_o)\right]/k - \frac{\partial^2 f}{\partial x\partial y}(x_o,y_o)\right| \le \varepsilon$$

if $0 < |k| < \delta(\varepsilon)$. Therefore

$$\frac{\partial^2 f}{\partial y\partial x}(x_o,y_o) \ \exists = \frac{\partial^2 f}{\partial x\partial y}(x_o,y_o) \quad . \qquad \square$$

<u>Notation</u>: $D \subset R^n$.

$C(D)$: The set of continuous functions on D .

$C^k(D)$: The set of functions on D having all kth order partial derivatives continuous on D .

The range of the functions involved will be obvious from the context in which the notation is used.

<u>Theorem 4.13</u> (Taylor's Theorem).

(i) Let $a = (\alpha_1,\alpha_2)$, $b = (\beta_1,\beta_2) \in R^2$. If $f : R^2 \to R$, $f \in C^n(U)$, where U is a convex open subset of R^2 , and $a,b \in U$ then

$$f(b) = \sum_{k=0}^{n-1} \frac{1}{k!}\left\{(\beta_1-\alpha_1)\frac{\partial}{\partial x} + (\beta_2-\alpha_2)\frac{\partial}{\partial y}\right\}^k f(a) + R_n$$

where $R_n = \frac{1}{n!}\left\{(\beta_1-\alpha_1)\frac{\partial}{\partial x} + (\beta_2-\alpha_2)\frac{\partial}{\partial y}\right\}^n f(c)$ and c is some point on the line segment between a and b , $c \neq a,b$.

(ii) Let $a = (\alpha_1,\ldots,\alpha_m)$, $b = (\beta_1,\ldots,\beta_m) \in R^m$. If $f \in C^n(U)$, U a convex open subset of R^m and $a,b \in U$ then

$$f(b) = \sum_{k=0}^{n-1} \frac{1}{k!} \{(\beta_1-\alpha_1) \frac{\partial}{\partial x_1} + \ldots + (\beta_m-\alpha_m) \frac{\partial}{\partial x_m}\}^k f(a) + R_n$$

where $R_n = \frac{1}{n!} \{(\beta_1-\alpha_1) \frac{\partial}{\partial x_1} + \ldots + (\beta_m-\alpha_m) \frac{\partial}{\partial x_m}\}^n f(c)$ and c is some point on the line segment between a and b , $c \neq a,b$.

<u>Proof of (i)</u>: As in Theorem 4.10 (M.V. Th.) let

$$F(t) = f(\lambda(t)) \qquad \lambda(t) = a + t(b-a) \quad .$$

By the Chain Rule $F'(t) = D f(\lambda(t))(b-a)$

$$= (\beta_1-\alpha_1) \frac{\partial f}{\partial x} (\lambda(t)) + (\beta_2-\alpha_2) \frac{\partial f}{\partial y} (\lambda(t))$$

$$= \{(\beta-\alpha_1) \frac{\partial}{\partial x} + (\beta_2-\alpha_2) \frac{\partial}{\partial y}\} \ f(\lambda(t)) \ .$$

By induction

$$F^{(k)}(t) = \{(\beta_1-\alpha_1) \frac{\partial}{\partial x} + (\beta_2-\alpha_2) \frac{\partial}{\partial y}\}^k f(\lambda(t)) \quad .$$

Now Taylor's Theorem for F implies

$$F(1) = \sum_{k=0}^{n-1} \frac{1}{k!} F^{(k)}(0) + R_n \ , \ R_n = \frac{1}{n!} F^{(n)}(t_o) \ , \ t_o \in (0,1)$$

i.e.,

$$f(b) = \sum_{k=0}^{n-1} \frac{1}{k!} \{(\beta_1-\alpha_1) \frac{\partial}{\partial x} + (\beta_2-\alpha_2) \frac{\partial}{\partial y}\}^k f(a) + R_n$$

$$R_n = \frac{1}{n!} \{(\beta_1-\alpha_1) \frac{\partial}{\partial x} + (\beta_2-\alpha_2) \frac{\partial}{\partial y}\}^n f(\lambda(t_o)) \quad . \qquad \square$$

Application (Extrema of functions of several variables): The symmetric matrix

$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is positive (negative) semidefinite if

$$Q(x,y) = [x,y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 \geq 0 \quad (\leq 0)$$

for all $(x,y) \in R^2$. It is said to be positive (negative) definite if $Q(x,y) > 0$ (< 0) for all $(x,y) \neq (0,0)$. Otherwise it is called indefinite (i.e. if it is not a least semidefinite).

Lemma: $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ is

(i) positive (negative) semidefinite <=>

$$ac - b^2 \geq 0 \quad \text{and} \quad a \geq 0 (\leq 0) \quad , \quad c \geq 0 \ (\leq 0) \quad ,$$

(ii) positive (negative) definite <=>

$$ac - b^2 > 0 \quad \text{and} \quad a > 0 \ (< 0) \quad ,$$

(iii) indefinite <=> $ac - b^2 < 0$.

Proof: (a) $Q(x,0) = ax^2$ has the same sign as a

(b) $Q(0,y) = cy^2$ has the same sign as c

(c) $aQ(x,y) = (ax+by)^2 + (ac-b^2)y^2$.

The Lemma follows from (a), (b), (c).

<u>Definition</u>: (i) $f : R^n \to R$. If $\frac{\partial f}{\partial x_i}(c)$, $i = 1,\ldots,n$ exist and are all zero then c is called a <u>stationary</u> point of f.

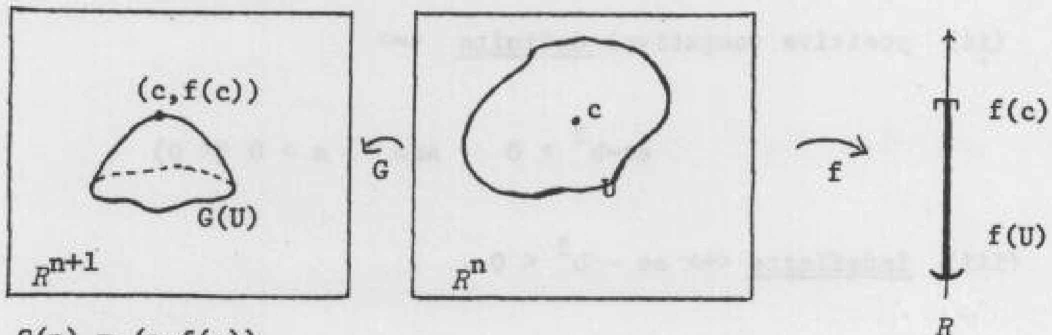
(ii) If c is an interior point of the domain of f and $f(p) \leq f(c)$ ($\geq f(c)$) for all p in some neighbourhood of c then f has an <u>interior relative maximum (minimum)</u> at c.

<u>Theorem 4.14</u>: If f has an interior relative maximum (minimum) at c then c is a stationary point of f.

<u>Proof</u>: Exercise.

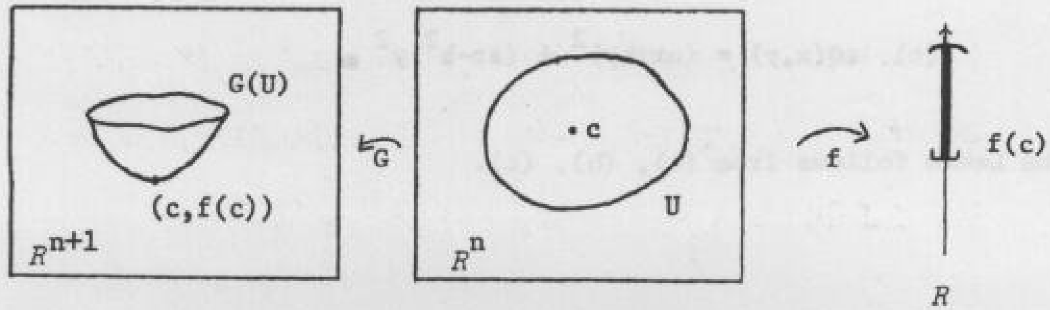
<u>Classification of Stationary Points</u>: An interior stationary point may be a

(a) <u>Relative Maximum</u> if $f(p) \leq f(c)$ $\forall$ $p \in U$, a neighbourhood of c.

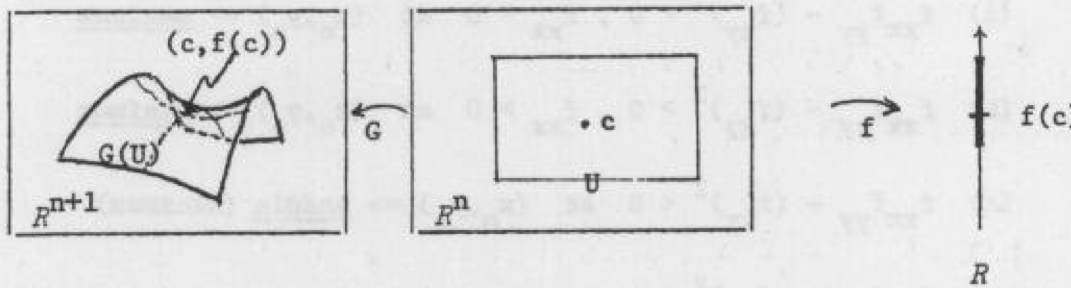


$G(p) = (p,f(p))$

(b) <u>Relative Minimum</u> if $f(p) \geq f(c)$ $\forall$ $p \in U$, a neighbourhood of c.



(c) Saddle Point (or Minimax) if there are points p_1 , p_2 in each neighbourhood U of c such that $f(p_1) < f(c)$, $f(p_2) > f(c)$.



Moreover c is called a strict relative maximum or minimum when the inequalities in (a), (b) are strict.

In the following the symbols f_x , f_y , f_{xy} , f_{yx} , f_{xx} , f_{yy} denote the functions $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$, $\frac{\partial^2 f}{\partial y \partial x}$, $\frac{\partial^2 f}{\partial x \partial y}$, $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial^2 f}{\partial y^2}$ respectively.

Theorem 4.15: $f : R^2 \to R$, $f \in C^2(U)$, U a neighbourhood of (x_o,y_o) . Suppose

(i) $f_x(x_o,y_o) = f_y(x_o,y_o) = 0$,

(ii) $A(x,y) = \begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{bmatrix}$ (symmetric. Why?)

Then, at (x_o,y_o) , f has a

(a) relative maximum if $A(x_o,y_o)$ is negative definite,

(b) relative minimum if $A(x_o,y_o)$ is positive definite,

(c) saddle point if $A(x_o,y_o)$ is indefinite, and

(d) <u>anything</u> can happen at (x_o,y_o) if $A(x_o,y_o)$ is properly <u>semi-definite</u> (i.e. det $A(x_o,y_o) = 0$) <u>critical case</u>

(a) $f_{xx}f_{yy} - (f_{xy})^2 > 0$, $f_{xx} < 0$ at (x_o,y_o) => <u>maximum</u>

(b) $f_{xx}f_{yy} - (f_{xy})^2 > 0$, $f_{xx} > 0$ at (x_o,y_o) => <u>minimum</u>

(c) $f_{xx}f_{yy} - (f_{xy})^2 < 0$ at (x_o,y_o) => <u>saddle</u> (minimax)

(d) $f_{xx}f_{yy} - (f_{xy})^2 = 0$ at (x_o,y_o) => ? i.e. <u>critical case</u>

<u>Proof</u>: By Taylor's Theorem, if (x,y) is sufficiently close to (x_o,y_o)

$$f(x,y) = f(x_o,y_o) + \frac{1}{1!}\{(x-x_o)\frac{\partial}{\partial x} + (y-y_o)\frac{\partial}{\partial y}\} f(x_o,y_o) + R_2(x,y)$$

$$= f(x_o,y_o) + R_2(x,y) \text{ , by (i), where}$$

$$R_2(x,y) = \frac{1}{2!}\{(x-x_o)\frac{\partial}{\partial x} + (y-y_o)\frac{\partial}{\partial y}\}^2 f(x_1,y_1)$$

where (x_1,y_1) is a point on the line segment between (x_o,y_o) and (x,y) . Thus

$$R_2(x,y) = \frac{1}{2!}\{(x-x_o)^2\frac{\partial^2 f}{\partial x^2}(x_1,y_1) + 2(x-x_o)(y-y_o)\frac{\partial^2 f}{\partial x\partial y}(x_1,y_1) + (y-y_o)^2\frac{\partial^2 f}{\partial y^2}(x_1,y_1)\}$$

$$= \frac{1}{2!}[x-x_o,y-y_o]\begin{bmatrix} f_{xx}(x_1,y_1) & f_{xy}(x_1,y_1) \\ f_{yx}(x_1,y_1) & f_{yy}(x_1,y_1)\end{bmatrix}\begin{bmatrix} x-x_o \\ y-y_o\end{bmatrix}$$

$$= \frac{1}{2!}[x-x_o,y-y_o][A(x_1,y_1)]\begin{bmatrix} x-x_o \\ y-y_o\end{bmatrix} \qquad (*)$$

(a),(b) : If the conditions of (a) ((b)) hold at (x_o,y_o) then by continuity $(f \in C^2(U))$ they hold in a neighbourhood of (x_o,y_o) so that if (x,y) is close to (x_o,y_o) then $A(x,y)$ is still negative (positive) definite, i.e. from (*) $R_2(x,y) < 0$ (> 0) for all (x,y) near (x_o,y_o).

(c) : If $[\mu,\nu]\,[A(x_o,y_o)]\,\begin{bmatrix}\mu\\ \nu\end{bmatrix} < 0$ (> 0) for some fixed (μ,ν) then it holds for each scalar multiple of (μ,ν). Furthermore, by continuity, the same inequality holds with $A(x_o,y_o)$ replaced by $A(x,y)$ if (x,y) is close to (x_o,y_o). Thus if $A(x_o,y_o)$ is indefinite $R_2(x,y)$ takes both positive and negative values in each neighbourhood of (x_o,y_o).

(d) : The critical case $f_{xx}f_{yy} - (f_{xy})^2 = 0$ is illustrated by Examples 4, 5 below.

Remarks:

(1) It may be evident which behaviour the function has at a stationary point simply from consideration of $f(x,y)$ near (x_o,y_o) (Examples 4, 5).

(2) The sign of the remainder $R_2(x,y)$ may be clear by considering the matrix $A(x_1,y_1)$ directly rather than from continuity as was done in the proof of the preceding theorem.

Examples:

(1) $f(x,y) = x^2 + xy + y^2$, $\dfrac{\partial f}{\partial x} = 2x + y = 0$

$$\frac{\partial f}{\partial y} = x + 2y = 0$$

$(x,y) = (0,0)$ is the only stationary point.

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad 2 > 0\ ,\ 2.2-1.1 = 3 > 0\ \ ,\ \text{positive definite}$$

f has a <u>relative</u> <u>minimum</u> at $(0,0)$. In fact it is a global minimum since the matrix above is positive definite for all (x,y) and so $R_2(x,y) > 0$ for all $(x,y) \neq (0,0)$.

(2) $f(x,y) = x^2 + 4xy + y^2$, $\frac{\partial f}{\partial x} = 2x + 4y = 0$

$$\frac{\partial f}{\partial y} = 4x + 2y = 0$$

$(x,y) = (0,0)$ is the only stationary point

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix} \quad 2.2-4.4 = -12 < 0 \text{ , } \underline{\text{indefinite}}$$

f has a <u>saddle</u> <u>point</u> at $(0,0)$. Alternatively we can see this by observing that if $(x,y) \neq (0,0)$, $f(x,y) > 0$ on the x and y axes and $f(x,y) < 0$ on that portion of the line $y = x$ which lies in the third quadrant.

