MATHEMATICAL ANALYSIS

Elias Zakon
University of Windsor

LibreTexts™

University of Windsor

Book: Mathematical Analysis (Zakon)

The LibreTexts libraries are Powered by MindTouch® and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739. Unless otherwise noted, LibreTexts content is licensed by CC BY-NC-SA 3.0.

This text was compiled on 11/16/2022

TABLE OF CONTENTS

8: Measurable Functions and Integration

About the Author

Elias Zakon was born in Russia under the czar in 1908, and he was swept along in the turbulence of the great events of twentieth-century Europe. Zakon studied mathematics and law in Germany and Poland, and later he joined his father's law practice in Poland. Fleeing the approach of the German Army in 1941, he took his family to Barnaul, Siberia, where, with the rest of the populace, they endured five years of hardship. The Leningrad Institute of Technology was also evacuated to Barnaul upon the siege of Leningrad, and there he met the mathematician I. P. Natanson; with Natanson's encouragement, Zakon again took up his studies and research in mathematics.

Zakon and his family spent the years from 1946 to 1949 in a refugee camp in Salzburg, Austria, where he taught himself Hebrew, one of the six or seven languages in which he became fluent. In 1949, he took his family to the newly created state of Israel and he taught at the Technion in Haifa until 1956. In Israel he published his first research papers in logic and analysis. Throughout his life, Zakon maintained a love of music, art, politics, history, law, and especially chess; it was in Israel that he achieved the rank of chess master.

In 1956, Zakon moved to Canada. As a research fellow at the University of Toronto, he worked with Abraham Robinson. In 1957, he joined the mathematics faculty at the University of Windsor, where the first degrees in the newly established Honors program in Mathematics were awarded in 1960. While at Windsor, he continued publishing his research results in logic and analysis. In this post-McCarthy era, he often had as his house-guest the prolific and eccentric mathematician Paul Erdös, who was then banned from the United States for his political views. Erdös would speak at the University of Windsor, where mathematicians from the University of Michigan and other American universities would gather to hear him and to discuss mathematics.

While at Windsor, Zakon developed three volumes on mathematical analysis, which were bound and distributed to students. His goal was to introduce rigorous material as early as possible; later courses could then rely on this material. We are publishing here the latest complete version of the second of these volumes, which was used in a two-semester class required of all secondyear Honours Mathematics students at Windsor.

Licensing

A detailed breakdown of this resource's licensing can be found in ***Back Matter/Detailed Licensing***.

CHAPTER OVERVIEW

1: Set Theory

1.1: Sets and Operations on Sets. Quantifiers

Sets and Operations on Sets

A *set* is a collection of objects of any specified kind. Sets are usually denoted by capitals. The objects belonging to a set are called its *elements* or *members*. We write $x \in A$ if x is a member of A, and $x \notin A$ if it is not.

$A = \{a, b, c, \dots\}$ means that A consists of the elements $a, b, c, \dots$. In particular, $A = \{a, b\}$ consists of a and b; $A = \{p\}$ consists of p alone. The *empty* or *void* set, $\emptyset$, has *no* elements. Equality $(=)$ means *logical identity.*

If all members of A are also in B, we call A a *subset* of B (and B is a *superset* of A), and write $A \subseteq B$ or $B \supseteq A$. It is an axiom that *the sets A and B are equal* $(A = B)$ *if they have the same members*, i.e.,

$$A \subseteq B \quad \text{and} \quad B \subseteq A. \tag{1.1.1}$$

If, however, $A \subseteq B$ but $B \nsubseteq A$ (i.e., B has some elements not in A), we call A a proper subset of B and write $A \subset B$ or $B \supset A$. "$\subseteq$" is called the inclusion relation.

Set equality is not affected by the order in which elements appear. Thus $a, b = b, a$. Not so for ordered pairs (a, b). For such pairs,

$$(a, b) = (x, y) \quad \text{iff} \quad a = x \quad \text{and} \quad b = y, \tag{1.1.2}$$

but not if $a = y$ and $b = x$. Similarly, for *ordered n-tuples*,

$$(a_1, a_2, \dots, a_n) = (x_1, x_2, \dots, x_n) \quad \text{iff} \quad a_k = x_k, k = 1, 2, \dots, n. \tag{1.1.3}$$

We write $x|P(x)$ for "the set of all x satisfying the condition $P(x)$." Similarly, $(x, y)|P(x, y)$ is the set of all *ordered pairs* for which $P(x, y)$ holds; $x \in A|P(x)$ is the set of those x *in* A for which $P(x)$ is true.

For any sets A and B, we define their *union* $A \cup B$, *intersection* $A \cap B$, *difference* $A = B$, and *Cartesian product* (or *cross product*) $A \times B$, as follows:

$A \cup B$ is the set of all members of A and B taken *together*:

$$\{x | x \in A \text{ or } x \in B\}. \tag{1.1.4}$$

$A \cap B$ is the set of all *common* elements of A and B:

$$\{x \in A | x \in B\}. \tag{1.1.5}$$

$A - B$ consists of those $x \in A$ that are *not* in B:

$$\{x \in A | x \notin B\}. \tag{1.1.6}$$

$A \times B$ is the set of all *ordered pairs* (x, y), with $x \in A$ and $y \in B$:

$$\{(x, y)\ x \in A, y \in B\}. \tag{1.1.7}$$

Similarly, $A_1 \times A_2 \times \dots \times A_n$ is the set of all *ordered n-tuples* $(x_1, \dots, x_n)$ such that $x_k \in A_k, k = 1, 2, \dots, n$. We write A^n for $A \times A \times \dots \times A$ (n factors).

A and B are said to be *disjoint* iff $A \cap B = \emptyset$ (no common elements). Otherwise, we say that A *meets* B ($A \cap B \neq \emptyset$). Usually all sets involved are subsets of a "*master set*" S, called the *space*. Then we write $-X$ for $S - X$, and call $-X$ the *complement* of X (in S). Various other notations are likewise in use.

Example 1.1.1

Let $A = \{1, 2, 3\}, B = \{2, 4\}$. Then

$$A \cup B = \{1, 2, 3, 4\}, \quad A \cap B - \{2\}, \quad A - B = \{1, 3\}, \tag{1.1.8}$$

$$A \times B = \{(1, 2), (1, 4), (2, 2), (2, 4), (3, 2), (3, 4)\}. \tag{1.1.9}$$

If N is the set of all *naturals* (positive integers), we could also write

$$A = \{x \in N | x < 4\}. \tag{1.1.10}$$

Theorem 1.1.1

a. $A \cup A = A; A \cap A = A$;
b. $A \cup B = B \cup A, A \cap B = B \cap A$;
c. $(A \cup B) \cup C = A \cup (B \cup C); (A \cap B) \cap C = A \cap (B \cap C)$;
d. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$;
e. $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.

Proof

The proof of (d) is sketched out in Problem 1. The rest is left to the reader.

Because of (c), we may omit brackets in $A \cup B \cup C$ and $A \cap B \cap C$; similarly for four or more sets. More generally, we may consider whole *families* of sets, i.e., collections of many (possibly infinitely many) sets. If $\mathcal{M}$ is such a family, we define its *union*, $\bigcup \mathcal{M}$, to be the set of all elements x, each belonging to *at least one* set of the family. The intersection of $\mathcal{M}$, denoted $\bigcap \mathcal{M}$, consists of those x that belong to all sets of the family *simultaneously*. Instead, we also write

$$\bigcup\{X | X \in \mathcal{M}\} \quad \text{and} \quad \bigcap\{X | X \in \mathcal{M}\}, \quad \text{respectively.} \tag{1.1.11}$$

Often we can *number* the sets of a given family:

$$A_1, A_2, \ldots, A_n, \ldots \tag{1.1.12}$$

More generally, we may denote all sets of a family $\mathcal{M}$ by some letter (say, X) with indices i attached to it (the indices may, but *need not*, be numbers). The family $\mathcal{M}$ then is denoted by $\{X_i\}$ or $\{X_i | i \in I$, where i is a variable index ranging over a suitable set I of indices ("index notation"). In this case, the union and intersection of $\mathcal{M}$ are denoted by such symbols as

$$\bigcup\{X_i | i \in I\} = \bigcup_i X_i = \bigcup X_i = \bigcup_{i \in I} X_i; \tag{1.1.13}$$

$$\bigcap\{X_i | i \in I\} = \bigcap_i X_i = \bigcap X_i = \bigcap_{i \in I} X_i; \tag{1.1.14}$$

If the indices are *integers*, we may write

$$\bigcup_{n=1}^{m} X_n, \bigcup_{n=1}^{\infty} X_n, \bigcup_{n=k}^{m} X_n, \text{etc.} \tag{1.1.15}$$

Theorem 1.1.1: De Morgan's Duality Laws

For any sets S and A_i $(i \in I)$, the following are true:

$$(i) \;\; S - \bigcup_i A_i = \bigcap_i (S - A); \quad (ii) \;\; S - \bigcap_i A_i = \cup_i (S - A_i). \tag{1.1.16}$$

(If S is the entire space, we may write $-A_i$ for $S - A_i$, $-\bigcup A_i$ for $S - \bigcup A_i$, etc.

Before proving these laws, we introduce some useful notation.

Logical Quantifiers

From logic we borrow the following abbreviations.

"$(\forall x \in A)$..." means "For each member x of A, it is true that . . ."

"$(\exists x \in A)$..." means "There is at least one x in A such that . . ."

"$(\exists! x \in A)$..." means "There is a *unique* x in A such that . . ."

The symbols "$(\forall x \in A)$" and "$(\exists x \in A)$" are called the *universal* and *existential quantifiers*, respectively. If confusion is ruled out, we simply write "$(\forall x)$," "$(\exists x)$," and "$(\exists! x)$" instead. For example, if we agree that m, and n denote *naturals*, then

$$\text{"} \; (\forall n)(\exists m) \;\; m > n \, \text{"} \tag{1.1.17}$$

means "For each natural n, there is a natural m such that $m > n$." We give some more examples.

Let $\mathcal{M} = \{A_i | i \in I\}$ be an indexed set family. By definition, $x \in \bigcup A_i$ means that x is in *at least one* of the sets A_i; in symbols,

$$(\exists i \in I) \;\; x \in A_i. \tag{1.1.18}$$

Thus we note that

$$x \in \bigcup_{i \in I} A_i \;\text{ iff }\; [(\exists i \in I) x \in A_i]. \tag{1.1.19}$$

Similarly,

$$x \in \bigcap_{i} A_i \;\text{ iff }\; [(\forall i \in I) x \in A_i]. \tag{1.1.20}$$

Also note that $x \notin \bigcup A_i$ iff x is in *none* of the A_i, i.e.,

$$(\forall i) \;\; x \notin A_i. \tag{1.1.21}$$

Similarly, $x \notin \bigcap A_i$ iff x fails to be in *some* A_i. i.e.,

$$(\exists i) \;\; x \notin A_i. \;\; (Why?) \tag{1.1.22}$$

We now use these remarks to prove Theorem 2(i). We have to show that $S - \bigcup A_i$ has the same elements as $\bigcap(S - A_i)$, i.e., that $x \in S - \bigcup A_i$ iff $x \in \bigcap(S - A_i)$. But, by our definitions, we have

$$\begin{aligned} x \in S - \bigcup A_i &\iff [x \in S, x \notin \bigcup A_i] \\ &\iff (\forall i)[x \in S, x \notin A_i] \\ &\iff (\forall i) x \in S - A_i \\ &\iff x \in \bigcap (S - A_i), \end{aligned} \tag{1.1.23}$$

as required.

One proves part (ii) of Theorem 2 quite similarly. (Exercise!)

We shall now dwell on quantifiers more closely. Sometimes a formula $P(x)$ holds not for all $x \in A$, but only for those with an additional property $Q(x)$.

This will be written as

$$(\forall x \in A | Q(x)) \;\; P(x), \tag{1.1.24}$$

where the vertical stroke stands for "such that." For example, if N is again the naturals, then the formula

$$(\forall x \in N | x > 3) \;\; x \geq 4 \tag{1.1.25}$$

means "for each $x \in N$ *such that* $x \geq 4$." In other words, for naturals, $x > 3 \implies x \geq 4$ (the arrow stands for "implies"). Thus (1) can also be written as

$$(\forall x \in N) \;\; x > 3 \implies x \geq 4. \tag{1.1.26}$$

In mathematics, we often have to form the *negation* of a formula that starts with one or several quantifiers. It is noteworthy, then, that *each universal quantifier is replaced by an existential one (and vice versa),* followed by the negation of the subsequent part of the formula. For example, in calculus, a real number p is called the *limit* of a sequence $x_1, x_2, \ldots, x_n, \ldots$ iff the following is true:

For every real $\epsilon > 0$, there is a natural k (depending on ϵ) such that, for all natural $n > k$, we have $|x_n - p| < \epsilon|$.

If we agree that lower case letters (possibly with subscripts) denote real numbers, and that n, k denote naturals $(n, k \in \mathbb{N})$, this sentence can be written as

$$(\forall \epsilon > 0)(\exists k)(\forall n > k) \;\; |x_n - p| < \epsilon. \tag{1.1.27}$$

Here the expressions "$(\forall \epsilon > 0)$" and "$(\forall n > k)$" stand for "$(\forall \epsilon | \epsilon > 0)$" and "$(\forall n | n > k)$," respectively (such self-explanatory abbreviations will also be used in other similar cases).

Now, since (2) states that "for all $\epsilon > 0$" something (i.e., the rest of (2)) is true, the negation of (2) starts with "*there is an* $\epsilon > 0$" (for which the rest of the formula *fails*). Thus we start with "$(\exists \epsilon > 0)$," and form the negation of what follows, i.e., of

$$(\exists k)(\forall n > k) \ \ |x_n - p| < \epsilon. \tag{1.1.28}$$

This negation, in turn, starts with "$(\forall k)$," etc. Step by step, we finally arrive at

$$(\exists \epsilon > 0)(\forall k)(\exists n > k) \ \ |x_n - p| \geq \epsilon. \tag{1.1.29}$$

Note that here *the choice of* $n > k$ *may depend on k.* To stress it, we often write n_k for n. Thus the negation of (2) finally emerges as

$$(\exists \epsilon > 0)(\forall k)(\exists n_k > k) \ \ |x_{n_k} - p| \geq \epsilon. \tag{1.1.30}$$

The *order* in which the quantifiers follow each other is *essential.* For example, the formula

$$(\forall n \in N)(\exists m \in N) \ \ m > n \tag{1.1.31}$$

("each $n \in N$ is exceeded by some $m \in N$") is true, but

$$(\exists m \in N)(\forall n \in N) \ \ m > n \tag{1.1.32}$$

is false. However, two *consecutive* universal quantifiers (or two *consecutive* existential ones) may be interchanged. We briefly write

$$\text{"} \ (\forall x, y \in A) \text{"} \ \text{ for } \ \text{"} \ (\forall x \in A)(\forall y \in A), \text{"} \tag{1.1.33}$$

and

$$\text{"} \ (\exists x, y \in A) \text{"} \ \text{ for } \ \text{"} \ (\exists x \in A)(\exists y \in A), \text{"} \ \text{ etc.} \tag{1.1.34}$$

does not imply the existence of an x for which $P(x)$ is true. It is only meant to imply that *there is no* x *in* A *for which* $P(x)$ *fails.*

The latter is true even if $A = \emptyset$; we then say that "$(\forall x \in A) \ \ P(x)$" is *vacuously true.* For examplek the formula $\emptyset \subseteq B$, i.e.,

$$(\forall x \in \emptyset) \ \ x \in B \tag{1.1.35}$$

is always true (vacuously).

1.1.E: Problems in Set Theory (Exercises)

Exercise 1.1.E. 1

Prove Theorem 1 (show that x is in the left-hand set iff it is in the right-hand set). For example, for (d),

$$\begin{aligned} x \in (A \cup B) \cap C & \Longleftrightarrow [x \in (A \cup B) \text{ and } x \in C] \\ & \Longleftrightarrow [(x \in A \text{ or } x \in B), \text{ and } x \in C] \\ & \Longleftrightarrow [(x \in A, x \in C) \text{ or } (x \in B, x \in C)]. \end{aligned}$$

Exercise 1.1.E. 2

Prove that
(i) $-(-A)=A$;
(ii) $A \subseteq B$ iff $-B \subseteq -A$.

Exercise 1.1.E. 3

Prove that

$$A-B=A \cap(-B)=(-B)-(-A)=-[(-A) \cup B]. \tag{1.1.E.1}$$

Also, give three expressions for $A \cap B$ and $A \cup B$, in terms of complements.

Exercise 1.1.E. 4

Prove the second duality law (Theorem 2(ii)).

Exercise 1.1.E. 5

Describe geometrically the following sets on the real line:

$$\begin{array}{ll} \text{(i) } \{x \mid x<0\}; & \text{(ii) } \{x \mid |x|<1\}; \\ \text{(iii) } \{x \mid |x-a|<\varepsilon\}; & \text{(iv) } \{x \mid a<x \leq b\}; \\ \text{(v) } \{x \mid |x|<0\}. & \end{array} \tag{1.1.E.2}$$

Exercise 1.1.E. 6

Let (a, b) denote the set

$$\{\{a\},\{a, b\}\} \tag{1.1.E.3}$$

(Kuratowski's definition of an ordered pair).
(i) Which of the following statements are true?

$$\begin{array}{ll} \text{(a) } a \in(a, b); & \text{(b) } \{a\} \in(a, b); \\ \text{(c) } (a, a)=\{a\}; & \text{(d) } b \in(a, b); \\ \text{(e) } \{b\} \in(a, b); & \text{(f) } \{a, b\} \in(a, b). \end{array} \tag{1.1.E.4}$$

(ii) Prove that $(a, b)=(u, v)$ if $a=u$ and $b=v$.
[Hint: Consider separately the two cases $a=b$ and $a \neq b$, noting that $\{a, a\}=\{a\}$. Also note that $\{a\} \neq a$.]

? Exercise $1.1.E.7$

Describe geometrically the following sets in the xy-plane.
(i) $\{(x,y)|x<y\}$;
(ii) $\left\{(x,y)|x^2+y^2<1\right\}$;
(iii) $\{(x,y)|\max(|x|,|y|)<1\}$;
(iii) $\left\{(x,y)|y>x^2\right\}$;
(iv) $\left\{(x,y)|y>x^2\right\}$;
(vii) $\{(x,y)||x|+|y|<4\}$;
(vii) $\left\{(x,y)|(x-2)^2+(y+5)^2\leq 9\right\}$;
(viii) $\left\{(x,y)|x^2-2xy+y^2<0\right\}$;
(ix) $\left\{(x,y)|x^2-2xy+y^2=0\right\}$.

? Exercise $1.1.E.8$

Prove that
(i) $(A\cup B)\times C=(A\times C)\cup(B\times C)$;
(ii) $(A\cap B)\times(C\cap D)=(A\times C)\cap(B\times D)$;
(iii) $(X\times Y)-(X'\times Y')=[(X\cap X')\times(Y-Y')]\cup[(X-X')\times Y]$;
[Hint: In each case, show that an ordered pair (x,y) is in the left-hand set iff it is in the right-hand set, treating (x,y) as one element of the Cartesian product.]

? Exercise $1.1.E.9$

Prove the distributive laws
(i) $A\cap\cup X_i=\bigcup(A\cap X_i)$;
(ii) $A\cup\cap X_i=\bigcap(A\cup X_i)$;
(iii) $(\cap X_i)-A=\cap(X_i-A)$;
(iv) $(\cup X_i)-A=\cup(X_i-A)$;
(v) $\cap X_i\cup\cap Y_j=\cap_{i,j}(X_i\cup Y_j)$;
(vi) $\cup X_i\cap\cup Y_j=\cup_{i,j}(X_i\cap Y_j)$.

? Exercise $1.1.E.10$

Prove that
(i) $(\cup A_i)\times B=\bigcup(A_i\times B)$;
(ii) $(\cap A_i)\times B=\cap(A_i\times B)$;
(iii) $(\cap_i A_i)\times(\cap_j B_j)=\bigcap_{i,j}(A_i\times B_i)$;
(iv) $(\cup_i A_i)\times\left(\bigcup_j B_j\right)=\bigcup_{i,j}(A_i\times B_j)$.

1.2: Relations. Mappings

Relations

In §1, we have already considered *sets of ordered pairs*, such as Cartesian products $A \times B$ or sets of the form $\{(x, y) | P(x, y)\}$(cf. §§1–3, Problem 7). If the pair (x, y) is an element of such a set R, we write

$$(x, y) \in R \tag{1.2.1}$$

treating (x, y) as *one* thing. Note that this *does not* imply that x and y taken *separately* are members of R (in which case we would write $x, y \in R$). We call x, y the *terms* of (x, y).

In mathematics, it is customary to call any set of ordered pairs a relation. For example, all sets listed in Problem 7 of §§1–3 are relations. Since relations are sets, equality $R = S$ for relations means that they consist of the same elements (ordered pairs), i.e., that

$$(x, y) \in R \Longleftrightarrow (x, y) \in S \tag{1.2.2}$$

If $(x, y) \in R$, we call y an R-relative of x; we also say that y is R-related to x or that the relation R holds between x and y (in this order). Instead of $(x, y) \in R$, we also write xRy, and often replace "R" by special symbols like $<$, $\sim$, etc. Thus, in case (i) of Problem 7 above, "xRy" means that $x < y$.

Replacing all pairs $(x, y) \in R$ by the inverse pairs (y, x), we obtain a new relation, called the inverse of R and denoted R^{-1}. Clearly, $xR^{-1}y$ iff yRx; thus

$$R^{-1} = \{(x, y) | yRx\} = \{(y, x) | xRy\}. \tag{1.2.3}$$

Hence R, in turn, is the inverse of R^{-1}; i.e.,

$$(R^{-1})^{-1} = R. \tag{1.2.4}$$

For example, the relations $<$ and $>$ between numbers are inverse to each other; so also are the relations $\subseteq$ and $\supseteq$ between sets. (We may treat "$\subseteq$" as the name of *the set of all pairs* (X, Y) *such that* $X \subseteq Y$ in a given space.)

If R contains the pairs $(x, x'), (y, y'), (z, z'), \ldots$, we shall write

$$R = \begin{pmatrix} x & y & z & \ldots \\ x' & y' & z' & \end{pmatrix}; \quad e.g., \quad R = \begin{pmatrix} 1 & 4 & 1 & 3 \\ 2 & 2 & 1 & 1 \end{pmatrix}. \tag{1.2.5}$$

To obtain R^{-1}, we simply interchange the upper and lower rows in Equation 1.2.1.

Definition 1

The set of all *left* terms x of pairs $(x, y) \in R$ is called the *domain* of R, denoted D_R. The set of all *right* terms of these pairs is called the *range* of R, denoted D'_R. Clearly, $x \in D_R$ iff xRy for some y. In symbols,

$$x \in D_R \Longleftrightarrow (\exists y) \ xRy; \ \text{similarly,} \ y \in D'_R \Longleftrightarrow (\exists x) \ xRy. \tag{1.2.6}$$

In Equation 1.2.1, D_R is the upper row, and D'_R is the lower row. Clearly,

$$D_{R^{-1}} = D'_R \text{ and } D'_{R^{-1}} = D_R. \tag{1.2.7}$$

For example, if

$$R = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 2 & 1 \end{pmatrix}, \tag{1.2.8}$$

then

$$D_R = D'_{R^{-1}} = \{1, 4\} \text{ and } D'_R = D_{R^{-1}} = \{1, 2\}. \tag{1.2.9}$$

Definition 2

The *image* of a set A under a relation R (briefly, the $R-image$ of A) is the set of all R-relatives of elements of A, denoted $R[A]$. The *inverse image* of A under R is the image of A under the *inverse* relation, i.e., $R^{-1}[A]$. If A consists of a single element, $A=x$, then $R[A]$ and $R^{-1}[A]$ are also written $R[x]$ and $R^{-1}[x]$, respectively, instead of $R[x]$ and $R^{-1}[x]$.

Example 1.2.1

Let

$$R=\begin{pmatrix}1&1&1&2&2&3&3&3&3&7\\1&3&4&5&3&4&1&3&5&1\end{pmatrix},\quad A=\{1,2\},\quad B=\{2,4\}. \tag{1.2.10}$$

Then

$$\begin{array}{lll}R[1]=\{1,3,4\}; & R[2]=\{3,5\}; & R[3]=\{1,3,4,5\};\\ R[5]=\emptyset; & R^{-1}[1]=\{1,3,7\}; & R^{-1}[2]=\emptyset;\\ R^{-1}[3]=\{1,2,3\}; & R^{-1}[4]=\{1,3\}; & R[A]=\{1,3,4,5\};\\ R^{-1}[A]=\{1,3,7\}; & R[B]=\{3,5\}. & \end{array} \tag{1.2.11}$$

By definition, $R[x]$ is the set of all R-relatives of x. Thus

$$y\in R[x] \text{ iff } (x,y)\in R; \text{ i.e., } xRy. \tag{1.2.12}$$

More generally, $y\in R[A]$ means that $(x,y)\in R$ for some $x\in A$. In symbols,

$$y\in R[A] \iff (\exists x\in A)(x,y)\in R. \tag{1.2.13}$$

Note that $R[A]$ is always defined.

Mappings

We shall now consider an especially important kind of relation.

Definition 3

A relation R is called a *mapping (map)*, or a *function*, or a *transformation*, iff every element $x\in D_R$ has a *unique* R-relative, so that $R[x]$ consists of a *single* element. This unique element is denoted by $R(x)$ and is called the *function value* at x (under R). Thus $R(x)$ is the only member of $R[x]$.

If, in addition, different elements of D_R have *different* images, R is called a *one-to-one* (or *one-one*) map. In this case,

$$x\neq y\,(x,y\in D_R) \text{ implies } R(x)\neq R(y); \tag{1.2.14}$$

equivalently,

$$R(x)=R(y) \text{ implies } x=y. \tag{1.2.15}$$

In other words, no two pairs belonging to R have the same left, or the same right, terms. This shows that R is *one to one iff* R^{-1}, *too, is a map*. Mappings are often denoted by the letters f,g,h,F,ψ, etc.

A mapping f is said to be "*from* A *to* B" iff $D_f=A$ and $D'_f\subseteq B$; we then write

$$f:A\rightarrow B \quad (\text{" } f \text{ maps } A \text{ into } B\text{"}) \tag{1.2.16}$$

If, in particular, $D_f=A$ and $D'_f=B$, we call f a map of A *onto* B, and we write

$$f:A\underset{\text{onto}}{\longrightarrow}B \quad (\text{" } f \text{ maps } A \text{ onto } B\text{"}) \tag{1.2.17}$$

If f is both onto and one to one, we write

$$f:A\longleftrightarrow B \tag{1.2.18}$$

($f : A \longleftrightarrow B$ means that f is one to one).

All pairs belonging to a mapping f have the form $(x, f(x))$ where $f(x)$ is the function value at x, i.e., the unique f-relative of $x, x \in D_f$. Therefore, in order to define some function f, *it suffices to specify its domain D_f and the function value $f(x)$ for each* $x \in D_f$. We shall often use such definitions. It is customary to say that f *is defined on* A (or "f is a function *on* A") iff $A = D_f$.

Example 1.2.2

(a) The relation

$$R = \{(x, y) | x \text{ is the wife of } y\} \tag{1.2.19}$$

is a one-to-one map of the set of all wives onto the set of all husbands. R^{-1} is here a one-to-one map of the set of all husbands $\left(= D_R'\right)$ onto
the set of all wives $(= D_R)$.

(b) The relation

$$f = \{(x, y) | y \text{ is the father of } x\} \tag{1.2.20}$$

is a map of the set of all people onto the set of their fathers. It is not one to one since several persons may have the same father (f-relative), and so $x \neq x'$ does not imply $f(x) \neq f\left(x'\right)$.

(c) Let

$$g = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 8 \end{pmatrix} \tag{1.2.21}$$

Then g is a map of $D_g = \{1, 2, 3, 4\}$ onto $D_g' = \{2, 3, 8\}$, with

$$g(1) = 2, g(2) = 2, g(3) = 3, g(4) = 8 \tag{1.2.22}$$

(As noted above, these formulas may serve to *define g.*) It is not one to one since $g(1) = g(2)$, so g^{-1} is not a map.

(d) Consider

$$f : N \rightarrow N, \text{ with } f(x) = 2x \text{ for each } x \in N \tag{1.2.23}$$

By what was said above, f is well defined. It is one to one since $x \neq y$ implies $2x \neq 2y$. Here $D_f = N($ the naturals $)$, but D_f' consists of *even* naturals only. Thus f is not *onto* N (it is onto a *smaller* set, the *even* naturals); f^{-1} maps the even naturals onto all of N.

The domain and range of a relation may be quite arbitrary sets. In particular, we can consider functions f in which each element of the domain D_f is itself an ordered pair (x, y) or n-tuple $(x_1, x_2, \ldots, x_n)$. Such mappings are called *functions of two* (*respectively*, n) *variables*. To any n -tuple $(x_1, \ldots, x_n)$ that belongs to D_f, the function f assigns a unique function value, denoted by $f\left(x_1, \ldots, x_n\right)$. It is convenient to regard $x_1, x_2, \ldots, x_n$ as certain variables; then the function value, too, becomes a variable depending on the $x_1, \ldots, x_n$. Often D_f consists of all ordered n-tuples of elements taken from a set A, i.e., $D_f = A^n$ (cross-product of n sets, each equal to A). The *range* may be an arbitrary set B; so $f : A^n \rightarrow B$. Similarly, $f : A \times B \rightarrow C$ is a function of *two* variables, with $D_f = A \times B, D_f' \subseteq C$.

Functions of two variables are also called *(binary) operations.* For example, addition of natural numbers may be treated as a map $f : N \times N \rightarrow N$, with $f(x, y) = x + y$.

Definition 4

A relation R is said to be
(i) *reflexive* iff we have xRx for each $x \in D_R$;
(ii) *symmetric* iff xRy always implies yRx;
(iii) *transitive* iff xRy combined with yRz always implies xRz.

R is called an *equivalence relation* on a set A iff $A = D_R$ and R has all the three properties (i), (ii), and (iii). For example, such is the *equality relation* on A (also called the *identity map* on A) denoted

$$I_A = \{(x, y) | x \in A, x = y\} \tag{1.2.24}$$

Equivalence relations are often denoted by special symbols resembling equality, such as $\equiv, \approx, \sim$, etc. The formula xRy, where R is such a symbol, is read

$$\text{" } x \text{ is equivalent (or} R\text{-equivalent) to } y, \text{"} \tag{1.2.25}$$

and $R[x] = \{y | xRy\}$ (i.e., the R-image of x) is called the *R-equivalence class* (briefly *R-class*) of x in A; it consists of all elements that are R-equivalent to x and *hence to each other* (for xRy and xRz imply first yRx, by symmetry, and hence yRz, by transitivity). Each such element is called a *representative* of the given R-class, or its *generator*. We often write $[x]$ for $R[x]$.

Example 1.2.3

(a') The inequality relation $<$ between real numbers is transitive since

$$x < y \text{ and } y < z \text{ implies } x < z \tag{1.2.26}$$

it is neither reflexive nor symmetric. (Why?)

(b') The inclusion relation $\subseteq$ between sets is reflexive (for $A \subseteq A$) and transitive (for $A \subseteq B$ and $B \subseteq C$ implies $A \subseteq C$), but it is not symmetric.

(c') The membership relation $\in$ between an element and a set is neither reflexive nor symmetric nor transitive ($x \in A$ and $A \in \mathcal{M}$ *does not* imply $x \in \mathcal{M}$).

(d') Let R be the *parallelism* relation between lines in a plane, i.e., the set of all pairs (X, Y), where X and Y are parallel lines. Writing $\|$ for R, we have $X \| X, X \| Y$ implies $Y \| X$, and $(X \| Y$ and $Y \setminus\| Z)$ implies $X \| Z$, so R is an equivalence relation. An R -class here consists of all lines parallel to a given line in the plane.

(e') Congruence of triangles is an equivalence relation. (Why?)

Theorem 1.2.1

If R (*also written* $\equiv$) is an *equivalence relation on* A, *then all* R-*classes are disjoint from each other, and* A *is their union*.

Proof

Take two R-classes, $[p] \neq [q]$. Seeking a contradiction, suppose they are *not* disjoint, so

$$(\exists x) \quad x \in [p] \text{ and } x \in [q]; \tag{1.2.27}$$

i.e., $p \equiv x \equiv q$ and hence $p \equiv q$. But then, by symmetry and transitivity,

$$y \in [p] \Leftrightarrow y \equiv p \Leftrightarrow y \equiv q \Leftrightarrow y \in [q]; \tag{1.2.28}$$

i.e., $[p]$ and $[q]$ *consist of the same elements* y, contrary to assumption $[q] \neq [q]$. Thus, indeed, any two (distinct) R-classes are disjoint.

Also, by reflexivity,

$$(\forall x \in A) \quad x \equiv x \tag{1.2.29}$$

i.e., $x \in [x]$. Thus each $x \in A$ is in some R-class (namely, in $[x]$); of A is in the *union* of such classes,

$$A \subseteq \bigcup_x R[x] \tag{1.2.30}$$

Conversely,

$$(\forall x) \quad R[x] \subseteq A \tag{1.2.31}$$

since

$$y \in R[x] \Rightarrow xRy \Rightarrow yRx \Rightarrow (y, x) \in R \Rightarrow y \in D_R = A \tag{1.2.32}$$

by definition. Thus A contains all $R[x]$, hence their union, and so

$$A = \bigcup_x R[x]. \square \tag{1.2.33}$$

1.2.E: Problems on Relations and Mappings (Exercises)

Exercise $1.2.E.1$

For the relations specified in Problem 7 of §§1-3, find D_R, D'_R, and R^{-1}. Also, find $R[A]$ and $R^{-1}[A]$ if

$$\begin{array}{ll}\text{(a) } A=\left\{\frac{1}{2}\right\}; & \text{(b) } A=\{1\} \\ \text{(c) } A=\{0\}; & \text{(d) } A=\emptyset; \\ \text{(e) } A=\{0,3,-15\}; & \text{(f) } A=\{3,4,7,0,-1,6\} \\ \text{(g) } A=\{x \mid -20<x<5\} & \end{array} \tag{1.2.E.1}$$

Exercise $1.2.E.2$

Prove that if $A \subseteq B$, then $R[A] \subseteq R[B]$. Disprove the converse by a counterexample.

Exercise $1.2.E.3$

Prove that
(i) $R[A \cup B]=R[A] \cup R[B]$;
(ii) $R[A \cap B] \subseteq R[A] \cap R[B]$;
(iii) $R[A-B] \supseteq R[A]-R[B]$.
Disprove reverse inclusions in (ii) and (iii) by examples. Do (i) and (ii) with A, B replaced by an arbitrary set family $\{A_i \mid i \in I\}$.

Exercise $1.2.E.4$

Under which conditions are the following statements true?

$$\begin{array}{ll}\text{(i) } R[x]=\emptyset; & \text{(ii) } R^{-1}[x]=\emptyset; \\ \text{(iii) } R[A]=\emptyset; & \text{(iv) } R^{-1}[A]=\emptyset;\end{array} \tag{1.2.E.2}$$

Exercise $1.2.E.5$

Let $f: N \rightarrow N(N=\{\text{ naturals }\})$. For each of the following functions, specify $f[N]$, i.e., D'_f, and determine whether f is one to one and onto N, given that for all $x \in N$,

$$\begin{array}{l}\text{(i) } f(x)=x^3; \quad \text{(ii) } f(x)=1; \quad \text{(iii) } f(x)=|x|+3 \\ \text{(iv) } f(x)=x^2; \quad \text{(v) } f(x)=4x+5\end{array} \tag{1.2.E.3}$$

Do all this also if N denotes
(a) the set of all integers;
(b) the set of all reals.

Exercise $1.2.E.6$

Prove that for any mapping f and any sets $A, B, A_i (i \in I)$,
(a) $f^{-1}[A \cup B]=f^{-1}[A] \cup f^{-1}[B]$;
(b) $f^{-1}[A \cap B]=f^{-1}[A] \cap f^{-1}[B]$;
(c) $f^{-1}[A-B]=f^{-1}[A]-f^{-1}[B]$;
(d) $f^{-1}\left[\bigcup_i A_i\right]=\bigcup_i f^{-1}\left[A_i\right]$;
(e) $f^{-1}\left[\bigcap_i A_i\right]=\bigcap_i f^{-1}\left[A_i\right]$.

Compare with Problem 3.
[Hint: First verify that $x \in f^{-1}[A]$ iff $x \in D_f$ and $f(x) \in A$.]

Exercise $1.2.E.7$

Let f be a map. Prove that
(a) $f\left[f^{-1}[A]\right] \subseteq A$;
(b) $f\left[f^{-1}[A]\right] = A$ if $A \subseteq D'_f$;
(c) if $A \subseteq D_f$ and f is one to one, $A = f^{-1}[f[A]]$/
Is $f[A] \cap B \subseteq f\left[A \cap f^{-1}[B]\right]$?

Exercise $1.2.E.8$

Is R an equivalence relation on the set J of all integers, and, if so, what are the R -classes, if
(a) $R = \{(x, y) | x - y \text{ is divisible by a fixed } n\}$;
(b) $R = \{(x, y) | x - y \text{ is odd }\}$;
(c) $R = \{(x, y) | x - y \text{ is a prime }\}$.
$(x, y, n \text{ denote integers.})$

Exercise $1.2.E.9$

Is any relation in Problem 7 of §§1-3 reflexive? Symmetric? Transitive?

10. Show by examples that R may be
(a) reflexive and symmetric, without being transitive;
(b) reflexive and transitive without being symmetric.
Does symmetry plus transitivity imply reflexivity? Give a proof or counterexample.

1.3: Sequences

By an *infinite sequence* (briefly *sequence*) we mean a mapping (call it u) whose domain is N (all natural numbers $1, 2, 3, \ldots$); D_u may also contain 0.

A *finite sequence* is a map u in which D_u consists of all positive (or non-negative) integers *less than a fixed integer* p. The *range* D_u' of any sequence u may be an arbitrary set B; we then call u a sequence of elements of B, or in B. For example,

$$u=\left(\begin{array}{lllllll}1 & 2 & 3 & 4 & \ldots & n & \ldots \\ 2 & 4 & 6 & 8 & \ldots & 2 n & \ldots\end{array}\right) \tag{1.3.1}$$

is a sequence with

$$D_u=N=\{1,2,3,\ldots\} \tag{1.3.2}$$

and with function values

$$u(1)=2, u(2)=4, u(n)=2 n, \quad n=1,2,3, \ldots \tag{1.3.3}$$

Instead of $u(n)$ we usually write u_n ("index notation"), and call u_n the n^{th} *term* of the sequence. If n is treated as a *variable*, u_n is called the *general term* of the sequence, and $\{u_n\}$ is used to denote the entire (infinite) sequence, as well as its range D_u' (whichever is meant, will be clear from the context). The formula $\{u_n\} \subseteq B$ means that $D_u' \subseteq B$, i.e., that u is a sequence in B. To
determine a sequence, it suffices to define its general term u_n by some formula or rule. **In** (1) **above**, $u_n=2n$.

Often we omit the mention of $D_u=N$ (since it is *known*) and give only the range D_u'. Thus instead of (1), we briefly write

$$2,4,6, \ldots, 2 n, \ldots \tag{1.3.4}$$

or, more generally,

$$u_1, u_2, \ldots, u_n, \ldots \tag{1.3.5}$$

Yet it should be remembered that u is a set of *pairs* (a map).

If all u_n are *distinct* (different from each other), u is a *one-to-one* map. However, this need not be the case. It may even occur that all u_n are equal (then u is said to be constant); e.g., $u_n=1$ yields the sequence $1, 1, 1, \ldots, 1, \ldots$,i.e.

$$u=\left(\begin{array}{cccccc}1 & 2 & 3 & \ldots & n & \ldots \\ 1 & 1 & 1 & \ldots & 1 & \ldots\end{array}\right) \tag{1.3.6}$$

Note that here$u$ is an *infinite* sequence (since $D_u=N$), even though its range D_u' has only one element, $D_u'=\{1\}$. (In *sets*, repeated terms count as *one* element; but the *sequence* u consists of infinitely many distinct *pairs* $(n, 1)$. *) If all u_n are real numbers, we call u a real sequence. For such sequences, we have the following definitions.*

Definition 1

A real sequence $\{u_n\}$ is said to be *monotone* (or *monotonic*) iff it is either *nondecreasing*, i.e.

$$(\forall n) \quad u_n \leq u_{n+1} \tag{1.3.7}$$

or *nonincreasing*, i.e.,

$$(\forall n) \quad u_n \geq u_{n+1} \tag{1.3.8}$$

Notation: $\{u_n\} \uparrow$ and $\{u_n\} \downarrow$, respectively. If instead we have the *strict* inequalities $u_n<u_{n+1}$ (respectively, $u_n>u_{n+1}$), we call $\{u_n\}$ *strictly monotone* (increasing or decreasing).

A similar definition applies to sequences of *sets.*

Definition 2

A sequence of sets $A_1, A_2, \ldots, A_n, \ldots$ is said to be *monotone* iff it is either *expanding*, i.e.,

$$(\forall n) \quad A_n \subseteq A_{n+1} \tag{1.3.9}$$

or *contracting*, i.e.,

$$(\forall n) \quad A_n \supseteq A_{n+1} \tag{1.3.10}$$

Notation: $\{A_n\} \uparrow$ and $\{A_n\} \downarrow$, respectively. For example, any sequence of concentric solid spheres (treated as *sets of points*), with increasing radii, is expanding; if the radii decrease, we obtain a contracting sequence.

Definition 3

Let $\{u_n\}$ be any sequence, and let

$$n_1 < n_2 < \cdots < n_k < \cdots \tag{1.3.11}$$

be a *strictly increasing* sequence of natural numbers. Select from $\{u_n\}$ those terms whose subscripts are $n_1, n_2, \ldots, n_k, \ldots$ Then the sequence $\{u_{n_k}\}$ so selected (with k th term equal to u_{n_k}), is called the *subsequence* of $\{u_n\}$, determined by the subscripts $n_k, k = 1, 2, 3, \ldots$.

Thus (roughly) a subsequence is any sequence obtained from $\{u_n\}$ by dropping some terms, *without changing the order of the remaining terms* (this is ensured by the inequalities $n_1 < n_2 < \cdots < n_k < \cdots$ where the n_k are the subscripts of the remaining terms). For example, let us select from (1) the subsequence of terms whose subscripts are *primes* (including 1). Then the subsequence is

$$2, 4, 6, 10, 14, 22, \ldots \tag{1.3.12}$$

i.e.,

$$u_1, u_2, u_3, u_5, u_7, u_{11}, \ldots \tag{1.3.13}$$

All these definitions apply to finite sequences accordingly. Observe that every sequence arises by "numbering" the elements of its range (the terms): u_1 is the *first* term, u_2 is the *second* term, and so on. By so numbering, we put the terms in a certain *order*, determined by their subscripts $1, 2, 3, \ldots$(like the numbering of buildings in a street, of books in a library, etc.). The question now arises: Given a set A, is it always possible to "number" its elements *by integers*? As we shall see in \$4, this is not always the case. This leads us to the following definition.

Definition 4

A set A is said to be *countable* iff A is contained in the range of some sequence (briefly, the *elements* of A *can be put in a sequence*).

If, in particular, this sequence can be chosen finite, we call A a *finite* set. (The empty set is finite.)

Sets that are not finite are said to be *infinite*.

Sets that are not countable are said to be *uncountable*.

Note that all finite sets are countable. The simplest example of an infinite countable set is $N = \{1, 2, 3, \ldots\}$.

1.4: Some Theorems on Countable Sets

We now derive some corollaries of Definition 4 in § 3.

COROLLARY 1.4.1

If a set A is countable or finite, so is any subset $B \subseteq A$.

For if $A \subset D_u'$ for a sequence u, then certainly $B \subseteq A \subseteq D_u'$

COROLLARY 1.4.2

If A is uncountable (or just infinite), so is any superset $B \supseteq A$.

For, if B were countable or finite, so would be $A \subseteq B$, by Corollary 1

Theorem 1.4.1

If A and B are countable, so is their cross product $A \times B$

Proof

If A or B is $\emptyset$, then $A \times B = \emptyset$, and there is nothing to prove.

Thus let A and B be nonvoid and countable. We may assume that they *fill* two infinite sequences, $A = \{a_n\}, B = \{b_n\}$ (*repeat* terms if necessary). Then, by definition, $A \times B$ is the set of all ordered pairs of the form

$$(a_n, b_m), \quad n, m \in N \tag{1.4.1}$$

Call $n+m$ the *rank* of the pair (a_n, b_m). For each $r \in N$, there are $r-1$ pairs of rank r:

$$(a_1, b_{r-1}), (a_2, b_{r-2}), \ldots, (a_{r-1}, b_1) \tag{1.4.2}$$

We now put all pairs (a_n, b_m) in *one* sequence as follows. We start with

$$(a_1, b_1) \tag{1.4.3}$$

as the first term; then take the two pairs of rank three,

$$(a_1, b_2), (a_2, b_1) \tag{1.4.4}$$

then the three pairs of rank four, and so on. At the $(r-1)$ st step, we take all pairs of rank r, in the order indicated in (1).

Repeating this process for all ranks ad infinitum, we obtain the sequence of pairs

$$(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_1, b_3), (a_2, b_2), (a_3, b_1), \ldots \tag{1.4.5}$$

in which $u_1 = (a_1, b_1), u_2 = (a_1, b_2)$, etc.

By construction, this sequence contains all pairs of all ranks r, hence all pairs that form the set $A \times B$ (for every such pair has some rank r and so it must eventually occur in the sequence). Thus $A \times B$ can be put in a sequence. □

COROLLARY 1.4.3

The set R of all rational numbers is countable.

Proof

Consider first the set Q of all positive rationals, i.e.,

$$\text{fractions } \frac{n}{m}, \text{ with } n, m \in N \tag{1.4.6}$$

We may formally identify them with ordered pairs (n, m), i.e., with $N \times N$ We call $n+m$ the *rank* of (n, m). As in Theorem 1, we obtain the sequence

$$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{2}{2}, \frac{3}{1}, \frac{1}{4}, \frac{2}{3}, \frac{3}{2}, \frac{4}{1}, \ldots \tag{1.4.7}$$

By dropping reducible fractions and inserting also 0 and the negative rationals, we put R into the sequence

$$0, 1, -1, \frac{1}{2}, -\frac{1}{2}, 2, -2, \frac{1}{3}, -\frac{1}{3}, 3, -3, \ldots, \text{ as required. } \square \tag{1.4.8}$$

Theorem 1.4.2

The union of any sequence $\{A_n\}$ *of countable sets is countable.*

Proof

As each A_n is countable, we may put

$$A_n = \{a_{n1}, a_{n2}, \ldots, a_{nm}, \ldots\} \tag{1.4.9}$$

(The double subscripts are to distinguish the sequences representing different sets A_n.) As before, we may assume that all sequences are infinite. Now, $U_n A_n$ obviously consists of the elements of *all* A_n *combined*, i.e., *all* $a_{nm}(n, m \in N)$. We call $n+m$ the $rank$ of a_{nm} and proceed as in Theorem 1, thus obtaining

$$\bigcup_n A_n = \{a_{11}, a_{12}, a_{21}, a_{13}, a_{22}, a_{31}, \ldots\} \tag{1.4.10}$$

Thus $\cup_n A_n$ can be put in a sequence. $\square$

Note 1: Theorem 2 is briefly expressed as

"*Any countable union of countable sets is a countable set.*"

(The term "*countable union*" means "union of a *countable* family of sets", i.e., a family of sets whose elements can be put in a sequence $\{A_n\}$.) In particular, if A and B are countable, so are $A \cup B, A \cap B,$ and $A - B$ (by Corollary 1).

Note 2: From the proof it also follows that *the range of any double sequence*$\{a_{nm}\}$ *is countable*. (A *double sequence* is a function u whose domain D_u is $N \times N;$ say, $u: N \times N \rightarrow B.$ If $n, m \in N$, we write u_{nm} for $u(n, m)$ here $u_{nm} = a_{nm}.$)

To prove the existence of *uncountable* sets, we shall now show that the interval

$$[0, 1) = \{x | 0 \leq x < 1\} \tag{1.4.11}$$

of the real axis is uncountable.

We assume as known the fact that each real number $x \in [0, 1)$ has a unique infinite decimal expansion

$$0.\, x_1, x_2, \ldots, x_n, \ldots \tag{1.4.12}$$

where the x_n are the decimal digits (possibly zeros), and the sequence $\{x_n\}$ does not terminate in *nines* (this ensures *uniqueness*)."

Theorem 1.4.3

The interval$[0, 1)$ *of the real axis is uncountable.*

Proof

We must show that no sequence can comprise *all* of $[0, 1)$. Indeed, given any $\{u_n\}$, write each term u_n as an infinite fraction; say,

$$u_n = 0.\, a_{n1}, a_{n2}, \ldots, a_{nm}, \ldots \tag{1.4.13}$$

Next, construct a *new* decimal fraction

$$z=0.x_1,x_2,\ldots,x_n,\ldots \tag{1.4.14}$$

choosing its digits x_n as follows.

If a_{nn} (i.e., the n th digit of u_n) is 0, put $x_n=1$; if, however, $a_{nn}\neq 0$, put $x_n=0$. Thus, in all cases, $x_n\neq a_{nn}$, i.e., z *differs from each* u_n *in at least one decimal digit* (namely, the n th digit). It follows that z is different from all u_n and hence is not in $\{u_n\}$, even though $z\in[0,1)$.

Thus, no matter what the choice of $\{u_n\}$ was, we found some $z\in[0,1)$ not in the range of that sequence. Hence $no\{u_n\}$ contains all of $[0,1)$. □

Note 3: By Corollary 2, any superset of $[0,1)$, e.g., the entire real axis, is *uncountable*.

Note 4: Observe that the numbers a_{nn} used in the proof of Theorem 3 form the diagonal of the infinitely extending square composed of all a_{nm}. Therefore, the method used above is called the diagonal process (due to G. Cantor).

1.4.E: Problems on Countable and Uncountable Sets (Exercises)

Exercise $1.4.E.1$

Prove that if A is countable but B is not, then $B-A$ is uncountable.
[Hint: If $B-A$ were countable, so would be

$$(B-A) \cup A \supseteq B. \quad \text{(Why?)} \tag{1.4.E.1}$$

Use Corollary $1.$]

Exercise $1.4.E.2$

Let f be a mapping, and $A \subseteq D_f$. Prove that
(i) if A is countable, so is $f[A]$;
(ii) if f is one to one and A is uncountable, so is $f[A]$.
[Hints: (i) If $A=\{u_n\}$, then

$$f[A]=\{f(u_1), f(u_2), \ldots, f(u_n), \ldots\} \tag{1.4.E.2}$$

(ii) If $f[A]$ were countable, so would be $f^{-1}[f[A]]$, by (i). Verify that

$$f^{-1}[f[A]]=A \tag{1.4.E.3}$$

here; cf. Problem 7 in §§4-7.]

Exercise $1.4.E.3$

Let a, b be real numbers $(a<b)$. Define a map f on $[0,1)$ by

$$f(x)=a+x(b-a). \tag{1.4.E.4}$$

Show that f is one to one and onto the interval $[a, b)=\{x \mid a \leq x<b\}$. From Problem 2, deduce that $[a, b)$ is uncountable. Hence, by Problem 1, so $is (a, b)=\{x \mid a<x<b\}$.

Exercise $1.4.E.4$

Show that between any real numbers $a, b(a<b)$ there are uncountably many irrationals, i.e., numbers that are not rational.
[Hint: By Corollary 3 and Problems 1 and 3, the set $(a, b)-R$ is uncountable. Explain in detail.

Exercise $1.4.E.5$

Show that every infinite set A contains a countably infinite set, i.e., an infinite sequence of distinct terms.
[Hint: Fix any $a_1 \in A$; A cannot consist of a_1 alone, so there is another element

$$a_2 \in A-\{a_1\}. \quad \text{(Why?)} \tag{1.4.E.5}$$

Again, $A \neq\{a_1, a_2\}$, so there is an $a_3 \in A-\{a_1, a_2\}$. (Why?) Continue thusly ad infinitum to obtain the required sequence $\{a_n\}$. Why are all a_n distinct?]

Exercise $1.4.E.6$

From Problem 5, prove that if A is infinite, there is a map $f : A \rightarrow A$ that is one to one but not onto A.
[Hint: With a_n as in Problem 5, define $f\left(a_n\right) = a_{n+1}$. If, however, x is none of the a_n, put $f(x) = x$. Observe that $f(x) = a_1$ is never true, so f is not onto A. Show, however, that f is one to one.

Exercise $1.4.E.7$

Conversely (cf. Problem 6), prove that if there is a map $f : A \rightarrow A$ that is one to one but not onto A, then A contains an infinite sequence $\{a_n\}$ of distinct terms.
[Hint: As f is not onto A, there is $a_1 \in A$ such that $a_1 \notin f[A]$. (Why?) Fix a_1 and define

$$a_2 = f\left(a_1\right), a_3 = f\left(a_2\right), \ldots, a_{n+1} = f\left(a_n\right), \ldots \text{ ad infinitum.} \tag{1.4.E.6}$$

To prove distinctness, show that each a_n is distinct from all a_m with $m > n$. For a_1, this is true since $a_1 \notin f[A]$, whereas $a_m \in f[A] (m > 1)$. Then proceed inductively.]

CHAPTER OVERVIEW

2: Real Numbers and Fields

2.1: Axioms and Basic Definitions

Real numbers can be constructed step by step: first the integers, then the rationals, and finally the irrationals. Here, however, we shall assume the set of all real numbers, denoted E^1, as *already given*, without attempting to reduce this notion to simpler concepts. We shall also accept without definition (as *primitive* concepts) the notions of the *sum* $(a+b)$ and the *product*, $(a \cdot b)$ or (ab), of two real numbers, as well as the *inequality relation* $<$ (read "*less than*"). Note that $x \in E^1$ means "x *is in* E^{1}", i.e., "*x is a real number.*"

It is an important fact that all arithmetic properties of reals can be deduced from several simple axioms, listed (and named) below.

Axioms of Addition and Multiplication

Definition

1. (closure laws) The *sum* $x+y$, *and the product* xy, *any real numbers are real numbers themselves.* In symbols,

$$\left(\forall x, y \in E^{1}\right) \quad (x+y) \in E^{1} \text{ and } (xy) \in E^{1} \tag{2.1.1}$$

2. (commutative laws)

$$\left(\forall x, y \in E^{1}\right) \quad x+y=y+x \text{ and } xy=yx \tag{2.1.2}$$

3. (associative laws)

$$\left(\forall x, y, z \in E^{1}\right) \quad (x+y)+z=x+(y+z) \text{ and } (xy)z=x(yz) \tag{2.1.3}$$

4. (existence of neutral elements)

(a) *There is a* (*unique*) *real number, called zero* (0), *such that, for all real* x, $x+0=x$.

(b) *There is a* (*unique*) *real number, called one* (1), *such that* $1\ 1 \neq 0$ *and, for all real* x, $x \cdot 1=x$.

In symbols,

(a)

$$\left(\exists! 0 \in E^{1}\right)\left(\forall x \in E^{1}\right) \quad x+0=x; \tag{2.1.4}$$

(b)

$$\left(\exists 1 \in E^{1}\right)\left(\forall x \in E^{1}\right) \quad x \cdot 1=x, 1 \neq 0. \tag{2.1.5}$$

(The real numbers 0 and 1 are called the *neutral elements* of addition and multiplication, respectively.)

5. (existence of inverse elements)

(a) *For every real* x, *there is a* (*unique*) *real, denoted* $-x$, *such that* $x+(-x)=0$.

(b) *For every real* x *other than* 0, *there is a* (*unique*) *real, denoted* x^{-1}, *such that* $x \cdot x^{-1}=1$.

In symbols,

(a)

$$\left(\forall x \in E^{1}\right)\left(\exists!-x \in E^{1}\right) \quad x+(-x)=0; \tag{2.1.6}$$

(b)

$$\left(\forall x \in E^{1} | x \neq 0\right)\left(\exists! x^{-1} \in E^{1}\right) \quad x x^{-1}=1. \tag{2.1.7}$$

(The real numbers $-x$ and x^{-1} are called, respectively, the additive inverse (or the symmetric) and the multiplicative inverse (or the reciprocal) of x.)

6. (distributive law)

$$\left(\forall x, y, z \in E^{1}\right) \quad (x+y)z=xz+yz \tag{2.1.8}$$

Axioms of Order

Definition

7. (trichotomy) *For any real* x *and* y, *we have*

$$\text{either} x<y \text{ or } y<x \text{ or } x=y \tag{2.1.9}$$

but never two of these relations together.

8. (transitivity)

$$\left(\forall x, y, z \in E^{1}\right) \quad x<y \text{ and } y<z \text{ implies } x<z \tag{2.1.10}$$

9. (monotonicity of addition and multiplication) *For any* $x, y, z \in E^{1}$, *we have*

(a)

$$x<y \text{ implies } x+z<y+z; \tag{2.1.11}$$

(b) \[x<y \text{ and } z>0 \text{ implies] x z<y z.\]

Note 1: The uniqueness assertions in Axioms 4 and 5 are actually redundant since they can be deduced from other axioms. We shall not dwell on this.

Note 2: Zero has no reciprocal; i.e., for no x is $0x=1$. In fact, $0x=0$. For, by Axioms VI and IV,

$$0x+0x=(0+0)x=0x=0x+0. \tag{2.1.12}$$

Cancelling $0x$(i.e., adding $-0x$ on both sides), we obtain $0x=0$, by Axioms 3 and 5 (a).

Note 3: Due to Axioms 7 and 8, real numbers may be regarded as given in a certain *order* under which smaller numbers precede the larger ones. (This is why we speak of "axioms of *order*.") The ordering of real numbers can be visualized by "plotting" them as points on a directed line ("the real axis") in a well-known manner. Therefore, E^{1} is also often called "*the real axis*," and real numbers are called "*points*"; we say "the *point x* instead of "*the number x*.

Observe that the axioms only state certain properties of real numbers *without specifying what these numbers are*. Thus we may treat the reals as just *any* mathematical objects satisfying our axioms, but otherwise arbitrary. Indeed, our theory also applies to any other set of objects (numbers or not), provided they satisfy our axioms with respect to a certain relation of order $(<)$ and certain operations $(+)$ and $(\cdot)$, which may, but need not, be ordinary addition and multiplication. Such sets exist indeed. We now give them a name.

Definition 1

A *field* is any set F of objects, with two operations $(+)$ and $(.)$ defined in it in such a manner that they satisfy Axioms 1-6 listed above (with E^{1} replaced by F, of course).

If F is also endowed with a relation $<$ satisfying Axioms 7 to 9, we call F an ordered field.

In this connection, postulates 1 to 9 are called axioms of an (ordered) field.
By Definition $1, E^{1}$ is an ordered field. Clearly, whatever follows from the axioms must hold not only in E^{1} but also in any other ordered field. Thus
we shall henceforth state our definitions and theorems in a more general way, speaking of ordered fields in general instead of E^{1} alone.

Definition 2

An element x of an ordered field is said to be *positive* if $x>0$ or *negative* if $x<0$.

Here and below, $" x>y "$ means the same as $" y<x . "$ We also write $" x \leq y "$ for $" x<y$ or $x=y'' ;$ similarly for $" x \geq y . "$

Definition 3

For any elements x, y of a field, we define their *difference*

$$x-y=x+(-y) \tag{2.1.13}$$

If $y \neq 0$, we also define the *quotient* of x by y

$$\frac{x}{y}=x y^{-1} \tag{2.1.14}$$

also denoted by x/y.

Note 4: *Division by* 0 *remains undefined*.

Definition 4

For any element x of an ordered field, we define its *absolute value*,

$$|x|=\begin{cases}x & \text{if } x \geq 0 \text{ and} \\ -x & \text{if } x<0\end{cases} \tag{2.1.15}$$

It follows that $|x| \geq 0$ *always*; for if $x \geq 0$, then

$$|x|=x \geq 0 \tag{2.1.16}$$

and if $x<0$, then

$$|x|=-x>0. \quad (\text{ Why? }) \tag{2.1.17}$$

Moreover,

$$-|x| \leq x \leq|x|, \tag{2.1.18}$$

for,

$$\text{if } x \geq 0, \text{ then } |x|=x; \tag{2.1.19}$$

and

$$\text{if } x<0, \text{ then } x<|x| \text{ since } |x|>0. \tag{2.1.20}$$

Thus, in all cases,

$$x \leq|x|. \tag{2.1.21}$$

Similarly one shows that

$$-|x| \leq x. \tag{2.1.22}$$

As we have noted, all rules of arithmetic (dealing with the four arithmetic operations and inequalities) can be deduced from Axioms 1 through 9 and thus apply to *all* ordered fields, along with E^{1}. We shall not dwell on their deduction, limiting ourselves to a few simple corollaries as examples.

Corollary 2.1.1

(i) $a(-b)=(-a)b=-(ab)$;

(ii) $\quad(-a)(-b)=ab$.

Proof

By Axiom 6,

$$a(-b)+ab=a[(-b)+b]=a \cdot 0=0. \tag{2.1.23}$$

Thus

$$a(-b)+ab=0. \tag{2.1.24}$$

By definition, then, $a(-b)$ is the *additive inverse* of ab, i.e.,

$$a(-b)=-(ab). \tag{2.1.25}$$

Similarly, we show that

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

and that

$$-(-a)=a. \tag{2.1.26}$$

Finally, (ii) is obtained from (i) when a is replaced by $-a$. $\square$

Corollary 2.1.2

In an ordered field, $a \neq 0$ *implies*

$$a^2=(a \cdot a)>0 \tag{2.1.27}$$

(Hence $1=1^2>0.$ *)*

Proof

If $a>0$, we may multiply by $a($ Axiom 9(b) to obtain

$$a \cdot a>0 \cdot a=0, \text{ i.e., } a^2>0. \tag{2.1.28}$$

If $a<0$, then $-a>0$; so we may multiply the inequality $a<0$ by $-a$ and obtain

$$a(-a)<0(-a)=0; \tag{2.1.29}$$

i.e., by Corollary 1,

$$-a^2<0, \tag{2.1.30}$$

whence

$$a^2>0 \ \square \tag{2.1.31}$$

2.2: Natural Numbers. Induction

The element 1 was introduced in Axiom 4(b). Since addition is also assumed known, we can use it to define, step by step, the elements

$$2=1+1, 3=2+1, 4=3+1, \text{ etc.} \tag{2.2.1}$$

If this process is continued indefinitely, we obtain what is called the set N of all *natural elements* in the given field F. In particular, the natural elements of E^1 are called *natural numbers*. Note that

$$(\forall n \in N) \quad n+1 \in N \tag{2.2.2}$$

*A more precise approach to natural elements is as follows. A subset S of a field F is said to be inductive iff

$$(i)\ \ 1 \in S \text{ and} \tag{2.2.3}$$

$$(ii)\ \ (\forall x \in S)\ x+1 \in S \tag{2.2.4}$$

Such subsets certainly exist; e.g., the entire field F is inductive since

$$1 \in F \text{ and } (\forall x \in F)\ x+1 \in F. \tag{2.2.5}$$

Define N as the intersection of *all* inductive sets in F.

Theorem 2.2.1

The set N so defined is inductive itself. In fact, it is the "smallest" inductive subset of F (i . e ., contained in any other such subset).

Proof

We have to show that

$$(i)\ \ 1 \in N, \text{ and} \tag{2.2.6}$$

$$(ii)\ \ (\forall x \in N)\ x+1 \in N. \tag{2.2.7}$$

Now, by definition, the unity 1 is in *each* inductive set; hence it also belongs to the intersection of such sets, i.e., to N. Thus $1 \in N$, as claimed.

Next, take any $x \in N$. Then, by our definition of N, x is in *each* inductive set S; thus, by property (ii) of such sets, also $x+1$ is in each such S; hence $x+1$ is in the intersection of all inductive sets, i.e.,

$$x+1 \in N \tag{2.2.8}$$

and so N is inductive, indeed.

Finally, by definition, N is the *common part* of all such sets and hence contained in each. □

For applications, Theorem 1 is usually expressed as follows.

Theorem 2.2.1′

(first induction law). *A proposition $P(n)$ involving a natural n holds for all $n \in N$ in a field F if*

$$(i) \text{ it holds for } n=1, \text{ i.e., } P(1) \text{ is true; and} \tag{2.2.9}$$

$$(ii)\ \ \text{whenever } P(n) \text{ holds for } n=m, \text{ it holds for } n=m+1, \text{ i.e.,} \tag{2.2.10}$$

$$P(m) \Longrightarrow P(m+1). \tag{2.2.11}$$

Proof

Let S be the set of all those $n \in N$ for which $P(n)$ is true,

$$S=\{n \in N | P(n)\} \tag{2.2.12}$$

We have to show that actually each $n \in N$ is in S, i.e., $N \subseteq S$

First, we show that S is *inductive*.

Indeed, by assumption (i), $P(1)$ is true; so 1$\in S$.

Next, let $x \in S$. This means that $P(x)$ is true. By assumption (ii), however, this implies $P(x+1)$, i.e., $x+1 \in S$. Thus

$$1 \in S \text{ and } (\forall x \in S) \ x+1 \in S \tag{2.2.13}$$

S is inductive.

Then, by Theorem 1 (second clause), $N \subseteq S$, and all is proved. $\square$

This theorem is used to prove various properties of N "*by induction*."

Example 2.2.1

(a) If $m, n \in N$, then also $m+n \in N$ and $mn \in N$.

To prove the first property, fix any $m \in N$. Let $P(n)$ mean

$$m+n \in N \quad (n \in N) \tag{2.2.14}$$

Then

(i) $P(1)$ is true, for as $m \in N$, the definition of N yields $m+1 \in N$, i.e., $P(1)$.

(ii) $P(k) \Rightarrow P(k+1)$ for $k \in N$. Indeed,

$$P(k) \Rightarrow \ m+k \in N \Rightarrow (m+k)+1 \in N \tag{2.2.15}$$
$$\Rightarrow m+(k+1) \in N \Rightarrow P(k+1) \tag{2.2.16}$$

Thus, by Theorem $1'$, $P(n)$ holds for all n; i.e.,

$(\forall n \in N) \quad m+n \in N$

for any $m \in N$.

To prove the same for mn, we let $P(n)$ mean

$mn \in N \quad (n \in N)$

and proceed similarly.

(b) If $n \in N$, then $n-1=0$ or $n-1 \in N$.

For an inductive proof, let $P(n)$ mean

$$n-1=0 \text{ or } n-1 \in N \quad (n \in N) \tag{2.2.17}$$

Then proceed as in (a).

(c) *In an ordered field, all naturals are* ≥ 1.

Indeed, let $P(n)$ mean that

$$n \geq 1 \quad (n \in N). \tag{2.2.18}$$

Then

(i) $P(1)$ *holds* since $1=1$
(ii) $P(m) \Rightarrow P(m+1)$ for $m \in N$, since

$$P(m) \Rightarrow m \geq 1 \Rightarrow (m+1) > 1 \Rightarrow P(m+1) \tag{2.2.19}$$

Thus Theorem $1'$ yields the result.

(d) *In an ordered field,* $m, n \in N$ and $m > n$ implies $m-n \in N$
For an inductive proof, fix any $m \in N$ and let $P(n)$ mean

$$m-n \leq 0 \text{ or } m-n \in N \quad (n \in N). \tag{2.2.20}$$

Use (b).

(e) In an ordered field, $m, n \in N$ and $m<n+1$ implies $m \leq n$
For, by $(\mathrm{d}), m>n$ would imply $m-n \in N,$ hence $m-n \geq 1,$ or $m \geq n+1,$ contrary to $m<n+1$.

Our next theorem states the so-called *well-ordering property of* N.

Theorem 2.2.2 (well-ordering of N)

In an ordered field, each nonvoid set $A \subseteq N$ has a least member (i.e., one that exceeds no other element of A).

Proof

The following is just an outline of a proof:

Given $\emptyset \neq A \subseteq N,$ let $P(n)$ be the proposition "Any subset of A containing elements $\leq n$ has a least member" $(n \in N)$. Use Theorem $1^{\prime}$ and Example (e) . $\square$

This theorem yields a new form of the induction law.

Theorem 2.2.2′ (second induction law)

A proposition $P(n)$ holds for all $n \in N$

(i') $P(1)$ holds and
(ii') whenever $P(n)$ holds for all naturals less than some $m \in N,$ then $P(n)$
also holds for $n=m$.

Proof

Assume $\left(\mathrm{i}^{\prime}\right)$ and $\left(\mathrm{ii}^{\prime}\right)$. Seeking a contradiction, suppose there are some $n \in N($ call them "bad") for which $P(n)$ fails. Then these "bad" naturals form a nonvoid subset of N, call it A.

By Theorem 2, A has a least member m. Thus m is the least natural for which $P(n)$ fails. It follows that all n less than m do satisfy $P(n)$. But then, by our assumption (ii'), $P(n)$ also holds for $n=m$, which is impossible for, by construction, m is "bad" (it is in A). This contradiction shows that there are no "bad" naturals. Thus all is proved. $\square$

Note 1: All the preceding arguments hold also if, in our definition of N and all formulations, the unity 1 is replaced by 0 or by some $k(\pm k \in N)$. Then, however, the conclusions must be changed to say that $P(n)$ holds for all integers $n \geq k$ (instead of "n \geq 1 "). We then say that "induction starts with k."

An analogous induction law also applies to definitions of concepts $C(n)$.

A notion $C(n)$ involving a natural n is regarded as defined for each $n \in N$ $\left(inE^{1}\right)$ if

(i) it is defined for $n=1$ and

(ii) some rule is given that expresses $C(n+1)$ in terms of $C(1), \ldots, C(n)$.

(Note 1 applies here, too.)

$C(n)$ itself need not be a *number*; it may be of quite general nature.

We shall adopt this principle as a kind of logical axiom, without proof (though it can be proved in a similar manner as Theorems $1^{\prime}$ and $2^{\prime}$). The underlying intuitive idea is a "step-by-step" process - first, we define $C(1)$; then, as $C(1)$ is known, we may use it to define $C(2)$; next, once both are known, we may use them to define $C(3)$; and so on, ad infinitum. Definitions based on that principle are called *inductive* or *recursive*. The following examples are important.

Example 2.2.1 (continued)

(f) For any element x of a field, we define its n^{th} *power* x^n and its n-*multiple* nx by

(i) $x^1 = 1x = x$
(ii) $x^{n+1} = x^n x$ (respectively, $(n+1)x = nx + x$) .

We may think of it as a step-by-step definition:

$$x^1 = x, x^2 = x^1 x, x^3 = x^2 x, \text{ etc.} \tag{2.2.21}$$

(g) For each natural number n, we define its *factorial* $n!$ by

$$1! = 1, (n+1)! = n!(n+1); \tag{2.2.22}$$

e.g. , $2! = 1!(2) = 2, 3! = 2!(3) = 6,$ etc. We also define $0! = 1$.

(h) The sum and product of n field elements $x_1, x_2, \ldots, x_n$, denoted by

$$\sum_{k=1}^{n} x_k \text{ and } \prod_{k=1}^{n} x_k \tag{2.2.23}$$

or

$$x_1 + x_2 + \cdots + x_n \text{ and } x_1 x_2 \cdots x_n, \text{ respectively} \tag{2.2.24}$$

are defined recursively.

Sums are defined by

$$(i) \quad \sum_{k=1}^{1} x_k = x_1 \tag{2.2.25}$$

$$(ii) \quad \sum_{k=1}^{n+1} x_k = \left(\sum_{k=1}^{n} x_k\right) + x_{n+1}, n = 1, 2, \ldots \tag{2.2.26}$$

Thus

$$x_1 + x_2 + x_3 = (x_1 + x_2) + x_3 \tag{2.2.27}$$

$$x_1 + x_2 + x_3 + x_4 = (x_1 + x_2 + x_3) + x_4, \text{ etc.} \tag{2.2.28}$$

Products are defined by

$$(i) \quad \prod_{k=1}^{1} x_k = x_1 \tag{2.2.29}$$

$$(ii) \quad \prod_{k=1}^{n+1} x_k = \left(\prod_{k=1}^{n} x_k\right) \cdot x_{n+1} \tag{2.2.30}$$

(i) Given any objects $x_1, x_2, \ldots, x_n, \ldots$, the *ordered n-tuple*

$$(x_1, x_2, \ldots, x_n) \tag{2.2.31}$$

is defined inductively by

(i) $(x_1) = x_1$ *(i.e., the ordered "one-tuple"* (x_1) *is* x_1 *itself) and*

(ii) $(x_1, x_2, \ldots, x_{n+1}) = ((x_1, \ldots, x_n), x_{n+1})$ *, i.e., the ordered* $(n+1)-$ *tuple is a pair* (y, x_{n+1}) *in which the first term* y *is itself an ordered* n *-tuple,* $(x_1, \ldots, x_n)$ *; for example,*

$$(x_1, x_2, x_3) = ((x_1, x_2), x_3), \text{ etc.} \tag{2.2.32}$$

detailed edit history is available upon request.

2.2.E: Problems on Natural Numbers and Induction (Exercises)

Exercise $2.2.E.1$

Complete the missing details in Examples (a), (b), and (d).

Exercise $2.2.E.2$

Prove Theorem 2 in detail.

Exercise $2.2.E.3$

Suppose $x_k < y_k, k = 1, 2, \ldots,$ in an ordered field. Prove by induction on n that
(a) $\sum_{k=1}^{n} x_k < \sum_{k=1}^{n} y_k$
(b) if all x_k, y_k are greater than zero, then

$$\prod_{k=1}^{n} x_k < \prod_{k=1}^{n} y_k \tag{2.2.E.1}$$

Exercise $2.2.E.4$

Prove by induction that
(i) $1^n = 1$;
(ii) $a < b \Rightarrow a^n < b^n$ if $a > 0$.
Hence deduce that
(iii) $0 \leq a^n < 1$ if $0 \leq a < 1$;
(iv) $a^n < b^n \Rightarrow a < b$ if $b > 0$; proof by contradiction.

Exercise $2.2.E.5$

Prove the Bernoulli inequalities: For any element ε of an ordered field,
(i) $(1+\varepsilon)^n \geq 1 + n\varepsilon$ if $\varepsilon > -1$;
(ii) $(1-\varepsilon)^n \geq 1 - n\varepsilon$ if $\varepsilon < 1; n = 1, 2, 3, \ldots$

Exercise $2.2.E.6$

For any field elements a, b and natural numbers m, n, prove that

$$\begin{array}{ll} \text{(i) } a^m a^n = a^{m+n}; & \text{(ii) } (a^m)^n = a^{mn} \\ \text{(iii) } (ab)^n = a^n b^n; & \text{(iv) } (m+n)a = ma + na \\ \text{(v) } n(ma) = (nm) \cdot a; & \text{(vi) } n(a+b) = na + nb \end{array} \tag{2.2.E.2}$$

[Hint: For problems involving two natural numbers, fix m and use induction on n].

Exercise $2.2.E.7$

Prove that in any field,

$$a^{n+1} - b^{n+1} = (a-b) \sum_{k=0}^{n} a^k b^{n-k}, \quad n = 1, 2, 3, \ldots \tag{2.2.E.3}$$

Hence for $r \neq 1$

$$\sum_{k=0}^{n} ar^k = a\frac{1-r^{n+1}}{1-r} \tag{2.2.E.4}$$

(sum of n terms of a geometric series).

Exercise 2.2.E.8

For $n>0$ define

$$\binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!}, & 0 \leq k \leq n \\ 0, & \text{otherwise} \end{cases} \tag{2.2.E.5}$$

Verify Pascal's law,

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}. \tag{2.2.E.6}$$

Then prove by induction on n that
(i) $(\forall k \mid 0 \leq k \leq n) \binom{n}{k} \in N$; and
(ii) for any field elements a and b,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}, \quad n \in N \text{ (the binomial theorem).} \tag{2.2.E.7}$$

What value must 0^0 take for (ii) to hold for all a and b?

Exercise 2.2.E.9

Show by induction that in an ordered field F any finite sequence $x_1, \ldots, x_n$ has a largest and a least term (which need not be x_1 or x_n). Deduce that all of N is an infinite set, in any ordered field.

Exercise 2.2.E.10

Prove in E^1 that
(i) $\sum_{k=1}^{n} k = \frac{1}{2} n(n+1)$;
(ii) $\sum_{k=1}^{n} k^2 = \frac{1}{6} n(n+1)(2n+1)$;
(iii) $\sum_{k=1}^{n} k^3 = \frac{1}{4} n^2 (n+1)^2$;
(iv) $\sum_{k=1}^{n} k^4 = \frac{1}{30} n(n+1)(2n+1)\left(3n^2+3n-1\right)$.

2.3: Integers and Rationals

All natural elements of a field F, their additive inverses, and 0 are called the *integral elements* of F, briefly *integers*.

An element $x \in F$ is said to be *rational* iff $x = \frac{p}{q}$ for some integers p and q $(q \neq 0)$; x is *irrational* iff it is not rational.

We denote by J the set of all integers, and by R the set of all rationals, in F. *Every integer* p is also a *rational* since p can be written as p/q with $q = 1$
Thus

$$R \supseteq J \supset N \tag{2.3.1}$$

In an *ordered* field,

$$N = \{x \in J | x > 0\}. \text{ (Why?)} \tag{2.3.2}$$

Theorem 2.3.1

If a and b are integers (or rationals) in F, so are $a+b$ and ab.

Proof

For integers, this follows from Examples (a) and (d) in Section 2; one only has to distinguish three cases:

(i) $a, b \in N$;

(ii) $-a \in N, b \in N$;

(iii) $a \in N, -b \in N$.

The details are left to the reader (see Basic Concepts of Mathematics, Chapter 2, §7, Theorem 1).

Now let a and b be rationals, say,

$$a = \frac{p}{q} \text{ and } b = \frac{r}{s} \tag{2.3.3}$$

where $qs \neq 0$; and qs and pr are $integers$ by the first part of the proof (since $p, q, r, s \in J$).

$$a \pm b = \frac{ps \pm qr}{qs} \text{ and } ab = \frac{pr}{qs} \tag{2.3.4}$$

where $qs \neq 0$; and qs and pr are integers by the first part of the proof (since $p, q, r, s \in J$).

Thus $a \pm b$ and ab are fractions with integral numerators and denominators. Hence, by definition, $a \pm b \in R$ and $ab \in R$.
□

Theorem 2.3.2

In any field F, the set R of all rationals is a field itself, under the operations defined in F, with the same neutral elements 0 and 1. Moreover, R is an ordered field if F is. (We call R the rational subfield of F.)

Proof

We have to check that R satisfies the field axioms.

The closure law 1 follows from Theorem 1.

Axioms 2, 3, and 6 hold for rationals because they hold for *all* elements of F; similarly for Axioms 7 to 9 if F is ordered.

Axiom 4 holds in R because the neutral elements 0 and 1 *belong* to R; indeed, they are integers, hence certainly rationals.

To verify Axiom 5, we must show that $-x$ and x^{-1} *belong to* R if x does. If, however,

$$x = \frac{p}{q} \quad (p, q \in J, q \neq 0) \tag{2.3.5}$$

then

$$-x=\frac{-p}{q} \tag{2.3.6}$$

where again $-p \in J$ by the definition of J; thus $-x \in R$.

If, in addition, $x \neq 0$, then $p \neq 0$, and

$$x=\frac{p}{q} \text{ implies } x^{-1}=\frac{q}{p}. \text{ (Why?)} \tag{2.3.7}$$

Thus $x^{-1} \in R$. $\square$

Note. The representation

$$x=\frac{p}{q} \quad (p, q \in J) \tag{2.3.8}$$

is not unique in general; in an *ordered* field, however, we can always choose $q>0$, i.e., $q \in N($ take $p \leq 0$ if $x \leq 0)$.

Among all such q there is a least one by Theorem 2 of $\$85-6$. If $x=p/q$, with this *minimal* $q \in N$, we say that the rational x is given in *lowest* terms.

2.4: Upper and Lower Bounds. Completeness

A subset A of an ordered field F is said to be *bounded below* (or *left bounded*) iff there is $p \in F$ such that

$$(\forall x \in A) \quad p \leq x \tag{2.4.1}$$

A is *bounded above* (or *right bounded*) iff there is $q \in F$ such that

$$(\forall x \in A) \quad x \leq q \tag{2.4.2}$$

In this case, p and q are called, respectively, a *lower* (or *left*) bound and an *upper* (or *right*) bound, of A. If *both* exist, we simply say that A is *bounded* (by p and q). The empty set $\emptyset$ is regarded as ("vacuously") bounded by *any* p and q (cf. the end of Chapter $1, \S 3$).

The bounds p and q may, but *need not*, belong to A. If a left bound p is itself in A, we call it the *least element* or *minimum* of A, denoted min A. Similarly, if A *contains* an upper bound q, we write $q = \max A$ and call q the *largest element* or *maximum* of A. However, A may well have no minimum or
maximum.

Note 1. A finite set $A \neq \emptyset$ always has a minimum and a maximum (see Problem 9 of §§ 5-6)\).

Note 2. A set A can have at most one maximum and at most one minimum. For if it had two maxima q, q', then

$$q \leq q' \tag{2.4.3}$$

(since $q \in A$ and q' is a right bound); similarly

$$q' \leq q; \tag{2.4.4}$$

so $q = q'$ after all. Uniqueness of $\min A$ is proved in the same manner.

Note 3. If A has *one* lower bound p, it has *many* (e.g., take any $p' < p)$.

Similarly, if A has *one* upper bound q, it has *many* (take any $q' > q)$.

Geometrically, on the real axis, all lower (upper) bounds lie to the left (right) of A; see Figure 1.

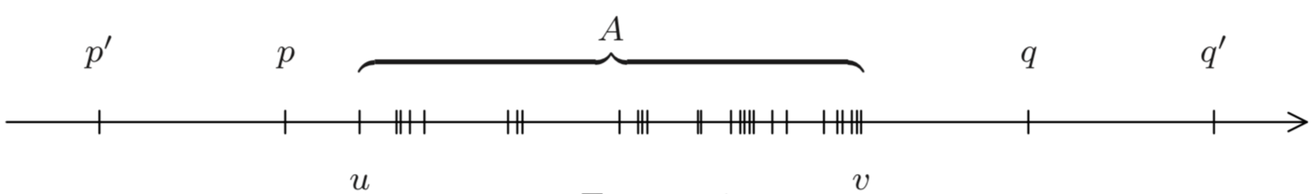

FIGURE 1

✔ Examples

(1) Let

$$A = \{1, -2, 7\}. \tag{2.4.5}$$

Then A is bounded above (e.g. , by $7, 8, 10, \ldots$)and below (e.g. , by $-2, -5, -12, \ldots$)

We have $\min A = -2, \max A = 7$.

(2) The set N of all naturals is bounded below (e.g., by $1, 0, \frac{1}{2}, -1, \ldots$) and $1 = \min N;$ N has no maximum, for each $q \in N$ is *exceeded* by some $n \in N$ (e.g. , $n = q + 1$).

(3) Given $a, b \in F (a \leq b),$ we define in F the *open interval*

$$(a, b) = \{x | a < x < b\}; \tag{2.4.6}$$

the *closed interval*

$$[a, b] = \{x | a \leq x \leq b\}; \tag{2.4.7}$$

the *half-open interval*

$$(a, b] = \{x | a < x \leq b\}; \tag{2.4.8}$$

and the *half-closed interval*

$$[a, b) = \{x | a \leq x < b\}. \tag{2.4.9}$$

Clearly, each of these intervals is bounded by the endpoints a and b; moreover, $a \in [a, b]$ and $a \in [a, b)$ (the latter provided $[a, b) \neq \emptyset$, i.e., $a < b$), and $a = \min[a, b] = \min[a, b)$; similarly, $b = \max[a, b] = \max(a, b]$. But $[a, b)$ has no maximum, $(a, b]$ has no minimum, and (a, b) has neither. (Why?)

Geometrically, it seems plausible that among all left and right bounds of A (if any) there are some "closest" to A, such as u and v in Figure 1, i.e., a *least upper bound* v and a *greatest lower bound* u. These are abbreviated

$$\operatorname{lub} A \text{ and } \operatorname{glb} A \tag{2.4.10}$$

and are also called the *supremum* and *infimum* of A, respectively; briefly,

$$v = \sup A, u = \inf A \tag{2.4.11}$$

However, this assertion, though valid in E^1, fails to materialize in many other fields such as the field R of all rationals (cf. §§$11-12$). Even for E^1, it cannot be *proved* from Axioms 1 through 9.

On the other hand, this property is of utmost importance for mathematical analysis; so we introduce it as an *axiom* (for E^1), called the *completeness axiom*. It is convenient first to give a general definition.

Definition

An ordered field F is said to be *complete* iff every nonvoid right-bounded subset $A \subset F$ has a supremum (i.e., a lub) in F.

Note that we use the term "complete" only for *ordered* fields.

With this definition, we can give the tenth and final axiom for E^1.

The Completeness Axiom

Definition

The real field E^1 is complete in the above sense. That is, each right-bounded set $A \subset E^1$ has a supremum $(\sup A)$ in E^1, provided $A \neq \emptyset$.

The corresponding assertion for *infima* can now be proved as a theorem.

Theorem 2.4.1

In a complete field F (such as E^1), every nonvoid left-bounded subset $A \subset F$ has an infimum $(i.e.,$ a glb$)$.

Proof

Let B be the (nonvoid) set of all lower bounds of A (such bounds *exist* since A is left bounded). Then, clearly, no member of B exceeds any member of A, and so B *is right bounded* by an element of A. Hence, by the assumed completeness of F, B *has a supremum in* F, call it p.

We shall show that p is also the required infimum of A, thus completing the proof.

Indeed, we have

(i) *p is a lower bound of A.* For, by definition, p is the *least* upper bound of B. But, as shown above, each $x \in A$ is an upper bound of B. Thus

$$(\forall x \in A) \quad p \leq x \tag{2.4.12}$$

(ii) p *is the greatest lower bound of* A. For $p=\sup B$ is not exceeded by any member of B. But, by definition, B contains *all* lower bounds of A; so p is not exceeded by any of them, i.e.,

$$p=\mathrm{g1b} A=\inf A \tag{2.4.13}$$

Note 4. The lub and glb of A (if they exist) are *unique*. For inf A is, by definition, the maximum of the set B of all lower bounds of A, and hence unique, by Note 2; similarly for the uniqueness of sup A.

Note 5. Unlike min A and max A, the glb and lub of A *need not* belong to A. For example, if A is the interval (a, b) in $E^1 (a<b)$ then, as is easily seen,

$$a=\inf A \text{ and } b=\sup A \tag{2.4.14}$$

though $a, b \notin A$. *Thus* sup A *and* inf A *may exist, though* max A *and* min A *do not.*

On the other hand, if

$$q=\max A(p=\min A) \tag{2.4.15}$$

then also

$$q=\sup A(p=\inf A). \quad \text{(Why?)} \tag{2.4.16}$$

Theorem 2.4.2

In an ordered field F, *we have* $q=\sup A(A \subset F)$ *iff*

(i) $(\forall x \in A) \quad x \leq q$ *and*

(ii) *each field element* $p<q$ *is exceeded by some* $x \in A$; *i.e.,*

$$(\forall p<q)(\exists x \in A) \quad p<x. \tag{2.4.17}$$

Equivalently,

(ii')

$$(\forall \varepsilon>0)(\exists x \in A) \quad q-\varepsilon<x; \quad (\varepsilon \in F) \tag{2.4.18}$$

Similarly, $p=\inf A$ *iff*

$$(\forall x \in A) \quad p \leq x \quad \text{ and } \quad (\forall \varepsilon>0)(\exists x \in A) \quad p+\varepsilon>x. \tag{2.4.19}$$

Proof

Condition (i) states that q is an upper bound of A, while (ii) implies that no *smaller* element p is such a bound (since it is *exceeded* by some x in A). When combined, (i) and (ii) state that q is the *least* upper bound.

Moreover, any element $p<q$ can be written as $q-\varepsilon(\varepsilon>0)$. Hence (ii) can be rephrased as (ii').

The proof for inf A is quite analogous. □

Corollary 2.4.1

Let $b \in F$ *and* $A \subset F$ *in an ordered field* F. *If each element* x *of* A *satisfies* $x \leq b(x \geq b)$, *so does sup* A, *respectively), provided it exists in* F.

In fact, the condition

$$(\forall x \in A) \quad x \leq b \tag{2.4.20}$$

means that b is a right bound of A. However, sup A is the *least* right bound, so sup $A \leq b$; similarly for inf A.

Corollary 2.4.2

In any ordered field, $\emptyset \neq A \subseteq B$ *implies*

$$\sup A \leq \sup B \text{ and } \inf A \geq \inf B \tag{2.4.21}$$

as well as

$$inf A \leq \sup A \tag{2.4.22}$$

provided the suprema and infima involved exist.

Proof

Let $p=\inf B$ and $q=\sup B$.

As q is a right bound of B,

$$x \leq q \text{ for all } x \in B. \tag{2.4.23}$$

But $A \subseteq B$, so B contains all elements of A. Thus

$$x \in A \Rightarrow x \in B \Rightarrow x \leq q \tag{2.4.24}$$

so, by Corollary 1, also

$$\sup A \leq q=\sup B, \tag{2.4.25}$$

as claimed.

Similarly, one gets inf $A \geq \inf B$.

Finally, if $A \neq \emptyset$, we can fix some $x \in A$. Then

$$\inf A \leq x \leq \sup A \tag{2.4.26}$$

and all is proved. □

2.4.E: Problems on Upper and Lower Bounds (Exercises)

Exercise $2.4.E.1$

Complete the proofs of Theorem 2 and Corollaries 1 and 2 for infima.
Prove the last clause of Note 4.

Exercise $2.4.E.2$

Prove that F is complete iff each nonvoid left-bounded set in F has an infimum.

Exercise $2.4.E.3$

Prove that if $A_1, A_2, \ldots, A_n$ are right bounded (left bounded) in F, so is

$$\bigcup_{k=1}^{n} A_k \tag{2.4.E.1}$$

Exercise $2.4.E.4$

Prove that if $A=(a, b)$ is an open interval $(a<b)$, then

$$a=\inf A \text{ and } b=\sup A. \tag{2.4.E.2}$$

Exercise $2.4.E.5$

In an ordered field F, let $\emptyset \neq A \subset F$. Let $c \in F$ and let cA denote the set of all products $cx(x \in A)$; i.e.,

$$cA=\{cx \mid x \in A\}. \tag{2.4.E.3}$$

$$\begin{aligned} &\text{(i) if } c \geq 0 \text{ , then} \\ &\quad \sup (cA)=c \cdot \sup A \text{ and } \inf (cA)=c \cdot \inf A \\ &\quad \text{(ii) if } c<0 \text{ , then} \\ &\quad\quad \sup (cA)=c \cdot \inf A \text{ and } \inf (cA)=c \cdot \sup A \end{aligned} \tag{2.4.E.4}$$

In both cases, assume that the right-side sup A (respectively, inf A) exists.

Exercise $2.4.E.6$

From Problem 5(ii) with $c=-1$, obtain a new proof of Theorem 1.
[Hint: If A is left bounded, show that $(-1)A$ is right bounded and use its supremum.]

Exercise $2.4.E.7$

Let A and B be subsets of an ordered field F. Assuming that the required lub and glb exist in F, prove that
(i) if $(\forall x \in A)(\forall y \in B) x \leq y$, then $\sup A \leq \inf B$;
(ii) if $(\forall x \in A)(\exists y \in B) x \leq y$, then $\sup A \leq \sup B$;
(iii) if $(\forall y \in B)(\exists x \in A) x \leq y$, then $\inf A \leq \inf B$.
$[$ Hint for (i) : By Corollary 1, $(\forall y \in B) \sup A \leq y$, so $\sup A \leq \inf B$. (Why?)$]$

Exercise 2.4.E.8

For any two subsets A and B of an ordered field F, let $A+B$ denote the set of all sums $x+y$ with $x\in A$ and $y\in B$; i.e.,

$$A+B=\{x+y \mid x\in A, y\in B\}. \tag{2.4.E.5}$$

Prove that if $\sup A=p$ and $\sup B=q$ exist in F, then

$$p+q=\sup(A+B); \tag{2.4.E.6}$$

similarly for infima.
[Hint for sup: By Theorem 2, we must show that
(i) $(\forall x\in A)(\forall y\in B)\, x+y\leq p+q(\text{ which is easy })$ and
(ii')$(\forall\varepsilon>0)(\exists x\in A)(\exists y\in B)\, x+y>(p+q)-\varepsilon$.
Fix any $\varepsilon>0$. By Theorem 2,

$$(\exists x\in A)(\exists y\in B)\quad p-\frac{\varepsilon}{2}<x \text{ and } q-\frac{\varepsilon}{2}<y. \text{ (Why?)} \tag{2.4.E.7}$$

Then

$$x+y>\left(p-\frac{\varepsilon}{2}\right)+\left(q-\frac{\varepsilon}{2}\right)=(p+q)-\varepsilon, \tag{2.4.E.8}$$

as required.]

Exercise 2.4.E.9

In Problem 8 let A and B consist of positive elements only, and let

$$AB=\{xy \mid x\in A, y\in B\}. \tag{2.4.E.9}$$

Prove that if $\sup A=p$ and $\sup B=q$ exist in F, then

$$pq=\sup(AB); \tag{2.4.E.10}$$

similarly for infima.
[Hint: Use again Theorem $2(\text{ii}')$. For $\sup(AB)$, take

$$0<\varepsilon<(p+q)\min\{p,q\} \tag{2.4.E.11}$$

and

$$x>p-\frac{\varepsilon}{p+q} \text{ and } y>q-\frac{\varepsilon}{p+q}; \tag{2.4.E.12}$$

show that

$$xy>pq-\varepsilon+\frac{\varepsilon^2}{(p+q)^2}>pq-\varepsilon. \tag{2.4.E.13}$$

For $\inf(AB)$, let $s=\inf B$ and $r=\inf A$; choose $d<1$, with

$$0<d<\frac{\varepsilon}{1+r+s}. \tag{2.4.E.14}$$

Now take $x \in A$ and $y \in B$ with

$$x < r + d \text{ and } y < s + d, \tag{2.4.E.15}$$

and show that

$$xy < rs + \varepsilon. \tag{2.4.E.16}$$

Explain!

Exercise $2.4.E.10$

Prove that
(i) if $(\forall \varepsilon > 0) a \geq b - \varepsilon$, then $a \geq b$;
(ii) if $(\forall \varepsilon > 0) a \leq b + \varepsilon$, then $a \leq b$.

Exercise $2.4.E.11$

Prove the principle of nested intervals: If $[a_n, b_n]$ are closed intervals in a complete ordered field F, with

$$[a_n, b_n] \supseteq [a_{n+1}, b_{n+1}], \quad n = 1, 2, \ldots \tag{2.4.E.17}$$

then

$$\bigcap_{n=1}^{\infty} [a_n, b_n] \neq \emptyset. \tag{2.4.E.18}$$

[Hint: Let

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}. \tag{2.4.E.19}$$

Show that A is bounded above by each b_n.
Let $p = \sup A$. (Does it exist?)
Show that

$$(\forall n) \quad a_n \leq p \leq b_n, \tag{2.4.E.20}$$

i.e.,

$$p \in [a_n, b_n].] \tag{2.4.E.21}$$

Exercise $2.4.E.12$

Prove that each bounded set $A \neq \emptyset$ in a complete field F is contained in a smallest closed interval $[a, b]$ (so $[a, b]$ is contained in any other $[c, d] \supseteq A$).
Show that this fails if "closed" is replaced by "open."
[Hint: Take $a = \inf A, b = \sup A$] .

Exercise $2.4.E.13$

Prove that if A consists of positive elements only, then $q=\sup A$ iff
(i) $(\forall x \in A) x \leq q$ and
(ii) $(\forall d>1)(\exists x \in A) q / d<x$.
[Hint: Use Theorem 2.]

2.5: Some Consequences of the Completeness Axiom

The ancient Greek geometer and scientist Archimedes was first to observe that even a large distance y can be measured by a small yardstick x; one only has to mark x off sufficiently many times. Mathematically, this means that, given any $x>0$ and any y, there is an $n \in N$ such that $nx>y$. This fact, known as the *Archimedean property*, holds not only in E^1 but also in many other ordered fields. Such fields are called *Archimedean*. In particular, we have the following theorem.

Theorem 2.5.1

Any complete field F $\left(e.g.,E^1\right)$ is Archimedean.

That is, given any $x, y \in F(x>0)$ *in such a field, there is a natural* $n \in F$ *such that* $nx>y$.

Proof

(by contradiction) Suppose this fails. Thus, given $y, x \in F(x>0)$, assume that there is no$n \in N$ with $nx>y$.

Then

$$(\forall n \in N) \quad nx \leq y \tag{2.5.1}$$

i.e., y is an upper bound of the set of *all* products $nx(n \in N)$. Let

$$A=\{nx \mid n \in N\} \tag{2.5.2}$$

Clearly, A is bounded above (by y) and $A \neq \emptyset$; so, by the assumed completeness of F, A has a supremum, say, $q=\sup A$.

As q is an upper bound, we have (by the definition of A) that $nx \leq q$ for all $n \in N$, hence also $(n+1)x \leq q$; i.e.,

$$nx \leq q-x \tag{2.5.3}$$

for all $n \in N($ since $n \in N \Rightarrow n+1 \in N)$.

Thus $q-x$ (which is less than q for $x>0$) *is another upper bound of all* nx i.e., of the set A.

This is impossible, however, since $q=\sup A$ is the least upper bound of A.

This contradiction completes the proof. □

corollary 2.5.1

In any Archimedean (*hence also in any complete*) *field* F, *the set* N *of all natural elements has no upper bounds, and the set* J *of all integers*
has neither upper nor lower bounds. Thus

$$(\forall y \in F)(\exists m, n \in N) \quad -m<y<n \tag{2.5.4}$$

Proof

Given any $y \in F$, one can use Archimedean property (with $x=1$) to find an $n \in N$ such that

$$n \cdot 1>y, \text{ i.e., } n>y. \tag{2.5.5}$$

Similarly, there is an $m \in N$ such that

$$m>-y, \text{ i.e., } -m<y. \tag{2.5.6}$$

This proves our last assertion and shows that *no*$y \in F$ can be a right bound of $N($ for $y<n \in N)$, or a left bound of $J($ for $y>-m \in J)$. □

Theorem 2.5.2

In any Archimedean (hence also in any complete) field F, *each left (right) bounded set* A *of integers* $(\emptyset \neq A \subset J)$ *has a minimum (maximum, respectively\)).*

Proof

Suppose $\emptyset \neq A \subseteq J$, and A has a *lower* bound y.

Then Corollary 1 (last part) yields a *natural* m, with $-m < y$, so that

$$(\forall x \in A) \quad -m < x, \tag{2.5.7}$$

and so $x + m > 0$.

Thus, by adding m to each $x \in A$, we obtain a set (call it $A + m$) of *naturals*.

Now, by Theorem 2 of §§$5-6$, $A+m$ has a minimum; call it p. As p is the least of all sums $x+m, p-m$ is the least of all $x \in A$; so $p-m = \min A$ exists, as claimed.

Next, let A have a *right* bound z. Then look at the set of all *additive inverses* $-x$ of points $x \in A$; call it B.

Clearly, B is left bounded (by $-z$), so it has a minimum, say, $u = \min B$. Then $-u = \max A$. (Verify!) □

In particular, given any $x \in F(F$ Archimedean), let $[x]$ denote the greatest integer $\leq x$ (called the integral part of x). We thus obtain the following corollary.

corollary 2.5.2

Any element x *of an Archimedean field* F *has an integral part* $[x]$. *It is the unique integer* n *such that*

$$n \leq x < n+1 \tag{2.5.8}$$

(*It exists, by Theorem* 2.)

Any ordered field has the so-called *density property:*

If $a < b$ in F, there is $x \in F$ such that $a < x < b$; e.g. take

$$x = \frac{a+b}{2}. \tag{2.5.9}$$

We shall now show that, in *Archimedean* fields, x can be chosen *rational*, even if a and b are not. We refer to this as the *density of rationals* in an Archimedean field

Theorem 2.5.3

(density of rationals) Between any elements a and b $(a < b)$ of an Archimedean field F (such as E^1), there is a rational $r \in F$ with

$$a < r < b. \tag{2.5.10}$$

Let $p = [a]$ (the integral part of a). The idea of the proof is to start with p and to mark off a small "yardstick"

$$\frac{1}{n} < b - a \tag{2.5.11}$$

several (m) times, until

$$p + \frac{m}{n} \text{ lands inside } (a, b) \tag{2.5.12}$$

then $r = p + \frac{m}{n}$ is the desired rational.

We now make it precise. As F is Archimedean, there are $m, n \in N$ such that

$$n(b-a)>1 \text{ and } m\left(\frac{1}{n}\right)>a-p \tag{2.5.13}$$

We fix the least such m (it exists, by Theorem 2 in $\S\S 5-6$). Then

$$a-p<\frac{m}{n}, \text{ but } \frac{m-1}{n} \leq a-p \tag{2.5.14}$$

(by the *minimality of* m). Hence

$$a<p+\frac{m}{n} \leq a+\frac{1}{n}<a+(b-a), \tag{2.5.15}$$

since $\frac{1}{n}<b-a$. Setting

$$r=p+\frac{m}{n}, \tag{2.5.16}$$

we find

$$a<r<a+b-a=b . \square \tag{2.5.17}$$

Note. Having found one rational r_1,

$$a<r_1<b, \tag{2.5.18}$$

we can apply Theorem 3 to find another $r_2 \in R$,

$$r_1<r_2<b, \tag{2.5.19}$$

then a third $r_3 \in R$,

$$r_2<r_3<b, \tag{2.5.20}$$

and so on. Continuing this process indefinitely, we obtain *infinitely many* rationals in (a, b).

2.6: Powers with Arbitrary Real Exponents. Irrationals

In *complete* fields, one can define a^r for any $a>0$ and $r \in E^1$ (for $r \in N$, see §§5-6, Example (f)). First of all, we have the following theorem.

Theorem 2.6.1

Given $a \geq 0$ in a complete field F, and a natural number $n \in E^1$, there always is a unique element $p \in F, p \geq 0$, such that

$$p^n = a. \tag{2.6.1}$$

It is called the n th root of a, denoted

$$\sqrt[n]{a} \text{ or } a^{1/n}. \tag{2.6.2}$$

(*Note that $\sqrt[n]{a} \geq 0$, by definition.*)

Proof

A direct proof, from the completeness axiom, is sketched in Problems 1 and 2 below. We shall give a simpler proof in Chapter 4,§9, Example (a). At present, we omit it and temporarily take Theorem 1 for granted. Hence we obtain the following result.

Theorem 2.6.2

Every complete field F (such as E^1) has irrational elements, i.e., elements that are not rational.

In particular, $\sqrt{2}$ is irrational.

Proof

By Theorem 1, F has the element

$$p = \sqrt{2} \text{ with } p^2 = 2 \tag{2.6.3}$$

Seeking a contradiction, suppose $\sqrt{2}$ is rational, i.e.,

$$\sqrt{2} = \frac{m}{n} \tag{2.6.4}$$

for some $m, n \in N$ in lowest terms (see §7, final note).

Then m and n are not both even (otherwise, reduction by 2 would yield a smaller n). From $m/n = \sqrt{2}$, we obtain

$$m^2 = 2n^2; \tag{2.6.5}$$

so m^2 is *even*.

Only even elements have even squares, however. Thus m itself must be even; i.e., $m = 2r$ for some $r \in N$. It follows that

$$4r^2 = m^2 = 2n^2, \text{ i.e., } 2r^2 = n^2 \tag{2.6.6}$$

and, by the same argument, n *must be even*.

This contradicts the fact that m and n are not *both* even, and this contradiction shows that $\sqrt{2}$ must be irrational. □

Note 1. Similarly, one can prove the irrationality of $\sqrt{a}$ where $a \in N$ and a is not the square of a natural. See Problem 3 below for a hint.

Note 2. Theorem 2 shows that *the field R of all rationals is not complete* (for it contains no irrationals), even though it is *Archimedean* (see Problem 6). Thus the *Archimedean property does not impleteness* (but see Theorem 1 of §10).

Next, we define a^r for any *rational* number $r > 0$.

Definition

Given $a \geq 0$ in a complete field F, and a rational number

$$r=\frac{m}{n} \quad \left(m, n \in N \subseteq E^{1}\right) \tag{2.6.7}$$

we define

$$a^{r}=\sqrt[n]{a^{m}}. \tag{2.6.8}$$

Here we must clarify two facts.

(1) If $n=1,$ we have

$$a^{r}=a^{m/1}=\sqrt[1]{a^{m}}=a^{m}. \tag{2.6.9}$$

If $m=1$, we get

$$a^{r}=a^{1/n}=\sqrt[n]{a}. \tag{2.6.10}$$

Thus Definition 1 agrees with our previous definitions of a^{m} and $\sqrt[n]{a}\ (m, n \in N)$.

(2) If r is written as a fraction *in two different ways*,

$$r=\frac{m}{n}=\frac{p}{q}, \tag{2.6.11}$$

then, as is easily seen,

$$\sqrt[n]{a^{m}}=\sqrt[q]{a^{p}}=a^{r}, \tag{2.6.12}$$

and so our definition is *unambiguous* (independent of the particular representation of r).

Indeed,

$$\frac{m}{n}=\frac{p}{q} \text{ implies } mq=np, \tag{2.6.13}$$

whence

$$a^{mq}=a^{pn}, \tag{2.6.14}$$

i.e.,

$$\left(a^{m}\right)^{q}=\left(a^{p}\right)^{n}; \tag{2.6.15}$$

cf. §§5-6, Problem 6.

By definition, however,

$$\left(\sqrt[n]{a^{m}}\right)^{n}=a^{m} \text{ and } \left(\sqrt[q]{a^{p}}\right)^{q}=a^{p}. \tag{2.6.16}$$

Substituting this in $\left(a^{m}\right)^{q}=\left(a^{p}\right)^{n},$ we get

$$\left(\sqrt[n]{a^{m}}\right)^{nq}=\left(\sqrt[q]{a^{p}}\right)^{nq}, \tag{2.6.17}$$

whence

$$\sqrt[n]{a^{m}}=\sqrt[q]{a^{p}}. \tag{2.6.18}$$

Thus Definition 1 is valid, indeed.

By using the results of Problems 4 and 6 of §§5-6, the reader will easily obtain analogous formulas for powers with positive *rational* exponents, namely,

$$\begin{aligned} a^{r} a^{s} &=a^{r+s} ; \left(a^{r}\right)^{s}=a^{rs} ; (ab)^{r}=a^{r} b^{r} ; a^{r}<a^{s} \text{ if } 0<a<1 \text{ and } r>s \\ a &<b \text{ iff } a^{r}<b^{r}\,(a, b, r>0) ; a^{r}>a^{s} \text{ if } a>1 \text{ and } r>s ; 1^{r}=1 \end{aligned}$$

Henceforth we assume these formulas known, for *rational* $r, s > 0$.

Next, we define a^r for any real $r > 0$ and any element $a > 1$ in a complete field F.

Let A_{ar} denote the set of all members of F of the form a^x, with $x \in R$ and $0 < x \le r$; i.e.,

$$A_{ar} = \{a^x | 0 < x \le r, x \text{ rational}\}. \tag{2.6.19}$$

By the density of rationals in E^1 (Theorem 3 of §10), such rationals x do exist; thus $A_{\text{ar}} \neq \emptyset$.

Moreover, A_{ar} is *right bounded* in F. Indeed, fix any rational number $y > r$. By the formulas in (1), we have, for any positive *rational* $x \le r$,

$$a^y = a^{x+(y-x)} = a^x a^{y-x} > a^x \tag{2.6.20}$$

since $a > 1$ and $y - x > 0$ implies

$$a^{y-x} > 1. \tag{2.6.21}$$

Thus a^y is an *upper bound* of all a^x in A_{ar}.

Hence, by the assumed completeness of F, sup A_{ar} *exists*. So we may define

$$a^r = \sup A_{ar}. \tag{2.6.22}$$

We also put

$$a^{-r} = \frac{1}{a^r}. \tag{2.6.23}$$

If $0 < a < 1$ (so that $\frac{1}{a} > 1$), we put

$$a^r = \left(\frac{1}{a}\right)^{-r} \text{ and } a^{-r} = \frac{1}{a^r}, \tag{2.6.24}$$

where

$$\left(\frac{1}{a}\right)^r = \sup A_{1/a,r}, \tag{2.6.25}$$

as above.

Summing up, we have the following definitions.

Definition

Given $a > 0$ in a complete field F, and $r \in E^1$, we define the following.(i) If $r > 0$ and $a > 1$, then

$$a^r = \sup A_{ar} = \sup\{a^x | 0 < x \le r, x \text{ rational }\} \tag{2.6.26}$$

(ii) If $r > 0$ and $0 < a < 1$, then $a^r = \frac{1}{(1/a)^r}$, also written $(1/a)^{-r}$.

(iii) $a^{-r} = 1/a^r$. (This defines powers with negative exponents as well.)

We also define $0^r = 0$ for any real $r > 0$, and $a^0 = 1$ for any $a \in F$, $a \neq 0$; 0^0 remains undefined.

The power a^r is also defined if $a < 0$ and r is a rational $\frac{m}{n}$ with n because $a^r = \sqrt[n]{a^m}$ has sense in this case. (Why?) This does not work for other values of r. Therefore, in general, we assume $a > 0$.

Again, it is easy to show that the formulas in (1) remain also valid for powers with real exponents (see Problems } 8-13 below), provided F is complete.

2.6.E: Problems on Roots, Powers, and Irrationals (Exercises)

Exercise $2.6.E.1$

Let $n \in N$ in E^1; let $p>0$ and $a>0$ be elements of an ordered field F.
Prove that
(i) if $p^n>a,$ then $(\exists x \in F) p>x>0$ and $x^n>a$;
(ii) if $p^n<a,$ then $(\exists x \in F) x>p$ and $x^n<a$.
[Hint: For (i), put

$$x=p-d, \text{ with } 0<d<p. \tag{2.6.E.1}$$

Use the Bernoulli inequality (Problem 5 (ii) in §§5-6) to find d such that

$$x^n=(p-d)^n>a, \tag{2.6.E.2}$$

i.e.,

$$\left(1-\frac{d}{p}\right)^n>\frac{a}{p^n}. \tag{2.6.E.3}$$

Solving for d, show that this holds if

$$0<d<\frac{p^n-a}{np^{n-1}}<p. \quad \text{(Why does such a } d \text{ exist?)} \tag{2.6.E.4}$$

For (ii), if $p^n<a,$ then

$$\frac{1}{p^n}>\frac{1}{a}. \tag{2.6.E.5}$$

Use (i) with a and p replaced by $1/a$ and $1/p$.]

Exercise $2.6.E.2$

Prove Theorem 1 assuming that
(i) $a>1$;
(ii) $0<a<1$ (the cases $a=0$ and $a=1$ are trivial).
[Hints: (i) Let

$$A=\left\{x \in F \mid x \geq 1, x^n>a\right\}. \tag{2.6.E.6}$$

Show that A is bounded below (by 1) and $A \neq \emptyset$ (e.g., $a+1 \in A-$ why?).
By completeness, put $p=\inf A$.
Then show that $p^n=a$ (i.e., p is the required $\sqrt[n]{a}$).
Indeed, if $p^n>a,$ then Problem 1 would yield an $x \in A$ with

$$x<p=\inf A. \text{ (Contradiction!)} \tag{2.6.E.7}$$

Similarly, use Problem 1 to exclude $p^n<a$.
To prove uniqueness, use Problem 4(ii) of §§5-6.
Case (ii) reduces to (i) by considering $1/a$ instead of a.]

Exercise 2.6.E.3

Prove Note 1.
[Hint: Suppose first that a is not divisible by any square of a prime, i.e.,

$$a=p_1 p_2 \cdots p_m, \tag{2.6.E.8}$$

where the p_k are distinct primes. (We assume it known that each $a \in N$ is the product of [possibly repeating] primes.) Then proceed as in the proof of Theorem 2, replacing "even" by "divisible by p_k."
The general case, $a=p^2 b,$ reduces to the previous case since $\sqrt{a}=p\sqrt{b}.$]

Exercise 2.6.E.4

Prove that if r is rational and q is not, then $r \pm q$ is irrational; so also are rq, q/r, and r/q if $r \neq 0$.
[Hint: Assume the opposite and find a contradiction.]

Exercise 2.6.E.5

$\Rightarrow 5$. Prove the density of irrationals in a complete field F : If $a<b(a, b \in F)$, there is an irrational $x \in F$ with

$$a<x<b \tag{2.6.E.9}$$

(hence infinitely many such irrationals x). See also Chapter 1, §9, Problem 4.
[Hint: By Theorem 3 of §10,

$$(\exists r \in R) \quad a\sqrt{2}<r<b\sqrt{2}, r \neq 0. \text{ (Why?)} \tag{2.6.E.10}$$

Put $x=r/\sqrt{2}$; see Problem 4].

Exercise 2.6.E.6

Prove that the rational subfield R of any ordered field is Archimedean.
[Hint: If

$$x=\frac{k}{m} \text{ and } y=\frac{p}{q} \quad (k, m, p, q \in N), \tag{2.6.E.11}$$

then $nx>y$ for $n=mp+1]$.

Exercise 2.6.E.7

Verify the formulas in (1) for powers with positive rational exponents r, s.

Exercise 2.6.E.8

Prove that
(i) $a^{r+s}=a^r a^s$ and
(ii) $a^{r-s}=a^r/a^s$ for $r, s \in E^1$ and $a \in F(a>0)$.
[Hints: For (i), if $r, s>0$ and $a>1,$ use Problem 9 in §§8-9 to get
Verify that

$$\begin{aligned} A_{ar}A_{as} &=\{a^x a^y \mid x, y \in R, 0<x \leq r, 0<y \leq s\} \\ &=\{a^z \mid z \in R, 0<z \leq r+s\}=A_{a,r+s} \end{aligned}$$

Hence deduce that

$$a^r a^s = \sup\left(A_{a,r+s}\right) = a^{r+s} \tag{2.6.E.12}$$

by Definition 2.
For (ii), if $r > s > 0$ and $a > 1$, then by (i),

$$a^{r-s} a^s = a^r \tag{2.6.E.13}$$

so

$$a^{r-s} = \frac{a^r}{a^s}. \tag{2.6.E.14}$$

For the cases $r < 0$ or $s < 0$, or $0 < a < 1$, use the above results and Definition 2(ii)(iii).]

Exercise $2.6.E.9$

From Definition 2 prove that if $r > 0\ \left(r \in E^1\right)$, then

$$a > 1 \Longleftrightarrow a^r > 1 \tag{2.6.E.15}$$

for $a \in F(a > 0)$.

Exercise $2.6.E.10$

Prove for $r, s \in E^1$ that
(i) $r < s \Leftrightarrow a^r < a^s$ if $a > 1$;
(ii) $r < s \Leftrightarrow a^r > a^s$ if $0 < a < 1$.
[Hints: (i) By Problems 8 and 9,

$$a^s = a^{r+(s-r)} = a^r a^{s-r} > a^r \tag{2.6.E.16}$$

since $a^{s-r} > 1$ if $a > 1$ and $s - r > 0$.
(ii) For the case $0 < a < 1$, use Definition 2(ii).]

Exercise $2.6.E.11$

Prove that

$$(a \cdot b)^r = a^r b^r \text{ and } \left(\frac{a}{b}\right)^r = \frac{a^r}{b^r} \tag{2.6.E.17}$$

for $r \in E^1$ and positive $a, b \in F$.
[Hint: Proceed as in Problem 8.]

Exercise $2.6.E.12$

Given $a, b > 0$ in F and $r \in E^1$, prove that
(i) $a > b \Leftrightarrow a^r > b^r$ if $r > 0$, and
(ii) $a > b \Leftrightarrow a^r < b^r$ if $r < 0$.
[Hint:

$$a>b \Longleftrightarrow \frac{a}{b}>1 \Longleftrightarrow\left(\frac{a}{b}\right)^r>1 \tag{2.6.E.18}$$

if $r>0$ by Problems 9 and 11].

Exercise $2.6.E.13$

Prove that

$$\left(a^r\right)^s=a^{rs} \tag{2.6.E.19}$$

for $r, s \in E^1$ and $a \in F(a>0)$.
[Hint: First let $r, s>0$ and $a>1$. To show that

$$\left(a^r\right)^s=a^{rs}=\sup A_{a, rs}=\sup \left\{a^{xy} \mid x, y \in R, 0<xy \leq rs\right\}, \tag{2.6.E.20}$$

use Problem 13 in §§8-9. Thus prove that
(i) $(\forall x, y \in R \mid 0<xy \leq rs) a^{xy} \leq\left(a^r\right)^s$, which is easy, and
(ii) $(\forall d>1)(\exists x, y \in R \mid 0<xy \leq rs)\left(a^r\right)^s<d a^{xy}$.
Fix any $d>1$ and put $b=a^r$. Then

$$\left(a^r\right)^s=b^s=\sup A_{bs}=\sup \left\{b^y \mid y \in R, 0<y \leq s\right\}. \tag{2.6.E.21}$$

Fix that y. Now

$$a^r=\sup A_{ar}=\sup \left\{a^x \mid x \in R, 0<x \leq r\right\}; \tag{2.6.E.22}$$

so

$$(\exists x \in R \mid 0<x \leq r) \quad a^r<d^{\frac{1}{2y}} a^x . \quad (\text{Why?}) \tag{2.6.E.23}$$

Combining all and using the formulas in (1) for rationals x, y, obtain

$$\left(a^r\right)^s<d^{\frac{1}{2}}\left(a^r\right)^y<d^{\frac{1}{2}}\left(d^{\frac{1}{2y}} a^x\right)^y=d a^{xy}, \tag{2.6.E.24}$$

thus proving (ii)].

2.7: The Infinities. Upper and Lower Limits of Sequences

The Infinities

As we have seen, a set $A \neq \emptyset$ in E^{1} has a lub (\mathrm{glb}) if A is bounded above (respectively, below), but not otherwise.

In order to avoid this inconvenient restriction, we now add to E^1 two new objects of arbitrary nature, and call them "minus infinity" $(-\infty)$ and "plus infinity" $(+\infty)$, with the convention that $-\infty < +\infty$ and $-\infty < x < +\infty$ for all $x \in E^1$.
It is readily seen that with this convention, the laws of transitivity and trichotomy (Axioms 7 and 8) remain valid.

The set consisting of all reals and the two infinities is called the extended real number system. We denote it by E^* and call its elements extended real numbers. The ordinary reals are also called finite numbers, while $\pm\infty$ are the only two infinite elements of E^*. (Caution: They are not real numbers.)

At this stage we do not define any operations involving $\pm\infty$. (This will be done later. However, the notions of upper and lower bound, maximum, minimum, supremum, and infimum are defined in E^* exactly as in E^1. In particular,

$$-\infty = \min E^* \text{ and } +\infty = \max E^* \tag{2.7.1}$$

Thus in E^* all sets are bounded.

It follows that in E^* every set $A \neq \emptyset$ has a lub and a glb. For if A has none in E^1, it still has the upper bound $+\infty$ in E^*, which in this case is the unique (hence also the least) upper bound; thus sup $A = +\infty$. I Similarly, inf $A = -\infty$ if there is no other lower bound.? As is readily seen, all properties of lub and glb stated in §§8-9 remain valid in E^* (with the same proof). The only exception is Theorem $2(\mathrm{ii}')$ in the case $q = +\infty$ (respectively, $p = -\infty$) $\sin\mathrm{ce} + \infty - \varepsilon$ and $-\infty + \varepsilon$ make no sense. Part (ii) of Theorem 2 is valid.

We can now define intervals in E^* exactly as in E^1 §§8-9, Example (3), allowing also infinite values of a, b, x. For example,

$$\begin{aligned} (-\infty, a) &= \{x \in E^* \mid -\infty < x < a\} = \{x \in E^1 \mid x < a\} \\ (a, +\infty) &= \{x \in E^1 \mid a < x\} \\ (-\infty, +\infty) &= \{x \in E^* \mid -\infty < x < +\infty\} = E^1 \\ [-\infty, +\infty] &= \{x \in E^* \mid -\infty \leq x \leq +\infty\}; \text{ etc.} \end{aligned}$$

Intervals with finite endpoints are said to be finite; all other intervals are called infinite. The infinite intervals

$$(-\infty, a), (-\infty, a], (a, +\infty), [a, +\infty), \quad a \in E^1 \tag{2.7.2}$$

are actually subsets of E^1, as is $(-\infty, +\infty)$. Thus we shall speak of infinite intervals in E^1 as well.

Upper and Lower Limits

In Chapter 1, §§1-3 we already mentioned that a real number p is called the limit of a sequence $\{x_n\} \subseteq E^1$ $(p = \lim x_n)$ iff

$$(\forall \varepsilon > 0)(\exists k)(\forall n > k) \quad |x_n - p| < \varepsilon, \text{ i.e., } p - \varepsilon < x_n < p - \varepsilon \tag{2.7.3}$$

where $\varepsilon \in E^1$ and $n, k \in N$.

This may be stated as follows:

For sufficiently large $n (n > k)$, x_n becomes and stays as close to p as we like ("ε-close").

We also define (in E^1 and E^*)

$$\lim_{n\to\infty} x_n = +\infty \Longleftrightarrow \left(\forall a \in E^1\right)(\exists k)(\forall n > k) \quad x_n > a \text{ and} \tag{2.7.4}$$

$$\lim_{n\to\infty} x_n = -\infty \Longleftrightarrow \left(\forall b \in E^1\right)(\exists k)(\forall n > k) \quad x_n < b. \tag{2.7.5}$$

Note that (2) and (3) make sense in E^1, too, since the symbols $\pm\infty$ do not occur on the right side of the formulas. Formula (2) means that x_n becomes arbitrarily larger than any $a \in E^1$ given in advance) for sufficiently large $n (n > k)$. The interpretation of (3) is analogous. A more general and unified approach will now be developed for E^* (allowing infinite terms x_n, too).

Let $\{x_n\}$ be any sequence in E^*. For each n, let A_n be the set of all terms from x_n onward, i.e.,

$$\{x_n, x_{n+1}, \ldots\} \tag{2.7.6}$$

For example,

$$A_1 = \{x_1, x_2, \ldots\}, A_2 = \{x_2, x_3, \ldots\}, \text{ etc.} \tag{2.7.7}$$

The A_n form a contracting sequence (see Chapter 1, §8) since

$$A_1 \supseteq A_2 \supseteq \cdots. \tag{2.7.8}$$

Now, for each n, let

$$p_n = \inf A_n \text{ and } q_n = \sup A_n \tag{2.7.9}$$

also denoted

$$p_n = \inf_{k \geq n} x_k \text{ and } q_n = \sup_{k \geq n} x_k. \tag{2.7.10}$$

(These infima and suprema always exist in E^*, as noted above.) Since $A_n \supseteq A_{n+1}$, Corollary 2 of §§8-9 yields

$$\inf A_n \leq \inf A_{n+1} \leq \sup A_{n+1} \leq \sup A_n \tag{2.7.11}$$

Thus

$$p_1 \leq p_2 \leq \cdots \leq p_n \leq p_{n+1} \leq \cdots \leq q_{n+1} \leq q_n \leq \cdots \leq q_2 \leq q_1 \tag{2.7.12}$$

and so $\{p_n\} \uparrow$, while $\{q_n\} \downarrow$ in E^*. We also see that each q_m is an upper bound of all p_n and hence

$$q_m \geq \sup_n p_n \,(= \text{lub of all } p_n). \tag{2.7.13}$$

This, in turn, shows that this sup (call it $\underline{L}$) is a lower bound of all q_m, and so

$$\underline{L} \leq \inf_m q_m. \tag{2.7.14}$$

We put

$$\inf_m q_m = \overline{L}. \tag{2.7.15}$$

Definition

For each sequence $\{x_n\} \subseteq E^*$, we define its upper limit $\overline{L}$ and its lower limit $\underline{L}$, denoted

$$\overline{L} = \overline{\lim} x_n = \limsup_{n \to \infty} x_n \text{ and } \underline{L} = \underline{\lim} x_n = \liminf_{n \to \infty} x_n \tag{2.7.16}$$

as follows.

We put $(\forall n)$

$$q_n = \sup_{k \geq n} x_k \text{ and } p_n = \inf_{k \geq n} x_k, \tag{2.7.17}$$

as before. Then we set

$$\overline{L} = \overline{\lim} x_n = \inf_n q_n \text{ and } \underline{L} = \underline{\lim} x_n = \sup_n p_n, \text{ all in } E^*. \tag{2.7.18}$$

Here and below, inf $_n q_n$ is the inf of all q_n, and $\sup_n p_n$ is the sup of all p_n.

Corollary 2.7.1

For any sequence in E^*,

$$\inf_n x_n \leq \underline{\lim} x_n \leq \overline{\lim} x_n \leq \sup_n x_n. \tag{2.7.19}$$

For, as we noted above,

$$\underline{L}=\sup_n p_n \leq \inf_m q_m=\overline{L}. \tag{2.7.20}$$

Also,

$$\underline{L} \geq p_n=\inf A_n \geq \inf A_1=\inf_n x_n \text{ and} \tag{2.7.21}$$

$$\overline{L} \leq q_n=\sup A_n \leq \sup A_1=\sup_n x_n, \tag{2.7.22}$$

with A_n as above.

Example 2.7.1

(a) Let

$$x_n=\frac{1}{n}. \tag{2.7.23}$$

Here

$$q_1=\sup\left\{1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots\right\}=1, q_2=\frac{1}{2}, q_n=\frac{1}{n}. \tag{2.7.24}$$

Hence

$$\overline{L}=\inf_n q_n=\inf\left\{1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots\right\}=0, \tag{2.7.25}$$

as easily follows by Theorem 2 in §§8-9 and the Archimedean property. (Verify!) Also,

$$p_1=\inf_{k \geq 1} \frac{1}{k}=0, p_2=\inf_{k \geq 2} \frac{1}{k}=0, \ldots, p_n=\inf_{k \geq n} \frac{1}{k}=0. \tag{2.7.26}$$

since all p_n are 0, so is $\overline{L}=\sup_n p_n$. Thus here $\underline{L}=\overline{L}=0$.

(b) Consider the sequence

$$1,-1,2,-\frac{1}{2}, \ldots, n,-\frac{1}{n}, \ldots \tag{2.7.27}$$

Here

$$p_1=-1=p_2, p_3=-\frac{1}{2}=p_4, \ldots ; p_{2n-1}=-\frac{1}{n}=p_{2n}. \tag{2.7.28}$$

Thus

$$\underline{\lim}_n x_n=\sup_n p_n=\sup\left\{-1,-\frac{1}{2}, \ldots,-\frac{1}{n}, \ldots\right\}=0. \tag{2.7.29}$$

On the other hand, $q_n=+\infty$ for all n. (Why?) Thus

$$\overline{\lim} x_n=\inf_n q_n=+\infty. \tag{2.7.30}$$

Theorem 2.7.1

(i) If $x_n \geq b$ for infinitely many n, then

$$\overline{\lim} x_n \geq b \quad \text{as well}. \tag{2.7.31}$$

(ii) If $x_n \leq a$ for all but finitely many n, then

$$\overline{\lim} x_n \leq a \quad \text{as well} . \tag{2.7.32}$$

Similarly for lower limits (with all inequalities reversed).

Proof

(i) If $x_n \geq b$ for infinitely many n, then such n must occur in each set

$$A_m = \{x_m, x_{m+1}, \ldots\} . \tag{2.7.33}$$

Hence

$$(\forall m) \quad q_m = \sup A_m \geq b; \tag{2.7.34}$$

so $\overline{L} = \inf_m q_m \geq b$, by Corollary 1 of §§8-9.

(ii) If $x_n \leq a$ except finitely many n, let n_0 be the last of these "exceptional" values of n.

Then for $n > n_0, x_n \leq a$, i.e., the set

$$A_n = \{x_n, x_{n+1}, \ldots\} \tag{2.7.35}$$

is bounded above by a; so

$$(\forall n > n_0) \quad q_n = \sup A_n \leq a. \tag{2.7.36}$$

Hence certainly $\overline{L} = \inf_n q_n \leq a$. □

corollary 2.7.2

(i) If $\overline{\lim} x_n > a$, then $x_n > a$ for infinitely many n.

(ii) If $\overline{\lim} x_n < b$, then $x_n < b$ for all but finitely many n.

Similarly for lower limits (with all inequalities reversed).

Proof

Assume the opposite and find a contradiction to Theorem 1. □

To unify our definitions, we now introduce some useful notions.

By a neighborhood of p, briefly G_p, we mean, for $p \in E^1$, any interval of the form

$$(p-\varepsilon, p+\varepsilon), \quad \varepsilon > 0. \tag{2.7.37}$$

If $p = +\infty($ respectively $, p = -\infty), G_p$ is an infinite interval of the form

$$(a, +\infty] \text{ (respectively, } [-\infty, b)), \text{ with } a, b \in E^1. \tag{2.7.38}$$

We can now combine formulas (1)-(3) into one equivalent definition.

Definition

An element $p \in E^*$ (finite or not) is called the limit of a sequence $\{x_n\}$ in E^* iff each G_p (no matter how small it is) contains all but finitely many x_n, i.e. all x_n from some x_k onward. In symbols,

$$(\forall G_p)\,(\exists k)(\forall n > k) \quad x_n \in G_p. \tag{2.7.39}$$

We shall use the notation

$$p = \lim x_n \text{ or } \lim_{n \rightarrow \infty} x_n. \tag{2.7.40}$$

Indeed, if $p \in E^1$, then $x_n \in G_p$ means

$$p-\varepsilon < x_n < p+\varepsilon, \tag{2.7.41}$$

as in (1). If, however, $p = \pm\infty$, it means

$$x_n > a \,(\text{ respectively, } x_n < b)\,, \tag{2.7.42}$$

as in (2) and (3).

Theorem 2.7.2

We have $q = \overline{\lim}\, x_n$ in E^* iff

(i) each neighborhood G_q contains x_n for infinitely many n, and

(ii') if $q < b$, then $x_n \geq b$ for at most finitely many n.

Proof

If $q = \overline{\lim}\, x_n,$ Corollary 2 yields (ii')

It also shows that any interval (a, b), with $a < q < b,$ contains infinitely many x_n (for there are infinitely many $x_n > a$, and only finitely many $x_n \geq b$ by (ii')).

Now if $q \in E^1$,

$$G_q = (q - \varepsilon, q + \varepsilon) \tag{2.7.43}$$

is such an interval, so we obtain (i') . The cases $q = \pm\infty$ are analogous; we leave them to the reader.

Conversely, assume (i') and (ii') .

Seeking a contradiction, let $q < \overline{L};$ say,

$$q < b < \overline{\lim}\, x_n. \tag{2.7.44}$$

Then Corollary $2(\text{i})$ yields $x_n > b$ for infinitely many n, contrary to our assumption $(\text{i}i')$.

Similarly, $q > \overline{\lim}\, x_n$ would contradict (i') .

Thus necessarily $q = \overline{\lim}\, x_n. \square$

Theorem 2.7.3

We have $q = \lim x_n$ in E^* iff

$$\underline{\lim}\, x_n = \overline{\lim}\, x_n = q. \tag{2.7.45}$$

Proof

Suppose

$$\underline{\lim}\, x_n = \overline{\lim}\, x_n = q. \tag{2.7.46}$$

If $q \in E^1$, then every G_q is an interval $(a, b), a < q < b$; therefore, Corollary 2(ii) and its analogue for $\underline{\lim}\, x_n$ imply (with q treated as both $\overline{\lim}\, x_n$ and $\underline{\lim}\, x_n$ that

$$a < x_n < b \text{ for all but finitely many n.} \tag{2.7.47}$$

Thus by Definition $2, q = \lim x_n$, as claimed.

Conversely, if so, then any G_q (no matter how small) contains all but finitely many x_n. Hence so does any interval (a, b) with $a < q < b,$ for it contains some small G_q.

Now, exactly as in the proof of Theorem 2, one excludes

$$q \neq \underline{\lim}\, x_n \text{ and } q \neq \overline{\lim}\, x_n. \tag{2.7.48}$$

This settles the case $q \in E^{1}$. The cases $q = \pm\infty$ are quite analogous. □

2.7.E: Problems on Upper and Lower Limits of Sequences in E ∗ E∗ (Exercises)

Exercise 2.7.E.1

Complete the missing details in the proofs of Theorems 2 and 3, Corollary 1, and Examples (a) and (b).

Exercise 2.7.E.2

State and prove the analogues of Theorems 1 and 2 and Corollary 2 for
$\underline{\lim} x_n$.

Exercise 2.7.E.3

Find $\overline{\lim} x_n$ and $\underline{\lim} x_n$ if
(a) $x_n = c$ (constant);
(b) $x_n = -n$;
(c) $x_n = n$; and
(d) $x_n = (-1)^n n - n$
Does $\lim x_n$ exist in each case?

Exercise 2.7.E.4

$\Rightarrow 4$. A sequence $\{x_n\}$ is said to cluster at $q \in E^*$, and q is called its cluster point, iff each G_q contains x_n for infinitely many values of n.

Show that both $\underline{L}$ and $\overline{L}$ are cluster points ($\underline{L}$ the least and $\overline{L}$ the largest).
[Hint: Use Theorem 2 and its analogue for $\underline{L}$.
To show that no $p < \underline{L}$ (or $q > \overline{L}$) is a cluster point, assume the opposite and find a contradiction to Corollary 2.]

Exercise 2.7.E.5

$\Rightarrow 5$. Prove that
(i) $\overline{\lim}(-x_n) = -\underline{\lim} x_n$ and
(ii) $\overline{\lim}(a x_n) = a \cdot \overline{\lim} x_n$ if $0 \leq a < +\infty$.

Exercise 2.7.E.6

Prove that

$$\overline{\lim} x_n < +\infty \, (\underline{\lim} x_n > -\infty) \tag{2.7.E.1}$$

iff $\{x_n\}$ is bounded above (below) in E^1.

Exercise 2.7.E.7

Prove that if $\{x_n\}$ and $\{y_n\}$ are bounded in E^1, then

$$\overline{\lim} x_n + \overline{\lim} y_n \geq \overline{\lim}(x_n + y_n) \geq \overline{\lim} x_n + \underline{\lim} y_n \geq \underline{\lim}(x_n + y_n) \geq \underline{\lim} x_n + \underline{\lim} y_n . \tag{2.7.E.2}$$

[Hint: Prove the first inequality and then use that and Problem 5(i) for the others.]

Exercise $2.7.E.8$

$\Rightarrow 8$. Prove that if $p=\lim x_n$ in E^1, then

$$\underline{\lim}(x_n+y_n)=p+\underline{\lim}\, y_n; \tag{2.7.E.3}$$

similarly for $\overline{L}$.

Exercise $2.7.E.9$

$\Rightarrow 9$. Prove that if $\{x_n\}$ is monotone, then $\lim x_n$ exists $in E^*$. Specifically, if $\{x_n\}\uparrow$, then

$$\lim x_n=\sup_n x_n, \tag{2.7.E.4}$$

and if $\{x_n\}\downarrow$, then

$$\lim x_n=\inf_n x_n. \tag{2.7.E.5}$$

Exercise $2.7.E.10$

$\Rightarrow 10$. Prove that
(i) if $\lim x_n=+\infty$ and $(\forall n) x_n\leq y_n$, then also $\lim y_n=+\infty$, and
(ii) if $\lim x_n=-\infty$ and $(\forall n) y_n\leq x_n$, then also $\lim y_n=-\infty$.

Exercise $2.7.E.11$

Prove that if $x_n\leq y_n$ for all n, then

$$\underline{\lim}\, x_n\leq\underline{\lim}\, y_n \text{ and } \overline{\lim}\, x_n\leq\overline{\lim}\, y_n. \tag{2.7.E.6}$$

CHAPTER OVERVIEW

3: Vector Spaces and Metric Spaces

3.1: The Euclidean n-Space, E^n

By definition, the *Euclidean* n*-space* E^n is the set of all possible ordered n-tuples of real numbers, i.e., the Cartesian product

$$E^1 \times E^1 \times \cdots \times E^1 (n \text{ times}). \tag{3.1.1}$$

In particular, $E^2 = E^1 \times E^1 = \{(x,y) | x, y \in E^1\}$,

$$E^3 = E^1 \times E^1 \times E^1 = \{(x,y,z) | x, y, z \in E^1\}, \tag{3.1.2}$$

and so on. E^1 itself is a special case of $E^n (n = 1)$. In a familiar way, pairs (x, y) can be plotted as points of the xy -plane, or as "vectors" (directed line segments) joining $(0, 0)$ to such points. Therefore, the pairs (x, y) themselves are called points or vectors in E^2; similarly for E^3.

In $E^n (n > 3)$, there is no actual geometric representation, but it is convenient to use geometric language in this case, too. Thus any ordered n-tuple $(x_1, x_2, \ldots, x_n)$ of real numbers will also be called a point or vector in E^n, and the single numbers $x_1, x_2, \ldots, x_n$ are called its coordinates or components. A point in E^n is often denoted by a single letter (preferably with a bar or an arrow above it), and then its n components are denoted by the same letter, with subscripts (but without the bar or arrow). For example,

$$\overline{x} = (x_1, \ldots, x_n), \vec{u} = (u_1, \ldots, u_n), etc.; \tag{3.1.3}$$

$\overline{x} = (0, -1, 2, 4)$ is a point (vector) in E^4 with coordinates $0, -1, 2$, and 4 (in this order). The formula $\overline{x} \in E^n$ means that $\overline{x} = (x_1, \ldots, x_n)$ is a point (vector) in E^n. since such "points" are ordered n -tuples, $\overline{x}$ and $\overline{y}$ are equal $(\overline{x} = \overline{y})$ iff the corresponding coordinates are the same, i.e., $x_1 = y_1, x_2 = y_2 \ \ldots, x_n = y_n$ (see Problem 1 below).

The point whose coordinates are all 0 is called the z ero-vector or the origin, denoted $\overrightarrow{0}$ or $\overline{0}$. The vector whose k th component is 1, and the other components are 0, is called the k th basic unit vector, denoted $\vec{e}_k$. There are exactly n such vectors,

$$\vec{e}_1 = (1, 0, 0, \ldots, 0), \vec{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \vec{e}_n = (0, \ldots, 0, 1) \tag{3.1.4}$$

In E^3, we often write $\vec{i}, \vec{j}$, and $\vec{k}$ for $\vec{e}_1$, and (x, y, z) for (x_1, x_2, x_3). Similarly in E^2. Single real numbers are called scalars (as opposed to vectors).

Definition

Given $\overline{x} = (x_1, \ldots, x_n)$ and $\overline{y} = (y_1, \ldots, y_n)$ in E^n, we define the following.

1. The sum of $\overline{x}$ and $\overline{y}$,

$$\overline{x} + \overline{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \text{ (hence } \overline{x} + \overline{0} = \overline{x}). \tag{3.1.5}$$

2. The dot product, or inner product, of $\overline{x}$ and $\overline{y}$,

$$\overline{x} \cdot \overline{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n. \tag{3.1.6}$$

3. The distance between $\overline{x}$ and $\overline{y}$,

$$\rho(\overline{x}, \overline{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}. \tag{3.1.7}$$

4. The absolute value, or length, of $\overline{x}$,

$$|\overline{x}| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \rho(\overline{x}, \overline{0}) = \sqrt{\overline{x} \cdot \overline{x}} \tag{3.1.8}$$

(three formulas that are all equal by Definitions 2 and 3).

5. The inverse of $\overline{x}$,

$$-\overline{x} = (-x_1, -x_2, \ldots, -x_n). \tag{3.1.9}$$

6. The product of $\overline{x}$ by a scalar $c \in E^1$,

$$c\overline{x} = \overline{x}c = (cx_1, cx_2, \ldots, cx_n); \tag{3.1.10}$$

in particular, $(-1)\overline{x}=(-x_1,-x_2,\ldots,-x_n)=-\overline{x}, 1\overline{x}=\overline{x},$ and $0\overline{x}=\overline{0}$.

7. The difference of $\overline{x}$ and $\overline{y}$,

$$\overline{x}-\overline{y}=\overrightarrow{yx}=(x_1-y_1,x_2-y_2,\ldots,x_n-y_n). \tag{3.1.11}$$

In particular, $\overline{x}-\overline{0}=\overline{x}$ and $\overline{0}-\overline{x}=-\overline{x}$. (Verify!)

Note 1. Definitions $2-4$ yield *scalars*, while the rest are *vectors*.

Note 2. We shall not define *inequalities* $(<)$ in $E^n (n \geq 2)$, nor shall we define vector products other than the *dot product* (2), which is a *scalar*.

Note 3. From Definitions 3, 4, and 7, we obtain $\rho(\overline{x},\overline{y})=|\overline{x}-\overline{y}|$. (Verify!)

Note 4. We often write $\overline{x}/c$ for $(1/c)\overline{x}$, where $c\in E^1, c\neq 0$.

Note 5. In $E^1, \overline{x}=(x_1)=x_1$. Thus, by Definition 4,

$$|\overline{x}|=\sqrt{x_1^2}=|x_1|, \tag{3.1.12}$$

where $|x_1|$ is defined as in Chapter 2, §§1, Definition 4. Thus the two definitions *agree*.

We call $\overline{x}$ a unit vector iff its length is 1, i.e., $|x|=1$. Note that if $\overline{x}\neq\overline{0}$, then $\overline{x}/|\overline{x}/$ is a unit vector, since

$$\left|\frac{\overline{x}}{|\overline{x}|}\right|=\sqrt{\frac{x_1^2}{|\overline{x}|^2}+\cdots+\frac{x_n^2}{|\overline{x}|^2}}=1. \tag{3.1.13}$$

The vectors $\overline{x}$ and $\overline{y}$ are said to be orthogonal or perpendicular $(\overline{x}\perp\overline{y})$ iff $\overline{x}\cdot\overline{y}=0$ and parallel$(\overline{x}\|\overline{y})$ iff $\overline{x}=t\overline{y}$ or $\overline{y}=t\overline{x}$ for some $t\in E^1$. Note that $\overline{x}\perp\overline{0}$ and $\overline{x}\|\overline{0}$.

Example 3.1.1

If $\overline{x}=(0,-1,4,2)$ and $\overline{y}=(2,2,-3,2)$ are vectors in E^4, then

$$\begin{aligned}\overline{x}+\overline{y}&=(2,1,1,4);\\ \overline{x}-\overline{y}&=(-2,-3,7,0);\\ \rho(\overline{x},\overline{y})=|\overline{x}-\overline{y}|&=\sqrt{2^2+3^2+7^2+0^2}=\sqrt{62};\\ (\overline{x}+\overline{y})\cdot(\overline{x}-\overline{y})&=2(-2)+1(-3)+7+0=0.\end{aligned}$$

So $(\overline{x}+\overline{y})\perp(\overline{x}-\overline{y})$ here.

Theorem 3.1.1

For any vectors $\overline{x},\overline{y}$, and $\overline{z}\in E^n$ and any $a,b\in E^1$, we have

(a) $\overline{x}+\overline{y}$ and a $\overline{x}$ are vectors in E^n (closure laws);

(b) $\overline{x}+\overline{y}=\overline{y}+\overline{x}$ (commutativity of vector addition);

(c) $(\overline{x}+\overline{y})+\overline{z}=\overline{x}+(\overline{y}+\overline{z})$ (associativity of vector addition);

(d) $\overline{x}+\overline{0}=\overline{0}+\overline{x}=\overline{x}$, i.e., $\overline{0}$ is the neutral element of addition;

(e) $\overline{x}+(-\overline{x})=\overline{0}$, i.e., $-\overline{x}$ is the additive inverse of $\overline{x}$;

(f) $a(\overline{x}+\overline{y})=a\overline{x}+a\overline{y}$ and $(a+b)\overline{x}=a\overline{x}+b\overline{x}$ (distributive laws);

(g) $(ab)\overline{x}=a(b\overline{x})$;

(h) $1\overline{x}=\overline{x}$.

Proof

Assertion (a) is immediate from Definitions 1 and 6. The rest follows from corresponding properties of real numbers.

For example, to prove (b), let $\bar{x}=(x_1,\ldots,x_n), \bar{y}=(y_1,\ldots,y_n)$. Then by definition, we have

$$\bar{x}+\bar{y}=(x_1+y_1,\ldots,x_n+y_n) \text{ and } \bar{y}+\bar{x}=(y_1+x_1,\ldots,y_n+x_n). \tag{3.1.14}$$

The right sides in both expressions, however, coincide since addition is commutative $in E^1$. Thus $\bar{x}+\bar{y}=\bar{y}+\bar{x}$, as claimed; similarly for the rest, which we leave to the reader. □

Theorem 3.1.2

If $\bar{x}=(x_1,\ldots,x_n)$ is a vector in E^n, then, with $\bar{e}_k$ as above,

$$\bar{x}=x_1\bar{e}_1+x_2\bar{e}_2+\cdots+x_n\bar{e}_n=\sum_{k=1}^n x_k\bar{e}_k. \tag{3.1.15}$$

Moreover, if $\bar{x}=\sum_{k=1}^n a_k\bar{e}_k$ for some $a_k\in E^1$, then necessarily $a_k=x_k$, $k=1,\ldots,n$.

Proof

By definition,

$$\bar{e}_1=(1,0,\ldots,0),\bar{e}_2=(0,1,\ldots,0),\ldots,\bar{e}_n=(0,0,\ldots,1). \tag{3.1.16}$$

Thus

$$x_1\bar{e}_1=(x_1,0,\ldots,0), x_2\bar{e}_2=(0,x_2,\ldots,0),\ldots,x_n\bar{e}_n=(0,0,\ldots,x_n). \tag{3.1.17}$$

Adding up componentwise, we obtain

$$\sum_{k=1}^n x_k\bar{e}_k=(x_1,x_2,\ldots,x_n)=\bar{x}, \tag{3.1.18}$$

as asserted.

Moreover, if the x_k are replaced by any other $a_k\in E^1$, the same process yields

$$(a_1,\ldots,a_n)=\bar{x}=(x_1,\ldots,x_n), \tag{3.1.19}$$

i.e., the two n -tuples coincide, whence $a_k=x_k, k=1,\ldots,n$. □

Note 6. Any sum of the form

$$\sum_{k=1}^m a_k\bar{x}_k \quad \left(a_k\in E^1,\bar{x}_k\in E^n\right) \tag{3.1.20}$$

is called a linear combination of the vectors $\bar{x}_k$ (whose number m is arbitrary). Thus Theorem 2 shows that $any \bar{x}\in E^n$ can be expressed, in a unique way, as a linear combination of the n basic unit vectors. In E^3, we write

$$\bar{x}=x_1\bar{i}+x_2\bar{j}+x_3\bar{k}. \tag{3.1.21}$$

Note 7. If, as above, some vectors are numbered (e.g., $\bar{x}_1,\bar{x}_2,\ldots,\bar{x}_m$), we denote their components by attaching a second subscript; for example, the components of $\bar{x}_1$ are $x_{11},x_{12},\ldots,x_{1n}$.

Theorem 3.1.3

For any vectors $\bar{x},\bar{y}$, and $\bar{z}\in E^n$ and any $a,b\in E^1$, we have

(a) $\bar{x}\cdot\bar{x}\geq 0$, and $\bar{x}\cdot\bar{x}>0$ iff $\bar{x}\neq\overline{0}$;

(b) $(a\bar{x})\cdot(b\bar{y})=(ab)(\bar{x}\cdot\bar{y})$;

(c) $\bar{x}\cdot\bar{y}=\bar{y}\cdot\bar{x}$ (commutativity of inner products);

(d) $(\bar{x}+\bar{y})\cdot\bar{z}=\bar{x}\cdot\bar{z}+\bar{y}\cdot\bar{z}($ distributive law).

Proof

To prove these properties, express all in terms of the components of $\overline{x}$, $\overline{y}$, and $\overline{z}$, and proceed as in Theorem 1. □

Note that (b) implies $\overline{x} \cdot \overline{0} = 0$ (put $a = 1, b = 0$).

Theorem 3.1.4

For any vectors $\overline{x}$ and $\overline{y} \in E^n$ and any $a \in E^1$, we have the following properties:

(a') $|\overline{x}| \geq 0$, and $|\overline{x}| > 0$ iff $\overline{x} \neq \overline{0}$.

(b') $|a\overline{x}| = |a||\overline{x}|$.

(c') $|\overline{x} \cdot \overline{y}| \leq |\overline{x}||\overline{y}|$, or, in components,

$$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leq \left(\sum_{k=1}^{n} x_k^2\right)\left(\sum_{k=1}^{n} y_k^2\right) \quad \text{(Cauchy-Schwarz inequality)} \tag{3.1.22}$$

Equality, $|\overline{x} \cdot \overline{y}| = |\overline{x}||\overline{y}|$, holds iff $\overline{x} \| \overline{y}$.

(d') $|\overline{x}+\overline{y}| \leq |\overline{x}|+|\overline{y}|$ and $||\overline{x}|-|\overline{y}|| \leq |\overline{x}-\overline{y}|$ (triangle inequalities).

Proof

Property (a') follows from Theorem $3(\mathrm{a})$ since

$$|\overline{x}|^2 = \overline{x} \cdot \overline{x} (\text{see Definition } 4). \tag{3.1.23}$$

For (b'), use Theorem $3(\mathrm{b})$, to obtain

$$(a\overline{x}) \cdot (a\overline{x}) = a^2(\overline{x} \cdot \overline{x}) = a^2|\overline{x}|^2. \tag{3.1.24}$$

By Definition 4, however,

$$(a\overline{x}) \cdot (a\overline{x}) = |a\overline{x}|^2. \tag{3.1.25}$$

Thus

$$|a\overline{x}|^2 = a^2|x|^2 \tag{3.1.26}$$

so that $|a\overline{x}| = |a||\overline{x}|$, as claimed.

NOW we prove (c'). If $\overline{x}\|\overline{y}$ then $\overline{x} = t\overline{y}$ or $\overline{y} = t\overline{x}$; so $|\overline{x} \cdot \overline{y}| = |\overline{x}||\overline{y}|$ follows by (b'). (Verify!)

Otherwise, $\overline{x} \neq t\overline{y}$ and $\overline{y} \neq t\overline{x}$ for all $t \in E^1$. Then we obtain, for all $t \in E^1$

$$0 \neq |t\overline{x}-\overline{y}|^2 = \sum_{k=1}^{n}(tx_k - y_k)^2 = t^2\sum_{k=1}^{n} x_k^2 - 2t\sum_{k=1}^{n} x_k y_k + \sum_{k=1}^{n} y_k^2. \tag{3.1.27}$$

Thus, setting

$$A = \sum_{k=1}^{n} x_k^2, B = 2\sum_{k=1}^{n} x_k y_k, \text{ and } C = \sum_{k=1}^{n} y_k^2, \tag{3.1.28}$$

we see that the quadratic equation

$$0 = At^2 - Bt + C \tag{3.1.29}$$

has *no* real solutions in t, so its discriminant, $B^2 - 4AC$, must be negative; i.e.,

$$4\left(\sum_{k=1}^{n} x_k y_k\right)^2 - 4\left(\sum_{k=1}^{n} x_k^2\right)\left(\sum_{k=1}^{n} y_k^2\right) < 0, \tag{3.1.30}$$

proving (c').

To prove (d'), use Definition 2 and Theorem 3(d), to obtain

$$|\overline{x}+\overline{y}|^{2}=(\overline{x}+\overline{y})\cdot(\overline{x}+\overline{y})=\overline{x}\cdot\overline{x}+\overline{y}\cdot\overline{y}+2\overline{x}\cdot\overline{y}=|\overline{x}|^{2}+|\overline{y}|^{2}+2\overline{x}\cdot\overline{y}. \tag{3.1.31}$$

But $\overline{x}\cdot\overline{y}\leq|\overline{x}||\overline{y}|$ by (c'). Thus we have

$$|\overline{x}+\overline{y}|^{2}\leq|\overline{x}|^{2}+|\overline{y}|^{2}+2|\overline{x}||\overline{y}|=(|\overline{x}|+|\overline{y}|)^{2}, \tag{3.1.32}$$

whence $|\overline{x}+\overline{y}|\leq|\overline{x}|+|\overline{y}|,$ as required.

Finally, replacing here $\overline{x}$ by $\overline{x}-\overline{y},$ we have

$$|\overline{x}-\overline{y}|+|\overline{y}|\geq|\overline{x}-\overline{y}+\overline{y}|=|\overline{x}|,\text{ or }|\overline{x}-\overline{y}|\geq|\overline{x}|-|\overline{y}|, \tag{3.1.33}$$

Similarly, replacing $\overline{y}$ by $\overline{y}-\overline{x},$ we get $|\overline{x}-\overline{y}|-|\overline{y}|-|\overline{x}|.$ Hence

$$|\overline{x}-\overline{y}|\geq\pm(|\overline{x}|-|\overline{y}|), \tag{3.1.34}$$

i.e., $|\overline{x}-\overline{y}|\geq||\overline{x}|-|\overline{y}||,$ proving the second formula in (d'). $square$

Theorem 3.1.5

For any points $\overline{x},\overline{y},$ and $\overline{z}\in E^{n},$ we have

(i) $\rho(\overline{x},\overline{y})\geq 0$, and $\rho(\overline{x},\overline{y})=0$ iff $\overline{x}=\overline{y}$;

(ii) $\rho(\overline{x},\overline{y})=\rho(\overline{y},\overline{x})$;

(iii) $\rho(\overline{x},\overline{z})\leq\rho(\overline{x},\overline{y})+\rho(\overline{y},\overline{z})(\text{ triangle inequality});$

Proof

(i) By Definition 3 and Note 3, $\rho(\overline{x},\overline{y})=|\overline{x}-\overline{y}|;$ therefore, by Theorem $4(\mathrm{a}')$, $\rho(\overline{x},\overline{y})=|\overline{x}-\overline{y}|\geq 0$.

Also, $|\overline{x}-\overline{y}|>0$ iff $\overline{x}-\overline{y}\neq 0,$ i.e., iff $\overline{x}\neq\overline{y}.$ Hence $\rho(\overline{x},\overline{y})\neq 0$ iff $\overline{x}\neq\overline{y},$ and assertion (i) follows.

(ii) By Theorem 4 $\left(\mathrm{b}'\right)$, $|\overline{x}-\overline{y}|=|(-1)(\overline{y}-\overline{x})|=|\overline{y}-\overline{x}|,$ so (ii) follows.

(iii) By Theorem $4\left(\mathrm{d}'\right)$,

$\rho(\overline{x},\overline{y})+\rho(\overline{y},\overline{z})=|\overline{x}-\overline{y}|+|\overline{y}-\overline{z}|\geq|\overline{x}-\overline{y}+\overline{y}-\overline{z}|=\rho(\overline{x},\overline{z}).\square$

Note 8. We also have $|\rho(\overline{x},\overline{y})-\rho(\overline{z},\overline{y})|\leq\rho(\overline{x},\overline{z}).$ (Prove it!) The two triangle inequalities have a simple geometric interpretation (which explains their name). If $\overline{x},\overline{y},$ and $\overline{z}$ are treated as the vertices of a triangle, we obtain that the length of a side, $\rho(\overline{x},\overline{z})$ never exceeds the sum of the two other sides and is never less their difference.

As E^{1} is a special case of E^{n} (in which "vectors" are single numbers), all our theory applies to E^{1} as well. In particular, distances in E^{1} are defined by $\rho(x,y)=|x-y|$ and obey the three laws of Theorem 5. Dot products in E^{1} become ordinary products xy. (Why?) From Theorems 4 $\left(\mathrm{b}'\right)$ $\left(\mathrm{d}'\right)$, we have

$$|a||x|=|ax|;|x+y|\leq|x|+|y|;|x-y|\geq||x|-|y||\quad\left(a,x,y\in E^{1}\right). \tag{3.1.35}$$

3.1.E: Problems on Vectors in E n En (Exercises)

Exercise 3.1.E.1

Prove by induction on n that

$$\left(x_{1}, x_{2}, \ldots, x_{n}\right)=\left(y_{1}, y_{2}, \ldots, y_{n}\right) \text{ iff } x_{k}=y_{k}, k=1,2, \ldots, n. \tag{3.1.E.1}$$

[Hint: Use Problem 6(ii) of Chapter 1, §§1-3, and Example (i) in Chapter 2, §§5-6.]

Exercise 3.1.E.2

Complete the proofs of Theorems 1 and 3 and Notes 3 and 8.

Exercise 3.1.E.3

Given $\overline{x}=(-1,2,0,-7), \overline{y}=(0,0,-1,-2)$, and $\overline{z}=(2,4,-3,-3)$ in E^{4}, express $\overline{x}, \overline{y}$, and $\overline{z}$ as linear combinations of the basic unit vectors. Also, compute their absolute values, their inverses, as well as their mutual sums, differences, dot products, and distances. Are any of them orthogonal? Parallel?

Exercise 3.1.E.4

With $\overline{x}, \overline{y}$, and $\overline{z}$ as in Problem 3, find scalars a, b, and c such that

$$a \overline{x}+b \overline{y}+c \overline{z}=\overline{u}, \tag{3.1.E.2}$$

when

$$\begin{array}{llll}\text{(i)} \overline{u} & =\overline{e}_{1}; & \text{(ii)} \ \overline{u}=\overline{e}_{3}; \\ \text{(iii)} \ \overline{u} & =(-2,4,0,1); & \text{(iv)} \ \overline{u}=\overline{0}.\end{array} \tag{3.1.E.3}$$

Exercise 3.1.E.5

A finite set of vectors $\overline{x}, \overline{x}_{2}, \ldots, \overline{x}_{m}$ is said to be dependent iff there are scalars $a_{1}, \ldots, a_{m}$, not all zero, such that

$$\sum_{k=1}^{m} a_{k} \overline{x}_{k}=\overline{0}, \tag{3.1.E.4}$$

and independent otherwise. Prove the independence of the following sets of vectors:
(a) $\overline{e}_{1}, \overline{e}_{2}, \ldots, \overline{e}_{n}$ in E^{n};
(b) $(1,2,-3,4)$ and $(2,3,0,0)$ in E^{4};
(c) $(2,0,0),(4,-1,3)$,and $(0,4,1)$ in E^{3};
(d) the vectors $\overline{x}, \overline{y}$, and $\overline{z}$ of Problem 3.

Exercise 3.1.E.6

Prove (for E^{2} and E^{3}) that

$$\overline{x} \cdot \overline{y}=|\overline{x}||\overline{y}| \cos \alpha, \tag{3.1.E.5}$$

where α is the angle between the vectors $\overrightarrow{0x}$ and $\overrightarrow{0y}$; we denote α by $\langle\overline{x}, \overline{y}\rangle$.

[Hint: Consider the triangle $\overline{0}\overline{x}\overline{y}$, with sides $\overline{x}=\overrightarrow{0x}, \overline{y}=\overrightarrow{0y}$, and $\overrightarrow{xy}=\vec{y}-\vec{x}$ (see Definition 7). By the law of cosines,

$$|\vec{x}|^{2}+|\vec{y}|^{2}-2|\vec{x}||\vec{y}|\cos\alpha=|\vec{y}-\vec{x}|^{2}. \tag{3.1.E.6}$$

Now substitute $|\vec{x}|^{2}=\vec{x}\cdot\vec{x}, |\vec{y}|^{2}=\vec{y}\cdot\vec{y}$, and

$$|\vec{y}-\vec{x}|^{2}=(\vec{y}-\vec{x})\cdot(\vec{y}-\vec{x})=\vec{y}\cdot\vec{y}+\vec{x}\cdot\vec{x}-2\vec{x}\cdot\vec{y}.\text{ (Why?)} \tag{3.1.E.7}$$

Then simplify.]

Exercise $3.1.E.7$

Motivated by Problem 6, define in E^n

$$\langle\overline{x},\overline{y}\rangle=\arccos\frac{\overline{x}\cdot\overline{y}}{|\overline{x}||\overline{y}|}\text{ if }\overline{x}\text{ and }\overline{y}\text{ are nonzero.} \tag{3.1.E.8}$$

(Why does an angle with such a cosine exist?) Prove that
(i) $\overline{x}\perp\overline{y}$ iff $\cos\langle\overline{x},\overline{y}\rangle=0$, i.e., $\langle\overline{x},\overline{y}\rangle=\frac{\pi}{2}$;
(ii) $\sum_{k=1}^{n}\cos^{2}\langle\overline{x},\overline{e}_{k}\rangle=1$.

Exercise $3.1.E.8$

Continuing Problems 3 and 7, find the cosines of the angles between the sides, $\overrightarrow{xy}$, $\overrightarrow{yz}$, and $\overrightarrow{zx}$ of the triangle $\overline{x}\overline{y}\overline{z}$, with $\overline{x},\overline{y}$, and $\overline{z}$ as in Problem 3.

Exercise $3.1.E.9$

Find a unit vector in E^4, with positive components, that forms equal angles with the axes, i.e., with the basic unit vectors (see Problem 7).

Exercise $3.1.E.10$

Prove for E^n that if $\overline{u}$ is orthogonal to each of the basic unit vectors $\overline{e}_1,\overline{e}_2,\ldots,\overline{e}_n$, then $\overline{u}=\overline{0}$. Deduce that

$$\overline{u}=\overline{0}\text{ iff }\left(\forall\overline{x}\in E^{n}\right)\overline{x}\cdot\overline{u}=0. \tag{3.1.E.9}$$

Exercise $3.1.E.11$

Prove that $\overline{x}$ and $\overline{y}$ are parallel iff

$$\frac{x_1}{y_1}=\frac{x_2}{y_2}=\cdots=\frac{x_n}{y_n}=c \quad \left(c\in E^{1}\right), \tag{3.1.E.10}$$

where " $x_k/y_k=$" c is to be replaced by " $x_k=0$ " if $y_k=0$.

Exercise $3.1.E.12$

Use induction on n to prove the Lagrange identity (valid in any field),

$$\left(\sum_{k=1}^{n}x_k^2\right)\left(\sum_{k=1}^{n}y_k^2\right)-\left(\sum_{k=1}^{n}x_ky_k\right)^{2}=\sum_{1\leq i<k\leq n}\left(x_iy_k-x_ky_i\right)^{2}. \tag{3.1.E.11}$$

Hence find a new proof of Theorem $4(c')$.

Exercise $3.1.E.13$

Use Problem 7 and Theorem $4(c')$("equality") to show that two nonzero vectors $\bar{x}$ and $\bar{y}$ in E^n are parallel iff $\cos\langle\bar{x}, \bar{y}\rangle = \pm 1$.

Exercise $3.1.E.14$

(i) Prove that $|\bar{x}+\bar{y}| = |\bar{x}| + |\bar{y}| + |\bar{y}|$ iff $\bar{x} = t\bar{y}$ or $\bar{y} = t\bar{x}$ for some $t \geq 0$; equivalently, iff $\cos\langle\bar{x}, \bar{y}\rangle = 1$ (see Problem 7).
(ii) Find similar conditions for $|\bar{x}-\bar{y}| = |\bar{x}| + |\bar{y}|$.
[Hint: Look at the proof of Theorem $4(\mathrm{d}')$.]

3.2: Lines and Planes in E^n

To obtain a line in E^2 or E^3 passing through two points $\bar{a}$ and $\bar{b}$, we take the vector

$$\vec{u}=\overrightarrow{ab}=\bar{b}-\bar{a} \tag{3.2.1}$$

and, so to say, "stretch" it indefinitely in both directions, i.e., multiply $\vec{u}$ by all possible scalars $t \in E^1$. Then the set of all points $\bar{x}$ of the form

$$\bar{x}=\bar{a}+t\vec{u} \tag{3.2.2}$$

is the required line. It is natural to adopt this as a definition in E^n as well.

Below, $\bar{a} \neq \bar{b}$.

Definition: parametric equation of the line

The line $\overline{ab}$ through the points $\bar{a}, \bar{b} \in E^n$ (also called the line through $\bar{a}$, in the direction of the vector $\vec{u}=\bar{b}-\bar{a}$) is the set of all points $\bar{x} \in E^n$ of the form

$$\bar{x}=\bar{a}+t\vec{u}=\bar{a}+t(\bar{b}-\bar{a}), \tag{3.2.3}$$

where t varies over E^1. We call t a variable real parameter and $\vec{u}$ a direction vector for $\overline{ab}$. Thus

$$\text{Line } \overline{ab}=\left\{\bar{x} \in E^n | \bar{x}=\bar{a}+t\vec{u} \text{ for some } t \in E^1\right\}, \quad \vec{u}=\bar{b}-\bar{a} \neq \overline{0}. \tag{3.2.4}$$

The formula

$$\bar{x}=\bar{a}+t\vec{u}, \text{ or } \bar{x}=\bar{a}+t(\bar{b}-\bar{a}), \tag{3.2.5}$$

is called the **parametric equation of the line**. (We briefly say "the line $\bar{x}=\bar{a}+t\vec{u}.$") It is equivalent to n simultaneous equations in terms of coordinates, namely,

$$x_k=a_k+tu_k=a_k+t\left(b_k-a_k\right), \quad k=1,2,\ldots,n. \tag{3.2.6}$$

Note 1. As the vector $\vec{u}$ is anyway being multiplied by all real numbers t, the line (as a set of points) will not change if $\vec{u}$ is replaced by some c \vec{u}\left(c \in E^{1}\right, $c \neq 0)$. In particular, taking $c=1/|\vec{u}|$, we may replace $\vec{u}$ by $\vec{u}/|\vec{u}|$, $\quad a$ unit vector.. We may as well assume that $\vec{u}$ is a unit vector itself.

If we let t vary not over all of $E^{\perp}$ but only over some interval in $E^{\perp}$, we obtain what is called a line segment. In particular, we define the open line segment $L(\bar{a}, \bar{b})$, the closed line segment $L[\bar{a}, \bar{b}]$, the half-open line segment $L(\bar{a}, \bar{b}]$, and the half-closed line segment $L[\bar{a}, \bar{b})$, as we did for E^1.

Definition: endpoints of the segment

Given $\vec{u}=\bar{b}-\bar{a}$, we set

$$\begin{array}{ll}\text{(i) } L(\bar{a}, \bar{b})=\{\bar{a}+t\vec{u} | 0<t<1\}; & \text{(ii) } L[\bar{a}, \bar{b}]=\{\bar{a}+t\vec{u} | 0 \leq t \leq 1\} \\ \text{(iii) } L(\bar{a}, \bar{b}]=\{\bar{a}+t\vec{u} | 0<t \leq 1\}; & \text{(iv) } L[\bar{a}, \bar{b})=\{\bar{a}+t\vec{u} | 0 \leq t<1\}\end{array}$$

In all cases, $\bar{a}$ and $\bar{b}$ are called the **endpoints of the segment**; $\rho(\bar{a}, \bar{b})=|\bar{b}-\bar{a}|$ is its length; and $\frac{1}{2}(\bar{a}+\bar{b})$ is its midpoint.