(3) $f(x,y) = 3x^2 - y^2 + x^3$, $\frac{\partial f}{\partial x} = 6x + 3x^2 = 3x(2+x) = 0$

$$\frac{\partial f}{\partial y} = -2y = 0$$

$(0,0)$ and $(-2,0)$ are the stationary points.

$(0,0)$: $\begin{bmatrix} 6 & 0 \\ 0 & -2 \end{bmatrix}$ is indefinite => <u>saddle</u> at $(0,0)$

$(-2,0)$: $\begin{bmatrix} -6 & 0 \\ 0 & -2 \end{bmatrix}$ is negative definite => relative maximum at $(-2,0)$

(4) $f(x,y) = x^4 + y^4$, $\frac{\partial f}{\partial x} = 4x^3 = 0$

$$\frac{\partial f}{\partial y} = 4y^3 = 0$$

Stationary at $(0,0)$

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 12x^2 & 0 \\ 0 & 12y^2 \end{bmatrix}$$

This matrix at $(0,0)$ is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ which gives no information (critical case). However $12x^2 \geq 0$, $12y^2 \geq 0$, $144x^2y^2 \geq 0$ so that the matrix is positive semidefinite for all (x,y) (positive definite if $x \neq 0$ and $y \neq 0$). Thus the remainder $R_2(x,y) \geq 0$ for all (x,y) and f has a <u>minimum</u> at $(0,0)$. Of course elementary considerations yield $f(x,y) > f(0,0)$ if $(x,y) \neq (0,0)$ without any consideration of second order partials.

(5) $f(x,y) = x^4 - y^4$

Again the only stationary point is $(0,0)$ and

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 12x^2 & 0 \\ 0 & -12y^2 \end{bmatrix}$$

which gives the critical case at $(0,0)$. However the matrix is <u>postive semidefinite</u> on the line $y = 0$ and <u>negative semidefinite</u> on $x = 0$ so that $(0,0)$ is a saddle point. In this case it is also obvious without considering second order partials that

$$f(x,0) > f(0,0) \text{ if } x \neq 0 \text{ and } f(0,y) < f(0,0) \text{ if } y \neq 0 .$$

Examples 4 and 5 illustrate the fact that anything can happen in the case that $f_{xx}f_{yy} - (f_{xy})^2 = 0$ at the stationary point.

The problem is similar for functions of more than two variables. A symmetric matrix

$$A = \begin{bmatrix} a_{11} & , & \cdots & , & a_{1n} \\ \vdots & & & & \vdots \\ a_{n1} & , & \cdots & , & a_{nn} \end{bmatrix} , \quad a_{ij} = a_{ji}$$

has an associated quadratic form

$$Q(x,\ldots,x_n) = [x_1,\ldots,x_n]\, A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i,j=1}^{n} a_{ij}\, x_i x_j .$$

A is called <u>positive</u> (<u>negative</u>) <u>semidefinite</u> if

$$Q(x_1,\ldots,x_n) \geq 0 \; (\leq 0) \quad , \quad \forall\, (x_1,\ldots,x_n)$$

and is called <u>positive</u> (<u>negative</u>) <u>definite</u> if the inequality is strict for all $(x_1,\ldots,x_n) \neq (0,\ldots,0)$; otherwise (i.e. if the range of Q has both positive and negative values) A is said to be <u>indefinite</u>.

The following criteria may be found in the book Gantmacher: "Matrix Theory" pp. 306-308.

(a) A is <u>positive semidefinite</u> <=> the determinants of <u>all</u> $k \times k$ submatrices $(k=1,\ldots,n)$ of A symmetric about the main diagonal are ≥ 0 .

(b) A is <u>negative</u> <u>semidefinite</u> <=> the determinants of <u>all</u> $k \times k$ sub-matrices $(k=1,\ldots,n)$ of A symmetric about the main diagonal are ≤ 0 , ≥ 0 according as k is odd or even.

(c) A is <u>positive</u> <u>definite</u> <=>

$$a_{11} > 0 \;, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0 \;, \quad \ldots \;, \quad \begin{vmatrix} a_{11} \,, & \ldots \,, & a_{1n} \\ a_{n1} \,, & \ldots \,, & a_{nn} \end{vmatrix} > 0$$

(d) A is <u>negative</u> <u>definite</u> <=>

$$a_{11} < 0 \;, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0 \;, \quad \ldots \;, \quad (-1)^n \begin{vmatrix} a_{11} \,, & \ldots \,, & a_{1n} \\ a_{n1} \,, & \ldots \,, & a_{nn} \end{vmatrix} > 0$$

(e) A is <u>indefinite</u> <=> it satisfies none of (a), (b), (c), (d).

The nature of a stationary point can be determined as before by the use of Taylor's Theorem: $a = (\alpha_1,\ldots,\alpha_n)$, $b = (\beta_1,\ldots,\beta_n)$

$$f(b)-f(a) = \frac{1}{1!}\left\{(\beta_1-\alpha_1)\frac{\partial}{\partial x_1} + \ldots + (\beta_n-\alpha_n)\frac{\partial}{\partial x_n}\right\} f(a) + R_2(b)$$

$$= R_2(b) \quad \text{if } a \text{ is a stationary point (i.e. } \frac{\partial f}{\partial x_i}(a) = 0).$$

$$R_2(b) = \frac{1}{2!}\left\{(\beta_1-\alpha_1)\frac{\partial}{\partial x_1} + \ldots + (\beta_n-\alpha_n)\frac{\partial}{\partial x_n}\right\}^2 f(c)$$

$$= \frac{1}{2!}\sum_{i,j=1}^{n} \frac{\partial^2 f(c)}{\partial x_i \partial x_j}(\beta_i-\alpha_i)(\beta_j-\alpha_j)$$

$$= \frac{1}{2!}\,[\beta_1-\alpha_1,\ldots,\beta_n-\alpha_n]\begin{bmatrix} f_{x_1x_1}(c) & ,\ldots, & f_{x_1x_n}(c) \\ \vdots & & \vdots \\ f_{x_nx_1}(c) & ,\ldots, & f_{x_nx_n}(c) \end{bmatrix}\begin{bmatrix} \beta_1-\alpha_1 \\ \vdots \\ \beta_n-\alpha_n \end{bmatrix}$$

where c is some point on the line segment between a and b. We see that the matrix of second partials determines the nature of the stationary point.

Example:

$$f(x,y,z) = x^2 + y^2 + z^2 + 2xyz$$

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= 2x + 2yz \\ \frac{\partial f}{\partial y} &= 2y + 2zx \\ \frac{\partial f}{\partial z} &= 2z + 2xy \end{aligned}\right\} \quad \text{Stationary points at } (0,0,0),\ (-1,1,1),\ (1,-1,1),\ (1,1,-1),\ (-1,-1,-1).$$

$$\begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix} = \begin{bmatrix} 2 & 2z & 2y \\ 2z & 2 & 2x \\ 2y & 2x & 2 \end{bmatrix}$$

$(0,0,0)$: $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ is positive definite so f has a <u>relative</u> <u>minimum</u> at $(0,0,0)$.

Other stationary points

$$2 > 0 \quad , \quad \begin{vmatrix} 2 & 2z \\ 2z & 2 \end{vmatrix} = 4 - 4z^2 = 0 \quad \text{if} \quad |z| = 1$$

$$\begin{vmatrix} 2 & 2z & 2y \\ 2z & 2 & 2x \\ 2y & 2x & 2 \end{vmatrix} \begin{array}{l} = 8(1-x^2-y^2-z^2+2xyz) \\ < 0 \ , \quad \text{if} \quad |x| = |y| = |z| = 1 \ , \ xyz = -1 \ . \end{array}$$

Thus the matrix is <u>indefinite</u> at all other critical points so that these are <u>saddles</u>.

<u>Exercises</u>:

In these exercises assume all the differentiability you need unless otherwise specified.

4.31: Find V_x , V_y , V_z , V_u , V_{xy} , V_{xyz} , V_{xyzu} for the following functions V :

(i) $\dfrac{x^2y^2u}{a^2-z^2}$. (ii) $\dfrac{x}{y} + \dfrac{y}{z} + \dfrac{z}{u} + \dfrac{u}{x} + \dfrac{xy}{zu}$.

[Solution: (i) $V_x = \dfrac{2xy^2u}{a^2-z^2}$, $V_y = \dfrac{2x^2yu}{a^2-z^2}$, $V_z = \dfrac{2x^2y^2uz}{(a^2-z^2)^2}$, $V_u = \dfrac{x^2y^2}{a^2-z^2}$,

$V_{xy} = \dfrac{4xyu}{a^2-z^2}$, $V_{xyz} = \dfrac{8xyzu}{(a^2-z^2)^2}$, $V_{xyzu} = \dfrac{8xyz}{(a^2-z^2)^2}$.

(ii) $V_x = \dfrac{1}{y} - \dfrac{u}{x^2} + \dfrac{y}{zu}$, $V_y = \dfrac{-x}{y^2} + \dfrac{1}{z} + \dfrac{x}{zu}$, $V_z = \dfrac{-y}{z^2} + \dfrac{1}{u} - \dfrac{xy}{z^2u}$,

$V_u = \dfrac{-z}{u^2} + \dfrac{1}{x} - \dfrac{xy}{u^2z}$, $V_{xy} = -\dfrac{1}{y^2} + \dfrac{1}{zu}$, $V_{xyz} = -\dfrac{1}{uz^2}$, $V_{xyzu} = \dfrac{1}{u^2z^2}$.

4.32: If $V = x^2-y^2$, $x = 3u - 4v + 2$, $y = 2u + 3v - 1$, show that $V_u = 6x - 4y$, $V_v = -8x - 6y$.

4.33: If $V = x^3 + 8y^3 - 18xy$ find the points (x,y) at which $V_x = V_y = 0$ and investigate the nature of V at these points. (Solution: $(0,0)$ saddle, $(3,\frac{3}{2})$ minimum.)

4.34: If $P = x^2y - y^3 - y^2z$, $Q = xy^2 - x^3 - x^2z$, $R = xy^2 + x^2y$, show that $P(Q_z-R_y) + Q(R_x-P_z) + R(P_y-Q_x) = 0$.

4.35: If $u = x + y + z$, $v = x^2 + y^2 + z^2$, $w = x^3 + y^3 + z^3 - 3xyz$, show $\frac{\partial(u,v,w)}{\partial(x,y,z)} = 0$.

4.36: If V,P,Q,R,μ are C^1 functions of (x,y,z) satisfying the relations

$$V_x = \mu P \quad , \quad V_y = \mu Q \quad , \quad V_z = \mu R \quad (\mu \neq 0)$$

show that

$$P(Q_z-R_y) + Q(R_x-P_z) + R(P_y-Q_x) = 0 \quad .$$

4.37: (i) $P,Q : R^2 \to R$; P , Q , $\frac{\partial P}{\partial y}$, $\frac{\partial Q}{\partial x}$ continuous. Prove that there is a real valued function f on R^2 such that

$$f'(x,y) = [P(x,y),Q(x,y)]$$

if and only if $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.

[Hint: "Only if" easy. "If" - consider

$$f(x,y) = \int_{x_o}^{x} P(t,y_o)dt + \int_{y_o}^{y} Q(x,t)\, dt \; .$$

Use Exercise 4.30 (ii).]

(ii) <u>State</u> what you would consider to be a generalization of part (i) for functions of three variables.

(iii) Verify that each of the following is the Jacobian matrix $f'(x,y)$ of some function $f(x,y)$ and find $f(x,y)$.

(a) $[2xy, x^2+3y^2]$,

(b) $[2y\, e^{2x} + 2x \cos y\, ,\ e^{2x} - x^2 \sin y]$.

4.38: If $F = \dfrac{x}{x^2+y^2}$, $G = \dfrac{y}{x^2+y^2}$, show

(i) $F_y = G_x$, $F_x = -G_y$

(ii) $\nabla^2 F = \nabla^2 G = 0$ where $\nabla^2 = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$.

4.39: (i) If $V = V(x,y)$, $x = x(u,v)$, $y = y(u,v)$ then

(a) $$[V_u\ V_v] = [V_x\ V_y] \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}$$

(b) $$\begin{bmatrix} V_{uu} & V_{uv} \\ V_{vu} & V_{vv} \end{bmatrix} = \begin{bmatrix} x_u & y_u \\ x_v & y_v \end{bmatrix} \begin{bmatrix} V_{xx} & V_{xy} \\ V_{yx} & V_{yy} \end{bmatrix} \begin{bmatrix} x_u & x_v \\ y_u & u_v \end{bmatrix}$$
$$+ V_x \begin{bmatrix} x_{uu} & x_{uv} \\ x_{vu} & x_{vv} \end{bmatrix} + V_y \begin{bmatrix} y_{uu} & y_{uv} \\ y_{vu} & y_{vv} \end{bmatrix} .$$

(ii) If $U = U(x,y)$, $V = V(x,y)$, $x = x(u,v)$, $y = y(u,v)$, show $\dfrac{\partial(U,V)}{\partial(u,v)} = \dfrac{\partial(U,V)}{\partial(x,y)} \cdot \dfrac{\partial(x,y)}{\partial(u,v)}$. [Use (i)(a).]

(iii) If $x = r \cos\theta$, $y = r \sin\theta$ show

$$\frac{\partial(U,V)}{\partial(x,y)} = \frac{1}{r}\frac{\partial(U,V)}{\partial(r,\theta)} .$$

4.40: If $V = 3x^2 + 2y^2 + \phi(x^2-y^2)$ prove

$$y\, V_x + x\, V_y = 10xy \ .$$

4.41: If $V = x^2 + y^2 + \phi(xy) + \psi(\frac{y}{x})$ prove

$$x^2\, V_{xx} - y^2\, V_{yy} + x\, V_x - y\, V_y = 4(x^2-y^2) \ .$$

4.42: If $V = v(x,y)$, $x = \rho \cos \phi$, $y = \rho \sin \phi$ show

$$\nabla^2 V \overset{\text{def}}{=} V_{xx} + V_{yy} = V_{\rho\rho} + \frac{1}{\rho} V_\rho + \frac{1}{\rho^2} V_{\phi\phi}$$

[Hint: $\rho = \sqrt{x^2+y^2}$, $\phi = \tan^{-1}(\frac{y}{x})$.]

4.43: If $V = V(r,\theta)$, $\rho = \frac{c^2}{r}$, $\phi = 2\pi - \theta$ show

$$\rho^2\, V_{\rho\rho} + \rho\, V_\rho + V_{\phi\phi} = r^2\, V_{rr} + r\, V_r + V_{\theta\theta} \ .$$

4.44: If $V = V(x,y)$, $x = x(u,v)$, $y = y(u,v)$ and $x_u = y_v$, $x_v = -y_u$, show

$$\frac{V_{uu} + V_{vv}}{V_{xx} + V_{yy}} = x_u^2 + x_v^2 = y_u^2 + y_v^2 \ .$$

4.45: (<u>Euler's</u> <u>Theorem</u>, continued cf. Exercise 4.30.) If V is a homogeneous function of (x,y,z) of the mth degree show

$$x^2\, V_{xx} + y^2\, V_{yy} + z^2\, V_{zz} + 2xy\, V_{xy} + 2yz\, V_{yz} + 2zx\, V_{zx} = m(m-1)V \ .$$

4.46: If $z_y = F(z_x)$, $F' \neq 0$, then $z_{xx}\, z_{yy} = z_{xy}^2$.