Note that in E^1, line segments simply become intervals, (a, b), $[a, b]$, etc.

To describe a plane in E^3, we fix one of its points, $\bar{a}$, and a vector $\vec{u}=\overrightarrow{ab}$ perpendicular to the plane (imagine a vertical pencil standing at $\bar{a}$ on the horizontal plane of the table). Then a point $\bar{x}$ lies on the plane iff $\vec{u} \perp \overrightarrow{ax}$. It is natural to accept this as a definition in E^n as well.

Definition: plane

Given a point $\overline{a} \in E^n$ and a vector $u \neq \overrightarrow{0}$, we define the **plane** (also called hyperplane if $n > 3$) through $\overline{a}$, orthogonal to $\vec{u}$, to be the set of all $\overline{x} \in E^n$ such that $\vec{u} \perp \overrightarrow{ax}$, i.e., $\vec{u} \cdot (\overline{x} - \overline{a}) = 0$, or, in terms of components,

$$\sum_{k=1}^{n} u_k (x_k - a_k) = 0, \text{ where } \vec{u} \neq \overrightarrow{0} \text{ (i.e., not all values } u_k \text{ are } 0). \tag{3.2.7}$$

We briefly say

$$\text{"the plane } \vec{u} \cdot (\overline{x} - \overline{a}) = 0\text{" or "the plane } \sum_{k=1}^{n} u_k (x_k - a_k) = 0\text{"} \tag{3.2.8}$$

(this being the equation of the plane). Removing brackets in (3), we have

$$u_1 x_2 + u_2 x_2 + \cdots + u_n x_n = c, \text{ or } \vec{u} \cdot \overline{x} = c, \text{ where } c = \sum_{k-1}^{n} u_k a_k, \vec{u} \neq \overrightarrow{0}. \tag{3.2.9}$$

An equation of this form is said to be linear in $x_1, x_2, \ldots, x_n$.

Theorem 3.2.1

A set $A \subseteq E^n$ is a plane (hyperplane) iff A is exactly the set of all $\overline{x} \in E^n$ satisfying (4) for some fixed $c \in E^1$ and $\vec{u} = (u_1, \ldots, u_n) \neq \overline{0}$.

Proof

Add proof here and it will automatically be hidden Indeed, as we saw above, each plane has an equation of the form (4).

Conversely, any equation of that form (with, say, $u_1 \neq 0$) can be written as

$$u_1 \left(x_1 - \frac{c}{u_1}\right) + u_2 x_2 + u_3 x_3 + \cdots + u_n x_n = 0. \tag{3.2.10}$$

Then, setting $a_1 = c/u_1$ and $a_k = 0$ for $k \geq 2$, we transform it into (3), which is, by definition, the equation of a plane through $\overline{a} = (c/u_1, 0, \ldots, 0)$, orthogonal to $u = (u_1, \ldots, u_n)$. $\square$

Thus, briefly, planes are exactly all sets with linear equations (4). In this connection, equation (4) is called the general equation of a plane. The vector $\vec{u}$ is said to be normal to the plane. **Clearly, if both sides of (4) are multiplied by a nonzero scalar** q, one obtains an equivalent equation (representing the same set). Thus we may replace u_k by qu_k, i.e., $\vec{u}$ by $q\vec{u}$, without affecting the plane. In particular, we may replace $\vec{u}$ by the unit vector $\vec{u}/|\vec{u}|$, as in lines is called the normalization of the equation). Thus

$$\frac{\vec{u}}{|\vec{u}|} \cdot (\overline{x} - \overline{a}) = 0 \tag{3.2.11}$$

and

$$\overline{x} = \overline{a} + t\frac{\vec{u}}{|\vec{u}|} \tag{3.2.12}$$

are the normalized (or normal) equations of the plane (3) and line (1), respectively.

Note 2. The equation $x_k = c$ (for a fixed k) represents a plane orthogonal to the basic unit vector $\vec{e}_k$ or, as we shall say, to the kth axis. **The equation results from (4)** if we take $\vec{u} = \vec{e}_k$ so that $u_k = 1$, while $u_i = 0$ $(i \neq k)$. For example, $x_1 = c$ is the equation of a plane orthogonal to $\vec{e}_1$; it consists of all $\overline{x} \in E^n$, with $x_1 = c$ (while the other coordinates of $\overline{x}$ are arbitrary). In E^2, it is a line. In E^1, it consists of c alone.

Two planes (respectively, two lines) are said to be perpendicular to each other iff their normal vectors (respectively, direction vectors) are orthogonal; similarly for parallelism. A plane $\vec{u} \cdot \bar{x} = c$ is said to be perpendicular to a line $\bar{x} = \bar{a} + t\vec{v}$ iff $\vec{u} \| \vec{v}$; the line and the plane are parallel iff $\vec{u} \perp \vec{v}$.

Note 3. When normalizing, as in (5) or (6), we actually have two choices of a unit vector, namely, $\pm\vec{u}/|\vec{u}|$. If one of them is prescribed, we speak of a directed plane (respectively, line).

Example 3.2.1

(a) Let $\bar{a} = (0, -1, 2), \bar{b} = (1, 1, 1)$, and $\bar{c} = (3, 1, -1)$ in E^3. Then the line $\overline{ab}$ has the parametric equation $\bar{x} = \bar{a} + t(\bar{b} - \bar{a})$ or, in coordinates, writing x, y, z for x_1, x_2, x_3,

$$x = 0 + t(1 - 0) = t, y = -1 + 2t, z = 2 - t. \tag{3.2.13}$$

This may be rewritten

$$t = \frac{x}{1} = \frac{y+1}{2} = \frac{z-2}{-1}, \tag{3.2.14}$$

where $\vec{u} = (1, 2, -1)$ is the direction vector (composed of the denominators. Normalizing and dropping t, we have

$$\frac{x}{1/\sqrt{6}} = \frac{y+1}{2/\sqrt{6}} = \frac{z-2}{-1/\sqrt{6}} \tag{3.2.15}$$

(the so-called symmetric form of the normal equations).

Similarly, for the line $\overline{bc}$, we obtain

$$t = \frac{x-1}{2} = \frac{y-1}{0} = \frac{z-1}{-2}, \tag{3.2.16}$$

where "$t = (y-1)/0$ " stands for "$y - 1 = 0$." (It is customary to use this notation.)

(b) Let $\bar{a} = (1, -2, 0, 3)$ and $\vec{u} = (1, 1, 1, 1)$ in E^4. Then the plane normal to $\vec{u}$ through $\bar{a}$ has the equation $(\bar{x} - \bar{a}) \cdot \vec{u} = 0$, or

$$(x_1 - 1) \cdot 1 + (x_2 + 2) \cdot 1 + (x_3 - 0) \cdot 1 + (x_4 - 3) \cdot 1 = 0, \tag{3.2.17}$$

or $x_1 + x_2 + x_3 + x_4 = 2$. **Observe that, by formula (4),** the coefficients of x_1, x_2, x_3, x_4 are the components of the normal vector $\vec{u}$ (here $(1, 1, 1, 1)$).

Now define a map $f : E^4 \to E^1$ setting $f(\bar{x}) = x_1 + x_2 + x_3 + x_4$ (the left-hand side of the equation). This map is called the linear functional corresponding to the given plane. **(For another approach, see Problems 4-6 below.)**

(c) The equation $x + 3y - 2z = 1$ represents a plane in E^3, with $\vec{u} = (1, 3, -2)$. The point $\bar{a} = (1, 0, 0)$ lies on the plane (why?), so the plane equation may be written $(\bar{x} - \bar{a}) \cdot \vec{u} = 0$ or $\bar{x} \cdot \vec{u} = 1$, where $\bar{x} = (x, y, z)$ and $\bar{a}$ and $\vec{u}$ are as above.

3.2.E: Problems on Lines and Planes in E n En (Exercises)

Exercise $3.2.E.1$

Let $\bar{a}=(-1,2,0,-7), \bar{b}=(0,0,-1,2)$, and $\bar{c}=(2,4,-3,-3)$ be points in E^4. Find the symmetric normal equations (see Example (a)) of the lines $\overline{ab}, \overline{bc}$, and $\overline{ca}$. Are any two of the lines perpendicular? Parallel? On the line $\overline{ab}$, find some points inside $L(\bar{a}, \bar{b})$ and some outside $L[\bar{a}, \bar{b}]$. Also, find the symmetric equations of the line through $\bar{c}$ that is

$$\text{(i) parallel to } \overline{ab}; \quad \text{(ii) perpendicular to } \overline{ab}. \tag{3.2.E.1}$$

Exercise $3.2.E.2$

With $\bar{a}$ and $\bar{b}$ as in Problem 1, find the equations of the two planes that trisect, and are perpendicular to, the line segment $L[\bar{a}, \bar{b}]$.

Exercise $3.2.E.3$

Given a line $\bar{x}=\bar{a}+t\vec{u}(\vec{u}=\bar{b}-\bar{a} \neq \overrightarrow{0})$ in E^n, define $f: E^1 \rightarrow E^n$ by

$$f(t)=\bar{a}+t\vec{u} \text{ for } t \in E^1. \tag{3.2.E.2}$$

Show that $L[\bar{a}, \bar{b}]$ is exactly the f -image of the interval $[0,1]$ in E^1, with $f(0)=a$ and $f(1)=b$, while $f\left[E^1\right]$ is the entire line. Also show that f is one to one.
$[\text{ Hint: } t \neq t^{\prime} \text{ implies } |f(t)-f(t^{\prime})| \neq 0. \text{ Why? }]$

Exercise $3.2.E.4$

A map $f: E^n \rightarrow E^1$ is called a linear functional iff

$$(\forall \bar{x}, \bar{y} \in E^n)\left(\forall a, b \in E^1\right) \quad f(a\bar{x}+b\bar{y})=af(\bar{x})+bf(\bar{y}). \tag{3.2.E.3}$$

Show by induction that f preserves linear combinations; that is,

$$f\left(\sum_{k=1}^m a_k \bar{x}_k\right)=\sum_{k=1}^m a_k f\left(\bar{x}_k\right) \tag{3.2.E.4}$$

for any $a_k \in E^1$ and $\bar{x}_k \in E^n$.

Exercise $3.2.E.5$

From Problem 4 prove that a map $f: E^n \rightarrow E^1$ is a linear functional iff there is $\vec{u} \in E^n$ such that

$$(\forall \bar{x} \in E^n) f(\bar{x})=\vec{u} \cdot \bar{x}(\text{" representation theorem "}). \tag{3.2.E.5}$$

[Hint: If f is a linear functional, write each $\bar{x} \in E^n$ as $\bar{x}=\sum_{k=1}^n x_k \bar{e}_k (\S\S 1-3, Theorem 2). Then \[f(\bar{x})=f\left(\sum_{k=1}^m x_k \bar{e}_k\right)=\sum_{k=1}^n x_k f\left(\bar{e}_k\right). \] Setting \(u_k=f\left(\bar{e}_k\right)$ and $\in E^1$
$\vec{u}=(u_1, \ldots, u_n)$, obtain $f(\bar{x})=\vec{u} \cdot \bar{x}$, as required. For the converse, use Theorem 3 in §§1-3.]

Exercise $3.2.E.6$

Prove that a set $A \subseteq E^n$ is a plane iff there is a linear functional f (Problem 4), not identically zero, and some $c \in E^1$ such that

$$A=\{\overline{x} \in E^n | f(\overline{x})=c\}. \tag{3.2.E.6}$$

(This could serve as a definition of planes in E^n.)

[Hint: A is a plane iff $A=\{\overline{x} | \vec{u} \cdot \overline{x}=c\}$. Put $f(\overline{x})=\vec{u} \cdot \overline{x}$ and use Problem 5. Show that $f \neq 0$ iff $\vec{u} \neq \overrightarrow{0}$ by Problem 10 of §§1-3.]

Exercise $3.2.E.7$

Prove that the perpendicular distance of a point $\overline{p}$ to a plane $\vec{u} \cdot \overline{x}=c$ in E^n is

$$\rho(\overline{p}, \overline{x}_0)=\frac{|\vec{u} \cdot \overline{p}-c|}{|\vec{u}|}. \tag{3.2.E.7}$$

($\overline{x}_0$ is the orthogonal projection of $\overline{p}$, i.e., the point on the plane such that $\overrightarrow{p x_0} \| \vec{u}$.)
[Hint: Put $\vec{v}=\vec{u}/|\vec{u}|$. Consider the line $\overline{x}=\overline{p}+t\vec{v}$. Find t for which $\overline{p}+t\vec{v}$ lies on both the line and plane. Find $|t|$.]

Exercise $3.2.E.8$

A globe (solid sphere) in E^n, with center $\overline{p}$ and radius $\varepsilon>0$, is the set $\{\overline{x} | \rho(\overline{x}, \overline{p})<\varepsilon\}$, denoted $G_{\overline{p}}(\varepsilon)$. Prove that if $\overline{a}, \overline{b} \in G_{\overline{p}}(\varepsilon)$, then also $L[\overline{a}, \overline{b}] \subseteq G_{\tilde{p}}(\varepsilon)$. Disprove it for the sphere $S_{\overline{p}}(\varepsilon)=\{\overline{x} | \rho(\overline{x}, \overline{p})=\varepsilon\}$. [Hint: Take a line through $\overline{p}$.]

3.3: Intervals in Eⁿ

Consider the rectangle in E^2 shown in Figure 2. Its interior (without the perimeter consists of all points $(x, y) \in E^2$ such that

$$a_1 < x < b_1 \text{ and } a_2 < y < b_2; \tag{3.3.1}$$

i.e.,

$$x \in (a_1, b_1) \text{ and } y \in (a_2, b_2). \tag{3.3.2}$$

Thus it is the Cartesian product of two line intervals, (a_1, b_1) and (a_2, b_2). To include also all or some sides, we would have to replace open intervals by closed, half-closed, or half-open ones. Similarly, Cartesian products of three line intervals yield rectangular parallelepipeds in E^3. We call such sets in E^n intervals.

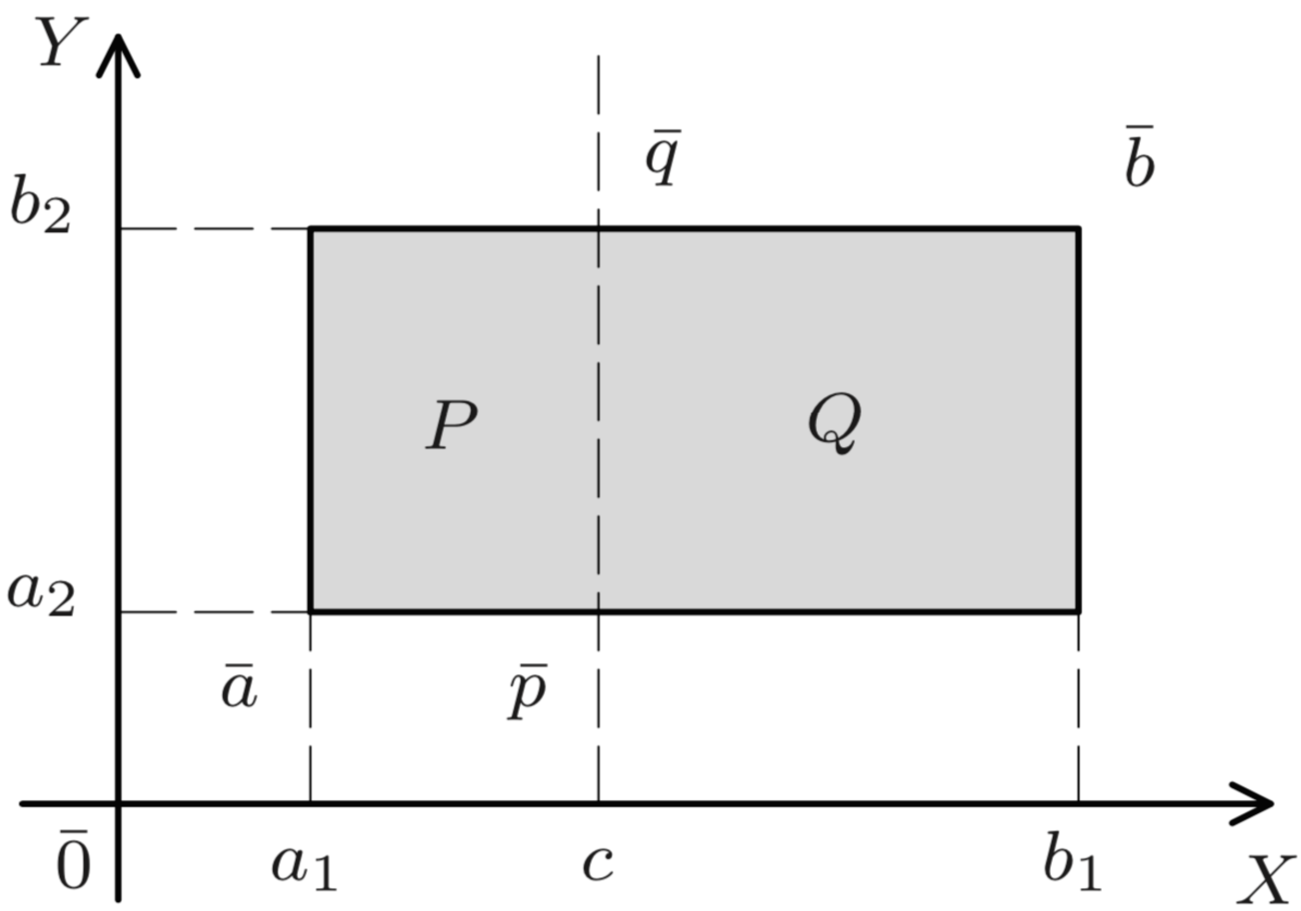

FIGURE 2

Definition

1. By an interval in E^n we mean the Cartesian product of any n intervals in E^1 (some may be open, some closed or half-open, etc.).

2. In particular, given

$$\bar{a} = (a_1, \ldots, a_n) \text{ and } \bar{b} = (b_1, \ldots, b_n) \tag{3.3.3}$$

with

$$a_k \leq b_k, \quad k=1,2,\ldots,n, \tag{3.3.4}$$

we define the open interval $(\bar{a}, \bar{b})$, the closed interval $[\bar{a}, \bar{b}]$, the half-open interval $(\bar{a}, \bar{b}]$, and the half-closed interval $[\bar{a}, \bar{b})$ as follows:

$$\begin{aligned} (\bar{a}, \bar{b}) &= \{\bar{x} \mid a_k < x_k < b_k, k=1,2,\ldots,n\} \\ &= (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n) \\ [\bar{a}, \bar{b}] &= \{\bar{x} \mid a_k \leq x_k \leq b_k, k=1,2,\ldots,n\} \\ &= [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \\ (\bar{a}, \bar{b}] &= \{\bar{x} \mid a_k < x_k \leq b_k, k=1,2,\ldots,n\} \\ &= (a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n] \\ [a, b) &= \{\bar{x} \mid a_k \leq x_k < b_k, k=1,2,\ldots,n\} \\ &= [a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_n, b_n) \end{aligned}$$

In all cases, $\bar{a}$ and $\bar{b}$ are called the endpoints of the interval. Their distance

$$\rho(\bar{a}, \bar{b}) = |\bar{b} - \bar{a}| \tag{3.3.5}$$

is called its diagonal. The n differences

$$b_k - a_k = \ell_k \quad (k=1,\ldots,n) \tag{3.3.6}$$

are called its n edge-lengths. Their product

$$\prod_{k=1}^{n} \ell_k = \prod_{k=1}^{n} (b_k - a_k) \tag{3.3.7}$$

is called the volume of the interval (in E^2 it is its area, in E^1 its length) .\) The point

$$\bar{c} = \frac{1}{2}(\bar{a} + \bar{b}) \tag{3.3.8}$$

is called its center or midpoint. The set difference

$$[\bar{a}, \bar{b}] - (\bar{a}, \bar{b}) \tag{3.3.9}$$

is called the boundary of any interval with endpoints $\bar{a}$ and $\vec{b}$; it consists of $2n$ "faces" defined in a natural manner. (How?)

We often denote intervals by single letters, e.g.. $A = (\bar{a}, \bar{b})$, and write dA for "diagonal of A'' and vA or vol A for "volume of $A.$" If all edge-lengths $b_k - a_k$ are equal, A is called a cube (in E^2, a square). The interval A is said to be degenerate iff $b_k = a_k$ for some k, in which case, clearly,

$$\operatorname{vol} A = \prod_{k=1}^{n} (b_k - a_k) = 0. \tag{3.3.10}$$

Note 1. We have $\bar{x} \in (\bar{a}, \bar{b})$ iff the inequalities $a_k < x_k < b_k$ hold simultaneously for all k. This is impossible if $a_k = b_k$ for some k; similarly for the inequalities $a_k < x_k \leq b_k$ or $a_k \leq x_k < b_k$. Thus a degenerate interval is empty, unless it is closed (in which case it contains $\bar{a}$ and $\bar{b}$ at least).

Note 2. In any interval A,

$$dA = \rho(\bar{a}, \bar{b}) = \sqrt{\sum_{k=1}^{n} (b_k - a_k)^2} = \sqrt{\sum_{k=1}^{n} \ell_k^2}. \tag{3.3.11}$$

In E^2, we can split an interval A into two subintervals P and Q by drawing a line (see Figure 2). In E^3, this is done by a plane orthogonal to one of the axes of the form x_{k}=c\left(see §§4-6, Note 2), with $a_k < c < b_k$. In particular, if \right.

$c=\frac{1}{2}(a_k+b_k)$, the plane bisects the k th edge of A; and so the k th edge-length of $P($ and $Q)$ equals $\frac{1}{2}\ell_k=\frac{1}{2}(b_k-a_k)$. If A is closed, so is P or Q, depending on our choice. (We may include the "partition" $x_k=c$ in P or Q.)[1]

Now, successively draw n planes $x_k=c_k, \quad c_k=\frac{1}{2}(a_k+b_k), \quad k=1,2,\ldots,n.$ The first plane bisects ℓ_j leaving the other edges of Aun$-$ changed. The resulting two subintervals P and Q then are cut by the plane $x_2=c_2$, bisecting the second edge in each of them. Thus we get four subintervals (see Figure 3 for E^2. Each successive plane doubles the number of subintervals. After n steps, we thus obtain 2^n disjoint intervals, with all edges ℓ_k bisected. Thus by Note 2, the diagonal of each of them is

$$\sqrt{\sum_{k=1}^{n}\left(\frac{1}{2}\ell_k\right)^2}=\frac{1}{2}\sqrt{\sum_{k=1}^{n}\ell_k^2}=\frac{1}{2}dA. \tag{3.3.12}$$

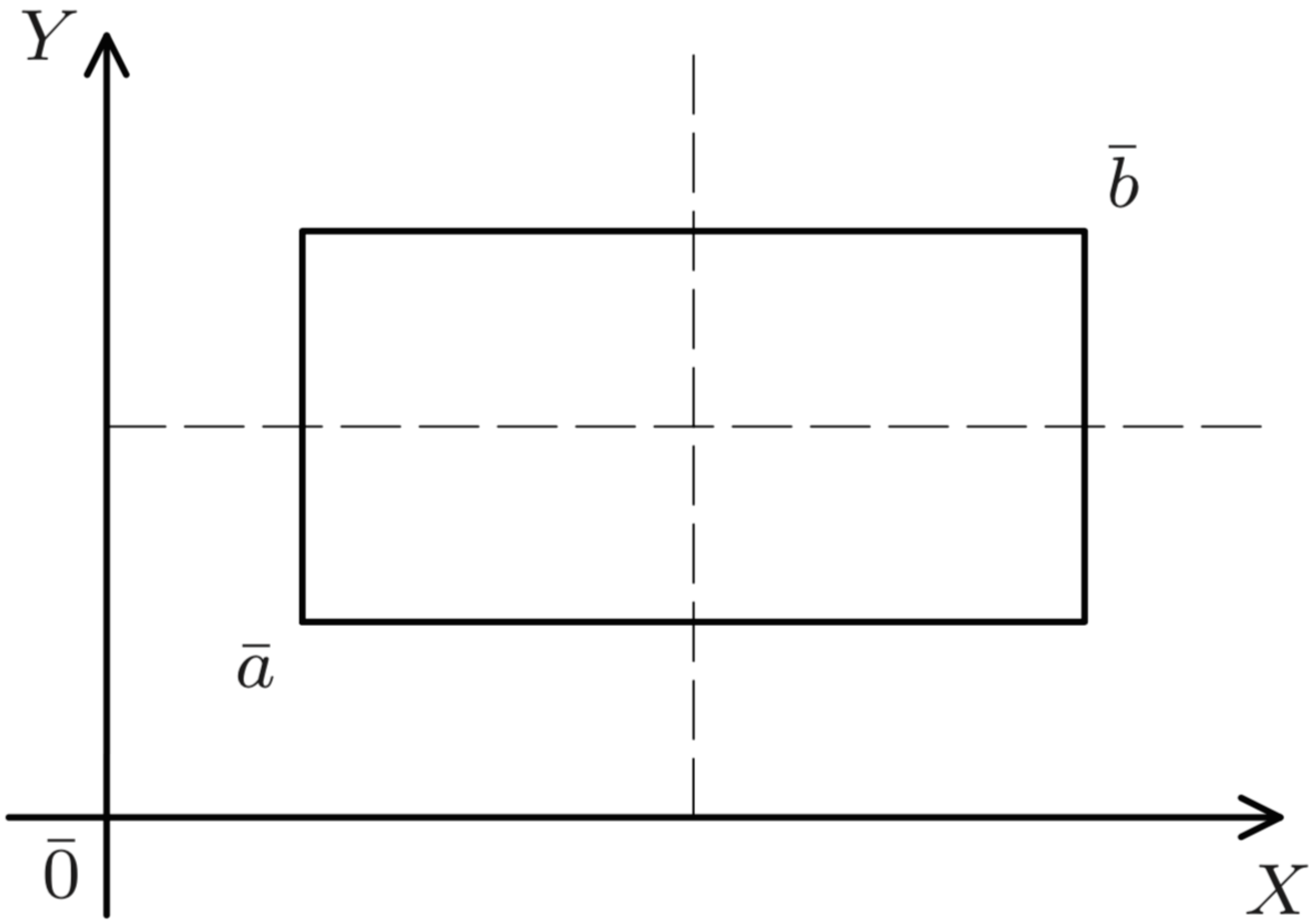

FIGURE 3

Note 3. If A is closed then, as noted above, we can make any one (but only one) of the 2^n subintervals closed by properly manipulating each step.

The proof of the following simple corollaries is left to the reader.

Corollary 3.3.1

No distance between two points of an interval A exceeds dA, its diagonal. That is, $(\forall \bar{x}, \bar{y} \in A)\rho(\bar{x}, \bar{y}) \leq dA$

Corollary 3.3.2

If an interval A contains $\overline{p}$ and $\overline{q}$, then also $L[\overline{p}, \overline{q}] \subseteq A$.

corollary 3.3.3

Every nondegenerate interval in E^n contains rational points, i.e., points whose coordinates are all rational.

(Hint: Use the density of rationals in E^1 for each coordinate separately.)

3.3.E: Problems on Intervals in E^n (Exercises)

(Here A and B denote intervals.)

Exercise $3.3.E.1$

Prove Corollaries 1-3.

Exercise $3.3.E.2$

Prove that if $A \subseteq B$, then $dA \leq dB$ and $vA \leq vB$.

Exercise $3.3.E.3$

Give an appropriate definition of a "face" and a "vertex" of A.

Exercise $3.3.E.4$

Find the edge-lengths of $A=(\bar{a}, \bar{b})$ in E^4 if

$$\bar{a}=(1,-2,4,0) \text{ and } \bar{b}=(2,0,5,3). \tag{3.3.E.1}$$

Is A a cube? Find some rational points in it. Find dA and vA.

Exercise $3.3.E.5$

Show that the sets P and Q as defined in footnote 1 are intervals, indeed. In particular, they can be made half-open (half-closed) if A is half-open (half-closed).
$[$ Hint: Let $A=(\bar{a}, \bar{b}]$,

$$P=\left\{\bar{x} \in A | x_k \leq c\right\}, \text{ and } Q=\left\{\bar{x} \in A | x_k>c\right\}. \tag{3.3.E.2}$$

To fix ideas, let $k=1$, i.e., cut the first edge. Then let

$$\bar{p}=\left(c, a_2, \ldots, a_n\right) \text{ and } \bar{q}=\left(c, b_2, \ldots, b_n\right) \text{ (see Figure 2)}, \tag{3.3.E.3}$$

and verify that $P=(\bar{a}, \bar{q}]$ and $Q=(\bar{p}, \bar{b}]$. Give a proof. $]$

Exercise $3.3.E.6$

In Problem 5, assume that A is closed, and make Q closed. (Prove it!)

Exercise $3.3.E.7$

In Problem 5 show that $(\text{with } k \text{ fixed})$ the k th edge-lengths of P and Q equal $c-a_k$ and b_k-c, respectively, while for $i \neq k$ the edge-length ℓ_i is the same in A, P, and Q, namely, $\ell_i=b_i-a_i$.
[Hint: If $k=1$, define $\bar{p}$ and $\bar{q}$ as in Problem 5. $]$

Exercise $3.3.E.8$

Prove that if an interval A is split into subintervals P and $Q(P \cap Q=\emptyset)$, then $vA=vP+vQ$.
[Hint: Use Problem 7 to compute vA, vP, and vQ. Add up.]
Give an example. (Take A as in Problem 4 and split it by the plane $x_4=1$.)

Exercise $3.3.E.9$

*9. Prove the additivity of the volume of intervals, namely, if A is subdivided, in any manner, into m mutually disjoint subintervals $A_1, A_2, \ldots, A_m \, inE^n$, then

$$vA=\sum_{i=1}^m vA_i. \tag{3.3.E.4}$$

(This is true also if some A_i contain common faces).
[Proof outline: For $m=2$, use Problem 8.
Then by induction, suppose additivity holds for any number of intervals smaller than a certain m $(m>1)$. Now let

$$A=\bigcup_{i=1}^m A_i \quad (A_i \text{ disjoint }). \tag{3.3.E.5}$$

One of the A_i (say, $A_1=[\bar{a}, \bar{p}]$) must have some edge-length smaller than the corresponding edge-length of A (say, ℓ_1). Now cut all of A into $P=[\bar{a}, \bar{d}]$ and $Q=A-P(\text{ Figure } 4)$ by the plane $x_1=c\,(c=p_1)$ so that $A_1 \subseteq P$ while $A_2 \subseteq Q$. For simplicity, assume that the plane cuts each A_i into two subintervals $A_i^{\prime}$ and $A_i^{\prime\prime}$. (One of them may be empty.)
Then

$$P=\bigcup_{i=1}^m A_i^{\prime} \text{ and } Q=\bigcup_{i=1}^m A_i^{\prime\prime}. \tag{3.3.E.6}$$

Actually, however, P and Q are split into fewer than m (nonempty) intervals since $A_1^{\prime\prime}=\emptyset=A_2^{\prime}$ by construction. Thus, by our inductive assumption,

$$vP=\sum_{i=1}^m vA_i^{\prime} \text{ and } vQ=\sum_{i=1}^m vA_i^{\prime\prime}, \tag{3.3.E.7}$$

where $vA_1^{\prime\prime}=0=vA_2^{\prime},$ and $vA_i=vA_i^{\prime}+vA_i^{\prime\prime}$ by Problem 8. Complete the inductive proof by showing that

$$vA=vP+vQ=\sum_{i=1}^m vA_i.] \tag{3.3.E.8}$$

3.4: Complex Numbers

With all the operations defined in §§1-3, $E^n (n>1)$ is not yet a field because of the lack of a vector multiplication satisfying the field axioms. We shall now define such a multiplication, but only for E^2. Thus E^2 will become a field, which we shall call the complex field, C.

We make some changes in notation and terminology here. Points ofE^2, when regarded as elements of C, will be called complex numbers (each being an ordered pair of real numbers). We denote them by single letters (preferably z) without a bar or an arrow. For example, $z=(x,y)$. We preferably write (x,y) for (x_1,x_2). If $z=(x,y)$, then x and y are called the real and imaginary parts of z, respectively, [1] and $\bar{z}$ denotes the complex number $(x,-y)$, called the conjugate of z (see Figure 5).

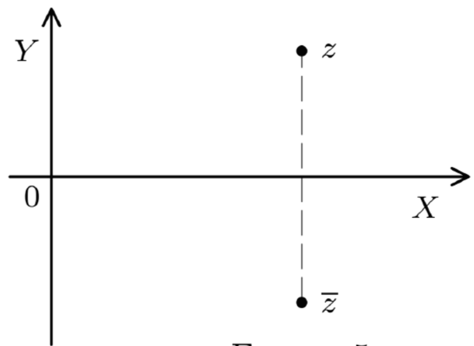

FIGURE 5

Complex numbers with vanishing imaginary part, $(x,0)$, are called real points of C. For brevity, we simply write x for $(x,0)$; for example, $2=(2,0)$. In particular, $1=(1,0)=\bar{\theta}_1$ is called the real unit in C. Points with vanishing real part, $(0,y)$, are called (purely) imaginary numbers. In particular, $\bar{\theta}_2=(0,1)$ is such a number; we shall now denote it by i and call it the imaginary unit in C. Apart from these peculiarities, all our former definitions of §§1-3 remain valid in $E^2=C$. In particular, if $z=(x,y)$ and $z'=(x',y')$, we have

$$z\pm z'=(x,y)\pm(x',y')=(x\pm x',y\pm y'), \tag{3.4.1}$$

$$\rho(z,z')=\sqrt{(x-x')^2+(y-y')^2}, \text{ and} \tag{3.4.2}$$

$$|z|=\sqrt{x^2+y^2}. \tag{3.4.3}$$

All theorems of §§1-3 are valid.

We now define the new multiplication in C, which will make it a field.

Definition

$$\text{If } z=(x,y) \text{ and } z'=(x',y'), \text{ then } zz'=(xx'-yy',xy'+yx'). \tag{3.4.4}$$

Theorem 3.4.1

$E^2=C$ is a field, with zero element $0=(0,0)$ and unity $1=(1,0)$, under addition and multiplication as defined above.

Proof

We only must show that multiplication obeys Axioms 1-6 of the field axioms. Note that for addition, all is proved in Theorem 1 of §§1-3.

Axiom 1 (closure) is obvious from our definition, for if z and z' are in C, so is zz'.

To prove commutativity, take any complex numbers

$$z=(x,y) \text{ and } z'=(x',y') \tag{3.4.5}$$

and verify that $zz' = z'z$. Indeed, by definition,

$$zz' = (xx' - yy', xy' + yx') \text{ and } z'z = (x'x - y'y, x'y + y'x); \tag{3.4.6}$$

but the two expressions coincide by the commutative laws for real numbers. Associativity and distributivity are proved in a similar manner.

Next, we show that $1 = (1, 0)$ satisfies Axiom 4(b), i.e., that $1z = z$ for any complex number $z = (x, y)$. In fact, by definition, and by axioms for E^1,

$$1z = (1, 0)(x, y) = (1x - 0y, 1y + 0x) = (x - 0, y + 0) = (x, y) = z. \tag{3.4.7}$$

It remains to verify Axiom 5(b), i.e., to show that each complex number $z = (x, y) \neq (0, 0)$ has an inverse z^{-1} such that $zz^{-1} = 1$. It turns out that the inverse is obtained by setting

$$z^{-1} = \left(\frac{x}{|z|^2}, -\frac{y}{|z|^2}\right). \tag{3.4.8}$$

In fact, we then get

$$zz^{-1} = \left(\frac{x^2}{|z|^2} + \frac{y^2}{|z|^2}, -\frac{xy}{|z|^2} + \frac{yx}{|z|^2}\right) = \left(\frac{x^2 + y^2}{|z|^2}, 0\right) = (1, 0) = 1 \tag{3.4.9}$$

since $x^2 + y^2 = |z|^2$, by definition. This completes the proof. □

corollary 3.4.1

$i^2 = -1; i.e., (0, 1)(0, 1) = (-1, 0)$.

Proof

By definition, $(0, 1)(0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0)$.

Thus C has an element i whose square is -1, while E^1 has no such element, by Corollary 2 in Chapter $2, 8\{1 - 4$. This is no contradiction since that corollary holds in ordered fields only. It only shows that C cannot be made an ordered field.

However, the "real points" in C form a subfield that can be ordered by setting

$$(x, 0) < (x', 0) \text{ iff } x < x' \text{ in } E^1. \tag{3.4.10}$$

Then this subfield behaves exactly like E^1. Therefore, it is customary not to distinguish between "real points in C'' and "real numbers," identifying $(x, 0)$ with x. With this convention, E^1 simply is a subset $($and a subfield $)$ of C. Henceforth, we shall simply say that "x is real" or "x $\in E^{1\prime}$ instead of "$x = (x, 0)$ is a real point." We then obtain the following result.

Theorem 3.4.2

Every $z \in C$ has a unique representation as

$$z = x + yi, \tag{3.4.11}$$

where x and y are real and $i = (0, 1)$. Specifically,

$$z = x + yi \text{ iff } z = (x, y). \tag{3.4.12}$$

Proof

By our conventions, $x = (x, 0)$ and $y = (y, 0)$, so

$$x + yi = (x, 0) + (y, 0)(0, 1). \tag{3.4.13}$$

Computing the right-hand expression from definitions, we have for any $x, y \in E^1$ that

$$x+yi=(x,0)+(y\cdot 0-0\cdot 1, y\cdot 1+0\cdot 1)=(x,0)+(0,y)=(x,y). \tag{3.4.14}$$

Thus $(x,y)=x+yi$ for any $x,y\in E^1$. In particular, if (x,y) is the given number $z\in C$ of the theorem, we obtain $z=(x,y)=x+yi$, as required.

To prove uniqueness, suppose that we also have

$$z=x'+y'i \text{ with } x'=(x',0) \text{ and } y'=(y',0). \tag{3.4.15}$$

Then, as shown above, $z=(x',y')$. since also $z=(x,y)$, we have $(x,y)=(x',y')$, i.e., the two ordered pairs coincide, and so $x=x'$ and $y=y'$ after all. □

Geometrically, instead of Cartesian coordinates (x,y), we may also use polar coordinates r,θ, where

$$r=\sqrt{x^2+y^2}=|z| \tag{3.4.16}$$

and θ is the (counterclockwise) rotation angle from the x-axis to the directed line $\overrightarrow{0z}$; see Figure 6. Clearly, z is uniquely determined by r and θ but θ is not uniquely determined by z; indeed, the same point of E^2 results if θ is replaced by $\theta+2n\pi (n=1,2,\ldots)$. (If $z=0$, then θ is not defined at all.) The values r and θ are called, respectively, the modulus and argument of $z=(x,y)$. By elementary trigonometry, $x=r\cos\theta$ and $y=r\sin\theta$. Substituting in $z=x+yi$, we obtain the following corollary.

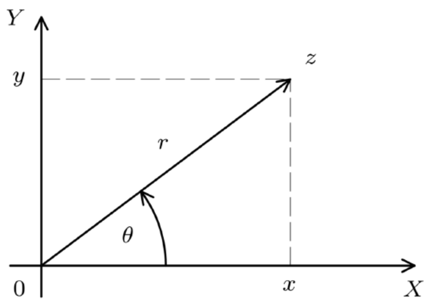

FIGURE 6

corollary 3.4.2

$$z=r(\cos\theta+i\sin\theta) (\text{trigonometric or polar form of } z). \tag{3.4.17}$$

3.4.E: Problems on Complex Numbers (Exercises)

Exercise $3.4.E.1$

Complete the proof of Theorem 1 (associativity, distributivity, etc.).

Exercise $3.4.E.1'$

Verify that the "real points" in C form an ordered field.

Exercise $3.4.E.2$

Prove that $z\bar{z}=|z|^2$. Deduce that $z^{-1}=\bar{z}/|z|^2$ if $z\neq 0$.[4]

Exercise $3.4.E.3$

Prove that

$$\overline{z+z'}=\bar{z}+\overline{z'} \text{ and } \overline{zz'}=\bar{z}\cdot\overline{z'} \tag{3.4.E.1}$$

Hence show by induction that
$\overline{z^n}=(\bar{z})^n, n=1,2,\ldots,$ and $\overline{\sum_{k=1}^n a_k z^k}=\sum_{k=1}^n \bar{a}_k \bar{z}^k$

Exercise $3.4.E.4$

Define

$$e^{\theta i}=\cos\theta+i\sin\theta. \tag{3.4.E.2}$$

Describe $e^{\theta i}$ geometrically. Is $\left|e^{\theta i}\right|=1$?

Exercise $3.4.E.5$

Compute
(a) $\frac{1+2i}{3-i}$;
(b) $(1+2i)(3-i)$; and
(c) $\frac{x+1+i}{x+1-i}, x\in E^1$.
Do it in two ways: (i) using definitions only and the notation (x,y) for $x+yi$; and (ii) using all laws valid in a field.

Exercise $3.4.E.6$

Solve the equation $(2,-1)(x,y)=(3,2)$ for x and y in E^1.

Exercise $3.4.E.7$

Let

$$\begin{aligned} z &= r(\cos\theta+i\sin\theta) \\ z' &= r'\left(\cos\theta'+i\sin\theta'\right), \text{ and} \\ z'' &= r''\left(\cos\theta''+i\sin\theta''\right) \end{aligned}$$

as in Corollary 2. Prove that $z=z'z''$ if

$$r=|z|=r'r'', \text{ i.e., } |z'z''|=|z'|\,|z''|\,, \text{ and } \theta=\theta'+\theta''. \tag{3.4.E.3}$$

Discuss the following statement: To multiply z' by z'' means to rotate $\overrightarrow{0z'}$ counterclockwise by the angle θ'' and to multiply it by the scalar $r''=|z''|$. Consider the cases $z''=i$ and $z''=-1$.
[Hint: Remove brackets in

$$r(\cos\theta+i\sin\theta)=r'\left(\cos\theta'+i\sin\theta'\right)\cdot r''\left(\cos\theta''+i\sin\theta''\right) \tag{3.4.E.4}$$

and apply the laws of trigonometry.]

Exercise $3.4.E.8$

By induction, extend Problem 7 to products of n complex numbers, and derive de Moivre's formula, namely, if $z=r(\cos\theta+i\sin\theta)$, then

$$z^n=r^n(\cos(n\theta)+i\sin(n\theta)). \tag{3.4.E.5}$$

Use it to find, for $n=1,2,\ldots$

$$(a) i^n; \quad (b)(1+i)^n; \quad (c)\frac{1}{(1+i)^n}. \tag{3.4.E.6}$$

Exercise $3.4.E.9$

From Problem 8, prove that for every complex number $z\neq 0$, there are exactly n complex numbers w such that

$$w^n=z; \tag{3.4.E.7}$$

they are called the n th roots of z
[Hint: If

$$z=r(\cos\theta+i\sin\theta) \text{ and } w=r'\left(\cos\theta'+i\sin\theta'\right), \tag{3.4.E.8}$$

the equation $w^n=z$ yields, by Problem 8

$$\left(r'\right)^n=r \text{ and } n\theta'=\theta, \tag{3.4.E.9}$$

and conversely.
While this determines r' uniquely, θ may be replaced by $\theta+2k\pi$ without affecting z. Thus

$$\theta'=\frac{\theta+2k\pi}{n}, \quad k=1,2,\ldots \tag{3.4.E.10}$$

Distinct points w result only from $k=0,1,\ldots,n-1$ (then they repeat cyclically).
Thus n values of w are obtained.]

Exercise $3.4.E.10$

Use Problem 9 to find in C

$$\text{(a) all cube roots of } 1; \quad \text{(b) all fourth roots of } 1 \tag{3.4.E.11}$$

Describe all n th roots of 1 geometrically.

3.5: Vector Spaces. The Space Cⁿ. Euclidean Spaces

I. We shall now follow the pattern of E^{n} to obtain the general notion of a vector space (just as we generalized E^{1} to define fields).

Let V be a set of arbitrary elements (not necessarily n -tuples), called "vectors" or "points," with a certain operation (call it "addition," $+$) somehow defined in V. Let F be any field (e.g. , E^{1} or C); its elements will be called scalars; its zero and unity will be denoted by 0 and 1, respectively. Suppose that yet another operation ("multiplication of scalars by vectors") has been defined that assigns to every scalar $c \in F$ and every vector $x \in V$ a certain vector, denoted cx or xc and called the c -multiple of x. Furthermore, suppose that this multiplication and addition in V satisfy the nine laws specified in Theorem 1 of §§1-3. That is, we have closure:

$$(\forall x, y \in V)(\forall c \in F) \quad x+y \in V \text{ and } cx \in V \tag{3.5.1}$$

Vector addition is commutative and associative. There is a unique zero-vector, $\overrightarrow{0}$, such that

$$(\forall x \in V) \quad x+\overrightarrow{0}=x \tag{3.5.2}$$

and each $x \in V$ has a unique inverse, $-x$, such that

$$x+(-x)=\overrightarrow{0}. \tag{3.5.3}$$

We have distributivity:

$$a(x+y)=ax+ay \text{ and } (a+b)x=ax+bx. \tag{3.5.4}$$

Finally, we have

$$1x=x \tag{3.5.5}$$

and

$$(ab)x=a(bx) \tag{3.5.6}$$

$(a, b \in F; x, y \in V)$.

In this case, V together with these two operations is called a vector space (or a linear space) over the field F; F is called its scalar field, and elements of F are called the scalars of V.

Example 3.5.1

(a) E^{n} is a vector space over E^{1} (its scalar field).

(a') R^{n}, the set of all rational points of E^{n} (i.e., points with rational coordinates is a vector space over R, the rationals in E^{1}. (Note that we could take R as a scalar field for all of E^{n}; this would yield another vector space, E^{n} over R, not to be confused with E^{n} over E^{1}, i.e., the ordinary E^{n}.)

(b) Let F be any field, and let F^{n} be the set of all ordered n -tuples of elements of F, with sums and scalar multiples defined as in E^{n} (with F playing the role of E^{1}). Then F^{n} is a vector space over $F($ proof as in Theorem 1 of §§1-3).

(c) Each field F is a vector space (over itself) under the addition and multiplication defined in F. Verify!

(d) Let V be a vector space over a field F, and let W be the set of all possible mappings

$$f: A \rightarrow V \tag{3.5.7}$$

from some arbitrary set $A \neq \emptyset$ into V. Define the sum $f+g$ of two such maps by setting

$$(f+g)(x)=f(x)+g(x) \text{ for all } x \in A. \tag{3.5.8}$$

Similarly, given $a \in F$ and $f \in W$, define the map af by

$$(af)(x)=af(x). \tag{3.5.9}$$

Vector spaces over E^{1} (respectively, C) are called real (respectively, complex) linear spaces. Complex spaces can always be transformed into real ones by restricting their scalar field C to E^{1} (treated as a subfield of C).

II. An important example of a complex linear space is C^n, the set of all ordered n-tuples

$$x=\left(x_1, \ldots, x_n\right) \tag{3.5.10}$$

of complex numbers x_k (now treated as scalars), with sums and scalar multiples defined as in E^n. In order to avoid confusion with conjugates of complex numbers, we shall not use the bar notation $\bar{x}$ for a vector in this section, writing simply x for it. Dot products in C^n are defined by

$$x \cdot y=\sum_{k=1}^n x_k \bar{y}_k, \tag{3.5.11}$$

where $\bar{y}_k$ is the conjugate of the complex number y_k (see §8), and hence a scalar in C. Note that $\bar{y}_k=y_k$ if $y_k \in E^1$. Thus, for vectors with real components,

$$x \cdot y=\sum_{k=1}^n x_k y_k, \tag{3.5.12}$$

as in E^n. The reader will easily verify (exactly as for E^n) that, for $x, y \in C^n$ and $a, b \in C$, we have the following properties:

1. $x \cdot y \in C$; thus $x \cdot y$ is a scalar, not a vector.

2. $x \cdot x \in E^1$, and $x \cdot x \geq 0$; moreover, $x \cdot x=0$ iff $x=\overrightarrow{0}$. (Thus the dot product of a vector by itself is a real number ≥ 0.)

3. $x \cdot y=\overline{y \cdot x}(=$ conjugate of $y \cdot x)$. Commutativity fails in general.

4. $(a x) \cdot(b y)=(a \bar{b})(x \cdot y)$. Hence (iv$'$) $(a x) \cdot y=a(x \cdot y)=x \cdot(\bar{a} y)$.

5. $(x+y) \cdot z=x \cdot z+y \cdot z$ and $\left(5'\right) z \cdot(x+y)=z \cdot x+z \cdot y$.

Observe that (5') follows from (5) by (3). (Verify!)

III. Sometimes (but not always) dot products can also be defined in real or complex linear spaces other than E^n or C^n, in such a manner as to satisfy the laws (1)-(5), hence also (5'), listed above, with C replaced by E^1 if the space is real. If these laws hold, the space is called Euclidean. For example, E^n is a real Euclidean space and C^n is a complex one.

In every such space, we define absolute values of vectors by

$$|x|=\sqrt{x \cdot x}. \tag{3.5.13}$$

(This root exists in E^1 by formula (ii).) In particular, this applies to E^n and C^n. Then given any vectors x, y and a scalar a, we obtain as before the following properties:

(a') $|x| \geq 0$; and $|x|=0$ iff $x=\overrightarrow{0}$.

(b') $|a x|=|a||x|$.

(c') Triangle inequality: $|x+y| \leq|x|+|y|$.

(d') Cauchy-Schwarz inequality: $|x \cdot y| \leq|x||y|$, and $|x \cdot y|=|x||y|$ iff $x \| y$ (i.e., $x=a y$ or $y=a x$ for some scalar a).

We prove only (d') ;\) the rest is proved as in Theorem 4 of §§1-3.

If $x \cdot y=0$, all is trivial, so let $z=x \cdot y=r c \neq 0$, where $r=|x \cdot y|$ and c has modulus 1, and let $y'=c y$. For any (variable)$t \in E^1$, consider $\left|t x+y'\right|$. By definition and (5),(3), and (4),

$$\begin{aligned}\left|t x+y'\right|^2 &=\left(t x+y'\right) \cdot\left(t x+y'\right) \\ &=t x \cdot t x+y' \cdot t x+t x \cdot y'+y' \cdot y' \\ &=t^2(x \cdot x)+t\left(y' \cdot x\right)+t\left(x \cdot y'\right)+\left(y' \cdot y'\right)\end{aligned}$$

since $\bar{t}=t$. Now, since $c \bar{c}=1$,

$$x \cdot y'=x \cdot(c y)=(\bar{c} x) \cdot y=\bar{c} r c=r=|x \cdot y|. \tag{3.5.14}$$

Similarly,

$$y' \cdot x = \overline{x \cdot y'} = \bar{r} = r = |x \cdot y|, x \cdot x = |x|^2, \text{ and } y' \cdot y' = y \cdot y = |y|^2. \tag{3.5.15}$$

Thus we obtain

$$(\forall t \in E^1) \quad |tx + cy|^2 = t^2|x|^2 + 2t|x \cdot y| + |y|^2. \tag{3.5.16}$$

Here $|x|^2, 2|x \cdot y|$, and $|y|^2$ are fixed real numbers. We treat them as coefficients in t of the quadratic trinomial

$$f(t) = t^2|x|^2 + 2t|x \cdot y| + |y|^2. \tag{3.5.17}$$

Now if x and y are not parallel, then $cy \neq -tx$, and so

$$|tx + cy| = |tx + y'| \neq 0 \tag{3.5.18}$$

for any $t \in E^1$. Thus by (1), the quadratic trinomial has no real roots; hence its discriminant,

$$4|x \cdot y|^2 - 4(|x||y|)^2, \tag{3.5.19}$$

is negative, so that $|x \cdot y| < |x||y|$.

If, however, $x \| y$, one easily obtains $|x \cdot y| = |x||y|$, by (b'). (Verify.)

Thus $|x \cdot y| = |x||y|$ or $|x \cdot y| < |x||y|$ according to whether $x \| y$ or not. □

In any Euclidean space, we define distances by $\rho(x, y) = |x - y|$. Planes, lines, and line segments are defined exactly as in E^n. Thus

$$\text{line } \overline{pq} = \{p + t(q - p) | t \in E^1\} \text{ (in real and complex spaces alike)}. \tag{3.5.20}$$

3.5.E: Problems on Linear Spaces (Exercises)

Exercise $3.5.E.1$

Prove that F^n in Example (b) is a vector space, i.e., that it satisfies all laws stated in Theorem 1 in §§1-3; similarly for W in Example (d).

Exercise $3.5.E.2$

Verify that dot products in C^n obey the laws $(\mathrm{i})-(\mathrm{v}')$. Which of these laws would fail if these products were defined by

$$x \cdot y=\sum_{k=1}^{n} x_{k} y_{k} \text{ instead of } x \cdot y=\sum_{k=1}^{n} x_{k} \bar{y}_{k} ? \tag{3.5.E.1}$$

How would this affect the properties of absolute values given in $(\mathrm{a}')-(\mathrm{d}')$?

Exercise $3.5.E.3$

Complete the proof of formulas $(\mathrm{a}')-(\mathrm{d}')$ for Euclidean spaces. What change would result if property (ii) of dot products were restated as

$$" x \cdot x \geq 0 \text{ and } \overrightarrow{0} \cdot \overrightarrow{0}=0'' ? \tag{3.5.E.2}$$

Exercise $3.5.E.4$

Define orthogonality, parallelism and angles in a general Euclidean space following the pattern of §§1-3 (text and Problem 7 there). Show that $u=\overrightarrow{0}$ iff u is orthogonal to all vectors of the space.

Exercise $3.5.E.5$

Define the basic unit vectors e_k in C^n exactly as in E^n, and prove
Theorem 2 in §§1-3 for C^n $\left(\text{replacing } E^1 \text{ by } C\right)$. Also, do Problem $5(\mathrm{a})$ of §§1-3 for C^n.

Exercise $3.5.E.6$

Define hyperplanes in C^n as in Definition 3 of §§4-6, and prove Theorem 1 stated there, for C^n. Do also Problems $4-6$ there for C^n (replacing E^1 by C) and Problem 4 there for vector spaces in general (replacing E^1 by the scalar field F).

Exercise $3.5.E.7$

Do Problem 3 of §§4-6 for general Euclidean spaces (real or complex). Note: Do not replace E^1 by C in the definition of a line and a line segment.

Exercise $3.5.E.8$

A finite set of vectors $B=\{x_1, \ldots, x_m\}$ in a linear space V over F is said to be independent iff

$$(\forall a_1, a_2, \ldots, a_m \in F) \quad \left(\sum_{i=1}^{m} a_i x_i=\overrightarrow{0} \Longrightarrow a_1=a_2=\cdots=a_m=0\right). \tag{3.5.E.3}$$

Prove that if B is independent, then

(i) $\overrightarrow{0} \notin B$;
(ii) each subset of B is independent ($\emptyset$ counts as independent); and
(iii) if for some scalars $a_i, b_i \in F$,

$$\sum_{i=1}^{m} a_i x_i = \sum_{i=1}^{m} b_i x_i, \tag{3.5.E.4}$$

then $a_i = b_i, i = 1, 2, \ldots, m$.

Exercise $3.5.E.9$

Let V be a vector space over F and let $A \subseteq V$. By the span of A in V, denoted $\operatorname{span}(A)$, is meant the set of all "linear combinations" of vectors from A, i.e., all vectors of the form

$$\sum_{i=1}^{m} a_i x_i, \quad a_i \in F, x_i \in A, m \in N. \tag{3.5.E.5}$$

Show that $\operatorname{span}(A)$ is itself a vector space $V' \subseteq V$ (a subspace of V) over the same field F, with the operations defined in V. (We say that A spans V'. Show that in E^n and C^n, the basic unit vectors span the entire space.

3.6: Normed Linear Spaces

By a normed linear space (briefly normed space) is meant a real or complex vector space E in which every vector x is associated with a real number $|x|$, called its absolute value or norm, in such a manner **that the properties** $(\mathbf{a}')-(\mathbf{c}')$ of §9 hold. That is, for any vectors $x, y \in E$ and scalar a, we have

(i) $|x| \geq 0$;

(i$'$) $|x|=0$ iff $x=\overrightarrow{0}$;

(ii) $|a x|=|a||x|$; and

(iii) $|x+y| \leq|x|+|y|$ (triangle inequality).

Mathematically, the existence of absolute values in E amounts to that of a map (called a norm map) $x \rightarrow|x|$ on E, i.e., a map $\varphi: E \rightarrow E^{1}$, with function values $\varphi(x)$ written as $|x|$, **satisfying the laws (i)-(iii) above**. Often such a map can be chosen in many ways (not necessarily via dot products, which may not exist in E, thus giving rise to different norms on E. Sometimes we write $\|x\|$ for $|x|$ or use other similar symbols.

Note 1. From (iii), we also obtain $|x-y| \geq||x|-|y||$ exactly as in E^{n}.

Example 3.6.1

(A) Each Euclidean space (§9) such as E^{n} or C^{n}, is a normed space, with norm defined by

$$|x|=\sqrt{x \cdot x}, \tag{3.6.1}$$

as follows from formulas (a')-(c') in §9. In E^{n} and C^{n}, one can also equivalently define

$$|x|=\sqrt{\sum_{k=1}^{n}\left|x_{k}\right|^{2}}, \tag{3.6.2}$$

where $x=\left(x_{1}, \ldots, x_{n}\right)$. This is the so-called standard norm, usually presupposed in $E^{n}\left(C^{n}\right)$.

(B) One can also define other, "nonstandard," norms on E^{n} and C^{n}. For example, fix some real $p \geq 1$ and put

$$|x|_{p}=\left(\sum_{k=1}^{n}\left|x_{k}\right|^{p}\right)^{\frac{1}{p}}. \tag{3.6.3}$$

One can show that $|x|_{p}$ so defined satisfies $(\mathrm{i})-($ iii) and thus is a norm (see Problems 5-7 below).

(C) Let W be the set of all bounded maps

$$f: A \rightarrow E \tag{3.6.4}$$

from a set $A \neq \emptyset$ into a normed space E, i.e., such that

$$(\forall t \in A) \quad|f(t)| \leq c \tag{3.6.5}$$

for some real constant $c>0$ (dependent on f but not on t). Define $f+g$ and $a f$ **as in Example (d)** of §9 so that W becomes a vector space. Also, put

$$\|f\|=\sup_{t \in A}|f(t)|, \tag{3.6.6}$$

i.e., the supremum of all $|f(t)|$, with $t \in A$. Due to boundedness, this supremum exists in E^{1}, so $\|f\| \in E^{1}$.

It is easy to show that $\|f\|$ is a norm on W. For example, we verify (iii) as follows.

By definition, we have for $f, g \in W$ and $x \in A$,

$$\begin{aligned} |(f+g)(x)| &= |f(x)+g(x)| \\ &\leq |f(x)|+|g(x)| \\ &\leq \sup_{t \in A}|f(t)|+\sup_{t \in A}|g(t)| \\ &= \|f\|+\|g\|. \end{aligned}$$

(The first inequality is true because (iii) holds in the normed space E to which $f(x)$ and $g(x)$ belong.) **By (1)**, $\|f\|+\|g\|+\|g\|$ is an upper bound of all expressions $|(f+g)(x)|, x \in A$. Thus

$$\|f\|+\|g\| \geq \sup_{x \in A}|(f+g)(x)|=\|f+g\|. \tag{3.6.7}$$

Note 2. Formula (1) also shows that the map $f+g$ is bounded and hence is a member of W. Quite similarly we see that $af \in W$ for any scalar a and $f \in W$. Thus we have the closure laws for W. The rest is easy.

In every normed (in particular, in each Euclidean) space E, we define distances by

$$\rho(x, y)=|x-y| \quad \text{for all x, y } \in E. \tag{3.6.8}$$

Such distances depend, of course, on the norm chosen for E; thus we call them norm-induced distances. In particular, using the standard norm in E^n and C^n (Example (A)), we have

$$\rho(x, y)=\sqrt{\sum_{k=1}^{n}\left|x_{k}-y_{k}\right|^{2}}. \tag{3.6.9}$$

Using the norm of Example (B), we get

$$\rho(x, y)=\left(\sum_{k=1}^{n}\left|x_{k}-y_{k}\right|^{p}\right)^{\frac{1}{p}} \tag{3.6.10}$$

instead. In the space W of **Example (C)** we have

$$\rho(f, g)=\|f-g\|=\sup_{x \in A}|f(x)-g(x)|. \tag{3.6.11}$$

Proceeding exactly as in the proof of **Theorem 5** in §§1-3, we see that norm- induced distances obey the three laws stated there. (Verify!) Moreover, by definition,

$$\rho(x+u, y+u)=|(x+u)-(y+u)|=|x-y|=\rho(x, y). \tag{3.6.12}$$

Thus we have

$$\rho(x, y)=\rho(x+u, y+u) \text{ for norm-induced distances;} \tag{3.6.13}$$

i.e., the distance $\rho(x, y)$ does not change if both x and y are "translated" by one and the same vector u. We call such distances translation-invariant.

A more general theory of distances will be given in §§11ff.

3.6.E: Problems on Normed Linear Spaces (Exercises)

Exercise $3.6.E.1$

Show that distances in normed spaces obey the laws stated in Theorem 5 of §§1-3.

Exercise $3.6.E.2$

Complete the proof of assertions made in Example (C) and Note 2.

Exercise $3.6.E.3$

Define $|x|=x_1$ for $x=(x_1,\ldots,x_n)$ in C^n or E^n. Is this a norm? Which (if any) of the laws (i) - (iii) does it obey? How about formula (2)?

Exercise $3.6.E.4$

Do Problem 3 in §§4-6 for a general normed space E, with lines defined as in E^n (see also Problem 7 in §9). Also, show that contracting sequences of line segments in E are f-images of contracting sequences of intervals in E^1. Using this fact, deduce from Problem 11 in Chapter 2 §§8-9, an analogue for line segments in E, namely, if

$$L\left[a_n,b_n\right]\supseteq L\left[a_{n+1},b_{n+1}\right],\quad n=1,2,\ldots \tag{3.6.E.1}$$

then

$$\bigcap_{n=1}^{\infty} L\left[a_n,b_n\right]\neq\emptyset. \tag{3.6.E.2}$$

Exercise $3.6.E.5$

Take for granted the lemma that

$$a^{1/p}b^{1/q}\leq\frac{a}{p}+\frac{b}{q} \tag{3.6.E.3}$$

if $a,b,p,q\in E^1$ with $a,b\geq 0$ and $p,q>0$, and

$$\frac{1}{p}+\frac{1}{q}=1. \tag{3.6.E.4}$$

(A proof will be suggested in Chapter 5, §6, Problem $11.$) Use it to prove Hölder's inequality, namely, if $p>1$ and $\frac{1}{p}+\frac{1}{q}=1,$ then

$$\sum_{k=1}^{n}\left|x_k y_k\right|\leq\left(\sum_{k=1}^{n}\left|x_k\right|^p\right)^{\frac{1}{p}}\left(\sum_{k=1}^{n}\left|y_k\right|^q\right)^{\frac{1}{q}}\text{ for any } x_k,y_k\in C. \tag{3.6.E.5}$$

[Hint: Let

$$A=\left(\sum_{k=1}^{n}\left|x_k\right|^p\right)^{\frac{1}{p}}\text{ and } B=\left(\sum_{k=1}^{n}\left|y_k\right|^q\right)^{\frac{1}{q}}. \tag{3.6.E.6}$$

If $A=0$ or $B=0$, then all x_k or all y_k vanish, and the required inequality is trivial. Thus assume $A \neq 0$ and $B \neq 0$. Then, setting

$$a=\frac{\left|x_k\right|^p}{A^p} \text { and } b=\frac{\left|y_k\right|^q}{B^q} \tag{3.6.E.7}$$

in the lemma, obtain

$$\frac{\left|x_k y_k\right|}{A B} \leq \frac{\left|x_k\right|^p}{p A^p}+\frac{\left|y_k\right|^q}{q B^q}, k=1,2, \ldots, n. \tag{3.6.E.8}$$

Now add up these n inequalities, substitute the values of A and B, and simplify.]

? Exercise $3.6.E.6$

Prove the Minkowski inequality,

$$\left(\sum_{k=1}^n\left|x_k+y_k\right|^p\right)^{\frac{1}{p}} \leq\left(\sum_{k=1}^n\left|x_k\right|^p\right)^{\frac{1}{p}}+\left(\sum_{k=1}^n\left|y_k\right|^p\right)^{\frac{1}{p}} \tag{3.6.E.9}$$

for any real $p \geq 1$ and $x_k, y_k \in C$.
[Hint: If $p=1$, this follows by the triangle inequality in C. If $p>1$, let

$$A=\sum_{k=1}^n\left|x_k+y_k\right|^p \neq 0. \tag{3.6.E.10}$$

(If $A=0$, all is trivial.) Then verify (writing "$\sum$" for "$\sum_{k=1}^n$" for simplicity)

$$A=\sum\left|x_k+y_k\right|\left|x_k+y_k\right|^{p-1} \leq \sum\left|x_k\right|\left|x_k+y_k\right|^{p-1}+\sum\left|y_k\right|\left|x_k+y_k\right|^{p-1} \tag{3.6.E.11}$$

Now apply Hölder's inequality (Problem 5) to each of the last two sums, with $q=p /(p-1)$, so that $(p-1) q=p$ and $1 / p=1-1 / q$. Thus obtain

$$A \leq\left(\sum\left|x_k\right|^p\right)^{\frac{1}{p}}\left(\sum\left|x_k+y_k\right|^p\right)^{\frac{1}{q}}+\left(\sum\left|y_k\right|^p\right)^{\frac{1}{p}}\left(\sum\left|x_k+y_k\right|^p\right)^{\frac{1}{q}}. \tag{3.6.E.12}$$

Then divide by $A^{\frac{1}{q}}=\left(\sum\left|x_k+y_k\right|^p\right)^{\frac{1}{q}}$ and simplify.]

? Exercise $3.6.E.7$

Show that Example (B) indeed yields a norm for C^n and E^n.
[Hint: For the triangle inequality, use Problem 6. The rest is easy.]

? Exercise $3.6.E.8$

A sequence $\left\{x_m\right\}$ of vectors in a normed space E (e.g. , in E^n or C^n) is said to be bounded iff

$$\left(\exists c \in E^1\right)(\forall m) \quad\left|x_m\right|<c, \tag{3.6.E.13}$$

i.e., iff $\sup_m\left|x_m\right|$ is finite.

Denote such sequences by single letters, $x=\{x_m\}$, $y=\{y_m\}$, etc. and define

$$x+y=\{x_m+y_m\}, \text{ and } ax=\{ax_m\} \text{ for any scalar } a. \tag{3.6.E.14}$$

Also let

$$|x|=\sup_m |x_m|. \tag{3.6.E.15}$$

Show that, with these definitions, the set M of all bounded infinite sequences in E becomes a normed space (in which every such sequence is to be treated as a single vector, and the scalar field is the same as that of E).

3.7: Metric Spaces

I. In §§1-3, we defined distances $\rho(\overline{x}, \overline{y})$ for points $\overline{x}, \overline{y}$ in E^n using the formula

$$\rho(\overline{x}, \overline{y})=\sqrt{\sum_{k=1}^{n}\left(x_k-y_k\right)^2}=|\overline{x}-\overline{y}|. \tag{3.7.1}$$

This actually amounts to defining a certain function ρ of two variables $\overline{x}, \overline{y} \in E^n$. We also showed that ρ obeys the three laws of Theorem 5 there. (We call them metric laws.)

Now, as will be seen, such functions ρ can also be defined in other sets, using quite different defining formulas. In other words, given any set $S \neq \emptyset$ of arbitrary elements, one can define in it, so to say, "fancy distances" $\rho(x, y)$ satisfying the same three laws. It turns out that it is not the particular formula used to define ρ but rather the preservation of the three laws that is most important for general theoretical purposes.

Thus we shall assume that a function ρ with the same three properties has been defined, in some way or other, for a set $S \neq \emptyset$, and propose to study the consequences of the three metric laws alone, without assuming anything else. (In particular, no operations other than ρ, or absolute values, or inequalities $<$ need be defined in S.) All results so obtained will, of course, apply to distances in E^n (since they obey the metric laws), but they will also apply to other cases where the metric laws hold.

The elements of S (though arbitrary) will be called "points," usually denoted by p, q, x, y, z (sometimes with bars, etc.); ρ is called a metric for S. We symbolize it by

$$\rho: S \times S \rightarrow E^1 \tag{3.7.2}$$

since it is function defined on $S \times S$ (pairs of elements of S) into E^1. Thus we are led to the following definition.

Definition

A metric space is a set $S \neq \emptyset$ together with a function

$$\rho: S \times S \rightarrow E^1 \tag{3.7.3}$$

(called a metric for S) satisfying the metric laws (axioms):

For any x, y, and z in S, we have

i. $\rho(x, y) \geq 0$, and (i$'$) $\rho(x, y)=0$ iff $x=y$;
ii. $\rho(x, y)=\rho(y, x)$ (symmetry law); and
iii. $\rho(x, z) \leq \rho(x, y)+\rho(y, z)($ triangle law).

Thus a metric space is a pair (S, ρ), namely, a set S and a metric ρ for it. In general, one can define many different metrics

$$\rho, \rho^{\prime}, \rho^{\prime \prime}, \ldots \tag{3.7.4}$$

for the same S. The resulting spaces

$$(S, \rho),\left(S, \rho^{\prime}\right),\left(S, \rho^{\prime \prime}\right), \ldots \tag{3.7.5}$$

then are regarded as different. However, if confusion is unlikely, we simply write S for (S, ρ). We write "$p \in(S, \rho)$" for "$p \in S$ with metric ρ," and "$A \subseteq(S, \rho)$" for "$A \subseteq S$ in (S, ρ)."

Example 3.7.1

(1) In E^n, we always assume

$$\rho(\overline{x}, \overline{y})=|\overline{x}-\overline{y}| \text { (the "standard metric") } \tag{3.7.6}$$

unless stated otherwise. **By Theorem 5** in §§1-3, $\left(E^n, \rho\right)$ is a metric space.

(2) However, one can define for E^n many other "nonstandard" metrics. For example,

$$\rho'(\bar{x}, \bar{y}) = \left(\sum_{k=1}^{n} |x_k - y_k|^p\right)^{1/p} \text{ for any real } p \geq 1 \tag{3.7.7}$$

likewise satisfies the metric laws (a proof is suggested in §10, Problems 5-7; similarly for C^n.

(3) Any set $S \neq \emptyset$ can be "metrized" (i.e., endowed with a metric) by setting

$$\rho(x, y) = 1 \text{ if } x \neq y, \text{ and } \rho(x, x) = 0. \tag{3.7.8}$$

(Verify the metric laws!) This is the so-called discrete metric. The space (S, ρ) so defined is called a discrete space.

(4) Distances ("mileages") on the surface of our planet are actually measured along circles fitting in the curvature of the globe (not straight lines). One can show that they obey the metric laws and thus define a (nonstandard) metric for $S =$ (surface of the globe).

(5) A mapping $f : A \rightarrow E^1$ is said to be bounded iff

$$\left(\exists K \in E^1\right)(\forall x \in A) \quad |f(x)| \leq K. \tag{3.7.9}$$

For a fixed $A \neq \emptyset$, let W be the set of all such maps (each being treated as a single "point" of W). Metrize W by setting, for $f, g \in W$,

$$\rho(f, g) = \sup_{x \in A} |f(x) - g(x)|. \tag{3.7.10}$$

(Verify the metric laws; see a similar proof in §10.)

II. We now define "balls" in any metric space (S, ρ).

Definition

Given $p \in (S, \rho)$ and a real $\varepsilon > 0$, we define the open ball or globe with center p and radius ε (briefly "ε-globe about p", denoted

$$G_p \text{ or } G_p(\varepsilon) \text{ or } G(p; \varepsilon), \tag{3.7.11}$$

to be the set of all $x \in S$ such that

$$\rho(x, p) < \varepsilon. \tag{3.7.12}$$

Similarly, the closed ε-globe about p is

$$\overline{G}_p = \overline{G}_p(\varepsilon) = \{x \in S | \rho(x, p) \leq \varepsilon\}. \tag{3.7.13}$$

The ε -sphere about p is defined by

$$S_p(\varepsilon) = \{x \in S | \rho(x, p) = \varepsilon\}. \tag{3.7.14}$$

Note. An open globe in E^3 is an ordinary solid sphere (without its surface $S_p(\varepsilon)$), as known from geometry. In E^2, an open globe is a disc (the interior of a circle). In E^1, the globe $G_p(\varepsilon)$ is simply the open interval $(p - \varepsilon, p + \varepsilon)$, while $\overline{G}_p(\varepsilon)$ is the closed interval $[p - \varepsilon, p + \varepsilon]$.

The shape of the globes and spheres depends on the metric ρ. It may become rather strange for various unusual metrics. For example, in the discrete space (Example (3)), any globe of radius < 1 consists of its center alone, while $G_p(2)$ contains the entire space. (Why?) See also Problems 1 ,2, and 4.

III. Now take any nonempty set $A \subseteq (S, \rho)$.

The distances $\rho(x, y)$ in S are, of course, also defined for points of A (since $A \subseteq S$, and the metric laws remain valid in A. Thus A is likewise a (smaller) metric space under the metric ρ "inherited" from S; we only have to restrict the domain of ρ to $A \times A$ (pairs of points from A). The set A with this metric is called a subspace of S. We shall denote it by (A, ρ), using the same letter ρ or simply by A. Note that A with some other metric ρ' is not called a subspace of (S, ρ).

By definition, points in (A, ρ) have the same distances as in (S, ρ). However, globes and spheres in (A, ρ) must consist of points from A only, with centers in A. Denoting such a globe by

$$G_{p}^{*}(\varepsilon)=\{x \in A | \rho(x, p)<\varepsilon\}, \tag{3.7.15}$$

we see that it is obtained by restricting $G_{p}(\varepsilon)$ (the corresponding globe in S) to points of A, i.e., removing all points not in A. Thus

$$G_{p}^{*}(\varepsilon)=A \cap G_{p}(\varepsilon); \tag{3.7.16}$$

similarly for closed globes and spheres. $A \cap G_{p}(\varepsilon)$ is often called the relativized (to A) globe $G_{p}(\varepsilon)$. Note that $p \in G_{p}^{*}(\varepsilon)$ since $\rho(p, p)=0<\varepsilon,$ and $p \in A$.

For example, let R be the subspace of E^{1} consisting of rationals only. Then the relativized globe $G_{p}^{*}(\varepsilon)$ consists of all rationals in the interval

$$G_{p}(\varepsilon)=(p-\varepsilon, p+\varepsilon), \tag{3.7.17}$$

and it is assumed here that p is rational itself.

IV. A few remarks are due on the extended real number system E^{*} (see Chapter 2, §§13) . As we know, E^{*} consists of all reals and two additional elements, $\pm \infty$, with the convention that $-\infty<x<+\infty$ for all $x \in E^{1}$. The standard metric ρ does not apply to E^{*}. However, one can metrize E^{*} in various other ways. The most common metric $\rho^{\prime}$ is suggested in Problems 5 and 6 below. Under that metric, globes turn out to be finite and infinite intervals in E^{*}.

Instead of metrizing E^{*}, we may simply adopt the convention that intervals of the form

$$(a,+\infty] \text { and }[-\infty, a), a \in E^{1}, \tag{3.7.18}$$

will be called "globes" about $+\infty$ and $-\infty$, respectively (without specifying any "radii"). Globes about finite points may remain as they are in E^{1}. This convention suffices for most purposes of limit theory. We shall use it often (as we did in Chapter 2, §13).

3.7.E: Problems on Metric Spaces (Exercises)

The "arrowed" problems should be noted for later work.

Exercise $3.7.E.1$

Show that E^2 becomes a metric space if distances $\rho(\overline{x}, \overline{y})$ are defined by
(a) $\rho(\overline{x}, \overline{y}) = |x_1 - y_1| + |x_2 - y_2|$ or
(b) $\rho(\overline{x}, \overline{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$,
where $\overline{x} = (x_1, x_2)$ and $\overline{y} = (y_1, y_2)$. In each case, describe $G_{\overline{0}}(1)$ and $S_{\overline{0}}(1)$. Do the same for the subspace of points with nonnegative coordinates.

Exercise $3.7.E.2$

Prove the assertions made in the text about globes in a discrete space. Find an empty sphere in such a space. Can a sphere contain the entire space?

Exercise $3.7.E.3$

Show that ρ in Examples (3) and (5) obeys the metric axioms.

Exercise $3.7.E.4$

Let M be the set of all positive integers together with the "point" ∞. Metrize M by setting

$$\rho(m, n) = \left|\frac{1}{m} - \frac{1}{n}\right|, \text{ with the convention that } \frac{1}{\infty} = 0. \tag{3.7.E.1}$$

Verify the metric axioms. Describe $G_\infty\left(\frac{1}{2}\right)$, $S_\infty\left(\frac{1}{2}\right)$, and $G_1(1)$.

Exercise $3.7.E.5$

$\Rightarrow 5.$ Metrize the extended real number system E^* by

$$\rho'(x, y) = |f(x) - f(y)|, \tag{3.7.E.2}$$

where the function

$$f : E^* \underset{\text{onto}}{\longrightarrow} [-1, 1] \tag{3.7.E.3}$$

is defined by

$$f(x) = \frac{x}{1 + |x|} \text{ if } x \text{ is finite, } f(-\infty) = -1, \text{ and } f(+\infty) = 1. \tag{3.7.E.4}$$

Compute $\rho'(0, +\infty)$, $\rho'(0, -\infty)$, $\rho'(-\infty, +\infty)$, $\rho'(0, 1)$, $\rho'(1, 2)$, and $\rho'(n, +\infty)$. Describe $G_0(1)$, $G_{+\infty}(1)$, and $G_{-\infty}\left(\frac{1}{2}\right)$. Verify the metric axioms (also when infinities are involved).

Exercise $3.7.E.6$

$\Rightarrow 6.$ In Problem 5, show that the function f is one to one, onto $[-1, 1]$, and increasing; i.e.

$$x < x' \text{ implies } f(x) < f\left(x'\right) \text{ for } x, x' \in E^*. \tag{3.7.E.5}$$

Also show that the f -image of an interval $(a, b) \subseteq E^*$ is the interval $(f(a), f(b))$. Hence deduce that globes in E^* (with ρ' as in Problem 5) are intervals in E^* (possibly infinite).
[Hint: For a finite x, put

$$y = f(x) = \frac{x}{1+|x|}. \tag{3.7.E.6}$$

Solving for x (separately in the cases $x \geq 0$ and $x < 0$), show that

$$(\forall y \in (-1,1)) \quad x = f^{-1}(y) = \frac{y}{1-|y|}; \tag{3.7.E.7}$$

thus x is uniquely determined by y, i.e., f is one to one and onto-each $y \in (-1,1)$ corresponds to some $x \in E^1$. (How about ± 1?)
To show that f is increasing, consider separately the three cases $x < 0 < x'$, $x < x' < 0$ and $0 < x < x'$ (also for infinite x and x').]

Exercise $3.7.E.7$

Continuing Problems 5 and 6, consider (E^1, ρ') as a subspace of (E^*, ρ') with ρ' as in Problem 5. Show that globes in (E^1, ρ') are exactly all open intervals in E^*. For example, $(0,1)$ is a globe. What are its center and radius under ρ' and under the standard metric ρ?

Exercise $3.7.E.8$

Metrize the closed interval $[0, +\infty]$ in E^* by setting

$$\rho(x, y) = \left| \frac{1}{1+x} - \frac{1}{1+y} \right|, \tag{3.7.E.8}$$

with the conventions $1 + (+\infty) = +\infty$ and $1/(+\infty) = 0$. Verify the metric axioms. Describe $G_p(1)$ for arbitrary $p \geq 0$.

Exercise $3.7.E.9$

Prove that if ρ is a metric for S, then another metric ρ' for S is given by
(i) $\rho'(x, y) = \min\{1, \rho(x, y)\}$;
(ii) $\rho'(x, y) = \frac{\rho(x,y)}{1+\rho(x,y)}$.
In case (i), show that globes $G_p(\varepsilon)$ of radius $\varepsilon \leq 1$ are the same under ρ and ρ'. In case (ii), prove that any $G_p(\varepsilon)$ in (S, ρ) is also a globe $G_p(\varepsilon')$ in (S, ρ') of radius

$$\varepsilon' = \frac{\varepsilon}{1+\varepsilon}, \tag{3.7.E.9}$$

and any globe of radius $\varepsilon' < 1$ in (S, ρ') is also a globe in (S, ρ). (Find the converse formula for ε as well!)
[Hint for the triangle inequality in (ii): Let $a = \rho(x, z), b = \rho(x, y)$, and $c = \rho(y, z)$ so that $a \leq b + c$. The required inequality is

$$\frac{a}{1+a} \leq \frac{b}{1+b} + \frac{c}{1+c}. \tag{3.7.E.10}$$

Simplify it and show that it follows from $a \leq b + c$.]

Exercise $3.7.E.10$

Prove that if (X, ρ') and (Y, ρ'') are metric spaces, then a metric ρ for the set $X \times Y$ is obtained by setting, for $x_1, x_2 \in X$ and $y_1, y_2 \in Y$,
(i) $\rho\left(\left(x_1, y_1\right), \left(x_2, y_2\right)\right) = \max\left\{\rho'\left(x_1, x_2\right), \rho''\left(y_1, y_2\right)\right\}$; or
(ii) $\rho\left(\left(x_1, y_1\right), \left(x_2, y_2\right)\right) = \sqrt{\rho'(x_1, x_2)^2 + \rho''(y_1, y_2)^2}$.
[Hint: For brevity, put $\rho'_{12} = \rho'\left(x_1, x_2\right), \rho''_{12} = \rho''\left(y_1, y_2\right)$, etc. The triangle inequality in (ii),

$$\sqrt{\left(\rho'_{13}\right)^2 + \left(\rho''_{13}\right)^2} \leq \sqrt{\left(\rho'_{12}\right)^2 + \left(\rho''_{12}\right)^2} + \sqrt{\left(\rho'_{23}\right)^2 + \left(\rho''_{23}\right)^2}, \tag{3.7.E.11}$$

is verified by squaring both sides, isolating the remaining square root on the right side, simplifying, and squaring again. Simplify by using the triangle inequalities valid in X and Y, i.e.,

$$\rho'_{13} \leq \rho'_{12} + \rho'_{23} \text{ and } \rho''_{13} \leq \rho''_{12} + \rho''_{23}. \tag{3.7.E.12}$$

Reverse all steps, so that the required inequality becomes the last step.]

Exercise $3.7.E.11$

Prove that

$$|\rho(y, z) - \rho(x, z)| \leq \rho(x, y) \tag{3.7.E.13}$$

in any metric space (S, ρ).
[Caution: The formula $\rho(x, y) = |x - y|$, valid in E^n, cannot be used in (S, ρ). Why?]

Exercise $3.7.E.12$

Prove that

$$\rho\left(p_1, p_2\right) + \rho\left(p_2, p_3\right) + \cdots + \rho\left(p_{n-1}, p_n\right) \geq \rho\left(p_1, p_n\right). \tag{3.7.E.14}$$

[Hint: Use induction.]

3.8: Open and Closed Sets. Neighborhoods

I. Let A be an open globe in (S, ρ) or an open interval $(\bar{a}, \bar{b})$ in E^n. Then every $p \in A$ can be enclosed in a small globe $G_p(\delta) \subseteq A($ Figures 7 and 8$)$. (This would fail for "boundary" points; but there are none inside an open G_q or $(\bar{a}, \bar{b})$.).