4.47: If $z = x\,F(x+y) + G(x+y)$ show that

$$z_{xx} - 2z_{xy} + z_{yy} = 0 \ .$$

4.48: (a) If $z = F(y+m_1x) + G(y+m_2x)$ and m_1 , m_2 are the roots of the quadratic equation $am^2 + 2hm + b = 0$ then $a\,z_{xx} + 2h\,z_{xy} + b\,z_{yy} = 0$. Show that this equation is satisfied by $z = x\,F(y+m_1x) + G(y+m_1x)$ if m is a double root of the quadratic equation.

(b) The equation of lateral vibration of a taut string is $\frac{\partial^2 z}{\partial t^2} - c^2\,\frac{\partial^2 z}{\partial x^2} = 0$. Deduce from (a) that $z = F(x+ct) + G(x-ct)$ is a solution.

(c) A vibrating string for which initial displacement and velocity are specified is governed by relations of the form $\frac{\partial^2 z}{\partial t^2} - c^2\,\frac{\partial^2 z}{\partial x^2} = 0$,

$z(x,0) = f(x)$, $z_t(x,0) = g(x)$. Show that

$$z(x,t) = \frac{1}{2}\,\{f(x+ct) + f(x-ct)\} + \frac{1}{2c}\int_{x-ct}^{x+ct} g$$

is a solution to this problem.

4.49: (a) If $z = x\,\phi(\frac{y}{x}) + \psi(\frac{y}{x})$ then

$$x^2\,z_{xx} + 2xy\,z_{xy} + y^2\,z_{yy} = 0 \ .$$

(b) If $z = x^3\,\phi(\frac{y}{x}) + \frac{1}{x^3}\,\psi(\frac{y}{x})$ then

$$x^2\,z_{xx} + 2xy\,z_{xy} + y^2\,z_{yy} + x\,z_x + y\,z_y = 9z \ .$$

[If you are observant you don't have to do all that differentiation.]

4.50: Discuss the nature of the stationary points of the functions

(a) $34x^2 - 24xy + 41y^2$ [Solution: (0,0) minimum.]

(b) $3x^2 + 4xy - 4y^2$ [Solution: (0,0) saddle.]

(c) $x^2y - 4x^2 - y^2$ [Solution: (0,0) max., $(\pm 2\sqrt{2},4)$ saddles.]

(d) $x^2y + 2x^2 - 2xy + 3y^2 - 4x + 7y$

[Solution: $(1,-1)$ min., $(1\pm\sqrt{6},-2)$ saddles.]

(e) $x^2yz - 2xyz + x^2z + x^2 + y^2 + z^2 + yz - 2xz - 2x + 2y + z$

[Solution: $(1,-1,0)$ min., $(1\pm\sqrt{2},-2,1)$, $(1\pm\sqrt{2},0,-1)$ saddles.]

4.51: Show that the function

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

does not have a relative extremum at (0,0) even though it has a relative minimum at this point when the domain is restricted to any line $x = t$, $y = \alpha t$. [In the (x,y) - plane sketch the curves $f(x,y) = 0$, then determine the regions $f(x,y) < 0$, $f(x,y) > 0$; check that on each line through (0,0) $f(x,y) > 0 = f(0,0)$ if (x,y) is close enough to (0,0).]

4.52: Show that the box of maximum volume which can be inscribed in the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

has sides of length $\frac{2a}{\sqrt{3}}$, $\frac{2b}{\sqrt{3}}$, $\frac{2c}{\sqrt{3}}$.

4.53: Suppose we are given n points $(x_j, y_j) \in R^2$ and desire to find a function $F(x) = Ax + B$ for which the quantity $\sum_{j=1}^{n} (F(x_j) - y_j)^2$ is minimized. Show that this problem leads to the equations

$$A \sum_{j=1}^{n} x_j^2 + B \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} x_j y_j$$

$$A \sum_{j=1}^{n} x_j + n\,B = \sum_{j=1}^{n} y_j$$

which are easily solved for A and B. The line $y = Ax + B$ is the line which best fits the given set of points in the sense of <u>least squares</u>.

4.54: Let $\{\phi_n\}$ be a sequence of real-valued continuous functions on $[a,b]$ such that

$$\int_a^b \phi_n \phi_m = \begin{cases} 1 , & \text{if } n = m \\ 0 , & \text{if } n \neq m . \end{cases}$$

Let f be a real-valued continuous function on $[a,b]$. Prove that the choice of constants $\gamma_1, \ldots, \gamma_\ell$ which minimizes the quantity $\int_a^b (f - \sum_{k=1}^{\ell} \gamma_k \phi_k)^2$, for any ℓ, is

$$\gamma_k = \int_a^b f \phi_k , \quad k = 1, \ldots, \ell .$$

This problem arises in the theory of <u>Fourier</u> <u>Series</u>.

4.55: The capacity of a condenser formed by two concentric spherical conductors of radii a and b , $0 < a < b$ is

$$C(a,b) = \frac{ab}{b-a} .$$

Suppose that in measuring the radii a and b of the spheres the measurements are subject to errors of amount δa and δb respectively. Given that products of the errors are negligible relative to the other quantities involved derive the following approximation on the error in C .

$$\frac{\delta C}{C} \simeq \frac{\delta a}{a}\frac{b}{b-a} - \frac{\delta b}{b}\frac{a}{b-a} .$$

4.56: The breaking weight W of a cantilever beam is given by the formula $W\ell = kbd^2$ where b : breadth , ℓ : length , d : depth , k : constant depending on the material in the beam. If the breadth is increased by 2 per cent and the depth by 5 per cent, show that the length should be increased by about 12 per cent if the breaking weight is to remain unchanged.

4.57: In a triangle ABC the area is calculated from the elements a , B , C , the measurements being subject to errors δa , δB , δC . Show that the error δS in the area is approximately given by

$$\frac{\delta S}{S} \simeq 2\,\frac{\delta a}{a} + \frac{c\ \delta B}{a \sin B} + \frac{b\ \ \delta C}{a \sin C} .$$

4.58: Let $f : R^2 \to R$, f of class C^2 on a convex subset D of R^2 . Suppose that at each point $p \in D$ the matrix $\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$ is positive semidefinite. Prove that if $p,q \in D$ then $f(\frac{p+q}{2}) \leq \frac{1}{2}[f(p)+f(q)]$. What does the result mean geometrically?

LOCAL PROPERTIES OF C^1 FUNCTIONS

Recall that if $f : R^n \to R^n$ (i.e. $f(p) = (f_1(p),\ldots,f_n(p))$, $p = (x_1,\ldots,x_n)$) which is differentiable at p_o then the Jacobian of f at p_o is

$$\det f'(p_o) = \left|\begin{matrix} \frac{\partial f_1}{\partial x_1} , & \cdots , & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_n}{\partial x_1} , & \cdots , & \frac{\partial f_n}{\partial x_n} \end{matrix}\right|_{p_o} = \frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)}(p_o) = J_f(p_o)$$

Definition: A function f is locally one-to-one on a subset D of its domain if there is a neighbourhood of each point $p \in D$ on which f is one-to-one.

Lemma 4.16: If $f : R^n \to R^n$ is of class C^1 at p_o and $J_f(p_o) \neq 0$ then there is a neighbourhood U of p_o on which f is one-to-one.

Proof: Since the partials of f are continuous at p_o and $J_f(p_o) \neq 0$ there is a convex neighbourhood U of p_o such that if $p_i \in U$, $i = 1,\ldots,n$ then

$$(1) \qquad \left|\begin{matrix} \frac{\partial f_1}{\partial x_1}(p_1) , & \cdots , & \frac{\partial f_1}{\partial x_n}(p_1) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \frac{\partial f_n}{\partial x_1}(p_n) , & \cdots , & \frac{\partial f_n}{\partial x_n}(p_n) \end{matrix}\right| \neq 0$$

If $a,b \in U$ then, applying the Mean-Value Theorem to each component of f , we find

$$
(2) \qquad
\begin{aligned}
f_1(b) - f_1(a) &= \sum_{i=1}^{n} (\beta_i - \alpha_i) \frac{\partial f_1}{\partial x_i}(p_1) \\
&\cdots\cdots \\
f_n(b) - f_n(a) &= \sum_{i=1}^{n} (\beta_i - \alpha_i) \frac{\partial f_n}{\partial x_i}(p_n)
\end{aligned}
$$

where $a = (\alpha_1,\dots,\alpha_n)$, $b = (\beta_1,\dots,\beta_n)$ and p_i , $i = 1,\dots,n$ are some points on the line segment between a and b . But (1) and (2) imply $f_i(b) - f_i(a) = 0$, $i = 1,\dots,n$ $\Leftrightarrow$ $\beta_i - \alpha_i = 0$, $i = 1,\dots,n$, i.e. $f(b) = f(a) \Leftrightarrow b = a$. Thus f is one-to-one on U . $\square$

The following theorem is an immediate consequence of this lemma.

<u>Theorem 4.16</u>: $f : R^n \to R^m$, $m \geq n$. Suppose D is an open subset of R^n and

(i) $f \in C^1(D)$

(ii) $\text{rk } f'(p) = n \ \forall p \in D$

then f is locally one-to-one on D .

<u>Proof</u>: Suppose $f = (f_1,\dots,f_m)$ $(m \geq n)$; we wish to show that f is one-to-one on a neighbourhood U of each point $p_o \in D$. Since $\text{rk } f'(p_o) = n$ we may assume without loss of generality that

$$\frac{\partial(f_1,\dots,f_n)}{\partial(x_1,\dots,x_n)}(p_o) \neq 0 \quad .$$

(This can always be achieved by relabeling the f's .) Thus, if $\tilde{f} : R^n \to R^n$ is defined by $\tilde{f} = (f_1,\dots,f_n)$, $\tilde{f} \in C^1(D)$ and $J_{\tilde{f}}(p_o) \neq 0$ so, by the Lemma, $\tilde{f}$ is one-to-one on a neighbourhood of p_o . But since $f(p_1) = f(p_2)$

$$\Rightarrow \quad \tilde{f}(p_1) = \tilde{f}(p_2)\ ,\ \tilde{f}(1\text{-}1) \Rightarrow f(1\text{-}1)\ .$$

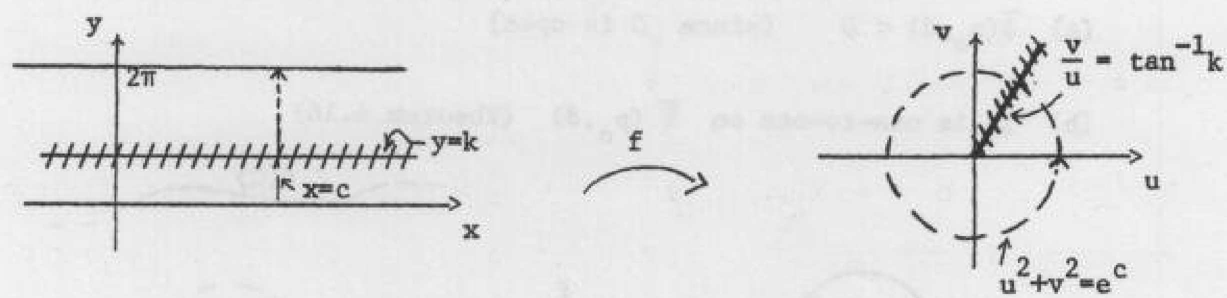
f is not necessarily globally one-to-one on D as the following example shows.

Example: $f(x,y) = (e^x \cos y\ ,\ e^x \sin y)$

$$J_f(x,y) = \begin{vmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{vmatrix} = e^{2x} \neq 0\ ,\ \forall (x,y) \in R^2\ .$$

Thus, from Theorem 4.16, f is locally one-to-one on R^2. However f is not globally one-to-one on R^2 since

$$f(x,y+2\pi) = f(x,y) \quad \forall\ (x,y) \in R^2$$



$$f(\{(x,y) : x \in R\ ,\ y \in [\alpha,\alpha+2\pi)\}) = R^2 - \{(0,0)\}\ ,\ \forall\ \alpha \in R.$$

Exercises:

4.59: In Exercise 4.28 a sufficient condition was given for $f : R^n \to R^n$ to be globally one-to-one. Verify that this condition is not satisfied by the preceding example. [If you did not do Exercise 4.28 a proof is contained in the preceding lemma.]

4.60: Show that if $f : R \to R$ and $f'(x) \neq 0$ for each $x \in R$ then f is one-to-one globally on R.

<u>Lemma 4.17</u>: $f : R^n \to R^n$. Suppose

(i) D open $\subset R^n$, $f \in C'(D)$.

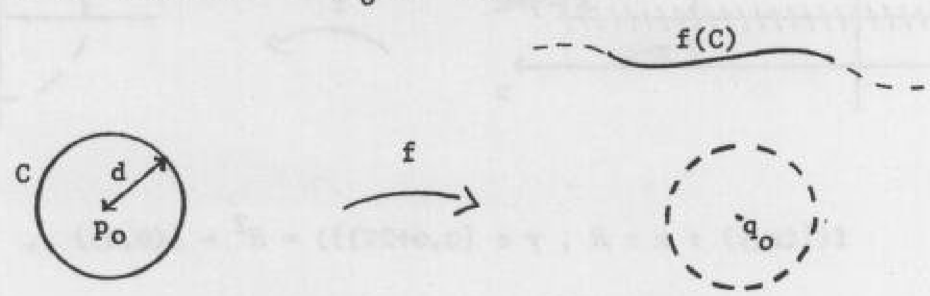
(ii) $J_f(p) \neq 0$, $\forall\, p \in D$,

then $f(D)$ is open.

<u>Proof</u>: In order to prove $f(D)$ open we must show if $q_o \in f(D)$, $\exists\ \delta_0 > 0 \ni B(q_o,\delta_o) \subset f(D)$. Let $q_o = f(p_o) \in f(D)$; $\exists\ d > 0 \ni$

(a) $\overline{B}(p_o,d) \subset D$ (since D is open)

(b) f is one-to-one on $\overline{B}(p_o,d)$ (Theorem 4.16)



Let $C = \partial B(p_o,d) = \{p : |p-p_o| = d\}$. $f(C)$ is <u>compact</u> (Why?) and $q_o = f(p_o) \notin f(C)$ (Why?).

If $\delta = \inf \{|q-q_o| : q \in f(C)\}$, $\delta > 0$ (Why?).

Claim: $B(q_o,\frac{\delta}{3}) \subset f(D)$ so $f(D)$ is open.

To prove this claim suppose $q_1 \in B(q_o,\frac{\delta}{3})$; we wish to show $q_1 \in f(D)$. Consider

$$\phi(p) = |f(p)-q_1|^2 \quad , \quad p \in \overline{B}(p_o,d)$$

(1) $\quad \phi \in C(D) \Rightarrow \exists\ p_* \in \overline{B}(p_o,d) \ni \phi(p_*) \leq \phi(p)\ ,\ \forall\ p \in \overline{B}(p_o,d)$ (Why?).

$p_* \notin C$ since $p_* \in C \Rightarrow \sqrt{\phi(p_*)} \geq \frac{2}{3}\delta > \frac{1}{3}\delta > \sqrt{\phi(p_o)}$ contradicting (1). Therefore ϕ has an __interior__ minimum at p_* so the partial derivatives of ϕ are all zero at p_* (Theorem 4.14)

$$\phi(p) = |f(p)-q_1|^2 = (f(p)-q_1)\cdot(f(p)-q_1)$$

$$= \sum_{i=1}^{n} (f_i(p)-y_i)^2$$

$$q_1 = (y_1,\ldots,y_n) \qquad p = (x_1,\ldots,x_n)$$

$$\therefore \quad 0 = 2\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_j}(p_*)(f_i(p_*)-y_i) \quad , \quad j = 1,\ldots,n \quad .$$

Since $\det[\frac{\partial f_i}{\partial x_j}(p_*)] = J_f(p_*) \neq 0$ we conclude

$$f_i(p_*) = y_i \quad , \quad i = 1,\ldots,n$$

$$\text{i.e.} \quad f(p_*) = q_1$$

$$\text{i.e.} \quad q_1 \in f(D) \quad .$$

Therefore $B(q_o,\frac{\delta}{3}) \subset f(D)$ and $f(D)$ is open. $\square$

<u>Remark:</u> The condition $J_f(p) \neq 0$ may not be dropped even at a single point in D ; e.g. if $f(x) = x^2$, $x \in R$ then $f(R) = [0,\infty)$ is not open even though $f'(x) \neq 0$ except when $x = 0$.

<u>Theorem 4.17</u>: $f : R^n \to R^m$ $(m \leq n)$. Suppose

(i) D open $\subset R^n$, $f \in C^1(D)$,

(ii) $\operatorname{rk} f'(p) = m \;\; \forall\, p \in D$,

then $f(D)$ is open.

<u>Proof</u>: Given $c = (\gamma_1,\ldots,\gamma_n) \in D$, we must prove $f(c) \in$ interior $f(D)$. We may assume $\dfrac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_m)}(c) \neq 0$ from (ii) (relabel the x's if necessary); since $f \in C^1(D)$, $\dfrac{\partial(f_1,\ldots,f_m)}{\partial(x_1,\ldots,x_m)}(p) \neq 0$ $\forall\, p$ in some open neighbourhood U of c . Let $\tilde{f} : R^m \to R^m$ be defined by $\tilde{f}(x_1,\ldots,x_m) = f(x_1,\ldots,x_m,\gamma_{m+1},\ldots,\gamma_m)$ on the open set $\tilde{U} = \{(x_1,\ldots,x_m) : (x_1,\ldots,x_m,\gamma_{m+1},\ldots,\gamma_n) \in U\} \subset R^m$. Since $\tilde{f} \in C^1(\tilde{U})$ and $J_{\tilde{f}}(x_1,\ldots,x_m) \neq 0$, $\forall\, (x_1,\ldots,x_m) \in \tilde{U}$, Lemma 4.17 => $\tilde{f}(\tilde{U})$ open

$$\therefore \quad f(c) = \tilde{f}(c) \in \tilde{f}(\tilde{U}) \text{ open} \subset f(U) \subset f(D)$$

$$\therefore \quad f(c) \in \operatorname{int} f(D) \text{ , } \forall\, c \in D \text{ so } f(D) \text{ is open.}$$

<u>Question</u>: Why is $\tilde{U}$ open?

<u>Lemma 4.18.1</u>: If f is continuous and one-to-one on a <u>compact</u> set S then f^{-1} is continuous on $f(S)$.

Proof: Let $q_o \in f(S)$; we wish to show that f^{-1} is continuous at q_o . If f^{-1} is not continuous at q_o then for some $\varepsilon_o > 0$ there is a sequence of points $q_n \in f(S)$ for which

(1) $$\lim\{q_n\} = q_o \quad \text{and} \quad |f^{-1}(q_n) - f^{-1}(q_o)| \geq \varepsilon_o .$$

$\{f^{-1}(q_n)\} \subset S$ (compact) so there is a convergent subsequence $\{f^{-1}(q_{n_k})\}$ and

(2) $$\lim\{f^{-1}(q_{n_k})\} = p_1 \in S \ , \ p_1 \neq f^{-1}(q_o) \quad \text{from (1)} .$$

But f is continuous at p_1 and f is one-to-one so, from (2),

$$\lim\{q_{n_k}\} = \lim\{f(f^{-1}(q_{n_k}))\} = f(p_1) \neq f(f^{-1}(q_o)) = q_o$$

i.e. $\lim\{q_{n_k}\}$ exists and is different from q_o contradicting (1). Thus f^{-1} must be continuous at each $q_o \in f(S)$. □

In general the requirement that S be compact cannot be dropped; it cannot be replaced even by the requirement that $f(S)$ be compact. Consider $f : [0,2\pi) \to \{(x,y) : x^2 + y^2 = 1\}$, $f(\theta) = (\cos\theta, \sin\theta)$, $0 \leq \theta < 2\pi$, $f \in C[0,2\pi)$, f is (1-1) and $f([0,2\pi)) = \{(x,y) : x^2 + y^2 = 1\}$ is compact but f^{-1} is discontinuous at $(1,0)$.

Exercise:

4.61: $f : R \to R$. Show that if f is continuous and one-to-one on a connected subset S of R then f^{-1} is continuous on $f(S)$.

<u>Lemma 4.18.2</u>: $f : R^n \to R^n$. If $f \in C^1$ at p_o and $J_f(p_o) \neq 0$ then there is a neighbourhood U of p_o and $m > 0$ such that if $p \in U$ then

$$|f(p)-f(p_o)| \geq m|p-p_o| \ .$$

<u>Proof</u>:

$$J_f(p_o) \neq 0 \Rightarrow D\, f(p_o) : R^n \xrightarrow{(1\text{-}1)} R^n$$

$$\Rightarrow \exists\ m > 0 \ni |D\, f(p_o)(u)| \geq 2m|u|, \forall\ u \in R^n \quad \text{(Theorem 4.3)}$$

(1) $$\therefore\ |D\, f(p_o)(p-p_o)| \geq 2m|p-p_o| \ , \ \forall\ p \in R^n \ .$$

But from the definition of $D\, f(p_o)$ there is a neighbourhood U of $p_o \ni$ $p \in U \Rightarrow$

(2) $$|f(p)-f(p_o) - D\, f(p_o)(p-p_o)| \leq m|p-p_o| \ .$$

But

(3) $$|f(p)-f(p_o) - D\, f(p_o)(p-p_o)| \geq |D\, f(p_o)(p-p_o)| - |f(p)-f(p_o)|$$

$$\geq 2m|p-p_o| - |f(p)-f(p_o)| \ , \ \text{ from (1)} \ .$$

(2),(3) $\Rightarrow 2m|p-p_o| - |f(p)-f(p_o)| \leq m|p-p_o|$, if $p \in U$

$$\Rightarrow m|p-p_o| \leq |f(p) - f(p_o)| \ .$$

<u>Theorem 4.18</u>: $f : R^n \to R^n$. Suppose

(i) D open $\subset R^n$, $f \in C^1(D)$, $J_f(p) \neq 0$, $\forall\, p \in D$,

(ii) f is one-to-one on D .

Then $f^{-1} \in C^1(f(D))$ and $D\, f^{-1} = (D\, f)^{-1}$.

Proof: Let $u : R^n \to R^n$ be defined by

$$(4) \qquad u(p) \overset{\text{def}}{=} \frac{1}{|p-p_o|}\,[f(p) - f(p_o) - D\, f(p_o)(p-p_o)]$$

$$(5) \qquad \lim_{p \to p_o} u(p) = 0 \qquad \text{(definition of } D\, f(p_o)).$$

Since $J_f(p_o) \neq 0$, $(D\, f(p_o))^{-1}$ exists and, from (4),

$$(6) \qquad |p-p_o|(D\, f(p_o))^{-1}(u(p)) = (D\, f(p_o))^{-1}(f(p)-f(p_o) - D\, f(p_o)(p-p_o))$$

$$= (D\, f(p_o))^{-1}(f(p)-f(p_o)) - (p-p_o) \quad .$$

By Theorem 4.17 $f(D)$ is open and in particular it is a neighbourhood of $f(p_o)$. Hence, with $q_o = f(p_o)$, $q = f(p)$, we may deduce from (6) and Lemma 4.18.2

$$(7) \qquad \frac{1}{m}\,|q-q_o|\,|(D\, f(p_o))^{-1}(u(p))| \geq |f^{-1}(q) - f^{-1}(q_o) - (D\, f(p_o))^{-1}(q-q_o)|$$

if p is in some neighbourhood of p_o . Furthermore from Lemma 4.18.1 f and f^{-1} are continuous at p_o and q_o so $q \to q_o \Rightarrow f^{-1}(q) \to f^{-1}(q_o)$, i.e. $p \to p_o$. Therefore from (7) and (5)

$$\lim_{q \to q_o} |f^{-1}(q) - f^{-1}(q_o) - (D\, f_o(p_o))^{-1}(q-q_o)|/|q-q_o| = 0$$

since

$$\lim_{q\to q_o} (D\, f(p_o))^{-1}(u(p)) = \lim_{p\to p_o} (D\, f(p_o))^{-1}(u(p)) = 0 \ .$$

Thus, since $(D\, f(p_o))^{-1}$ is linear, $D\, f^{-1}(p_o) \exists = (D\, f(p_o))^{-1}$.

$\underline{f^{-1} \in C'(f(D))}$ follows from the fact that since $D\, f^{-1} = (D\, f)^{-1}$ the partials of f^{-1} are rational functions of the partials of f in which the denominator $J_f(p)$ does not vanish. $\square$

<u>Example</u>: We have seen that $f(x,y) = (e^x \cos y\ ,\ e^x \sin y)$ is one-to-one on the strip $0 \leq y < 2\pi$

$$f'(x,y) = \begin{bmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{bmatrix} \ , \quad J_f(x,y) = e^{2x} \neq 0$$

$$\therefore \quad [f'(x,y)]^{-1} = e^{-2x} \begin{bmatrix} e^x \cos y & e^x \sin y \\ -e^x \sin y & e^x \cos y \end{bmatrix}$$

To find f^{-1} solve $(u,v) = (e^x \cos y\ ,\ e^x \sin y)$

$$u^2 + v^2 = e^{2x} \quad , \quad \tan y = \frac{v}{u}$$

$$\therefore \quad x = \frac{1}{2}\log(u^2+v^2) \quad , \quad y = \tan^{-1}\frac{v}{u}$$

$$\therefore \quad f^{-1}(u,v) = \left(\frac{1}{2}\log(u^2+v^2)\ ,\ \tan^{-1}\frac{v}{u}\right)$$

and

$$(f^{-1})'(u,v) = \begin{bmatrix} \frac{u}{u^2+v^2} & \frac{v}{u^2+v^2} \\ \frac{-v}{u^2+v^2} & \frac{u}{u^2+v^2} \end{bmatrix} = \frac{1}{u^2+v^2}\begin{bmatrix} u & v \\ -v & u \end{bmatrix}$$

$$= e^{-2x}\begin{bmatrix} e^x \cos y & e^x \sin y \\ -e^x \sin y & e^x \cos y \end{bmatrix} = [f'(x,y)]^{-1} .$$

Thus we see that $\mathcal{D}\, f^{-1}(u,v) = (\mathcal{D}\, f(x,y))^{-1}$.

<u>**Exercises:**</u>

4.62: If $y_i = y_i(x_1,\ldots,x_n) \in C'$, $i = 1,\ldots,n$ and $\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)} \neq 0$ show $\frac{\partial(y_1,\ldots,y_n)}{\partial(x_1,\ldots,x_n)} \cdot \frac{\partial(x_1,\ldots,x_n)}{\partial(y_1,\ldots,y_n)} = 1$.

4.63: Let $f(x,y) = (x^2,\frac{y}{x})$, $x \neq 0$. Find $f'(x,y)$. Show f is one-to-one on its domain by finding f^{-1} . Check that $[f']^{-1} = (f^{-1})'$.

4.64: Let $f(x,y) = (\frac{x}{(x^2+y^2)^{1/2}}, \frac{y}{(x^2+y^2)^{1/2}})$; show $J_f(x,y) = 0 \;\forall\; (x,y) \in \mathcal{R}^2 - \{(0,0)\}$ and f is not locally one-to-one anywhere on its domain. Show that the range of f is the circle $u^2 + v^2 = 1$ and thus contains no open subset.

THE IMPLICIT FUNCTION THEOREM

Question: Under what circumstances does a system of equations

$$f_i(x_1,\ldots,x_m,y_1,\ldots,y_n) = 0 \ , \ i = 1,\ldots,n$$

uniquely define y_i , $i = 1,\ldots,n$, implicitly as functions of $(x_1,\ldots,x_n)$ i.e.

$$y_i = \phi_i(x_1,\ldots,x_n) \ , \ i = 1,\ldots,n \ ?$$

If f_i were linear in $y_1,\ldots,y_n$ then the requirement is, of course, that the determinant of the coefficients of $y_1,\ldots,y_n$ be non-zero. An analogous answer in the general case is provided by the following important theorem.

Theorem 4.19 (Implicit Function Theorem):

$$f = (f_1,\ldots,f_n) \ , \ p = (x_1,\ldots,x_m) \ , \ q = (y_1,\ldots,y_n) \ .$$

Suppose

(i) $f : R^{n+m} \to R^n$, $f \in C^1(D)$, D open $\subset R^{n+m}$,

(ii) $(p_o,q_o) \in D$, $f(p_o,q_o) = 0$,

(iii) $\dfrac{\partial(f_1,\ldots,f_n)}{\partial(y_1,\ldots,y_n)}(p_o,q_o) \neq 0$.

I. Then there is a unique function $\phi : R^m \to R^n$ defined in a neighbourhood U of p_o such that $\phi \in C^1(U)$

(a) $\phi(p_o) = q_o$

(b) $f(p,\phi(p)) = 0, \forall \ p \in U$.

II. Moreover if $y_i = \phi_i(p)$, $i = 1,\ldots,n$,

(a) $[\frac{\partial f}{\partial x}] = -[\frac{\partial f}{\partial y}][\frac{\partial \phi}{\partial x}]$

(b) $$\frac{\partial \phi_i}{\partial x_j} = \frac{\partial y_i}{\partial x_j} = -\frac{\partial(f_1,\ldots,f_n)}{\partial(y_1,\ldots,\underset{\substack{\uparrow \\ \text{ith place}}}{x_j},\ldots,y_n)} \Big/ \frac{\partial(f_1,\ldots,f_n)}{\partial(y_1,\ldots,y_n)} , \quad \begin{array}{l} i = 1,\ldots,n \\ j = 1,\ldots,m \end{array}$$

<u>Example</u>: $f(x,y) = x^2 - y$.

$$\left.\begin{array}{l} f(0,0) = 0 \\ \frac{\partial f}{\partial y}(0,0) = -1 \end{array}\right\} \quad \begin{array}{l} f(x,y) = 0 \text{ is solved by } y = \phi(x) = x^2 \\ \text{in a neighbourhood of } x = 0 . \end{array}$$

Notice that

$$\frac{dy}{dx} = 2x = \frac{-2x}{-1} = -\frac{\partial f}{\partial x} \Big/ \frac{\partial f}{\partial y} .$$

How about solving for x as a function of y ?

$$\left.\begin{array}{l} f(0,0) = 0 \\ \frac{\partial f}{\partial x}(0,0) = 0 \end{array}\right\} \quad \text{Condition (iii) of the IFT is not satified.}$$

$f(x,y) = 0$ cannot be solved in the form $x = \psi(y)$ with ψ defined in a full neighbourhood of $y = 0$. Notice that there are <u>two</u> solutions of the form $x = \pm\sqrt{y}$ defined and continuous on $[0,\infty)$ but even these are not C^1 at 0 . There are infinitely many discontinuous solutions defined on $[0,\infty)$, e.g. $\psi(y) = \sqrt{y}$, $y \in Q$, $\psi(y) = -\sqrt{y}$, $y \notin Q$.