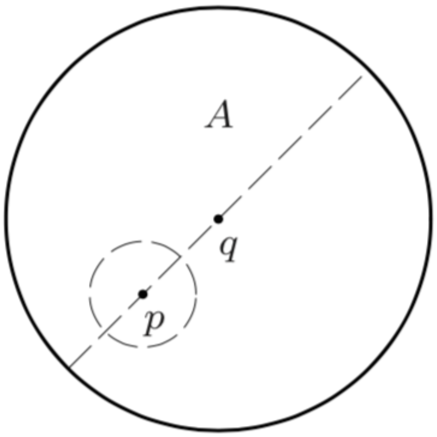

FIGURE 7

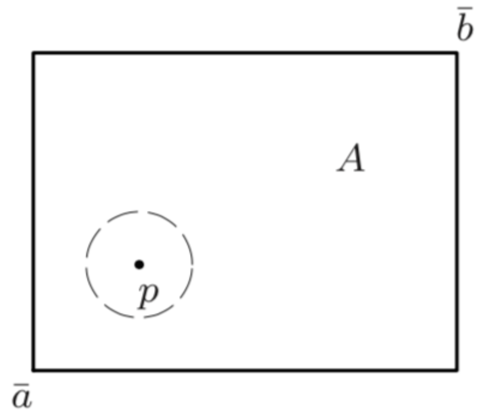

FIGURE 8

This suggests the following ideas, for any (S, ρ).

Definition

A point p is said to be interior to a set $A \subseteq (S, \rho)$ iff A contains some G_p; i.e., p, together with some globe G_p, belongs to A. We then also say that A is a neighborhood of p. The set of all interior points of A ("the interior of A'' is denoted A^0. Note: $\emptyset^0 = \emptyset$ and $S^0 = S$!

Definition

A set $A \subseteq (S, \rho)$ is said to be open iff A coincides with its interior $\left(A^0 = A\right)$. Such are $\emptyset$ and S.

Example 3.8.1

(1) As noted above, an open globe $G_q(r)$ has interior points only, and thus is an open set in the sense of Definition 2. (See Problem 1 for a proof.)

(2) The same applies to an open interval $(\bar{a}, \bar{b})$ in E^n. (See Problem 2.)

(3) The interior of any interval in E^n never includes its endpoints $\bar{a}$ and $\bar{b}$. In fact, it coincides with the open interval $(\bar{a}, \bar{b})$. (See Problem 4.)

(4) The set R of all rationals in E^1 has no interior points at all $\left(R^0 = \emptyset\right)$ because it cannot contain any $G_p = (p - \varepsilon, p + \varepsilon)$. Indeed, any such G_p contains irrationals (see Chapter 2, §§11-12, Problem 5), so it is not entirely contained in R.

Theorem 3.8.1

(Hausdorff property). Any two points p and q $(p \neq q)$ in (S, ρ) are centers of two disjoint globes.

More precisely,

$$(\exists \varepsilon > 0) \quad G_p(\varepsilon) \cap G_q(\varepsilon) = \emptyset. \tag{3.8.1}$$

Proof

As $p \neq q$, we have $\rho(p,q)>0$ by metric axiom (i$'$). Thus we may put

$$\varepsilon=\frac{1}{2}\rho(p,q)>0. \tag{3.8.2}$$

It remains to show that with this $\varepsilon, G_p(\varepsilon) \cap G_q(\varepsilon)=\emptyset$.

Seeking a contradiction, suppose this fails. Then there is $x \in G_p(\varepsilon) \cap G_q(\varepsilon)$ so that $\rho(p,x)<\varepsilon$ and $\rho(x,q)<\varepsilon$. By the triangle law,

$$\rho(p,q) \leq \rho(p,x)+\rho(x,q)<\varepsilon+\varepsilon=2\varepsilon; \text{ i.e., } \rho(p,q)<2\varepsilon, \tag{3.8.3}$$

which is impossible since $\rho(p,q)=2\varepsilon.$ □

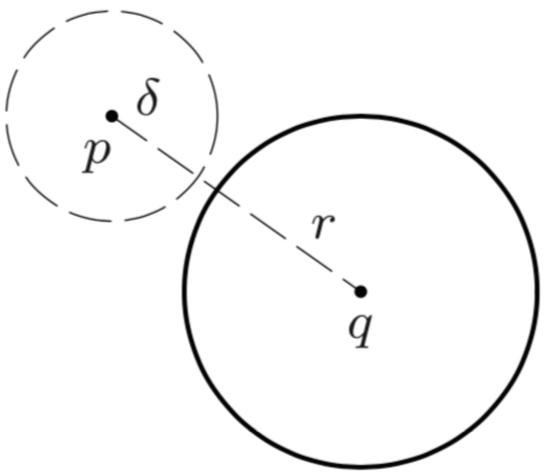

FIGURE 9

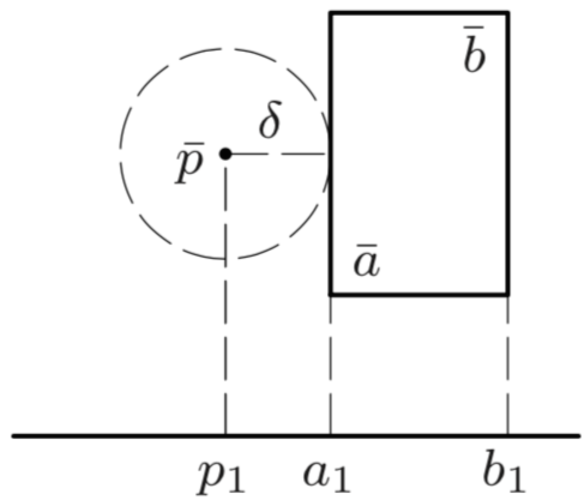

FIGURE 10

Note. A look at Figure 9 explains the idea of this proof, namely, to obtain two disjoint globes of equal radius, it suffices to choose $\varepsilon \leq \frac{1}{2}\rho(p,q)$. The reader is advised to use such diagrams in E^2 as a guide.

II. We can now define closed sets in terms of open sets.

Definition

A set $A \subseteq (S,\rho)$ is said to be closed iff its complement $-A=S-A$ is open, i.e., has interior points only.

That is, each $p \in -A$ (outside A) is in some globe $G_p \subseteq -A$ so that

$$A \cap G_p=\emptyset. \tag{3.8.4}$$

Example 3.8.1

(Continued).

(5) The sets $\emptyset$ and S are closed, for their complements, S and $\emptyset$, are open, as noted above. Thus a set may be both closed and open ("clopen").

(6) All closed globes in (S,ρ) and all closed intervals in E^n are closed sets by Definition 3. Indeed (see Figures 9 and 10), if $A=\overline{G}_q(r)$ or $A=[\bar{a},\bar{b}]$, then any point p outside A can be enclosed in a globe $G_p(\delta)$ disjoint from A; so, by Definition 3, A

is closed (see Problem 12).

(7) A one-point set $\{q\}$ (also called "singleton") in (S, ρ) is always closed, for any p outside $\{q\}(p \neq q)$ is in a globe disjoint from $\{q\}$ by Theorem 1 In a discrete space (§§11,) Example (3)), $\{q\}$ is also open since it is an open g lobe, $\{q\} = G_q\left(\frac{1}{2}\right)($ why?); so it is "clopen. " Hence, in such a space, all sets are "clopen". For $p \in A$ implies $\{p\} = G_p\left(\frac{1}{2}\right) \subseteq A;$ similarly for $-A.$ Thus A and $-A$ have interior points only, so both are open.

(8) The interval $(a, b]$ in E^1 is neither open nor closed. (Why?)

III. (The rest of this section may be deferred until Chapter (4, §10.)

Theorem 3.8.2

The union of any finite or infinite family of open sets $A_i(i \in I)$, denoted

$$\bigcup_{i \in I} A_i, \tag{3.8.5}$$

is open itself. So also is

$$\bigcap_{i=1}^{n} A_i \tag{3.8.6}$$

for finitely many open sets. (This fails for infinitely many sets A_i; see Problem 11 below.)

Proof

We must show that any point p of $A = \bigcup_i A_i$ is interior to A.

Now if $p \in \bigcup_i A_i, p$ is in some A_i, and it is an interior point of A_i (for A_i is open, by assumption). Thus there is a globe

$$G_p \subseteq A_i \subseteq A, \tag{3.8.7}$$

as required.

For finite intersections, it suffices to consider two open sets A and B (for n sets, all then follows by induction). We must show that each $p \in A \cap B$ is interior to $A \cap B$.

Now as $p \in A$ and A is open, we have some $G_p\left(\delta'\right) \subseteq A.$ Similarly, there is $G_p\left(\delta''\right) \subseteq B.$ Then the smaller of the two globes, call it G_p, is in both A and B, so

$$G_p \subseteq A \cap B \tag{3.8.8}$$

and p is interior to $A \cap B,$ indeed. □

Theorem 3.8.3

If the sets $A_i(i \in I)$ are closed, so is

$$\bigcap_{i \in I} A_i \tag{3.8.9}$$

(even for infinitely many sets). So also is

$$\bigcup_{i=1}^{n} A_i \tag{3.8.10}$$

for finitely many closed sets A_i. (Again, this fails for infinitely many sets A_i.)

Proof

Let $A = \bigcap_{i \in I} A_i.$ To prove that A is closed, we show that $-A$ is open.

Now by set theory (see Chapter 1, §§1-3, Theorem 2) ,

$$-A=-\bigcap_i A_i=\bigcup_i\left(-A_i\right), \tag{3.8.11}$$

where the $\left(-A_i\right)$ are open (for the A_i are closed). Thus by Theorem $2, -A$ is open, as required.

The second assertion (as to $\bigcup_{i=1}^n A_i$) follows quite similarly. $\square$

corollary 3.8.1

A nonempty set $A \subseteq(S, \rho)$ is open iff A is a union of open globes.

For if A is such a union, it is open by Theorem 2. Conversely, if A is open, then each $p \in A$ is in some $G_p \subseteq A$. All such $G_p(p \in A)$ cover all of A, so $A \subseteq \bigcup_{p \in A} G_p$. Also, $\bigcup_{p \in A} G_p \subseteq A$ since all G_p are in A. Thus

$$A=\bigcup_{p \in A} G_p . \tag{3.8.12}$$

corollary 3.8.2

Every finite set F in a metric space (S, ρ) is closed.

Proof

If $F=\emptyset, F$ is closed by Example (5). If $F \neq \emptyset$, let

$$F=\left\{p_1, \ldots, p_n\right\}=\bigcup_{k=1}^n\left\{p_k\right\} . \tag{3.8.13}$$

Now by Example (7), each $\left\{p_k\right\}$ is closed; hence so is F by by theorem $3 .$ $\square$

Note. The family of all open sets in a given space (S, ρ) is denoted by $\mathcal{G}$; that of all closed sets, by $\mathcal{F}$. Thus " $A \in \mathcal{G}''$ means that A is open; "A $\in \mathcal{F}''$ means that A is closed. By Theorems 2 and 3, we have

$$(\forall A, B \in \mathcal{G}) \quad A \cup B \in \mathcal{G} \text{ and } A \cap B \in \mathcal{G}; \tag{3.8.14}$$

similarly for $\mathcal{F}$. This is a kind of "closure law." We say that $\mathcal{F}$ and $\mathcal{G}$ are "closed under finite unions and intersections."

In conclusion, consider any subspace (A, ρ) of (S, ρ). As we know from §11 it is a metric space itself, so it has its own open and closed sets (which must consist of points of A only). We shall now show that they are obtained from those of (S, ρ) by intersecting the latter sets with A.

Theorem 3.8.4

Let (A, ρ) be a subspace of (S, ρ). Then the open (closed) sets in (A, ρ) are exactly all sets of the form $A \cap U$, with U open (closed) in S.

Proof

Let G be open in (A, ρ). By Corollary $1, G$ is the union of some open globes $G_i^*(i \in I)$ in (A, ρ). (For brevity, we omit the centers and radii; we also omit the trivial case $G=\emptyset$.

$$G=\bigcup_i G_i^*=\bigcup_i\left(A \cap G_i\right)=A \cap \bigcup_i G_i, \tag{3.8.15}$$

by set theory (see Chapter 1, §§1-3,\) Problem 9).

Again by Corollary $1, U=\bigcup_i G_i$ is an open set in (S, ρ). Thus G has the form

$$A \cap \bigcup_i G_i=A \cap U, \tag{3.8.16}$$

with U open in S, as asserted.

Conversely, assume the latter, and let $p \in G$. Then $p \in A$ and $p \in U$. As U is open in (S, ρ), there is a globe G_p in (S, ρ) such that $p \in G_p \subseteq U$. As $p \in A$, we have

$$p \in A \cap G_p \subseteq A \cap U. \tag{3.8.17}$$

However, $A \cap G_p$ is a globe in (A, ρ), call it G_p^*. Thus

$$p \in G_p^* \subseteq A \cap U = G; \tag{3.8.18}$$

i.e., p is an interior point of G in (A, ρ). We see that each $p \in G$ is interior to G, as a set in (A, ρ), so G is open in (A, ρ).

This proves the theorem for open sets. Now let F be closed in (A, ρ). Then by Definition 3, $A - F$ is open in (A, ρ). (Of course, when working in (A, ρ), we replace S by A in taking complements.) Let $G = A - F$, so $F = A - G$, and G is open in (A, ρ). By what was shown above, $G = A \cap U$ with U open in S.

Thus

$$F = A - G = A - (A \cap U) = A - U = A \cap (-U) \tag{3.8.19}$$

by set theory. Here $-U = S - U$ is closed in (S, ρ) since U is open there. Thus $F = A \cap (-U)$, as required.

The proof of the converse (for closed sets) is left as an exercise. □

3.8.E: Problems on Neighborhoods, Open and Closed Sets (Exercises)

Exercise $3.8.E.1$

$\Rightarrow 1.$ Verify Example (1).
[Hint: Given $p \in G_q(r)$, let

$$\delta = r - \rho(p, q) > 0. \quad \text{(Why } > 0?) \tag{3.8.E.1}$$

Use the triangle law to show that

$$x \in G_p(\delta) \Rightarrow \rho(x, q) < r \Rightarrow x \in G_q(r).] \tag{3.8.E.2}$$

Exercise $3.8.E.2$

$\Rightarrow 2.$ Check Example (2); see Figure 8.
[Hint: If $\overline{p} \in (\overline{a}, \overline{b})$, choose δ less than the $2n$ numbers

$$p_k - a_k \text{ and } b_k - p_k, \quad k = 1, \ldots, n; \tag{3.8.E.3}$$

then show that $G_{\overline{p}}(\delta) \subseteq (\overline{a}, \overline{b})$.]

Exercise $3.8.E.3$

Prove that if $\overline{p} \in G_{\overline{q}}(r)$ in E^n, then $G_{\overline{q}}(r)$ contains a cube $[\overline{c}, \overline{d}]$ with $\overline{c} \neq \overline{d}$ and with center $\overline{p}$.
[Hint: By Example (1), there is $G_{\overline{p}}(\delta) \subseteq G_{\overline{q}}(r)$. Inscribe in $G_{\overline{p}}\left(\frac{1}{2}\delta\right)$ a cube of diagonal δ. Find its edge-length $(\delta/\sqrt{n})$. Then use it to find the coordinates of the endpoints, $\overline{c}$ and $\overline{d}($ given $\overline{p},$ the center$)$. Prove that $[\overline{c}, \overline{d}] \subseteq G_{\overline{p}}(\delta)$.]

Exercise $3.8.E.4$

Verify Example (3).
[Hint: To show that no interior points of $[\overline{a}, \overline{b}]$ are outside $(\overline{a}, \overline{b})$, let $\overline{p} \notin (\overline{a}, \overline{b})$. Then at least one of the inequalities $a_k < p_k$ or $p_k < b_k$ fails. (Why?) Let it be $a_1 < p_1$, say, so $p_1 \leq a_1$.
Now take any globe $G_{\overline{p}}(\delta)$ about $\overline{p}$ and prove that it is not contained in $[\overline{a}, \overline{b}]$ (so $\overline{p}$ cannot be an interior point). For this purpose, as in Problem 3, show that $G_{\overline{p}}(\delta) \supseteq [\overline{c}, \overline{d}]$ with $c_1 < p_1 \leq a_1$. Deduce that $\overline{c} \in G_{\overline{p}}(\delta)$, but $\overline{c} \notin [\overline{a}, \overline{b}]$; so $G_{\overline{p}}(\delta) \nsubseteq [\overline{a}, \overline{b}]$.]

Exercise $3.8.E.5$

Prove that each open globe $G_{\overline{q}}(r)$ in E^n is a union of cubes (which can be made open, closed, half-open, etc., as desired). Also, show that each open interval $(\overline{a}, \overline{b}) \neq \emptyset$ in E^n is a union of open (or closed) globes.
[Hint for the first part: By Problem 3, each $\overline{p} \in G_{\overline{q}}(r)$ is in a cube $C_p \subseteq G_{\overline{q}}(r)$. Show that $G_{\overline{q}}(r) = \bigcup C_p$.]

Exercise $3.8.E.6$

Show that every globe in E^n contains rational points, i.e., those with rational coordinates only (we express it by saying that the set R^n of such points is dense in E^n); similarly for the set I^n of irrational points (those with irrational coordinates).
[Hint: First check it with globes replaced by cubes $(\overline{c}, \overline{d})$; see §$7$, Corollary 3. Then use Problem 3 above.]

? Exercise 3.8.E.7

Prove that if $\overline{x} \in G_{\overline{q}}(r)$ in E^n, there is a rational point $\overline{p}$ (Problem 6) and a rational number $\delta > 0$ such that $\overline{x} \in G_{\overline{p}}(\delta) \subseteq G_{\overline{q}}(r)$. Deduce that each globe $G_{\overline{q}}(r)$ in E^n is a union of rational globes (those with rational centers and radii). Similarly, show that $G_{\overline{q}}(r)$ is a union of intervals with rational endpoints.
[Hint for the first part: Use Problem 6 and Example (1).]

? Exercise 3.8.E.8

Prove that if the points $p_1, \ldots, p_n$ in (S, ρ) are distinct, there is an $\varepsilon > 0$ such that the globes $G(p_k; \varepsilon)$ are disjoint from each other, for $k = 1, 2, \ldots, n$.

? Exercise 3.8.E.9

Do Problem 7, with $G_{\overline{q}}(r)$ replaced by an arbitrary open set $G \neq \emptyset$ in E^n.

? Exercise 3.8.E.10

Show that every open set $G \neq \emptyset$ in E^n is infinite ($*$ even uncountable; see Chapter 1, §9).
[Hint: Choose $G_{\overline{q}}(r) \subseteq G$. By Problem 3, $G_{\overline{p}}(r) \supset L[\overline{c}, \overline{d}]$, a line segment.]

? Exercise 3.8.E.11

Give examples to show that an infinite intersection of open sets may not be open, and an infinite union of closed sets may not be closed.
[Hint: Show that

$$\bigcap_{n=1}^{\infty}\left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\} \tag{3.8.E.4}$$

and

$$\bigcup_{n=2}^{\infty}\left[\frac{1}{n}, 1-\frac{1}{n}\right] = (0,1).] \tag{3.8.E.5}$$

? Exercise 3.8.E.12

Verify Example (6) as suggested in Figures 9 and 10.
[Hints: (i) For $\overline{G}_q(r)$, take

$$\delta = \rho(p, q) - r > 0. \quad (\text{Why} > 0?) \tag{3.8.E.6}$$

(ii) If $\overline{p} \notin [\overline{a}, \overline{b}]$, at least one of the $2n$ inequalities $a_k \leq p_k$ or $p_k \leq b_k$ fails (why?), say, $p_1 < a_1$. Take $\delta = a_1 - p_1$.
In both (i) and (ii) prove that $A \cap G_p(\delta) = \emptyset$ (proceed as in Theorem 1).]

? Exercise 3.8.E.$*13$

Prove the last parts of Theorems 3 and 4.

Exercise 3.8.E.*14

Prove that A^0, the interior of A, is the union of all open globes contained in A (assume $A^0 \neq \emptyset$). Deduce that A^0 is an open set, the largest contained in A.

Exercise 3.8.E.*15

For sets $A, B \subseteq (S, \rho)$, prove that
(i) $(A \cap B)^0 = A^0 \cap B^0$;
(ii) $\left(A^0\right)^0 = A^0$; and
(iii) if $A \subseteq B$ then $A^0 \subseteq B^0$.
$\left[\text{Hint for (ii) : } A^0 \text{ is open by Problem 14.}\right]$

Exercise 3.8.E.16

Is $A^0 \cup B^0 = (A \cup B)^0$?
[Hint: See Example (4). Take $A = R, B = E^1 - R$.]

Exercise 3.8.E.17

Prove that if M and N are neighborhoods of p in (S, ρ), then
(a) $p \in M \cap N$;
(b) $M \cap N$ is a neighborhood of p;
*(c) so is M^0; and
(d) so also is each set $P \subseteq S$ such that $P \supseteq M$ or $P \supseteq N$.
$[\text{Hint for (c) : See Problem 14.}]$

Exercise 3.8.E.18

The boundary of a set $A \subseteq (S, \rho)$ is defined by

$$\operatorname{bd} A = -\left[A^0 \cup (-A)^0\right]; \tag{3.8.E.7}$$

thus it consists of points that fail to be interior in A or in $-A$.
Prove that the following statements are true:
(i) $S = A^0 \cup \operatorname{bd} A \cup (-A)^0$, all disjoint.
(ii) $\operatorname{bd} S = \emptyset$, $\operatorname{bd} \emptyset = \emptyset$.
$*(\text{ iii })A$ is open iff $A \cap \operatorname{bd} A = \emptyset$; A is closed iff $A \supseteq \operatorname{bd} A$.
$(\text{iv}) \operatorname{In} E^n$,

$$\operatorname{bd} G_{\bar{p}}(r) = \operatorname{bd} \overline{G_{\bar{p}}}(r) = S_{\bar{p}}(r) \tag{3.8.E.8}$$

(the sphere with center $\bar{p}$ and radius r). Is this true in all metric spaces?
[Hint: Consider $G_p(1)$ in a discrete space (S, ρ) with more than one point in S; see §11, Example (3).]
$(\text{v}) \operatorname{In} E^n$, if $(\bar{a}, \bar{b}) \neq \emptyset$, then
$\operatorname{bd}(\bar{a}, \bar{b}] = \operatorname{bd}[\bar{a}, \bar{b}) = \operatorname{bd}(\bar{a}, \bar{b}) = \operatorname{bd}[\bar{a}, \bar{b}] = [\bar{a}, \bar{b}] - (\bar{a}, \bar{b})$.
(vi) $\operatorname{In} E^n, \left(R^n\right)^0 = \emptyset$; hence $\operatorname{bd} R^n = E^n$ $\left(R^n \text{ as in Problem } 6\right)$.

Exercise $3.8.E.19$

Verify Example (8) for intervals in E^n.

3.9: Bounded Sets. Diameters

I. Geometrically, the diameter of a closed globe in E^n could be defined as the maximum distance between two of its points. In an open globe in E^n, there is no "maximum" distance (why?), but we still may consider the supremum of all distances inside the globe. Moreover, this makes sense in any set $A \subseteq (S, \rho)$. Thus we accept it as a general definition, for any such set.

Definition

The diameter of a set $A \neq \emptyset$ in a metric space (S, ρ), denoted dA, is the supremum (in E^*) of all distances $\rho(x, y)$, with $x, y \in A$;[1] in symbols,

$$dA = \sup_{x,y \in A} \rho(x, y). \tag{3.9.1}$$

If $A = \emptyset$, we put $dA = 0$. If $dA < +\infty$, A is said to be bounded (in (S, ρ)).

Equivalently, we could define a bounded set as in the statement of the following theorem.

Theorem 3.9.1

A set $A \subseteq (S, \rho)$ is bounded iff A is contained in some globe. If so, the center p of this globe can be chosen at will.

Proof

If $A = \emptyset$, all is trivial.

Thus let $A \neq \emptyset$; let $q \in A$, and choose any $p \in S$. Now if A is bounded, then $dA < +\infty$, so we can choose a real $\varepsilon > \rho(p, q) + dA$ as a suitable radius for a globe $G_p(\varepsilon) \supseteq A($ see Figure 11 for motivation$)$. Now if $x \in A$, then by the definition of dA $\rho(q, x) \leq dA$; so by the triangle law,

$$\begin{aligned} \rho(p, x) &\leq \rho(p, q) + \rho(q, x) \\ &\leq \rho(p, q) + dA < \varepsilon; \end{aligned}$$

i.e., $x \in G_p(\varepsilon)$. Thus $(\forall x \in A) x \in G_p(\varepsilon)$ as required.

Conversely, if $A \subseteq G_p(\varepsilon)$, then any $x, y \in A$ are also in $G_p(\varepsilon)$; so $\rho(x, p) < \varepsilon$ and $\rho(p, y) < \varepsilon$, whence

$$\rho(x, y) \leq \rho(x, p) + \rho(p, y) < \varepsilon + \varepsilon = 2\varepsilon. \tag{3.9.2}$$

Thus 2ε is an upper bound of all $\rho(x, y)$ with $x, y \in A$. Therefore,

$$dA = \sup \rho(x, y) \leq 2\varepsilon < +\infty; \tag{3.9.3}$$

i.e., A is bounded, and all is proved. □

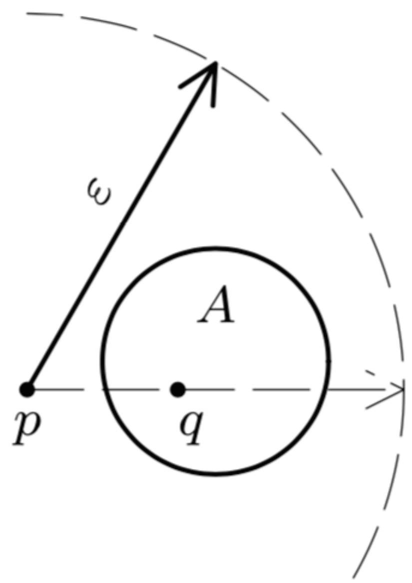

FIGURE 11

As a special case we obtain the following.

Theorem 3.9.1

A set $A \subseteq E^{n}$ is bounded iff there is a real $K>0$ such that

$$(\forall \overline{x} \in A) \quad|\overline{x}|<K \tag{3.9.4}$$

(*similarly in C^{n} and other normed spaces).

Proof

By Theorem 1(choosing $\overline{0}$ for p), A is bounded iff A is contained in some globe $G_{\overline{0}}(\varepsilon)$ about $\overline{0}$. That is,

$$(\forall \overline{x} \in A) \quad \overline{x} \in G_{\overline{0}}(\varepsilon) \text { or } \rho(\overline{x}, \overline{0})=|\overline{x}|<\varepsilon. \tag{3.9.5}$$

Thus ε is the required K.(*The proof for normed spaces is the same.) □

Note 1. In E^{1}, this means that

$$(\forall x \in A) \quad-K<x<K; \tag{3.9.6}$$

i.e., A is bounded by $-K$ and K. This agrees with our former definition, given in Chapter 2, §§8-9.

Caution: Upper and lower bounds are not defined in (S, ρ), in general.

Example 3.9.1

(1) $\emptyset$ is bounded, with $d \emptyset=0,$ by definition.

(2) Let $A=[\overline{a}, \overline{b}]$ in E^{n}, with $d=\rho(\overline{a}, \overline{b})$ its diagonal. By Corollary 1 in §7 d is the largest distance in A. In nonclosed intervals, we still have

$$d=\sup _{x, y \in A} \rho(x, y)=d A<+\infty \text { (see Problem 10 (ii)). } \tag{3.9.7}$$

Thus all intervals in E^{n} are bounded.

(3) Each globe $G_p(\varepsilon)$ in (S, ρ) is bounded, with $dG_p(\varepsilon) \leq 2\varepsilon < +\infty$, as was shown in the proof of Theorem 1. See, however, Problems 5 and 6 below.

(4) All of E^n is not bounded, under the standard metric, for if E^n had a finite diameter d, no distance in E^n would exceed d; but $\rho\left(-d\bar{e}_1, d\bar{e}_1\right) = 2d$, a contradiction!

(5) On the other hand, under the discrete metric §11, Example (3)), any set (even the entire space) is contained in $G_p(3)$ and hence bounded. The same applies to the metric ρ' defined for E^* in Problem 5 of §§11, since distances under that metric never exceed 2, and so $E^* \subseteq G_p(3)$ for any choice of p.

Note 2. This shows that boundedness depends on the metric ρ. A set may be bounded under one metric and not bounded under another. A metric ρ is said to be bounded iff all sets are bounded under ρ (as in Example (5)).

Problem 9 of §11 shows that any metric ρ can be transformed into a bounded one, even preserving all sufficiently small globes; in part (i) of the problem, even the radii remain the same if they are ≤ 1.

Note 3. An idea similar to that of diameter is often used to define distances between sets. If $A \neq \emptyset$ and $B \neq \emptyset$ in (S, ρ), we define $\rho(A, B)$ to be the infimum of all distances $\rho(x, y)$, with $x \in A$ and $y \in B$. In particular, if $B = \{p\}$ (a singleton), we write $\rho(A, p)$ for $\rho(A, B)$. Thus

$$\rho(A, p) = \inf_{x \in A} \rho(x, p). \tag{3.9.8}$$

II. The definition of boundedness extends, in a natural manner, to sequences and functions. We briefly write $\{x_m\} \subseteq (S, \rho)$ for a sequence of points in (S, ρ), and $f : A \rightarrow (S, \rho)$ for a mapping of an arbitrary set A into the space S. Instead of "infinite sequence with general term x_m," we say "the sequence x_m."

Definition

A sequence $\{x_m\} \subseteq (S, \rho)$ is said to be bounded iff its range is bounded in (S, ρ), i.e., iff all its terms x_m are contained in some globe in (S, ρ).

In E^n, this means (by Theorem 2) that

$$(\forall m) \quad |x_m| < K \tag{3.9.9}$$

for some fixed $K \in E^1$.

Definition

A function $f : A \rightarrow (S, \rho)$ is said to be bounded on a set $B \subseteq A$ iff the image set $f[B]$ is bounded in (S, ρ); i.e. iff all function values $f(x)$, with $x \in B$, are in some globe in (S, ρ).

In E^n, this means that

$$(\forall x \in B) \quad |f(x)| < K \tag{3.9.10}$$

for some fixed $K \in E^1$.

If $B = A$, we simply say that f is bounded.

Note 4. If $S = E^1$ or $S = E^*$, we may also speak of upper and lower bounds. It is customary to call sup $f[B]$ also the supremum of f on B and denote it by symbols like

$$\sup_{x \in B} f(x) \text{ or } \sup\{f(x) | x \in B\}. \tag{3.9.11}$$

In the case of sequences, we often write sup $_m x_m$ or sup x_m instead; similarly for infima, maxima, and minima.

Example 3.9.1

(a) The sequence

$$x_m = \frac{1}{m} \quad \text{in } E^1 \tag{3.9.12}$$

is bounded since all terms x_m are in the interval $(0,2) = G_1(1)$. We have $\inf x_m = 0$ and $\sup x_m = \max x_m = 1$.

(b) The sequence

$$x_m = m \quad \text{in } E^1 \tag{3.9.13}$$

is bounded below (by 1) but not above. We have $\inf x_m = \min x_m = 1$ and $\sup x_m = +\infty$ (in E^*).

(c) Define $f : E^1 \rightarrow E^1$ by

$$f(x) = 2x. \tag{3.9.14}$$

This map is bounded on each finite interval $B = (a,b)$ since $f[B] = (2a, 2b)$ is itself an interval and hence bounded. However, f is not bounded on all of E^1 since $f\left[E^1\right] = E^1$ is not a bounded set.

(d) Under a bounded metric ρ, all functions $f : A \rightarrow (S, \rho)$ are bounded.

(e) The so-called identity map on S, $f : S \rightarrow (S, \rho)$, is defined by

$$f(x) = x. \tag{3.9.15}$$

Clearly, f carries each set $B \subseteq S$ onto itself; i.e., $f[B] = B$. Thus f is bounded on B iff B is itself a bounded set in (S, ρ).

(f) Define $f : E^1 \rightarrow E^1$ by

$$f(x) = \sin x. \tag{3.9.16}$$

Then $f\left[E^1\right] = [-1, 1]$ is a bounded set in the range space E^1. Thus f is bounded on E^1 (briefly, bounded).

3.9.E: Problems on Boundedness and Diameters (Exercises)

Exercise $3.9.E.1$

Show that if a set A in a metric space is bounded, so is each subset $B \subseteq A$.

Exercise $3.9.E.2$

Prove that if the sets $A_1, A_2, \ldots, A_n$ in (S, ρ) are bounded, so is

$$\bigcup_{k=1}^{n} A_k. \tag{3.9.E.1}$$

Disprove this for infinite unions by a counterexample.
[Hint: By Theorem 1, each A_k is in some $G_p\left(\varepsilon_k\right)$, with one and the same center p. If the number of the globes is finite, we can put max $\left(\varepsilon_1, \ldots, \varepsilon_n\right)=\varepsilon$, so $G_p(\varepsilon)$ contains all A_k. Verify this in detail.

Exercise $3.9.E.3$

$\Rightarrow 3.$ From Problems 1 and 2 show that a set A in (S, ρ) is bounded iff it is contained in a finite union of globes,

$$\bigcup_{k=1}^{n} G\left(p_k ; \varepsilon_k\right). \tag{3.9.E.2}$$

Exercise $3.9.E.4$

A set A in (S, ρ) is said to be totally bounded iff for every $\varepsilon>0$ (no matter how small), A is contained in a finite union of globes of radius ε. By Problem 3, any such set is bounded. Disprove the converse by a counterexample.
[Hint: Take an infinite set in a discrete space.]

Exercise $3.9.E.5$

Show that distances between points of a globe $\overline{G_p}(\varepsilon)$ never exceed 2ε. (Use the triangle inequality!) Hence infer that $dG_p(\varepsilon) \leq 2\varepsilon$. Give an example where $dG_p(\varepsilon)<2\varepsilon$. Thus the diameter of a globe may be less than twice its radius.
[Hint: Take a globe $G_p\left(\frac{1}{2}\right)$ in a discrete space.]

Exercise $3.9.E.6$

Show that in E^n ($*$ as well as in C^n and any other normed linear space $\neq\{0\}$), the diameter of a globe $G_p(\varepsilon)$ always equals 2ε (twice its radius).
[Hint: By Problem $5, 2\varepsilon$ is an upper bound of all $\rho(\bar{x}, \bar{y})$ with $\bar{x}, \bar{y} \in G_p(\varepsilon)$.
To show that there is no smaller upper bound, prove that any number

$$2\varepsilon-2r \quad(r>0) \tag{3.9.E.3}$$

is exceeded by some $\rho(\bar{x}, \bar{y})$; e.g., take $\bar{x}$ and $\bar{y}$ on some line through $\bar{p}$,

$$\overline{\boldsymbol{x}}=\overline{\boldsymbol{p}}+\boldsymbol{t} \overrightarrow{\boldsymbol{u}}, \tag{3.9.E.4}$$

choosing suitable values for t to get $\rho(\bar{x}, \bar{y})=|\bar{x}-\bar{y}|>2\varepsilon-2r.$]

Exercise $3.9.E.7$

Prove that in E^n, a set A is bounded iff it is contained in an interval.

Exercise $3.9.E.8$

Prove that for all sets A and B in (S, ρ) and each $p \in S$

$$\rho(A, B) \leq \rho(A, p)+\rho(p, B). \tag{3.9.E.5}$$

Disprove

$$\rho(A, B)<\rho(A, p)+\rho(p, B) \tag{3.9.E.6}$$

by an example.

Exercise $3.9.E.9$

Find $\sup x_n, \inf x_n, \max x_n$, and $\min x_n$ (if any) for sequences with general term
(a) n;
(b) $(-1)^n\left(2-2^{2-n}\right)$;
(c) $1-\frac{2}{n}$;
(d) $\frac{n(n-1)}{(n+2)^2}$.
Which are bounded in E^1?

Exercise $3.9.E.10$

Prove the following about lines and line segments.
(i) Show that any line segment in E^n is a bounded set, but the entire line is not.
(ii) Prove that the diameter of $L(\bar{a}, \bar{b})$ and of $(\bar{a}, \bar{b})$ equals $\rho(\bar{a}, \bar{b})$.

Exercise $3.9.E.11$

Let $f: E^1 \rightarrow E^1$ be given by

$$f(x)=\frac{1}{x} \text { if } x \neq 0, \text { and } f(0)=0. \tag{3.9.E.7}$$

Show that f is bounded on an interval $[a, b]$ iff $0 \notin[a, b]$. Is f bounded on $(0,1)$?

Exercise $3.9.E.12$

Prove the following:
(a) If $A \subseteq B \subseteq(S, \rho)$, then $d A \leq d B$.
(b) $d A=0$ iff A contains at most one point.
(c) If $A \cap B \neq \emptyset$, then

$$d(A \cup B) \leq d A+d B. \tag{3.9.E.8}$$

Show by an example that this may fail if $A \cap B=\emptyset$.

3.10: Cluster Points. Convergent Sequences

This page is a draft and is under active development.

Consider the set

$$A=\left\{1, \frac{1}{2}, \ldots, \frac{1}{m}, \ldots\right\} ; \tag{3.10.1}$$

we may as well let A denote the sequence $x_m=1/m$ in $E^{1.1}$ Plotting it on the axis, we observe a remarkable fact: The points x_m "cluster" close to 0, approaching 0 as m increases-see Figure 12.

FIGURE 12

To make this more precise, take any globe about 0 in E^1, $G_0(\varepsilon)=(-\varepsilon, \varepsilon)$. No matter how small, it contains infinitely many (even all but finitely many) points x_m, namely, all from some x_k onward, so that

$$(\forall m>k) \quad x_m \in G_0(\varepsilon). \tag{3.10.2}$$

Indeed, take $k>1/\varepsilon$, so $1/k<\varepsilon$. Then

$$(\forall m>k) \quad \frac{1}{m}<\frac{1}{k}<\varepsilon; \tag{3.10.3}$$

i.e., $x_m \in(-\varepsilon, \varepsilon)=G_0(\varepsilon)$.
This suggests the following generalizations.

Definition: cluster at a point

A set, or sequence, $A \subseteq(S, \rho)$ is said to **cluster at a point** $p \in S$ (not necessarily $p \in A$), and p is called its cluster point or accumulation point, iff every globe G_p about p contains infinitely many points (respectively, terms of A. (Thus only infinite sets can cluster.

Note 1. In sequences (unlike sets) an infinitely repeating term counts as infinitely many terms. For example, the sequence $0,1,0,1$, clusters at 0 and 1 (why?); but its range, $\{0,1\}$, has no cluster points (being finite). This distinction is, however, irrelevant if all terms x_m are distinct, i.e., different from each other. Then we may treat sequences and sets alike.

Definition

A sequence $\{x_m\} \subseteq(S, \rho)$ is said to converge or tend to a point p in S, and p is called its limit, iff every globe $G_p(\varepsilon)$ about p (no matter how small) contains all but finitely many terms x_m.[2] In symbols,

$$(\forall \varepsilon>0)(\exists k)(\forall m>k) \quad x_m \in G_p(\varepsilon), \text{ i.e., } \rho\left(x_m, p\right)<\varepsilon \tag{3.10.4}$$

If such a p exists, we call $\{x_m\}$ a convergent sequence in (S, ρ)); otherwise, a divergent one. The notation is

$$x_m \rightarrow p, \text{ or } \lim x_m=p, \text{ or } \lim_{m \rightarrow \infty} x_m=p. \tag{3.10.5}$$

In E^n, $\rho\left(\bar{x}_m, \bar{p}\right)=\left|\bar{x}_m-\bar{p}\right|$; thus formula (1) turns into

$$\bar{x}_m \rightarrow \bar{p} \text{ in } E^n \text{ iff } (\forall \varepsilon>0)(\exists k)(\forall m>k) \quad \left|\bar{x}_m-\bar{p}\right|<\varepsilon \tag{3.10.6}$$

Since "all but finitely many" (as in Definition 2) implies "infinitely many" (as in Definition 1), any limit is also a cluster point. Moreover, we obtain the following result.

corollary 3.10.1

If $x_m \rightarrow p$, then p is the unique cluster point of $\{x_m\}$. (Thus a sequence with two or more cluster points, or none at all, diverges.) For if $p \neq q$, the Hausdorff property (Theorem 1 of §12) yields an ε such that

$$G_p(\varepsilon) \cap G_q(\varepsilon)=\emptyset. \tag{3.10.7}$$

As $x_m \rightarrow p$, $G_p(\varepsilon)$ leaves out at most finitely many x_m, and only these can possibly be in $G_q(\varepsilon)$. (Why?) Thus q fails to satisfy Definition 1 and hence is no cluster point. Hence $\lim x_m$ (if it exists) is unique.

corollary 3.10.2

(i) We have $x_m \rightarrow p$ in (S, ρ) iff $\rho\left(x_m, p\right) \rightarrow 0$ in E^1.
Hence
(ii) $\overline{x}_m \rightarrow \overline{p}$ in E^n iff $\left|\overline{x}_m-\overline{p}\right| \rightarrow 0$ and
(iii) $\overline{x}_m \rightarrow \overline{0}$ in E^n iff $\left|\overline{x}_m\right| \rightarrow 0$.

Proof

By (2), we have $\rho\left(x_m, p\right) \rightarrow 0$ in E^1 if

$$(\forall \varepsilon>0)(\exists k)(\forall m>k) \quad\left|\rho\left(x_m, p\right)-0\right|=\rho\left(x_m, p\right)<\varepsilon. \tag{3.10.8}$$

By (1), however, this means that $x_m \rightarrow p$, proving our first assertion. The rest easily follows from it, since $\rho\left(\overline{x}_m, \overline{p}\right)=\left|\overline{x}_m-\overline{p}\right|$ in E^n. $\square$

corollary 3.10.3

If x_m tends to p, then so does each subsequence x_{m_k}

For $x_m \rightarrow p$ means that each G_p leaves out at most finitely many x_m. This certainly still holds if we drop some terms, passing to $\left\{x_{m_k}\right\}$.

Note 2. A similar argument shows that the convergence or divergence of $\left\{x_m\right\}$, and its limit or cluster points, are not affected by dropping or adding
a finite number of terms; similarly for cluster points of sets. For example, if $\{x_m\}$ tends to p, so does $\{x_{m+1}\}$ (the same sequence without x_1).

We leave the following two corollaries as exercises.

corollary 3.10.4

If $\{x_m\}$ splits into two subsequences, each tending to the same limit p, then also $x_m \rightarrow p$.

corollary 3.10.5

If $\{x_m\}$ converges in (S, ρ), it is bounded there.

Of course, the convergence or divergence of $\{x_m\}$ and its clustering depend on the metric ρ and the space S. Our theory applies to any (S, ρ). In particular, it applies to E^*, with the metric $\rho^{\prime}$ of Problem 5 in §11. Recall that under that metric, globes about $\pm \infty$ have the form $(a,+\infty]$ and $[-\infty, a)$, respectively. Thus limits and cluster points in $\left(E^*, \rho^{\prime}\right)$ coincide with those defined in Chapter 2, §13, (formulas $(1)-(3)$ and Definition 2 there). Our theory then applies to infinite limits as well, and generalizes Chapter 2, §13.

Example 3.10.1

(a) Let

$$x_m = p \quad \text{for all } m \tag{3.10.9}$$

(such sequences are called constant). As $p \in G_p$, any G_p contains all x_m. Thus $x_m \rightarrow p$, by Definition 2. We see that each constant sequence converges to the common value of its terms.

(b) In our introductory example, we showed that

$$\lim_{m \rightarrow \infty} \frac{1}{m} = 0 \quad \text{in } E^1 \tag{3.10.10}$$

and that 0 is the (unique) cluster point of the set $A = \left\{1, \frac{1}{2}, \ldots\right\}$. Here $0 \notin A$.

(c) The sequence

$$0, 1, 0, 1, \ldots \tag{3.10.11}$$

has two cluster points, 0 and 1, so it diverges by Corollary 1. (It "oscillates" from 0 to 1.) This shows that a bounded sequence may diverge. The converse to Corollary 5 fails.

(d) The sequence

$$x_m = m \tag{3.10.12}$$

(or the set N of all naturals) has *no* cluster points in E^1, for a globe of radius $< \frac{1}{2}$ (with any center $p \in E^1$) contains at most one x_m, and hence no p satisfies Definition 1 or 2.

However, $\{x_m\}$ does cluster in (E^*, ρ'), and even has a limit there,
namely $+\infty$. (Prove it!)

(e) The set R of all rationals in E^1 clusters at each $p \in E^1$. Indeed, any globe

$$G_p(\varepsilon) = (p - \varepsilon, p + \varepsilon) \tag{3.10.13}$$

contains infinitely many rationals (see Chapter 2, §10, Theorem 3), and this means that each $p \in E^1$ is a cluster point of R.

(f) The sequence

$$1, 1, 2, \frac{1}{2}, 3, \frac{1}{3}, \ldots \quad \left(\text{with } x_{2k} = \frac{1}{k} \text{ and } x_{2k-1} = k\right) \tag{3.10.14}$$

has only one cluster point, 0, in E^1; yet it diverges, being unbounded (see Corollary 5). In (E^*, ρ'), it has two cluster points, 0 and $+\infty$. (Verify!)

(g) The lim and lim of any sequence in E^* are cluster points (cf. Chapter 2, §13, Theorem 2 and Problem 4). Thus in E^*, all sequences cluster.

(h) Let

$$A = [a, b], \quad a < b. \tag{3.10.15}$$

Then A clusters exactly at all its points, for if $p \in A$, then any globe

$$G_p(\varepsilon) = (p - \varepsilon, p + \varepsilon) \tag{3.10.16}$$

overlaps with A (even with (a, b)) and so contains infinitely many points of A, as required. Even the endpoints a and b are cluster points of A (and of (a, b), $(a, b]$, and $[a, b)$). On the other hand, no point outside A is a cluster point. (Why?)

(i) In a discrete space (§11, Example (3)), no set can cluster, since small globes, such as $G_p\left(\frac{1}{2}\right)$, are singletons. (Explain!)

Example (h) shows that a set A may equal the set of its cluster points (call it A'); i. e.

$$A = A'. \tag{3.10.17}$$

Such sets are said to be perfect. Sometimes we have $A \subseteq A', A' \subseteq A, A' = S$ (as in Example (e)), or $A' = \emptyset$. We conclude with the following result.

corollary 3.10.6

A set $A \subseteq (S, \rho)$ clusters at p iff each globe G_p (about p) contains at least one point of A other than p.

Indeed, assume the latter. Then, in particular, each globe

$$G_p\left(\frac{1}{n}\right), \quad n = 1, 2, \ldots \tag{3.10.18}$$

contains some point of A other than p; call it x_n. We can make the x_n distinct by choosing each time x_{n+1} closer to p than x_n is. It easily follows that each $G_p(\varepsilon)$ contains infinitely many points of A (the details are left to the reader), as required. The converse is obvious.

3.10.E: Problems on Cluster Points and Convergence (Exercises)

Exercise $3.10.E.1$

Is the Archimedean property (see Chapter 2, §10) involved in the proof that

$$\lim_{m \rightarrow \infty} \frac{1}{m}=0? \tag{3.10.E.1}$$

Exercise $3.10.E.2$

Prove Note 2 and Corollaries 4 and 6.

Exercise $3.10.E.3$

Verify Example (c) in detail.

Exercise $3.10.E.4$

Prove Corollary 5.
[Hint: Fix some $G_{p}(\varepsilon)$. Use Definition 2. If $G_{p}(\varepsilon)$ leaves out $x_{1}, x_{2}, \ldots, x_{k}$, take a larger radius r greater than

$$\rho\left(x_{m}, p\right), \quad m=1,2, \ldots, k. \tag{3.10.E.2}$$

Then the enlarged globe $G_{p}(r)$ contains all x_{m}. Use Theorem 1 in §13.]

Exercise $3.10.E.5$

Show that $x_{m}=m$ tends to $+\infty$ in E^{*}. Does it contradict Corollary 5?

Exercise $3.10.E.6$

Show that E^{1} is a perfect set $in E^{1}: E^{1}=\left(E^{1}\right)^{\prime}$. Is E^{1} a perfect set in E^{*}? Why?

Exercise $3.10.E.7$

$\Rightarrow 7.$ Review Problems 2 and 4 of Chapter 2, §13. (Do them if not done before.)

Exercise $3.10.E.8$

Verify Examples (f) and (h).

Exercise $3.10.E.9$

Explain Example (i) in detail.

Exercise $3.10.E.10$

In the following cases find the set $A^{\prime}$ of all cluster points of A in E^{1}. Is $A^{\prime} \subseteq A$? Is $A \subseteq A^{\prime}$? Is A perfect? Give a precise proof.
(a) A consists of all points of the form

$$\frac{1}{n} \text{ and } 1+\frac{1}{n}, \quad n=1,2,\ldots; \tag{3.10.E.3}$$

i.e., A is the sequence

$$\left\{1,2,\frac{1}{2},1\frac{1}{2},\ldots,\frac{1}{n},1+\frac{1}{n},\ldots\right\}. \tag{3.10.E.4}$$

(b) A is the set of all rationals in $(0,1)$. Answer: $A'=[0,1]$. Why?
(c) A is the union of the intervals

$$\left[\frac{2n}{2n+1},\frac{2n+1}{2n+2}\right], \quad n=0,1,2,\ldots \tag{3.10.E.5}$$

(d) A consists of all points of the form

$$2^{-n} \text{ and } 2^{-n}+2^{-n-k}, \quad n,k\in N. \tag{3.10.E.6}$$

? Exercise $3.10.E.11$

Can a sequence $\{x_m\}\subseteq E^1$ cluster at each $p\in E^1$?
[Hint: See Example (e).]

? Exercise $3.10.E.12$

Prove that if

$$p=\sup A \text{ or } p=\inf A \text{ in } E^1 \tag{3.10.E.7}$$

$\left(\emptyset\neq A\subseteq E^1\right)$, and if $p\notin A$, then p is a cluster point of A.
[Hint: Take $G_p(\varepsilon)=(p-\varepsilon,p+\varepsilon)$. Use Theorem 2 of Chapter 2, §§8-9.]

? Exercise $3.10.E.13$

Prove that a set $A\subseteq(S,\rho)$ clusters at p iff every neighborhood of p (see §12, Definition 1) contains infinitely many points of A; similarly for sequences. How about convergence? State it in terms of cubic neighborhoods in E^n.

? Exercise $3.10.E.14$

Discuss Example (h) for nondegenerate intervals in E^n. Give a proof.

? Exercise $3.10.E.15$

Prove that a set $A\neq\emptyset$ clusters at $p(p\notin A)$ iff $\rho(p,A)=0$. (See §13, Note $3.$)

? Exercise $3.10.E.16$

Show that in $E^n(*\text{ and in any other normed space }\neq\{\overline{0}\})$, the cluster points of any globe $G_{\bar{p}}(\varepsilon)$ form exactly the closed globe $\overline{G}_{\bar{p}}(\varepsilon)$, and that $\overline{G}_{\bar{p}}(\varepsilon)$ is perfect. Is this true in other spaces? (Consider a discrete space!)

Exercise $3.10.E.17$

(Cantor's set.) Remove from $[0,1]$ the open middle third

$$\left(\frac{1}{3}, \frac{2}{3}\right). \tag{3.10.E.8}$$

From the remaining closed intervals

$$\left[0, \frac{1}{3}\right] \text{ and } \left[\frac{2}{3}, 1\right], \tag{3.10.E.9}$$

remove their open middles,

$$\left(\frac{1}{9}, \frac{2}{9}\right) \text{ and } \left(\frac{7}{9}, \frac{8}{9}\right). \tag{3.10.E.10}$$

Do the same with the remaining four closed intervals, and so on, ad infinitum. The set P which remains after all these (infinitely many) removals is called Cantor's set.
Show that P is perfect.
[Hint: If $p \notin P$, then either p is in one of the removed open intervals, or $p \notin [0,1]$. In both cases, p is no cluster point of P. (Why?) Thus no p outside P is a cluster point.
On the other hand, if $p \in P$, show that any $G_p(\varepsilon)$ contains infinitely many endpoints of removed open intervals, all in P; thus $p \in P'$. Deduce that $P = P'$]

3.11: Operations on Convergent Sequences

This page is a draft and is under active development.

Sequences in E^{1} and C can be added and multiplied termwise; for example, adding $\{x_{m}\}$ and $\{y_{m}\}$, one obtains the sequence with general term $x_{m}+y_{m}$. This leads to important theorems, valid also for E^{n}(* and other normed spaces). Theorem 1 below states, roughly, that the limit of the sum $\{x_{m}+y_{m}\}$ equals the sum of lim x_{m} and lim y_{m} (if these exist), and similarly for products and quotients (when they are defined).

Theorem 3.11.1

Let $x_{m}\rightarrow q, y_{m}\rightarrow r,$ and $a_{m}\rightarrow a$ in E^{1} or C (the complex field). Then

(i) $x_{m}\pm y_{m}\rightarrow q\pm r$;

(ii) $a_{m}x_{m}\rightarrow aq$;

(iii) $\frac{x_{m}}{a_{m}}\rightarrow\frac{q}{a}$ if $a\neq 0$ and for all $m\geq 1, a_{m}\neq 0$.

This also holds if the $x_{m}, y_{m}, q,$ and r are vectors in E^{n} ("or in another normed space), while the a_{m} and a are scalars for that space.

Proof

(i) By formula (2) of §14, we must show that

$$(\forall\varepsilon>0)(\exists k)(\forall m>k)\quad|x_{m}\pm y_{m}-(q\pm r)|<\varepsilon. \tag{3.11.1}$$

Thus we fix an arbitrary $\varepsilon>0$ and look for a suitable k. since $x_{m}\rightarrow q$ and $y_{m}\rightarrow r,$ there are k' and k'' such that

$$(\forall m>k')\quad|x_{m}-q|<\frac{\varepsilon}{2} \tag{3.11.2}$$

and

$$(\forall m>k'')\quad|y_{m}-r|<\frac{\varepsilon}{2} \tag{3.11.3}$$

(as ε is arbitrary, we may as well replace it by $\frac{1}{2}\varepsilon$). Then both inequalities hold for $m>k, k=\max(k',k'')$. Adding them, we obtain

$$(\forall m>k)\quad|x_{m}-q|+|y_{m}-r|<\varepsilon. \tag{3.11.4}$$

Hence by the triangle law,

$$|x_{m}-q\pm(y_{m}-r)|<\varepsilon,\text{ i.e., }|x_{m}\pm y_{m}-(q\pm r)|<\varepsilon\text{ for }m>k, \tag{3.11.5}$$

as required. □

This proof of (i) applies to sequences of vectors as well, without any change.

The proof of (ii) and (iii) is sketched in Problems 1-4 below.

Note 1. By induction, parts (i) and (ii) hold for sums and products of any finite (but fixed) number of suitable convergent sequences.

Note 2. The theorem does not apply to infinite limits q, r, a.

Note 3. The assumption $a\neq 0$ in Theorem 1(iii) is important. It ensures not only that q/a is defined but also that at most finitely many a_{m} can vanish (see Problem 3). Since we may safely drop a finite number of terms (see Note 2 in §14), we can achieve that noa_{m} is 0, so that x_{m}/a_{m} is defined. It is with this understanding that part (iii) of the theorem has been formulated. The next two theorems are actually special cases of more general propositions to be proved in Chapter 4, §§3 and 5. Therefore, we only state them here, leaving the proofs as exercises, with some hints provided.

Theorem 3.11.2

(componentwise convergence). We have $\bar{x}_m \to \bar{p}$ in E^n $({}^*C^n)$ iff each of the n components of $\bar{x}_m$ tends to the corresponding component of $\bar{p}$, i.e., iff $x_{mk} \to p_k, k = 1, 2, \ldots, n,$ in $E^1(C)$. (See Problem 8 for hints.)

Theorem 3.11.3

Every monotone sequence $\{x_n\} \subseteq E^*$ has a finite or infinite limit, which equals sup_ $_n x_n$ if $\{x_n\} \uparrow$ and inf $_n x_n$ if $\{x_n\} \downarrow$. If $\{x_n\}$ is monotone and bounded in E^1, its limit is finite $(by$ Corollary 1 of Chapter 2, §13).

The proof was requested in Problem 9 of Chapter 2, §13. See also Chapter 4, §5, Theorem 1. An important application is the following.

Example 3.11.1

(the number e).

Let $x_n = \left(1+\frac{1}{n}\right)^n$ in E^1. By the binomial theorem,

$$\begin{aligned} x_n &= 1+1+\frac{n(n-1)}{2!n^2}+\frac{n(n-1)(n-2)}{3!n^3}+\cdots+\frac{n(n-1)\cdots(n-(n-1))}{n!n^n} \\ &= 2+\left(1-\frac{1}{n}\right)\frac{1}{2!}+\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\frac{1}{3!}+\cdots+\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-1}{n}\right)\frac{1}{n!} \end{aligned}$$

If n is replaced by $n+1$, all terms in this expansion increase, as does their number. Thus $x_n < x_{n+1}$, i.e., $\{x_n\} \uparrow$. Moreover, for $n > 1$,

$$\begin{aligned} 2 < x_n &< 2+\frac{1}{2!}+\cdots+\frac{1}{n!} \leq 2+\frac{1}{2}+\cdots+\frac{1}{2^{n-1}} \\ &= 2+\frac{1}{2}\left(1+\cdots+\frac{1}{2^{n-2}}\right) = 2+\frac{1}{2}\frac{1-\left(\frac{1}{2}\right)^{n-1}}{\frac{1}{2}} < 2+1 = 3 \end{aligned}$$

Thus $2 < x_n < 3$ for $n > 1$. Hence $2 < \sup_n x_n \leq 3$; and by Theorem 3, $\sup_n x_n = \lim x_n$. This limit, denoted by e, plays an important role in analysis. It can be shown that it is irrational, and (to within 10^{-20}) $e = 2.71828182845904523536..$ In any case,

$$2 < e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \leq 3. \tag{3.11.6}$$

The following corollaries are left as exercises for the reader.

corollary 3.11.1

Suppose $\lim x_m = p$ and $\lim y_m = q$ exist in E^*.

(a) If $p > q$, then $x_m > y_m$ for all but finitely many m.

(b) If $x_m \leq y_m$ for infinitely many m, then $p \leq q$; i.e., $\lim x_m \leq \lim y_m$.

This is known as passage to the limit in inequalities. Caution: The strict inequalities $x_m < y_m$ do not imply $p < q$ but only $p \leq q$. For example, let

$$x_m = \frac{1}{m} \text{ and } y_m = 0. \tag{3.11.7}$$

Then

$$(\forall m) \quad x_m > y_m; \tag{3.11.8}$$

yet $\lim x_m = \lim y_m = 0$.

corollary 3.11.2

Let $x_m \rightarrow p$ in E^*, and let $c \in E^*$ (finite or not). Then the following are true:

(a) If $p > c$ (respectively,$p < c$), we have $x_m > c (x_m < c)$ for all but finitely many m.

(b) If $x_m \leq c$ (respectively, $x_m \geq c$) for infinitely many m, then $p \leq c$ $(p \geq c)$.

One can prove this from Corollary 1, with $y_m = c$ (or $x_m = c$) for all m.

corollary 3.11.3

(rule of intermediate sequence). If $x_m \rightarrow p$ and $y_m \rightarrow p$ in E^* and if $x_m \leq z_m \leq y_m$ for all but finitely many m, then also $z_m \rightarrow p$.

Theorem 3.11.4

(continuity of the distance function). If

$$x_m \rightarrow p \text{ and } y_m \rightarrow q \text{ in a metric space } (S, \rho), \tag{3.11.9}$$

then

$$\rho\left(x_m, y_m\right) \rightarrow \rho(p, q) \text{ in } E^1. \tag{3.11.10}$$

Proof

Hint: Show that

$$\left|\rho\left(x_m, y_m\right) - \rho(p, q)\right| \leq \rho\left(x_m, p\right) + \rho\left(q, y_m\right) \rightarrow 0 \tag{3.11.11}$$

by Theorem 1.

3.11.E: Problems on Limits of Sequences (Exercises)

See also Chapter 2, §13.

Exercise $3.11.E.1$

Prove that if $x_m \rightarrow 0$ and if $\{a_m\}$ is bounded in E^1 or C, then

$$a_m x_m \rightarrow 0. \tag{3.11.E.1}$$

This is true also if the x_m are vectors and the a_m are scalars (or vice versa).
[Hint: If $\{a_m\}$ is bounded, there is a $K \in E^1$ such that

$$(\forall m) \quad |a_m| < K. \tag{3.11.E.2}$$

As $x_m \rightarrow 0$,

$$(\forall \varepsilon > 0)(\exists k)(\forall m > k) \quad |x_m| < \frac{\varepsilon}{K} (\text{why?}), \tag{3.11.E.3}$$

so $|a_m x_m| < \varepsilon.]$

Exercise $3.11.E.2$

Prove Theorem $1(\text{ ii })$.
[Hint: By Corollary 2(ii)(iii) in §14, we must show that $a_m x_m - aw \rightarrow 0$. Now

$$a_m x_m - aq = a_m (x_m - q) + (a_m - a) q. \tag{3.11.E.4}$$

where $x_m - q \rightarrow 0$ and $a_m - a \rightarrow 0$ by Corollary 2 of §14. Hence by Problem 1,

$$a_m (x_m - q) \rightarrow 0 \text{ and } (a_m - a) q \rightarrow 0 \tag{3.11.E.5}$$

(treat q as a constant sequence and use Corollary 5 in §14). Now apply Theorem $1(\mathrm{i}).]$

Exercise $3.11.E.3$

Prove that if $a_m \rightarrow a$ and $a \neq 0$ in E^1 or C, then

$$(\exists \varepsilon > 0)(\exists k)(\forall m > k) \quad |a_m| \geq \varepsilon. \tag{3.11.E.6}$$

(We briefly say that the a_m are bounded away from 0, for $m > k.)$ Hence prove the boundedness of $\left\{\frac{1}{a_m}\right\}$ for $m > k$.
[Hint: For the first part, proceed as in the proof of Corollary 1 in $\S 14,$ with $x_m = a_m, p = a,$ and $q = 0$.
For the second part, the inequalities

$$(\forall m > k) \quad \left|\frac{1}{a_m}\right| \leq \frac{1}{\varepsilon} \tag{3.11.E.7}$$

lead to the desired result. $]$

Exercise $3.11.E.4$

Prove that if $a_m \rightarrow a \neq 0$ in E^1 or C, then

$$\frac{1}{a_m} \rightarrow \frac{1}{a}. \tag{3.11.E.8}$$

Use this and Theorem $1($ ii$)$ to prove Theorem $1($ iii$)$, noting that

$$\frac{x_m}{a_m} = x_m \cdot \frac{1}{a_m}. \tag{3.11.E.9}$$

[Hint: Use Note 3 and Problem 3 to find that

$$(\forall m > k) \quad \left|\frac{1}{a_m} - \frac{1}{a}\right| = \frac{1}{|a|}\left|a_m - a\right| \frac{1}{\left|a_m\right|}, \tag{3.11.E.10}$$

where $\left\{\frac{1}{a_m}\right\}$ is bounded and $\frac{1}{|a|}\left|a_m - a\right| \rightarrow 0.$ (Why?)
Hence, by Problem $1, \left|\frac{1}{a_m} - \frac{1}{a}\right| \rightarrow 0.$ Proceed.]

Exercise $3.11.E.5$

Prove Corollaries 1 and 2 in two ways:
(i) Use Definition 2 of Chapter 2, §13 for Corollary $1(a)$, treating infinite limits separately; then prove (b) by assuming the opposite and exhibiting a contradiction to (a).
(ii) Prove (b) first by using Corollary 2 and Theorem 3 of Chapter 2, §13; then deduce (a) by contradiction.

Exercise $3.11.E.6$

Prove Corollary 3 in two ways (cf. Problem 5).

Exercise $3.11.E.7$

Prove Theorem 4 as suggested, and also without using Theorem $1(\mathrm{i})$.

Exercise $3.11.E.8$

Prove Theorem 2.
[Hint: If $\bar{x}_m \rightarrow \bar{p},$ then

$$(\forall \varepsilon > 0)(\exists q)(\forall m > q) \quad \varepsilon > \left|\bar{x}_m - \bar{p}\right| \geq \left|x_{mk} - p_k\right|. \quad \text{(Why?)} \tag{3.11.E.11}$$

Thus by definition $x_{mk} \rightarrow p_k, k = 1, 2, \ldots, n.$
Conversely, if so, use Theorem $1(\mathrm{i})($ ii $)$ to obtain

$$\sum_{k=1}^n x_{mk} \vec{e}_k \rightarrow \sum_{k=1}^n p_k \vec{e}_k, \tag{3.11.E.12}$$

with $\vec{e}_k$ as in Theorem 2 of §§1-3].

Exercise $3.11.E.8'$

In Problem 8, prove the converse part from definitions. (Fix $\varepsilon>0$, etc.)

Exercise $3.11.E.9$

Find the following limits in E^1, in two ways: (i) using Theorem 1, justifying each step; (ii) using definitions only.

$$\begin{array}{ll}\text{(a) } \lim_{m\to\infty}\frac{m+1}{m}; & \text{(b) } \lim_{m\to\infty}\frac{3m+2}{2m-1}\\ \text{(c) } \lim_{n\to\infty}\frac{1}{1+n^2}; & \text{(d) } \lim_{n\to\infty}\frac{n(n-1)}{1-2n^2}\end{array} \tag{3.11.E.13}$$

[Solution of (a) by the first method: Treat

$$\frac{m+1}{m}=1+\frac{1}{m} \tag{3.11.E.14}$$

as the sum of $x_m=1$ (constant) and

$$y_m=\frac{1}{m}\rightarrow 0 \text{ (proved in §14).} \tag{3.11.E.15}$$

Thus by Theorem 1(i),

$$\frac{m+1}{m}=x_m+y_m\rightarrow 1+0=1. \tag{3.11.E.16}$$

Second method: Fix $\varepsilon>0$ and find k such that

$$(\forall m>k)\quad\left|\frac{m+1}{m}-1\right|<\varepsilon. \tag{3.11.E.17}$$

Solving for m, show that this holds if $m>\frac{1}{\varepsilon}$. Thus take an integer $k>\frac{1}{\varepsilon}$, so

$$(\forall m>k)\quad\left|\frac{m+1}{m}-1\right|<\varepsilon. \tag{3.11.E.18}$$

Caution: One cannot apply Theorem 1 (iii) directly, treating $(m+1)/m$ as the quotient of $x_m=m+1$ and $a_m=m$, because x_m and a_m diverge in E^1. (Theorem 1 does not apply to infinite limits.) As a remedy, we first divide the numerator and denominator by a suitable power of m(or n).]

Exercise $3.11.E.10$

Prove that

$$|x_m|\rightarrow+\infty \text{ in } E^* \text{ iff } \frac{1}{x_m}\rightarrow 0\quad(x_m\neq 0). \tag{3.11.E.19}$$

Exercise $3.11.E.11$

Prove that if

$$x_m\rightarrow+\infty \text{ and } y_m\rightarrow q\neq-\infty \text{ in } E^*, \tag{3.11.E.20}$$

then

$$x_m + y_m \to +\infty. \tag{3.11.E.21}$$

This is written symbolically as

$$\text{" } +\infty + q = +\infty \text{ if } q \neq -\infty. \text{ "} \tag{3.11.E.22}$$

Do also

$$\text{" } -\infty + q = -\infty \text{ if } q \neq +\infty. \text{ "} \tag{3.11.E.23}$$

Prove similarly that

$$\text{" } (+\infty) \cdot q = +\infty \text{ if } q > 0 \text{ "} \tag{3.11.E.24}$$

and

$$\text{" } (+\infty) \cdot q = -\infty \text{ if } q < 0. \text{ "} \tag{3.11.E.25}$$

[Hint: Treat the cases $q \in E^1, q = +\infty,$ and $q = -\infty$ separately. Use definitions.]

? Exercise $3.11.E.12$

Find the limit (or $\underline{\lim}$ and $\overline{\lim}$) of the following sequences in E^* :
(a) $x_n = 2 \cdot 4 \cdots 2n = 2^n n!$;
(b) $x_n = 5n - n^3$;
(c) $x_n = 2n^4 - n^3 - 3n^2 - 1$;
(d) $x_n = (-1)^n n!$;
(e) $x_n = \frac{(-1)^n}{n!}$.
[Hint for (b) : $x_n = n\left(5 - n^2\right)$; use Problem 11.]

? Exercise $3.11.E.13$

Use Corollary 4 in §14, to find the following:
(a) $\lim_{n \to \infty} \frac{(-1)^n}{1+n^2}$;
(b) $\lim_{n \to \infty} \frac{1-n+(-1)^n}{2n+1}$.

? Exercise $3.11.E.14$

Find the following.
(a) $\lim_{n \to \infty} \frac{1+2+\cdots+n}{n^2}$;
(b) $\lim_{n \to \infty} \sum_{k=1}^n \frac{k^2}{n^3+1}$;
(c) $\lim_{n \to \infty} \sum_{k=1}^n \frac{k^3}{n^4-1}$.
[Hint: Compute $\sum_{k=1}^n k^m$ using Problem 10 of Chapter 2, §§5-6.]
What is wrong with the following "solution" of (a) : $\frac{1}{n^2} \to 0, \frac{2}{n^2} \to 0,$ etc.; hence the limit is 0?

Exercise $3.11.E.15$

For each integer $m \geq 0$, let

$$S_{mn} = 1^m + 2^m + \cdots + n^m. \tag{3.11.E.26}$$

Prove by induction on m that

$$\lim_{n\to\infty} \frac{S_{mn}}{(n+1)^{m+1}} = \frac{1}{m+1}. \tag{3.11.E.27}$$

[Hint: First prove that

$$(m+1)S_{mn} = (n+1)^{m+1} - 1 - \sum_{i=0}^{m-1} \binom{m+1}{i} S_{mi} \tag{3.11.E.28}$$

by adding up the binomial expansions of $(k+1)^{m+1}, k = 1, \ldots, n.$]

Exercise $3.11.E.16$

Prove that

$$\lim_{n\to\infty} q^n = +\infty \text{ if } q > 1; \quad \lim_{n\to\infty} q^n = 0 \text{ if } |q| < 1; \quad \lim_{n\to\infty} 1^n = 1. \tag{3.11.E.29}$$

[Hint: If $q > 1$, put $q = 1 + d, d > 0$. By the binomial expansion,

$$q^n = (1+d)^n = 1 + nd + \cdots + d^n > nd \rightarrow +\infty. \quad \text{(Why?)} \tag{3.11.E.30}$$

If $|q| < 1$, then $\left|\frac{1}{q}\right| > 1$; so $\lim \left|\frac{1}{q}\right|^n = +\infty$; use Problem 10.]

Exercise $3.11.E.17$

Prove that

$$\lim_{n\to\infty} \frac{n}{q^n} = 0 \text{ if } |q| > 1, \text{ and } \lim_{n\to\infty} \frac{n}{q^n} = +\infty \text{ if } 0 < q < 1. \tag{3.11.E.31}$$

[Hint: If $|q| > 1$, use the binomial as in Problem 16 to obtain

$$|q|^n > \frac{1}{2} n(n-1) d^2, n \geq 2, \text{ so } \frac{n}{|q|^n} < \frac{2}{(n-1)d^2} \rightarrow 0. \tag{3.11.E.32}$$

Use Corollary 3 with

$$x_n = 0, |z_n| = \frac{n}{|q|^n}, \text{ and } y_n = \frac{2}{(n-1)d^2} \tag{3.11.E.33}$$

to get $|z_n| \rightarrow 0$; hence also $z_n \rightarrow 0$ by Corollary 2(iii) of §14. In case $0 < q < 1$, use 10.]

Exercise 3.11.E.18

Let $r, a \in E^1$. Prove that

$$\lim_{n\to\infty} n^r a^{-n} = 0 \text{ if } |a| > 1. \tag{3.11.E.34}$$

[Hint: If $r > 1$ and $a > 1$, use Problem 17 with $q = a^{1/r}$ to get $na^{-n/r} \to 0$. As

$$0 < n^r a^{-n} = \left(na^{-n/r}\right)^r \leq na^{-n/r} \to 0, \tag{3.11.E.35}$$

obtain $n^r a^{-n} \to 0$.
If $r < 1$, then $n^r a^{-n} < na^{-n} \to 0$. What if $a < -1$?]

Exercise 3.11.E.19

(Geometric series.) Prove that if $|q| < 1$, then

$$\lim_{n\to\infty}\left(a + aq + \cdots + aq^{n-1}\right) = \frac{a}{1-q}. \tag{3.11.E.36}$$

[Hint:

$$a\left(1 + q + \cdots + q^{n-1}\right) = a\frac{1-q^n}{1-q}, \tag{3.11.E.37}$$

where $q^n \to 0$, by Problem 16.]

Exercise 3.11.E.20

Let $0 < c < +\infty$. Prove that

$$\lim_{n\to\infty} \sqrt[n]{c} = 1. \tag{3.11.E.38}$$

[Hint: If $c > 1$, put $\sqrt[n]{c} = 1 + d_n, d_n > 0$. Expand $c = (1 + d_n)^n$ to show that

$$0 < d_n < \frac{c}{n} \to 0, \tag{3.11.E.39}$$

so $d_n \to 0$ by Corollary 3.]

Exercise 3.11.E.21

Investigate the following sequences for monotonicity, $\underline{\lim}$, $\overline{\lim}$, and lim. (In each case, find suitable formula, or formulas, for the general term.)

(a) $2, 5, 10, 17, 26, \ldots;$
(b) $2, -2, 2, -2, \ldots;$
(c) $2, -2, -6, -10, -14, \ldots;$
(d) $1, 1, -1, -1, 1, 1, -1, -1, \ldots;$
(e) $\frac{3\cdot 2}{1}, \frac{4\cdot 6}{4}, \frac{5\cdot 10}{9}, \frac{6\cdot 14}{16}, \ldots.$

Exercise $3.11.E.22$

Do Problem 21 for the following sequences.
(a) $\frac{1}{2\cdot3}, \frac{-8}{3\cdot4}, \frac{27}{4\cdot5}, \frac{-64}{5\cdot6}, \frac{125}{6\cdot7}, \ldots;$
(b) $\frac{2}{9}, -\frac{5}{9}, \frac{8}{9}, -\frac{13}{9}, \ldots;$
(c) $\frac{2}{3}, -\frac{2}{5}, \frac{4}{7}, -\frac{4}{9}, \frac{6}{11}, -\frac{6}{13}, \ldots$
(d) $1,3,5,1,1,3,5,2,1,3,5,3, \ldots, 1,3,5,n, \ldots;$
(e) $0.9, 0.99, 0.999, \ldots;$
(f) $+\infty, 1, +\infty, 2, +\infty, 3, \ldots;$
(g) $-\infty, 1, -\infty, \frac{1}{2}, \ldots, -\infty, \frac{1}{n}, \ldots.$

Exercise $3.11.E.23$

Do Problem 20 as follows: If $c \geq 1, \{\sqrt[n]{c}\} \downarrow$. (Why?) By Theorem 3, $p=\lim_{n\rightarrow\infty}\sqrt[n]{c}$ exists and

$$(\forall n) \quad 1 \leq p \leq \sqrt[n]{c}, \text{ i.e., } 1 \leq p^n \leq c. \tag{3.11.E.40}$$

By Problem 16, p cannot be >1, so $p=1$.
In case $0<c<1$, consider $\sqrt[n]{1/c}$ and use Theorem $1(\text{ iii})$.

Exercise $3.11.E.24$

Prove the existence of $\lim x_n$ and find it when x_n is defined inductively by
(i) $x_1=\sqrt{2}, x_{n+1}=\sqrt{2x_n};$
(ii) $x_1=c>0, x_{n+1}=\sqrt{c^2+x_n};$
(iii) $x_1=c>0, x_{n+1}=\frac{cx_n}{n+1};$ hence deduce that $\lim_{n\rightarrow\infty}\frac{c^n}{n!}=0$.
[Hint: Show that the sequences are monotone and bounded in E^1 (Theorem 3).
For example, in (ii) induction yields

$$x_n<x_{n+1}<c+1. \quad (\text{ Verify! }) \tag{3.11.E.41}$$

Thus $\lim x_n = \lim x_{n+1} = p$ exists. To find p, square the equation

$$x_{n+1}=\sqrt{c^2+x_n} \quad (\text{ given }) \tag{3.11.E.42}$$

and use Theorem 1 to get

$$p^2=c^2+p. \quad (\text{Why?}) \tag{3.11.E.43}$$

Solving for p (noting that $p>0$), obtain

$$p=\lim x_n=\frac{1}{2}\left(1+\sqrt{4c^2+1}\right); \tag{3.11.E.44}$$

similarly in cases (i) and (iii).]

Exercise $3.11.E.25$

Find $\lim x_n$ in E^1 or E^* (if any), given that
(a) $x_n=(n+1)^q-n^q, 0<q<1;$
(b) $x_n=\sqrt{n}(\sqrt{n+1}-\sqrt{n});$

(c) $x_n = \frac{1}{\sqrt{n^2+k}}$;
(d) $x_n = n(n+1)c^n$, with $|c| < 1$;
(e) $x_n = \sqrt[n]{\sum_{k=1}^m a_k^n}$, with $a_k > 0$;
(f) $x_n = \frac{3\cdot 5\cdot 7\cdots(2n+1)}{2\cdot 5\cdot 8\cdots(3n-1)}$.
[Hints:
(a) $0 < x_n = n^q\left[\left(1+\frac{1}{n}\right)^q - 1\right] < n^q\left(1+\frac{1}{n}-1\right) = n^{q-1} \to 0$. (Why?)
(b) $x_n = \frac{1}{1+\sqrt{1+1/n}}$, where $1 < \sqrt{1+\frac{1}{n}} < 1+\frac{1}{n} \to 1$, so $x_n \to \frac{1}{2}$. (Why?)
(c) Verify that

$$\frac{n}{\sqrt{n^2+n}} \leq x_n \leq \frac{n}{\sqrt{n^2+1}}, \tag{3.11.E.45}$$

so $x_n \to 1$ by Corollary 3. (Give a proof.)
(d) See Problems 17 and 18.
(e) Let $a = \max(a_1, \ldots, a_m)$. Prove that $a \leq x_n \leq a\sqrt[n]{m}$. Use Problem 20.]
The following are some harder but useful problems of theoretical importance.
The explicit hints should make them not too hard.

Exercise $3.11.E.26$

Let $\{x_n\} \subseteq E^1$. Prove that if $x_n \to p$ in E^1, then also

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n x_i = p \tag{3.11.E.46}$$

(i.e., p is also the limit of the sequence of the arithmetic means of the x_n).
[Solution: Fix $\varepsilon > 0$. Then

$$(\exists k)(\forall n > k) \quad p - \frac{\varepsilon}{4} < x_n < p + \frac{\varepsilon}{4}. \tag{3.11.E.47}$$

Adding $n-k$ inequalities, get

$$(n-k)\left(p-\frac{\varepsilon}{4}\right) < \sum_{i=k+1}^n x_i < (n-k)\left(p+\frac{\varepsilon}{4}\right). \tag{3.11.E.48}$$

With k so fixed, we thus have

$$(\forall n > k) \quad \frac{n-k}{n}\left(p-\frac{\varepsilon}{4}\right) < \frac{1}{n}\left(x_{k+1}+\cdots+x_n\right) < \frac{n-k}{n}\left(p+\frac{\varepsilon}{4}\right). \tag{3.11.E.49}$$

Here, with k and ε fixed,

$$\lim_{n\to\infty} \frac{n-k}{n}\left(p-\frac{\varepsilon}{4}\right) = p - \frac{\varepsilon}{4}. \tag{3.11.E.50}$$

Hence, as $p - \frac{1}{2}\varepsilon < p - \frac{1}{4}\varepsilon$, there is k' such that

$$(\forall n > k') \quad p - \frac{\varepsilon}{2} < \frac{n-k}{n}\left(p-\frac{\varepsilon}{4}\right). \tag{3.11.E.51}$$

Similarly,

$$(\exists k'')\,(\forall n>k'')\quad \frac{n-k}{n}\left(p+\frac{\varepsilon}{4}\right)<p+\frac{\varepsilon}{2}. \tag{3.11.E.52}$$

Combining this with (i), we have, for $K'=\max\left(k,k',k''\right)$,

$$(\forall n>K')\quad p-\frac{\varepsilon}{2}<\frac{1}{n}(x_{k+1}+\cdots+x_n)<p+\frac{\varepsilon}{2}. \tag{3.11.E.53}$$

Now with k fixed,

$$\lim_{n\to\infty}\frac{1}{n}(x_1+x_2+\cdots+x_k)=0. \tag{3.11.E.54}$$

Hence

$$(\exists K'')\,(\forall n>K'')\quad -\frac{\varepsilon}{2}<\frac{1}{n}(x_1+\cdots+x_k)<\frac{\varepsilon}{2}. \tag{3.11.E.55}$$

Let $K=\max\left(K',K''\right)$. Then combining with (ii), we have

$$(\forall n>K)\quad p-\varepsilon<\frac{1}{n}(x_1+\cdots+x_n)<p+\varepsilon, \tag{3.11.E.56}$$

and the result follows.

Exercise $3.11.E.26'$

Show that the result of Problem 26 holds also for infinite limits $p=\pm\infty\in E^*$.

Exercise $3.11.E.27$

Prove that if $x_n\to p$ in $E^*\left(x_n>0\right)$, then

$$\lim_{n\to\infty}\sqrt[n]{x_1x_2\cdots x_n}=p. \tag{3.11.E.57}$$

[Hint: Let first $0<p<+\infty$. Given $\varepsilon>0$, use density to fix $\delta>1$ so close to 1 that

$$p-\varepsilon<\frac{p}{\delta}<p<p\delta<p+\varepsilon. \tag{3.11.E.58}$$

As $x_n\to p$,

$$(\exists k)(\forall n>k)\quad \frac{p}{\sqrt[4]{\delta}}<x_n<p\sqrt[4]{\delta}. \tag{3.11.E.59}$$

Continue as in Problem 26, replacing ε by δ, and multiplication by addition (also subtraction by division, etc., as shown above). Find a similar solution for the case $p=+\infty$. Note the result of Problem 20.]

Exercise $3.11.E.28$

Disprove by counterexamples the converse implications in Problems 26 and 27. For example, consider the sequences

$$1,-1,1,-1, \ldots \tag{3.11.E.60}$$

and

$$\frac{1}{2}, 2, \frac{1}{2}, 2, \frac{1}{2}, 2, \ldots \tag{3.11.E.61}$$

Exercise $3.11.E.29$

Prove the following.
(i) If $\left\{x_n\right\} \subset E^1$ and $\lim _{n \rightarrow \infty}\left(x_{n+1}-x_n\right)=p$ in E^*, then $\frac{x_n}{n} \rightarrow p$.
(ii) If $\left\{x_n\right\} \subset E^1\left(x_n>0\right)$ and if $\frac{x_{n+1}}{x_n} \rightarrow p \in E^*$, then $\sqrt[n]{x_n} \rightarrow p$.
Disprove the converse statements by counterexamples.
[Hint: For (i), let $y_1=x_1$ and $y_n=x_n-x_{n-1}, n=2,3, \ldots$ Then $y_n \rightarrow p$ and

$$\frac{1}{n} \sum_{i=1}^n y_i=\frac{x_n}{n}, \tag{3.11.E.62}$$

so Problems 26 and $26^{\prime}$ apply.
For (ii), use Problem 27. See Problem 28 for examples.]

Exercise $3.11.E.30$

From Problem 29 deduce that
(a) $\lim _{n \rightarrow \infty} \sqrt[n]{n!}=+\infty$;
(b) $\lim _{n \rightarrow \infty} \frac{n+1}{n!}=0$;
(c) $\lim _{n \rightarrow \infty} \sqrt[n]{\frac{n^n}{n!}}=e$;
(d) $\lim _{n \rightarrow \infty} \frac{1}{n} \sqrt[n]{n!}=\frac{1}{e}$;
(e) $\lim _{n \rightarrow \infty} \sqrt[n]{n}=1$.

Exercise $3.11.E.31$

Prove that

$$\lim _{n \rightarrow \infty} x_n=\frac{a+2 b}{3}, \tag{3.11.E.63}$$

given

$$x_0=a, x_1=b, \text { and } x_{n+2}=\frac{1}{2}\left(x_n+x_{n+1}\right). \tag{3.11.E.64}$$

[Hint: Show that the differences $d n=x_n-x_{n-1}$ form a geometric sequence, with ratio $q=-\frac{1}{2}$, and $x_n=a+\sum_{k=1}^n d_k$. Then use the result of Problem 19.]

Exercise 3.11.E.32

⇒ 32. For any sequence $\{x_n\} \subseteq E^1$, prove that

$$\underline{\lim} x_n \leq \underline{\lim} \frac{1}{n} \sum_{i=1}^{n} x_i \leq \overline{\lim} \frac{1}{n} \sum_{i=1}^{n} x_i \leq \overline{\lim} x_n. \tag{3.11.E.65}$$

Hence find a new solution of Problems 26 and $26'$.
[Proof for $\overline{\lim}$: Fix any $k \in N$. Put

$$c = \sum_{i=1}^{k} x_i \text{ and } b = \sup_{i \geq k} x_i. \tag{3.11.E.66}$$

Verify that

$$(\forall n > k) \quad x_{k+1} + x_{k+2} + \cdots + x_n \leq (n-k)b. \tag{3.11.E.67}$$

Add c on both sides and divide by n to get

$$(\forall n > k) \quad \frac{1}{n} \sum_{i=1}^{n} x_i \leq \frac{c}{n} + \frac{n-k}{n} b. \tag{3.11.E.68}$$

Now fix any $\varepsilon > 0$, and first let $|b| < +\infty$. As $\frac{c}{n} \rightarrow 0$ and $\frac{n-k}{n} b \rightarrow b$, there is $n_k > k$ such that

$$(\forall n > n_k) \quad \frac{c}{n} < \frac{\varepsilon}{2} \text{ and } \frac{n-k}{n} b < b + \frac{\varepsilon}{2}. \tag{3.11.E.69}$$

Thus by (i^*),

$$(\forall n > n_k) \quad \frac{1}{n} \sum_{i=1}^{n} x_i \leq \varepsilon + b. \tag{3.11.E.70}$$

This clearly holds also if $b = \sup_{i \geq k} x_i = +\infty$. Hence also

$$\sup_{n \geq n_k} \frac{1}{n} \sum_{i=1}^{n} x_i \leq \varepsilon + \sup_{i \geq k} x_i. \tag{3.11.E.71}$$

As k and ε were arbitrary, we may let first $k \rightarrow +\infty$, then $\varepsilon \rightarrow 0$, to obtain

$$\overline{\lim} \frac{1}{n} \sum_{i=1}^{n} x_i \leq \lim_{k \rightarrow \infty} \sup_{i \geq k} x_i = \overline{\lim} x_n. \quad (\text{ Explain! })] \tag{3.11.E.72}$$

Exercise 3.11.E.33

⇒ 33. Given $\{x_n\} \subseteq E^1, x_n > 0$, prove that

$$\underline{\lim} x_n \leq \underline{\lim} \sqrt[n]{x_1 x_2 \cdots x_n} \text{ and } \overline{\lim} \sqrt[n]{x_1 x_2 \cdots x_n} \leq \overline{\lim} x_n. \tag{3.11.E.73}$$

Hence obtain a new solution for Problem 27.
[Hint: Proceed as suggested in Problem 32, replacing addition by multiplication.]

Exercise 3.11.E.34

Given $x_n, y_n \in E^1\,(y_n > 0)$, with

$$x_n \rightarrow p \in E^* \text{ and } b_n = \sum_{i=1}^{n} y_i \rightarrow +\infty, \tag{3.11.E.74}$$

prove that

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i} = p. \tag{3.11.E.75}$$

Note that Problem 26 is a special case of Problem 34 (take all $y_n = 1$). [Hint for a finite p: Proceed as in Problem 26. However, before adding the $n-k$ inequalities, multiply by y_i and obtain

$$\left(p - \frac{\varepsilon}{4}\right) \sum_{i=k+1}^{n} y_i < \sum_{i=k+1}^{n} x_i y_i < \left(p + \frac{\varepsilon}{4}\right) \sum_{i=k+1}^{n} y_i. \tag{3.11.E.76}$$

Put $b_n = \sum_{i=1}^{n} y_i$ and show that

$$\frac{1}{b_n} \sum_{i=k+1}^{n} x_i y_i = 1 - \frac{1}{b_n} \sum_{i=1}^{k} x_i y_i, \tag{3.11.E.77}$$

where $b_n \rightarrow +\infty(\text{ by assumption }),$ so

$$\frac{1}{b_n} \sum_{i=1}^{k} x_i y_i \rightarrow 0 \quad (\text{for a fixed } k). \tag{3.11.E.78}$$

Proceed. Find a proof for $p = \pm\infty$.]

Exercise 3.11.E.35

Do Problem 34 by considering $\underline{\lim}$ and $\overline{\lim}$ as in Problem 32.
$\left[\text{Hint: Replace } \frac{c}{n} \text{ by } \frac{c}{b_n}, \text{ where } b_n = \sum_{i=1}^{n} y_i \rightarrow +\infty.\right]$

Exercise 3.11.E.36

Prove that if $u_n, v_n \in E^1$, with $\{v_n\} \uparrow$ (strictly) and $v_n \rightarrow +\infty$, and if

$$\lim_{n \rightarrow \infty} \frac{u_n - u_{n-1}}{v_n - v_{n-1}} = p \quad (p \in E^*), \tag{3.11.E.79}$$

then also

$$\lim_{n \rightarrow \infty} \frac{u_n}{v_n} = p, \tag{3.11.E.80}$$

[Hint: The result of Problem 34, with

$$x_n = \frac{u_n - u_{n-1}}{v_n - v_{n-1}} \text{ and } y_n = v_n - v_{n-1}. \tag{3.11.E.81}$$

leads to the final result.]

Exercise 3.11.E.37

From Problem 36 obtain a new solution for Problem 15. Also prove that

$$\lim_{n\to\infty}\left(\frac{S_{mn}}{n^{m+1}}-\frac{1}{m+1}\right)=\frac{1}{2}. \tag{3.11.E.82}$$

[Hint: For the first part, put

$$u_n=S_{mn} \text{ and } v_n=n^{m+1}. \tag{3.11.E.83}$$

For the second, put

$$u_n=(m+1)S_{mn}-n^{m+1} \text{ and } v_n=n^m(m+1).] \tag{3.11.E.84}$$

Exercise 3.11.E.38

Let $0<a<b<+\infty$. Define inductively: $a_1=\sqrt{ab}$ and $b_1=\frac{1}{2}(a+b)$;

$$a_{n+1}=\sqrt{a_nb_n} \text{ and } b_{n+1}=\frac{1}{2}(a_n+b_n), n=1,2,\ldots \tag{3.11.E.85}$$

Then $a_{n+1}<b_{n+1}$ for

$$b_{n+1}-a_{n+1}=\frac{1}{2}(a_n+b_n)-\sqrt{a_nb_n}=\frac{1}{2}\left(\sqrt{b_n}-\sqrt{a_n}\right)^2>0. \tag{3.11.E.86}$$

Deduce that

$$a<a_n<a_{n+1}<b_{n+1}<b_n<b, \tag{3.11.E.87}$$

so $\{a_n\}\uparrow$ and $\{b_n\}\downarrow$. By Theorem 3, $a_n\to p$ and $b_n\to q$ for some $p,q\in E^1$. Prove that $p=q$, i.e.,

$$\lim a_n=\lim b_n. \tag{3.11.E.88}$$

(This is Gauss's arithmetic-geometric mean of a and b.)
[Hint: Take limits of both sides in $b_{n+1}=\frac{1}{2}(a_n+b_n)$ to get $q=\frac{1}{2}(p+q)$.]

Exercise 3.11.E.39

Let $0<a<b$ in E^1. Define inductively $a_1=a, b_1=b$,

$$a_{n+1}=\frac{2a_nb_n}{a_n+b_n}, \text{ and } b_{n+1}=\frac{1}{2}(a_n+b_n), \quad n=1,2,\ldots \tag{3.11.E.89}$$

Prove that

$$\sqrt{ab}=\lim_{n\to\infty}a_n=\lim_{n\to\infty}b_n. \tag{3.11.E.90}$$

[Hint: Proceed as in Problem 38.]

Exercise $3.11.E.40$

Prove the continuity of dot multiplication, namely, if

$$\bar{x}_n \rightarrow \bar{q} \text{ and } \bar{y}_n \rightarrow \bar{r} \text{ in } E^n \tag{3.11.E.91}$$

(*or in another Euclidean space; see §9), then

$$\bar{x}_n \cdot \bar{y}_n \rightarrow \bar{q} \cdot \bar{r}. \tag{3.11.E.92}$$

3.12: More on Cluster Points and Closed Sets. Density

This page is a draft and is under active development.

I. The notions of cluster point and closed set (§§12, 14) can be characterized in terms of convergent sequences. We start with cluster points.

Theorem 3.12.1

(i) A sequence $\{x_m\} \subseteq (S, \rho)$ clusters at a point $p \in S$ iff it has a subsequence $\{x_{m_n}\}$ converging to p.

(ii) A set $A \subseteq (S, \rho)$ clusters at $p \in S$ iff p is the limit of some sequence $\{x_n\}$ of points of A other than p; if so, the terms x_n can be made distinct.

Proof

(i) If $p = \lim_{n\to\infty} x_{m_n}$, then by definition each globe about p contains all but finitely many x_{m_n}, hence infinitely many x_m. Thus p is a cluster point.

Conversely, if so, consider in particular the globes

$$G_p\left(\frac{1}{n}\right), \quad n = 1, 2, \ldots \tag{3.12.1}$$

By assumption, $G_p(1)$ contains some x_m. Thus fix

$$x_{m_1} \in G_p(1). \tag{3.12.2}$$

Next, choose a term

$$x_{m_2} \in G_p\left(\frac{1}{2}\right) \text{ with } m_2 > m_1. \tag{3.12.3}$$

(Such terms exist since $G_p\left(\frac{1}{2}\right)$ contains infinitely many x_m.) Next, fix

$$x_{m_3} \in G_p\left(\frac{1}{3}\right), \text{ with } m_3 > m_2 > m_1, \tag{3.12.4}$$

and so on.

Thus, step by step (inductively), select a sequence of subscripts

$$m_1 < m_2 < \cdots < m_n < \cdots \tag{3.12.5}$$

that determines a subsequence (see Chapter 1, §8) such that

$$(\forall n) \quad x_{m_n} \in G_p\left(\frac{1}{n}\right), \text{ i.e., } \rho\left(x_{m_n}, p\right) < \frac{1}{n} \rightarrow 0, \tag{3.12.6}$$

whence $\rho\left(x_{m_n}, p\right) \rightarrow 0,$ or $x_{m_n} \rightarrow p.$ (Why?) Thus we have found a subsequence $x_{m_n} \rightarrow p,$ and assertion (i) is proved.

Assertion (ii) is proved quite similarly - proceed as in the proof of Corollary 6 in §§14; the inequalities $m_1 < m_2 < \cdots$ are not needed here. □

Example 3.12.1

(a) Recall that the set R of all rationals clusters at each $p \in E^1$ (§§14, Example (e)). Thus by Theorem 1(ii), each real p is the limit of a sequence of rationals. See also Problem 6 of §§12 for $\overline{p}$ in E^n.

(b) The sequence

$$0,1,0,1,\ldots \tag{3.12.7}$$

has two convergent subsequences,

$$x_{2n}=1\rightarrow 1 \text{ and } x_{2n-1}=0\rightarrow 0. \tag{3.12.8}$$

Thus by Theorem 1(i), it clusters at 0 and 1.

Interpret Example (f) and Problem 10(a) in §14 similarly.

As we know, even infinite sets may have no cluster points (take N in E^1). However, a bounded infinite set or sequence in E^n (*or C^n) must cluster. This important theorem (due to Bolzano and Weierstrass) is proved next.

Theorem 3.12.1 (Bolzano-Weierstrass).

(i) Each bounded infinite set or sequence A in E^n (* or C^n) has at least one cluster point $\overline{p}$ there (possibly outside A.

(ii) Thus each bounded sequence in E^n (* or C^n) has a convergent subsequence.

Proof

Take first a bounded sequence $\{z_m\}\subseteq[a,b]$ in E^1. Let

$$p=\overline{\lim} z_m. \tag{3.12.9}$$

By Theorem 2(i) of Chapter 2, §13, $\{z_m\}$ clusters at p. Moreover, as

$$a\leq z_m\leq b, \tag{3.12.10}$$

we have

$$a\leq\inf z_m\leq p\leq\sup z_m\leq b \tag{3.12.11}$$

by Corollary 1 of Chapter 2, §13. Thus

$$p\in[a,b]\subseteq E^1, \tag{3.12.12}$$

and so $\{z_m\}$ clusters in E^1.

Assertion (ii) now follows - for E^1- by Theorem 1(i) above.

Next, take

$$\{\overline{z}_m\}\subseteq E^2, \overline{z}_m=(x_m,y_m); x_m,y_m\in E^1. \tag{3.12.13}$$

If $\{\overline{z}_m\}$ is bounded, all $\overline{z}_m$ are in some square $[\overline{a},\overline{b}]$. (Why?) Let

$$\overline{a}=(a_1,a_2) \text{ and } \overline{b}=(b_1,b_2). \tag{3.12.14}$$

Then

$$a_1\leq x_m\leq b_1 \text{ and } a_2\leq y_m\leq b_2 \text{ in } E^1. \tag{3.12.15}$$

Thus by the first part of the proof, $\{x_m\}$ has a convergent subsequence

$$x_{m_k}\rightarrow p_1 \text{ for some } p_1\in[a_1,b_1]. \tag{3.12.16}$$

For simplicity, we henceforth write x_m for x_{m_k}, y_m for y_{m_k}, and $\overline{z}_m$ for $\overline{z}_{m_k}$. Thus $\overline{z}_m=(x_m,y_m)$ is now a subsequence, with $x_m\rightarrow p_1$, and $a_2\leq y_m\leq b_2$, as before.

We now reapply this process to $\{y_m\}$ and obtain a subsubsequence

$$y_{m_i}\rightarrow p_2 \text{ for some } p_2\in[a_2,b_2]. \tag{3.12.17}$$

The corresponding terms x_{m_i} still tend to p_1 by Corollary 3 of §14. Thus we have a subsequence

$$\overline{z}_{m_i}=(x_{m_i},y_{m_i})\rightarrow(p_1,p_2) \quad \text{ in } E^2 \tag{3.12.18}$$

by Theorem 2 in §15. Hence $\overline{p}=(p_1,p_2)$ is a cluster point of $\{\bar{z}_m\}$. Note that $\overline{p}\in[\overline{a},\overline{b}]$ (see above). This proves the theorem for sequences in E^2 (hence in C).

The proof for E^n is similar; one only has to take subsequences n times. (*The same applies to C^n with real components replaced by complex ones.)

Now take a bounded infinite set $A\subset E^n\,(^*C^n)$. Select from it an infinite sequence $\{\bar{z}_m\}$ of distinct points (see Chapter 1, §9, Problem 5). By what was shown above, $\{\bar{z}_m\}$ clusters at some point $\overline{p}$, so each $G_{\overline{p}}$ contains infinitely many distinct points $\bar{z}_m\in A$. Thus by definition, A clusters at $\overline{p}$. □

Note 1. We have also proved that if $\{\bar{z}_m\}\subseteq[\overline{a},\overline{b}]\subset E^n$, then $\{\bar{z}_m\}$ has a cluster point in $[\overline{a},\overline{b}]$. (This applies to closed intervals only.)

Note 2. The theorem may fail in spaces other than $E^n\,(^*C^n)$. For example, in a discrete space, all sets are bounded, but no set can cluster.

II. Cluster points are closely related to the following notion.

Definition

The closure of a set $A\subseteq(S,\rho)$, denoted $\overline{A}$, is the union of A and the set of all cluster points of A call it A'. Thus $\overline{A}=A\cup A'$.

Theorem 3.12.1

We have $p\in\overline{A}$ in (S,ρ) iff each globe $G_p(\delta)$ about p meets A, i. e.,

$$(\forall\delta>0)\quad A\cap G_p(\delta)\neq\emptyset. \tag{3.12.19}$$

Equivalently, $p\in\overline{A}$ iff

$$p=\lim_{n\to\infty}x_n \text{ for some } \{x_n\}\subseteq A. \tag{3.12.20}$$

Proof

The proof is as in Corollary 6 of §14 and Theorem 1. (Here, however, the x_n need not be distinct or different from p.) The details are left to the reader.

This also yields the following new characterization of closed sets (cf. §12).

Theorem 3.12.1

A set $A\subseteq(S,\rho)$ is closed iff one of the following conditions holds.

(i) A contains all its cluster points (or has none); i.e., $A\supseteq A'$.

(ii) $A=\overline{A}$.

(iii) A contains the limit of each convergent sequence $\{x_n\}\subseteq A$ (if any).

Proof

Parts (i) and (ii) are equivalent since

$$A\supseteq A'\Longleftrightarrow A=A\cup A'=\overline{A}.\quad\text{(Explain!)} \tag{3.12.21}$$

Now let A be closed. If $p\notin A$, then $p\in-A$; therefore, by Definition 3 in §12, some G_p fails to meet $A\,(G_p\cap A=\emptyset)$. Hence no $p\in-A$ is a cluster point, or the limit of a sequence $\{x_n\}\subseteq A$. (This would contradict Definitions 1 and 2 of §14.) Consequently, all such cluster points and limits must be in A, as claimed.

Conversely, suppose A is not closed, so $-A$ is not open. Then $-A$ has a noninterior point p; i.e., $p \in -A$ but $no G_p$ is entirely in $-A$. This means that each G_p meets A. Thus

$$p \in \overline{A} \text{ (by Theorem 3)}, \tag{3.12.22}$$

and

$$p = \lim_{n\to\infty} x_n \text{ for some } \{x_n\} \subseteq A \text{ (by the same theorem)}, \tag{3.12.23}$$

even though $p \notin A($ for $p \in -A)$.

We see that (iii) and (ii), hence also (i), fail if A is not closed and hold if A is closed. (See the first part of the proof.) Thus the theorem is proved. □

Corollary 1. $\overline{\emptyset} = \emptyset$.

Corollary 2. $A \subseteq B \Longrightarrow \overline{A} \subseteq \overline{B}$.

Corollary 3. $\overline{A}$ is always a closed set $\supseteq A$.

Corollary 4. $\overline{A \cup B} = \overline{A} \cup \overline{B}$ (the closure of $A \cup B$ equals the union of $\overline{A}$ and $\overline{B}$).

III. As we know, the rationals are dense in E^1 (Theorem 3 of Chapter 2, §10). This means that every globe $G_p(\delta) = (p-\delta, p+\delta)$ in E^1 contains rationals. Similarly (see Problem 6 in §12), the set R^n of all rational points is dense in E^n. We now generalize this idea for arbitrary sets in a metric space (S, ρ).

Definition

Given $A \subseteq B \subseteq (S, \rho)$, we say that A is dense in B iff each globe $G_p\; p \in B$, meets A. By Theorem 3, this means that each $p \in B$ is in $\overline{A}$; i.e.,

$$p = \lim_{n\to\infty} x_n \quad \text{for some } \{x_n\} \subseteq A. \tag{3.12.24}$$

Equivalently, $A \subseteq B \subseteq \overline{A}$.[3] .

Summing up, we have the following:

$$A \text{ is open iff } A = A^0. \tag{3.12.25}$$

$$A \text{ is closed iff } A = \overline{A}\text{; equivalently, iff } A \supseteq A'. \tag{3.12.26}$$

$$A \text{ is dense in } B \text{ iff } A \subseteq B \subseteq \overline{A}. \tag{3.12.27}$$

$$A \text{ is perfect iff } A = A'. \tag{3.12.28}$$

3.12.E: Problems on Cluster Points, Closed Sets, and Density

Exercise $3.12.E.1$

Complete the proof of Theorem $1(\text{ ii })$.

Exercise $3.12.E.2$

Prove that $\overline{R}=E^{1}$ and $\overline{R^{n}}=E^{n}(\text { Example (a))}$.

Exercise $3.12.E.3$

Prove Theorem 2 for E^{3}. Prove it for E^{n} $\left(^{*} \text { and } C^{n}\right)$ by induction on n.

Exercise $3.12.E.4$

Verify Note 2.

Exercise $3.12.E.5$

Prove Theorem 3.

Exercise $3.12.E.6$

Prove Corollaries 1 and 2.

Exercise $3.12.E.7$

Prove that $(A \cup B)^{\prime}=A^{\prime} \cup B^{\prime}$.
[Hint: Show by contradiction that $p \notin\left(A^{\prime} \cup B^{\prime}\right)$ excludes $p \in(A \cup B)^{\prime}$. Hence $(A \cup B)^{\prime} \subseteq A^{\prime} \cup B^{\prime}$. Then show that $A^{\prime} \subseteq(A \cup B)^{\prime}, \text { etc. }]$

Exercise $3.12.E.8$

From Problem 7, deduce that $A \cup B$ is closed if A and B are. Then prove Corollary 4. By induction, extend both assertions to any finite number of sets.

Exercise $3.12.E.9$

From Theorem 4, prove that if the sets $A_{i}(i \in I)$ are closed, so is $\bigcap_{i \in I} A_{i}$.

Exercise $3.12.E.10$

Prove Corollary 3 from Theorem 3. Deduce that $\overline{\overline{A}}=\overline{A}$ and prove footnote 3.
[Hint: Consider Figure 7 and Example (1) in §12 when using Theorem 3 (twice).]

Exercise $3.12.E.11$

Prove that $\overline{A}$ is contained in any closed superset of A and is the intersection of all such supersets.
[Hint: Use Corollaries 2 and 3.]

Exercise $3.12.E.12$

(i) Prove that a bounded sequence $\{\overline{x}_m\} \subseteq E^n\ ({}^*C^n)$ converges to $\overline{p}$ iff $\overline{p}$ is its only cluster point.
(ii) Disprove it for
(a) unbounded $\{\overline{x}_m\}$ and
(b) other spaces.
[Hint: For (i), if $\overline{x}_m \rightarrow \overline{p}$ fails, some $G_{\overline{p}}$ leaves out infinitely many $\overline{x}_m$. These $\overline{x}_m$ form a bounded subsequence that, by Theorem 2, clusters at some $\overline{q} \neq \overline{p}$. (Why?) Thus $\overline{q}$ is another cluster point (contradiction!)
For (ii), consider (a) Example (f) in §14 and (b) Problem 10 in §14, with (0,2] as a subspace of E^1.]

Exercise $3.12.E.13$

In each case of Problem 10 in §14, find $\overline{A}$. Is A closed? (Use Theorem 4.)

Exercise $3.12.E.14$

Prove that if $\{b_n\} \subseteq B \subseteq \overline{A}$ in (S, ρ), there is a sequence $\{a_n\} \subseteq A$ such that $\rho(a_n, b_n) \rightarrow 0$. Hence $a_n \rightarrow p$ iff $b_n \rightarrow p$.
[Hint: Choose $a_n \in G_{b_n}(1/n)$.]

Exercise $3.12.E.15$

We have, by definition,

$$p \in A^0 \text{ iff } (\exists \delta > 0) G_p(\delta) \subseteq A; \tag{3.12.E.1}$$

hence

$$p \notin A^0 \text{ iff } (\forall \delta > 0) G_p(\delta) \nsubseteq A, \text{ i.e., } G_p(\delta) - A \neq \emptyset. \tag{3.12.E.2}$$

(See Chapter 1, §§$1-3$.) Find such quantifier formulas for $p \in \overline{A}, p \notin \overline{A}, p \in A'$, and $p \notin A'$.
[Hint: Use Corollary 6 in §14, and Theorem 3 in §16.]

Exercise $3.12.E.16$

Use Problem 15 to prove that
(i) $-(\overline{A}) = (-A)^0$ and
(ii) $-\left(A^0\right) = \overline{-A}$.

Exercise $3.12.E.17$

Show that

$$\overline{A} \cap (\overline{-A}) = \operatorname{bd} A(\text{ boundary of } A); \tag{3.12.E.3}$$

cf. §12, Problem 18. Hence prove again that A is closed iff $A \supseteq$ bd A.
[Hint: Use Theorem 4 and Problem 16 above.]

Exercise $3.12.E.*18$

A set A is said to be nowhere dense in (S, ρ) iff $(\overline{A})^0 = \emptyset$. Show that Cantor's set $P(\text{§}14,\ \text{Problem } 17)$ is nowhere dense.
[Hint: P is closed, so $\overline{P} = P$.]

Exercise $3.12.E.*19$

Give another proof of Theorem 2 for E^1.
[Hint: Let $A \subseteq [a, b]$. Put

$$Q = \{x \in [a, b] | x \text{ exceeds infinitely many points (or terms) of } A\}. \tag{3.12.E.4}$$

Show that Q is bounded and nonempty, so it has a glb, say, $p = \inf A$. Show that A clusters at p.]

Exercise $3.12.E.*20$

For any set $A \subseteq (S, \rho)$ define

$$G_A(\varepsilon) = \bigcup_{x \in A} G_x(\varepsilon). \tag{3.12.E.5}$$

Prove that

$$\overline{A} = \bigcap_{n=1}^{\infty} G_A\left(\frac{1}{n}\right). \tag{3.12.E.6}$$

Exercise $3.12.E.*21$

Prove that

$$\overline{A} = \{x \in S | \rho(x, A) = 0\}; \text{ see } \$13, \text{ Note } 3. \tag{3.12.E.7}$$

Hence deduce that a set A in (S, ρ) is closed iff
$(\forall x \in S) \quad \rho(x, A) = 0 \Longrightarrow x \in A$.

3.13: Cauchy Sequences. Completeness

This page is a draft and is under active development.

A convergent sequence is characterized by the fact that its terms x_m become (and stay) arbitrarily close to its limit, as $m \rightarrow +\infty$. Due to this, however, they also get close to each other; in fact, $\rho\left(x_m, x_n\right)$ can be made arbitrarily small for sufficiently large m and n. It is natural to ask whether the latter property, in turn, implies the existence of a limit. This problem was first studied by Augustin-Louis Cauchy $(1789-1857)$. Thus we shall call sequences Cauchy sequences. More precisely, we formulate the following.

Definition

A sequence $\left\{x_m\right\} \subseteq(S, \rho)$ is called a Cauchy sequence (we briefly say that " $\left\{x_m\right\}$ is Cauchy") iff, given any $\varepsilon>0$ (no matter how small), we have $\rho\left(x_m, x_n\right)<\varepsilon$ for all but finitely many m and n. In symbols,

$$(\forall \varepsilon>0)(\exists k)(\forall m, n>k) \quad \rho\left(x_m, x_n\right)<\varepsilon. \tag{3.13.1}$$

Observe that here we only deal with terms x_m, x_n, not with any other point. The limit (if any) is not involved, and we do not have to know it in advance. We shall now study the relationship between property (1) and convergence.

Theorem 3.13.1

Every convergent sequence $\left\{x_m\right\} \subseteq(S, \rho)$ is Cauchy.

Proof

Let $x_m \rightarrow p$. Then given $\varepsilon>0$, there is a k such that

$$(\forall m>k) \quad \rho\left(x_m, p\right)<\frac{\varepsilon}{2}. \tag{3.13.2}$$

As this holds for any $m>k$, it also holds for any other term x_n with $n>k$.

Thus

$$(\forall m, n>k) \quad \rho\left(x_m, p\right)<\frac{\varepsilon}{2} \text{ and } \rho\left(p, x_n\right)<\frac{\varepsilon}{2}. \tag{3.13.3}$$

Adding and using the triangle inequality, we get

$$\rho\left(x_m, x_n\right) \leq \rho\left(x_m, p\right)+\rho\left(p, x_n\right)<\varepsilon, \tag{3.13.4}$$

and (1) is proved. □

Theorem 3.13.2

Every Cauchy sequence $\left\{x_m\right\} \subseteq(S, \rho)$ is bounded.

Proof

We must show that all x_m are in some globe. First we try an arbitrary radius ε. Then by (1), there is k such that $\rho\left(x_m, x_n\right)<\varepsilon$ for $m, n>k$. Fix some $n>k$. Then

$$(\forall m>k) \rho\left(x_m, x_n\right)<\varepsilon, \text{ i.e., } x_m \in G_{x_n}(\varepsilon). \tag{3.13.5}$$

Thus the globe $G_{x_n}(\varepsilon)$ contains all x_m except possibly the k terms $x_1, \ldots, x_k$. To include them as well, we only have to take a larger radius r, greater than $\rho\left(x_m, x_n\right), m=1, \ldots, k$. Then all x_m are in the enlarged globe $G_{x_n}(r)$. □

Note 1. In E^{1}, under the standard metric, only sequences with finite limits are regarded as convergent. If $x_{n} \rightarrow \pm\infty$, then $\{x_{n}\}$ is not even a Cauchy sequence in $E^{1}(\text{ in view of Theorem } 2)$; but in E^{*}, under a suitable metric (cf. Problem 5 in §11, it is convergent (hence also Cauchy and bounded).

Theorem 3.13.3

If a Cauchy sequence $\{x_{m}\}$ clusters at a point p, then $x_{m} \rightarrow p$.

Proof

We want to show that $x_{m} \rightarrow p$, i.e., that

$$(\forall \varepsilon>0)(\exists k)(\forall m>k) \quad \rho\left(x_{m}, p\right)<\varepsilon. \tag{3.13.6}$$

Thus we fix $\varepsilon>0$ and look for a suitable k. Now as $\{x_{m}\}$ is Cauchy, there is a k such that

$$(\forall m, n>k) \quad \rho\left(x_{m}, x_{n}\right)<\frac{\varepsilon}{2}. \tag{3.13.7}$$

Also, as p is a cluster point, the globe $G_{p}\left(\frac{\varepsilon}{2}\right)$ contains infinitely many x_{n}, so we can fix one with $n>k\,(k \text{ as above})$. Then $\rho\left(x_{n}, p\right)<\frac{\varepsilon}{2}$ and, as noted above, also $\rho\left(x_{m}, x_{n}\right)<\frac{\varepsilon}{2}$ for $m>k$. Hence

$$(\forall m>k) \quad \rho\left(x_{m}, x_{n}\right)+\rho\left(x_{n}, p\right)<\varepsilon, \tag{3.13.8}$$

implying $\rho\left(x_{m}, p\right) \leq \rho\left(x_{m}, x_{n}\right)+\rho\left(x_{n}, p\right)<\varepsilon,$ as required. □

Note 2. It follows that a Cauchy sequence can have at most one cluster point p, for p is also its limit and hence unique; see §14, Corollary 1.

These theorems show that Cauchy sequences behave very much like convergent ones. Indeed, our next theorem (a famous result by Cauchy) shows that, in E^{n} $\left({}^{*} \text{ and } C^{n}\right)$ the two kinds of sequences coincide.

Theorem 3.13.4

(Cauchy's convergence criterion). A sequence $\{\overline{x}_{m}\}$ in E^{n} (*or C^{n}) converges if and only if it is a Cauchy sequence.

Proof

Conversely, let $\{x_{m}\}$ be a Cauchy sequence. Then by Theorem 2, it is bounded. Hence by the Bolzano-Weierstrass theorem (Theorem 2 of §16, it has a cluster point $\overline{p}$. Thus by Theorem 3 above, it converges to $\overline{p}$, and all is proved. □

Unfortunately, this theorem (along with the Bolzano-Weierstrass theorem used in its proof) does not hold in all metric spaces. It even fails in some subspaces of E^{1}. For example, we have

$$x_{m}=\frac{1}{m} \rightarrow 0 \text{ in } E^{1}. \tag{3.13.9}$$

By Theorem 1 , this sequence, being convergent, is also a Cauchy sequence. Moreover, it still preserves (1) even if we remove the point 0 from E^{1} since the distances $\rho\left(x_{m}, x_{n}\right)$ remain the same. However, in the resulting subspace $S=E^{1}-\{0\},$ the sequence no longer converges because its limit (and unique cluster point) 0 has disappeared, leaving a "gap" in its place. Thus we have a Cauchy sequence in S, without a limit or cluster points, so Theorem 4 fails in S (along with the Bolzano-Weierstrass theorem).

Quite similarly, both theorems fail in $(0,1)$ (but not in $[0,1]$) as a subspace of E^{1}. By analogy to incomplete ordered fields, it is natural to say that S is "incomplete" because of the missing cluster point 0, and call a space (or subspace) "complete" if it has no such "gaps," i.e., if Theorem 4 holds in it.

Thus we define as follows.

Definition

A metric space (or subspace) (S, ρ) is said to be complete iff every Cauchy sequence in S converges to some point p in S.

Similarly, a set $A \subseteq (S, \rho)$ is called complete iff each Cauchy sequence $\{x_m\} \subseteq A$ converges to some point p in A, i.e., iff (A, ρ) is complete as a metric subspace of (S, ρ).

In particular, $E^n\left({}^* \text{and}\, C^n\right)$ are complete by Theorem 4. The sets $(0,1)$ and $E^1 - \{0\}$ are incomplete in E^1, but $[0,1]$ is complete. Indeed, we have the following theorem.

Theorem 3.13.5

(i) Every closed set in a complete space is complete itself.

(ii) Every complete set $A \subseteq (S, \rho)$ is necessarily closed.

Proof

(i) Let A be a closed set in a complete space (S, ρ). We have to show that Theorem 4 holds in $A(\text{ as it does in } S)$. Thus we fix any Cauchy sequence $\{x_m\} \subseteq A$ and prove that it converges to some p in A.

Now, since S is complete, the Cauchy sequence $\{x_m\}$ has a limit p in S. As A is closed, however, that limit must be in A by Theorem 4 in §16. Thus (i) is proved.

(ii) Now let A be complete in a metric space (S, ρ). To prove that A is closed, we again use Theorem 4 of §16. Thus we fix any convergent sequence $\{x_m\} \subseteq A, x_m \rightarrow p \in S,$ and show that p must be $in A$.

Now, since $\{x_m\}$ converges in S, it is a Cauchy sequence, in S as well as in A. Thus by the assumed completeness of A, it has a limit q in A. Then, however, the uniqueness of $\lim_{x_m} x_m(\text{ in } S)$ implies that $p = q \in A,$ so that p is in A, indeed. $\square$

3.13.E: Problems on Cauchy Sequences

Exercise $3.13.E.1$

Without using Theorem 4, prove that if $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences in E^1(or C), so also are

$$\text{(i) } \{x_n+y_n\} \quad \text{ and } \quad \text{(ii) } \{x_n y_n\}. \tag{3.13.E.1}$$

Exercise $3.13.E.2$

Prove that if $\{x_m\}$ and $\{y_m\}$ are Cauchy sequences in (S, ρ), then the sequence of distances

$$\rho\left(x_m, y_m\right), \quad m=1,2, \ldots, \tag{3.13.E.2}$$

converges in E^1.
[Hint: Show that this sequence is Cauchy in E^1; then use Theorem 4.]

Exercise $3.13.E.3$

Prove that a sequence $\{x_m\}$ is Cauchy in (S, ρ) iff

$$(\forall \varepsilon>0)(\exists k)(\forall m>k) \quad \rho\left(x_m, x_k\right)<\varepsilon. \tag{3.13.E.3}$$

Exercise $3.13.E.4$

Two sequences $\{x_m\}$ and $\{y_m\}$ are called concurrent iff

$$\rho\left(x_m, y_m\right) \rightarrow 0. \tag{3.13.E.4}$$

Notation: $\{x_m\} \approx \{y_m\}$. Prove the following.
(i) If one of them is Cauchy or convergent, so is the other, and
$\lim x_m=\lim y_m$ (if it exists).
(ii) If any two sequences converge to the same limit, they are concurrent.

Exercise $3.13.E.5$

Show that if $\{x_m\}$ and $\{y_m\}$ are Cauchy sequences in (S, ρ), then

$$\lim_{m \rightarrow \infty} \rho\left(x_m, y_m\right) \tag{3.13.E.5}$$

does not change if $\{x_m\}$ or $\{y_m\}$ is replaced by a concurrent sequence (see Problems 4 and 2).
Call

$$\lim_{m \rightarrow \infty} \rho\left(x_m, y_m\right) \tag{3.13.E.6}$$

the "distance"

$$\rho\left(\{x_m\},\{y_m\}\right) \tag{3.13.E.7}$$

between $\{x_m\}$ and $\{y_m\}$. Prove that such "distances" satisfy all metric axioms, except that $\rho\left(\{x_m\},\{y_m\}\right)$ may be 0 even for different sequences. (When?)
Also, show that if

$$(\forall m) \quad x_m = a \text{ and } y_m = b(\text{ constant }), \tag{3.13.E.8}$$

then $\rho(\{x_m\}, \{y_m\}) = \rho(a, b)$.

Exercise $3.13.E.5'$

Continuing Problems 4 and 5, show that the concurrence relation $(\approx)$ is reflexive, symmetric, and transitive (Chapter $1, \S\S 4-7$), i.e., an equivalence relation. That is, given $\{x_m\}, \{y_m\}$ in S, prove that
(a) $\{x_m\} \approx \{x_m\}$ (reflexivity);
(b) if $\{x_m\} \approx \{y_m\}$ then $\{y_m\} \approx \{x_m\}$ (symmetry);
(c) if $\{x_m\} \approx \{y_m\}$ and $\{y_m\} \approx \{z_m\}$, then $\{x_m\} \approx \{z_m\}$ (transitivity).

Exercise $3.13.E. *5''$

From Problem 4 deduce that the set of all sequences in (S, ρ) splits into disjoint equivalence classes (as defined in Chapter $1, \S\S 4-7$) under the relation of concurrence $(\approx)$. Show that all sequences of one and the same class either converge to the same limit or have no limit at all, and either none of them is Cauchy or all are Cauchy.

Exercise $3.13.E.6$

Give examples of incomplete metric spaces possessing complete subspaces.

Exercise $3.13.E.7$

Prove that if a sequence $\{x_m\} \subseteq (S, \rho)$ is Cauchy then it has a subsequence $\{x_{m_k}\}$ such that

$$(\forall k) \quad \rho(x_{m_k}, x_{m_{k+1}}) < 2^{-k}. \tag{3.13.E.9}$$

Exercise $3.13.E.8$

Show that every discrete space (S, ρ) is complete.

Exercise $3.13.E. *9$

Let C be the set of all Cauchy sequences in (S, ρ); we denote them by capitals, e.g., $X = \{x_m\}$. Let

$$X^* = \{Y \in C | Y \approx X\} \tag{3.13.E.10}$$

denote the equivalence class of X under concurrence, $\approx$ (see Problems 2, $5'$, and $5''$ "). We define

$$\sigma(X^*, Y^*) = \rho(\{x_m\}, \{y_m\}) = \lim_{m \to \infty} \rho(x_m, y_m). \tag{3.13.E.11}$$

By Problem 5, this is unambiguous, for $\rho(\{x_m\}, \{y_m\})$ does not depend on the particular choice of $\{x_m\} \in X^*$ and $\{y_m\} \in Y^*$; and $\lim \rho(x_m, y_m)$ exists by Problem 2.
Show that σ is a metric for the set of all equivalence classes X^* $(X \in C)$; call this set C^*.

Exercise $3.13.E. *10$

Continuing Problem 9, let x^* denote the equivalence class of the sequence with all terms equal to x ; let C' be the set of all such "constant" equivalence classes (it is a subset of C^*).
Show that C' is dense in (C^*, σ), i.e., $\overline{C'} = C^*$ under the metric σ. (See §16, Definition 2.)

[Hint: Fix any "point" $X^{*} \in C^{*}$ and any globe $G\left(X^{*} ; \varepsilon\right)$ about X^{*} in $\left(C^{*}, \sigma\right)$. We must show that it contains some $x^{*} \in C^{\prime}$. By definition, X^{*} is the equivalence class of some Cauchy sequence $X=\left\{x_{m}\right\}$ in (S, ρ), so

$$(\exists k)(\forall m, n>k) \quad \rho\left(x_{m}, x_{n}\right)<\frac{\varepsilon}{2}. \tag{3.13.E.12}$$

Fix some $x=x_{n}(n>k)$ and consider the equivalence class x^{*} of the sequence $\{x, x, \ldots, x, \ldots\}$; thus, $x^{*} \in C^{\prime}$, and

$$\sigma\left(X^{*}, x^{*}\right)=\lim_{m \rightarrow \infty} \rho\left(x_{m}, x\right) \leq \frac{\varepsilon}{2}. \quad \text{(Why?)} \tag{3.13.E.13}$$

Thus $x^{*} \in G\left(X^{*}, \varepsilon\right)$, as required.]

Exercise $3.13.E.*11$

Two metric spaces (S, ρ) and (T, σ) are said to be $isometric$ iff there is a map $f: S \longleftrightarrow_{onto} T$ such that

$$(\forall x, y \in S) \quad \rho(x, y)=\sigma(f(x), f(y)). \tag{3.13.E.14}$$

Show that the spies (S, ρ) and $\left(C^{\prime}, \sigma\right)$ of Problem 10 are $isometric$. Note that it is customary not to distinguish between two isometric spaces, treating each of them as just an "isometric copy" of the other. Indeed, distances in each of them are alike.
[Hint: Define \(f(x) = x*.]

Exercise $3.13.E.*12$

Continuing Problems 9 to 11, show that the space $\left(C^{*}, \sigma\right)$ is complete. Thus prove that for every metric space (S, ρ), there is a complete metric space $\left(C^{*}, \sigma\right)$ containing an isometric copy $C^{\prime}$ of S, with $C^{\prime}$ dense in C^{*}. C^{*} is called a completion of (S, ρ).
[Hint: Take a Cauchy sequence $\left\{X_{m}^{*}\right\}$ in $\left(C^{*}, \sigma\right)$. By Problem 10, each globe $G\left(X_{m}^{*} ; \frac{1}{m}\right)$ contains some $x_{m}^{*} \in C^{\prime}$, where x_{m}^{*} is the equivalence class of

$$\left\{x_{m}, x_{m}, \ldots, x_{m}, \ldots\right\} \tag{3.13.E.15}$$

and $\sigma\left(X_{m}^{*}, x_{m}^{*}\right)<\frac{1}{m} \rightarrow 0$. Thus by Problem 4, $\left\{x_{m}^{*}\right\}$ is Cauchy in $\left(C^{*}, \sigma\right)$, as is $\left\{X_{m}^{*}\right\}$. Deduce that $X=\left\{x_{m}\right\} \in C$, and $X^{*}=\lim_{m \rightarrow \infty} X_{m}^{*}$ in $\left(C^{*}, \sigma\right)$.]

CHAPTER OVERVIEW

4: Function Limits and Continuity

4.1: Basic Definitions

We shall now consider functions whose domains and ranges are sets in some fixed (but otherwise arbitrary) metric spaces (S,ρ) and (T,ρ'), respectively. We write

$$f : A \rightarrow (T,\rho') \tag{4.1.1}$$

for a function f with $D_f = A \subseteq (S,\rho)$ and $D'_f \subseteq (T,\rho')$. S is called the domain space, and T the range space, of f.

I. Given such a function, we often have to investigate its "local behavior" near some point $p \in S$. In particular, if $p \in A = D_f($ so that $f(p)$ is defined$)$ we may ask: Is it possible to make the function values $f(x)$ as near as we like ("$\varepsilon-$ near") to $f(p)$ by keeping x sufficiently close $\left(\text{"close"}\right)$ to p, i.e., inside some sufficiently small globe $G_p(\delta)$? If this is the case, we say that f is continuous at p. More precisely, we formulate the following definition.

Definition

A function $f : A \rightarrow (T,\rho')$, with $A \subseteq (S,\rho)$, is said to be continuous at p iff $p \in A$ and, moreover, for each $\varepsilon > 0$ (no matter how small) there is $\delta > 0$ such that $\rho'(f(x), f(p)) < \varepsilon$ for all $x \in A \cap G_p(\delta)$. In symbols,

$$(\forall \varepsilon > 0)(\exists \delta > 0)\,(\forall x \in A \cap G_p(\delta)) \begin{cases} \rho'(f(x), f(p)) < \varepsilon, \text{ or} \\ f(x) \in G_{f(p)}(\varepsilon) \end{cases} \tag{4.1.2}$$

If (1) fails, we say that f is discontinuous at p and call p a discontinuity point of f. This is also the case if $p \notin A$ (since $f(p)$ is not defined).

If (1) holds for each p in a set $B \subseteq A$, we say that f is continuous on B. If this is the case for $B = A$, we simply say that f is continuous.

Sometimes we prefer to keep x near p but different from p. We then replace $G_p(\delta)$ in (1) by the set $G_p(\delta) - \{p\}$, i.e., the globe without its center, denoted $G_{\neg p}(\delta)$ and called the deleted δ-globe about p. This is even necessary if $p \notin D_f$. Replacing $f(p)$ in (1) by some $q \in T$, we then are led to the following definition.

Definition

Given $f : A \rightarrow (T,\rho'), A \subseteq (S,\rho), p \in S,$ and $q \in T$, we say that $f(x)$ tends to q as x tends to $p(f(x) \rightarrow q \text{ as } x \rightarrow p)$ iff for each $\varepsilon > 0$ there is $\delta > 0$ such that $\rho'(f(x), q) < \varepsilon$ for all $x \in A \cap G_{\neg p}(\delta)$. In symbols,

$$(\forall \varepsilon > 0)(\exists \delta > 0)\,(\forall x \in A \cap G_{\neg p}(\delta)) \quad \begin{cases} \rho'(f(x), q) < \varepsilon, \text{ i.e.} \\ f(x) \in G_q(\varepsilon) \end{cases} \tag{4.1.3}$$

This means that $f(x)$ is ε-close to q when x is δ-close to p and $x \neq p$.

If (2) holds for some q, we call q a limit of f at p. There may be no such q. We then say that f has no limit at p, or that this limit does not exist. If there is only one such $q($ for a given $p)$, we write $q = \lim_{x \rightarrow p} f(x)$.

Note 1. Formula (2) holds "vacuously" (see Chapter 1,8 §§1-3, end remark) if $A \cap G_{\neg p}(\delta) = \emptyset$ for some $\delta > 0$. Then any $q \in T$ is a limit at p, so a limit exists but is not unique. (We discard the case where T is a singleton.)

Note 2. However, uniqueness is ensured if $A \cap G_{\neg p}(\delta) \neq \emptyset$ for all $\delta > 0$, as we prove below.

Observe that by Corollary 6 of Chapter 3, §14, the set A clusters at p iff

$$(\forall \delta > 0) \quad A \cap G_{\neg p}(\delta) \neq \emptyset. \quad (\text{ Explain! }) \tag{4.1.4}$$

Thus we have the following corollary.

corollary 4.1.1

If A clusters at p in (S,ρ), then a function $f: A \rightarrow (T,p')$ can have at most one limit at p; i.e.

$$\lim_{x\rightarrow p} f(x) \text{ is unique (if it exists).} \tag{4.1.5}$$

In particular, this holds if $A \supseteq (a,b) \subset E^1 (a<b)$ and $p \in [a,b]$.

Proof

Suppose f has two limits, q and r, at p. By the Hausdorff property,

$$G_q(\varepsilon) \cap G_r(\varepsilon) = \emptyset \quad \text{for some } \varepsilon > 0. \tag{4.1.6}$$

Also, by (2), there are $\delta', \delta'' > 0$ such that

$$\begin{array}{ll}(\forall x \in A \cap G_{\neg p}(\delta')) & f(x) \in G_q(\varepsilon) \text{ and} \\ (\forall x \in A \cap G_{\neg p}(\delta'')) & f(x) \in G_r(\varepsilon)\end{array} \tag{4.1.7}$$

Let $\delta = \min(\delta', \delta'')$. Then for $x \in A \cap G_{\neg p}(\delta), f(x)$ is in both $G_q(\varepsilon)$ and $G_r(\varepsilon)$, and such an x exists since $A \cap G_{\neg p}(\delta) \neq \emptyset$ by assumption.

But this is impossible since $G_q(\varepsilon) \cap G_r(\varepsilon) = \emptyset$ (a contradiction!). $\square$

For intervals, see Chapter 3, §14, Example (h).

corollary 4.1.2

f is continuous at p $(p \in D_f)$ iff $f(x) \rightarrow f(p)$ as $x \rightarrow p$.

Proof

The straightforward proof from definitions is left to the reader.

Note 3. In formula (2), we excluded the case $x=p$ by assuming that $x \in A \cap G_{\neg p}(\delta)$. This makes the behavior of f at p itself irrelevant. Thus for the existence of a limit q at p, it does not matter whether $p \in D_f$ or whether $f(p)=q$. But both conditions are required for continuity at p (see Corollary 2 and Definition 1).

Note 4. Observe that if (1) or (2) holds for some δ, it certainly holds for any $\delta' \leq \delta$. Thus we may always choose δ as small as we like. Moreover, as x is limited to $G_p(\delta)$, we may disregard, or change at will, the function values $f(x)$ for $x \notin G_p(\delta)$ ("local character of the limit notion").

II. Limits in E*. If S or T is E^* $\left(\text{or } E^1\right)$, we may let $x \rightarrow \pm\infty$ or $f(x) \rightarrow \pm\infty$. For a precise definition, we rewrite (2) in terms of $globes$ G_p and G_q :

$$(\forall G_q)(\exists G_p)(\forall x \in A \cap G_{\neg p}) \quad f(x) \in G_q. \tag{4.1.8}$$

This makes sense also if $p=\pm\infty$ or $q=\pm\infty$. We only have to use our conventions as to $G_{\pm\infty}$, or the metric ρ' for E^*, as explained in Chapter 3, §11.

For example, consider

$$''f(x) \rightarrow q \text{ as } x \rightarrow +\infty'' \left(A \subseteq S=E^*, p=+\infty, q \in (T,\rho')\right). \tag{4.1.9}$$

Here G_p has the form $(a,+\infty], a \in E^1$, and $G_{\neg p}=(a,+\infty)$, while $G_q=G_q(\varepsilon)$, as usual. Noting that $x \in G_{\neg p}$ means $x>a$ $\left(x \in E^1\right)$, we can rewrite $(2')$ as

$$(\forall \varepsilon>0)\left(\exists a \in E^1\right)(\forall x \in A | x>a) \quad f(x) \in G_q(\varepsilon), \text{ or } \rho'(f(x),q)<\varepsilon. \tag{4.1.10}$$

This means that $f(x)$ becomes arbitrarily close to q for large $x(x>a)$.

Next consider "$f(x) \rightarrow +\infty$ as $x \rightarrow -\infty$" Here $G_{\neg p}=(-\infty, a)$ and $G_q=(b,+\infty]$. Thus formula $(2')$ yields (with $S=T=E^*$, and x varying over E^1)

$$\left(\forall b \in E^1\right)\left(\exists a \in E^1\right)(\forall x \in A \mid x<a) \quad f(x)>b; \tag{4.1.11}$$

similarly in other cases, which we leave to the reader.

Note 5. In (3), we may take $A=N$ (the naturals). Then $f: N \rightarrow (T, \rho')$ is a sequence in T. Writing m for x, set $u_m=f(m)$ and $a=k \in N$ to obtain

$$(\forall \varepsilon>0)(\exists k)(\forall m>k) \quad u_m \in G_q(\varepsilon); \text{ i.e., } \rho'\left(u_m, q\right)<\varepsilon. \tag{4.1.12}$$

This coincides with our definition of the limit q of a sequence $\{u_m\}$ (see Chapter 3, §14). Thus limits of sequences are a special case of function limits. Theorems on sequences can be obtained from those on functions $f: A \rightarrow (T, \rho')$ by simply taking $A=N$ and $S=E^*$ as above.

Note 6. Formulas (3) and (4) make sense sense also if $S=E^1$ (respectively, $S=T=E^1$) since they do not involve any mention of $\pm\infty$. We shall use such formulas also for functions $f: A \rightarrow T$, with $A \subseteq S \subseteq E^1$ or $T \subseteq E^1$, as the case may be.

III. Relative Limits and Continuity. Sometimes the desired result (1) or (2) does not hold in full, but only with A replaced by a smaller set $B \subseteq A$. Thus we may have

$$(\forall \varepsilon>0)(\exists \delta>0)\left(\forall x \in B \cap G_{\neg p}(\delta)\right) \quad f(x) \in G_q(\varepsilon). \tag{4.1.13}$$

In this case, we call q a relative limit of f at p over B and write

$$\text{" } f(x) \rightarrow q \text{ as } x \rightarrow p \text{ over } B \text{"} \tag{4.1.14}$$

or

$$\lim_{x \rightarrow p, x \in B} f(x)=q \quad (\text{ if } q \text{ is unique }); \tag{4.1.15}$$

B is called the path over which x tends to p. If, in addition, $p \in D_f$ and $q=f(p)$, we say that f is relatively continuous at p over B; then (1) holds with A replaced by B. Again, if this holds for every $p \in B$, we say that f is relatively continuous on B. Clearly, if $B=A=D_f$, this yields ordinary (nonrelative) limits and continuity. Thus relative limits and continuity are more general.

Note that for limits over a path B, x is chosen from B or $B-\{p\}$ only. Thus the behavior of f outside B becomes irrelevant, and so we may arbitrarily redefine f on $-B$. For example, if $p \notin B$ but $\lim_{x \rightarrow p, x \in B} f(x)=q$ exists, we may define $f(p)=q$, thus making f relatively continuous at $p($ over $B)$. We also may replace (S, ρ) by $(B, \rho)($ if $p \in B)$, or restrict f to B, i.e., replace f by the function $g: B \rightarrow (T, \rho')$ defined by $g(x)=f(x)$ for $x \in B$ (briefly, $g=f$ on B).

A particularly important case is

$$A \subseteq S \subseteq E^*, \text{ e.g., } S=E^1. \tag{4.1.16}$$

Then inequalities are defined in S, so we may take

$$B=\{x \in A \mid x<p\} \text{ (points in } A, \text{ preceding } p). \tag{4.1.17}$$

Then, writing G_q for $G_q(\varepsilon)$ and $a=p-\delta$, we obtain from formula (2)

$$\left(\forall G_q\right)(\exists a<p)(\forall x \in A \mid a<x<p) \quad f(x) \in G_q. \tag{4.1.18}$$

If (5) holds, we call q a left limit of f at p and write

$$\text{" } f(x) \rightarrow q \text{ as } x \rightarrow p^{-} \text{"} \quad \left(\text{" } x \text{ tends to } p \text{ from the left'}\right). \tag{4.1.19}$$

If, in addition, $q=f(p)$, we say that f is left continuous at p. Similarly, taking

$$B=\{x \in A \mid x>p\}, \tag{4.1.20}$$

we obtain right limits and continuity. We write

$$f(x) \rightarrow q \text{ as } x \rightarrow p^{+} \tag{4.1.21}$$

iff q is a right limit of f at p, i.e., if (5) holds with all inequalities reversed.

If the set B in question clusters at p, the relative limit (if any) is unique. We then denote the left and right limit, respectively, by $f\left(p^{-}\right)$ and $f\left(p^{+}\right)$, and we write

$$\lim_{x \rightarrow p^{-}} f(x)=f\left(p^{-}\right) \text{ and } \lim_{x \rightarrow p^{+}} f(x)=f\left(p^{+}\right). \tag{4.1.22}$$

corollary 4.1.3

With the previous notation, if $f(x) \rightarrow q$ as $x \rightarrow p$ over a path B, and also over D, then $f(x) \rightarrow q$ as $x \rightarrow p$ over $B \cup D$.

Hence if $D_f \subseteq E^*$ and $p \in E^*$, we have

$$q=\lim_{x \rightarrow p} f(x) \text{ iff } q=f\left(p^{-}\right)=f\left(p^{+}\right). \quad (\text{ Exercise! }) \tag{4.1.23}$$

We now illustrate our definitions by a diagram in E^2 representing a function $f: E^1 \rightarrow E^1$ by its graph, i.e., points (x, y) such that $y=f(x)$.

Here

$$G_q(\varepsilon)=(q-\varepsilon, q+\varepsilon) \tag{4.1.24}$$

is an interval on the y -axis. The dotted lines show how to construct an interval

$$(p-\delta, p+\delta)=G_p \tag{4.1.25}$$

on the x -axis, satisfying formula (1) in Figure 13, formulas (5) and (6) in Figure 14, or formula (2) in Figure 15. The point Q in each diagram belongs to the graph; i.e., $Q=(p, f(p))$. In Figure 13, f is continuous at $p(\text{ and also at } p_1)$. However, it is only left-continuous at p in Figure 14, and it is discontinuous at p in Figure 15, though $f\left(p^{-}\right)$ and $f\left(p^{+}\right)$ exist. (Why?)

Example 4.1.1

(a) Let $f: A \rightarrow T$ be constant on $B \subseteq A$; i.e.

$$f(x)=q \text{ for a fixed } q \in T \text{ and all } x \in B. \tag{4.1.26}$$

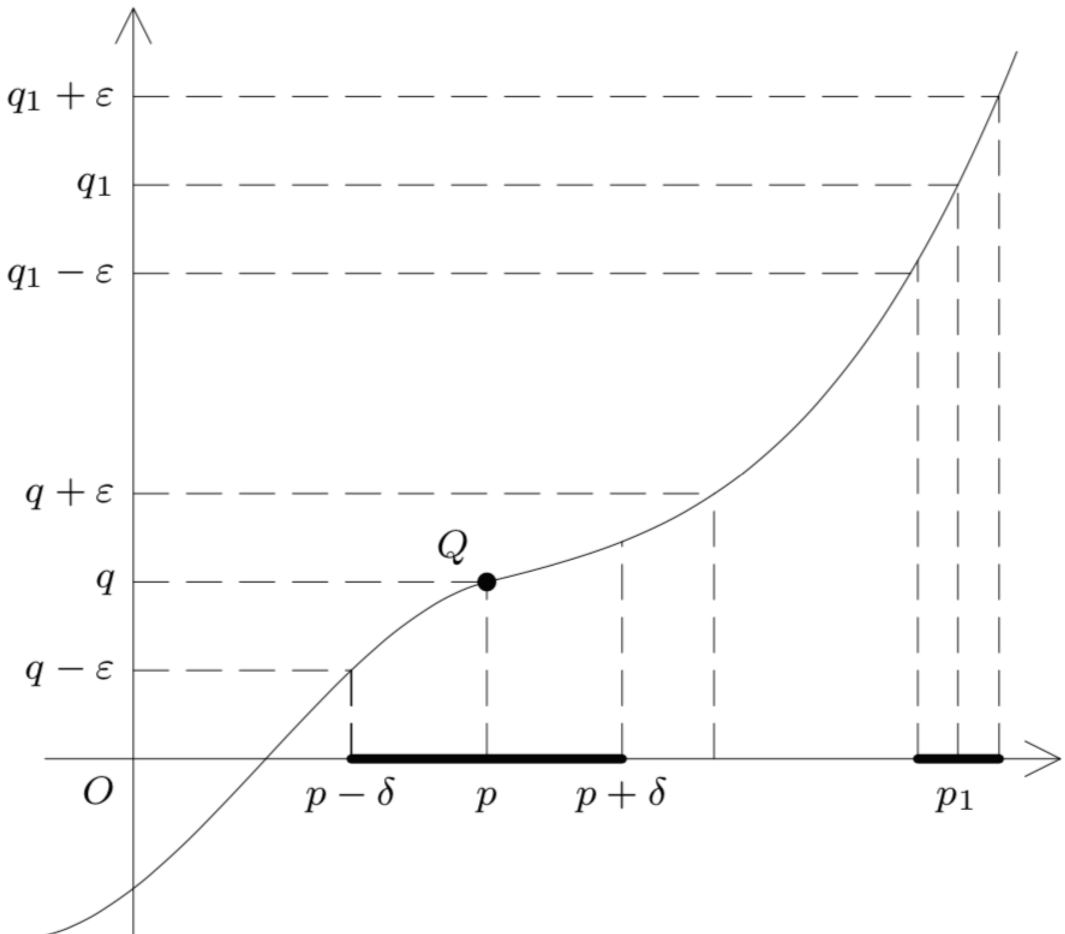
$q_1+\varepsilon$
q_1
$q_1-\varepsilon$
$q+\varepsilon$
Q
q
$q-\varepsilon$
O
$p-\delta$
p
$p+\delta$
p_1

FIGURE 13

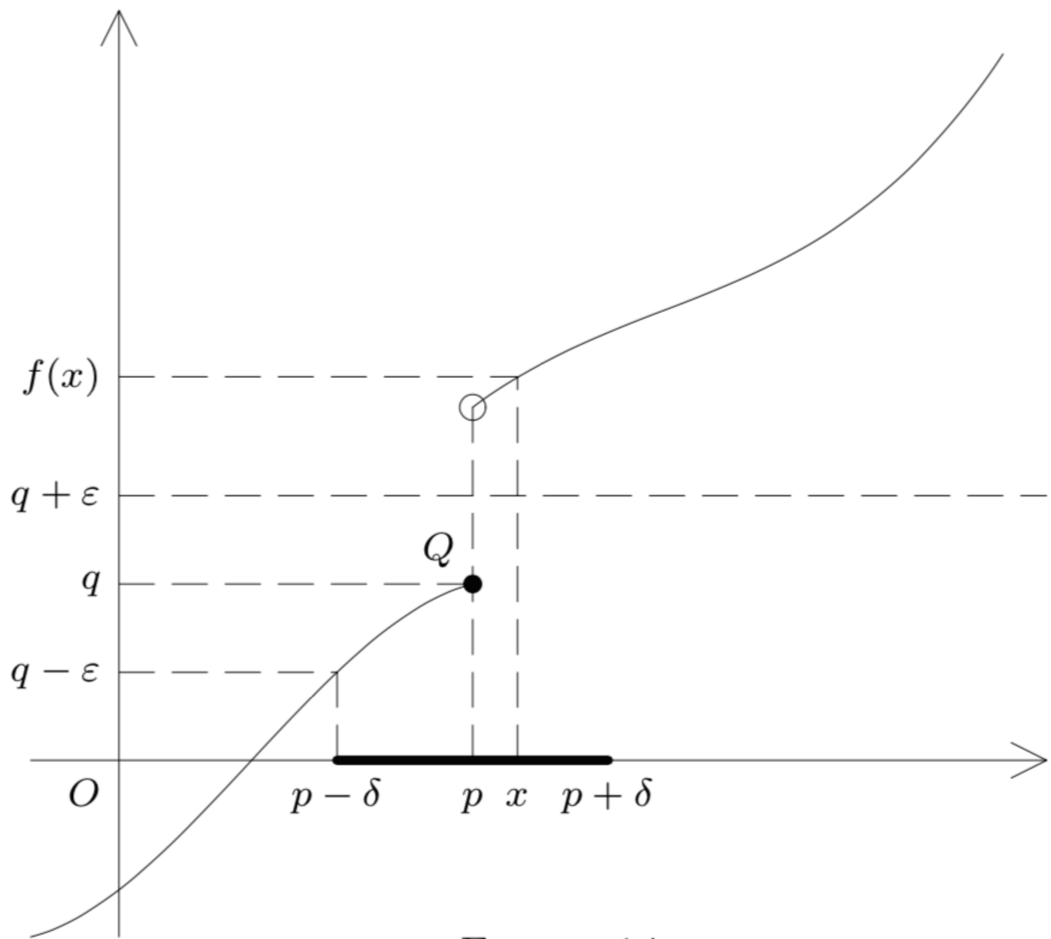

FIGURE 14

Then f is relatively continuous on B, and $f(x) \rightarrow q$ as $x \rightarrow p$ over B, at each p. (Given $\varepsilon>0,$ take an arbitrary $\delta>0$. Then

$$\left(\forall x \in B \cap G_{\neg p}(\delta)\right) \quad f(x)=q \in G_{q}(\varepsilon), \tag{4.1.27}$$

as required; similarly for continuity.)

(b) Let f be the i dentity map on $A \subset(S, \rho)$; i.e.,

$$(\forall x \in A) \quad f(x)=x. \tag{4.1.28}$$

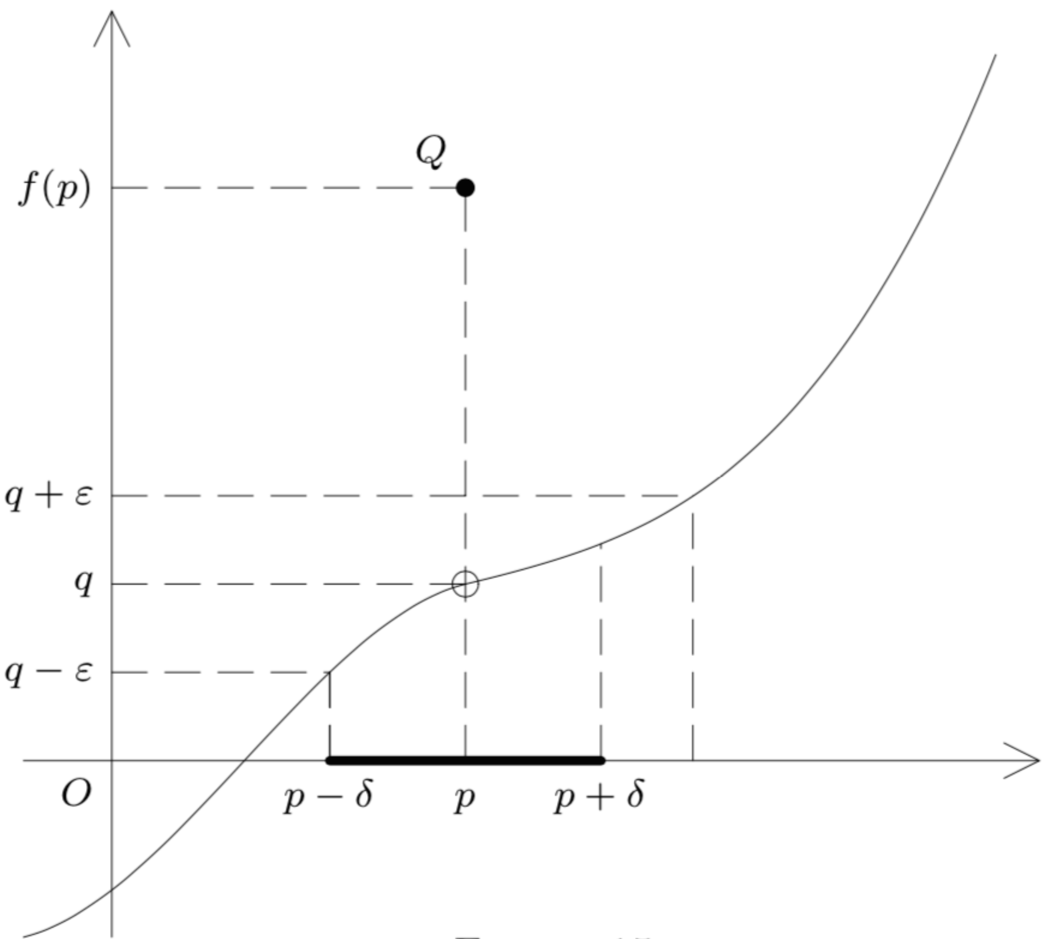

FIGURE 15

Then, given $\varepsilon>0,$ take $\delta=\varepsilon$ to obtain, for $p\in A$,

$$(\forall x\in A\cap G_{p}(\delta))\quad \rho(f(x),f(p))=\rho(x,p)<\delta=\varepsilon. \tag{4.1.29}$$

Thus by (1), f is continuous at any $p\in A$, hence on A.

(c) Define $f:E^{1}\rightarrow E^{1}$ by

$$f(x)=1 \text{ if } x \text{ is rational, and } f(x)=0 \text{ otherwise.} \tag{4.1.30}$$

(This is the Dirichlet function, so named after Johann Peter Gustav Lejeune Dirichlet.)

No matter how small δ is, the globe

$$G_{p}(\delta)=(p-\delta,p+\delta) \tag{4.1.31}$$

(even the deleted globe) contains both rationals and irrationals. Thus as x varies over $G_{\neg p}(\delta)$, $f(x)$ takes on both values, 0 and 1, many times and so gets out of any $G_{q}(\varepsilon)$, with $q\in E^{1},\varepsilon<\frac{1}{2}$.

Hence for any $q,p\in E^{1}$, formula (2) fails if we take $\varepsilon=\frac{1}{4}$, say. Thus f has no limit at any $p\in E^{1}$ and hence is discontinuous everywhere! However, f is relatively continuous on the set R of all rationals by Example (a).

(d) Define $f:E^{1}\rightarrow E^{1}$ by

$$f(x)=[x](= \text{ the integral part of } x; \text{ see Chapter } 2, \S 10). \tag{4.1.32}$$

Thus $f(x)=0$ for $x \in[0,1), f(x)=1$ for $x \in[1,2)$, etc. Then f is discontinuous at p if p is an integer (why?) but continuous at any other p (restrict f to a small $G_{p}(\delta)$ so as to make it constant)

However, left and right limits exist at each $p \in E^{1}$, even if $p=n($ an integer $)$. In fact,

$$f(x)=n, x \in(n, n+1) \tag{4.1.33}$$

and

$$f(x)=n-1, x \in(n-1, n), \tag{4.1.34}$$

hence $f\left(n^{+}\right)=n$ and $f\left(n^{-}\right)=n-1 ; f$ is right continuous on E^{1}. See Figure 16.

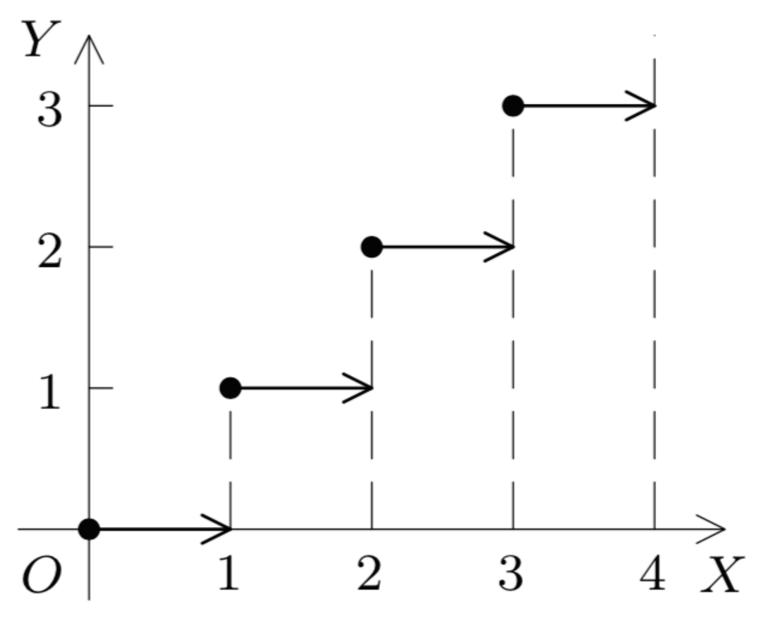

FIGURE 16

(e) Define $f: E^{1} \rightarrow E^{1}$ by

$$f(x)=\frac{x}{|x|} \text { if } x \neq 0, \text { and } f(0)=0. \tag{4.1.35}$$

(This is the so-called signum function, often denoted by sgn.)

Then (Figure 17)

$$f(x)=-1 \text { if } x<0 \tag{4.1.36}$$

and

$$f(x)=1 \text { if } x>0. \tag{4.1.37}$$

Thus, as in (d), we infer that f is discontinuous at 0, but continuous at each $p \neq 0$. Also, $f\left(0^{+}\right)=1$ and $f\left(0^{-}\right)=-1$. Redefining $f(0)=1$ or $f(0)=-1$, we can make f right (respectively, left) continuous at 0, but not both.

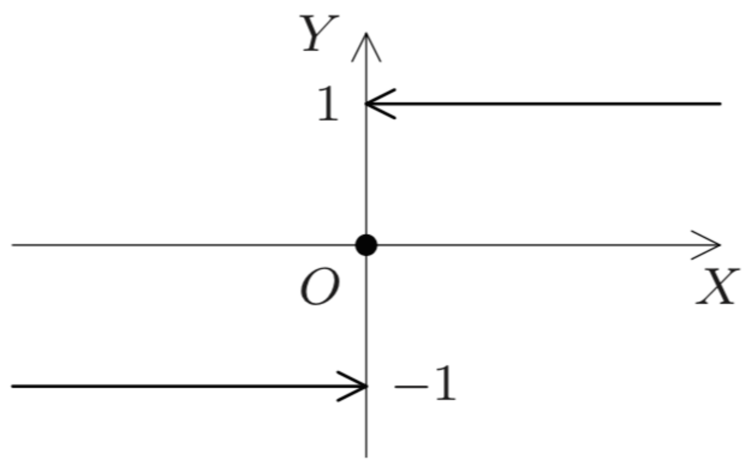

FIGURE 17

(f) Define $f: E^1 \rightarrow E^1$ by (see Figure 18)

$$f(x)=\sin \frac{1}{x} \text { if } x \neq 0, \text { and } f(0)=0. \tag{4.1.38}$$

Any globe $G_0(\delta)$ about 0 contains points at which $f(x)=1$, as well as those at which $f(x)=-1$ or $f(x)=0$ (take $x=2/(n\pi)$ for large integers n); in fact, the graph "oscillates" infinitely many times between -1 and 1. Thus by the same argument as in (c), f has no limit at 0 (not even a left or right limit) and hence is discontinuous at 0. No attempt at redefining f at 0 can restore even left or right continuity, let alone ordinary continuity, at 0.

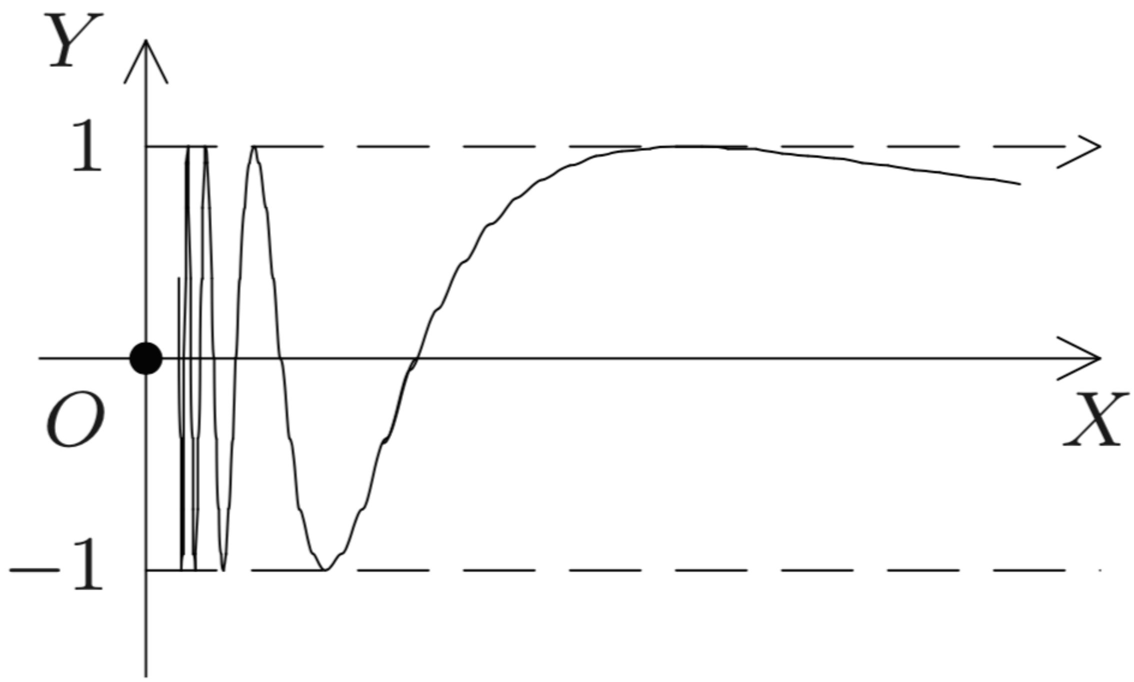

FIGURE 18

(g) Define $f: E^2 \rightarrow E^1$ by

$$f(\overline{0})=0 \text{ and } f(\overline{x})=\frac{x_1 x_2}{x_1^2+x_2^2} \text{ if } \overline{x}=(x_1, x_2) \neq \overline{0}. \tag{4.1.39}$$

Let B be any line in E^2 through $\overline{0}$, given parametrically by

$$\overline{x}=t\vec{u}, \quad t \in E^1, \vec{u} \text{ fixed (see Chapter 3, §§4-6)}, \tag{4.1.40}$$

so $x_1=tu_1$ and $x_2=tu_2$. As is easily seen, for $\overline{x} \in B$, $f(\overline{x})=f(\overline{u})$ (constant) if $\overline{x} \neq \overline{0}$. Hence

$$\left(\forall \overline{x} \in B \cap G_{\neg \overline{0}}(\delta)\right) \quad f(\overline{x})=f(\overline{u}), \tag{4.1.41}$$

i.e., $\rho(f(\overline{x}), f(\overline{u}))=0<\varepsilon,$ for any $\varepsilon>0$ and any deleted globe about $\overline{0}$.

By $\left(2^{\prime}\right)$, then, $f(\overline{x}) \rightarrow f(\overline{u})$ as $\overline{x} \rightarrow \overline{0}$ over the path B. Thus f has a relative limit $f(\overline{u})$ at $\overline{0}$, over any line $\overline{x}=t\overline{u},$ but this limit is different for various choices of $\overline{u}$, i.e., for different lines through $\overline{0}$. No ordinary limit at $\overline{0}$ exists (why?); f is not even relatively continuous at $\overline{0}$ over the line $\overline{x}=t\vec{u}$ unless $f(\overline{u})=0$ (which is the case only if the line is one of the coordinate axes (why?)).

4.1.E: Problems on Limits and Continuity

Exercise 4.1.E.1

Prove Corollary 2. Why can one interchange $G_p(\delta)$ and $G_{\neg p}(\delta)$ here?

Exercise 4.1.E.2

Prove Corollary 3. By induction, extend its first clause to unions of n paths. Disprove it for infinite unions of paths (see Problem 9 in §3).

Exercise 4.1.E.2′

Prove that a function $f : E^1 \rightarrow (T, \rho')$ is continuous at p iff

$$f(p) = f\left(p^{-}\right) = f\left(p^{+}\right). \tag{4.1.E.1}$$

Exercise 4.1.E.3

Show that relative limits and continuity at p (over B) are equivalent to the ordinary ones if B is a neighborhood of p (Chapter 3, §12); for example, if it is some G_p.

Exercise 4.1.E.4

Discuss Figures $13-15$ in detail, comparing $f(p), f\left(p^{-}\right)$, and $f\left(p^{+}\right)$; see Problem $2'$.
Observe that in Figure 13, different values of δ result at p and p_1 for the same ε. Thus δ depends on both ε and the choice of p.

Exercise 4.1.E.5

Complete the missing details in Examples $(\mathrm{d})-(\mathrm{g})$. In(d), redefine $f(x)$ to be the least integer $\geq x$. Show that f is then left-continuous on E^1.

Exercise 4.1.E.6

Give explicit definitions (such as (3)) for

$$\begin{array}{ll} \text{(a) } \lim_{x \rightarrow +\infty} f(x) = -\infty; & \text{(b) } \lim_{x \rightarrow -\infty} f(x) = q; \\ \text{(c) } \lim_{x \rightarrow p} f(x) = +\infty; & \text{(d) } \lim_{x \rightarrow p} f(x) = -\infty; \\ \text{(e) } \lim_{x \rightarrow p^{-}} f(x) = +\infty; & \text{(f) } \lim_{x \rightarrow p^{+}} f(x) = -\infty. \end{array} \tag{4.1.E.2}$$

In each case, draw a diagram (such as Figures $13-15$) and determine whether the domain and range of f must both be in E^*.

Exercise 4.1.E.7

Define $f : E^1 \rightarrow E^1$ by

$$f(x) = \frac{x^2-1}{x-1} \text{ if } x \neq 1, \text{ and } f(1) = 0. \tag{4.1.E.3}$$

Show that $\lim_{x \rightarrow 1} f(x) = 2$ exists, yet f is discontinuous at $p = 1$. Make it continuous by redefining $f(1)$.
[Hint: For $x \neq 1, f(x) = x + 1$. Proceed as in Example (b), using the deleted globe $G_{\neg p}(\delta)$.]

Exercise 4.1.E.8

Find $\lim_{x\to p} f(x)$ and check continuity at p in the following cases, assuming that $D_f = A$ is the set of all $x \in E^1$ for which the given expression for $f(x)$ has sense. Specify that set.

$$\begin{array}{ll}\text{(a) } \lim_{x\to 2}\left(2x^2-3x-5\right); \quad \text{(b) } \lim_{x\to 1}\frac{3x+2}{2x-1} & \\ \text{(c) } \lim_{x\to -1}\left(\frac{x^2-4}{x+2}-1\right); & \text{(d) } \lim_{x\to 2}\frac{x^3-8}{x-2} \\ \text{(e) } \lim_{x\to a}\frac{x^4-a^4}{x-a}; & \text{(f) } \lim_{x\to 0}\left(\frac{x}{x+1}\right)^3 \\ \text{(g) } \lim_{x\to -1}\left(\frac{1}{x^2+1}\right)^2 & \end{array} \tag{4.1.E.4}$$

[Example solution: Find $\lim_{x\to 1}\frac{5x^2-1}{2x+3}$.
Here

$$f(x)=\frac{5x^2-1}{2x+3}; A=E^1-\left\{-\frac{3}{2}\right\}; p=1. \tag{4.1.E.5}$$

We show that f is continuous at p, and so (by Corollary 2)

$$f(x)=\frac{5x^2-1}{2x+3}; A=E^1-\{-\frac{3}{2}\}; p=1. \tag{4.1.E.6}$$

We show that f is continuous at p, and so (by Corollary 2)

$$\lim_{x\to p} f(x)=f(p)=f(1)=\frac{4}{5}. \tag{4.1.E.7}$$

Using formula (1), we fix an arbitrary $\varepsilon>0$ and look for a δ such that

$$(\forall x\in A\cap G_p(\delta)) \quad \rho(f(x), f(1))=|f(x)-f(1)|<\varepsilon, \text{ i.e., }, \left|\frac{5x^2-1}{2x+3}-\frac{4}{5}\right|<\varepsilon; \tag{4.1.E.8}$$

or, by putting everything over a common denominator and using properties of absolute values,

$$|x-1|\frac{|25x+17|}{5|2x+3|}<\varepsilon \text{ whenever } |x-1|<\delta \text{ and } x\in A. \tag{4.1.E.9}$$

(Usually in such problems, it is desirable to factor out $x-p$.)
By Note 4, we may assume $0<\delta\leq 1$. Then $|x-1|<\delta$ implies $-1\leq x-1\leq 1$ i.e., $0\leq x\leq 2$, so

$$5|2x+3|\geq 15 \text{ and } |25x+17|\leq 67. \tag{4.1.E.10}$$

Hence (6) will certainly hold if

$$|x-1|\frac{67}{15}<\varepsilon, \text{ i.e., if } |x-1|<\frac{15\varepsilon}{67}. \tag{4.1.E.11}$$

To achieve it, we choose $\delta=\min(1, 15\varepsilon/67)$. Then, reversing all steps, we obtain (6), and hence $\lim_{x\to 1} f(x)=f(1)=4/5$.]