You will see from the simple exercises at the end of this section that when condition (iii) is not satisfied there may be no solution, or no C^1 solution, or indeed infinitely many solutions.

Proof of Theorem 4.19: Consider $F : R^{n+m} \to R^{n+m}$

$$F(p,q) = (p,f(p,q))$$

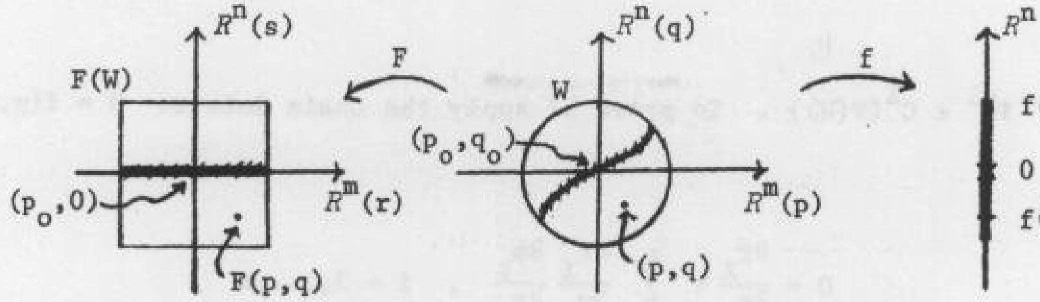
$$= (x_1,\dots,x_m,f_1(x_1,\dots,x_m,y_1,\dots,y_n),\dots,f_n(x_1,\dots,x_m,y_1,\dots,y_n))$$

$$f \in C^1(D) \Rightarrow F \in C^1(D)$$

and

$$J_F(p_o,q_o) = \left|\begin{array}{ccc|ccc} 1 & & 0 & & & \\ & 1 \cdot & \cdot & & 0 & \\ 0 & & \cdot\ 1 & & & \\ \hline \frac{\partial f_1}{\partial x_1}, & \cdots, & \frac{\partial f_1}{\partial x_m} & \frac{\partial f_1}{\partial y_1}, & \cdots, & \frac{\partial f_1}{\partial y_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}, & \cdots, & \frac{\partial f_n}{\partial x_m} & \frac{\partial f_n}{\partial y_1}, & \cdots, & \frac{\partial f_n}{\partial y_n} \end{array}\right|_{(p_o,q_o)}$$

$$= \frac{\partial(f_1,\dots,f_n)}{\partial(y_1,\dots,y_n)}(p_o,q_o) \neq 0 \ .$$



By Theorems 4.16 and 4.18 F has an inverse function in a neighbourhood W of (p_o,q_o) and $F^{-1} \in C^1(F(W))$ and $F(W)$ is a neighbourhood of $(p_o,0) = F(p_o,q_o)$ by Theorem 4.17.

$$F(p,q) = (p,f(p,q)) = (r,s) \qquad \text{(Notice } p=r\text{)}$$

$$\text{(1)} \qquad \therefore \quad (p,q) = F^{-1}(r,s) = (r,\theta(r,s))$$

$$\therefore \quad p = r \; , \; q = \theta(r,s)$$

$$\therefore \quad (r,s) = F(F^{-1}(r,s)) = F(r,\theta(r,s)) = (r,f(r,\theta(r,s)))$$

for all (r,s) in the neighbourhood $F(W)$ of $(p_o,0)$. In particular

$$s = f(r,\theta(r,s)) \qquad (r = p)$$

and thus

$$0 = f(p,\theta(p,0))$$

for all p in a neighbourhood U of p_o

$$\therefore \quad 0 = f(p,\phi(p)) \quad \text{if} \quad \phi(p) = \theta(p,0) \; ,$$

and $\phi(p_o) = \theta(p_o,0) = q_o$ so that I (a), (b) hold. $\phi \in C^1(U)$ follows from

(1) since $F^{-1} \in C^1(F(W))$. To prove II apply the Chain Rule to $0 = f(p,\phi(p))$ to obtain

II (a) $$0 = \frac{\partial f_i}{\partial x_j} + \sum_{k=1}^{n} \frac{\partial f_i}{\partial y_k}\frac{\partial \phi_k}{\partial x_j} \quad , \quad i = 1,\dots,n \ . \quad j = 1,\dots,m$$

Solving this system of equations for $\frac{\partial \phi_k}{\partial x_j}$, $k = 1,\dots,n$, by Cramer's Rule gives

II (b) $$\frac{\partial \phi_k}{\partial x_j} = \begin{vmatrix} \frac{\partial f_1}{\partial y_1} , \dots , -\frac{\partial f_1}{\partial x_j} , \dots , \frac{\partial f_1}{\partial y_n} \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \frac{\partial f_n}{\partial y_1} , \dots , -\frac{\partial f_n}{\partial x_j} , \dots , \frac{\partial f_n}{\partial y_n} \end{vmatrix} \Bigg/ \frac{\partial(f_1,\dots,f_n)}{\partial(y_1,\dots,y_n)}$$

$\uparrow$ kth column

$$= -\frac{\partial(f_1,\dots\dots\dots,f_n)}{\partial(y_1,\dots,x_j,\dots,y_n)} \Bigg/ \frac{\partial(f_1,\dots,f_n)}{\partial(y_1,\dots,y_n)}$$

$\uparrow$ kth place

<u>Remark</u>: ϕ is the unique solution to $f(p,\phi(p)) = 0$, $\phi(p_o) = q_o$ since if $f(p,\phi_1(p)) = f(p,\phi_2(p)) = 0$ and $\phi_1(p) \neq \phi_2(p)$ for some $p \in U$ then

$$F(p,\phi_1(p)) = F(p,\phi_2(p)) = (p,0)$$

contradicting the fact that F is one-to-one on W . This means that the graph of ϕ i.e. $\{(p,\phi(p)) : p \in U\}$ is the whole set $f^{-1}(0) \cap W$.

Corollary 4.19.1: $f : R^2 \to R$. If f is of class C^1 in a neighbourhood of (x_o,y_o) and

$$\text{(i)} \quad f(x_o,y_o) = 0 \qquad\qquad \text{(ii)} \quad \frac{\partial f}{\partial y}(x_o,y_o) \neq 0$$

then the equation $f(x,y) = 0$ has a unique solution $y = \phi(x)$, $\phi(x_o) = y_o$, which exists and is continuously differentiable in a neighbourhood U of x_o and

$$\phi'(x) = \frac{dy}{dx} = -\frac{\partial f}{\partial x} \Big/ \frac{\partial f}{\partial y} \ .$$

Corollary 4.19.2: $f : R^3 \to R$. If f is of class C' in a neighbourhood of (x_o,y_o,z_o) and

$$\text{(i)} \quad f(x_o,y_o,z_o) = 0 \qquad\qquad \text{(ii)} \quad \frac{\partial f}{\partial z}(x_o,y_o,z_o) \neq 0$$

then the equation $f(x,y,z) = 0$ has a unique solution $z = \phi(x,y)$, $\phi(x_o,y_o) = z_o$, which exists and is continuously differentiable in a neighbourhood U of (x_o,y_o) and

$$\frac{\partial z}{\partial x} = -\frac{\partial f}{\partial x} \Big/ \frac{\partial f}{\partial z} \ , \quad \frac{\partial z}{\partial y} = -\frac{\partial f}{\partial y} \Big/ \frac{\partial f}{\partial z} \ .$$

Corollary 4.19.3: $f = (f_1,f_2) : R^5 \to R^2$. If f is of class C^1 in a neighbourhood of (x_o,y_o,z_o,u_o,v_o) and

$$\text{(i)} \quad f_1(x_o,y_o,z_o,u_o,v_o) = 0 \qquad\qquad \text{(ii)} \quad \frac{\partial(f_1,f_2)}{\partial(u,v)}(x_o,y_o,z_o,u_o,v_o) \neq 0$$

$$f_2(x_o,y_o,z_o,u_o,v_o) = 0$$

then the equations

$$f_1(x,y,z,u,v) = 0$$

$$f_2(x,y,z,u,v) = 0$$

have a unique solution

$$u = \phi_1(x,y,z) \qquad u_o = \phi_1(x_o,y_o,z_o)$$

$$v = \phi_2(x,y,z) \qquad v_o = \phi_2(x_o,y_o,z_o)$$

with

$$\frac{\partial\phi_1}{\partial x} = \frac{\partial u}{\partial x} = -\frac{\partial(f_1,f_2)}{\partial(x,v)} \Big/ \frac{\partial(f_1,f_2)}{\partial(u,v)} , \quad \frac{\partial\phi_2}{\partial x} = \frac{\partial v}{\partial x} = -\frac{\partial(f_1,f_2)}{\partial(u,x)} \Big/ \frac{\partial(f_1,f_2)}{\partial(u,v)}$$

$$\frac{\partial\phi_1}{\partial y} = \frac{\partial u}{\partial y} = -\frac{\partial(f_1,f_2)}{\partial(y,v)} \Big/ \frac{\partial(f_1,f_2)}{\partial(u,v)} , \quad \frac{\partial\phi_2}{\partial y} = \frac{\partial v}{\partial y} = -\frac{\partial(f_1,f_2)}{\partial(u,y)} \Big/ \frac{\partial(f_1,f_2)}{\partial(u,v)}$$

$$\frac{\partial\phi_1}{\partial z} = \frac{\partial u}{\partial z} = -\frac{\partial(f_1,f_2)}{\partial(z,v)} \Big/ \frac{\partial(f_1,f_2)}{\partial(u,v)} , \quad \frac{\partial\phi_2}{\partial z} = \frac{\partial v}{\partial z} = -\frac{\partial(f_1,f_2)}{\partial(u,z)} \Big/ \frac{\partial(f_1,f_2)}{\partial(u,v)}$$

__Exercises__:

4.65: Prove Corollary 4.19.1, i.e. work through the proof of Theorem 4.19 in this special case.

4.66: The equation $y^2 - x^2 = 0$ has two C^1 solutions. $y = \pm x$ in a neighboorhood of $x = 0$ i.e. uniqueness does not hold. What condition of the Implicit Function Theorem does not hold? Check that there are four solutions of class C and infinitely many real-valued solutions.

4.67: Show that the equations

$$\begin{matrix} x^2 - yu = 0 \\ xy + uv = 0 \end{matrix} \quad , \quad u = u_o \ , \ v = v_o$$

define u,v implicitly as functions of x and y in a neighb ourhood of any point (x_o,y_o) if $y_o u_o \neq 0$ and $x_o^2 - y_o u_o = 0$, $x_o y_o + u_o v_o = 0$. Check that the Implicit Function Theorem gives

$$\frac{\partial u}{\partial x} = \frac{2x}{y} \ , \ \frac{\partial u}{\partial y} = -\frac{u}{y} \ , \ \frac{\partial v}{\partial x} = -\frac{y}{u} - \frac{2vx}{uy} \ , \ \frac{\partial v}{\partial y} = -\frac{x}{u} + \frac{v}{y}$$

and verify these results by solving the equations directly.

4.68: Consider $f(x,y,u,v) = (x^3 + yu + v \ , \ xv + y^3 - u)$. At what points (x,y,u,v) can one solve $f(x,y,u,v) = (0,0)$ in the form $(x,y) = (\phi_1(u,v) \ ,$ $\phi_2(u,v))$? Find the differential matrix

$$\begin{bmatrix} \frac{\partial \phi_1}{\partial u} & \frac{\partial \phi_1}{\partial v} \\ \frac{\partial \phi_2}{\partial u} & \frac{\partial \phi_2}{\partial v} \end{bmatrix} .$$

4.69: Suppose $\phi_i(u_1,\ldots,u_n,x_1,\ldots,x_n) = 0$, $i = 1,\ldots,n$ is satisfied by $u_j = u_j(x_1,\ldots,x_n)$, the ϕ's and u's being C^1 functions; prove

$$\frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(u_1,\ldots,u_n)} \cdot \frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)} = (-1)^n \frac{\partial(\phi_1,\ldots,\phi_n)}{\partial(x_1,\ldots,x_n)} .$$

[Use the Chain Rule and think matrices.]

4.70: If $u^2 + v^2 + 2uvx + y = 0$, $uv + (u+v)y + x^2 = 0$ prove that

$$\frac{\partial(u,v)}{\partial(x,y)} = \frac{uv(u+v) - x}{(u-v)\{(u+v)+y(1-x)\}} \quad .$$

4.71: If $u_1 = x_1 + x_2 + x_3 + x_4$, $u_1u_2 = x_2 + x_3 + x_4$, $u_1u_2u_3 = x_3 + x_4$, $u_1u_2u_3u_4 = x_4$, show

$$\frac{\partial(x_1,x_2,x_3,x_4)}{\partial(u_1,u_2,u_3,u_4)} = u_1^3\, u_2^2\, u_3 \ .$$

4.72: (a) If $V = \psi(u,v)$, $\phi(u,v) = E(x,y)$, $\chi(u,v) = F(x,y)$ and $\frac{\partial(\phi,\chi)}{\partial(u,v)} \neq 0$ prove

$$\frac{\partial V}{\partial x} \cdot \frac{\partial(\phi,\chi)}{\partial(u,v)} = \frac{\partial E}{\partial x}\,\frac{\partial(\psi,\chi)}{\partial(u,v)} + \frac{\partial F}{\partial x} \cdot \frac{\partial(\phi,\psi)}{\partial(u,v)} \quad .$$

(b) If $V = u^2 + v^2 + uv$, $u+v = x^2 + y^2$, $u^3 + v^3 = 2xy$ prove

$$3\,\frac{\partial V}{\partial y} = \frac{2x(x^2-y^2)}{(x^2+y^2)^2} + 8y(x^2+y^2) \quad .$$

4.73: If $V = \psi(u_1,\dots,u_n)$, $\phi_k(u_1,\dots,u_n) = f_k(x_1,\dots,x_m)$, $k = 1,\dots,n$ and $\frac{\partial(\phi_1,\dots,\phi_n)}{\partial(u_1,\dots,u_n)} \neq 0$ show

$$\frac{\partial V}{\partial x_j} \cdot \frac{\partial(\phi_1,\dots,\phi_n)}{\partial(u_1,\dots,u_n)} = \sum_{k=1}^{n} \frac{\partial f_k}{\partial x_j} \cdot \frac{\partial(\phi_1,\dots,\overset{\substack{\text{kth position}\\ \uparrow}}{\psi},\dots,\phi_n)}{\partial(u_1,\dots\dots\dots,u_n)} \quad .$$

4.74: If $F \in C^1$ show that the equation

$$F(F(x,y),y) = 0$$

may be solved for y as a function of x near $(0,0)$ provided $F(0,0) = 0$, $F_x(0,0) \neq -1$ and $F_y(0,0) \neq 0$.