Exercise $4.1.E.9$

Find (using definitions, such as (3))

$$\begin{array}{ll}\text{(a) } \lim_{x\to+\infty}\frac{1}{x}; & \text{(b) } \lim_{x\to-\infty}\frac{3x+2}{2x-1};\\ \text{(c) } \lim_{x\to+\infty}\frac{x^3}{1-x^2}; & \text{(d) } \lim_{x\to3^+}\frac{x-1}{x-3};\\ \text{(e) } \lim_{x\to3^-}\frac{x-1}{x-3}; & \text{(f) } \lim_{x\to3}\left|\frac{x-1}{x-3}\right|.\end{array} \tag{4.1.E.12}$$

Exercise $4.1.E.10$

Prove that if

$$\lim_{x\to p} f(x)=\overline{q}\in E^n\left({}^*C^n\right), \tag{4.1.E.13}$$

then for each scalar c,

$$\lim_{x\to p} cf(x)=c\overline{q}. \tag{4.1.E.14}$$

Exercise $4.1.E.11$

Define $f: E^1\to E^1$ by

$$f(x)=x\cdot\sin\frac{1}{x} \text{ if } x\neq 0, \text{ and } f(0)=0. \tag{4.1.E.15}$$

Show that f is continuous at $p=0$, i.e.,

$$\lim_{x\to 0} f(x)=f(0)=0. \tag{4.1.E.16}$$

Draw an approximate graph (it is contained between the lines $y=\pm x$).
$\left[\text{Hint: } \left|x\cdot\sin\frac{1}{x}-0\right|\leq|x|.\right]$

Exercise $4.1.E.*12$

Discuss the statement: f is continuous at p iff

$$\left(\forall G_{f(p)}\right)\left(\exists G_p\right) \quad f\left[G_p\right]\subseteq G_{f(p)}. \tag{4.1.E.17}$$

Exercise $4.1.E.13$

Define $f: E^1\to E^1$ by

$$f(x)=x \text{ if } x \text{ is rational} \tag{4.1.E.18}$$

and

$$f(x)=0 \text{ otherwise }. \tag{4.1.E.19}$$

Show that f is continuous at 0 but nowhere else. How about relative continuity?

Exercise $4.1.E.14$

Let $A=(0,+\infty)\subset E^1$. Define $f:A\rightarrow E^1$ by

$$f(x)=0 \text{ if } x \text{ is irrational} \tag{4.1.E.20}$$

and

$$f(x)=\frac{1}{n} \text{ if } x=\frac{m}{n} (\text{in lowest terms }) \tag{4.1.E.21}$$

for some natural m and n. Show that f is continuous at each irrational, but at no rational, point $p\in A$.
[Hints: If p is irrational, fix $\varepsilon>0$ and an integer $k>1/\varepsilon$. In $G_p(1)$, there are only finitely many irreducible fractions

$$\frac{m}{n}>0 \text{ with } n\leq k, \tag{4.1.E.22}$$

so one of them, call it r, is closest to p. Put

$$\delta=\min(1,|r-p|) \tag{4.1.E.23}$$

and show that

$$(\forall x\in A\cap G_p(\delta)) \quad |f(x)-f(p)|=f(x)<\varepsilon, \tag{4.1.E.24}$$

distinguishing the cases where x is rational and irrational.
If p is rational, use the fact that each $G_p(\delta)$ contains irrationals x at which

$$f(x)=0\Longrightarrow|f(x)-f(p)|=f(p). \tag{4.1.E.25}$$

Take $\varepsilon<f(p)$.]

Exercise $4.1.E.15$

Given two reals, $p>0$ and $q>0$, define $f:E^1\rightarrow E^1$ by

$$f(0)=0 \text{ and } f(x)=\left(\frac{x}{p}\right)\cdot\left[\frac{q}{x}\right] \text{ if } x\neq 0. \tag{4.1.E.26}$$

here $[q/x]$ is the integral part of q/x.
(i) Is f left or right continuous at 0?
(ii) Same question with $f(x)=[x/p](q/x)$.

Exercise $4.1.E.16$

Prove that if (S,ρ) is discrete, then all functions $f:S\rightarrow(T,\rho')$ are continuous. What if (T,ρ') is discrete but (S,ρ) is not?

4.2: Some General Theorems on Limits and Continuity

I. In §1 we gave the so-called "ε, δ" definition of continuity. Now we present another (equivalent) formulation, known as the sequential one. Roughly, it states that f is continuous iff it carries convergent sequences $\{x_m\} \subseteq D_f$ into convergent "image sequences" $\{f(x_m)\}$. More precisely, we have the following theorem.

Theorem 4.2.1 (sequential criterion of continuity).

(i) A function

$$f : A \rightarrow (T, \rho'), \text{ with } A \subseteq (S, \rho), \tag{4.2.1}$$

is continuous at a point $p \in A$ iff for every sequence $\{x_m\} \subseteq A$ such that $x_m \rightarrow p$ in (S, ρ), we have $f(x_m) \rightarrow f(p)$ in (T, ρ') . In symbols,

$$(\forall \{x_m\} \subseteq A | x_m \rightarrow p) \quad f(x_m) \rightarrow f(p). \tag{4.2.2}$$

(ii) Similarly, a point $q \in T$ is a limit of f at $p (p \in S)$ iff

$$(\forall \{x_m\} \subseteq A - \{p\} | x_m \rightarrow p) \quad f(x_m) \rightarrow q. \tag{4.2.3}$$

Note that in (2') we consider only sequences of terms other than p.

Proof

We first prove (ii). Suppose q is a limit of f at p, i.e. (see §1),

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in A \cap G_{\neg p}(\delta)) \quad f(x) \in G_q(\varepsilon). \tag{4.2.4}$$

Thus, given $\varepsilon > 0$, there is $\delta > 0$ (henceforth fixed) such that

$$f(x) \in G_q(\varepsilon) \text{ whenever } x \in A, x \neq p, \text{ and } x \in G_p(\delta). \tag{4.2.5}$$

We want to deduce (2'). Thus we fix any sequence

$$\{x_m\} \subseteq A - \{p\}, x_m \rightarrow p. \tag{4.2.6}$$

Then

$$(\forall m) \quad x_m \in A \text{ and } x_m \neq p, \tag{4.2.7}$$

and $G_p(\delta)$ contains all but finitely many x_m. Then these x_m satisfy the conditions stated in **(3)**. Hence $f(x_m) \in G_q(\varepsilon)$ for all but finitely many m. As ε is arbitrary, this implies $f(x_m) \rightarrow q$ (by the definition of $\lim_{m \rightarrow \infty} f(x_m)$), as is required in **(2')**. Thus **(2)** $\Longrightarrow$ **(2')**.

Conversely, suppose **(2)** fails, i.e., its negation holds. (See the rules for forming negations of such formulas in Chapter 1, §§1-3.) Thus

$$(\exists \varepsilon > 0)(\forall \delta > 0)(\exists x \in A \cap G_{\neg p}(\delta)) \quad f(x) \notin G_q(\varepsilon) \tag{4.2.8}$$

by the rules for quantifiers. We fix an ε **satisfying (4),** and let

$$\delta_m = \frac{1}{m}, \quad m = 1, 2, \ldots \tag{4.2.9}$$

By (4), for each δ_m there is x_m (depending on δ_m) such that

$$x_m \in A \cap G_{\neg p}\left(\frac{1}{m}\right) \tag{4.2.10}$$

and

$$f(x_m) \notin G_q(\varepsilon), \quad m = 1, 2, 3, \ldots \tag{4.2.11}$$

We fix these x_m. As $x_m \in A$ and $x_m \neq p$, we obtain a sequence

$$\{x_m\} \subseteq A - \{p\}. \tag{4.2.12}$$

Also, as $x_m \in G_p\left(\frac{1}{m}\right)$, we have $\rho(x_m, p) < 1/m \to 0$, and hence $x_m \to p$. On the other hand, **by (6)**, the image sequence $\{f(x_m)\}$ canverge to q (why?), i.e., **(2')** fails. Thus we see that **(2')** fails or holds accordingly as **(2)** does.

This proves assertion (ii). Now, by setting $q = f(p)$ in **(2) and (2')**, we also obtain the first clause of the theorem, as to continuity. □

Note 1. The theorem also applies to relative limits and continuity over a path B (just replace A by B in the proof), as well as to the cases $p = \pm\infty$ and $q = \pm\infty$ in E^* (for E^* can be treated as a metric space; see the end of Chapter 3, §11).

If the range space (T, ρ') is complete (Chapter 3, §17), then the image sequences $\{f(x_m)\}$ converge iff they are Cauchy. This leads to the following corollary.

Corollary 1. Let (T, ρ') be complete, such as E^n. Let a map $f : A \to T$ with $A \subseteq (S, \rho)$ and a point $p \in S$ be given. Then for f to have a limit at p, it suffices that $\{f(x_m)\}$ be Cauchy in (T, ρ') whenever $\{x_m\} \subseteq A - \{p\}$ and $x_m \to p$ in (S, ρ).

Indeed, as noted above, all such $\{f(x_m)\}$ converge. Thus it only remains to show that they tend to one and the same limit q, as is required in part (ii) of Theorem 1. We leave this as an exercise (Problem 1 below).

Theorem 4.2.2 (Cauchy criterion for functions).

With the assumptions of Corollary 1, the function f has a limit at p iff for each $\varepsilon > 0$, there is $\delta > 0$ such that

$$\rho'(f(x), f(x')) < \varepsilon \text{ for all } x, x' \in A \cap G_{\neg p}(\delta). \tag{4.2.13}$$

In symbols,

$$(\forall \varepsilon > 0)(\exists \delta > 0)\left(\forall x, x' \in A \cap G_{\neg p}(\delta)\right) \quad \rho'(f(x), f(x')) < \varepsilon. \tag{4.2.14}$$

Proof

Assume (7). To show that f has a limit at p, we use Corollary 1. Thus we take any sequence

$$\{x_m\} \subseteq A - \{p\} \text{ with } x_m \to p \tag{4.2.15}$$

and show that $\{f(x_m)\}$ is Cauchy, i.e.,

$$(\forall \varepsilon > 0)(\exists k)(\forall m, n > k) \quad \rho'(f(x_m), f(x_n)) < \varepsilon. \tag{4.2.16}$$

To do this, fix an arbitrary $\varepsilon > 0$. **By (7), we have**

$$\left(\forall x, x' \in A \cap G_{\neg p}(\delta)\right) \quad \rho'(f(x), f(x')) < \varepsilon, \tag{4.2.17}$$

for some $\delta > 0$. Now as $x_m \to p$, there is k such that

$$(\forall m, n > k) \quad x_m, x_n \in G_p(\delta). \tag{4.2.18}$$

As $\{x_m\} \subseteq A - \{p\}$, we even have $x_m, x_n \in A \cap G_{\neg p}(\delta)$. **Hence by (7'),**

$$(\forall m, n > k) \quad \rho'(f(x_m), f(x_n)) < \varepsilon; \tag{4.2.19}$$

i.e., $\{f(x_m)\}$ is Cauchy, as required in Corollary 1, and so f has a limit at p. **This shows that (7)** implies the existence of that limit.

The easy converse proof is left to the reader. (See Problem 2.) □

II. Composite Functions. The composite of two functions

$$f : S \to T \text{ and } g : T \to U, \tag{4.2.20}$$

denoted

$$g \circ f \quad \text{(in that order)}, \tag{4.2.21}$$

is by definition a map of S into U given by

$$(g \circ f)(x)=g(f(x)), \quad x \in S. \tag{4.2.22}$$

Our next theorem states, roughly, that $g \circ f$ is continuous if g and f are. We shall use Theorem 1 to prove it.

Theorem 4.2.3

Let $(S, \rho),(T, \rho')$, and (U, ρ'') be metric spaces. If a function $f: S \rightarrow T$ is continuous at a point $p \in S$, and if $g: T \rightarrow U$ is continuous at the point $q=f(p)$, then the composite function $g \circ f$ is continuous at p.

Proof

The domain of $g \circ f$ is S. So take any sequence

$$\left\{x_m\right\} \subseteq S \text{ with } x_m \rightarrow p. \tag{4.2.23}$$

As f is continuous at p, **formula (1')** yields $f\left(x_m\right) \rightarrow f(p)$, where $f\left(x_m\right)$ is in $T=D_g$. Hence, as g is continuous at $f(p)$, we have

$$g\left(f\left(x_m\right)\right) \rightarrow g(f(p)), \text{ i.e., }(g \circ f)\left(x_m\right) \rightarrow(g \circ f)(p), \tag{4.2.24}$$

and this holds for any $\left\{x_m\right\} \subseteq S$ with $x_m \rightarrow p$. Thus $g \circ f$ satisfies condition (1') and is continuous at p. □

Caution: The fact that

$$\lim_{x \rightarrow p} f(x)=q \text{ and } \lim_{y \rightarrow q} g(y)=r \tag{4.2.25}$$

does not imply

$$\lim_{x \rightarrow p} g(f(x))=r \tag{4.2.26}$$

(see Problem 3 for counterexamples).

Indeed, if $\left\{x_m\right\} \subseteq S-\{p\}$ and $x_m \rightarrow p$, we obtain, as before, $f\left(x_m\right) \rightarrow q$, but not $f\left(x_m\right) \neq q$. Thus we cannot r**e-apply formula (2')** to obtain $g\left(f\left(x_m\right)\right) \rightarrow r$ **since (2')** requires that $f\left(x_m\right) \neq q$. The argument still works if g is continuous at q (**then (1') applies**) or if $f(x)$ never equals q then $f(x_m) \neq q$. It even suffices that $f(x) \neq q$ for x in some deleted globe about p((see §1, Note 4). Hence we obtain the following corollary.

Corollary 2. With the notation of Theorem 3, suppose

$$f(x) \rightarrow q \backslash) as \backslash(x \rightarrow p, \text{ and } g(y) \rightarrow r \text{ as } y \rightarrow q. \tag{4.2.27}$$

Then

$$g(f(x)) \rightarrow r \text{ as } x \rightarrow p, \tag{4.2.28}$$

provided, however, that

(i) g is continuous at q, or

(ii) $f(x) \neq q$ for x in some deleted globe about p, or

(iii) f is one to one, at least when restricted to some $G_{\neg p}(\delta)$.

Indeed, (i) and (ii) suffice, as was explained above. Thus assume (iii). Then f can take the value q at most once, say, at some point

$$x_0 \in G_{\neg p}(\delta). \tag{4.2.29}$$

As $x_0 \neq p$, let

$$\delta'=\rho\left(x_0, p\right)>0. \tag{4.2.30}$$

Then $x_0 \notin G_{\neg p}\left(\delta'\right)$, so $f(x) \neq q$ on $G_{\neg p}\left(\delta'\right)$, and case (iii) reduces to (ii).

We now show how to apply Corollary 2.

Note 2. Suppose we know that

$$r=\lim_{y \rightarrow q} g(y) \text{ exists.} \tag{4.2.31}$$

Using this fact, we often pass to another variable x, setting $y=f(x)$ where f is such that $q=\lim_{x \rightarrow p} f(x)$ for some p. We shall say that the substitution (or "change of variable") $y=f(x)$ is admissible if one of the conditions (i), (ii), or (iii) of Corollary 2 holds. Then by Corollary 2,

$$\lim_{y \rightarrow q} g(y)=r=\lim_{x \rightarrow p} g(f(x)) \tag{4.2.32}$$

(yielding the second limit).

Example 4.2.1

(A) Let

$$h(x)=\left(1+\frac{1}{x}\right)^x \text{ for }|x| \geq 1. \tag{4.2.33}$$

Then

$$\lim_{x \rightarrow+\infty} h(x)=e. \tag{4.2.34}$$

For a proof, let $n=f(x)=[x]$ be the integral part of x. Then for $x>1$,

$$\left(1+\frac{1}{n+1}\right)^n \leq h(x) \leq\left(1+\frac{1}{n}\right)^{n+1} . \quad(\text { Verify! }) \tag{4.2.35}$$

As $x \rightarrow+\infty, n$ tends to $+\infty$ over integers, and by rules for sequences,

$$\lim_{n \rightarrow \infty}\left(1+\frac{1}{n}\right)^{n+1}=\lim_{n \rightarrow \infty}\left(1+\frac{1}{n}\right)\left(1+\frac{1}{n}\right)^n=1 \cdot \lim_{n \rightarrow \infty}\left(1+\frac{1}{n}\right)^n=1 \cdot e=e, \tag{4.2.36}$$

with e as in Chapter 3, §15. Similarly one shows that also

$$\lim_{n \rightarrow \infty}\left(1+\frac{1}{n+1}\right)^n=e. \tag{4.2.37}$$

Thus (8) implies that also $\lim_{x \rightarrow+\infty} h(x)=e$ (see Problem 6 below).

Remark. Here we used Corollary 2(ii) with

$$f(x)=[x], q=+\infty, \text{ and } g(n)=\left(1+\frac{1}{n}\right)^n . \tag{4.2.38}$$

The substitution $n=f(x)$ is admissible since $f(x)=n$ never equals $+\infty$, its limit, thus satisfying Corollary 2(ii).

(B) Quite similarly, one shows that also

$$\lim_{x \rightarrow-\infty}\left(1+\frac{1}{x}\right)^x=e. \tag{4.2.39}$$

See Problem 5.

(C) In Examples (A) and (B), we now substitute $x=1 / z$. This is admissible by Corollary 2(ii) since the dependence between x and z is one to one. Then

$$z=\frac{1}{x} \rightarrow 0^{+} \text{ as } x \rightarrow+\infty, \text{ and } z \rightarrow 0^{-} \text{ as } x \rightarrow-\infty. \tag{4.2.40}$$

Thus (A) and (B) yield

$$\lim_{z \rightarrow 0^{+}}(1+z)^{1 / z}=\lim_{z \rightarrow 0^{-}}(1+z)^{1 / z}=e. \tag{4.2.41}$$

Hence by Corollary 3 of §1, we obtain

$$\lim_{z \rightarrow 0}(1+z)^{1 / z}=e. \tag{4.2.42}$$

4.2.E: More Problems on Limits and Continuity

Exercise $4.2.E.1$

Complete the proof of Corollary 1.
$[$ Hint: Consider $\{f(x_m)\}$ and $\{f(x_m')\}$, with

$$x_m \rightarrow p \text{ and } x_m' \rightarrow p. \tag{4.2.E.1}$$

By Chapter 3, §14, Corollary 4, p is also the limit of

$$x_1, x_1', x_2, x_2', \ldots, \tag{4.2.E.2}$$

so, by assumption,

$$f(x_1), f(x_1'), \ldots \text{ converges (to } q, \text{ say).} \tag{4.2.E.3}$$

Hence $\{f(x_m)\}$ and $\{f(x_m')\}$ must have the same limit q. (Why?)]

Exercise $4.2.E.*2$

Complete the converse proof of Theorem 2 (cf. proof of Theorem 1 in Chapter 3, §17).

Exercise $4.2.E.3$

Define $f, g: E^1 \rightarrow E^1$ by setting
(i) $f(x)=2; g(y)=3$ if $y \neq 2$, and $g(2)=0$; or
(ii) $f(x)=2$ if x is rational and $f(x)=2x$ otherwise; g as in (i).
In both cases, show that

$$\lim_{x \rightarrow 1} f(x)=2 \text{ and } \lim_{y \rightarrow 2} g(y)=3 \text{ but not } \lim_{x \rightarrow 1} g(f(x))=3. \tag{4.2.E.4}$$

Exercise $4.2.E.4$

Prove Theorem 3 from "ε, δ " definitions. Also prove (both ways) that if f is relatively continuous on B, and g on $f[B]$, then $g \circ f$ is relatively continuous on B.

Exercise $4.2.E.5$

Complete the missing details in Examples (A) and (B).
[Hint for (B): Verify that

$$\left(1-\frac{1}{n+1}\right)^{-n-1}=\left(\frac{n}{n+1}\right)^{-n-1}=\left(\frac{n+1}{n}\right)^{n+1}=\left(1+\frac{1}{n}\right)\left(1+\frac{1}{n}\right)^n \rightarrow e.] \tag{4.2.E.5}$$

Exercise $4.2.E.6$

$\Rightarrow 6.$ Given $f, g, h: A \rightarrow E^*, A \subseteq (S, \rho)$, with

$$f(x) \leq h(x) \leq g(x) \tag{4.2.E.6}$$

for $x \in G_{\neg p}(\delta) \cap A$ for some $\delta > 0$. Prove that if

$$\lim_{x \rightarrow p} f(x) = \lim_{x \rightarrow p} g(x) = q, \tag{4.2.E.7}$$

also then

$$\lim_{x \rightarrow p} h(x) = q. \tag{4.2.E.8}$$

Use Theorem 1.
[Hint: Take any

$$\{x_m\} \subseteq A - \{p\} \text{ with } x_m \rightarrow p. \tag{4.2.E.9}$$

Then $f\left(x_m\right) \rightarrow q, g\left(x_m\right) \rightarrow q,$ and

$$\left(\forall x_m \in A \cap G_{\neg p}(\delta)\right) \quad f\left(x_m\right) \leq h\left(x_m\right) \leq g\left(x_m\right). \tag{4.2.E.10}$$

Now apply Corollary 3 of Chapter $3, \S 15.$]

Exercise $4.2.E.7$

$\Rightarrow 7.$ Given $f, g : A \rightarrow E^*, A \subseteq (S, \rho),$ with $f(x) \rightarrow q$ and $g(x) \rightarrow r$ as $x \rightarrow p$ $(p \in S)$, prove the following:
(i) If $q > r$, then

$$(\exists \delta > 0)\left(\forall x \in A \cap G_{\neg p}(\delta)\right) \quad f(x) > g(x). \tag{4.2.E.11}$$

(ii) (Passage to the limit in inequalities.) If

$$(\forall \delta > 0)\left(\exists x \in A \cap G_{\neg p}(\delta)\right) \quad f(x) \leq g(x), \tag{4.2.E.12}$$

then $q \leq r$. (Observe that here A clusters at p necessarily, so the limits are unique.)
[Hint: Proceed as in Problem 6; use Corollary 1 of Chapter $3, \S 15.$]

Exercise $4.2.E.8$

Do Problems 6 and 7 using only Definition 2 of §1.
$[$ Hint: Here prove $7($ ii $)$ first. $]$

Exercise $4.2.E.9$

Do Examples $(a)-(d)$ of §1 using Theorem 1.
[Hint: For (c), use also Example (a) in Chapter $3, \S 16.$]

Exercise $4.2.E.10$

Addition and multiplication in E^1 may be treated as functions

$$f, g : E^2 \rightarrow E^1 \tag{4.2.E.13}$$

with

$$f(x, y)=x+y \text{ and } g(x, y)=x y . \tag{4.2.E.14}$$

Show that f and g are continuous on E^2 (see footnote 2 in Chapter 3 §15). Similarly, show that the standard metric

$$\rho(x, y)=|x-y| \tag{4.2.E.15}$$

is a continuous mapping from E^2 to E^1.
$[\text{Hint: Use Theorems } 1, 2, \text{ and, } 4 \text{ of Chapter } 3, \S 15 \text{ and the sequential criterion.}]$

Exercise $4.2.E.11$

Using Corollary 2 and formula (9), find $\lim_{x \rightarrow 0}(1 \pm m x)^{1/x}$ for a fixed $m \in N$.

Exercise $4.2.E.12$

$\Rightarrow 12$. Let $a>0$ in E^1. Prove that $\lim_{x \rightarrow 0} a^x=1$.
$\left[\text{Hint: Let } n=f(x) \text{ be the integral part of } \frac{1}{x}(x \neq 0).\right.$ Verify that

$$a^{-1/(n+1)} \leq a^x \leq a^{1/n} \text{ if } a \geq 1, \tag{4.2.E.16}$$

with inequalities reversed if $0<a<1$. Then proceed as in Example (A), noting that

$$\lim_{n \rightarrow \infty} a^{1/n}=1=\lim_{n \rightarrow \infty} a^{-1/(n+1)} \tag{4.2.E.17}$$

by Problem 20 of Chapter $3, \S 15$. $(\text{Explain!})]$

Exercise $4.2.E.13$

$\Rightarrow 13$. Given $f, g: A \rightarrow E^*, A \subseteq(S, \rho)$, with

$$f \leq g \quad \text{ for } x \text{ in } G_{\neg p}(\delta) \cap A. \tag{4.2.E.18}$$

Prove that
(a) if $\lim_{x \rightarrow p} f(x)=+\infty$, then also $\lim_{x \rightarrow p} g(x)=+\infty$;
(b) if $\lim_{x \rightarrow p} g(x)=-\infty$, then also $\lim_{x \rightarrow p} f(x)=-\infty$.
Do it it two ways:
(i) Use definitions only, such as $\left(2^{\prime}\right)$ in $\S 1$.
(ii) Use Problem 10 of Chapter 2, §13 and the sequential criterion.

Exercise $4.2.E.14$

$\Rightarrow 14$. Prove that
(i) if $a>1$ in E^1, then

$$\lim_{x \rightarrow+\infty} \frac{a^x}{x}=+\infty \text{ and } \lim_{x \rightarrow+\infty} \frac{a^{-x}}{x}=0; \tag{4.2.E.19}$$

(ii) if $0<a<1$, then

$$\lim_{x \rightarrow+\infty} \frac{a^x}{x}=0 \text{ and } \lim_{x \rightarrow+\infty} \frac{a^{-x}}{x}=+\infty; \tag{4.2.E.20}$$

(iii) if $a>1$ and $0 \leq q \in E^{1}$, then

$$\lim_{x \rightarrow+\infty} \frac{a^{x}}{x^{q}}=+\infty \text { and } \lim_{x \rightarrow+\infty} \frac{a^{-x}}{x^{q}}=0; \tag{4.2.E.21}$$

(iv) if $0<a<1$ and $0 \leq q \in E^{1}$, then

$$\lim_{x \rightarrow+\infty} \frac{a^{x}}{x^{q}}=0 \text { and } \lim_{x \rightarrow+\infty} \frac{a^{-x}}{x^{q}}=+\infty. \tag{4.2.E.22}$$

[Hint: (i) From Problems 17 and 10 of Chapter $3, \S 15$, obtain

$$\lim \frac{a^{n}}{n}=+\infty. \tag{4.2.E.23}$$

Then proceed as in Examples $(\mathrm{A})-(\mathrm{C})$; (iii) reduces to (i) by the method used in Problem 18 of Chapter $3, \S 15$.]

Exercise $4.2.E.15$

$\Rightarrow *15$. For a map $f:(S, \rho) \rightarrow\left(T, \rho^{\prime}\right)$, show that the following statements are equivalent:
(i) f is continuous on S.
(ii) $(\forall A \subseteq S) f[\overline{A}] \subseteq \overline{f[A]}$.
(iii) $(\forall B \subseteq T) f^{-1}[\overline{B}] \supseteq \overline{f^{-1}[B]}$.
(iv) $f^{-1}[B]$ is closed in (S, ρ) whenever B is closed in $\left(T, \rho^{\prime}\right)$.
(v) $f^{-1}[B]$ is open in (S, ρ) whenever B is open in $\left(T, \rho^{\prime}\right)$.
[Hints: $(i) \Longrightarrow$ (ii): Use Theorem 3 of Chapter } 3, §16 and the sequential criterion to show that

$$p \in \overline{A} \Longrightarrow f(p) \in \overline{f[A]}. \tag{4.2.E.24}$$

(ii) $\Longrightarrow$ (iii) : Let $A=f^{-1}[B]$. Then $f[A] \subseteq B$, so by (ii),

$$f[\overline{A}] \subseteq \overline{f[A]} \subseteq \overline{B}. \tag{4.2.E.25}$$

Hence

$$\overline{f^{-1}[B]}=\overline{A} \subseteq f^{-1}[f[\overline{A}]] \subseteq f^{-1}[\overline{B}]. \quad \text{(Why?)} \tag{4.2.E.26}$$

(iii) $\Longrightarrow$ (iv) : If B is closed, $B=\overline{B}$ (Chapter 3, §16, Theorem 4(ii)), so by (iii),

$$f^{-1}[B]=f^{-1}[\overline{B}] \supseteq \overline{f^{-1}[B]}; \text{ deduce (iv) }. \tag{4.2.E.27}$$

(iv) $\Longrightarrow$ (v) : Pass to complements in (iv).
(v) $\Longrightarrow$ (i) : Assume (v). Take any $p \in S$ and use Definition 1 in $\S 1$.]

Exercise $4.2.E.16$

Let $f: E^{1} \rightarrow E^{1}$ be continuous. Define $g: E^{1} \rightarrow E^{2}$ by

$$g(x)=(x, f(x)). \tag{4.2.E.28}$$

Prove that
(a) g and g^{-1} are one to one and continuous;
(b) the range of g, i.e., the set

$$D_g' = \left\{(x, f(x)) | x \in E^1\right\}, \tag{4.2.E.29}$$

is closed in E^2.
[Hint: Use Theorem 2 of Chapter 3, §15, Theorem 4 of Chapter 3, §16, and the sequential criterion.]

4.3: Operations on Limits. Rational Functions

I. A function $f: A \rightarrow T$ is said to be real if its range D_f' lies in E^1, complex if $D_f' \subseteq C$, vector valued if D_f' is a subset of E^n, and scalar valued if D_f' lies in the scalar field of E^n. ('ln the latter two cases, we use the same terminology if E^n is replaced by some other (fixed) normed space under consideration.) The domain A may be arbitrary.

For such functions one can define various operations whenever they are defined for elements of their ranges, to which the function values $f(x)$ belong. Thus as in Chapter 3, §9, we define the functions $f \pm g, fg$, and f/g "pointwise," setting

$$(f \pm g)(x)=f(x) \pm g(x), \quad(f g)(x)=f(x) g(x), \text { and }\left(\frac{f}{g}\right)(x)=\frac{f(x)}{g(x)} \tag{4.3.1}$$

whenever the right side expressions are defined. We also define $|f|: A \rightarrow E^1$ by

$$(\forall x \in A) \quad|f|(x)=|f(x)|. \tag{4.3.2}$$

In particular, $f \pm g$ is defined if f and g are both vector valued or both scalar valued, and fg is defined if f is vector valued while g is scalar valued; similarly for f/g. (However, the domain of f/g consists of those $x \in A$ only for which $g(x) \neq 0.$)

In the theorems below, all limits are at some (arbitrary, but fixed) point p of the domain space (S, ρ). For brevity, we often omit " $x \rightarrow p.$ "

Theorem 4.3.1

For any functions $f, g, h: A \rightarrow E^1(C), A \subseteq(S, \rho),$ we have the following:

i. (i) If f, g, h are continuous at $p(p \in A)$, so are $f \pm g$ and fh. So also is f/h, provided $h(p) \neq 0$; similarly for relative continuity over $B \subseteq A$.
ii. (ii) If $f(x) \rightarrow q, g(x) \rightarrow r,$ and $h(x) \rightarrow a($ all, as $x \rightarrow p$ over $B \subseteq A),$ then
    a. $f(x) \pm g(x) \rightarrow q \pm r$
    b. $f(x) h(x) \rightarrow q a$; and
    c. $\frac{f(x)}{h(x)} \rightarrow \frac{q}{a}$, provided $a \neq 0$

All this holds also if f and g are vector valued and h is scalar valued.

For a simple proof, one can use Theorem 1 of Chapter 3, §15. (An independent proof is sketched in Problems 1-7 below.)

We can also use the sequential criterion (Theorem 1 in §2). To prove (ii), take any sequence

$$\left\{x_m\right\} \subseteq B-\{p\}, x_m \rightarrow p \tag{4.3.3}$$

Then by the assumptions made,

$$f\left(x_m\right) \rightarrow q, g\left(x_m\right) \rightarrow r, \text { and } h\left(x_m\right) \rightarrow a \tag{4.3.4}$$

Thus by Theorem 1 of Chapter 3, §15,

$$f\left(x_m\right) \pm g\left(x_m\right) \rightarrow q \pm r, f\left(x_m\right) g\left(x_m\right) \rightarrow q a, \text { and } \frac{f\left(x_m\right)}{g\left(x_m\right)} \rightarrow \frac{q}{a} \tag{4.3.5}$$

As this holds for any sequence $\left\{x_m\right\} \subseteq B-\{p\}$ with $x_m \rightarrow p$, our assertion (ii) follows by the sequential criterion; similarly for (i).

Note 1. By induction, the theorem also holds for sums and products of any finite number of functions (whenever such products are defined).

Note 2. Part (ii) does not apply to infinite limits q, r, a; but it does apply to limits at $p=\pm \infty$ (take E^* with a suitable metric for the space S).

Note 3. The assumption $h(x) \rightarrow a \neq 0($ as $x \rightarrow p$ over $B)$ implies that $h(x) \neq 0$ for x in $B \cap G_{\neg p}(\delta)$ for some $\delta>0$; see Problem 5 below. Thus the quotient function f/h is defined on $B \cap G_{\neg p}(\delta)$ at least.

II. If the range space of f is E^n (* or C^n), then each function value $f(x)$ is a vector in that space; thus n real (* respectively, complex) components, denoted

$$f_k(x), \quad k=1,2,\ldots,n. \tag{4.3.6}$$

Here we may treat f_k as a mapping of $A=D_f$ into E^1($*$ or C); it carries each point $x \in A$ into $f_k(x)$, the k th component of $f(x)$. In this manner, each function

$$f: A \rightarrow E^n\left({ }^* C^n\right) \tag{4.3.7}$$

uniquely determines n scalar-valued maps

$$f_k: A \rightarrow E^1(C) \tag{4.3.8}$$

called the components of f. Notation: $f=(f_1,\ldots,f_n)$.

Conversely, given n arbitrary functions

$$f_k: A \rightarrow E^1(C), \quad k=1,2,\ldots,n, \tag{4.3.9}$$

one can define $f: A \rightarrow E^n\left({ }^* C^n\right)$ by setting

$$f(x)=\left(f_1(x), f_2(x), \ldots, f_n(x)\right). \tag{4.3.10}$$

Then obviously $f=\left(f_1, f_2, \ldots, f_n\right)$. Thus the f_k in turn determine f uniquely. To define a function $f: A \rightarrow E^n\left({ }^* C^n\right)$ means to give its n components f_k. Note that

$$f(x)=\left(f_1(x), \ldots, f_n(x)\right)=\sum_{k=1}^n \bar{e}_k f_k(x), \quad \text{i.e., } f=\sum_{k=1}^n \bar{e}_k f_k \tag{4.3.11}$$

where the $\bar{e}_k$ are the n basic unit vectors; see Chapter 3, §1-3, Theorem 2. Our next theorem shows that the limits and continuity of f reduce to those of the f_k.

Theorem 4.3.2

(componentwise continuity and limits). For any function $f: A \rightarrow E^n\left(* C^n\right)$, with $A \subseteq(S, \rho)$ and with $f=\left(f_1, \ldots, f_n\right)$, we have that

(i) f is continuous at $p (p \in A)$ iff all its components f_k are, and

(ii) $f(x) \rightarrow \bar{q}$ as $x \rightarrow p (p \in S)$ iff

$$f_k(x) \rightarrow q_k \text{ as } x \rightarrow p \quad (k=1,2,\ldots,n), \tag{4.3.12}$$

i.e., iff each f_k has, as its limit at p, the corresponding component of $\bar{q}$.

Similar results hold for relative continuity and limits over a path $B \subseteq A$.

We prove (ii). If $f(x) \rightarrow \bar{q}$ as $x \rightarrow p$ then, by definition,

$$(\forall \varepsilon>0)(\exists \delta>0)\left(\forall x \in A \cap G_{\neg p}(\delta)\right) \quad \varepsilon>|f(x)-\bar{q}|=\sqrt{\sum_{k=1}^n\left|f_k(x)-q_k\right|^2}; \tag{4.3.13}$$

in turn, the right-hand side of the inequality given above is no less than each

$$\left|f_k(x)-q_k\right|, \quad k=1,2,\ldots,n. \tag{4.3.14}$$

Thus

$$(\forall \varepsilon>0)(\exists \delta>0)\left(\forall x \in A \cap G_{\neg p}(\delta)\right) \quad\left|f_k(x)-q_k\right|<\varepsilon; \tag{4.3.15}$$

i.e., $f_k(x) \rightarrow q_k, k=1,\ldots,n$.

Conversely, if each $f_k(x) \rightarrow q_k$, then Theorem 1 (ii) yields

$$\sum_{k=1}^{n} \bar{e}_k f_k(x) \rightarrow \sum_{k=1}^{n} \bar{e}_k q_k. \tag{4.3.16}$$

By formula (1), then, $f(x) \rightarrow \bar{q}$ (for $\sum_{k=1}^{n} \bar{e}_k q_k = \bar{q}$). Thus (ii) is proved; similarly for (i) and for relative limits and continuity.

Note 4. Again, Theorem 2 holds also for $p = \pm\infty$ (but not for infinite q).

Note 5. A complex function $f: A \rightarrow C$ may be treated as $f: A \rightarrow E^2$. Thus it has two real components: $f = (f_1, f_2)$. Traditionally, f_1 and f_2 are called the real and imaginary parts of f, also denoted by f_{re} and f_{im}, so

$$f = f_{\text{re}} + i \cdot f_{\text{im}}. \tag{4.3.17}$$

By Theorem 2, f is continuous at p iff f_{re} and f_{im} are.

Example 4.3.1

The complex exponential is the function $f: E^1 \rightarrow C$ defined by

$$f(x) = \cos x + i \cdot \sin x, \text{ also written } f(x) = e^{xi}. \tag{4.3.18}$$

As we shall see later, the sine and the cosine functions are continuous. Hence so is f by Theorem 2.

III. Next, consider functions whose domain is a set in E^n (* or C^n). We call them functions of n real ($*$ or complex) variables, treating $\bar{x} = (x_1, \ldots, x_n)$ as a variable n-tuple. The range space may be arbitrary.

In particular, a monomial in n variables is a map on E^n (* or C^n) given by a formula of the form

$$f(\bar{x}) = a x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n} = a \cdot \prod_{k=1}^{n} x_k^{m_k}, \tag{4.3.19}$$

where the m_k are fixed integers ≥ 0 and $a \in E^1$ (* or $a \in C$).[2] If $a \neq 0$, the sum $m = \sum_{k=1}^{n} m_k$ is called the degree of the monomial. Thus

$$f(x, y, z) = 3x^2 y z^3 = 3x^2 y^1 z^3 \tag{4.3.20}$$

defines a monomial of degree 6, in three real (or complex) variables x, y, z. (We often write x, y, z for x_1, x_2, x_3.)

A polynomial is any sum of a finite number of monomials; its degree is, by definition, that of its leading term, i.e., the one of highest degree. (There may be several such terms, of equal degree.) For example,

$$f(x, y, z) = 3x^2 y z^3 - 2xy^7 \tag{4.3.21}$$

defines a polynomial of degree 8 in x, y, z. Polynomials of degree 1 are sometimes called linear.

A rational function is the quotient f/g of two polynomials f and g on E^n (*orC^n). Its domain consists of those points at which g does not vanish. For example,

$$h(x, y) = \frac{x^2 - 3xy}{xy - 1} \tag{4.3.22}$$

defines a rational function on points (x, y), with $xy \neq 1$. Polynomials and monomials are rational functions with denominator 1.

Theorem 4.3.1

Any rational function (in particular, every polynomial) in one or several variables is continuous on all of its domain.

Proof

Consider first a monomial of the form

$$f(\bar{x}) = x_k \quad (k \text{ fixed}); \tag{4.3.23}$$

it is called the k th projection map because it "projects" each $\bar{x} \in E^n$ ($^*C^n$) onto its k th component x_k.

Given any $\varepsilon>0$ and $\overline{p}$, choose $\delta=\varepsilon$. Then

$$\left(\forall \overline{x} \in G_{\overline{p}}(\delta)\right) \quad |f(\overline{x})-f(\overline{p})|=\left|x_k-p_k\right| \leq \sqrt{\sum_{i=1}^n\left|x_i-p_i\right|^2}=\rho(\overline{x}, \overline{p})<\varepsilon . \tag{4.3.24}$$

Hence by definition, f is continuous at each $\overline{p}$. Thus the theorem holds for projection maps.

However, any other monomial, given by

$$f(\overline{x})=a x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}, \tag{4.3.25}$$

is the product of finitely many (namely of $m=m_1+m_2+\ldots+m_n$) projection maps multiplied by a constant a. Thus by Theorem 1, it is continuous. So also is any finite sum of monomials (i.e., any polynomial), and hence so is the quotient f/g of two polynomials (i.e., any rational function) wherever it is defined, i.e., wherever the denominator does not vanish. □

IV. For functions on E^n (* or C^n), we often consider relative limits over a line of the form

$$\overline{x}=\overline{p}+t \vec{e}_k \text{ (parallel to the } k^{th} \text{ axis, through } \overline{p}); \tag{4.3.26}$$

see Chapter 3, §§4-6, Definition 1. If f is relatively continuous at $\overline{p}$ over that line, we say that f is continuous at $\overline{p}$ in the k th variable x_k (because the other components of $\overline{x}$ remain constant, namely, equal to those of $\overline{p}$, as $\overline{x}$ runs over that line). As opposed to this, we say that f is continuous at $\overline{p}$ in all n variables jointly if it is continuous at $\overline{p}$ in the ordinary (not relative) sense. Similarly, we speak of limits in one variable, or in all of them jointly.

since ordinary continuity implies relative continuity over any path, joint continuity in all n variables always implies that in each variable separately, but the converse fails (see Problems 9 and 10 below); similarly for limits at $\overline{p}$.

4.3.E: Problems on Continuity of Vector-Valued Functions

Exercise $4.3.E.1$

Give an "ε, δ " proof of Theorem 1 for $f \pm g$.
[Hint: Proveed as in Theorem 1 of Chapter 3, §15, replacing $\max\left(k^{\prime}, k^{\prime\prime}\right)$ by $\delta=\min\left(\delta^{\prime}, \delta^{\prime\prime}\right)$. Thus fix $\varepsilon>0$ and $p \in S$. If $f(x) \rightarrow q$ and $g(x) \rightarrow r$ as $x \rightarrow p$ over B, then $\left(\exists \delta^{\prime}, \delta^{\prime\prime}>0\right)$ such that

$$\left(\forall x \in B \cap G_{\neg p}\left(\delta^{\prime}\right)\right) \quad|f(x)-q|<\frac{\varepsilon}{2} \text{ and } \left(\forall x \in B \cap G_{\neg p}\left(\delta^{\prime\prime}\right)\right) \quad|g(x)-r|<\frac{\varepsilon}{2}. \tag{4.3.E.1}$$

Put $\delta=\min\left(\delta^{\prime}, \delta^{\prime\prime}\right)$, etc.]

In Problems $2, 3,$ and 4, $E=E^{n}(*$ or another normed space $)$, F is its scalar field, $B \subseteq A \subseteq(S, \rho),$ and $x \rightarrow p$ over B.

Exercise $4.3.E.2$

For a function $f: A \rightarrow E$ prove that

$$f(x) \rightarrow q \Longleftrightarrow|f(x)-q| \rightarrow 0, \tag{4.3.E.2}$$

$$\text{equivalently, iff } f(x)-q \rightarrow \overline{0}. \quad \text{[Hint: Proceed as in Chapter 3, §14, Corollary 2.]} \tag{4.3.E.3}$$

Exercise $4.3.E.3$

Given $f: A \rightarrow\left(T, \rho^{\prime}\right)$, with $f(x) \rightarrow q$ as $x \rightarrow p$ over B. Show that for some $\delta>0$, f is bounded on $B \cap G_{\neg p}(\delta)$, i.e.,

$$f\left[B \cap G_{\neg p}(\delta)\right] \text{ is a bounded set in } \left(T, \rho^{\prime}\right). \tag{4.3.E.4}$$

Thus if $T=E,$ there is $K \in E^{1}$ such that

$$\left(\forall x \in B \cap G_{\neg p}(\delta)\right) \quad|f(x)|<K \tag{4.3.E.5}$$

(Chapter 3, §13, Theorem 2).

Exercise $4.3.E.4$

Given $f, h: A \rightarrow E^{1}(C)$ (or $f: A \rightarrow E, h: A \rightarrow F$), prove that if one of f and h has limit 0 (respectively, $\overline{0}$), while the other is bounded on $B \cap G_{\neg p}(\delta),$ then $h(x) f(x) \rightarrow 0(\overline{0})$.

Exercise $4.3.E.5$

Given $h: A \rightarrow E^{1}(C),$ with $h(x) \rightarrow a$ as $x \rightarrow p$ over B, and $a \neq 0$.
Prove that

$$(\exists \varepsilon, \delta>0)\left(\forall x \in B \cap G_{\neg p}(\delta)\right) \quad|h(x)| \geq \varepsilon, \tag{4.3.E.6}$$

i.e., $h(x)$ is bounded away from 0 on $B \cap G_{\neg p}(\delta)$. Hence show that $1 / h$ is bounded on $B \cap G_{\neg p}(\delta)$.
[Hint: Proceed as in the proof of Corollary 1 in §1, with $q=a$ and $r=0$. Then use

$$(\forall x \in B \cap G_{\neg p}(\delta)) \quad \left|\frac{1}{h(x)}\right| \leq \frac{1}{\varepsilon}.] \tag{4.3.E.7}$$

Exercise $4.3.E.6$

Using Problems 1 to 5, give an independent proof of Theorem 1.
[Hint: Proceed as in Problems 2 and 4 of Chapter 3, §15 to obtain Theorem 1(ii). Then use Corollary 2 of $1.]

Exercise $4.3.E.7$

Deduce Theorems 1 and 2 of Chapter 3, §15 from those of the present section, setting $A=B=N, S=E^{*},$ and $p=+\infty$.
[Hint: See §1, Note 5.]

Exercise $4.3.E.8$

Redo Problem 8 of §1 in two ways:
(i) Use Theorem 1 only.
(ii) Use Theorem 3.
$\left[\text{Example for (i) : Find } \lim_{x\to 1}\left(x^2+1\right)\right.$.
Here $f(x)=x^2+1$, or $f=gg+h$, where $h(x)=1$ (constant) and $g(x)=x$ (identity map). As h and g are continuous $(\S 1,$ Examples (a) and (b)), so is f by Theorem 1. Thus $\lim_{x\to 1} f(x)=f(1)=1^2+1=2$.
Or, using Theorem 1(ii) , $\lim_{x\to 1}\left(x^2+1\right)=\lim_{x\to 1} x^2+\lim_{x\to 1} 1,$ etc. $]$

Exercise $4.3.E.9$

Define $f: E^2 \rightarrow E^1$ by

$$f(x, y)=\frac{x^2 y}{\left(x^4+y^2\right)}, \text{ with } f(0,0)=0. \tag{4.3.E.8}$$

Show that $f(x, y) \rightarrow 0$ as $(x, y) \rightarrow(0,0)$ along any straight line through $\overline{0}$, but not over the parabola $y=x^2$ (then the limit is $\frac{1}{2}$). Deduce that f is continuous at $\overline{0}=(0,0)$ in x and y separately, but not jointly.

Exercise $4.3.E.10$

Do Problem 9, setting

$$f(x, y)=0 \text{ if } x=0, \text{ and } f(x, y)=\frac{|y|}{x^2} \cdot 2^{-|y|/x^2} \text{ if } x \neq 0. \tag{4.3.E.9}$$

Exercise $4.3.E.11$

Discuss the continuity of $f: E^2 \rightarrow E^1$ in x and y jointly and separately,
at $\overline{0}$, when
(a) $f(x, y)=\frac{x^2 y^2}{x^2+y^2}, f(0,0)=0$;
(b) $f(x, y)=$ integral part of $x+y$;
(c) $f(x, y)=x+\frac{x y}{|x|}$ if $x \neq 0, f(0, y)=0$;
(d) $f(x, y)=\frac{x y}{|x|}+x \sin \frac{1}{y}$ if $x y \neq 0$, and $f(x, y)=0$ otherwise;
(e) $f(x, y)=\frac{1}{x} \sin \left(x^2+|x y|\right)$ if $x \neq 0$, and $f(0, y)=0$.
[Hints: In (c) and (d), $|f(x, y)| \leq|x|+|y|;$ in (e), use $|\sin \alpha| \leq|\alpha| \cdot]$.

4.4: Infinite Limits. Operations in E*

As we have noted, Theorem 1 of §3 does not apply to infinite limits, even if the function values $f(x), g(x), h(x)$ remain finite (i.e., $in E^1$). Only in certain cases (stated below) can we prove some analogues.

There are quite a few such separate cases. Thus, for brevity, we shall adopt a kind of mathematical shorthand. The letter q will not necessarily denote a constant; it will stand for

$$\text{"a function } f: A \rightarrow E^1, A \subseteq (S, \rho), \text{ such that } f(x) \rightarrow q \in E^1 \text{ as } x \rightarrow p.\text{"} \tag{4.4.1}$$

Similarly, "0" and "$\pm\infty$" will stand for analogous expressions, with q replaced by 0 and $\pm\infty$, respectively.

For example, the "shorthand formula" $(+\infty)+(+\infty)=+\infty$ means

$$\text{"The sum of two real functions, with limit} +\infty \text{ at } p\ (p \in S), \text{ is itself a function with limit} +\infty \text{ at } p.\text{"} \tag{4.4.2}$$

The point p is fixed, possibly $\pm\infty$ (if $A \subseteq E^*$). With this notation, we have the following theorems.

Theorems

1. $(\pm\infty)+(\pm\infty)=\pm\infty$.

2. $(\pm\infty)+q=q+(\pm\infty)=\pm\infty$.

3. $(\pm\infty) \cdot(\pm\infty)=+\infty$.

4. $(\pm\infty) \cdot(\mp\infty)=-\infty$.

5. $|\pm\infty|=+\infty$.

6. $(\pm\infty) \cdot q=q \cdot(\pm\infty)=\pm\infty$ if $q>0$.

7. $(\pm\infty) \cdot q=q \cdot(\pm\infty)=\mp\infty$ if $q<0$.

8. $-(\pm\infty)=\mp\infty$.

9. $\frac{(\pm\infty)}{q}=(\pm\infty) \cdot \frac{1}{q}$ if $q \neq 0$.

10. $\frac{q}{(\pm\infty)}=0$.

11. $(+\infty)^{+\infty}=+\infty$.

12. $(+\infty)^{-\infty}=0$.

13. $(+\infty)^q=+\infty$ if $q>0$.

14. $(+\infty)^q=0$ if $q<0$.

15. If $q>1$, then $q^{+\infty}=+\infty$ and $q^{-\infty}=0$.

16. If $0<q<1$, then $q^{+\infty}=0$ and $q^{-\infty}=+\infty$.

Proof

We prove Theorems 1 and 2, leaving the rest as problems. (Theorems 11-16 are best postponed until the theory of logarithms is developed.)

1. Let $f(x)$ and $g(x) \rightarrow+\infty$ as $x \rightarrow p$. We have to show that

$$f(x)+g(x) \rightarrow+\infty, \tag{4.4.3}$$

i.e., that

$$\left(\forall b \in E^1\right)(\exists \delta>0)\left(\forall x \in A \cap G_{\neg p}(\delta)\right) \quad f(x)+g(x)>b \tag{4.4.4}$$

(we may assume $b>0$). Thus fix $b>0$. As $f(x)$ and $g(x) \rightarrow+\infty$, there are $\delta^{\prime}, \delta^{\prime \prime}>0$ such that

$$\left(\forall x \in A \cap G_{\neg p}\left(\delta^{\prime}\right)\right) f(x)>b \text{ and } \left(\forall x \in A \cap G_{\neg p}\left(\delta^{\prime \prime}\right)\right) g(x)>b. \tag{4.4.5}$$

Let $\delta=\min(\delta',\delta'')$. Then

$$(\forall x\in A\cap G_{\neg p}(\delta))\quad f(x)+g(x)>b+b>b, \tag{4.4.6}$$

as required; similarly for the case of $-\infty$.

2. Let $f(x)\rightarrow+\infty$ and $g(x)\rightarrow q\in E^1$. Then there is $\delta'>0$ such that for x in $A\cap G_{\neg p}(\delta')$, $|q-g(x)|<1$, so that $g(x)>q-1$.

Also, given any $b\in E^1$, there is δ'' such that

$$(\forall x\in A\cap G_{-p}(\delta''))\quad f(x)>b-q+1. \tag{4.4.7}$$

Let $\delta=\min(\delta',\delta'')$. Then

$$(\forall x\in A\cap G_{\neg p}(\delta))\quad f(x)+g(x)>(b-q+1)+(q-1)=b, \tag{4.4.8}$$

as required; similarly for the case of $f(x)\rightarrow-\infty$.

Caution: No theorems of this kind exist for the following cases (which therefore are called *indeterminate expressions*):

$$(+\infty)+(-\infty),\quad(\pm\infty)\cdot 0,\quad\frac{\pm\infty}{\pm\infty},\quad\frac{0}{0},\quad(\pm\infty)^0,\quad 0^0,\quad 1^{\pm\infty}. \tag{4.4.9}$$

In these cases, it does not suffice to know only the limits of f and g. It is necessary to investigate the functions themselves to give a definite answer, since in each case the answer may be different, depending on the properties of f and g. **The expressions (1*)** remain indeterminate even if we consider the simplest kind of functions, namely sequences, as we show next.

Examples

(a) Let

$$u_m=2m\text{ and }v_m=-m. \tag{4.4.10}$$

(This corresponds to $f(x)=2x$ and $g(x)=-x$.) Then, as is readily seen,

$$u_m\rightarrow+\infty, v_m\rightarrow-\infty,\text{ and }u_m+v_m=2m-m=m\rightarrow+\infty. \tag{4.4.11}$$

If, however, we take $x_m=2m$ and $y_m=-2m$, then

$$x_m+y_m=2m-2m=0; \tag{4.4.12}$$

thus x_m+y_m is constant, with limit 0 (for the limit of a constant function equals its value; see §1, Example (a)).

Next, let

$$u_m=2m\text{ and }z_m=-2m+(-1)^m. \tag{4.4.13}$$

Then again

$$u_m\rightarrow+\infty\text{ and }z_m\rightarrow-\infty,\text{ but }u_m+z_m=(-1)^m; \tag{4.4.14}$$

u_m+z_m "oscillates" from -1 to 1 as $m\rightarrow+\infty$, so it has no limit at all.

These examples show that $(+\infty)+(-\infty)$ is indeed an indeterminate expression since the answer depends on the nature of the functions involved. No general answer is possible.

(b) We now show that $1^{+\infty}$ is indeterminate.

Take first a constant $\{x_m\}$, $x_m=1$, and let $y_m=m$. Then

$$x_m\rightarrow 1, y_m\rightarrow+\infty,\text{ and }x_m^{y_m}=1^m=1=x_m\rightarrow 1. \tag{4.4.15}$$

If, however, $x_m=1+\frac{1}{m}$ and $y_m=m$, then again $y_m\rightarrow+\infty$ and $x_m\rightarrow 1$ (by Theorem 10 above and Theorem 1 of Chapter 3, §15), but

$$x_m^{y_m} = \left(1 + \frac{1}{m}\right)^m \tag{4.4.16}$$

does not tend to 1; it tends to $e > 2,$ as shown in Chapter 3, §15. Thus again the result depends on $\{x_m\}$ and $\{y_m\}$.

In a similar manner, one shows that **the other cases (1*)** are indeterminate.

Note 1. It is often useful to introduce additional "shorthand" conventions. Thus the symbol ∞ (unsigned infinity) might denote a function f such that

$$|f(x)| \rightarrow +\infty \text{ as } x \rightarrow p; \tag{4.4.17}$$

we then also write $f(x) \rightarrow \infty$. The symbol 0^+ (respectively, 0^-) denotes a function f such that

$$f(x) \rightarrow 0 \text{ as } x \rightarrow p \tag{4.4.18}$$

and, moreover

$$f(x) > 0\ (f(x) < 0, \text{ respectively}) \text{ on some } G_{\neg p}(\delta). \tag{4.4.19}$$

We then have the following additional formulas:

(i) $\frac{(\pm\infty)}{0^+} = \pm\infty, \frac{(\pm\infty)}{0^-} = \mp\infty$.

(ii) If $q > 0$, then $\frac{q}{0^+} = +\infty$ and $\frac{q}{0^-} = -\infty$.

(iii) $\frac{\infty}{0} = \infty$.

(iv) $\frac{q}{\infty} = 0$.

The proof is left to the reader.

Note 2. All these formulas and theorems hold for relative limits, too.

So far, we have defined no arithmetic operations in E^*. To fill this gap (at least partially), we shall henceforth treat Theorems 1-16 above not only as certain limit statements (in "shorthand") but also as definitions of certain operations in E^*. For example, the formula $(+\infty) + (+\infty) = +\infty$ shall be treated as the definition of the actual sum of $+\infty$ and $+\infty$ in E^*, with $+\infty$ regarded this time as an element of E^* (not as a function). This convention defines the arithmetic operations for certain cases only; the indeterminate **expressions (1*)** remain undefined, unless we decide to assign them some meaning.

In higher analysis, it indeed proves convenient to assign a meaning to at least some of them. We shall adopt these (admittedly arbitrary) conventions:

$$\left\{\begin{array}{l}(\pm\infty) + (\mp\infty) = (\pm\infty) - (\pm\infty) = +\infty; 0^0 = 1; \\ 0 \cdot (\pm\infty) = (\pm\infty) \cdot 0 = 0 \text{ (even if 0 stands for the zero-vector).}\end{array}\right.$$

Caution: These formulas must not be treated as limit theorems (in "short-hand"). Sums and products of **the form (2*)** will be called "*unorthodox*."

4.4.E: Problems on Limits and Operations in E ∗ E∗

Exercise $4.4.E.1$

Show by examples that all expressions (1^*) are indeterminate.

Exercise $4.4.E.2$

Give explicit definitions for the following "unsigned infinity" limit statements:

$$\text{(a)} \lim_{x \rightarrow p} f(x)=\infty ; \quad \text{(b)} \lim_{x \rightarrow p^{+}} f(x)=\infty ; \quad \text{(c)} \lim_{x \rightarrow \infty} f(x)=\infty . \tag{4.4.E.1}$$

Exercise $4.4.E.3$

Prove at least some of Theorems $1-10$ and formulas $(\mathrm{i})-(\mathrm{iv})$ in Note 1.

Exercise $4.4.E.4$

In the following cases, find $\lim f(x)$ in two ways: (i) use definitions only; (ii) use suitable theorems and justify each step accordingly.

$$\begin{array}{l}\text{(a) } \lim_{x \rightarrow \infty} \frac{1}{x}(=0). \quad \text{(b) } \lim_{x \rightarrow \infty} \frac{x(x-1)}{1-3 x^{2}} \\ \text{(c) } \lim_{x \rightarrow 2^{+}} \frac{x^{2}-2 x+1}{x^{2}-3 x+2} \text{ (d) } \lim_{x \rightarrow 2^{-}} \frac{x^{2}-2 x+1}{x^{2}-3 x+2} \\ \text{(e) } \lim_{x \rightarrow 2} \frac{x^{2}-2 x+1}{x^{2}-3 x+2}(=\infty)\end{array} \tag{4.4.E.2}$$

[Hint: Before using theorems, reduce by a suitable power of x.]

Exercise $4.4.E.5$

Let

$$f(x)=\sum_{k=0}^{n} a_{k} x^{k} \text{ and } g(x)=\sum_{k=0}^{m} b_{k} x^{k}\left(a_{n} \neq 0, b_{m} \neq 0\right). \tag{4.4.E.3}$$

Find $\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)}$ if (i)$n>m$; (ii)$n<m$; and (iii) $n=m(n, m \in N)$.

Exercise $4.4.E.6$

Verify commutativity and associativity of addition and multiplication in E^*, treating Theorems $1-16$ and formulas (2^*) as definitions. Show by examples that associativity and commutativity (for three terms or more) would fail if, instead of (2^*), the formula $(\pm \infty)+(\mp \infty)=0$ were adopted.
[Hint: For sums, first suppose that one of the terms in a sum is $+\infty$; then the sum is + ∞. For products, single out the case where one of the factors is 0; then consider the infinite cases.]

Exercise $4.4.E.7$

Continuing Problem 6, verify the distributive law $(x+y) z=x z+y z$ in E^*, assuming that x and y have the same sign (if infinite), or that $z \geq 0$. Show by examples that it may fail in other cases; e.g., if $x=-y=+\infty, \ z=-1$.

4.5: Monotone Function

A function $f : A \rightarrow E^{*},$ with $A \subseteq E^{*},$ is said to be nondecreasing on a set $B \subseteq A$ iff

$$x \leq y \text{ implies } f(x) \leq f(y) \text{ for } x, y \in B. \tag{4.5.1}$$

It is said to be nonincreasing on B iff

$$x \leq y \text{ implies } f(x) \geq f(y) \text{ for } x, y \in B. \tag{4.5.2}$$

Notation: $f \uparrow$ and $f \downarrow (\text{ on } B),$ respectively.

In both cases, f is said to be monotone or monotonic on B. If f is also one to one on B (i.e., when restricted to B), we say that it is strictly monotone (increasing if $f \uparrow$ and decreasing if $f \downarrow$).

Clearly, f is nondecreasing iff the function $-f=(-1) f$ is nonincreasing. Thus in proofs, we need consider only the case $f \uparrow$. The case $f \downarrow$ reduces to it by applying the result to $-f$.

Theorem 4.5.1

If a function $f : A \rightarrow E^{*}\left(A \subseteq E^{*}\right)$ is monotone on A, it has a left and a right (possibly infinite) limit at each point $p \in E^{*}$.

In particular, if $f \uparrow$ on an interval $(a, b) \neq \emptyset$, then

$$f\left(p^{-}\right)=\sup _{a<x<p} f(x) \text{ for } p \in(a, b] \tag{4.5.3}$$

and

$$f\left(p^{+}\right)=\inf _{p<x<b} f(x) \text{ for } p \in[a, b). \tag{4.5.4}$$

(In case $f \downarrow$, interchange "sup" and "inf.")

Proof

To fix ideas, assume $f \uparrow$.

Let $p \in E^{*}$ and $B=\{x \in A | x<p\}$. Put $q=\sup f[B]$ (this sup always exists in E^{*}; see Chapter 2, §13). We shall show that q is a left limit of f at p (i.e., a left limit over B).

There are three possible cases:

(1) If q is finite, any globe G_{q} is an interval $(c, d), c<q<d,$ in E^{1}. As $c<q=\sup f[B], c$ cannot be an upper bound of $f[B]$ (why?, so c is exceeded by some $f\left(x_{0}\right), x_{0} \in B.$ Thus

$$c<f\left(x_{0}\right), x_{0}<p. \tag{4.5.5}$$

Hence as $f \uparrow$, we certainly have

$$c<f\left(x_{0}\right) \leq f(x) \text{ for all } x>x_{0} \ (x \in B). \tag{4.5.6}$$

Moreover, as $f(x) \in f[B]$, we have

$$f(x) \leq \sup f[B]=q<d, \tag{4.5.7}$$

so $c<f(x)<d;$ i.e., $f(x) \in(c, d)=G_{q}$.

We have thus shown that

$$\left(\forall G_{q}\right)\left(\exists x_{0}<p\right)\left(\forall x \in B | x_{0}<x\right) \quad f(x) \in G_{q}, \tag{4.5.8}$$

so q is a left limit at p.

(2) If $q=+\infty,$ the same proof works with $G_{q}=(c,+\infty]$. Verify!

(3) If $q=-\infty,$ then

$$(\forall x \in B) \quad f(x) \leq \sup f[B]=-\infty, \tag{4.5.9}$$

i.e., $f(x) \leq -\infty$, so $f(x) = -\infty$ (constant) on B. Hence q is also a left limit at p (§1, Example (a)).

In particular, if $f \uparrow$ on $A = (a, b)$ with $a, b \in E^*$ and $a < b$, then $B = (a, p)$ for $p \in (a, b]$. Here p is a cluster point of the path B (Chapter 3, §14, Example (h)), so a unique left limit $f(p^-)$ exists. By what was shown above,

$$q = f\left(p^-\right) = \sup f[B] = \sup_{a<x<p} f(x), \text{ as claimed.} \tag{4.5.10}$$

Thus all is proved for left limits.

The proof for right limits is quite similar; one only has to set

$$B = \{x \in A | x > p\}, q = \inf f[B]. \quad \square \tag{4.5.11}$$

Note 1. The second clause of Theorem 1 holds even if (a, b) is only a subset of A, for the limits in question are not affected by restricting f to (a, b). (Why?) The endpoints a and b may be finite or infinite.

Note 2. If $D_f = A = N$ (the naturals), then by definition, $f : N \rightarrow E^*$ is a sequence with general term $x_m = f(m), m \in N$ (see §1, Note 2). Then setting $p = +\infty$ in the proof of Theorem 1, we obtain Theorem 3 of Chapter 3, §15. (Verify!)

Example 4.5.1

The exponential function $F : E^1 \rightarrow E^1$ to the base $a > 0$ is given by

$$F(x) = a^x. \tag{4.5.12}$$

It is monotone (Chapter 2, §§11-12, formula (1)), so $F(0^-)$ and $F(0^+)$ exist. By the sequential criterion (Theorem 1 of §2), we may use a suitable sequence to find $F(0^+)$, and we choose $x_m = \frac{1}{m} \rightarrow 0^+$. Then

$$F\left(0^+\right) = \lim_{m\to\infty} F\left(\frac{1}{m}\right) = \lim_{m\to\infty} a^{1/m} = 1 \tag{4.5.13}$$

(see Chapter 3, §15, Problem 20).

Similarly, taking $x_m = -\frac{1}{m} \rightarrow 0^-$, we obtain $F(0^-) = 1$. Thus

$$F\left(0^+\right) = F\left(0^-\right) = \lim_{x\to 0} F(x) = \lim_{x\to 0} a^x = 1. \tag{4.5.14}$$

(See also Problem 12 of §2.)

Next, fix any $p \in E^1$. Noting that

$$F(x) = a^x = a^{p+x-p} = a^p a^{x-p}, \tag{4.5.15}$$

we set $y = x - p$. (Why is this substitution admissible?) Then $y \rightarrow 0$ as $x \rightarrow p$, so we get

$$\lim_{x\to p} F(x) = \lim a^p \cdot \lim_{x\to p} a^{x-p} = a^p \lim_{y\to 0} a^y = a^p \cdot 1 = a^p = F(p). \tag{4.5.16}$$

As $\lim_{x\to p} F(x) = F(p), F$ is continuous at each $p \in E^1$. Thus all exponentials are continuous.

Theorem 4.5.2

If a function $f : A \rightarrow E^* \, (A \subseteq E^*)$ is nondecreasing on a finite or infinite interval $B = (a, b) \subseteq A$ and if $p \in (a, b)$, then

$$f\left(a^+\right) \leq f\left(p^-\right) \leq f(p) \leq f\left(p^+\right) \leq f\left(b^-\right), \tag{4.5.17}$$

and for no $x \in (a, b)$ do we have

$$f\left(p^-\right) < f(x) < f(p) \text{ or } f(p) < f(x) < f\left(p^+\right); \tag{4.5.18}$$

similarly in case $f \downarrow$ (with all inequalities reversed).

Proof

By Theorem 1, $f \uparrow$ on (a, p) implies

$$f\left(a^{+}\right)=\inf_{a<x<p} f(x) \text{ and } f\left(p^{-}\right)=\sup_{a<x<p} f(x); \tag{4.5.19}$$

thus certainly $f\left(a^{+}\right) \leq f\left(p^{-}\right)$. As $f \uparrow$, we also have $f(p) \geq f(x)$ for all $x \in(a, p)$; hence

$$f(p) \geq \sup_{a<x<p} f(x)=f\left(p^{-}\right). \tag{4.5.20}$$

Thus

$$f\left(a^{+}\right) \leq f\left(p^{-}\right) \leq f(p); \tag{4.5.21}$$

similarly for **the rest of (1).**

Moreover, if $a<x<p$, then $f(x) \leq f\left(p^{-}\right)$ since

$$f\left(p^{-}\right)=\sup_{a<x<p} f(x). \tag{4.5.22}$$

If, however, $p \leq x<b$, then $f(p) \leq f(x)$ since $f \uparrow$. Thus we never have $f\left(p^{-}\right)<f(x)<f(p)$. similarly, one excludes $f(p)<f(x)<f(x)<f\left(p^{+}\right)$. This completes the proof. □

Note 3. If $f\left(p^{-}\right), f\left(p^{+}\right)$, and $f(p)$ exist (all finite), then

$$\left|f(p)-f\left(p^{-}\right)\right| \text{ and } \left|f\left(p^{+}\right)-f(p)\right| \tag{4.5.23}$$

are called, respectively, the left and right jumps of f at p; their sum is the (total) jump at p. If f is monotone, the jump equals $\left|f\left(p^{+}\right)-f\left(p^{-}\right)\right|$.

For a graphical example, consider Figure 14 in §1. Here $f(p)=f\left(p^{-}\right)$ (both finite), so the left jump is 0. However, $f\left(p^{+}\right)>f(p)$, so the right jump is greater than 0. Since

$$f(p)=f\left(p^{-}\right)=\lim_{x \rightarrow p^{-}} f(x), \tag{4.5.24}$$

f is left continuous (but not right continuous) at p.

Theorem 4.5.3

If $f: A \rightarrow E^{*}$ is monotone on a finite or infinite interval (a, b) contained in A, then all its discontinuities in (a, b), if any, are "jumps," that
is, points p at which $f\left(p^{-}\right)$ and $f\left(p^{+}\right)$ exist, but $f\left(p^{-}\right) \neq f(p)$ or $f\left(p^{+}\right) \neq f(p)$.

Proof

By Theorem 1, $f\left(p^{-}\right)$ and $f\left(p^{+}\right)$ exist at each $p \in(a, b)$.

If, in addition, $f\left(p^{-}\right)=f\left(p^{+}\right)=f(p)$, then

$$\lim_{x \rightarrow p} f(x)=f(p) \tag{4.5.25}$$

by Corollary 3 of §1, so f is continuous at p. Thus discontinuities occur only if $f\left(p^{-}\right) \neq f(p)$ or $f\left(p^{+}\right) \neq f(p)$. □

4.5.E: Problems on Monotone Functions

Exercise $4.5.E.1$

Complete the proofs of Theorems 1 and 2. Give also an independent (analogous) proof for nonincreasing functions.

Exercise $4.5.E.2$

Discuss Examples (d) and (e) of §1 again using Theorems $1-3$.

Exercise $4.5.E.3$

Show that Theorem 3 holds also if f is piecewise monotone on (a, b), i.e., monotone on each of a sequence of intervals whose union is (a, b).

Exercise $4.5.E.4$

Consider the monotone function f defined in Problems 5 and 6 of Chapter 3, §11. Show that under the standard metric in E^1, f is continuous on E^1 and f^{-1} is continuous on $(0,1)$. Additionally, discuss continuity under the metric ρ'.

Exercise $4.5.E.5$

$\Rightarrow$ 5. Prove that if f is monotone on $(a, b) \subseteq E^*$, it has at most countably many discontinuities in (a, b).
[Hint: Let $f \uparrow$. By Theorem 3, all discontinuities of f correspond to mutually disjoint intervals $\left(f\left(p^{-}\right), f\left(p^{+}\right)\right) \neq \emptyset$. (Why?) Pick a rational from each such interval, so these rationals correspond one to one to the discontinuities and form a countable set (Chapter 1, §9)].

Exercise $4.5.E.6$

Continuing Problem 17 of Chapter 3, §14, let

$$G_{11}=\left(\frac{1}{3}, \frac{2}{3}\right), \quad G_{21}=\left(\frac{1}{9}, \frac{2}{9}\right), G_{22}=\left(\frac{7}{9}, \frac{8}{9}\right), \text { and so on; } \tag{4.5.E.1}$$

that is, G_{mi} is the i th open interval removed from $[0,1]$ at the m th step of the process $(i=1,2, \ldots, 2^{m-1}, m=1,2, \ldots$ ad infinitum$)$.
Define $F:[0,1] \rightarrow E^1$ as follows:
(i) $F(0)=0$;
(ii) if $x \in G_{mi}$, then $F(x)=\frac{2i-1}{2^m}$; and
(iii) if x is in none of the G_{mi} (i.e., $x \in P$), then

$$F(x)=\sup \{F(y) \mid y \in \bigcup_{m, i} G_{mi}, y<x\}. \tag{4.5.E.2}$$

Show that F is nondecreasing and continuous on $[0,1]$. (F is called Cantor's function.)

Exercise $4.5.E.7$

Restate Theorem 3 for the case where f is monotone on A, where A is a (not necessarily open) interval. How about the endpoints of A?

4.6: Compact Sets

We now pause to consider a very important kind of sets. In Chapter 3, §16, we showed that every sequence $\{\bar{z}_m\}$ taken from a closed interval $[\bar{a}, \bar{b}]$ in E^n must cluster in it (Note 1 of Chapter 3, §16). There are other sets with the same remarkable property. This leads us to the following definition.

Definition: sequentially compact

A set $A \subseteq (S, \rho)$ is said to be **sequentially compact** (briefly compact) iff every sequence $\{x_m\} \subseteq A$ clusters at some point p in A.

If all of S is compact, we say that the metric space (S, ρ) is compact.

Example 4.6.1

(a) Each closed interval in E^n is compact (see above).

(a') However, nonclosed intervals, and E^n itself, are not compact.

For example, the sequence $x_n = 1/n$ is in $(0, 1] \subset E^1$, but clusters only at 0, outside $(0, 1]$. As another example, the sequence $x_n = n$ has no cluster points in E^1. Thus $(0, 1]$ and E^1 fail to be compact (even though E^1 is complete); similarly for E^n (* and C^n).

(b) Any finite set $A \subseteq (S, \rho)$ is compact. Indeed, an infinite sequence in such a set must have at least one infinitely repeating term $p \in A$. Then by definition, this p is a cluster point (see Chapter 3, §14, Note 1).

(c) The empty set is "vacuously" compact (it contains no sequences).

(d) E^* is compact. See Example (g) in Chapter 3, §14.

Other examples can be derived from the theorems that follow.

Theorem 4.6.1

If a set $B \subseteq (S, \rho)$ is compact, so is any closed subset $A \subseteq B$.

Proof

We must show that each sequence $\{x_m\} \subseteq A$ clusters at some $p \in A$. However, as $A \subseteq B, \{x_m\}$ is also in B, so by the compactness of B, it clusters at some $p \in B$. Thus it remains to show that $p \in A$ as well.

Now by Theorem 1 of Chapter 3, §16, $\{x_m\}$has a subsequence $x_{m_k} \rightarrow p$. As $\{x_{m_k}\} \subseteq A$ and A is closed, this implies $p \in A$ (Theorem 4 in Chapter 3, §16). □

Theorem 4.6.2

Every compact set $A \subseteq (S, \rho)$ is closed.

Proof

Given that A is compact, we must show (by Theorem 4 in Chapter 3, §16) that A contains the limit of each convergent sequence $\{x_m\} \subseteq A$.

Thus let $x_m \rightarrow p, \{x_m\} \subseteq A$. As A is compact, the sequence $\{x_m\}$ clusters at some $q \in A$, i.e., has a subsequence $x_{m_k} \rightarrow q \in A$. However, the limit of the subsequence must be the same as that of the entire sequence. Thus $p = q \in A$; i.e., p is in A, as required. □

Theorem 4.6.3

Every compact set $A \subseteq (S, \rho)$ is bounded.

Proof

By Problem 3 in Chapter 3, §13, it suffices to show that A is contained in some finite union of globes. Thus we fix some arbitrary radius $\varepsilon > 0$ and, seeking a contradiction, assume that A cannot by any finite number of globes of that radius.

Then if $x_1 \in A$, the globe $G_{x_1}(\varepsilon)$ does not cover A, so there is a point $x_2 \in A$ such that

$$x_2 \notin G_{x_1}(\varepsilon), \text{ i.e., } \rho(x_1, x_2) \geq \varepsilon \tag{4.6.1}$$

By our assumption, A is not even covered by $G_{x_1}(\varepsilon) \cup G_{x_2}(\varepsilon)$. Thus there is a point $x_3 \in A$ with

$$x_3 \notin G_{x_1}(\varepsilon) \text{ and } x_3 \notin G_{x_2}(\varepsilon), \text{ i.e., } \rho(x_3, x_1) \geq \varepsilon \text{ and } \rho(x_3, x_2) \geq \varepsilon. \tag{4.6.2}$$

Again, A is not covered by $\bigcup_{i=1}^{3} G_{x_i}(\varepsilon)$, so there is a point $x_4 \in A$ not in that union; its distances from x_1, x_2, and x_3 must therefore be $\geq \varepsilon$.

Since A is never covered by any finite number of ε -globes, we can continue this process indefinitely (by induction) and thus select an infinite sequence $\{x_m\} \subseteq A$, with all its terms at least ε -apart from each other.

Now as A is compact, this sequence must have a convergent subsequence $\{x_{m_k}\}$, which is then certainly Cauchy (by Theorem 1 of Chapter 3, §17). This is impossible, however, since its terms are at distances $\geq \varepsilon$ from each other, contrary to Definition 1 in Chapter 3, §17. This contradiction completes the proof. □

Note 1. We have actually proved more than was required, namely, that no matter how small $\varepsilon > 0$ is, A can be covered by finitely many globes of radius ε with centers in A. This property is called total boundedness (Chapter 3, §13, Problem 4).

Note 2. Thus all compact sets are closed and bounded. The converse fails in metric spaces in general (see Problem 2 below). In E^n (* and C^n), however, the converse is likewise true, as we show next.

Theorem 4.6.4

In E^n (* and C^n) a set is compact iff it is closed and bounded.

Proof

In fact, if a set $A \subseteq E^n$ ($^*C^n$) is bounded, then by the Bolzano-Weierstrass theorem, each sequence $\{x_m\} \subseteq A$ has a convergent subsequence $x_{m_k} \rightarrow p$. If A is also closed, the limit point p must belong to A itself.

Thus each sequence $\{x_m\} \subseteq A$ clusters at some p in A, so A is compact.

The converse is obvious. □

Note 3. In particular, every closed globe in E^n (* or C^n) is compact since it is bounded and closed (Chapter 3, §12, Example (6)), so theorem 4 applies.

The converse is obvious. □

Theorem 4.6.5

(Cantor's principle of nested closed sets). Every contracting sequence of nonvoid compact sets

$$F_1 \supseteq F_2 \supseteq \cdots \supseteq F_m \supseteq \cdots \tag{4.6.3}$$

in a metric space (S, ρ) has a nonvoid intersection; i.e., some p belongs to all F_m.

For complete sets F_m, this holds as well, provided the diameters of the sets F_m tend to $0 : dF_m \rightarrow 0$.

Proof

We prove the theorem for complete sets first.

As $F_m \neq \emptyset$, we can pick a point x_m from each F_m to obtain a sequence $\{x_m\}, x_m \in F_m$. since $dF_m \rightarrow 0$, it is easy to see that $\{x_m\}$ is a Cauchy sequence. (The details are left to the reader.) Moreover,

$$(\forall m) \quad x_m \in F_m \subseteq F_1. \tag{4.6.4}$$

Thus $\{x_m\}$ is a Cauchy sequence in F_1, a complete set (by assumption).

Therefore, by the definition of completeness (Chapter 3, §17), $\{x_m\}$ has a limit $p \in F_1$. This limit remains the same if we drop a finite number of terms, say, the first $m-1$ of them. Then we are left with the sequence $x_m, x_{m+1}, \ldots,$ which, by construction, is entirely contained in F_m (why?), with the same limit P. Then, however, the completeness of F_m implies that $p \in F_m$ as well. As m is arbitrary here, it follows that $(\forall m) p \in F_m,$ i.e.,

$$p \in \bigcap_{m=1}^{\infty} F_m, \text{ as claimed.} \tag{4.6.5}$$

The proof for compact sets is analogous and even simpler. Here $\{x_m\}$ need not be a Cauchy sequence. Instead, using the compactness of F_1, we select from $\{x_m\}$ a subsequence $x_{m_k} \rightarrow p \in F_1$ and then proceed as above. □

Note 4. In particular, in E^n we may let the sets F_m be closed intervals (since they are compact). Then Theorem 5 yields the principle of nested intervals: Every contracting sequence of closed intervals in E^n has a nonempty intersection. (For an independent proof, see Problem 8 below.)

4.6.E: Problems on Compact Sets

Exercise $4.6.E.1$

Complete the missing details in the proof of Theorem 5.

Exercise $4.6.E.2$

Verify that any infinite set in a discrete space is closed and bounded but not compact.
[Hint: In such a space no sequence of distinct terms clusters.]

Exercise $4.6.E.3$

Show that E^n is not compact, in three ways:
(i) from definitions (as in Example $(a'))$;
(ii) from Theorem 4; and
(iii) from Theorem 5, by finding in E^n a contracting sequence of infinite closed sets with a void intersection. For example, in E^1 take the closed sets $F_m = [m, +\infty), m = 1, 2, \ldots$ (Are they closed?)

Exercise $4.6.E.4$

Show that E^* is compact under the metric ρ' defined in Problems 5 and 6 in Chapter 3, §11. Is E^1 a compact set under that metric?
[Hint: For the first part, use Theorem 2 of Chapter 2, §13, noting that G_q is also a globe under ρ'. For the second, consider the sequence $x_n = n$.]

Exercise $4.6.E.5$

Show that a set $A \subseteq (S, \rho)$ is compact iff every infinite subset $B \subseteq A$ has a cluster point $p \in A$.
[Hint: Select from B a sequence $\{x_m\}$ of distinct terms. Then the cluster points of $\{x_m\}$ are also those of B. (Why?)]

Exercise $4.6.E.6$

Prove the following.
(i) If A and B are compact, so is $A \cup B$, and similarly for unions of n sets.
(ii) If the sets $A_i (i \in I)$ are compact, so is $\bigcap_{i \in I} A_i$, even if I is infinite.
Disprove (i) for unions of infinitely many sets by a counterexample.
[Hint: For (ii), verify first that $\bigcap_{i \in I} A_i$ is sequentially closed. Then use Theorem 1.]

Exercise $4.6.E.7$

Prove that if $x_m \to p$ in (S, ρ), then the set

$$B = \{p, x_1, x_2, \ldots, x_m, \ldots\} \tag{4.6.E.1}$$

is compact.
[Hint: If B is finite, see Example (b). If not, use Problem 5, noting that any infinite subset of B defines a subsequence $x_{m_k} \to p$, so it clusters at p.]

Exercise $4.6.E.8$

Prove, independently, the principle of nested intervals in E^n, i.e., Theorem 5 with

$$F_m=\left[\bar{a}_m, \bar{b}_m\right] \subseteq E^n, \tag{4.6.E.2}$$

where

$$\bar{a}_m=\left(a_{m1}, \ldots, a_{mn}\right) \text{ and } \bar{b}_m=\left(b_{m1}, \ldots, b_{mn}\right). \tag{4.6.E.3}$$

Fixing k, let A_k be the set of all $a_{mk}, m=1,2, \ldots$. Show that A_k is bounded above by each b_{mk}, so let $p_k=\sup A_k$ in E^1. Then

$$(\forall m) \quad a_{mk} \leq p_k \leq b_{mk}. \text{ (Why?)} \tag{4.6.E.4}$$

Unfixing k, obtain such inequalities for $k=1,2, \ldots, n$. Let $\bar{p}=\left(p_1, \ldots, p_k\right)$. Then

$$(\forall m) \quad \bar{p} \in\left[\bar{a}_m, \bar{b}_m\right], \text{ i.e., } \bar{p} \in \bigcap F_m, \text{ as required.} \tag{4.6.E.5}$$

Note that the theorem fails for nonclosed intervals, even in E^1; e.g., take $F_m=(0,1/m]$ and show that $\bigcap_m F_m=\emptyset$.]

Exercise $4.6.E.9$

From Problem 8, obtain a new proof of the Bolzano-Weierstrass theorem.
[Hint: Let $\{\bar{x}_m\} \in[\bar{a}, \bar{b}] \subseteq E^n$; put $F_0=[\bar{a}, \bar{b}]$ and set

$$dF_0=\rho(\bar{a}, \bar{b})=d \quad (\text{ diagonal of } F_0). \tag{4.6.E.6}$$

Bisecting the edges of F_0, subdivide F_0 into 2^n intervals of diagonal $d/2$; one of them must contain infinitely many x_m. (Why?) Let F_1 be one such inter val; make it closed and subdivide it into 2^n subintervals of diagonal $d/2^2$. One of them, F_2, contains infinitely many x_m; make it closed, etc.
Thus obtain a contracting sequence of closed intervals F_m with

$$dF_m=\frac{d}{2^m}, \quad m=1,2, \ldots \tag{4.6.E.7}$$

From Problem 8, obtain

$$\bar{p} \in \bigcap_{m=1}^{\infty} F_m. \tag{4.6.E.8}$$

Show that $\{\bar{x}_m\}$ clusters at $\bar{p}$.]

Exercise $4.6.E.10$

$\Rightarrow$ 10. Prove the Heine-Borel theorem: If a closed interval $F_0 \subset E^n$ is covered by a family of open sets $G_i (i \in I)$, i.e.,

$$F_0 \subseteq \bigcup_{i \in I} G_i, \tag{4.6.E.9}$$

then it can always be covered by a finite number of these G_i.

[Outline of proof: Let $dF_0 = d$. Seeking a contradiction, suppose F_0 cannot be covered by any finite number of the G_i.
As in Problem 9, subdivide F_0 into 2^n intervals of diagonal $d/2$. At least one of them cannot be covered by finitely many G_i. (Why?) Choose one such interval, make it closed, call it F_1, and subdivide it into 2^n subintervals of diagonal $d/2^2$. One of these, F_2, cannot be covered by finitely many G_i; make it closed and repeat the process indefinitely.
Thus obtain a contracting sequence of closed intervals F_m with

$$dF_m = \frac{d}{2^m}, \quad m = 1, 2, \ldots \tag{4.6.E.10}$$

From Problem 8 (or Theorem 5), get $\overline{p} \in \bigcap F_m$.
As $\overline{p} \in F_0, \overline{p}$ is in one of the G_i; call it G. As G is open, $\overline{p}$ is its interior point, so let $G \supseteq G_{\overline{p}}(\varepsilon)$. Now take m so large that $d/2^m = dF_m < \varepsilon$. Show that then

$$F_m \subseteq G_{\overline{p}}(\varepsilon) \subseteq G. \tag{4.6.E.11}$$

Thus (contrary to our choice of the F_m) F_m is covered by a single set G_i. This contradiction completes the proof.]

Exercise 4.6.E.11

Prove that if $\{x_m\} \subseteq A \subseteq (S, \rho)$ and A is compact, then $\{x_m\}$ converges iff it has a single cluster point.
[Hint: Proceed as in Problem 12 of Chapter 3, §16.]

Exercise 4.6.E.12

Prove that if $\emptyset \neq A \subseteq (S, \rho)$ and A is compact, there are two points $p, q \in A$ such that $dA = \rho(p, q)$.
[Hint: As A is bounded (Theorem 3), $dA < +\infty$. By the properties of suprema,

$$(\forall n)\left(\exists x_n, y_n \in A\right) \quad dA - \frac{1}{n} < \rho\left(x_n, y_n\right) \leq dA. \quad \text{(Explain!)} \tag{4.6.E.12}$$

By compactness, $\{x_n\}$ has a subsequence $x_{n_k} \rightarrow p \in A$. For brevity, put $x_k' = x_{n_k}$, $y_k' = y_{n_k}$. Again, $\{y_k'\}$ has a subsequence $y_{k_m}' \rightarrow q \in A$. Also,

$$dA - \frac{1}{n_{k_m}} < \rho\left(x_{k_m}', y_{k_m}'\right) \leq dA. \tag{4.6.E.13}$$

Passing to the limit (as $m \rightarrow +\infty$), obtain

$$dA \leq \rho(p, q) \leq dA \tag{4.6.E.14}$$

by Theorem 4 in Chapter 3, §15.]

Exercise 4.6.E.13

Given nonvoid sets $A, B \subseteq (S, \rho)$, define

$$\rho(A, B) = \inf\{\rho(x, y) | x \in A, y \in B\}. \tag{4.6.E.15}$$

Prove that if A and B are compact and nonempty, there are $p \in A$ and $q \in B$ such that $\rho(p, q) = \rho(A, B)$. Give an example to show that this may fail if A and B are not compact (even if they are closed in E^1) .
[Hint: For the first part, proceed as in Problem 12 .]

Exercise $4.6.E.14$

Prove that every compact set is complete. Disprove the converse by examples.

4.7: More on Compactness

This page is a draft and is under active development.

Another useful approach to compactness is based on the notion of a covering of a set (already encountered in Problem 10 in §6). We say that a set F is covered by a family of sets $G_i (i \in I)$ iff

$$F \subseteq \bigcup_{i \in I} G_i. \tag{4.7.1}$$

If this is the case, $\{G_i\}$ is called a covering of F. If the sets G_i are open, we call the set family $\{G_i\}$ an open covering. The covering $\{G_i\}$ is said to be finite (infinite, countable, etc.) iff the number of the sets G_i is.

If $\{G_i\}$ is an open covering of F, then each point $x \in F$ is in some G_i and is its interior point $(\text{ for } G_i \text{ is open }),$ so there is a globe $G_x\left(\varepsilon_x\right) \subseteq G_i.$ In general, the radii ε_x of these globes depend on x, i.e., are different for different points $x \in F$. If, however, they can be chosen all equal to some ε, then this ε is called a Lebesgue number for the covering $\{G_i\}$ (so named after Henri Lebesgue). Thus ε is a Lebesgue number iff for every $x \in F$, the globe $G_x(\varepsilon)$ is contained in some G_i. We now obtain the following theorem.

Theorem 4.7.1

(Lebesgue). Every open covering $\{G_j\}$ of a sequentially compact set $F \subseteq (S, \rho)$ has at least one Lebesgue number ε. In symbols,

$$(\exists \varepsilon>0)(\forall x \in F)(\exists i) \quad G_x(\varepsilon) \subseteq G_i. \tag{4.7.2}$$

Proof

Seeking a contradiction, assume that (1) fails, i.e., its negation holds. As was explained in Chapter 1, §§1-3, this negation is

$$(\forall \varepsilon>0)\left(\exists x_{\varepsilon} \in F\right)(\forall i) \quad G_{x_{\varepsilon}}(\varepsilon) \nsubseteq G_i \tag{4.7.3}$$

(where we write x_{ε} for x since here x may depend on ε). As this is supposed to hold for all $\varepsilon>0,$ we take successively

$$\varepsilon=1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots \tag{4.7.4}$$

Then, replacing " x_{ε} " by " x_n " for convenience, we obtain

$$(\forall n)\left(\exists x_n \in F\right)(\forall i) \quad G_{x_n}\left(\frac{1}{n}\right) \nsubseteq G_i. \tag{4.7.5}$$

Thus for each n, there is some $x_n \in F$ such that the globe $G_{x_n}\left(\frac{1}{n}\right)$ is not contained in any G_i. We fix such an $x_n \in F$ for each n, thus obtaining a sequence $\{x_n\} \subseteq F$. As F is compact (by assumption), this sequence clusters at some $p \in F$.

The point p, being in F, must be in some $G_i(\text{ call it } G),$ together with some globe $G_p(r) \subseteq G.$ As p is a cluster point, even the smaller globe $G_p\left(\frac{r}{2}\right)$ contains infinitely many x_n. Thus we may choose n so large that $\frac{1}{n}<\frac{r}{2}$ and $x_n \in G_p\left(\frac{r}{2}\right)$. For that $n, G_{x_n}\left(\frac{1}{n}\right) \subseteq G_p(r)$ because

$$\left(\forall x \in G_{x_n}\left(\frac{1}{n}\right)\right) \quad \rho(x, p) \leq \rho\left(x, x_n\right)+\rho\left(x_n, p\right)<\frac{1}{n}+\frac{r}{2}<\frac{r}{2}+\frac{r}{2}=r. \tag{4.7.6}$$

As $G_p(r) \subseteq G$ (by construction), we certainly have

$$G_{x_n}\left(\frac{1}{n}\right) \subseteq G_p(r) \subseteq G. \tag{4.7.7}$$

However, this is impossible since by (2) no $G_{x_n}\left(\frac{1}{n}\right)$ is contained in any G_i. This contradiction completes the proof. $\square$

Our next theorem might serve as an alternative definition of compactness. In fact, in topology (which studies more general than metric spaces), this is is the basic definition of compactness. It generalizes Problem 10 in §6.

Theorem 4.7.2

(generalized Heine-Borel theorem). A set $F \subseteq (S, \rho)$ is compact iff every open covering of F has a finite subcovering.

That is, whenever F is covered by a family of open sets $G_i (i \in I), F$ can also be covered by a finite number of these G_i.

Proof

Let F be sequentially compact, and let $F \subseteq \cup G_i,$ all G_i open. We have to show that $\{G_i\}$ reduces to a finite subcovering.

By Theorem $1, \{G_i\}$ has a Lebesgue number ε satisfying (1). We fix this $\varepsilon > 0.$ Now by Note 1 in §6, we can cover F by a finite number of ε -globes,

$$F \subseteq \bigcup_{k=1}^{n} G_{x_k}(\varepsilon), \quad x_k \in F. \tag{4.7.8}$$

Also by (1), each $G_{x_k}(\varepsilon)$ is contained in some G_i; call it G_{i_k}. With the G_{i_k} so fixed, we have

$$F \subseteq \bigcup_{k=1}^{n} G_{x_k}(\varepsilon) \subseteq \bigcup_{k=1}^{n} G_{i_k}. \tag{4.7.9}$$

Thus the sets G_{i_k} constitute the desired finite subcovering, and the "only if' in the theorem is proved.

Conversely, assume the condition stated in the theorem. We have to show that F is sequentially compact, i.e., that every sequence $\{x_m\} \subseteq F$ clusters at some $p \in F$.

Seeking a contradiction, suppose F contains no cluster points of $\{x_m\}$. Then by definition, each point $x \in F$ is in some globe G_x containing at most finitely many x_m. The set F is covered by these open globes, hence also by finitely many of them (by our assumption). Then, however, F contains at most finitely many x_m (namely, those contained in the so-selected globes), whereas the sequence $\{x_m\} \subseteq F$ was assumed infinite. This contradiction completes the proof. □

4.8: Continuity on Compact Sets. Uniform Continuity

This page is a draft and is under active development.

I. Some additional important theorems apply to functions that are continuous on a compact set (see §6).

Theorem 4.8.1

If a function $f : A \rightarrow (T, \rho'), A \subseteq (S, \rho),$ is relatively continuous on a compact set $B \subseteq A,$ then $f[B]$ is a compact set in (T, ρ'). Briefly,

$$\text{the continuous image of a compact set is compact.} \tag{4.8.1}$$

Proof

To show that $f[B]$ is compact, we take any sequence $\{y_m\} \subseteq f[B]$ and prove that it clusters at some $q \in f[B]$.

As $y_m \in f[B], y_m = f(x_m)$ for some x_m in B. We pick such an $x_m \in B$ for each y_m, thus obtaining a sequence $\{x_m\} \subseteq B$ with

$$f(x_m) = y_m, \quad m = 1, 2, \ldots \tag{4.8.2}$$

Now by the assumed compactness of B, the sequence $\{x_m\}$ must cluster at some $p \in B$. Thus it has a subsequence $x_{m_k} \rightarrow p$. As $p \in B$, the function f is relatively continuous at p over B (by assumption). Hence by the sequential criterion (§2), $x_{m_k} \rightarrow p$ implies $f(x_{m_k}) \rightarrow f(p)$; i.e.,

$$y_{m_k} \rightarrow f(p) \in f[B]. \tag{4.8.3}$$

Thus $q = f(p)$ is the desired cluster point of $\{y_m\}$. $\square$

This theorem can be used to prove the compactness of various sets.

Example 4.8.1

(1) A closed line segment $L[\overline{a}, \overline{b}]$ in E^n (* and in other normed spaces) is compact, for, by definition,

$$L[\overline{a}, \overline{b}] = \{\overline{a} + t\vec{u} | 0 \leq t \leq 1\}, \text{ where } \vec{u} = \overline{b} - \overline{a}. \tag{4.8.4}$$

Thus $L[\overline{a}, \overline{b}]$ is the image of the compact interval $[0, 1] \subseteq E^1$ under the map $f : E^1 \rightarrow E^n$, given by $f(t) = \overline{a} + t\vec{u}$, which is continuous by Theorem 3 of §3. (Why?)

(2) The closed solid ellipsoid in E^3,

$$\left\{(x, y, z) | \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1\right\}, \tag{4.8.5}$$

is compact, being the image of a compact globe under a suitable continuous map. The details are left to the reader as an exercise.

lemma 4.8.1

Every nonvoid compact set $F \subseteq E^1$ has a maximum and a minimum.

Proof

By Theorems 2 and 3 of §6, F is closed and bounded. Thus F has an infimum and a supremum in E^1 (by the completeness axiom), say, $p = \inf F$ and $q = \sup F$. It remains to show that $p, q \in F$.

Assume the opposite, say, $q \notin F$. Then by properties of suprema, each globe $G_q(\delta) = (q-\delta, q+\delta)$ contains some $x \in B$ (specifically, $q-\delta < x < q$) other than q(for $q \notin B$, while $x \in B$). Thus

$$(\forall \delta > 0) \quad F \cap G_{\neg q}(\delta) \neq \emptyset; \tag{4.8.6}$$

i.e., F clusters at q and hence must contain q (being closed). However, since $q \notin F$, this is the desired contradiction, and the lemma is proved. □

The next theorem has many important applications in analysis.

Theorem 4.8.2

(Weierstrass).

(i) If a function $f : A \rightarrow (T, \rho')$ is relatively continuous on a compact set $B \subseteq A$, then f is bounded on B; i.e., $f[B]$ is bounded.

(ii) If, in addition, $B \neq \emptyset$ and f is real $\left(f : A \rightarrow E^1\right)$, then $f[B]$ has a maximum and a minimum; i.e., f attains a largest and a least value at some points of B.

Proof

Indeed, by Theorem 1, $f[B]$is compact, so it is bounded, as claimed in (i).

If further $B \neq \emptyset$ and f is real, then $f[B]$ is a nonvoid compact set in E^1, so by Lemma 1, it has a maximum and a minimum in E^1. Thus all is proved. □

Note 1. This and the other theorems of this section hold, in particular, if B is a closed interval in E^n or a closed globe in E^n (* or C^n) (because these sets are compact - see the examples in §6). This may fail, however, if B is not compact, e.g., if $B = (\bar{a}, \bar{b})$. For a counterexample, see Problem 11 in Chapter 3, §13.

Theorem 4.8.3

If a function $f : A \rightarrow (T, \rho')$, $A \subseteq (S, \rho)$, is relatively continuous on a compact set $B \subseteq A$ and is one to one on B (i.e., when restricted to B), then its inverse, f^{-1}, is continuout on $f[B]$.

Proof

To show that f^{-1} is continuous at each point $q \in f[B]$, we apply the sequential criterion (Theorem 1 in §2). Thus we fix a sequence $\{y_m\} \subseteq f[B], y_m \rightarrow q \in f[B]$, and prove that $f^{-1}(y_m) \rightarrow f^{-1}(q)$.

Let $f^{-1}(y_m) = x_m$ and $f^{-1}(q) = p$ so that

$$y_m = f(x_m), q = f(p), \text{ and } x_m, p \in B. \tag{4.8.7}$$

We have to show that $x_m \rightarrow p$, i.e., that

$$(\forall \varepsilon > 0)(\exists k)(\forall m > k) \quad \rho(x_m, p) < \varepsilon. \tag{4.8.8}$$

Seeking a contradiction, suppose this fails, i.e., its negation holds. Then (see Chapter 1, §§1–3) there is an $\epsilon > 0$ such that

$$(\forall k)(\exists m_k > k) \quad \rho(x_{m_k}, p) \geq \varepsilon, \tag{4.8.9}$$

where we write "m_k" for "m" to stress that the m_k may be different for different k. Thus by (1), we fix some m_k for each k so that (1) holds, choosing step by step,

$$m_{k+1} > m_k, \quad k = 1, 2, \ldots \tag{4.8.10}$$

Then the x_{m_k} form a subsequence of $\{x_m\}$, and the corresponding $y_{m_k} = f(x_{m_k})$ form a subsequence of $\{y_m\}$. Henceforth, for brevity, let $\{x_m\}$ and $\{y_m\}$ themselves denote these two subsequences. Then as before, $x_m \in B, y_m = f(x_m) \in f[B]$, and $y_m \rightarrow q, q = f(p)$. Also,by(1),

$$(\forall m) \quad \rho(x_m, p) \geq \varepsilon \,(x_m \text{ stands for } x_{m_k}). \tag{4.8.11}$$

Now as $\{x_m\} \subseteq B$ and B is compact, $\{x_m\}$ has a (sub)subsequence

$$x_{m_i} \rightarrow p' \text{ for some } p' \in B. \tag{4.8.12}$$

As f is relatively continuous on B, this implies

$$f(x_{m_i}) = y_{m_i} \rightarrow f(p') \tag{4.8.13}$$

However, the subsequence $\{y_{m_i}\}$ must have the same limit as $\{y_m\}$, i.e., $f(p)$. Thus $f(p') = f(p)$ whence $p = p'$ (for f is one to one on B), so $x_{m_i} \rightarrow p' = p$.

This contradicts (2), however, and thus the proof is complete. □

Example 4.8.2

(3) For a fixed $n \in N$, define $f : [0, +\infty) \rightarrow E^1$ by

$$f(x) = x^n. \tag{4.8.14}$$

Then f is one to one (strictly increasing) and continuous (being a monomial; see §3). Thus by Theorem 3, f^{-1} (the nth root function) is relatively continuous on each interval

$$f = [a^n, b^n]. \tag{4.8.15}$$

hence on $[0, +\infty)$.

See also Example (a) in §6 and Problem 1 below.

II. Uniform Continuity. If f is relatively continuous on B, then by definition,

$$(\forall \varepsilon > 0)(\forall p \in B)(\exists \delta > 0)\,(\forall x \in B \cap G_p(\delta)) \quad \rho'(f(x), f(p)) < \varepsilon. \tag{4.8.16}$$

Here, in general, δ depends on both ϵ and p (see Problem 4 in §1); that is, given $\epsilon > 0$, some values of δ may fit a given p but fail (3) for other points.

It may occur, however, that one and the same δ (depending on ϵ only) satisfies (3) for all $p \in B$ simultaneously, so that we have the stronger formula

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall p, x \in B | \rho(x, p) < \delta) \quad \rho'(f(x), f(p)) < \varepsilon. \tag{4.8.17}$$

Definition

If (4) is true, we say that f is uniformly continuous on B.

Clearly, this implies (3), but the converse fails.

Theorem 4.8.4

If a function $f : A \rightarrow (T, \rho')$, $A \subseteq (S, \rho)$, is relatively continuous on a compact set $B \subset A$, then f is also uniformly continuous on B.

Proof

(by contradiction). Suppose f is relatively continuous on B, but (4) fails. Then there is an $\epsilon > 0$ such that

$$(\forall \delta > 0)(\exists p, x \in B) \quad \rho(x, p) < \delta, \text{ and yet } \rho'(f(x), f(p)) \geq \varepsilon; \tag{4.8.18}$$

here p and x on δ. We fix such an ϵ and let

$$\delta = 1, \frac{1}{2}, \ldots, \frac{1}{m}, \ldots \tag{4.8.19}$$

Then for each δ (i.e., each m), we get two points $x_m, p_m \in B$ with

$$\rho\left(x_m, p_m\right)<\frac{1}{m} \tag{4.8.20}$$

and

$$\rho^{\prime}\left(f\left(x_m\right), f\left(p_m\right)\right) \geq \varepsilon, \quad m=1,2, \ldots \tag{4.8.21}$$

Thus we obtain two sequences, $\{x_m\}$ and $\{p_m\}$, in B. As B is compact, $\{x_m\}$ has a subsequence $x_{m_k} \rightarrow q(q \in B)$. For simplicity, let it be $\{x_m\}$ itself; thus

$$x_m \rightarrow q, \quad q \in B. \tag{4.8.22}$$

Hence by (5), it easily follows that also $p_m \rightarrow q$ (because $\rho\left(x_m, p_m\right) \rightarrow 0$. By the assumed relative continuity of f on B, it follows that

$$f\left(x_m\right) \rightarrow f(q) \text{ and } f\left(p_m\right) \rightarrow f(q) \text{ in } \left(T, \rho^{\prime}\right). \tag{4.8.23}$$

This, in turn, implies that $\rho^{\prime}\left(f\left(x_m\right), f\left(p_m\right)\right) \rightarrow 0$, which is impossible, in view of (6). This contradiction completes the proof. □

Example 4.8.1

(a) A function $f: A \rightarrow\left(T, \rho^{\prime}\right), A \subseteq(S, \rho)$, ic called a contraction map (on A) iff

$$\rho(x, y) \geq \rho^{\prime}(f(x), f(y)) \text{ for all } x, y \in A. \tag{4.8.24}$$

Any such map is uniformly continuous on A. In fact, given $\varepsilon>0$, we simply take $\delta=\varepsilon$. Then $\forall x, p \in A$

$$\rho(x, p)<\delta \text{ implies } \rho^{\prime}(f(x), f(p)) \leq \rho(x, p)<\delta=\varepsilon, \tag{4.8.25}$$

as required in (3).

(b) As a special case, consider the absolute value map (norm map) given by

$$f(\overline{x})=|\overline{x}| \text{ on } E^n\left({ }^* \text{ or another normed space }\right). \tag{4.8.26}$$

It is uniformly continuous onE^n because

$$||\overline{x}|-|\overline{p}|| \leq|\overline{x}-\overline{p}|, \text{ i.e., } \rho^{\prime}(f(\overline{x}), f(\overline{p})) \leq \rho(\overline{x}, \overline{p}), \tag{4.8.27}$$

which shows that f is a contraction map, so Example (a) applies.

(c) Other examples of contraction maps are

(1) constant maps (see §1, Example (a)) and

(2) projection maps (see the proof of Theorem 3 in §3).

Verify!

(d) Define $f: E^1 \rightarrow E^1$ by

$$f(x)=\sin x \tag{4.8.28}$$

By elementary trigonometry, $|\sin x| \leq|x|$. Thus $\left(\forall x, p \in E^1\right)$

$$\begin{aligned} |f(x)-f(p)| & =|\sin x-\sin p| \\ & =2\left|\sin \frac{1}{2}(x-p) \cdot \cos \frac{1}{2}(x+p)\right| \\ & \leq 2\left|\sin \frac{1}{2}(x-p)\right| \\ & \leq 2 \cdot \frac{1}{2}|x-p|=|x-p| \end{aligned},$$

and f is a contraction map again. Hence the sine function is uniformly continuous on E^1; similarly for the cosine function.

(e) Given $\emptyset \neq A \subseteq(S, \rho)$, define $f: S \rightarrow E^1$ by

$$f(x)=\rho(x, A) \text{ where } \rho(x, A)=\inf_{y \in A} \rho(x, y) \tag{4.8.29}$$

It is easy to show that

$$(\forall x, p \in S) \quad \rho(x, A) \leq \rho(x, p)+\rho(p, A) \tag{4.8.30}$$

i.e.,

$$f(x) \leq \rho(p, x)+f(p), \text{ or } f(x)-f(p) \leq \rho(p, x) \tag{4.8.31}$$

Similarly, $f(p)-f(x) \leq \rho(p, x)$. Thus

$$|f(x)-f(p)| \leq \rho(p, x) \tag{4.8.32}$$

i.e., f is uniformly continuous (being a contraction map).

(f) The identity map $f:(S, \rho) \rightarrow(S, \rho)$, given by

$$f(x)=x \tag{4.8.33}$$

is uniformly continuous on S since

$$\rho(f(x), f(p))=\rho(x, p) \text{ (a contraction map!)} \tag{4.8.34}$$

However, even relative continuity could fail if the metric in the domain space S were not the same as in S when regarded as the range space (e.g., make ρ' discrete!)

(g) Define $f: E^{1} \rightarrow E^{1}$ by

$$f(x)=a+b x \quad(b \neq 0). \tag{4.8.35}$$

Then

$$\left(\forall x, p \in E^{1}\right) \quad|f(x)-f(p)|=|b||x-p|; \tag{4.8.36}$$

i.e.,

$$\rho(f(x), f(p))=|b| \rho(x, p). \tag{4.8.37}$$

Thus, given $\varepsilon>0,$ take $\delta=\varepsilon/|b|$. Then

$$\rho(x, p)<\delta \Longrightarrow \rho(f(x), f(p))=|b| \rho(x, p)<|b| \delta=\varepsilon, \tag{4.8.38}$$

proving uniform continuity.

(h) Let

$$f(x)=\frac{1}{x} \quad \text{ on } B=(0,+\infty). \tag{4.8.39}$$

Then f is continuous on B, but not uniformly so. Indeed, we can prove the negation of (4), i.e.

$$(\exists \varepsilon>0)(\forall \delta>0)(\exists x, p \in B) \quad \rho(x, p)<\delta \text{ and } \rho'(f(x), f(p)) \geq \varepsilon. \tag{4.8.40}$$

Take $\varepsilon=1$ and any$\delta>0$. We look for x, p such that

$$|x-p|<\delta \text{ and }|f(x)-f(p)| \geq \varepsilon, \tag{4.8.41}$$

i.e.,

$$\left|\frac{1}{x}-\frac{1}{p}\right| \geq 1, \tag{4.8.42}$$

This is achieved by taking

$$p=\min \left(\delta, \frac{1}{2}\right), x=\frac{p}{2}. \quad \text{(Verify!)} \tag{4.8.43}$$

Thus (4) fails on $B=(0,+\infty),$ yet it holds on $[a,+\infty)$ for any $a>0$.
(Verify!)

4.8.E: Problems on Uniform Continuity; Continuity on Compact Sets

Exercise $4.8.E.1$

Prove that if f is relatively continuous on each compact subset of D, then it is relatively continuous on D.
[Hint: Use Theorem 1 of §2 and Problem 7 in §6.]

Exercise $4.8.E.2$

Do Problem 4 in Chapter 3, §17, and thus complete the last details in the proof of Theorem 4.

Exercise $4.8.E.3$

Give an example of a continuous one-to-one map f such that f^{-1} is not continuous.
[Hint: Show that any map is continuous on a discrete space (S, ρ).]

Exercise $4.8.E.4$

Give an example of a continuous function f and a compact set $D \subseteq (T, \rho')$ such that $f^{-1}[D]$ is not compact.
[Hint: Let f be constant on E^1.]

Exercise $4.8.E.5$

Complete the missing details in Examples (1) and (2) and $(\mathrm{c})-(\mathrm{h})$.

Exercise $4.8.E.6$

Show that every polynomial of degree one on $E^n(^*\text{or } C^n)$ is uniformly continuous.

Exercise $4.8.E.7$

Show that the arcsine function is uniformly continuous on $[-1,1]$.
$[\text{Hint: Use Example (d) and Theorems 3 and 4.}]$

Exercise $4.8.E.8$

$\Rightarrow 8.$ Prove that if f is uniformly continuous on B, and if $\{x_m\} \subseteq B$ is a Cauchy sequence, so is $\{f(x_m)\}$. (Briefly, f preserves Cauchy sequences.) Show that this may fail if f is only continuous in the ordinary sense. (See Example (h).)

Exercise $4.8.E.9$

Prove that if $f : S \rightarrow T$ is uniformly continuous on $B \subseteq S$, and $g : T \rightarrow U$ is uniformly continuous on $f[B]$, then the composite function $g \circ f$ is uniformly continuous on B.

Exercise $4.8.E.10$

Show that the functions f and f^{-1} in Problem 5 of Chapter 3, §11 are contraction maps, 5 hence uniformly continuous. By Theorem 1, find again that (E^*, ρ') is compact.

Exercise 4.8.E.11

Let A' be the set of all cluster points of $A \subseteq (S, \rho)$. Let $f : A \rightarrow (T, \rho')$ be uniformly continuous on A, and let (T, ρ') be complete.
(i) Prove that $\lim_{x \rightarrow p} f(x)$ exists at each $p \in A'$.
(ii) Thus define $f(p) = \lim_{x \rightarrow p} f(x)$ for each $p \in A' - A$, and show that f so extended is uniformly continuous on the set $\overline{A} = A \cup A'$.
(iii) Consider, in particular, the case $A = (a, b) \subseteq E^1$, so that

$$\overline{A} = A' = [a, b]. \tag{4.8.E.1}$$

[Hint: Take any sequence $\{x_m\} \subseteq A, x_m \rightarrow p \in A'$. As it is Cauchy (why?), so is $\{f(x_m)\}$ by Problem 8. Use Corollary 1 in §2 to prove existence of $\lim_{x \rightarrow p} f(x)$. For uniform continuity, use definitions; in case (iii), use Theorem 4 .]

Exercise 4.8.E.12

Prove that if two functions f, g with values in a normed vector space are uniformly continuous on a set B, so also are $f \pm g$ and af for a fixed scalar a.
For real functions, prove this also for $f \vee g$ and $f \wedge g$ defined by

$$(f \vee g)(x) = \max(f(x), g(x)) \tag{4.8.E.2}$$

and

$$(f \wedge g)(x) = \min(f(x), g(x)). \tag{4.8.E.3}$$

[Hint: After proving the first statements, verify that

$$\max(a, b) = \frac{1}{2}(a + b + |b - a|) \text{ and } \min(a, b) = \frac{1}{2}(a + b - |b - a|) \tag{4.8.E.4}$$

and use Problem 9 and Example (b).]

Exercise 4.8.E.13

Let f be vector valued and h scalar valued, with both uniformly continuous on $B \subseteq (S, \rho)$.
Prove that
(i) if f and h are bounded on B, then hf is uniformly continuous on B;
(ii) the function f/h is uniformly continuous on B if f is bounded on B and h is "bounded away" from 0 on B, i.e.,

$$(\exists \delta > 0)(\forall x \in B) \quad |h(x)| \geq \delta. \tag{4.8.E.5}$$

Give examples to show that without these additional conditions, hf and f/h may not be uniformly continuous (see Problem 14 below).

Exercise 4.8.E.14

In the following cases, show that f is uniformly continuous on $B \subseteq E^1$, but only continuous (in the ordinary sense) on D, as indicated, with $0 < a < b < +\infty$.
(a) $f(x) = \frac{1}{x^2}; B = [a, +\infty); D = (0, 1)$.
(b) $f(x) = x^2; B = [a, b]; D = [a, +\infty)$.

(c) $f(x)=\sin\frac{1}{x}$; B and D as in (a).
(d) $f(x)=x\cos x$; B and D as in (b).

Exercise 4.8.E.15

Prove that if f is uniformly continuous on B, it is so on each subset $A\subseteq B$.

Exercise 4.8.E.16

For nonvoid sets $A, B\subseteq(S,\rho)$, define

$$\rho(A,B)=\inf\{\rho(x,y)|x\in A, y\in B\}. \tag{4.8.E.6}$$

Prove that if $\rho(A,B)>0$ and if f is uniformly continuous on each of A and B, it is so on $A\cup B$.
Show by an example that this fails if $\rho(A,B)=0$, even if $A\cap B=\emptyset$ $(\text{e. }g.,\text{ take } A=[0,1], B=(1,2] \text{ in } E^1, \text{ making } f \text{ constant on each of } A \text{ and } B)$.
Note, however, that if A and B are compact, $A\cap B=\emptyset$ implies $\rho(A,B)>0$. (Prove it using Problem 13 in §6.) Thus $A\cap B=\emptyset$ suffices in this case.

Exercise 4.8.E.17

Prove that if f is relatively continuous on each of the disjoint closed sets

$$F_1, F_2, \ldots, F_n, \tag{4.8.E.7}$$

it is relatively continuous on their union

$$F=\bigcup_{k=1}^{n} F_k; \tag{4.8.E.8}$$

hence (see Problem 6 of §6) it is uniformly continuous on F if the F_k are compact.
[Hint: Fix any $p\in F$. Then p is in some F_k, say, $p\in F_1$. As the F_k are disjoint, $p\notin F_2,\ldots,F_p$; hence p also is no cluster point of any of $F_2,\ldots,F_n$ (for they are closed).
Deduce that there is a globe $G_p(\delta)$ disjoint from each of $F_2,\ldots,F_n$, so that $F\cap G_p(\delta)=F_1\cap G_p(\delta)$. From this it is easy to show that relative continuity of f on F follows from relative continuity on F_1.]

Exercise 4.8.E.18

\Rightarrow 18 .\) Let $\overline{p}_0, \overline{p}_1, \ldots, \overline{p}_m$ be fixed points in E^n (* or in another normed space).
Let

$$f(t)=\overline{p}_k+(t-k)\left(\overline{p}_{k+1}-\overline{p}_k\right) \tag{4.8.E.9}$$

whenever $k\leq t\leq k+1, t\in E^1, k=0,1,\ldots,m-1$.
Show that this defines a uniformly continuous mapping f of the interval $[0,m]\subseteq E^1$ onto the "polygon"

$$\bigcup_{k=0}^{m-1} L[p_k, p_{k+1}]. \tag{4.8.E.10}$$

In what case is f one to one? Is f^{-1} uniformly continuous on each $L[p_k, p_{k+1}]$? On the entire polygon?
[Hint: First prove ordinary continuity on $[0,m]$ using Theorem 1 of §3. (For the points $1,2,\ldots,m-1$, consider left and right limits.) Then use Theorems $1-4$.]

Exercise $4.8.E.19$

Prove the sequential criterion for uniform continuity: A function $f : A \rightarrow T$ is uniformly continuous on a set $B \subseteq A$ iff for any two (not necessarily convergent) sequences $\{x_m\}$ and $\{y_m\}$ in B, with $\rho\left(x_m, y_m\right) \rightarrow 0$, we have $\rho'\left(f\left(x_m\right), f\left(y_m\right)\right) \rightarrow 0$ (i.e., f preserves con-current pairs of sequences; see Problem 4 in Chapter 3, §17).

4.9: The Intermediate Value Property

This page is a draft and is under active development.

Definition: intermediate value property

A function $f: A \rightarrow E^{*}$ is said to have the **intermediate value property**, or Darboux property, [1] on a set $B \subseteq A$ iff, together with any two function values $f(p)$ and $f\left(p_{1}\right)\left(p, p_{1} \in B\right)$, it also takes all intermediate values between $f(p)$ and $f\left(p_{1}\right)$ at some points of B.

In other words, the image set $f[B]$ contains the entire interval between $f(p)$ and $f\left(p_{1}\right)$ in E^{*}.

Note 1. It follows that $f[B]$ itself is a finite or infinite interval in E^{*}, with endpoints inf $f[B]$ and sup $f[B]$. (Verify!)

Geometrically, if $A \subseteq E^{1}$, this means that the curve $y=f(x)$ meets all horizontal lines $y=q$, for q between $f(p)$ and $f\left(p_{1}\right)$. For example, in Figure 13 in §1, we have a "smooth" curve that cuts each horizontal line $y=q$ between $f(0)$ and $f\left(p_{1}\right)$; so f has the Darboux property on $\left[0, p_{1}\right]$. In Figures 14 and 15, there is a "gap" at p; the property fails. In Example (f) of §1, the property holds on all of E^{1} despite a discontinuity at 0. Thus it does not imply continuity.

Intuitively, it seems plausible that a "continuous curve" must cut all intermediate horizontals. A precise proof for functions continuous on an interval, was given independently by Bolzano and Weierstrass (the same as in Theorem 2 of Chapter 3, §16). Below we give a more general version of Bolzano's proof based on the notion of a convex set and related concepts.

Definition: Convex

A set B in E^{n} (* or in another normed space) is said to be **convex** iff for each $\bar{a}, \bar{b} \in B$ the line segment $L[\bar{a}, \bar{b}]$ is a subset of B.

A polygon joining $\bar{a}$ and $\bar{b}$ is any finite union of line segments (a "broken line") of the form

$$\bigcup_{i=0}^{m-1} L\left[\bar{p}_{i}, \bar{p}_{i+1}\right] \text { with } \bar{p}_{0}=\bar{a} \text { and } \bar{p}_{m}=\bar{b}. \tag{4.9.1}$$

The set B is said to be polygon connected (or piecewise convex) iff any two points $\bar{a}, \bar{b} \in B$ can be joined by a polygon contained in B.

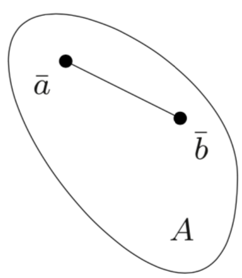

Figure 19

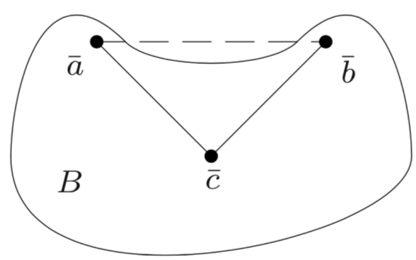

Figure 20

Example

Any globe in E^{n} (* or in another normed space) is convex, so also is any interval in E^{n} or in E^{*}. Figures 19 and 20 represent a convex set A and a polygon-connected set B in E^{2} is not convex; it has a "cavity").

We shall need a simple lemma that is noteworthy in its own right as well.

Lemma 4.9.1 (principle of nested line segments)

Every contracting sequence of closed line segments $L\left[\overline{p}_{m}, \overline{q}_{m}\right]$ in E^{n} (* or in any other normed space) has a nonvoid intersection; i.e., there is a point

$$\overline{p} \in \bigcap_{m=1}^{\infty} L[\overline{p}_{m}, \overline{q}_{m}]. \tag{4.9.2}$$

Proof

Use Cantor's theorem (Theorem 5 of §6) and Example (1) in §8. □

We are now ready for Bolzano's theorem. The proof to be used is typical of so-called "bisection proofs." (See also §6, Problems 9 and 10 for such proofs.)

Theorem 4.9.1: Bolzano's theorem

If $f : B \rightarrow E^{1}$ is relatively continuous on a polygon-connected set B in E^{n} (* or in another normed space), then f has the Darboux property on B.

In particular, if B is convex and if $f(\overline{p}) < c < f(\overline{q})$ for some $\overline{p}, \overline{q} \in B$, then there is a point $\overline{r} \in L(\overline{p}, \overline{q})$ such that $f(\overline{r}) = c$.

Proof

First, let B be convex. Seeking a contradiction, suppose $\overline{p}, \overline{q} \in B$ with

$$f(\overline{p}) < c < f(\overline{q}), \tag{4.9.3}$$

yet $f(\overline{x}) \neq c$ for all $\overline{x} \in L(\overline{p}, \overline{q})$.

Let P be the set of all those $\overline{x} \in L[\overline{p}, \overline{q}]$ for which $f(\overline{x}) < c$, i.e.,

$$P = \{\overline{x} \in L[\overline{p}, \overline{q}] | f(\overline{x}) < c\}, \tag{4.9.4}$$

and let

$$Q = \{\overline{x} \in L[\overline{p}, \overline{q}] | f(\overline{x}) > c\}. \tag{4.9.5}$$

Then $\overline{p} \in P, \overline{q} \in Q, P \cap Q = \emptyset,$ and $P \cup Q = L[\overline{p}, \overline{q}] \subseteq B$. (Why?)

Now let

$$\overline{r}_{0} = \frac{1}{2}(\overline{p} + \overline{q}) \tag{4.9.6}$$

be the midpoint on $L[\overline{p}, \overline{q}]$. Clearly, $\overline{r}_{0}$ is either in P or in Q. Thus it bisects $L[\overline{p}, \overline{q}]$ into two subsegments, one of which must have its left endpoint in P and its right endpoint in Q.

We denote this particular closed segment by $L\left[\overline{p}_{1}, \overline{q}_{1}\right], \overline{p}_{1} \in P, \overline{q}_{1} \in Q$. We then have

$$L[\overline{p}_{1}, \overline{q}_{1}] \subseteq L[\overline{p}, \overline{q}] \text{ and } |p_{1} - q_{1}| = \frac{1}{2}|\overline{p} - \overline{q}|. \text{ (Verify!)} \tag{4.9.7}$$

Now we bisect $L[\overline{p}_{1}, \overline{q}_{1}]$ and repeat the process. Thus let

$$\overline{r}_{1} = \frac{1}{2}(\overline{p}_{1} + \overline{q}_{1}). \tag{4.9.8}$$

By the same argument, we obtain a closed subsegment $L\left[\bar{p}_2, \bar{q}_2\right] \subseteq L\left[\bar{p}_1, \bar{q}_1\right]$ with $\bar{p}_2 \in P, \bar{q}_2 \in Q$, and

$$\left|\bar{p}_2-\bar{q}_2\right|=\frac{1}{2}\left|\bar{p}_1-\bar{q}_1\right|=\frac{1}{4}|\bar{p}-\bar{q}|. \tag{4.9.9}$$

Next, we bisect $L\left[\bar{p}_2, \bar{q}_2\right]$, and so on. Continuing this process indefinitely, we obtain an infinite contracting sequence of closed line segments $L\left[\bar{p}_m, \bar{q}_m\right]$ such that

$$(\forall m) \quad \bar{p}_m \in P, \bar{q}_m \in Q, \tag{4.9.10}$$

and

$$\left|\bar{p}_m-\bar{q}_m\right|=\frac{1}{2^m}|\bar{p}-\bar{q}| \rightarrow 0 \text{ as } m \rightarrow+\infty. \tag{4.9.11}$$

By Lemma 1, there is a point

$$\bar{r} \in \bigcap_{m=1}^{\infty} L\left[\bar{p}_m, \bar{q}_m\right]. \tag{4.9.12}$$

This implies that

$$(\forall m) \quad\left|\bar{r}-\bar{p}_m\right| \leq\left|\bar{p}_m-\bar{q}_m\right| \rightarrow 0, \tag{4.9.13}$$

whence $\bar{p}_m \rightarrow \bar{r}$. Similarly, we obtain $\bar{q}_m \rightarrow \bar{r}$.

Now since $\bar{r} \in L[\bar{p}, \bar{q}] \subseteq B$, the function f is relatively continuous at $\bar{r}$ over B (by assumption). By the sequential criterion, then,

$$f\left(\bar{p}_m\right) \rightarrow f(\bar{r}) \text{ and } f\left(\bar{q}_m\right) \rightarrow f(\bar{r}). \tag{4.9.14}$$

Moreover, $f\left(\bar{p}_m\right)<c<f\left(\bar{q}_m\right)\left(\text{for } \bar{p}_m \in P \text{ and } \bar{q}_m \in Q\right)$. Letting $m \rightarrow+\infty$, we pass to limits (Chapter 3, §15, Corollary 1) and get

$$f(\bar{r}) \leq c \leq f(\bar{r}), \tag{4.9.15}$$

so that $\bar{r}$ is neither in P nor in Q, which is a contradiction. This completes the proof for a convex B.

The extension to polygon-connected sets is left as an exercise (see Problem 2 below). Thus all is proved. □

Note 2. In particular, the theorem applies if B is a globe or an interval.

Thus continuity on an interval implies the Darboux property. The converse fails, as we have noted. However, for monotone functions, we obtain the following theorem.

Theorem 4.9.2

If a function $f: A \rightarrow E^1$ is monotone and has the Darboux property on a finite or infinite interval $(a, b) \subseteq A \subseteq E^1$, then it is continuous on (a, b).

Proof

Seeking a contradiction, suppose f is discontinuous at some $p \in(a, b)$.

For definiteness, let $f \uparrow$ on (a, b). Then by Theorems 2 and 3 in §5, we have either $f\left(p^{-}\right)<f(p)$ or $f(p)<f\left(p^{+}\right)$ or both, with no function values in between.

On the other hand, since f has the Darboux property, the function values $f(x)$ for x in (a, b) fill an entire interval (see Note 1). Thus it is impossible for $f(p)$ to be the only function value between $f\left(p^{-}\right)$ and $f\left(p^{+}\right)$ unless f is constant near p, but then it is also continuous at p, which we excluded. This contradiction completes the proof. □

Note 3. The theorem holds (with a similar proof) for nonopen intervals as well, but the continuity at the endpoints is relative (right at a, left at b).

Theorem 4.9.3

If $f: A \rightarrow E^1$ is strictly monotone and continuous when restricted to a finite or infinite interval $B \subseteq A \subseteq E^1$, then its inverse f^{-1} has the same properties on the set $f[B]$ (itself an interval, by Note 1 and Theorem 1).

Proof

It is easy to see that f^{-1} is increasing (decreasing) if f is; the proof is left as an exercise. Thus f^{-1} is monotone on $f[B]$ if f is so on B. To prove the relative continuity of f^{-1}, we use Theorem 2, i.e., show that f^{-1} has the Darboux property on $f[B]$.

Thus let $f^{-1}(p)<c<f^{-1}(q)$ for some $p, q \in f[B]$. We look for an $r \in f[B]$ such that $f^{-1}(r)=c$, i.e., $r=f(c)$. Now since $p, q \in f[B]$, the numbers $f^{-1}(p)$ and $f^{-1}(q)$ are in B, an interval. Hence also the intermediate value c is in B ; thus it belongs to the domain of f, and so the function value $f(c)$ exists. It thus suffices to put $r=f(c)$ to get the result. $\square$

ExampleS

(a) Define $f: E^1 \rightarrow E^1$ by

$$f(x)=x^n \text{ for a fixed } n \in N. \tag{4.9.16}$$

As f is continuous (being a monomial), it has the Darboux property on E^1. By Note 1, setting $B=[0,+\infty)$, we have $f[B]=[0,+\infty)$. (Why?) Also, f is strictly increasing on B. Thus by Theorem 3, the inverse function f^{-1} (i.e., the nth root function) exists and is continuous on $f[B]=[0,+\infty)$.

If n is odd, then f^{-1} has these properties on all of E^1, by a similar proof; thus $\sqrt[n]{x}$ exists for $x \in E^1$.

(b) Logarithmic functions. From the example in §5, we recall that the exponential function given by

$$F(x)=a^x \quad(a>0) \tag{4.9.17}$$

is continuous and strictly monotone on E^1. Its inverse, F^{-1}, is called the logarithmic function to the base a, denoted log. By Theorem 3, it is continuous and strictly monotone on $F\left[E^1\right]$.

To fix ideas, let $a>1$, so $F \uparrow$ and $\left(F^{-1}\right) \uparrow$. By Note 1, $F\left[E^1\right]$ is an interval with endpoints p and r, where

$$p=\inf F\left[E^1\right]=\inf \left\{a^x \mid-\infty<x<+\infty\right\} \tag{4.9.18}$$

and

$$r=\sup F\left[E^1\right]=\sup \left\{a^x \mid-\infty<x<+\infty\right\}. \tag{4.9.19}$$

Now by Problem 14 (iii) of §2 (with $q=0$),

$$\lim _{x \rightarrow+\infty} a^x=+\infty \text{ and } \lim _{x \rightarrow-\infty} a^x=0. \tag{4.9.20}$$

As $F \uparrow$, we use Theorem 1 in §5 to obtain

$$r=\sup a^x=\lim _{x \rightarrow+\infty} a^x=+\infty \text{ and } p=\lim _{x \rightarrow-\infty} a^x=0. \tag{4.9.21}$$

Thus $F\left[E^1\right]$, i.e., the domain of $\log_a$, is the interval $(p, r)=(0,+\infty)$. It follows that $\log_a x$ is uniquely defined for x in $(0,+\infty)$; it is called the logarithm of x to the base a.

The range of $\log_a$ (i.e. of F^{-1}) is the same as the domain of F, i.e., E^1. Thus if $a>1, \log_a x$ increases from $-\infty$ to $+\infty$ as x increases from 0 to $+\infty$. Hence

$$\lim _{x \rightarrow+\infty} \log_a x=+\infty \text{ and } \lim _{x \rightarrow 0+} \log_a x=-\infty, \tag{4.9.22}$$

provided $a>1$.

If $0<a<1$, the values of these limits are interchanged (since $F \downarrow$ in this case), but otherwise the results are the same.

If $a=e$, we write $\ln x$ or $\log x$ for $\log_a x$, and we call $\ln x$ the natural logarithm of x. Its inverse is, of course, the exponential $f(x)=e^x$, also written $\exp(x)$. Thus by definition, $\ln e^x = x$ and

$$x=\exp(\ln x)=e^{\ln x} \quad (0<x<+\infty). \tag{4.9.23}$$

(c) The power function $g:(0,+\infty)\rightarrow E^1$ is defined by

$$g(x)=x^a \text{ for a fixed real } a. \tag{4.9.24}$$

If $a>0$, we also define $g(0)=0$. For $x>0$, we have

$$x^a=\exp(\ln x^a)=\exp(a\cdot \ln x). \tag{4.9.25}$$

Thus by the rules for composite functions (Theorem 3 and Corollary 2 in §2), the continuity of g on $(0,+\infty)$ follows from that of exponential and log functions. If $a>0, g$ is also continuous at 0. (Exercise!)

4.9.E: Problems on the Darboux Property and Related Topics

Exercise $4.9.E.1$

Prove Note 1.

Exercise $4.9.E.1'$

Prove Note 3.

Exercise $4.9.E.1''$

Prove continuity at 0 in Example (c).

Exercise $4.9.E.2$

Prove Theorem 1 for polygon-connected sets.
[Hint: If

$$B \supseteq \bigcup_{i=0}^{m-1} L\left[\bar{p}_i, \bar{p}_{i+1}\right] \tag{4.9.E.1}$$

with

$$f\left(\bar{p}_0\right) < c < f\left(\bar{p}_m\right), \tag{4.9.E.2}$$

show that for at least one i, either $c = f\left(\bar{p}_i\right)$ or $f\left(\bar{p}_i\right) < c < f\left(\bar{p}_{i+1}\right)$. Then replace B in the theorem by the convex segment $L\left[\bar{p}_i, \bar{p}_{i+1}\right]$.]

Exercise $4.9.E.3$

Show that, if f is strictly increasing on $B \subseteq E$, then f^{-1} has the same property on $f[B]$, and both are one to one; similarly for decreasing functions.

Exercise $4.9.E.4$

For functions on $B = [a, b] \subset E^1$, Theorem 1 can be proved thusly: If

$$f(a) < c < f(b), \tag{4.9.E.3}$$

let

$$P = \{x \in B \mid f(x) < c\} \tag{4.9.E.4}$$

and put $r = \sup P$.
Show that $f(r)$ is neither greater nor less than c, and so necessarily $f(r) = c$.
[Hint: If $f(r) < c$, continuity at r implies that $f(x) < c$ on some $G_r(\delta)$ (§2, Problem 7), contrary to $r = \sup P$. (Why?)]

Exercise $4.9.E.5$

Continuing Problem 4, prove Theorem 1 in all generality, as follows.
Define

$$g(t)=\overline{p}+t(\overline{q}-\overline{p}), \quad 0 \leq t \leq 1. \tag{4.9.E.5}$$

Then g is continuous (by Theorem 3 in §3), and so is the composite function $h=f \circ g$, on $[0,1]$. By Problem 4, with $B=[0,1]$, there is a $t \in (0,1)$ with $h(t)=c$. Put $\overline{r}=g(t)$, and show that $f(\overline{r})=c$.

Exercise $4.9.E.6$

Show that every equation of odd degree, of the form

$$f(x)=\sum_{k=0}^{n} a_k x^k=0 \quad (n=2m-1, a_n \neq 0), \tag{4.9.E.6}$$

has at least one solution for x in E^1.
[Hint: Show that f takes both negative and positive values as $x \rightarrow -\infty$ or $x \rightarrow +\infty$; thus by the Darboux property, f must also take the intermediate value 0 for some $x \in E^1$.]

Exercise $4.9.E.7$

Prove that if the functions $f: A \rightarrow (0,+\infty)$ and $g: A \rightarrow E^1$ are both continuous, so also is the function $h: A \rightarrow E^1$ given by

$$h(x)=f(x)^{g(x)}. \tag{4.9.E.7}$$

[Hint: See Example (c)].

Exercise $4.9.E.8$

Using Corollary 2 in §2, and limit properties of the exponential and log functions, prove the "shorthand" Theorems $11-16$ of §4.

Exercise $4.9.E.8'$

Find $\lim_{x \rightarrow +\infty}\left(1+\frac{1}{x}\right)^{\sqrt{x}}$.

Exercise $4.9.E.8''$

Similarly, find a new solution of Problem 27 in Chapter 3, §15, reducing it to Problem 26.

Exercise $4.9.E.9$

Show that if $f: E^1 \rightarrow E^*$ has the Darboux property on B (e.g., if B is convex and f is relatively continuous on B) and if f is one to one on B, then f is necessarily strictly monotone on B.

Exercise $4.9.E.10$

Prove that if two real functions f, g are relatively continuous on $[a, b]$ $(a<b)$ and

$$f(x) g(x)>0 \text{ for } x \in[a, b], \tag{4.9.E.8}$$

then the equation

$$(x-a) f(x)+(x-b) g(x)=0 \tag{4.9.E.9}$$

has a solution between a and b; similarly for the equation

$$\frac{f(x)}{x-a}+\frac{g(x)}{x-b}=0 \quad\left(a, b \in E^{1}\right). \tag{4.9.E.10}$$

Exercise $4.9.E.10'$

Similarly, discuss the solutions of

$$\frac{2}{x-4}+\frac{9}{x-1}+\frac{1}{x-2}=0. \tag{4.9.E.11}$$

4.10: Arcs and Curves. Connected Sets

This page is a draft and is under active development.

A deeper insight into continuity and the Darboux property can be gained by generalizing the notions of a convex set and polygon-connected set to obtain so-called connected sets.

I. As a first step, we consider arcs and curves.

Definition

A set $A \subseteq (S, \rho)$ is called an arc iff A is a continuous image of a compact interval $[a, b] \subset E^1$, i.e., iff there is a continuous mapping

$$f : [a, b] \underset{\text{onto}}{\longrightarrow} A. \tag{4.10.1}$$

If, in addition, f is one to one, A is called a simple arc with endpoints $f(a)$ and $f(b)$.
If instead $f(a) = f(b)$, we speak of a closed curve.
A curve is a continuous image of any finite or infinite interval in E^1.

corollary 4.10.1

Each arc is a compact (hence closed and bounded) set (by Theorem 1 of §8).

Definition

A set $A \subseteq (S, \rho)$ is said to be arcwise connected iff every two points $p, q \in A$ are in some simple arc contained in A. (We then also say the p and q can be joined by an arc in A.)

Example 4.10.1

(a) Every closed line segment $L[\bar{a}, \bar{b}]$ in E^n (* or in any other normed space) is a simple arc (consider the map f in Example (1) of §8).
(b) Every polygon

$$A = \bigcup_{i=0}^{m-1} L\left[\bar{p}_i, \bar{p}_{i+1}\right] \tag{4.10.2}$$

is an arc (see Problem 18 in §8). It is a simple arc if the half-closed segments $L\left[\bar{p}_i, \bar{p}_{i+1}\right)$ do not intersect and the points $\bar{p}_i$ are distinct, for then the map f in Problem 18 of §8 is one to one.
(c) It easily follows that every polygon-connected set is also arcwise connected; one only has to show that every polygon joining two points $\bar{p}_0, \bar{p}_m$ can be reduced to a simple polygon (not a self-intersecting one). See Problem 2.
However, the converse is false. For example, two discs in E^2 connected by a parabolic arc form together an arcwise- (but not polygonwise-) connected set.
(d) Let $f_1, f_2, \ldots, f_n$ be real continuous functions on an interval $I \subseteq E^1$. Treat them as components of a function $f : I \rightarrow E^n$,

$$f = (f_1, \ldots, f_n). \tag{4.10.3}$$

Then f is continuous by Theorem 2 in §3. Thus the image set $f[I]$ is a

curve in E^n; it is an arc if I is a closed interval.
Introducing a parameter t varying over I, we obtain the parametric equations of the curve, namely,

$$x_k = f_k(t), \quad k = 1, 2, \ldots, n. \tag{4.10.4}$$

Then as t varies over I, the point $\bar{x} = (x_1, \ldots, x_n)$ describes the curve $f[I]$. This is the usual way of treating curves in E^n (* and C^n).

It is not hard to show that Theorem 1 in §9 holds also if B is only arcwise connected (see Problem 3 below). However, much more can be proved by introducing the general notion of a connected set. We do this next.

II. For this topic, we shall need Theorems 2-4 of Chapter 3, §12, and Problem 15 of Chapter 4, §2. The reader is advised to review them. In particular, we have the following theorem.

Theorem 4.10.1

A function $f : (A, \rho) \to (T, \rho')$ is continuous on A iff $f^{-1}[B]$ is closed in (A, ρ) for each closed set $B \subseteq (T, \rho')$; similarly for open sets.

Indeed, this is part of Problem 15 in §2 with (S, ρ) replaced by (A, ρ).

Definition

A metric space (S, ρ) is said to be connected iff S is not the union $P \cup Q$ of any two nonvoid disjoint closed sets; it is disconnected otherwise.
A set $A \subseteq (S, \rho)$ is called connected iff (A, ρ) is connected as a subspace of (S, ρ); i.e., iff A is not a union of two disjoint sets $P, Q \neq \emptyset$ that are closed (hence also open) in (A, ρ), as a subspace of (S, ρ).

Note 1. By Theorem 4 of Chapter 3, §12, this means that

$$P = A \cap P_1 \text{ and } Q = A \cap Q_1 \tag{4.10.5}$$

for some sets P_1, Q_1 that are closed in (S, ρ). Observe that, unlike compact sets, a set that is closed or open in (A, ρ) need not be closed or open in (S, ρ).

Example 4.10.1

(a') $\emptyset$ is connected.

(b') So is any one-point set $\{p\}$. (Why?)

(c') Any finite set of two or more points is disconnected. (Why?)

Other examples are provided by the theorems that follow.

Theorem 4.10.2

The only connected sets in E^1 are exactly all convex sets, i.e., finite and infinite intervals, including E^1 itself.

Proof

The proof that such intervals are exactly all convex sets in E^1 is left as an exercise.

Seeking a contradiction, suppose $p \notin A$ for some $p \in (a, b)$, $a, b \in A$. Let

$$P = A \cap (-\infty, p) \text{ and } Q = A \cap (p, +\infty). \tag{4.10.6}$$

Then $A = P \cup Q, a \in P, b \in Q,$ and $P \cap Q = \emptyset$. Moreover, $(-\infty, p)$ and $(p, +\infty)$ are open sets in E^1. (Why?) Hence P and Q are open in A, each being the intersection of A with a set open in E^1 (see Note 1 above). As $A = P \cup Q,$ with $P \cap Q = \emptyset,$ it follows that A is disconnected. This shows that if A is connected in E^1, it must be convex.

Conversely, let A be convex in E^1. The proof that A is connected is an almost exact copy of the proof given for Theorem 1 of §9, so we only briefly
sketch it here.

If A were disconnected, then $A=P \cup Q$ for some disjoint sets $P, Q \neq \emptyset$, both closed in A. Fix any $p \in P$ and $q \in Q$. Exactly as in Theorem 1 of §9, select a contracting sequence of line segments (intervals) $[p_m, q_m] \subseteq A$ such that $p_m \in P, q_m \in Q,$ and $|p_m - q_m| \rightarrow 0,$ and obtain a point

$$r \in \bigcap_{m=1}^{\infty}\left[p_m, q_m\right] \subseteq A \tag{4.10.7}$$

so that $p_m \rightarrow r, q_m \rightarrow r,$ and $r \in A$. As the sets P and Q are closed in (A, ρ), Theorem 4 of Chapter $3, \$16$ shows that both P and Q must contain the common limit r of the sequences $\{p_m\} \subseteq P$ and $\{q_m\} \subseteq Q$. This is impossible, however, since $P \cap Q=\emptyset,$ by assumption. This contradiction shows that A cannot be disconnected. Thus all is proved. $\square$

Note 2. By the same proof, any convex set in a normed space is connected. In particular, E^n and all other normed spaces are connected themselves.

Theorem 4.10.3

If a function $f: A \rightarrow (T, \rho')$ with $A \subseteq (S, \rho)$ is relatively continuous on a connected set $B \subseteq A,$ then $f[B]$ is a connected set in (T, ρ').

Proof

By definition (§1), relative continuity on B becomes ordinary continuity when f is restricted to B. Thus we may treat f as a mapping of B into $f[B]$, replacing S and T by their subspaces B and $f[B]$.

Seeking a contradiction, suppose $f[B]$ is disconnected, i.e.,

$$f[B]=P \cup Q \tag{4.10.8}$$

for some disjoint sets $P, Q \neq \emptyset$ closed in $(f[B], \rho')$. Then by Theorem 1, with T replaced by $f[B]$, the sets $f^{-1}[P]$ and $f^{-1}[Q]$ are closed in (B, ρ). They also are nonvoid and disjoint (as are P and Q) and satisfy

$$B=f^{-1}[P \cup Q]=f^{-1}[P] \cup f^{-1}[Q] \tag{4.10.9}$$

\[
(see Chapter 1, §4-7, Problem 6). Thus B is disconnected, contrary to assumption. $\square$

corollary 4.10.2

All arcs and curves are connected sets (by Definition 2 and Theorems 2 and 3).

lemma 4.10.1

A set $A \subseteq (S, \rho)$ is connected iff any two points $p, q \in A$ are in some connected subset $B \subseteq A.$ Hence any arcwise connected set is connected.

Proof

Seeking a contradiction, suppose the condition stated in Lemma 1 holds but A is disconnected, so $A=P \cup Q$ for some disjoint sets $P \neq \emptyset, Q \neq \emptyset$ both closed in (A, ρ).

Pick any $p \in P$ and $q \in Q$. By assumption, p and q are in some connected set $B \subseteq A.$ Treat (B, ρ) as a subspace of (A, ρ), and let

$$P'=B \cap P \text{ and } Q'=B \cap Q. \tag{4.10.10}$$

Then by Theorem 4 of Chapter 3, §12, P' and Q' are closed in B. Also, they are disjoint (for P and Q are) and nonvoid (for $p \in P', q \in Q'$), and

$$B=B \cap A=B \cap(P \cup Q)=(B \cap P) \cup(B \cap Q)=P' \cup Q'. \tag{4.10.11}$$

Thus B is disconnected, contrary to assumption. This contradiction proves the lemma (the converse proof is trivial).

In particular, if A is arcwise connected, then any points p, q in A are in some arc $B \subseteq A$, a connected set by Corollary 2. Thus all is proved. □

corollary 4.10.3

Any convex or polygon-connected set $(e.g., a$ globe$)$ in E^n (or in any other normed space) is arcwise connected, hence connected.

Proof

Use Lemma 1 and Example (c) in part I of this section. □

Caution: The converse fails. A connected set need not be arcwise connected, let alone polygon connected (see Problem 17). However, we have the following
theorem.

Theorem 4.10.4

Every open connected set A in E^n (* or in another normed space) is also arcwise connected and even polygon connected.

Proof

If $A=\emptyset$, this is "vacuously" true, so let $A \neq \emptyset$ and fix $\overline{a} \in A$.

Let P be the set of all $\overline{p} \in A$ that can be joined with $\overline{a}$ by a polygon $K \subseteq A$ Let $Q=A-P$. Clearly, $\overline{a} \in P$, so $P \neq \emptyset$. We shall show that P is open, i.e., that each $\overline{p} \in P$ is in a globe $G_{\overline{p}} \subseteq P$.

Thus we fix any $\overline{p} \in P$. As A is open and $\overline{p} \in A$, there certainly is a globe $G_{\overline{p}}$ contained in A. Moreover, as $G_{\overline{p}}$ is convex, each point $\overline{x} \in G_{\overline{p}}$ is joined with $\overline{p}$ by the line segment $L[\overline{x}, \overline{p}] \subseteq G_{\overline{p}}$. Also, as $\overline{p} \in P$, some polygon $K \subseteq A$ joins $\overline{p}$ with $\overline{a}$. Then

$$K \cup L[\overline{x}, \overline{p}] \tag{4.10.12}$$

is a polygon joining $\overline{x}$ and $\overline{a}$, and hence by definition $\overline{x} \in P$. Thus each $\overline{x} \in G_{\overline{p}}$ is in P, so that $G_{\overline{p}} \subseteq P$, as required, and P is open (also apen in A as a subspace).

Next, we show that the set $Q=A-P$ is open as well. As before, if $Q \neq \emptyset$, fix any $\overline{q} \in Q$ and a globe $G_{\overline{q}} \subseteq A$, and show that $G_{\overline{q}} \subseteq Q$. Indeed, if some $\overline{x} \in G_{\overline{q}}$ were *not* in Q, it would be in P, and thus it would be joined with $\overline{a}$ (fixed above) by a polygon $K \subseteq A$. Then, however, $\overline{q}$ itself could be so joined by the polygon

$$L[\overline{q}, \overline{x}] \cup K, \tag{4.10.13}$$

implying that $\overline{q} \in P$, not $\overline{q} \in Q$. This shows that $G_{\overline{q}} \subset Q$ indeed, as claimed.

Thus $A=P \cup Q$ with P, Q disjoint and open (hence clopen) in A. The connectedness of A then implies that $Q=\emptyset$. $(P$ is not empty, as has been noted.) Hence $A=P$. By the definition of P, then, each point $\overline{b} \in A$ can be joined to $\overline{a}$ by a polygon. As $\overline{a} \in A$ was arbitrary, A is polygon connected. □

Finally, we obtain a stronger version of the intermediate value theorem.

Theorem 4.10.5

If a function $f : A \rightarrow E^1$ is relatively continuous on a connected set $B \subseteq A \subseteq (S, \rho)$, then f has the Darboux property on B.

In fact, by Theorems 3 and 2, $f[B]$ is a connected set in E^1, i.e., an interval. This, however, implies the Darboux property.

4.10.E: Problems on Arcs, Curves, and Connected Sets

Exercise $4.10.E.1$

Discuss Examples (a) and (b) in detail. In particular, verify that $L[\bar{a}, \bar{b}]$ is a simple arc. (Show that the map f in Example (1) of §8 is one to one.)

Exercise $4.10.E.2$

Show that each polygon

$$K=\bigcup_{i=0}^{m-1} L\left[\bar{p}_i, \bar{p}_{i+1}\right] \tag{4.10.E.1}$$

can be reduced to a simple polygon $P(P \subseteq K)$ joining p_0 and p_m.
[Hint: First, show that if two line segments have two or more common points, they lie in one line. Then use induction on the number m of segments in K. Draw a diagram in E^2 as a guide.

Exercise $4.10.E.3$

Prove Theorem 1 of §9 for an arcwise connected $B \subseteq(S, \rho)$.
[Hint: Proceed as in Problems 4 and 5 in §9, replacing g by some continuous map $f:[a, b] \longrightarrow B$.]

Exercise $4.10.E.4$

Define f as in Example (f) of §1. Let

$$G_{ab}=\left\{(x, y) \in E^2 | a \leq x \leq b, y=f(x)\right\}. \tag{4.10.E.2}$$

$(G_{ab}$ is the graph of f over $[a, b].)$ Prove the following:
(i) If $a>0,$ then G_{ab} is a simple arc in E^2.
(ii) If $a \leq 0 \leq b, G_{ab}$ is not even arcwise connected.
[Hints: (i) Prove that f is continuous on $[a, b], a>0,$ using the continuity of the sine function. Then use Problem 16 in §2, restricting f to $[a, b]$.
(ii) For a contradiction, assume $\bar{0}$ is joined by a simple arc to some $\bar{p} \in G_{ab}.$]

Exercise $4.10.E.5$

Show that each arc is a continuous image of $[0,1]$.
[Hint: First, show that any $[a, b] \subseteq E^1$ is such an image. Then use a suitable composite mapping.

Exercise $4.10.E.*6$

Prove that a function $f: B \rightarrow E^1$ on a compact set $B \subseteq E^1$ must be continuous if its graph,

$$\left\{(x, y) \in E^2 | x \in B, y=f(x)\right\}, \tag{4.10.E.3}$$

is a compact set (e.g., an arc) in E^2.
[Hint: Proceed as in the proof of Theorem 3 of §8.]

Exercise $4.10.E.*7$

Prove that A is connected iff there is no continuous map

$$f : A \underset{\text{onto}}{\longrightarrow} \{0,1\}. \tag{4.10.E.4}$$

[Hint: If there is such a map, Theorem 1 shows that A is disconnected. (Why?)
Conversely, if $A = P \cup Q(P, Q$ as in Definition $3)$, put $f = 0$ on P and $f = 1$ on Q. Use again Theorem 1 to show that f so defined is continuous on A.]

Exercise $4.10.E.*8$

Let $B \subseteq A \subseteq (S, \rho)$. Prove that B is connected in S iff it is connected in (A, ρ).

Exercise $4.10.E.*9$

Suppose that no two of the sets $A_i (i \in I)$ are disjoint. Prove that if all A_i are connected, so is $A = \bigcup_{i \in I} A_i$.
[Hint: If not, let $A = P \cup Q(P, Q$ as in Definition $3)$. Let $P_i = A_i \cap P$ and $Q_i = A_i \cap Q$, so $A_i = P_i \cup Q_i, i \in I$.
At least one of the P_i, Q_i must be $\emptyset$ (why?); say, $Q_j = \emptyset$ for some $j \in I$. Then
$(\forall i) Q_i = \emptyset$, for $Q_i \neq \emptyset$ implies $P_i = \emptyset$, whence

$$A_i = Q_i \subseteq Q \Longrightarrow A_i \cap A_j = \emptyset \,(\text{ since } A_j \subseteq P), \tag{4.10.E.5}$$

contrary to our assumption. Deduce that $Q = \bigcup_i Q_i = \emptyset$. (Contradiction!)]

Exercise $4.10.E.*10$

Prove that if $\{A_n\}$ is a finite or infinite sequence of connected sets and if

$$(\forall n) \quad A_n \cap A_{n+1} \neq \emptyset, \tag{4.10.E.6}$$

then

$$A = \bigcup_n A_n \tag{4.10.E.7}$$

is connected.
[Hint: Let $B_n = \bigcup_{k=1}^n A_k$. Use Problem 9 and induction to show that the B_n are connected and no two are disjoint. Verify that $A = \bigcup_n B_n$ and apply Problem 9 to the sets B_n.]

Exercise $4.10.E.*11$

Given $p \in A, A \subseteq (S, \rho)$, let A_p denote the union of all connected subsets of A that contain p (one of them is $\{p\}$); A_p is called the p-component of A. Prove that
(i) A_p is connected (use Problem 9);
(ii) A_p is not contained in any other connected set $B \subseteq A$ with $p \in B$;
(iii) $(\forall p, q \in A) A_p \cap A_q = \emptyset$ iff $A_p \neq A_q$; and
(iv) $A = \cup \{A_p | p \in A\}$.
[Hint for (iii): If $A_p \cap A_q \neq \emptyset$ and $A_p \neq A_q$, then $B = A_p \cup A_q$ is a connected set larger than A_p, contrary to (ii).]

Exercise $4.10.E.*12$

Prove that if A is connected, so is its closure (Chapter 3, §16 Definition 1), and so is any set D such that $A \subseteq D \subseteq \overline{A}$.
[Hints: First show that D is the "least" closed set in (D, ρ) that contains A (Problem 11 in Chapter 3, §16 and Theorem 4 of Chapter 3, §12). Next, seeking a contradiction, let $D=P \cup Q, P \cap Q=\emptyset, P, Q \neq \emptyset,$ clopen in D. Then

$$A=(A \cap P) \cup(A \cap Q) \tag{4.10.E.8}$$

proves A disconnected, for if $A \cap P=\emptyset,$ say, then $A \subseteq Q \subset D$ (why?), contrary to the minimality of D; similarly for $A \cap Q=\emptyset$.]

Exercise $4.10.E.*13$

A set is said to be totally disconnected iff its only connected subsets are one-point sets and $\emptyset$.
Show that R (the rationals) has this property in E^1.

Exercise $4.10.E.*14$

Show that any discrete space is totally disconnected (see Problem 13).

Exercise $4.10.E.*15$

From Problems 11 and 12 deduce that each component A_p is closed $\left(A_p=\overline{A_p}\right)$.

Exercise $4.10.E.*16$

Prove that a set $A \subseteq(S, \rho)$ is disconnected iff $A=P \cup Q,$ with $P, Q \neq \emptyset$, and each of P, Q disjoint from the closure of the other: $P \cap \overline{Q}=\emptyset=\overline{P} \cap Q$.
[Hint: By Problem 12, the closure of P in (A, ρ) (i.e., the least closed set in (A, ρ) that contains P) is

$$A \cap \overline{P}=(P \cup Q) \cap \overline{P}=(P \cap \overline{P}) \cup(Q \cap \overline{P})=P \cup \emptyset=P, \tag{4.10.E.9}$$

so P is closed in A; similarly for Q. Prove the converse in the same manner.

Exercise $4.10.E.*17$

Give an example of a connected set that is not arcwise connected.
[Hint: The set $G_{0b}(a=0)$ in Problem 4 is the closure of $G_{0b}-\{\overline{0}\}$ (verify!), and the latter is connected (why?); hence so is G_{0b} by Problem 12.]

4.11: Product Spaces. Double and Iterated Limits

This page is a draft and is under active development.

Given two metric spaces (X, ρ_1) and (Y, ρ_2), we may consider the Cartesian product $X \times Y$, suitably metrized. Two metrics for $X \times Y$ are suggested in Problem 10 in Chapter 3, §11. We shall adopt the first of them as follows.

Definition

By the product of two metric spaces (X, ρ_1) and (Y, ρ_2) is meant the space $(X \times Y, \rho)$, where the metric ρ is defined by

$$\rho((x, y),(x', y')) = \max\{\rho_1(x, x'), \rho_2(y, y')\} \tag{4.11.1}$$

for $x, x' \in X$ and $y, y' \in Y$.

Thus the distance between (x, y) and (x', y') is the larger of the two distances

$$\rho_1(x, x') \text{ in } X \text{ and } \rho_2(y, y') \text{ in } Y. \tag{4.11.2}$$

The verification that ρ in (1) is, indeed, a metric is left to the reader. We now obtain the following theorem.

Theorem 4.11.1

(i) A globe $G_{(p,q)}(\varepsilon)$ in $(X \times Y, \rho)$ is the Cartesian product of the corresponding ε -globes in X and Y,

$$G_{(p,q)}(\varepsilon) = G_p(\varepsilon) \times G_q(\varepsilon). \tag{4.11.3}$$

(ii) Convergence of sequences $\{(x_m, y_m)\}$ in $X \times Y$ is componentwise. That is, we have

$$(x_m, y_m) \rightarrow (p, q) \text{ in } X \times Y \text{ iff } x_m \rightarrow p \text{ in } X \text{ and } y_m \rightarrow q \text{ in } Y. \tag{4.11.4}$$

Proof

Again, the easy proof is left as an exercise.

In this connection, recall that by Theorem 2 of Chapter 3, §15, convergence in E^2 is componentwise as well, even though the standard metric in E^2 is not the product metric (1); it is rather the metric (ii) of Problem 10 in Chapter 3, §11. We might have adopted this second metric for $X \times Y$ as well. Then part (i) of Theorem 1 would fail, but part (ii) would still follow by making

$$\rho_1(x_m, p) < \frac{\varepsilon}{\sqrt{2}} \text{ and } \rho_2(y_m, q) < \frac{\varepsilon}{\sqrt{2}}. \tag{4.11.5}$$

It follows that, as far as convergence is concerned, the two choices of ρ are equivalent.

Note 1. More generally, two metrics for a space S are said to be equivalent iff exactly the same sequences converge (to the same limits) under both metrics. Then also all function limits are the same since they reduential limits, by Theorem 1 of §2; similarly for such notions as continuity, compactness, completeness, closedness, openness, etc.

In view of this, we shall often call $X \times Y$ a product space (in the wider sense) even if its metric is not the ρ of formula (1) but equivalent to it. In this sense, E^2 is the product space $E^1 \times E^1$, and $X \times Y$ is its generalization.

Various ideas valid in E^2 extend quite naturally to $X \times Y$. Thus functions defined on a set $A \subseteq X \times Y$ may be treated as functions of two variables x, y such that $(x, y) \in A$. Given $(p, q) \in X \times Y$, we may consider ordinary or relative limits at (p, q), e.g. limits over a path

$$B = \{(x, y) \in X \times Y | y = q\} \tag{4.11.6}$$

(briefly called the "line $y = q''$). In this case, y remains fixed $(y = q)$ while $x \rightarrow p$; we then speak of limits and continuity in one variable x, as opposed to those in both variables jointly, i.e., the ordinary limits (cf. §3, part IV).

Some other kinds of limits are to be defined below. For simplicity, we consider only functions $f:(X \times Y) \rightarrow\left(T, \rho^{\prime}\right)$ defined on all of $X \times Y$. If confusion is unlikely, we write ρ for all metrics involved (such as $\rho^{\prime}$ in T). Below, p and q always denote cluster points of X and Y, respectively (this justifies the "lim" notation. Of course, our definitions apply in particular to E^{2} as the simplest special case of $X \times Y$.

Definition

A function $f:(X \times Y) \rightarrow\left(T, \rho^{\prime}\right)$ is said to have the double limit $s \in T$ at (p, q), denoted

$$s=\lim _{\substack{x \rightarrow p \\ y \rightarrow q}} f(x, y), \tag{4.11.7}$$

iff for each $\varepsilon>0,$ there is a $\delta>0$ such that $f(x, y) \in G_{s}(\varepsilon)$ whenever $x \in G_{\neg p}(\delta)$ and $y \in G_{\neg q}(\delta)$. In symbols,

$$(\forall \varepsilon>0)(\exists \delta>0)\left(\forall x \in G_{\neg p}(\delta)\right)\left(\forall y \in G_{-q}(\delta)\right) \quad f(x, y) \in G_{s}(\varepsilon). \tag{4.11.8}$$

Observe that this is the relative limit over the path

$$D=(X-\{p\}) \times(Y-\{q\}) \tag{4.11.9}$$

excluding the two "lines" $x=p$ and $y=q$. If f were restricted to D, this would coincide with the ordinary nonrelative limit (see §1), denoted

$$s=\lim _{(x, y) \rightarrow(p, q)} f(x, y), \tag{4.11.10}$$

where only the point (p, q) is excluded. Then we would have

$$(\forall \varepsilon>0)(\exists \delta>0)\left(\forall(x, y) \in G_{\neg(p, q)}(\delta)\right) \quad f(x, y) \in G_{s}(\varepsilon). \tag{4.11.11}$$

Now consider limits in one variable, say,

$$\lim _{y \rightarrow q} f(x, y) \text { with } x \text { fixed. } \tag{4.11.12}$$

If this limit exists for each choice of x from some set $B \subseteq X,$ it defines a function

$$g: B \rightarrow T \tag{4.11.13}$$

with value

$$g(x)=\lim _{y \rightarrow q} f(x, y), \quad x \in B. \tag{4.11.14}$$

This means that

$$(\forall x \in B)(\forall \varepsilon>0)(\exists \delta>0)\left(\forall y \in G_{\neg q}(\delta)\right) \quad \rho(g(x), f(x, y))<\varepsilon. \tag{4.11.15}$$

Here, in general, δ depends on both ε and x. However, in some cases (resembling uniform continuity), one and the same δ (depending on ε only) fits all choices of x from B. This suggests the following definition.

Definition

With the previous notation, suppose

$$\lim _{y \rightarrow q} f(x, y)=g(x) \text { exists for each } x \in B(B \subseteq X). \tag{4.11.16}$$

We say that this limit is uniform in x (on B), and we write

$$" g(x)=\lim _{y \rightarrow q} f(x, y) (\text {uniformly for } x \in B)," \tag{4.11.17}$$

iff for each $\varepsilon>0,$ there is a $\delta>0$ such that $\rho(g(x), f(x, y))<\varepsilon$ for all $x \in B$ and all $y \in G_{\neg q}(\delta)$. In symbols,

$$(\forall \varepsilon>0)(\exists \delta>0)(\forall x \in B)\left(\forall y \in G_{\neg q}(\delta)\right) \quad \rho(g(x), f(x, y))<\varepsilon. \tag{4.11.18}$$

Usually, the set B in formulas (4) and (5) is a deleted neighborhood of p in X, e.g.,

$$B=G_{\neg p}(r), \text { or } B=X-\{p\}. \tag{4.11.19}$$

Assume (4) for such a B, so

$$\lim _{y \rightarrow q} f(x, y)=g(x) \text { exists for each } x \in B. \tag{4.11.20}$$

If, in addition,

$$\lim _{x \rightarrow p} g(x)=s \tag{4.11.21}$$

exists, we call s the iterated limit of f at (p, q) (first in y, then in x), denoted

$$\lim _{x \rightarrow p} \lim _{y \rightarrow q} f(x, y). \tag{4.11.22}$$

This limit is obtained by first letting $y \rightarrow q$ (with x fixed) and then letting $x \rightarrow p$. Quite similarly, we define

$$\lim _{y \rightarrow q} \lim _{x \rightarrow p} f(x, y). \tag{4.11.23}$$

In general, the two iterated limits (if they exist) are different, and their existence does not imply that of the double limit (2), let alone (3), nor does it imply the equality of all these limits. (See Problems 4ff below.) However, we have the following theorem.

Theorem 4.11.2

(Osgood). Let $\left(T, \rho^{\prime}\right)$ be complete. Assume the existence of the following limits of the function $f: X \times Y \rightarrow T$:
(i) $\lim _{y \rightarrow q} f(x, y)=g(x)$ (uniformly for $x \in X-\{p\}$) and
(ii) $\lim _{x \rightarrow p} f(x, y)=h(y)$ for $y \in Y-\{q\}$.
Then the double limit and the two iterated limits of f at (p, q) exist and all three coincide.