4.75: Suppose $f(x,y) = 0$ where $f \in C^2$, $\frac{\partial f}{\partial y} \neq 0$; prove

$$\frac{d^2y}{dx^2} = \frac{1}{(\frac{\partial f}{\partial y})^3} \begin{vmatrix} 0 & \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x} & \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial f}{\partial y} & \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix}$$

Dimension:

We have a notion of dimension for linear objects, specifically vector spaces i.e., the number of vectors in a basis. If $L : R^k \to R^n$ (linear), $L(x) = Ax$, then the dimension of $L(R^k) \subset R^n$ is the rank of A ($\mathrm{rk}\, A$) ; in particular if $\mathrm{rk}\, A = k$ then $L(R^k)$ is k dimensional, the same dimension as R^k .

Definition:

(i) A subset S of R^n is a k-dimensional segment ($k > 0$) if there is an open connected set $D \subset R^k$ and a function $f : R^k \to R^n$ such that

(a) $f \in C^1(D)$, $f : D \xrightarrow{(1\text{-}1)} S$, $f(D) = S$,

(b) $\mathrm{rk}\, f'(p) = k$, $\forall\, p \in D$.

A o-<u>dimensional</u> <u>segment</u> in R^n is a single point in R^n .

(ii) A subset S of R^n is a k-<u>dimensional</u> <u>manifold</u> if $\forall\, q \in S\ \exists$ an open neighbourhood $V \subset R^n$ of q such that $V \cap S$ is a k-dimensional segment.

<u>Remarks</u>:

(1) $k \leq n$

(2) a k-dimensional segment is a k-dimensional manifold

(3) if the condition that f is (1-1) is dropped from (i) then f is still locally (1-1) by Theorem 4.16 so f(U) is a k-dimensional segment if U is a sufficiently small open subset of D .

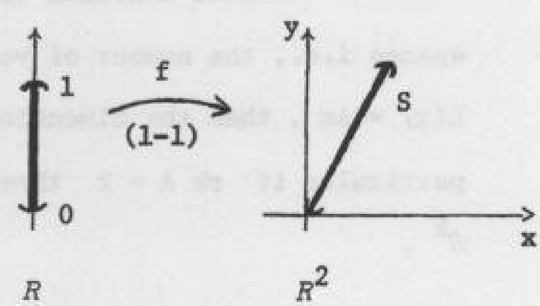
<u>Examples</u>:

(1) $S = \{(x,y) : y = 3x\ ,\ 0 < x < 1\}$ is a 1-dimensional manifold (segment) in R^2 .

Let $D = (0,1) \subset R$

$f(t) = (t,3t)\ ,\ 0 < t < 1$

$f'(t) = [{}^1_3]\ ,\ \ \text{rk}\, f' = 1$

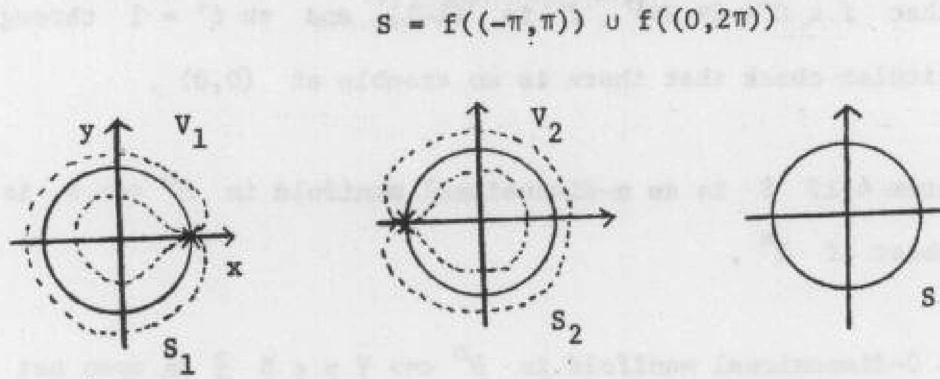


(2) $S = \{(x,y) : x^2 + y^2 = 1\}$ is a 1-dimensional manifold in R^2 .

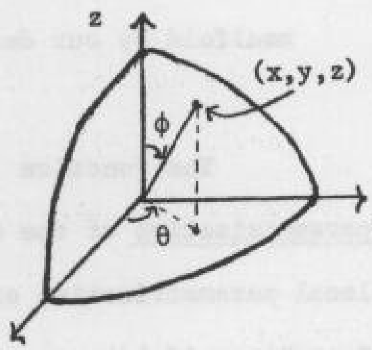
Consider the map $f(t) = (\cos t\ ,\ \sin t)\ ,\ t \in R$,

$\text{rk}\, f'(t) = \text{rk} \begin{bmatrix} -\sin t \\ \cos t \end{bmatrix} = 1$. Thus f is locally (1-1) and is in fact (1-1) on any open interval of length 2π .

$$S = f((-\pi,\pi)) \cup f((0,2\pi))$$



(3) $S = \{(x,y,z) : x^2 + y^2 + z^2 = 1 \ , \ x > 0 \ , \ y > 0 \ , \ z > 0\}$ is a 2-dimensional segment in R^3.



Consider

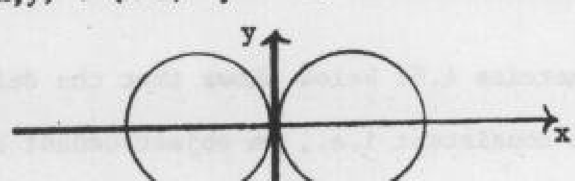
$$f(\theta,\phi) = (\cos\theta \sin\phi, \sin\theta \sin\phi, \cos\phi) = (x,y,z) \ , \quad \begin{array}{l} 0 < \theta < \frac{\pi}{2} \\ 0 < \phi < \frac{\pi}{2} \end{array}$$

$$f \in C^1((0,\tfrac{\pi}{2})\times(0,\tfrac{\pi}{2}))$$

$$f \text{ is } (1\text{-}1) \text{ and } f'(\theta,\phi) = \begin{bmatrix} -\sin\theta\sin\phi & \cos\theta\cos\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ 0 & -\sin\phi \end{bmatrix}$$

$$\operatorname{rk} f'(\theta,\phi) = 2 \text{ since e.g. } \frac{\partial(y,z)}{\partial(\theta,\phi)} = -\cos\theta\sin^2\phi \neq 0, \quad \begin{array}{l} 0 < \theta < \frac{\pi}{2} \\ 0 < \phi < \frac{\pi}{2} \end{array} .$$

(4) $S = \{(x,y) : (x+1)^2+y^2 = 1\} \cup \{(x,y) : (x-1)^2+y^2 = 1\}$ is a 1-dimensional segment in R^2.



Consider $f : (-2\pi,2\pi) \to R^2$ defined by

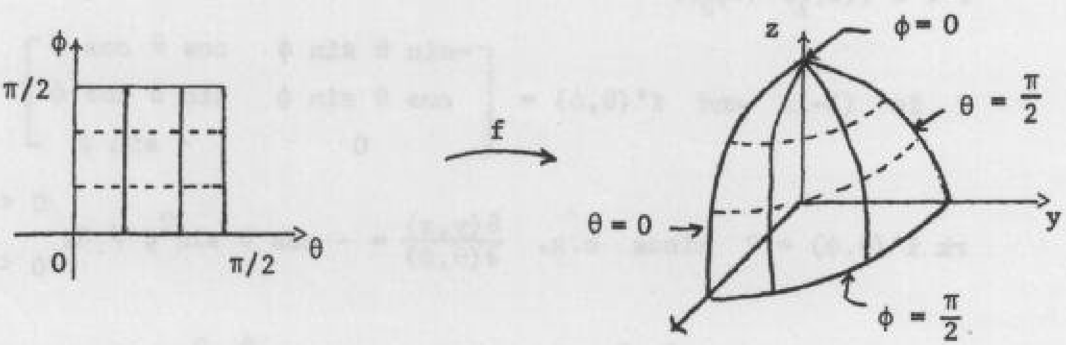
$$f(t) = \begin{cases} (-1+\cos t \ , \sin t) & , \ t \in (-2\pi,0] \\ (1+\cos(\pi-t) \ , \ \sin(\pi-t)) & , \ t \in (0,2\pi) \end{cases}$$

Check that $f \in C^1(-2\pi,2\pi)$, f is (1-1) and $\mathrm{rk}\, f' = 1$ throughout. In particular check that there is no trouble at $(0,0)$.

(5) By Theorem 4.17 S is an n-dimensional manifold in $R^n \iff S$ is an open subset of R^n .

(6) S is a 0-dimensional manifold in $R^n \iff \forall\, p \in S\ \exists$ an open set $V \ni V \cap S = \{p\}$ i.e. S consists of isolated points. Thus $\{\frac{1}{k} : k=1,2,\ldots\}$ is a 0-dimensional manifold in R but $\{\frac{1}{k} : k=1,2,\ldots\} \cup \{0\}$ is not a manifold by our definition.

The function f $(=f(p))$ in part (i) of the definition is called a <u>parametrization</u> of the segment (the variable p being called a parameter). A local parametrization of a manifold is essentially a local coordinate system for the manifold e.g., in Example 3:



Exercise 4.81 below shows that the definition of the dimension of a manifold is consistent i.e., an object cannot be both an r-dimensional and an s-dimensional manifold with $r \neq s$.

In R^2 and R^3 you are familiar with non-parametric representations of manifolds in the form of the equations $f(p) = 0$ (shorthand for $\{p : f(p) = 0\} = f^{-1}(0)$). For example we have seen (Example 2) that $x^2 + y^2 - 1 = 0$ represents a 1-dimensional manifold in R^2. The same equation represents a 2-dimensional manifold in R^3 (a cylinder). The two equations $x^2 + y^2 - 1 = 0$, $z = 0$ together represent a circle in R^3, the intersection of a cylinder and a plane perpendicular to the axis of the cylinder. We make this idea precise in the following theorem.

<u>Theorem 4.20</u>: $f : R^n \to R^k$. Suppose

(i) $f \in C^1(D)$, D open $\subset R^n$, $0 \in f(D)$,

(ii) $\operatorname{rk} f'(p) = k \leq n$, $\forall\, p \in D$.

Then $f^{-1}(0) = \{p : f(p) = 0\}$ is an $(n-k)$-dimensional manifold in R^n.

Less formally the system of equations

$$f_i(x_1,\ldots,x_n) = 0 \quad , \quad i = 1,\ldots,k$$

represents an $(n-k)$-dimensional manifold in R^n if the matrix $[\frac{\partial f_i}{\partial x_j}]$ has rank k.

<u>Proof</u>: If $k = n$ then by Theorem 4.16 f is locally (1-1) on D so if $f(p_o) = 0$ there is a neighbourhood U of $p_o \ni f(p) \neq 0$ if $p \in U$, $p \neq p_o$. Thus $f^{-1}(0)$ is a set of isolated points in R^n - a 0-dimensional manifold.

If $0 < k < n$ and $c = (\gamma_1,\ldots,\gamma_n) \in f^{-1}(0)$ we may assume, without loss of generality, $\frac{\partial(f_1,\ldots,f_k)}{\partial(x_1,\ldots,x_k)}(c) \neq 0$. By the Implicit Function Theorem, the system of equations

$$f_i(x_1,\ldots,x_k,x_{k+1},\ldots,x_n) = 0 \qquad i = 1,\ldots,k$$

may be solved uniquely in the form $x_i = \theta_i\ (x_{k+1},\ldots,x_n)$, $i = 1,\ldots,k$ (with $\gamma_i = \theta_i(\gamma_{k+1},\ldots,\gamma_n)$) in an open neighbourhood U of $(\gamma_{k+1},\ldots,\gamma_n)$. Thus for some neighbhourhood V of c the set $f^{-1}(0) \cap V$ is the range of the function $F : R^{n-k} \to R^n$

$$F(x_{k+1},\ldots,x_n) = (\theta_1(x_{k+1},\ldots,x_n),\ldots,\theta_k(x_{k+1},\ldots,x_n),x_{k+1},\ldots,x_n)$$

$$F' = \begin{array}{c} \left[\begin{array}{ccc} & \theta' & \\ \hline 1 & & 0 \\ & 1 & \\ & & \ddots \\ 0 & & 1 \end{array}\right] \end{array} \begin{array}{l} \}\ k \\ \\ \\ \}\ n-k \\ \end{array} , \ \mathrm{rk}\ F' = n-k \ .$$

F is (1-1) on U if U is sufficiently small, by Theorem 4.16. Thus $f^{-1}(0)$ is an $(n-k)$-dimensional manifold in R^n . □

<u>Remark</u>: The dimension of the range of f (i.e. the number of equations $f_i = 0$) is immaterial here. If $\mathrm{rk}\ f'(p) = k$ on an open set $D \supset f^{-1}(0)$ then $f^{-1}(0)$ is an $(n-k)$-dimensional manifold no matter how many equations there are. You are no doubt familiar with the corresponding statement for linear functions: if $L : R^n \to R^m$ (linear), $Lx = Ax$, where $\mathrm{rk}\ A = k$ then $L^{-1}(\{0\})$ is a vector space of dimension $n-k$ i.e. the solution set of the system $Ax = 0$ has dimension $n-k$. The number $n-k$ is called the <u>nullity</u> of the matrix A .

<u>Corollary 4.20.1</u>: $f : R \to R$. If f is C^1 on its domain, an open subset of R , then the graph of f is a 1-dimensional manifold (smooth curve) in R^2 .

<u>Proof</u>: $\{(x,y) : y = f(x)\} = F^{-1}(0)$, where $F(x,y) = f(x)-y$.

$$F'(x,y) = [f'(x),-1] \quad , \quad \text{rk } F' = 1 \ . \quad \square$$

<u>Corollary 4.20.2</u>: $f : R^2 \to R$. If f is C^1 on its domain, an open subset of R^2 , then the graph of f is a 2-dimensional manifold (smooth surface) in R^3 .

<u>Proof</u>: $\{(x,y,z) : z = f(x,y)\} = F^{-1}(0)$ where $F(x,y,z) = f(x,y)-z$.

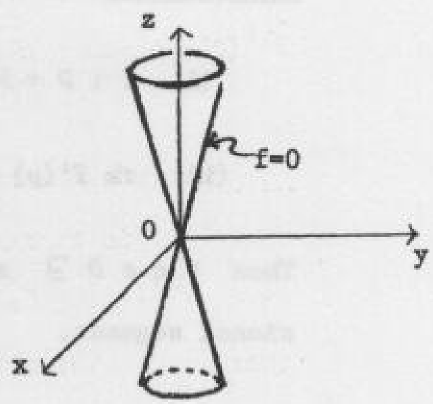
$$F'(x,y,z) = [f_x,f_y,-1] \quad , \quad \text{rk } F' = 1 \ . \quad \square$$

<u>Examples</u>.

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

$f'(x,y,z) = [2x,2y,2z]$

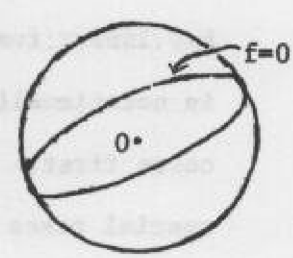
$f = 0$ is a 2-dimensional manifold (smooth surface) in R^3 if $(0,0,0)$ is omitted.



(2) $f = (f_1,f_2)$, $f_1(x,y,z) = x^2+y^2+z^2-1$

$f_2(x,y,z) = Ax+By+Cz$

$$f'(x,y,z) = \begin{bmatrix} 2x & 2y & 2z \\ A & B & C \end{bmatrix}$$



$f = 0$ is a 1-dimensional manifold in R^3 if $\operatorname{rk} f' = 2$ on an open set $D \supset f^{-1}(0)$. Thus if A,B,C are not all zero (i.e. $A^2 + B^2 + C^2 > 0$), we already know that x,y,z are not all zero $(x^2+y^2+z^2=1)$, so $\operatorname{rk} f' < 2 \iff (A,B,C) = \lambda(x,y,z)$, $\lambda \neq 0$. Substituting into $f_2 = 0$ we find $x^2+y^2+z^2 = 0$ if $\operatorname{rk} f' < 2$ contradicting $x^2+y^2+z^2 = 1$. Thus $\operatorname{rk} f' = 2$ at each point in $f^{-1}(0)$ so $\operatorname{rk} f' = 2$ on a neighbourhood of each point in $f^{-1}(0)$. Note that we have assumed $f^{-1}(0) \neq \phi$ here which is obvious geometrically in this case but may not be at all clear in general problems.

<u>Question</u>: Let $A(x) = \begin{bmatrix} a(x) & b(x) & c(x) \\ d(x) & e(x) & f(x) \end{bmatrix}$. Does $\operatorname{rk} A(x_o) = 2 \Rightarrow \operatorname{rk} A(x) = 2$ near x_o if the entries in A are continuous? Does $\operatorname{rk} A(x_o) = 1 \Rightarrow \operatorname{rk} A(x) = 1$ near x_o.

<u>Theorem 4.21</u>: Let $D \subset R^n$, open. Suppose

(i) $f : D \to R^m$, $f \in C^1(D)$,

(ii) $\operatorname{rk} f'(p) = k$, $\forall\, p \in D$.

Then $\forall\, c \in D\ \exists$ a neighbourhood U_o of c such that $f(U_o)$ is a k-dimensional segment.

This is analogous to the statement that the dimension of the range of any linear function $L(x) = Ax$ is the rank of A. The proof of this theorem is notationally quite complicated so we will consider a few examples and special cases first. Actually all the essential ideas of the proof are contained in the special cases so you may skip the proof if you wish.

Examples:

(1) $f : R^2 \to R^2$, $f(x,y) = (x+y,(x+y)^2)$

$$\operatorname{rk} f'(x,y) = \operatorname{rk}\begin{bmatrix} 1 & 1 \\ 2(x+y) & 2(x+y) \end{bmatrix} = 1 \quad \forall\ (x,y)$$

$f(R^2)$ is the 1-dimensional segment $\{(t,t^2) : t \in R\}$, a parabola. We have used $t = x+y$ to parametrize the curve.

(2) $f : R^3 \to R^3$, $f(x,y,z) = (x-y,y-z,x(x-2y)-z(z-2y))$.

$$\operatorname{rk} f'(x,y,z) = \operatorname{rk}\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 2(x-y) & 2(z-x) & 2(y-z) \end{bmatrix} = 2$$

$f(R^3)$ is the 2-dimensional segment $\{(u,v,u^2-v^2) : (u,v) \in R^2\}$ (i.e. the hyperboloid $w = u^2-v^2$) in R^3 . We have used $(u,v) = (x-y,y-z)$ to parametrize the surface.

Proposition: Let $f : R^2 \to R^2$, $f \in C^1(D)$, D open. If $\operatorname{rk} f'(p) = 1$ $\forall\ p \in D$ and $u = f_1(x,y)$, $v = f_2(x,y)$ then in a neighb ourhood of any point $(x_o,y_o) \in D$ either

$$u = \phi(v) \quad \text{or} \quad v = \psi(u) \quad , \quad \phi,\psi \in C^1$$

i.e. $f(D)$ is composed of smooth 1-dimensional segments.

Proof: If $\frac{\partial f_1}{\partial x}(x_o,y_o) \neq 0$ then $\frac{\partial f_1}{\partial x} \neq 0$ in an open neighbhourhood U of (x_o,y_o) so $f_1(U)$ is open (Theorem 4.17). By the IFT, near $(x_o,y_o$, the equation

$$u = f_1(x,y)$$

may be solved uniquely for x in the form

$$x = \theta(u,y) \quad , \quad \theta \in C^1$$

$$\therefore \quad v = f_2(\theta(u,y),y) \; .$$

We show that the function on the right is independent of y :

$$\frac{\partial v}{\partial y} = \frac{\partial f_2}{\partial x}\frac{\partial \theta}{\partial y} + \frac{\partial f_2}{\partial y} = \frac{\partial f_2}{\partial x}\{-\frac{\partial f_1}{\partial y} / \frac{\partial f_1}{\partial x}\} + \frac{\partial f_2}{\partial y}$$

$$= \{\frac{\partial f_1}{\partial x}\frac{\partial f_2}{\partial y} - \frac{\partial f_2}{\partial x}\frac{\partial f_1}{\partial y}\} / \frac{\partial f_1}{\partial x} = \frac{\partial(f_1,f_2)}{\partial(x,y)} / \frac{\partial f_1}{\partial x} = 0$$

since $\operatorname{rk} f' = 1$. Thus $v = \Psi(u)$ i.e. we have a local parametrization of $f(U_o)$ in the form

$$f(U_o) = \{(u,\Psi(u)) : u \in f_1(U_o)\}$$

where U_o is an open neighb ourhood of (x_o,y_o) . $\square$

<u>Proposition</u>: Let $f : R^3 \to R^3$, $f \in C^1(D)$, D open. If $\operatorname{rk} f'(p) = 2$ $\forall\, p \in D$ and

$$u = f_1(x,y,z) \; , \; v = f_2(x,y,z) \; , \; w = f_3(x,y,z)$$

then in some neighbhourhood U of each point $(x_o,y_o,z_o) \in D$ either

$$w = \phi(u,v) \; , \text{ or } \; u = \psi(v,w) \; , \text{ or } \; v = \eta(w,u) \; , \; \phi,\psi,\eta \in C^1 \; .$$

i.e. $f(U)$ is a 2-dimensional segment in R^3 .

<u>Proof</u>: If $\frac{\partial(f_1,f_2)}{\partial(x,y)}(x_o,y_o,z_o) \neq 0$ we may solve

$$u = f_1(x,y,z) \quad , \quad v = f_2(x,y,z)$$

uniquely for (x,y) in the form

$$x = \theta_1(u,v,z) \quad , \quad y = \theta_2(u,v,z) \quad , \quad \theta_1,\theta_2 \in C^1$$

and so $w = f_3(\theta_1(u,v,z),\theta_2(u,v,z),z)$. This function is independent of z

$$\begin{aligned}
\frac{\partial w}{\partial z} &= \frac{\partial f_3}{\partial x}\frac{\partial \theta_1}{\partial z} + \frac{\partial f_3}{\partial y}\frac{\partial \theta_2}{\partial z} + \frac{\partial f_3}{\partial z} \\
&= \frac{\partial f_3}{\partial x}\{-\frac{\partial(f_1,f_2)}{\partial(z,y)} / \frac{\partial(f_1,f_2)}{\partial(x,y)}\} + \frac{\partial f_3}{\partial y}\{-\frac{\partial(f_1,f_2)}{\partial(x,z)} / \frac{\partial(f_1,f_2)}{\partial(x,y)}\} + \frac{\partial f_3}{\partial z} \\
&= \{\frac{\partial f_3}{\partial x}\frac{\partial(f_1,f_2)}{\partial(y,z)} + \frac{\partial f_3}{\partial y}\frac{\partial(f_1,f_2)}{\partial(z,x)} + \frac{\partial f_3}{\partial z}\frac{\partial(f_1,f_2)}{\partial(x,y)}\} / \frac{\partial(f_1,f_2)}{\partial(x,y)} \\
&= \frac{\partial(f_1,f_2,f_3)}{\partial(x,y,z)} / \frac{\partial(f_1,f_2)}{\partial(x,y)} = 0 \text{ since } \operatorname{rk} f' = 2 .
\end{aligned}$$

Therefore $w = \phi(u,v)$ i.e. $f(U_o) = \{(u,v,\phi(u,v)) : (u,v) \in (f_1,f_2)(U_o)\}$ for some open neighb ourhood U_o of (x_o,y_o,z_o) . □

<u>Proof of Theorem 4.21</u> (Optional): If $k = n$ then f is locally (1-1) (Theorem 4.16) so the theorem reduces to the definition of a k-dimensional segment. If $k = 0$ then $\frac{\partial f_i}{\partial x_j} = 0 \ \forall\, i,j$ so f is constant on a convex neighbourhood U of c i.e. $f(U) = \{f(c)\}$ a 0-dimensional segment. When $0 < k < n$ consider

(1) $$u_i = f_i(x_1,\ldots,x_n) \quad , \quad i = 1,\ldots,m \ .$$

We are given $\operatorname{rk} f'(p) = k$ so if $c \in D$ and

(2) $$\frac{\partial(f_1,\ldots,f_k)}{\partial(x_1,\cdots,x_k)}(p) \neq 0 \ , \quad \text{at} \quad p = c \ ,$$

then (2) holds also in an open neighb ourhood U of c $(f \in C^1)$ so $V = (f_1,\ldots,f_k)(U)$ is open (Corollary 4.17.1). Condition (2) implies (by the IFT) that the first k equations in (1) may be solved for $(x_1,\ldots,x_k)$ in the form

(3) $$x_i = \theta_i(u_1,\ldots,u_k,x_{k+1},\ldots,x_k) \ , \quad i = 1,\ldots,k \ .$$

Substituting (3) into the remaining (n-k) equations in (1):

(4) $$u_i = f_i(\theta_1,\ldots,\theta_k,x_{k+1},\ldots,x_n) \ , \quad i = k+1,\ldots,n \ .$$

These are functions of $(u_1,\ldots,u_n)$ only e.g.

$$\frac{\partial u_i}{\partial x_{k+1}} = \sum_{j=1}^{k} \frac{\partial f_i}{\partial x_j}\left\{-\frac{\partial(f_1,\ldots\ldots\ldots\ldots,f_k)}{\partial(x_1,\ldots,x_{k+1},\ldots,x_k)} \Big/ \frac{\partial(f_1,\ldots,f_k)}{\partial(x_1,\ldots,x_k)}\right\} + \frac{\partial f_i}{\partial x_{k+1}}$$

$$= \left\{\sum_{j=1}^{k+1} (-1)^{k+1-j} \frac{\partial f_i}{\partial x_j} \frac{\partial(f_1,\ldots\ldots\ldots\ldots,f_k)}{\partial(x_1,\ldots,\hat{x}_j,\ldots,x_{k+1})}\right\} \Big/ \frac{\partial(f_1,\ldots,f_k)}{\partial(x_1,\ldots,x_k)}$$

$$= \frac{\partial(f_1,\ldots,f_k,f_i)}{\partial(x_1,\ldots,x_{k+1})} \Big/ \frac{\partial(f_1,\ldots,f_k)}{\partial(x_1,\ldots,x_k)} = 0 \quad (\text{since } \operatorname{rk} f' = k)$$

where $(x_1,\ldots,\hat{x}_j,\ldots,x_{k+1})$ denotes the k-tuple obtained by omitting x_j from $(x_1,\ldots,x_{k+1})$. Thus (4) may be written in the form

(5) $$u_i - \phi_i(u_1,\ldots,u_k) = 0 \ , \ i = k+1,\ldots,n \ , \ \phi_i \in C^1(V_o)$$

where $V_o = (f_1,\ldots,f_k)(U_o)$ is open. The Jacobian matrix of the $(n-k)$ equations (5) is of the form

$$
n-k\left\{\underbrace{\begin{bmatrix} 1 & & & 0 \\ & 1 & & \\ 0 & & \ddots & \\ & & & 1 \end{bmatrix}}_{n-k}\;\underbrace{\begin{bmatrix} \phi' \end{bmatrix}}_{k}\right.
$$

which has rank $(n-k)$. Therefore (by Theorem 4.20) $f(U_o)$ is a k-dimensional manifold.

[Notice that the pen never left the hand!!]

<u>Application</u> (The Lagrange Multiplier Rule):

<u>Definition</u>: $f : R^n \to R$, domain $D \subset R^n$. f has a <u>relative</u> <u>maximum</u> (<u>minimum</u>) with respect to the set S at $p_o \in S \cap D$ if, for some neighb ourhood U of p_o

$$f(p) \leq f(p_o) \quad (f(p) \geq f(p_o)) \quad \forall\, p \in U \cap S \ .$$

<u>Theorem 4.22</u> (Lagrange Multiplier Rule): Let $f : R^n \to R$ and $g = (g_1,\ldots,g_k) : R^n \to R^k$ $(k < n)$. If f and g are C^1 functions near p_o and

(i) f has a relative extremum at p_o with respect to the set

$$S = g^{-1}(0) = \{p : g_i(p) = 0\ ,\ i = 1,\ldots,k\}$$

(ii) $\operatorname{rk} g'(p_o) = k$

then there exist real numbers $\lambda_1,\ldots,\lambda_k$ such that

$$\frac{\partial f}{\partial x_i}(p_o) + \sum_{j=1}^{k} \lambda_j \frac{\partial g_j}{\partial x_i}(p_o) = 0 \ , \ i = 1,\ldots,n \ .$$

The equations $g_i(p) = 0$, $i = 1,\ldots,k$ are called <u>constraints</u>. The numbers λ_i are <u>Lagrange</u> <u>multipliers</u>.

<u>Examples</u>.

(1) Find the distance d from the point $(0,0)$ to the line $x+y = 1$, i.e. minimize $x^2 + y^2$ subject to the constraint $x+y-1 = 0$.

A minimum exists since $\lim(x^2+y^2) = \infty$, $(|(x,y)| \to \infty)$ and $x^2+y^2 > 0$. At the minimum LMR => $\exists \ \lambda \ni$

$$2x+\lambda = 0 \ , \ 2y+\lambda = 0 \ , \ x+y-1 = 0 \ .$$

Solving: $(x,y,\lambda) = (\frac{1}{2} , \frac{1}{2} , -1)$

$$\therefore \quad d^2 = (\tfrac{1}{2})^2 + (\tfrac{1}{2})^2 = \tfrac{1}{2} \quad \text{i.e. } d = \frac{1}{\sqrt{2}} \ .$$

(2) Find the maximum and minimum of $f(x,y) = 2x^2-3y^2-2x$ in the set $\{(x,y) : x^2 + y^2 \leq 1\}$.

If these occur in $\{(x,y) : x^2 + y^2 < 1\}$ then (Theorem 4.14)

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= 4x-2 = 0 \\ \frac{\partial f}{\partial y} &= -6y \ = 0 \end{aligned}\right\} \Rightarrow \ (x,y) = (\tfrac{1}{2},0) \ .$$

But

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & -6 \end{bmatrix}$$ is <u>indefinite</u>

so f has a minimax (saddle) at $(\frac{1}{2},0)$. The maximum and minimum occur in the set

$$\{(x,y) : x^2 + y^2 - 1 = 0\}$$

$$\text{LMR} \Rightarrow \quad 4x - 2 + 2\lambda x = 0$$

$$- 6y + 2\lambda y = 0$$

$$x^2 + y^2 - 1 = 0$$

Solutions are $(x,y,\lambda) = (1,0,-1)$, $(-1,0,-3)$ or $(\frac{1}{5}, \pm\sqrt{\frac{24}{25}}, 3)$.

$$f(1,0) = 0 \ , \ f(-1,0) = 4 \ , \ f(\tfrac{1}{5}, \pm\sqrt{\tfrac{24}{25}}) = -\frac{16}{5}$$

$$\therefore \text{Max } f = 4 \quad , \quad \text{Min } f = -\frac{16}{5} \quad .$$

Notice that the minimum is achieved at two points.

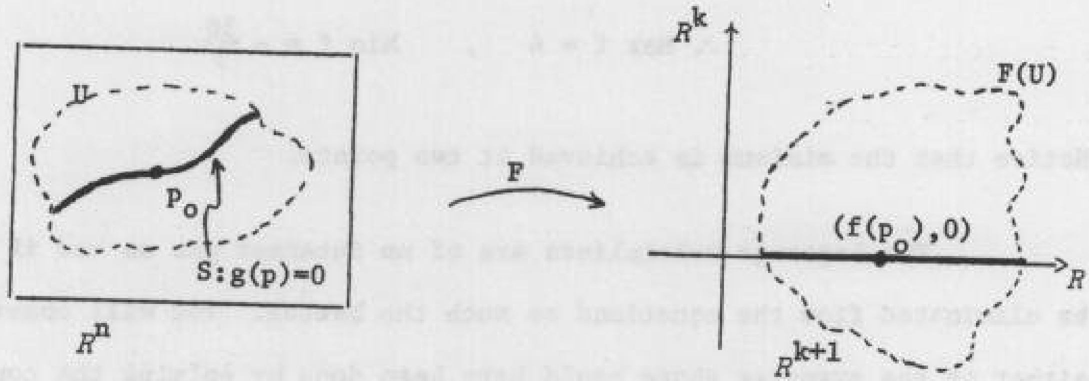
The Lagrange multipliers are of no interest per se and if they can be eliminated from the equations so much the better. You will observe that either of the examples above could have been done by solving the constraints to reduce the dimension of the problem and applying the standard methods e.g. in Example (1) $y = 1-x$ and $x^2+y^2 = x^2 + (1-x)^2 = 2x^2-2x+1$ which has a minimum $\frac{1}{2}$ at $x = \frac{1}{2}$. This procedure leads to unwieldy calculations in general but we will use it in one of the proofs of the rule.

<u>Proof 1 of Theorem 4.22</u>. Let f have a relative extremum at p_o with respect to the set $S = \{p : g(p) = 0\}$. Consider the function $F : R^n \to R^{k+1}$

$$F(p) = (f(p), g(p)) = (f(p), g_1(p), \ldots, g_k(p))$$

$$F'(p) = \begin{bmatrix} f'(p) \\ \hline g'(p) \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \end{bmatrix}_p$$

If $\text{rk } F'(p_o) = k+1$ then $\text{rk } F'(p) = k+1$ for each p in a sufficiently small open neighbourhood U of p_o. Then for each such U, $F(U)$ is a neighbourhood of $F(p_o) = (f(p_o, g(p_o)) = (f(p_o), 0)$ (4.17).



Hence every neighbourhood U of p_o contains points p_i, $i = 1,2$ such that $F(p_i) = (f(p_i), 0)$ i.e. $g(p_i) = 0$ and $f(p_1) > f(p_o)$, $f(p_2) < f(p_o)$ contradicting our assumption (i) that f has an extremum with respect to S at p_o. Thus $\text{rk } F'(p_o) = k+1$ leads to a contradiction so $\text{rk } F'(p_o) < k+1$ and in fact rk $F'(p_o) = k$ since rk $g'(p_o) = k$ by our assumption (ii). The rows of $F'(p_o)$ are therefore linearly dependent i.e. $\exists\ \lambda_i$, $i = 0,\ldots,k$, $\ni$

$$\lambda_o\left(\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n}\right)_{p_o} + \lambda_1\left(\frac{\partial g_1}{\partial x_1},\ldots,\frac{\partial g_1}{\partial x_n}\right)_{p_o} + \ldots + \lambda_k\left(\frac{\partial g_k}{\partial x_1},\ldots,\frac{\partial g_k}{\partial x_n}\right)_{p_o}$$

$$= (0,\ldots,0)\ ,\ \lambda_i \text{ not all zero.}$$

$\lambda_o \neq 0$ since $\lambda_o = 0 \Rightarrow$ rk $g'(p_o) < k$, contrary to (ii) so we may assume $\lambda_o = 1$ (divide by λ_o) to obtain

$$\frac{\partial f}{\partial x_i} + \lambda_1\frac{\partial g_1}{\partial x_i} + \ldots + \lambda_k\frac{\partial g_k}{\partial x_i} = 0 \text{ at } p_o\ ,\ i = 1,\ldots,n\ .$$

These n equations together with the k constraint equations

$$g_i(p_o) = 0\ ,\ i = 1,\ldots,k\ ,$$

serve to determine the $(n+k)$ numbers which are the coordinates of p_o and $\lambda_1,\ldots,\lambda_k$. □

<u>Proof 2 of Theorem 4.22</u>: This proof works in general but we prove just the case of one constraint. Suppose $f,g : R^n \to R$ are C^1 functions and $f(x_1,\ldots,x_n)$ has an extremum with respect to the constraint

(1) $$g(x_1,\ldots,x_n) = 0$$

at the point $c = (\gamma_1,\ldots,\gamma_n)$. If rk $g'(c) = 1$ then we may assume $\frac{\partial g}{\partial x_1}(c) \neq 0$ so the equation (1) may be solved for x_1 in a neighbourhood of $(\gamma_2,\ldots,\gamma_n)$ in the form $x_1 = \phi(x_2,\ldots,x_n)$ with $\phi \in C^1$ and $\gamma_1 = \phi(\gamma_2,\ldots,\phi_n)$. Therefore $f(\phi(x_2,\ldots,x_n),x_2,\ldots,x_n)$ has an interior relative extremum at $\tilde{c} = (\gamma_2,\ldots,\gamma_n)$ so, by the Chain Rule

$$\frac{\partial f}{\partial x_i}(c) + \frac{\partial f}{\partial x_1}(c)\,\frac{\partial \phi}{\partial x_i}(\tilde{c}) = 0 \quad , \quad i = 2,\ldots,n$$

i.e.,

$$\frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_1}\left\{-\frac{\partial g}{\partial x_i} \Big/ \frac{\partial g}{\partial x_1}\right\} = 0 \quad \text{at} \quad c \ , \ i = 2,\ldots,n \ .$$

This formula holds trivially for $i = 1$. Hence

$$\frac{\partial f}{\partial x_i}(c) + \lambda\,\frac{\partial g}{\partial x_i}(c) = 0 \quad , \quad i = 1,\ldots,n \ ,$$

where $\lambda = -\frac{\partial f}{\partial x_1}(c) \Big/ \frac{\partial g}{\partial x_1}(c)$. □

Exercises:

4.76: Let $f(x,y) = (x+y, 2x+ay) = (u,v)$.

(i) Show that $\mathrm{rk}\, f'(x,y) = 2 \iff a \neq 2$

(ii) Find the image of the square $K = \{(x,y) : 0 \leq x \leq 1 \ , \ 0 \leq y \leq 1\}$ in the three cases $a = 1$, $a = 2$, $a = 3$.

(iii) If a point moves around ∂K in the counterclockwise sense find what its image does in the cases $a = 1$, $a = 3$. What is the sign of J_f in each case?

4.77: Let $f(x,y) = (x, xy) = (u,v)$.

Draw some curves u = constant, v = constant in the (x,y) plane and some curves x = constant, y = constant in the (u,v) plane? Is this map (1-1) ? Into what region of the (u,v) plane does f map the rectangle $\{(x,y) : 1 \leq x \leq 2 \ , \ 0 \leq y \leq 2\}$? What points in the (x,y) plane map into the rectangle $\{(u,v) : 1 \leq u \leq 2 \ , \ 0 \leq v \leq 2\}$?

4.78: If $f(x,y) = (\cos(x+y), \sin(x+y)) = (u,v)$, show from $f'(x,y)$ that (u,v) lies on a <u>curve</u> in the (u,v) plane. What is its equation?

4.79: In each of the following show that (u,v,w) lies on a surface in R^3 and find an equation for the surface.

(i) $u = \frac{x+y}{z}$, $v = \frac{z+y}{x}$, $w = \frac{y(x+y+z)}{xz}$

(ii) $u = x+y+z$, $v = xy + yz + zx$, $w = x^3 + y^3 + z^3 - 3xyz$

(iii) $u = x+y+z$, $v = x^2 + y^2 + z^2$, $w = x^3 + y^3 + z^3 - 3xyz$.

[Solutions: (i) $w+1 = uv$, (ii) $w = u(u^2-3v)$, (iii) $w = \frac{u}{2}(3v-u^2)$.

4.80: Show that $\{(u \cos v, u \sin v, u) : 0 < u < 1 , 0 < v < 2\pi\}$ is a 2-dimensional segment in R^3 . Draw a picture of it.

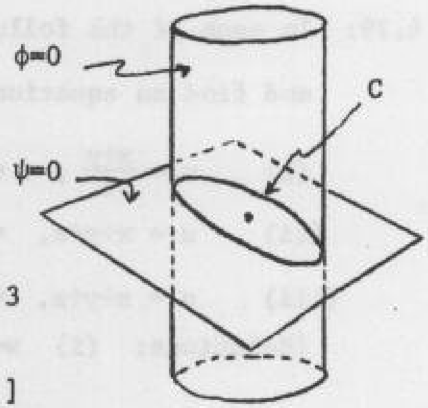
4.81: Show that a set S cannot be both an r dimensional segment and an s dimensional segment if $r \neq s$. [Suppose $S = \phi(U) = \psi(V)$ where ϕ, ψ are C^1 functions on U, V (resp.), open sets in R^r , R^s , $r < s$. Show that if ϕ, ψ are (1-1) and $\operatorname{rk} \phi' = r$, $\operatorname{rk} \psi' = s$ there is a contradiction. Use the Implicit Function Theorem.]

4.82: If ϕ_i are C^1 functions on an open set $D \subset R^n$ $(i = 1,\dots,k)$ then the set of points $(x_1,\dots,x_n,y_1,\dots,y_k)$ in R^{n+k} satisfying

$$y_i = \phi_i(x_1,\dots,x_n) \text{ , } i = 1,\dots,k$$

is an n-dimensional segment in R^{n+k} . [i.e. the graph of a C^1 function from $R^n \to R^k$ is an n-dimensional manifold if its domain is open.]

4.83: Let $V = x^2+y^2+z^2$, $\phi = x^2+y^2-1$, $\psi = x+y+z$. Find the maximum and minimum of V subject to the constraints $\phi = 0$, $\psi = 0$. This means: "Find the length (squared) of the major and minor semiaxes of the ellipse C which is the intersection of the cylinder $\phi = 0$ and the plane $\psi = 0$.



[Solution: $\max V = V(\mp \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}, \pm \sqrt{2}) = 3$

$\min V = V(\pm \frac{1}{\sqrt{2}}, \mp \frac{1}{\sqrt{2}}, 0) = 1$.]

4.84: Find the minimum value of $V = x^2+y^2+z^2$ when

(i) $x+y+z = 3a$ [Solution: $3a^2$ at (a,a,a).]

(ii) $xy+yz+zx = 3a^2$ [Solution: $3a^2$ at $(a,a,a),(-a,-a,-a)$.]

(iii) $xyz = a^3$ [Solution: $3a^2$ at $(a,a,a),(-a,-a,a),(-a,a,-a),(a,-a,-a)$.]

4.85: Find the maximum and minimum values of $x^2+y^2-3x+5y$ when $(x+y)^2 = 4(x-y)$.

4.86: Find the maximum and minimum values of $2x^2+y^2+2x$ if $|x| + |y| \leq 1$.

4.87: Do exercise 4.52 again using the Lagrange Multiplier Rule.

4.88: Prove that the stationary values of $V = x^2+y^2+z^2$, subject to the constraints $\ell x + my + nz = 0$, $ax^2 + by^2 + cz^2 = 1$ are given by the quadratic

$$\frac{\ell^2}{1-aV} + \frac{m^2}{1-bV} + \frac{n^2}{1-cV} = 0 \quad .$$

4.89: (a) Show that $(x_1\, x_2 \ldots x_n)^2 \leq \frac{1}{n^n}$ if $x_1^2 + x_2^2 + \ldots + x_n^2 = 1$.

(b) [Arithmetic-geometric mean inequality.] Show that

$$(a_1\, a_2 \ldots a_n)^{1/n} \leq \frac{1}{n}(a_1+a_2+\ldots+a_n)$$

if $a_i \geq 0$, $i = 1,\ldots,n$.

4.90: Given two smooth plane curves

$$f(x,y) = 0 \quad , \quad g(x,y) = 0$$

show that when the distance between points (α,β) and (ξ,η) lying on the respective curves has an extremum then

$$\frac{\alpha-\xi}{\beta-\eta} = \frac{f_x(\alpha,\beta)}{f_y(\alpha,\beta)} = \frac{g_x(\xi,\eta)}{g_y(\xi,\eta)} \quad .$$

Use this to find the shortest distance between the ellipse $x^2 + 2xy + 5y^2 - 16y = 0$ and the line $x + y - 8 = 0$.

4.91: Let f be a real valued function of class C^1 on R^3. Prove that there are at least two points on the sphere $x^2+y^2+z^2 = R^2 > 0$ at which the equations

$$y\frac{\partial f}{\partial x} - x\frac{\partial f}{\partial y} = 0$$

$$z\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial z} = 0$$

$$x\frac{\partial f}{\partial z} - z\frac{\partial f}{\partial x} = 0$$

are satisfied.

4.92: [The Hölder and Minkowski inequalities.]

Let $p > 1$, $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$.

(a) Show that the minimum of $f(x,y) = \frac{x^p}{p} + \frac{y^q}{q}$ subject to the constraints $x > 0$, $y > 0$, $xy = 1$, is 1.

(b) $a \geq 0$, $b \geq 0 \Rightarrow ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.

(c) $a_k \geq 0$, $b_k \geq 0$, $k = 1,\ldots,n \Rightarrow$

$$\sum_{k=1}^{n} a_k b_k \leq \left(\sum_{k=1}^{n} a_k^p\right)^{1/p} \left(\sum_{k=1}^{n} b_k^q\right)^{1/q} \quad \text{(Hölder's Inequality)}$$

[Hint: Let $A = \left(\sum_{k=1}^{n} a_k^p\right)^{1/p}$, $B = \left(\sum_{k=1}^{n} b_k^q\right)^{1/q}$ and consider $a = \frac{a_k}{A}$ $b = \frac{b_k}{B}$.]

(d) If a_k, $b_k \in R$, $k = 1,\ldots,n$ and $p \geq 1$ then

$$\left(\sum_{k=1}^{n} |a_k+b_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^{n} |a_k|^p\right)^{1/p} + \left(\sum_{k=1}^{n} |b_k|^p\right)^{1/p},$$

(Minkowski's Inequality).

[Hint: $|a+b|^p = |a+b||a+b|^{p/q} \leq |a||a+b|^{p/q} + |b||a+b|^{p/q}$ if $p > 1$.
Use Hölder's Inequality.]

References for Chapter IV

R.G. Bartle: Chapter V.

R.C. Buck: Chapter V.

Index of Math 217 lecture notes