Proof

Let $\varepsilon>0.$ By our assumption (i), there is a $\delta>0$ such that

$$(\forall x \in X-\{p\})\left(\forall y \in G_{\neg q}(\delta)\right) \quad \rho(g(x), f(x, y))<\frac{\varepsilon}{4} \quad(\text{cf. } (5)). \tag{4.11.24}$$

Now take any $y^{\prime}, y^{\prime \prime} \in G_{\neg q}(\delta).$ By assumption (ii), there is an $x^{\prime} \in X-\{p\}$ so close to p that

$$\rho\left(h\left(y^{\prime}\right), f\left(x^{\prime}, y^{\prime}\right)\right)<\frac{\varepsilon}{4} \text { and } \rho\left(h\left(y^{\prime \prime}\right), f\left(x^{\prime}, y^{\prime \prime}\right)\right)<\frac{\varepsilon}{4} \cdot(\text{Why?}) \tag{4.11.25}$$

Hence, using $(5')$ and the triangle law (repeatedly), we obtain for such $y^{\prime}, y^{\prime \prime}$

$$\begin{aligned}
\rho\left(h\left(y^{\prime}\right), h\left(y^{\prime \prime}\right)\right) \leq & \rho\left(h\left(y^{\prime}\right), f\left(x^{\prime}, y^{\prime}\right)\right)+\rho\left(f\left(x^{\prime}, y^{\prime}\right), g\left(x^{\prime}\right)\right) \\
& +\rho\left(g\left(x^{\prime}\right), f\left(x^{\prime}, y^{\prime \prime}\right)\right)+\rho\left(f\left(x^{\prime}, y^{\prime \prime}\right), h\left(y^{\prime \prime}\right)\right) \\
< & \frac{\varepsilon}{4}+\frac{\varepsilon}{4}+\frac{\varepsilon}{4}+\frac{\varepsilon}{4}=\varepsilon
\end{aligned}$$

It follows that the function h satisfies the Cauchy criterion of Theorem 2 in §2. (It does apply since T is complete.) Thus

$\lim_{y\to q} h(y)$ exists, and, by assumption (ii), it equals $\lim_{y\to q}\lim_{x\to p} f(x,y)$ (which therefore exists). Let then $H=\lim_{y\to q} h(y)$. With δ as above, fix some $y_0\in G_{\neg q}(\delta)$ so close to q that

$$\rho\left(h\left(y_0\right), H\right)<\frac{\varepsilon}{4}. \tag{4.11.26}$$

Also, using assumption (ii), choose a $\delta'>0\ (\delta'\leq\delta)$ such that

$$\rho\left(h\left(y_0\right), f\left(x, y_0\right)\right)<\frac{\varepsilon}{4} \quad \text{for } x\in G_{\neg p}\left(\delta'\right). \tag{4.11.27}$$

Combining with $(5')$, obtain $(\forall x\in G_{-p}(\delta'))$

$$\rho(H,g(x))\leq\rho\left(H,h\left(y_0\right)\right)+\rho\left(h\left(y_0\right), f\left(x,y_0\right)\right)+\rho\left(f\left(x,y_0\right), g(x)\right)<\frac{3\varepsilon}{4}. \tag{4.11.28}$$

Thus

$$\left(\forall x\in G_{\neg p}\left(\delta'\right)\right) \quad \rho(H,g(x))<\varepsilon. \tag{4.11.29}$$

Hence $\lim_{x\to p} g(x)=H$, i.e., the second iterated limit, $\lim_{x\to p}\lim_{y\to q} f(x,y)$, likewise exists and equals H. Finally, with the same $\delta'\leq\delta$, we combine (6) and $(5')$ to obtain

$$\left(\forall x\in G_{\neg p}\left(\delta'\right)\right)\left(\forall y\in G_{\neg q}\left(\delta'\right)\right) \rho(H,f(x,y))\leq\rho(H,g(x))+\rho(g(x),f(x,y))<\tfrac{3\varepsilon}{4}+\tfrac{\varepsilon}{4}=\varepsilon \tag{4.11.30}$$

Hence the double limit (2) also exists and equals H. $\square$

Note 2. The same proof works also with f restricted to$(X-\{p\})\times(Y-\{q\})$ so that the "lines" $x=p$ and $y=q$ are excluded from D_f. In this case,
formulas (2) and (3) mean the same; i.e.,

$$\lim_{\substack{x\to p\\ y\to q}} f(x,y)=\lim_{(x,y)\to(p,q)} f(x,y). \tag{4.11.31}$$

Note 3. In Theorem 2, we may take E^* (suitably metrized) for X or Y or T. Then the theorem also applies to limits at $\pm\infty$, and infinite limits. We may also take $X=Y=N\cup\{+\infty\}$ (the naturals together with $+\infty$), with the same E^*-metric, and consider limits at $p=+\infty$. Moreover, by Note 2, we may restrict f to $N\times N$, so that $f:N\times N\to T$ becomes a double sequence (Chapter 1, §9). Writing m and n for x and y, and u_{mn} for $f(x,y)$, we then obtain Osgood's theorem for double sequences (also called the Moore-Smith theorem) as follows.

Theorem 4.11.2$'$

Let $\{u_{mn}\}$ be a double sequence in a complete space (T,ρ'). If

$$\lim_{n\to\infty} u_{mn}=q_m \text{ exists for each } m \tag{4.11.32}$$

and if

$$\lim_{m\to\infty} u_{mn}=p_n (\text{uniformly in } n) \text{ likewise exists,} \tag{4.11.33}$$

then the double limit and the two iterated limits of $\{u_{mn}\}$ exist and

$$\lim_{\substack{m\to\infty\\ n\to\infty}} u_{mn}=\lim_{n\to\infty}\lim_{m\to\infty} u_{mn}=\lim_{m\to\infty}\lim_{n\to\infty} u_{mn}. \tag{4.11.34}$$

Here the assumption that $\lim_{m\to\infty} u_{mn} = p_n$ (uniformly in n) means, by (5), that

$$(\forall \varepsilon > 0)(\exists k)(\forall n)(\forall m > k) \quad \rho\left(u_{mn}, p_n\right) < \varepsilon. \tag{4.11.35}$$

Similarly, the statement " $\lim_{\substack{m\to\infty \\ n\to\infty}} u_{mn} = s''(\text{see}(2))$ is tantamount to

$$(\forall \varepsilon > 0)(\exists k)(\forall m, n > k) \quad \rho\left(u_{mn}, s\right) < \varepsilon. \tag{4.11.36}$$

Note 4. Given any sequence $\{x_m\} \subseteq (S, \rho)$, we may consider the double limit $\lim_{\substack{m\to\infty \\ n\to\infty}} \rho\left(x_m, x_n\right)$ in E^1. By using (8), one easily sees that

$$\lim_{\substack{m\to\infty \\ n\to\infty}} \rho\left(x_m, x_n\right) = 0 \tag{4.11.37}$$

iff

$$(\forall \varepsilon > 0)(\exists k)(\forall m, n > k) \quad \rho\left(x_m, x_n\right) < \varepsilon, \tag{4.11.38}$$

i.e., $iff\ \{x_m\}$is a Cauchy sequence. Thus Cauchy sequences are those for which $\lim_{\substack{n\to\infty \\ n\to\infty}} \rho\left(x_m, x_n\right) = 0$.

Theorem 4.11.3

In every metric space (S, ρ), the metric $\rho : (S \times S) \rightarrow E^1$ is a continuous function on the product space $S \times S$.

Proof

Fix any $(p, q) \in S \times S$. By Theorem 1 of §2, \rho\) is continuous at (p, q) iff

$$\rho\left(x_m, y_m\right) \rightarrow \rho(p, q) \text{ whenever } \left(x_m, y_m\right) \rightarrow (p, q), \tag{4.11.39}$$

i.e., whenever $x_m \rightarrow p$ and $y_m \rightarrow q$. However, this follows by Theorem 4 in Chapter 3, §15. Thus continuity is proved. □

LibreTexts

4.11.E: Problems on Double Limits and Product Spaces

Exercise $4.11.E.1$

Prove Theorem $1(\mathrm{i})$. Prove Theorem $1(\text{ ii })$ for both choices of ρ, as suggested.

Exercise $4.11.E.2$

Formulate Definitions 2 and 3 for the cases
(i) $p=q=s=+\infty$;
(ii) $p=+\infty, q \in E^{1}, s=-\infty$;
(iii) $p \in E^{1}, q=s=-\infty;$ and
(iv) $p=q=s=-\infty$.

Exercise $4.11.E.3$

Prove Theorem $2^{\prime}$ from Theorem 2 using Theorem 1 of §2. Give a direct proof as well.

Exercise $4.11.E.4$

Define $f: E^{2} \rightarrow E^{1}$ by

$$f(x, y)=\frac{x y}{x^{2}+y^{2}} \text { if }(x, y) \neq(0,0), \text { and } f(0,0)=0; \tag{4.11.E.1}$$

see §1, Example (g). Show that

$$\lim _{y \rightarrow 0} \lim _{x \rightarrow 0} f(x, y)=0=\lim _{x \rightarrow 0} \lim _{y \rightarrow 0} f(x, y), \tag{4.11.E.2}$$

but

$$\lim _{\substack{x \rightarrow 0 \\ y \rightarrow 0}} f(x, y) \text { does not exist. } \tag{4.11.E.3}$$

Explain the apparent failure of Theorem 2.

Exercise $4.11.E.4^{\prime}$

Define $f: E^{2} \rightarrow E^{1}$ by

$$f(x, y)=0 \text { if } x y=0 \text { and } f(x, y)=1 \text { otherwise. } \tag{4.11.E.4}$$

Show that f satisfies Theorem 2 at $(p, q)=(0,0)$, but

$$\lim _{(x, y) \rightarrow(p, q)} f(x, y) \tag{4.11.E.5}$$

does not exist.

Exercise $4.11.E.5$

Do Problem 4, with f defined as in Problems 9 and 10 of §3.

Exercise $4.11.E.6$

Define f as in Problem 11 of §3. Show that for (c), we have

$$\lim_{(x,y)\to(0,0)} f(x,y)=\lim_{\substack{x\to 0\\ y\to 0}} f(x,y)=\lim_{x\to 0}\lim_{y\to 0} f(x,y)=0, \tag{4.11.E.6}$$

but $\lim_{y\to 0}\lim_{x\to 0} f(x,y)$ does not exist; for (d),

$$\lim_{y\to 0}\lim_{x\to 0} f(x,y)=0, \tag{4.11.E.7}$$

but the iterated limits do not exist; and for (e), $\lim_{(x,y)\to(0,0)} f(x,y)$ fails to exist, but

$$\lim_{\substack{x\to 0\\ y\to 0}} f(x,y)=\lim_{y\to 0}\lim_{x\to 0} f(x,y)=\lim_{x\to 0}\lim_{y\to 0} f(x,y)=0. \tag{4.11.E.8}$$

Give your comments.

Exercise $4.11.E.7$

Find (if possible) the ordinary, the double, and the iterated limits of f at $(0,0)$ assuming that $f(x,y)$ is given by one of the expressions below, and f is defined at those points of E^2 where the expression has sense.

$$\begin{array}{ll} \text{(i) } \frac{x^2}{x^2+y^2}; & \text{(ii) } \frac{y\sin xy}{x^2+y^2} \\ \text{(iii) } \frac{x+2y}{x-y}; & \text{(iv) } \frac{x^3y}{x^6+y^2} \\ \text{(v) } \frac{x^2-y^2}{x^2+y^2}; & \text{(vi) } \frac{x^5+y^4}{(x^2+y^2)^2} \\ \text{(vii) } \frac{y+x\cdot 2^{-y^2}}{4+x^2}; & \text{(viii) } \frac{\sin xy}{\sin x\cdot\sin y} \end{array} \tag{4.11.E.9}$$

Exercise $4.11.E.8$

Solve Problem 7 with x and y tending to $+\infty$.

Exercise $4.11.E.9$

Consider the sequence u_{mn} in E^1 defined by

$$u_{mn}=\frac{m+2n}{m+n}. \tag{4.11.E.10}$$

Show that

$$\lim_{m\to\infty}\lim_{n\to\infty} u_{mn}=2 \text{ and } \lim_{n\to\infty}\lim_{m\to\infty} u_{mn}=1, \tag{4.11.E.11}$$

but the double limit fails to exist. What is wrong here? (See Theo rem $2'$.)

Exercise $4.11.E.10$

Prove Theorem 2, with (i) replaced by the weaker assumption ("subuni-form limit")

$$(\forall \varepsilon>0)(\exists \delta>0)\left(\forall x \in G_{\neg p}(\delta)\right)\left(\forall y \in G_{\neg q}(\delta)\right) \quad \rho(g(x), f(x, y))<\varepsilon \tag{4.11.E.12}$$

and with iterated limits defined by

$$s=\lim _{x \rightarrow p} \lim _{y \rightarrow q} f(x, y) \tag{4.11.E.13}$$

iff $(\forall \varepsilon>0)$

$$(\exists \delta^{\prime}>0)\left(\forall x \in G_{\neg p}\left(\delta^{\prime}\right)\right)\left(\exists \delta_{x}^{\prime \prime}>0\right)\left(\forall y \in G_{\neg q}\left(\delta_{x}^{\prime \prime}\right)\right) \quad \rho(f(x, y), s)<\varepsilon . \tag{4.11.E.14}$$

Exercise $4.11.E.11$

Does the continuity of f on $X \times Y$ imply the existence of (i) iterated limits? (ii) the double limit?
[Hint: See Problem 6.]

Exercise $4.11.E.12$

Show that the standard metric in E^{1} is equivalent to $\rho^{\prime}$ of Problem 7 in Chapter 3, §11.

Exercise $4.11.E.13$

Define products of n spaces and prove Theorem 1 for such product spaces.

Exercise $4.11.E.14$

Show that the standard metric in E^{n} is equivalent to the product metric for E^{n} treated as a product of n spaces E^{1}. Solve a similar problem for C^{n}.
[Hint: Use Problem 13.]

Exercise $4.11.E.15$

Prove that $\left\{\left(x_{m}, y_{m}\right)\right\}$ is a Cauchy sequence in $X \times Y$ iff $\left\{x_{m}\right\}$ and $\left\{y_{m}\right\}$ are Cauchy. Deduce that $X \times Y$ is complete iff X and Y are.

Exercise $4.11.E.16$

Prove that $X \times Y$ is compact iff X and Y are.
[Hint: See the proof of Theorem 2 in Chapter 3, §16, for E^{2}.]

Exercise $4.11.E.17$

(i) Prove the uniform continuity of projection maps P_{1} and P_{2} on $X \times Y$, given by $P_{1}(x, y)=x$ and $P_{2}(x, y)=y$.
(ii) Show that for each open set G in $X \times Y, P_{1}[G]$ is open in X and $P_{2}[G]$ is open in Y.
[Hint: Use Corollary 1 of Chapter $3, \{12.]$
(iii) Disprove (ii) for closed sets by a counterexample.
[Hint: Let $X \times Y=E^{2}$. Let G be the hyperbola $x y=1$. Use Theorem 4 of Chapter 3, §16 to prove that G is closed.]

Exercise $4.11.E.18$

Prove that if $X \times Y$ is connected, so are X and Y.
[Hint: Use Theorem 3 of §10 and the projection maps P_1 and P_2 of Problem 17.]

Exercise $4.11.E.19$

Prove that if X and Y are connected, so is $X \times Y$ under the product metric.
[Hint: Using suitable continuous maps and Theorem 3 in §10, show that any two "lines" $x=p$ and $y=q$ are connected sets in $X \times Y$. Then use Lemma 1 and Problem 10 in §10.]

Exercise $4.11.E.20$

Prove Theorem 2 under the weaker assumptions stated in footnote 1.

Exercise $4.11.E.21$

Prove the following:
(i) If

$$g(x)=\lim_{y \rightarrow q} f(x, y) \text{ and } H=\lim_{\substack{x \rightarrow p \\ y \rightarrow q}} f(x, y) \tag{4.11.E.15}$$

exist for $x \in G_{\neg p}(r)$ and $y \in G_{\neg q}(r)$, then

$$\lim_{x \rightarrow p} \lim_{y \rightarrow q} f(x, y)=H. \tag{4.11.E.16}$$

(ii) If the double limit and one iterated limit exist, they are necessarily equal.

Exercise $4.11.E.22$

In Theorem 2, add the assumptions

$$h(y)=f(p, y) \quad \text{ for } y \in Y-\{q\} \tag{4.11.E.17}$$

and

$$g(x)=f(x, q) \quad \text{ for } x \in X-\{p\}. \tag{4.11.E.18}$$

Then show that

$$\lim_{(x, y) \rightarrow(p, q)} f(x, y) \tag{4.11.E.19}$$

exists and equals the double limits.
[Hint: Show that here (5) holds also for $x=p$ and $y \in G_{\neg q}(\delta)$ and for $y=q$ and $x \in G_{\neg p}(\delta).]$

Exercise $4.11.E.23$

From Problem 22 prove that a function $f:(X \times Y) \rightarrow T$ is continuous at (p, q) if

$$f(p, y)=\lim_{x \rightarrow p} f(x, y) \text{ and } f(x, q)=\lim_{y \rightarrow q} f(x, y) \tag{4.11.E.20}$$

for (x, y) in some $G_{(p,q)}(\delta)$, and at least one of these limits is uniform.

4.12: Sequences and Series of Functions

This page is a draft and is under active development.

I. Let

$$f_1, f_2, \ldots, f_m, \ldots \tag{4.12.1}$$

be a sequence of mappings from a common domain A into a metric space (T, ρ'). For each (fixed) $x \in A$, the function values

$$f_1(x), f_2(x), \ldots, f_m(x), \ldots \tag{4.12.2}$$

form a sequence of points in the range space (T, ρ'). Suppose this sequence converges for each x in a set $B \subseteq A$. Then we can define a function $f : B \rightarrow T$ by setting

$$f(x) = \lim_{m \rightarrow \infty} f_m(x) \text{ for all } x \in B. \tag{4.12.3}$$

This means that

$$(\forall \varepsilon > 0)(\forall x \in B)(\exists k)(\forall m > k) \quad \rho'(f_m(x), f(x)) < \varepsilon. \tag{4.12.4}$$

Here k depends not only on ε but also on x, since each x yields a different sequence $\{f_m(x)\}$. However, in some cases (resembling uniform continuity), k depends on ε only; i.e., given $\varepsilon > 0$, one and the same k fits all x in B. In symbols, this is indicated by changing the order of quantifiers, namely,

$$(\forall \varepsilon > 0)(\exists k)(\forall x \in B)(\forall m > k) \quad \rho'(f_m(x), f(x)) < \varepsilon. \tag{4.12.5}$$

Of course, **(2) implies (1)**, but the converse fails (see examples below). This suggests the following definitions.

Definition 1

With the above notation, we call f the pointwise limit of a sequence of functions f_m on a set $B(B \subseteq A)$ iff

$$f(x) = \lim_{m \rightarrow \infty} f_m(x) \text{ for all } x \text{ in } B; \tag{4.12.6}$$

i.e., **formula (1)** holds. We then write

$$f_m \rightarrow f(\text{pointwise}) \text{ on } B. \tag{4.12.7}$$

In case (2), we call the limit uniform (on B) and write

$$f_m \rightarrow f(\text{uniformly}) \text{ on } B. \tag{4.12.8}$$

II. If the f_m are real, complex, or vector valued (§3), we can also define $s_m = \sum_{k=1}^{m} f_k$ (= sum of the first m functions) for each m, so

$$(\forall x \in A)(\forall m) \quad s_m(x) = \sum_{k=1}^{m} f_k(x). \tag{4.12.9}$$

The s_m form a new sequence of functions on A. The pair of sequences

$$(\{f_m\}, \{s_m\}) \tag{4.12.10}$$

is called the (infinite) series with general term f_m; s_m is called its m th partial sum. The series is often denoted by symbols like $\sum f_m, \sum f_m(x)$, etc.

Definition 2

The series $\sum f_m$ on A is said to converge (pointwise or uniformly) to a function f on a set $B \subseteq A$ iff the sequence $\{s_m\}$ of its partial sums does as well.

We then call f the sum of the series and write

$$f(x)=\sum_{k=1}^{\infty} f_k(x) \text{ or } f=\sum_{m=1}^{\infty} f_m=\lim s_m \tag{4.12.11}$$

(pointwise or uniformly) on B.

Note that series of constants, $\sum c_m$, may be treated as series of constant functions f_m, with $f_m(x)=c_m$ for $x \in A$.

If the range space is E^1 or E^*, we also consider infinite limits,

$$\lim_{m \rightarrow \infty} f_m(x)=\pm\infty. \tag{4.12.12}$$

However, a series for which

$$\sum_{m=1}^{\infty} f_m=\lim s_m \tag{4.12.13}$$

is infinite for some x is regarded as divergent (i.e., not convergent) at that x.

III. Since convergence of series reduces to that of sequences $\{s_m\}$, we shall first of all consider sequences. The following is a simple and useful test for uniform convergence of sequences $f_m : A \rightarrow (T, \rho')$.

Theorem 4.12.1

Given a sequence of functions $f_m : A \rightarrow (T, \rho')$, let $B \subseteq A$ and

$$Q_m=\sup_{x \in B} \rho'\left(f_m(x), f(x)\right). \tag{4.12.14}$$

Then $f_m \rightarrow f(\text{uniformly on } B)$ iff $Q_m \rightarrow 0$.

Proof

If $Q_m \rightarrow 0,$ then by definition

$$(\forall \varepsilon>0)(\exists k)(\forall m>k) \quad Q_m<\varepsilon. \tag{4.12.15}$$

However, Q_m is an upper bound of all distances $\rho'\left(f_m(x), f(x)\right), x \in B$. **Hence (2)** follows.

Conversely, if

$$(\forall x \in B) \quad \rho'\left(f_m(x), f(x)\right)<\varepsilon, \tag{4.12.16}$$

then

$$\varepsilon \geq \sup_{x \in B} \rho'\left(f_m(x), f(x)\right), \tag{4.12.17}$$

i.e., $Q_m \leq \varepsilon$. **Thus (2)** implies

$$(\forall \varepsilon>0)(\exists k)(\forall m>k) \quad Q_m \leq \varepsilon \tag{4.12.18}$$

and $Q_m \rightarrow 0$. $\square$

Examples

(a) We have

$$\lim_{n \rightarrow \infty} x^n=0 \text{ if } |x|<1 \text{ and } \lim_{n \rightarrow \infty} x^n=1 \text{ if } x=1. \tag{4.12.19}$$

Thus, setting $f_n(x)=x^n$, consider $B=[0,1]$ and $C=[0,1)$.

We have $f_n \rightarrow 0$ (pointwise) on C and $f_n \rightarrow f(\text{ pointwise })$ on B, with $f(x)=0$ for $x \in C$ and $f(1)=1$. However, the limit is not uniform on C, let alone on B. Indeed,

$$Q_n = \sup_{x \in C} |f_n(x) - f(x)| = 1 \text{ for each } n. \tag{4.12.20}$$

Thus Q_n does not tend to 0, and uniform convergence fails by Theorem 1.

(b) In Example (a), let $D = [0, a], 0 < a < 1$. Then $f_n \to f$ (uniformly) on D because, in this case,

$$Q_n = \sup_{x \in D} |f_n(x) - f(x)| = \sup_{x \in D} |x^n - 0| = a^n \to 0. \tag{4.12.21}$$

(c) Let

$$f_n(x) = x^2 + \frac{\sin nx}{n}, \quad x \in E^1. \tag{4.12.22}$$

For a fixed x,

$$\lim_{n \to \infty} f_n(x) = x^2 \quad \text{since} \left|\frac{\sin nx}{n}\right| \leq \frac{1}{n} \to 0. \tag{4.12.23}$$

Thus, setting $f(x) = x^2$, we have $f_n \to f$ (pointwise) on E^1. Also,

$$|f_n(x) - f(x)| = \left|\frac{\sin nx}{n}\right| \leq \frac{1}{n}. \tag{4.12.24}$$

Thus $(\forall n) Q_n \leq \frac{1}{n} \to 0$. By Theorem 1, the limit is uniform on all of E^1.

Theorem 4.12.2

Let $f_m : A \to (T, \rho')$ be a sequence of functions on $A \subseteq (S, \rho)$. If $f_m \to f$ (uniformly on a set $B \subseteq A$, and if the f_m are relatively (or uniformly) continuous on B, then the limit function f has the same property.

Proof

Fix $\varepsilon > 0$. As $f_m \to f$ (uniformly) on B, there is a k such that

$$(\forall x \in B)(\forall m \geq k) \quad \rho'(f_m(x), f(x)) < \frac{\varepsilon}{4}. \tag{4.12.25}$$

Take any f_m with $m > k$, and take any $p \in B$. By continuity, there is $\delta > 0$, with

$$(\forall x \in B \cap G_p(\delta)) \quad \rho'(f_m(x), f_m(p)) < \frac{\varepsilon}{4}. \tag{4.12.26}$$

Also, setting $x = p$ in (3) gives $\rho'(f_m(p), f(p)) < \frac{\varepsilon}{4}$. Combining this with (4) and (3), we obtain $(\forall x \in B \cap G_p(\delta))$

$$\begin{aligned} \rho'(f(x), f(p)) & \leq \rho'(f(x), f_m(x)) + \rho'(f_m(x), f_m(p)) + \rho'(f_m(p), f(p)) \\ & < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon. \end{aligned}$$

We thus see that for $p \in B$,

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in B \cap G_p(\delta)) \quad \rho'(f(x), f(p)) < \varepsilon, \tag{4.12.27}$$

i.e., f is relatively continuous at $p($ over $B)$, as claimed.

Quite similarly, the reader will show that f is uniformly continuous if the f_n are. □

Note 2. A similar proof also shows that if $f_m \to f$ (uniformly) on B, and if the f_m are relatively continuous at a point $p \in B$, so also is f.

Theorem 4.12.3 (Cauchy criterion for uniform convergence)

Let (T, ρ') be complete. Then a sequence $f_m : A \rightarrow T, A \subseteq (S, \rho),$ converges uniformly on a set $B \subseteq A$ iff

$$(\forall \varepsilon > 0)(\exists k)(\forall x \in B)(\forall m, n > k) \quad \rho'\left(f_m(x), f_n(x)\right) < \varepsilon. \tag{4.12.28}$$

Proof

If (5) holds then, for any (fixed) $x \in B$, $\{f_m(x)\}$ is a Cauchy sequence of points in T, so by the assumed completeness of T, it has a limit $f(x)$. Thus we can define a function $f : B \rightarrow T$ with

$$f(x) = \lim_{m \rightarrow \infty} f_m(x) \text{ on } B. \tag{4.12.29}$$

To show that $f_m \rightarrow f$ (uniformly) on B, we **use (5) again**. Keeping $\varepsilon, k,\ x$, and m temporarily fixed, we let $n \rightarrow \infty$ so that $f_n(x) \rightarrow f(x)$. Then by Theorem 4 of Chapter 3, §15, $\rho'\left(f_m(x), f_n(x)\right) \rightarrow p'\left(f(x), f_m(x)\right)$. Passing to the l**imit in (5), we thus obtain (2)**.

The easy proof of the converse is left to the reader (cf. Chapter 3, §17, Theorem 1). □

IV. If the range space (T, ρ') is E^1, C, or E^n (*or another normed space), the standard metric applies. In particular, for series we have

$$\begin{aligned} \rho'\left(s_m(x), s_n(x)\right) &= \left|s_n(x) - s_m(x)\right| \\ &= \left|\sum_{k=1}^{n} f_k(x) - \sum_{k=1}^{m} f_k(x)\right| \\ &= \left|\sum_{k=m+1}^{n} f_k(x)\right| \quad \text{ for } m < n. \end{aligned}$$

Replacing here m by $m-1$ and applying Theorem 3 to the sequence $\{s_m\}$, we obtain the following result.

Theorem 4.12.3′

Let the range space of $f_m, m = 1, 2, \ldots,$ be E^1, C, or E^n (*or another complete normed space). Then the series $\sum f_m$ converges uniformly on B iff

$$(\forall \varepsilon > 0)(\exists q)(\forall n > m > q)(\forall x \in B) \quad \left|\sum_{k=m}^{n} f_k(x)\right| < \varepsilon. \tag{4.12.30}$$

Similarly, via $\{s_m\}$, Theorem 2 extends to series of functions. (Observe that the s_m are continuous if the f_m are.) Formulate it!

V. If $\sum_{m=1}^{\infty} f_m$ exists on B, one may arbitrarily "group" the terms, i.e., replace every several consecutive terms by their sum. This property is stated more precisely in the following theorem.

Theorem 4.12.4

Let

$$f = \sum_{m=1}^{\infty} f_m (\text{pointwise}) \text{ on } B. \tag{4.12.31}$$

Let $m_1 < m_2 < \cdots < m_n < \cdots$ in N, and define

$$g_1 = s_{m_1}, \quad g_n = s_{m_n} - s_{m_{n-1}}, \quad n > 1. \tag{4.12.32}$$

(Thus $g_{n+1} = f_{m_n+1} + \cdots + f_{m_{n+1}}.$) Then

$$f=\sum_{n=1}^{\infty} g_n (\text{pointwise}) \text{ on } B \text{ as well;} \tag{4.12.33}$$

similarly for uniform convergence.

Proof

Let

$$s_n^{\prime}=\sum_{k=1}^{n} g_k, \quad n=1,2,\ldots \tag{4.12.34}$$

Then $s_n^{\prime}=s_{m_n}$ (verify!), so $\{s_n^{\prime}\}$ is a subsequence, $\{s_{m_n}\}$, of $\{s_m\}$. Hence $s_m \rightarrow f(\text{ pointwise })$ implies $s_n^{\prime} \rightarrow f$ (pointwise); i.e.,

$$f=\sum_{n=1}^{\infty} g_n \text{ (pointwise).} \tag{4.12.35}$$

For uniform convergence, see Problem 13 (cf. also Problem 19). □

4.12.E: Problems on Sequences and Series of Functions

Exercise $4.12.E.1$

Complete the proof of Theorems 2 and 3.

Exercise $4.12.E.2$

Complete the proof of Theorem 4.

Exercise $4.12.E.2'$

In Example (a), show that $f_n \rightarrow +\infty$ (pointwise) on $(1, +\infty)$, but not uniformly so. Prove, however, that the limit is uniform on any interval $[a, +\infty), a > 1.$ (Define "lim $f_n = +\infty$ (uniformly)" in a suitable manner.)

Exercise $4.12.E.3$

Using Theorem 1, discuss $\lim_{n\to\infty} f_n$ on B and $C($ as in Example $(a))$ for each of the following.
(i) $f_n(x) = \frac{x}{n}; B = E^1; C = [a, b] \subset E^1$.
(ii) $f_n(x) = \frac{\cos x + nx}{n}; B = E^1$.
(iii) $f_n(x) = \sum_{k=1}^{n} x^k; B = (-1, 1); C = [-a, a], |a| < 1$.
(iv) $f_n(x) = \frac{x}{1+nx}; C = [0, +\infty)$.
$[$Hint: Prove that $Q_n = \sup \frac{1}{n}\left(1 - \frac{1}{nx+1}\right) = \frac{1}{n}.]$
(v) $f_n(x) = \cos^n x; B = \left(0, \frac{\pi}{2}\right), C = \left[\frac{1}{4}, \frac{\pi}{2}\right)$.
(vi) $f_n(x) = \frac{\sin^2 nx}{1+nx}; B = E^1$.
(vii) $f_n(x) = \frac{1}{1+x^n}; B = [0, 1); C = [0, a], 0 < a < 1$.

Exercise $4.12.E.4$

Using Theorems 1 and 2, discuss $\lim f_n$ on the sets given below, with $f_n(x)$ as indicated and $0 < a < +\infty$. (Calculus rules for maxima and minima are assumed known in (v), (vi), and (vii).)
(i) $\frac{nx}{1+nx}; [a, +\infty), (0, a)$.
(ii) $\frac{nx}{1+n^3x^3}; (a, +\infty), (0, a)$.
(iii) $\sqrt[n]{\cos x}; \left(0, \frac{\pi}{2}\right), [0, a], a < \frac{\pi}{2}$.
(iv) $\frac{x}{n}; (0, a), (0, +\infty)$.
(v) $xe^{-nx}; [0, +\infty); E^1$.
(vi) $nxe^{-nx}; [a, +\infty), (0, +\infty)$.
(vii) $nxe^{-nx^2}; [a, +\infty), (0, +\infty)$.
[Hint: $\lim f_n$ cannot be uniform if the f_n are continuous on a set, but $\lim f_n$ is not.
[For (v), f_n has a maximum at $x = \frac{1}{n}$; hence find Q_n.]

Exercise $4.12.E.5$

Define $f_n : E^1 \rightarrow E^1$ by

$$f_n(x) = \begin{cases} nx & \text{if } 0 \leq x \leq \frac{1}{n} \\ 2 - nx & \text{if } \frac{1}{n} < x \leq \frac{2}{n}, \text{ and} \\ 0 & \text{otherwise} \end{cases} \tag{4.12.E.1}$$

Show that all f_n and $\lim f_n$ are continuous on each interval $(-a, a)$, though $\lim f_n$ exists only pointwise. (Compare this with Theorem 3.)

Exercise 4.12.E.6

The function f found in the proof of Theorem 3 is uniquely determined. Why?

Exercise 4.12.E.7

$\Rightarrow 7$. Prove that if each of the functions f_n is constant on B, or if B is finite, then a pointwise limit of the f_n on B is also a uniform limit; similarly for series.

Exercise 4.12.E.8

$\Rightarrow 8$. Prove that if $f_n \rightarrow f(\text{ uniformly })$ on B and if $C \subseteq B$, then $f_n \rightarrow f$ (uniformly) on C as well.

Exercise 4.12.E.9

$\Rightarrow 9$. Show that if $f_n \rightarrow f(\text{ uniformly })$ on each of $B_1, B_2, \ldots, B_m$, then $f_n \rightarrow f$ (uniformly) on $\bigcup_{k=1}^m B_k$.
Disprove it for infinite unions by an example. Do the same for series.

Exercise 4.12.E.10

$\Rightarrow 10$. Let $f_n \rightarrow f(\text{ uniformly })$ on B. Prove the equivalence of the following statements:
(i) Each f_n, from a certain n onward, is bounded on B.
(ii) f is bounded on B.
(iii) The f_n are ultimately uniformly bounded on B; that is, all function values $f_n(x), x \in B$, from a certain $n=n_0$ onward, are in one and the same globe $G_q(K)$ in the range space.
For real, complex, and vector-valued functions, this means that

$$\left(\exists K \in E^1\right)\left(\forall n \geq n_0\right)(\forall x \in B) \quad\left|f_n(x)\right|<K. \tag{4.12.E.2}$$

Exercise 4.12.E.11

$\Rightarrow 11$. Prove for real, complex, or vector-valued functions f_n, f, g_n, g that if

$$f_n \rightarrow f \text{ and } g_n \rightarrow g \text{ (uniformly) on } B, \tag{4.12.E.3}$$

then also

$$f_n \pm g_n \rightarrow f \pm g(\text{ uniformly }) \text{ on } B. \tag{4.12.E.4}$$

Exercise 4.12.E.12

$\Rightarrow 12$. Prove that if the functions f_n and g_n are real or complex (or if the g_n are vector valued and the f_n are scalar valued), and if

$$f_n \rightarrow f \text{ and } g_n \rightarrow g \text{ (uniformly) on } B, \tag{4.12.E.5}$$

then

$$f_n g_n \rightarrow f g \text{ (uniformly) on } B \tag{4.12.E.6}$$

provided that either f and g or the f_n and g_n are bounded on B (at least from some n onward); cf. Problem 11.

Disprove it for the case where only one of f and g is bounded.
[Hint: Let $f_n(x) = x$ and $g_n(x) = 1/n$ (constant) on $B = E^1$. Give some other examples.]

Exercise 4.12.E.13

⇒ 13. Prove that if $\{f_n\}$ tends to f (pointwise or uniformly), so does each subsequence $\{f_{n_k}\}$.

Exercise 4.12.E.14

⇒ 14. Let the functions f_n and g_n and the constants a and b be real or complex (or let a and b be scalars and f_n and g_n be vector valued). Prove that if

$$f = \sum_{n=1}^{\infty} f_n \text{ and } g = \sum_{n=1}^{\infty} g_n \text{ (pointwise or uniformly)}, \tag{4.12.E.7}$$

then

$$af + bg = \sum_{n=1}^{\infty} (af_n + bg_n) \text{ in the same sense.} \tag{4.12.E.8}$$

(Infinite limits are excluded.)
In particular,

$$f \pm g = \sum_{n=1}^{\infty} (f_n \pm g_n) \quad \text{(rule of termwise addition)} \tag{4.12.E.9}$$

and

$$af = \sum_{n=1}^{\infty} af_n. \tag{4.12.E.10}$$

[Hint: Use Problems 11 and 12.]

Exercise 4.12.E.15

⇒ 15. Let the range space of the functions f_m and g be E^n (*or C^n), and let $f_m = (f_{m1}, f_{m2}, \ldots, f_{mn})$, $g = (g_1, \ldots, g_n)$; see §3, part II. Prove that

$$f_m \to g \quad \text{(pointwise or uniformly)} \tag{4.12.E.11}$$

iff each component f_{mk} of f_m converges (in the same sense) to the corresponding component g_k of g; i.e.,

$$f_{mk} \to g_k \quad \text{(pointwise or uniformly)}, k = 1, 2, \ldots, n. \tag{4.12.E.12}$$

Similarly,

$$g = \sum_{m=1}^{\infty} f_m \tag{4.12.E.13}$$

iff

$$(\forall k \leq n) \quad g_k = \sum_{m=1}^{\infty} f_{mk}. \tag{4.12.E.14}$$

(See Chapter 3, §15, Theorem 2).

Exercise 4.12.E.16

⇒ 16. From Problem 15 deduce for complex functions that $f_m \rightarrow g$ (pointwise or uniformly) iff the real and imaginary parts of the f_m converge to those of g (pointwise or uniformly). That is, $(f_m)_{re} \rightarrow g_{re}$ and $(f_m)_{im} \rightarrow g_{im}$; similarly for series.

Exercise 4.12.E.17

⇒ 17. Prove that the convergence or divergence (pointwise or uniformly) of a sequence $\{f_m\}$, or a series $\sum f_m$, of functions is not affected by deleting or adding a finite number of terms. Prove also that $\lim_{m\to\infty} f_m$ (if any) remains the same, but $\sum_{m=1}^{\infty} f_m$ is altered by the difference between the added and deleted terms.

Exercise 4.12.E.18

⇒ 18. Show that the geometric series with ratio r,

$$\sum_{n=0}^{\infty} ar^n \quad \left(a, r \in E^1 \text{ or } a, r \in C\right), \tag{4.12.E.15}$$

converges iff $|r| < 1$, in which case

$$\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r} \tag{4.12.E.16}$$

(similarly if a is a vector and r is a scalar). Deduce that $\sum(-1)^n$ diverges. (See Chapter 3, §15, Problem 19.)

Exercise 4.12.E.19

Theorem 4 shows that a convergent series does not change its sum if every several consecutive terms are replaced by their sum. Show by an example that the reverse process (splitting each term into several terms) may affect convergence.
[Hint: Consider $\sum a_n$ with $a_n = 0$. Split $a_n = 1 - 1$ to obtain a divergent series: $\sum(-1)^{n-1}$, with partial sums $1, 0, 1, 0, 1, \ldots$]

Exercise 4.12.E.20

Find $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$.
[Hint: Verify: $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$. Hence find s_n, and let $n \rightarrow \infty$.]

Exercise 4.12.E.21

The functions $f_n : A \rightarrow (T, \rho')$, $A \subseteq (S, \rho)$ are said to be equicontinuous at $p \in A$ iff

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall n)\left(\forall x \in A \cap G_p(\delta)\right) \quad \rho'\left(f_n(x), f_n(p)\right) < \varepsilon. \tag{4.12.E.17}$$

Prove that if so, and if $f_n \rightarrow f$ (pointwise) on A, then f is continuous at p.

[Hint: "Imitate" the proof of Theorem 2 .]

4.13: Absolutely Convergent Series. Power Series

This page is a draft and is under active development.

I. A series $\sum f_m$ is said to be absolutely convergent on a set B iff the series $\sum |f_m(x)|$ (briefly, $\sum |f_m|$) of the absolute values of f_m converges on B (pointwise or uniformly). Notation:

$$f = \sum |f_m| \text{ (pointwise or uniformly) on } B. \tag{4.13.1}$$

In general, $\sum f_m$ may converge while $\sum |f_m|$ does not (see Problem 12). In this case, the convergence of $\sum f_m$ is said to be conditional. (It may be absolute for some x and conditional for others.) As we shall see, absolute convergence ensures the commutative law for series, and it implies ordinary convergence (i.e., that of $\sum f_m$), if the range space of the f_m is complete.

Note 1. Let

$$\sigma_m = \sum_{k=1}^{m} |f_k|. \tag{4.13.2}$$

Then

$$\sigma_{m+1} = \sigma_m + |f_{m+1}| \geq \sigma_m \quad \text{on } B; \tag{4.13.3}$$

i.e., the $\sigma_m(x)$ form a monotone sequence for each $x \in B$. Hence by Theorem 3 of Chapter 3, §15,

$$\lim_{m \to \infty} \sigma_m = \sum_{m=1}^{\infty} |f_m| \text{ always exists in } E^*; \tag{4.13.4}$$

$\sum |f_m|$ converges iff $\sum_{m=1}^{\infty} |f_m| < +\infty$.

For the rest of this section we consider only complete range spaces.

Theorem 4.13.1

Let the range space of the functions f_m (all defined on A) be E^1, C, or E^n (* or another complete normed space). Then for $B \subseteq A$, we have the following:

(i) If $\sum |f_m|$ converges on B (pointwise or uniformly), so does $\sum f_m$ itself. Moreover,

$$\left|\sum_{m=1}^{\infty} f_m\right| \leq \sum_{m=1}^{\infty} |f_m| \quad \text{on } B. \tag{4.13.5}$$

(ii) (Commutative law for absolute convergence.) If $\sum |f_m|$ converges (pointwise or uniformly on B, so does any series $\sum |g_m|$ obtained by rearranging the f_m in any different order. Moreover,

$$\sum_{m=1}^{\infty} f_m = \sum_{m=1}^{\infty} g_m \quad (\text{ both exist on } B)/ \tag{4.13.6}$$

Note 2. More precisely, a sequence $\{g_m\}$ is called a rearrangement of $\{f_m\}$ iff there is a map $u : N \longleftrightarrow_{onto} N$ such that

$$(\forall m \in N) \quad g_m = f_{u(m)}. \tag{4.13.7}$$

Proof

(i) If $\sum |f_m|$ converges uniformly on B, then by Theorem $3'$ of §12,

$$\begin{array}{c}(\forall \varepsilon > 0)(\exists k)(\forall n > m > k)(\forall x \in B) \\ \varepsilon > \sum_{i=m}^{n} |f_i(x)| \geq \left|\sum_{i=m}^{n} f_i(x)\right| \text{ (triangle law)}\end{array}. \tag{4.13.8}$$

However, this shows that $\sum f_n$ satisfies Cauchy's criterion (6) of §12, so it converges uniformly on B.

Moreover, letting $n \to \infty$ in the inequality

$$\left|\sum_{m=1}^{n} f_m\right| \leq \sum_{m=1}^{n} |f_m|, \tag{4.13.9}$$

we get

$$\left|\sum_{m=1}^{\infty} f_m\right| \leq \sum_{m=1}^{\infty} |f_m| < +\infty \quad \text{on } B, \text{ as claimed.} \tag{4.13.10}$$

By Note 1, this also proves the theorem for pointwise convergence.

(ii) Again, if $\sum |f_m|$ converges uniformly on B, the inequalities (1) hold for all f_i except (possibly) for $f_1, f_2, \ldots, f_k$. Now when $\sum f_m$ is rearranged, these k functions will be renumbered as certain g_i. Let q be the largest of their new subscripts i. Then all of them (and possibly some more functions) are among $g_1, g_2, \ldots, g_q$ (so that $q \geq k$). Hence if we exclude $g_1, \ldots, g_q$, the inequalities (1) will certainly hold for the remaining g_i $(i > q)$. Thus

$$(\forall \varepsilon > 0)(\exists q)(\forall n > m > q)(\forall x \in B) \quad \varepsilon > \sum_{i=m}^{n} |g_i| \geq \left|\sum_{i=m}^{n} g_i\right|. \tag{4.13.11}$$

By Cauchy's criterion, then, both $\sum g_i$ and $\sum |g_i|$ converge uniformly.

Moreover, by construction, the two partial sums

$$s_k = \sum_{i=1}^{k} f_i \text{ and } s_q' = \sum_{i=1}^{q} g_i \tag{4.13.12}$$

can differ only in those terms whose original subscripts (before the rearrangement) were $> k$. By (1), however, any finite sum of such terms is less than ε in absolute value. Thus $|s_q' - s_k| < \varepsilon$.

This argument holds also if k in (1) is replaced by a larger integer.
(Then also q increases, since $q \geq k$ as noted above.) Thus we may let $k \rightarrow +\infty($ hence also $q \rightarrow +\infty)$ in the inequality $|s_q' - s_k| < \varepsilon,$ with ε fixed. Then

$$s_k \rightarrow \sum_{m=1}^{\infty} f_m \text{ and } s_q' \rightarrow \sum_{i=1}^{\infty} g_i, \tag{4.13.13}$$

so

$$\left|\sum_{i=1}^{\infty} g_i - \sum_{m=1}^{\infty} f_m\right| \leq \varepsilon. \tag{4.13.14}$$

Now let $\varepsilon \rightarrow 0$ to get

$$\sum_{i=1}^{\infty} g_i = \sum_{m=1}^{\infty} f_m; \tag{4.13.15}$$

similarly for pointwise convergence. □

II. Next, we develop some simple tests for absolute convergence.

Theorem 4.13.2

(comparison test). Suppose

$$(\forall m) \quad |f_m| \leq |g_m| \text{ on } B. \tag{4.13.16}$$

Then

(i) $\sum_{m=1}^{\infty} |f_m| \leq \sum_{m=1}^{\infty} |g_m|$ on B;

(ii) $\sum_{m=1}^{\infty} |f_m| = +\infty$ implies $\sum_{m=1}^{\infty} |g_m| = +\infty$ on B; and

(iii) If $\sum|g_m|$ converges (pointwise or uniformly) on B, so does $\sum|f_m|$.

Proof

Conclusion (i) follows by letting $n \rightarrow \infty$ in

$$\sum_{m=1}^{n}|f_m| \leq \sum_{m=1}^{n}|g_m| . \tag{4.13.17}$$

In turn, (ii) is a direct consequence of (i).

Also, by (i),

$$\sum_{m=1}^{\infty}|g_m|<+\infty \text { implies } \sum_{m=1}^{\infty}|f_m|<+\infty. \tag{4.13.18}$$

This proves (iii) for the pointwise case (see Note $1)$. The uniform case follows exactly as in Theorem $1(\mathrm{i})$ on noting that

$$\sum_{k=m}^{n}|f_k| \leq \sum_{k=m}^{n}|g_k| \tag{4.13.19}$$

and that the functions $|f_k|$ and $|g_k|$ are real (so Theorem $3^{\prime}$ in §12 does apply). □

Theorem 4.13.3 (Weierstrass "M-test")

If $\sum M_n$ is a convergent series of real constants $M_n \geq 0$ and if

$$(\forall n) \quad|f_n| \leq M_n \tag{4.13.20}$$

on a set B, then $\sum|f_n|$ converges uniformly on B. Moreover,

$$\sum_{n=1}^{\infty}|f_n| \leq \sum_{n=1}^{\infty} M_n \quad \text { on } B. \tag{4.13.21}$$

Proof

Use Theorem 2 with $|g_n|=M_n$, noting that $\sum|g_n|$ converges uniformly since the $|g_n|$ are constant (§12, Problem 7). □

ExampleS

(a) Let

$$f_n(x)=\left(\frac{1}{2} \sin x\right)^n \text { on } E^1. \tag{4.13.22}$$

Then

$$(\forall n)\left(\forall x \in E^1\right) \quad|f_n(x)| \leq 2^{-n}, \tag{4.13.23}$$

and $\sum 2^{-n}$ converges (geometric series with ratio $\frac{1}{2}$; see §12, Problem 18). Thus, setting $M_n=2^{-n}$ in Theorem 3, we infer that the series $\sum\left|\frac{1}{2} \sin x\right|^n$ converges uniformly on E^1, as does $\sum\left(\frac{1}{2} \sin x\right)^n$; moreover,

$$\sum_{n=1}^{\infty}|f_n| \leq \sum_{n=1}^{\infty} 2^{-n}=1. \tag{4.13.24}$$

Theorem 4.13.4 (necessary condition of convergence)

If $\sum f_m$ or $\sum |f_m|$ converges on B (pointwise or uniformly), then $|f_m| \to 0$ on B (in the same sense).

Thus a series cannot converge unless its general term tends to 0 (respectively, $\overline{0}$).

Proof

If $\sum f_m = f$, say, then $s_m \to f$ and also $s_{m-1} \to f$. Hence

$$s_m - s_{m-1} \to f - f = \overline{0}. \tag{4.13.25}$$

However, $s_m - s_{m-1} = f_m$. Thus $f_m \to \overline{0}$, and $|f_m| \to 0$, as claimed.

This holds for pointwise and uniform convergence alike (see Problem 14 in §12) . □

Caution: The condition $|f_m| \to 0$ is necessary but not sufficient. Indeed, there are divergent series with general term tending to 0, as we show next.

Examples (Continued)

(b) $\sum_{n=1}^{\infty} \frac{1}{n} = +\infty$ (the so-called harmonic series).

Indeed, by Note 1,

$$\sum_{n=1}^{\infty} \frac{1}{n} \quad \text{exists (in } E^*\text{)}, \tag{4.13.26}$$

so Theorem 4 of §12 applies. We group the series as follows:

$$\begin{aligned} \sum \frac{1}{n} &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \cdots + \frac{1}{16}\right) + \cdots \\ &\geq \frac{1}{2} + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \left(\frac{1}{16} + \cdots + \frac{1}{16}\right) + \cdots. \end{aligned}$$

Each bracketed expression now equals $\frac{1}{2}$. Thus

$$\sum \frac{1}{n} \geq \sum g_m, \quad g_m = \frac{1}{2}. \tag{4.13.27}$$

As g_m does not tend to 0, $\sum g_m$ diverges, i.e., $\sum_{m=1}^{\infty} g_m$ is infinite, by Theorem 4. A fortiori, so is $\sum_{n=1}^{\infty} \frac{1}{n}$.

Theorem 4.13.5 (root and ratio tests)

A series of constants $\sum a_n$ $(|a_n| \neq 0)$ converges absolutely if

$$\overline{\lim} \sqrt[n]{|a_n|} < 1 \text{ or } \overline{\lim} \left(\frac{|a_{n+1}|}{|a_n|}\right) < 1. \tag{4.13.28}$$

It diverges if

$$\overline{\lim} \sqrt[n]{|a_n|} > 1 \text{ or } \underline{\lim} \left(\frac{|a_{n+1}|}{|a_n|}\right) > 1. \tag{4.13.29}$$

It may converge or diverge if

$$\overline{\lim} \sqrt[n]{|a_n|} = 1 \tag{4.13.30}$$

or if

$$\underline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right) \leq 1 \leq \overline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right). \tag{4.13.31}$$

(The a_n may be scalars or vectors.)

Proof

If $\overline{\lim} \sqrt[n]{\left|a_{n}\right|}<1,$ choose $r>0$ such that

$$\overline{\lim} \sqrt[n]{\left|a_{n}\right|}<r<1. \tag{4.13.32}$$

Then by Corollary 2 of Chapter 2, §13, $\sqrt[n]{\left|a_{n}\right|}<r$ for all but finitely many n. Thus, dropping a finite number of terms (§12, Problem 17), we may assume that

$$\left|a_{n}\right|<r^{n} \text{ for all } n. \tag{4.13.33}$$

As $0<r<1,$ the geometric series $\sum r^{n}$ converges. Hence so does $\sum\left|a_{n}\right|$ by Theorem 2.

In the case

$$\overline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right)<1, \tag{4.13.34}$$

we similarly obtain $(\exists m)(\forall n \geq m)\left|a_{n+1}\right|<\left|a_{n}\right| r;$ hence by induction,

$$(\forall n \geq m) \quad\left|a_{n}\right| \leq\left|a_{m}\right| r^{n-m}. \quad \text{(Verify!)} \tag{4.13.35}$$

Thus $\sum\left|a_{n}\right|$ converges, as before.

If $\overline{\lim} \sqrt[n]{\left|a_{n}\right|}>1,$ then by Corollary 2 of Chapter 2, §13,\left|a_{n}\right|>1\) for infinitely many n. Hence $\left|a_{n}\right|$ cannot tend to 0, and so $\sum a_{n}$ diverges by Theorem 4.

Similarly, if

$$\underline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right)>1, \tag{4.13.36}$$

then $\left|a_{n+1}\right|>\left|a_{n}\right|$ for all but finitely many n, so $\left|a_{n}\right|$ cannot tend to 0 again. □

Note 3. We have

$$\underline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right) \leq \underline{\lim} \sqrt[n]{\left|a_{n}\right|} \leq \overline{\lim} \sqrt[n]{\left|a_{n}\right|} \leq \overline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right). \tag{4.13.37}$$

Thus

$$\begin{array}{l}\overline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right)<1 \text{ implies } \overline{\lim} \sqrt[n]{\left|a_{n}\right|}<1; \text{ and } \\ \underline{\lim}\left(\frac{\left|a_{n+1}\right|}{\left|a_{n}\right|}\right)>1 \text{ implies } \overline{\lim} \sqrt[n]{\left|a_{n}\right|}>1.\end{array} \tag{4.13.38}$$

Hence whenever the ratio test indicates convergence or divergence, so certainly does the root test. On the other hand, there are cases where the root test yields a result while the ratio test does not. Thus the root test is stronger (but the ratio test is often easier to apply).

Examples (continued)

(c) Let $a_{n}=2^{-k}$ if $n=2k-1$ (odd) and $a_{n}=3^{-k}$ if $n=2k$ (even). Thus

$$\sum a_{n}=\frac{1}{2^{1}}+\frac{1}{3^{1}}+\frac{1}{2^{2}}+\frac{1}{3^{2}}+\frac{1}{2^{3}}+\frac{1}{3^{3}}+\frac{1}{2^{4}}+\frac{1}{3^{4}}+\cdots. \tag{4.13.39}$$

Here

$$\underline{\lim}\left(\frac{a_{n+1}}{a_n}\right)=\lim_{k\to\infty}\frac{3^{-k}}{2^{-k}}=0 \text{ and } \overline{\lim}\left(\frac{a_{n+1}}{a_n}\right)=\lim_{k\to\infty}\frac{2^{-k-1}}{3^{-k}}=+\infty, \tag{4.13.40}$$

while

$$\overline{\lim}\sqrt[n]{a_n}=\lim\sqrt[2n-1]{2^{-n}}=\frac{1}{\sqrt{2}}<1. \quad \text{(Verify!)} \tag{4.13.41}$$

Thus the ratio test fails, but the root test proves convergence.

Note 4. The assumption $|a_n|\neq 0$ is needed for the ratio test only.

III. Power Series. As an application, we now consider so-called power series,

$$\sum a_n(x-p)^n, \tag{4.13.42}$$

where $x, p, a_n \in E^1(C)$; the a_n may also be vectors.

Theorem 4.13.6

For any power series $\sum a_n(x-p)^n$, there is a unique $r\in E^*$ $(0\leq r\leq +\infty)$, called its convergence radius, such that the series converges absolutely for each x with $|x-p|<r$ and does not converge (even conditionally) if $|x-p|>r$.

Specifically,

$$r=\frac{1}{d}, \text{ where } d=\overline{\lim}\sqrt[n]{|a_n|} \quad (\text{with } r=+\infty \text{ if } d=0). \tag{4.13.43}$$

Proof

Fix any $x=x_0$. By Theorem 5, the series $\sum a_n(x_0-p)^n$ converges absolutely if $\overline{\lim}\sqrt[n]{|a_n|}\,|x_0-p|<1,$ i.e., if

$$|x_0-p|<r \quad \left(r=\frac{1}{\overline{\lim}\sqrt[n]{|a_n|}}=\frac{1}{d}\right), \tag{4.13.44}$$

and diverges if $|x_0-p|>r.$ (Here we assumed $d\neq 0$; but if $d=0,$ the condition $d\,|x_0-p|<1$ is trivial for any$x_0,$ so $r=+\infty$ in this case.) Thus r is the required radius, and clearly there can be only one such r. (Why?) □

Note 5. If $\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}$ exists, it equals $\lim_{n\to\infty}\sqrt[n]{|a_n|}$, by Note 3 (for $\overline{lim}$ and $\underline{lim}$ coincide here). In this case, one can use the ratio test to find

$$d=\lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|} \tag{4.13.45}$$

and hence (if $d\neq 0$)

$$r=\frac{1}{d}=\lim_{n\to\infty}\frac{|a_n|}{|a_{n+1}|}. \tag{4.13.46}$$

Theorem 4.13.7

If a power series $\sum a_n(x-p)^n$ converges absolutely for some $x=x_0\neq p,$ then $\sum|a_n(x-p)^n|$ converges uniformly on the closed globe $\overline{G}_p(\delta)$ $\delta=|x_0-p|$. So does $\sum a_n(x-p)^n$ if the range space is complete (Theorem 1).

Proof

Suppose $\sum|a_n(x_0-p)^n|$ converges. Let

$$\delta=|x_0-p| \text{ and } M_n=|a_n|\,\delta^n; \tag{4.13.47}$$

thus $\sum M_n$ converges.

Now if $x \in \overline{G}_p(\delta),$ then $|x-p| \leq \delta,$ so

$$\left|a_n(x-p)^n\right| \leq \left|a_n\right| \delta^n = M_n. \tag{4.13.48}$$

Hence by Theorem 3, $\sum \left|a_n(x-p)^n\right|$ converges uniformly on $\overline{G}_p(\delta)$. $\square$

Examples (Continued)

(d) Consider $\sum \frac{x^n}{n!}$ Here

$$p = 0 \text{ and } a_n = \frac{1}{n!}, \text{ so } \frac{\left|a_n\right|}{\left|a_{n+1}\right|} = n+1 \rightarrow +\infty. \tag{4.13.49}$$

By Note 5, then, $r = +\infty;$ i.e., the series converges absolutely on all of E^1. Hence by Theorem 7, it converges uniformly on any $\overline{G}_0(\delta)$, hence on any finite interval in E^1. (The pointwise convergence is on all of E^1.)

4.13.E: More Problems on Series of Functions

Exercise $4.13.E.1$

Verify Note 3 and Example (c) in detail.

Exercise $4.13.E.2$

Show that the so-called hyperharmonic series of order p,

$$\sum \frac{1}{n^{p}} \quad \left(p \in E^{1}\right), \tag{4.13.E.1}$$

converges iff $p>1$.
[Hint: If $p \leq 1$,

$$\sum_{n=1}^{\infty} \frac{1}{n^{p}} \geq \sum_{n=1}^{\infty} \frac{1}{n}=+\infty \quad (\text{ Example (b)}). \tag{4.13.E.2}$$

If $p>1$,

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{n^{p}} &=1+\left(\frac{1}{2^{p}}+\frac{1}{3^{p}}\right)+\left(\frac{1}{4^{p}}+\cdots+\frac{1}{7^{p}}\right)+\left(\frac{1}{8^{p}}+\cdots+\frac{1}{15^{p}}\right)+\cdots \\ & \leq 1+\left(\frac{1}{2^{p}}+\frac{1}{2^{p}}\right)+\left(\frac{1}{4^{p}}+\cdots+\frac{1}{4^{p}}\right)+\left(\frac{1}{8^{p}}+\cdots+\frac{1}{8^{p}}\right)+\cdots \\ &=\sum_{n=0}^{\infty} \frac{1}{\left(2^{p-1}\right)^{n}}. \end{aligned}$$

a convergent geometric series. Explain each step.]

Exercise $4.13.E.3$

$\Rightarrow 3$. Prove the refined comparison test:
(i) If two series of constants, $\sum\left|a_{n}\right|$ and $\sum\left|b_{n}\right|$, are such that the sequence $\left\{\left|a_{n}\right| /\left|b_{n}\right|\right\}$ is bounded in E^{1}, then

$$\sum_{n=1}^{\infty}\left|b_{n}\right|<+\infty \text{ implies } \sum_{n=1}^{\infty}\left|a_{n}\right|<+\infty. \tag{4.13.E.3}$$

(ii) If

$$0<\lim _{n \rightarrow \infty} \frac{\left|a_{n}\right|}{\left|b_{n}\right|}<+\infty, \tag{4.13.E.4}$$

then $\sum\left|a_{n}\right|$ converges if and only if $\sum\left|b_{n}\right|$ does.
What is

$$\lim _{n \rightarrow \infty} \frac{\left|a_{n}\right|}{\left|b_{n}\right|}=+\infty ? \tag{4.13.E.5}$$

[Hint: If (\forall n)|a_{n}| / |b_{n}| \leq K\), then $\left|a_{n}\right| \leq K\left|b_{n}\right|$.]

? Exercise 4.13.E.4

Test $\sum a_n$ for absolute convergence in each of the following. Use Problem 3 or Theorem 2 or the indicated references.
(i) $a_n=\frac{n+1}{\sqrt{n^4+1}}\left(\text{ take } b_n=\frac{1}{n}\right)$;
(ii) $a_n=\frac{\cos n}{\sqrt{n^3-1}}\left(\text{ take } b_n=\frac{1}{\sqrt{n^3}};\text{ use Problem } 2\right)$;
(iii) $a_n=\frac{(-1)^n}{n^p}(\sqrt{n+1}-\sqrt{n}), p \in E^1$;
(iv) $a_n=n^5 e^{-n}$ (use Problem 18 of Chapter 3, §15);
(v) $a_n=\frac{2^n+n}{3^n+1}$;
(vi) $a_n=\frac{(-1)^n}{(\log n)^q}; n \geq 2$;
(vii) $a_n=\frac{(\log n)^q}{n\left(n^2+1\right)}, q \in E^1$.
[Hint for (vi) and (vii): From Problem 14 in §2, show that

$$\lim _{y \rightarrow+\infty} \frac{y}{(\log y)^q}=+\infty \tag{4.13.E.6}$$

and hence

$$\lim _{n \rightarrow \infty} \frac{(\log n)^q}{n}=0. \tag{4.13.E.7}$$

Then select b_n.

? Exercise 4.13.E.5

Prove that $\sum_{n=1}^{\infty} \frac{n^n}{n!}=+\infty$.
[Hint: Show that $n^n / n!$ does not tend to 0.]

? Exercise 4.13.E.6

Prove that $\lim _{n \rightarrow \infty} \frac{x^n}{n!}=0$.
[Hint: Use Example (d) and Theorem 4.]

? Exercise 4.13.E.7

Use Theorems $3,5,6,$ and 7 to show that $\sum\left|f_n\right|$ converges uniformly on B, provided $f_n(x)$ and B are as indicated below, with $0<a<+\infty$ and $b \in E^1$. For parts (ix)–(xii), find $M_n=\max _{x \in B}\left|f_n(x)\right|$ and use Theorem 3. (Calculus rules for maxima are assumed known.)
(i) $\frac{x^{2n}}{(2n)!};[-a, b]$.
(ii) $(-1)^{n+1} \frac{x^{2n-1}}{(2n-1)!};[-a, b]$.
(iii) $\frac{x^n}{n^n};[-a, a]$.
(iv) $n^3 x^n;[-a, a](a<1)$.
(v) $\frac{\sin nx}{n^2}; B=E^1$ (use Problem 2).
(vi) $e^{-nx} \sin nx;[a,+\infty)$.
(vii) $\frac{\cos nx}{\sqrt{n^3+1}}; B=E^1$.
(viii) $a_n \cos nx$, with $\sum_{n=1}^{\infty}\left|a_n\right|<+\infty; B=E^1$.
(ix) $x^n e^{-nx};[0,+\infty)$.
(x) $x^n e^{nx};\left(-\infty, \frac{1}{2}\right]$.
(xi) $(x \cdot \log x)^n, f_n(0)=0;\left[-\frac{3}{2}, \frac{3}{2}\right]$.

(xii) $\left(\frac{\log x}{x}\right)^n$; $[1,+\infty)$.
(xiii) $\frac{q(q-1)\cdots(q-n+1)x^n}{n!}, q \in E^1$; $\left[-\frac{1}{2}, \frac{1}{2}\right]$.

Exercise $4.13.E.8$

$\Rightarrow 8$. (Summation by parts.) Let f_n, h_n, and g_n be real or complex functions (or let f_n and h_n be scalar valued and g_n be vector valued). Let $f_n = h_n - h_{n-1} (n \geq 2)$. Verify that $(\forall m > n > 1)$

$$\begin{aligned}\sum_{k=n+1}^{m} f_k g_k &= \sum_{k=n+1}^{m}\left(h_k - h_{k-1}\right) g_k \\ &= h_m g_m - h_n g_{n+1} - \sum_{k=n+1}^{m-1} h_k\left(g_{k+1} - g_k\right).\end{aligned}$$

[Hint: Rearrange the sum.]

Exercise $4.13.E.9$

$\Rightarrow 9$. (Abel's test.) Let the f_n, g_n, and h_n be as in Problem 8, with $h_n = \sum_{i=1}^{n} f_i$. Suppose that
(i) the range space of the g_n is complete;
(ii) $|g_n| \rightarrow 0$ (uniformly) on a set B; and
(iii) the partial sums $h_n = \sum_{i=1}^{n} f_i$ are uniformly bounded on B; i.e.,

$$\left(\exists K \in E^1\right)(\forall n) \quad |h_n| < K \text{ on } B. \tag{4.13.E.8}$$

Then prove that $\sum f_k g_k$ converges uniformly on B if $\sum |g_{n+1} - g_n|$ does.
(This always holds if the g_n are real and $g_n \geq g_{n+1}$ on B.)
[Hint: Let $\varepsilon > 0$. Show that

$$(\exists k)(\forall m > n > k) \quad \sum_{i=n+1}^{m} |g_{i+1} - g_i| < \varepsilon \text{ and } |g_n| < \varepsilon \text{ on } B. \tag{4.13.E.9}$$

Then use Problem 8 to show that

$$\left|\sum_{i=n+1}^{m} f_i g_i\right| < 3K\varepsilon. \tag{4.13.E.10}$$

Apply Theorem $3'$ of §12.]

Exercise $4.13.E.9'$

$\Rightarrow \mathbf{9'}$. Prove that if $\sum a_n$ is a convergent series of constants $a_n \in E^1$ and if $\{b_n\}$ is a bounded monotone sequence in E^1, then $\sum a_n b_n$ converges.
[Hint: Let $b_n \rightarrow b$. Write

$$a_n b_n = a_n\left(b_n - b\right) + a_n b \tag{4.13.E.11}$$

and use Problem 9 with $f_n = a_n$ and $g_n = b_n - b$.]

Exercise $4.13.E.10$

$\Rightarrow 10$. Prove the Leibniz test for alternating series: If $\{b_n\} \downarrow$ and $b_n \to 0$ in E^1, then $\sum(-1)^n b_n$ converges, and the sum $\sum_{n=1}^{\infty}(-1)^n b_n$ differs from $s_n = \sum_{k=1}^{n}(-1)^k b_k$ by b_{n+1} at most.

Exercise $4.13.E.11$

$\Rightarrow 11$. (Dirichlet test.) Let the f_n, g_n, and h_n be as in Problem 8 with $\sum_{n=0}^{\infty} f_n$ uniformly convergent on B to a function f, and with

$$h_n = -\sum_{i=n+1}^{\infty} f_i \text{ on } B. \tag{4.13.E.12}$$

Suppose that
(i) the range space of the g_n is complete; and
(ii) there is $K \in E^1$ such that

$$\left|g_0\right| + \sum_{n=0}^{\infty}\left|g_{n+1} - g_n\right| < K \text{ on } B. \tag{4.13.E.13}$$

Show that $\sum f_n g_n$ converges uniformly on B.
[Proof outline: We have

$$\left|g_n\right| = \left|g_0 + \sum_{i=0}^{n-1}\left(g_{i+1} - g_i\right)\right| \leq \left|g_0\right| + \sum_{i=0}^{n-1}\left|g_{i+1} - g_i\right| < K \quad \text{by (ii).} \tag{4.13.E.14}$$

Also,

$$\left|h_n\right| = \left|\sum_{i=0}^{n} f_i - f\right| \to 0 \text{ (uniformly) on } B \tag{4.13.E.15}$$

by assumption. Hence

$$(\forall \varepsilon > 0)(\exists k)(\forall n > k) \quad \left|h_n\right| < \varepsilon \text{ on } B. \tag{4.13.E.16}$$

Using Problem 8, obtain

$$(\forall m > n > k)\left|\sum_{i=n+i}^{m} f_i g_i\right| < 2K\varepsilon.] \tag{4.13.E.17}$$

Exercise $4.13.E.12$

Prove that if $0 < p \leq 1$, then $\sum \frac{(-1)^n}{n^p}$ converges conditionally.
[Hint: Use Problems 11 and 2 .]

Exercise $4.13.E.13$

$\Rightarrow 13$. Continuing Problem 14 in §12, prove that if $\sum\left|f_n\right|$ and $\sum\left|g_n\right|$ converge on B (pointwise or uniformly), then so do the series

$$\sum\left|a f_n + b g_n\right|, \sum\left|f_n \pm g_n\right|, \text{ and } \sum\left|a f_n\right|. \tag{4.13.E.18}$$

[Hint : $|af_n + bg_n| \leq |a|\,|f_n| + |b|\,|g_n|$. Use Theorem 2.]
For the rest of the section, we define

$$x^{+} = \max(x, 0) \text{ and } x^{-} = \max(-x, 0). \tag{4.13.E.19}$$

Exercise 4.13.E.14

$\Rightarrow 14$. Given $\{a_n\} \subset E^*$ show the following:
(i) $\sum a_n^+ + \sum a_n^- = \sum |a_n|$.
(ii) If $\sum a_n^+ < +\infty$ or $\sum a_n^- < +\infty$, then $\sum a_n = \sum a_n^+ - \sum a_n^-$.
(iii) If $\sum a_n$ converges conditionally, then $\sum a_n^+ = +\infty = \sum a_n^-$.
(iv) If $\sum |a_n| < +\infty$, then for any $\{b_n\} \subset E^1$,

$$\sum |a_n \pm b_n| < +\infty \text{ iff } \sum |b_n| < \infty; \tag{4.13.E.20}$$

$$\text{moreover, } \sum a_n \pm \sum b_n = \sum (a_n \pm b_n) \text{ if } \sum b_n \text{ exists.} \tag{4.13.E.21}$$

[Hint: Verify that $|a_n| = a_n^+ + a_n^-$ and $a_n = a_n^+ - a_n^-$. Use the rules of §4.]

Exercise 4.13.E.15

$\Rightarrow 15$. (Abel's theorem.) Show that if a power series

$$\sum_{n=0}^{\infty} a_n (x-p)^n \quad \left(a_n \in E, x, p \in E^1\right) \tag{4.13.E.22}$$

converges for some $x = x_0 \neq p$, it converges uniformly on $[p, x_0]$ (or $[x_0, p]$ if $x_0 < p$).
[Proof outline: First let $p = 0$ and $x_0 = 1$. Use Problem 11 with

$$f_n = a_n \text{ and } g_n(x) = x^n = (x-p)^n. \tag{4.13.E.23}$$

As $f_n = a_n 1^n = a_n (x_0 - p)^n$, the series $\sum f_n$ converges by assumption. The convergence is uniform since the f_n are constant. Verify that if $x = 1$, then

$$\sum_{k=1}^{\infty} |g_{k+1} - g_k| = 0, \tag{4.13.E.24}$$

and if $0 \leq x < 1$, then

$$\sum_{k=0}^{\infty} |g_{k+1} - g_k| = \sum_{k=0}^{\infty} x^k |x-1| = (1-x) \sum_{k=0}^{\infty} x^k = 1 \quad \text{(a geometric series).} \tag{4.13.E.25}$$

Also, $|g_0(x)| = x^0 = 1$. Thus by Problem 11(with $K = 2$), $\sum f_n g_n$ converges uniformly on $[0, 1]$, proving the theorem for $p = 0$ and $x_0 = 1$. The general case reduces to this case by the substitution $x - p = (x_0 - p)\,y$. Verify!]

Exercise 4.13.E.16

Prove that if

$$0 < \underline{\lim}\, a_n \leq \overline{\lim}\, a_n < +\infty, \tag{4.13.E.26}$$

then the convergence radius of $\sum a_n(x-p)^n$ is 1.

Exercise $4.13.E.17$

Show that a conditionally convergent series $\sum a_n\left(a_n \in E^1\right)$ can be rearranged so as to diverge, or to converge to any prescribed sum s. [Proof for $s \in E^1:$ Using Problem 14(iii), take the first partial sum

$$a_1^{+}+\cdots+a_m^{+}>s. \tag{4.13.E.27}$$

Then adjoin terms

$$-a_1^{-},-a_2^{-},\ldots,-a_n^{-} \tag{4.13.E.28}$$

until the partial sum becomes less than s. Then add terms a_k^{+} until it exceeds s. Then adjoin terms $-a_k^{-}$ until it becomes less than s, and so on.
As $a_k^{+} \rightarrow 0$ and $a_k^{-} \rightarrow 0$ (why?), the rearranged series tends to s. (Why?)
Give a similar proof for $s=\pm\infty$. Also, make the series oscillate, with no sum.]

Exercise $4.13.E.18$

Prove that if a power series $\sum a_n(x-p)^n$ converges at some $x=x_0 \neq p$, it converges absolutely (pointwise) on $G_p(\delta)$ if $\delta \leq\left|x_0-p\right|$.
[Hint: By Theorem $6, \delta \leq\left|x_0-p\right| \leq r\,(r=$ convergence radius$)$. Fix any $x \in G_p(\delta)$. Show that the line $\overrightarrow{p x}$, when extended, contains a point x_1 such that $|x-p|<\left|x_1-p\right|<\delta \leq r$. By Theorem 6, the series converges absolutely at x_1, hence at x as well, by Theorem 7.]

CHAPTER OVERVIEW

5: Differentiation and Antidifferentiation

5.1: Derivatives of Functions of One Real Variable

In this chapter, " E " will always denote any one of E^{1}, E^{*}, C (the complex field), $E^{n},^{*}$ or another normed space. We shall consider functions $f : E^{1} \rightarrow E$ of one real variable with values in E. Functions $f : E^{1} \rightarrow E^{*}$ (admitting finite and infinite values) are said to be extended real. Thus $f : E^{1} \rightarrow E$ may be real, extended real, complex, or vector valued.

Operations in E^{*} were defined in Chapter 4, §4. Recall, in particular, our conventions $\left(2^{*}\right)$ there. Due to them, addition, subtraction, and multiplication are always defined in E^{*} (with sums and products possibly "unorthodox").

To simplify formulations, we shall also adopt the convention that

$$f(x)=0 \text{ unless defined otherwise.} \tag{5.1.1}$$

(" 0 " stands also for the zero-vector in E if E is a vector space.) Thus each function f is defined on all of E^{1}. For convenience, we call $f(x)$ "finite" if $f(x) \neq \pm \infty($ also if it is a vector $)$.

Definition

For each function $f : E^{1} \rightarrow E,$ we define its derived function $f^{\prime} : E^{1} \rightarrow E$
by setting, for every point $p \in E^{1}$,

$$f^{\prime}(p)=\begin{cases}\lim_{x \rightarrow p} \frac{f(x)-f(p)}{x-p} \text{ if this limit exists (finite or not);} \\ 0, \text{ otherwise.}\end{cases} \tag{5.1.2}$$

Thus $f^{\prime}(p)$ is always defined.

If the limit in (1) exists, we call it the derivative of f at p.

If, in addition, this limit is finite, we say that f is differentiable at p.

If this holds for each p in a set $B \subseteq E^{1}$, we say that f has a derivative (respectively, is differentiable) on B, and we call the function $f^{\prime}$ thederivative of f on B.

If the limit in (1) is one sided (with $x \rightarrow p^{-}$ or $x \rightarrow p^{+}$), we call it a one-sided (left or right) derivative at p, denoted $f_{-}^{\prime}$ or $f_{+}^{\prime}$.

Definition

Given a function $f : E^{1} \rightarrow E,$ we define its n th derived function (or derived function of order n), denoted $f^{(n)} : E^{1} \rightarrow E,$ by induction:

$$f^{(0)}=f, f^{(n+1)}=\left[f^{(n)}\right]^{\prime}, \quad n=0,1,2, \ldots \tag{5.1.3}$$

Thus $f^{(n+1)}$ is the derived function of $f^{(n)}$. By our conventions, $f^{(n)}$ is defined on all of E^{1} for each n and each function $f : E^{1} \rightarrow E.$ We have $f^{(1)}=f^{\prime},$ and we write $f^{\prime \prime}$ for $f^{(2)}, f^{\prime \prime \prime}$ for $f^{(3)}$, etc. We say that f has n derivatives at a point p iff the limits

$$\lim_{x \rightarrow q} \frac{f^{(k)}(x)-f^{(k)}(q)}{x-q} \tag{5.1.4}$$

exist for all q in a neighborhood G_{p} of p and for $k=0,1, \ldots, n-2,$ and also

$$\lim_{x \rightarrow p} \frac{f^{(n-1)}(x)-f^{(n-1)}(p)}{x-p} \tag{5.1.5}$$

exists. If all these limits are finite, we say that f is n times differentiable on I; similarly for one-sided derivatives.

It is an important fact that differentiability implies continuity.

Theorem 5.1.1

If a function $f: E^1 \rightarrow E$ is differentiable at a point $p \in E^1$, it is continuous at p, and $f(p)$ is finite (even if $E=E^*$).

Proof

Setting $\Delta x = x-p$ and $\Delta f = f(x)-f(p)$, we have the identity

$$|f(x)-f(p)| = \left|\frac{\Delta f}{\Delta x} \cdot (x-p)\right| \quad \text{for } x \neq p. \tag{5.1.6}$$

By assumption,

$$f'(p) = \lim_{x \rightarrow p} \frac{\Delta f}{\Delta x} \tag{5.1.7}$$

exists and is finite. Thus as $x \rightarrow p$, the right side of (2) (hence the left side as well) tends to 0, so

$$\lim_{x \rightarrow p} |f(x)-f(p)| = 0, \text{ or } \lim_{x \rightarrow p} f(x) = f(p) \tag{5.1.8}$$

proving continuity at p.

Also, $f(p) \neq \pm\infty$, for otherwise $|f(x)-f(p)| = +\infty$ for all x, and so $|f(x)-f(p)|$ cannot tend to 0. □

Note 1. Similarly, the existence of a finite left (right) derivative at p implies left (right) continuity at p. The proof is the same.

Note 2. The existence of an infinite derivative does not imply continuity, nor does it exclude it. For example, consider the two cases

(i) $f(x) = \frac{1}{x}$, with $f(0) = 0$, and

(ii) $f(x) = \sqrt[3]{x}$.

Give your comments for $p = 0$.

Caution: A function may be continuous on E^1 without being differentiable anywhere (thus the converse to Theorem 1 fails). The first such function was indicated by Weierstrass. We give an example due to Olmsted (Advanced Calculus).

Example 5.1.1

(a) We first define a sequence of functions $f_n: E^1 \rightarrow E^1 (n=1,2,\ldots)$ as follows. For each $k = 0, \pm 1, \pm 2, \ldots,$ let

$$f_n(x) = 0 \text{ if } x = k \cdot 4^{-n}, \text{ and } f_n(x) = \frac{1}{2} \cdot 4^{-n} \text{ if } x = \left(k+\frac{1}{2}\right) \cdot 4^{-n}. \tag{5.1.9}$$

Between $k \cdot 4^{-n}$ and $\left(k \pm \frac{1}{2}\right) \cdot 4^{-n}, f_n$ is linear (see Figure 21), so it is continuous on E^1. The series $\sum f_n$ converges uniformly on E^1. (Verify!)

Let

$$f = \sum_{n=1}^{\infty} f_n. \tag{5.1.10}$$

Then f is continuous on E^1(why? yet it is nowhere differentiable.

To prove this fact, fix any $p \in E^1$. For each n, let

$$x_n = p + d_n, \text{ where } d_n = \pm 4^{-n-1}, \tag{5.1.11}$$

choosing the sign of d_n so that p and x_n are in the same half of a "sawtooth" in the graph of f_n (Figure 21). Then

$$f_n(x_n) - f_n(p) = \pm d_n = \pm(x_n - p). \quad (\text{Why ?}) \tag{5.1.12}$$

Also,

$$f_m(x_n) - f_m(p) = \pm d_n \text{ if } m \leq n \tag{5.1.13}$$

but vanishes for $m>n$. (Why?)

Thus, when computing $f\left(x_{n}\right)-f(p)$, we may replace

$$f=\sum_{m=1}^{\infty} f_{m} \text { by } f=\sum_{m=1}^{n} f_{m} . \tag{5.1.14}$$

Since

$$\frac{f_{m}\left(x_{n}\right)-f_{m}(p)}{x_{n}-p}=\pm 1 \text { for } m \leq n . \tag{5.1.15}$$

the fraction

$$\frac{f\left(x_{n}\right)-f(p)}{x_{n}-p} \tag{5.1.16}$$

is an integer, odd if n is odd and even if n is even. Thus this fraction cannot tend to a finite limit as $n \rightarrow \infty$, i.e., as $d_{n}=4^{-n-1} \rightarrow 0$ and $x_{n}=p+d_{n} \rightarrow p$. A fortiori, this applies to

$$\lim _{x \rightarrow p} \frac{f(x)-f(p)}{x-p} . \tag{5.1.17}$$

Thus f is not differentiable at any p.

The expressions $f(x)-f(p)$ and $x-p$, briefly denoted Δf and Δx, and Δx, are called the increments of f and x (at p), respectively. 2 We now show that for differentiable functions, Δf and Δx are "nearly proportional' when x approaches p; that is,

$$\frac{\Delta f}{\Delta x}=c+\delta(x) \tag{5.1.18}$$

with c constant and $\lim _{x \rightarrow p} \delta(x)=0$.

Theorem 5.1.2

A function $f: E^{1} \rightarrow E$ is differentiable at p, and $f^{\prime}(p)=c$, iff there is a finite $c \in E$ and a function $\delta: E^{1} \rightarrow E$ such that $\lim _{x \rightarrow p} \delta(x)=\delta(p)=0$, and such that

$$\Delta f=[c+\delta(x)] \Delta x \quad \text { for all } x \in E^{1} . \tag{5.1.19}$$

Proof

If f is differentiable at p, put $c=f^{\prime}(p)$. Define $\delta(p)=0$ and

$$\delta(x)=\frac{\Delta f}{\Delta x}-f^{\prime}(p) \text { for } x \neq p . \tag{5.1.20}$$

Then $\lim _{x \rightarrow p} \delta(x)=f^{\prime}(p)-f^{\prime}(p)=0=\delta(p)$. Also, (3) follows.

Conversely, if (3) holds, then

$$\frac{\Delta f}{\Delta x}=c+\delta(x) \rightarrow c \text { as } x \rightarrow p(\text { since } \delta(x) \rightarrow 0) . \tag{5.1.21}$$

Thus by definition,

$$c=\lim _{x \rightarrow p} \frac{\Delta f}{\Delta x}=f^{\prime}(p) \text { and } f^{\prime}(p)=c \text { is finite. } \square \tag{5.1.22}$$

Theorem 5.1.3

(chain rule). Let the functions $g: E^1 \rightarrow E^1(\text{ real })$ and $f: E^1 \rightarrow E$ (real or not) be differentiable at p and q, respectively, where $q=g(p)$. Then the composite function $h=f \circ g$ is differentiable at p, and

$$h^{\prime}(p)=f^{\prime}(q) g^{\prime}(p). \tag{5.1.23}$$

Proof

Setting

$$\Delta h=h(x)-h(p)=f(g(x))-f(g(p))=f(g(x))-f(q). \tag{5.1.24}$$

we must show that

$$\lim_{x \rightarrow p} \frac{\Delta h}{\Delta x}=f^{\prime}(q) g^{\prime}(p) \neq \pm \infty. \tag{5.1.25}$$

Now as f is differentiable at q, Theorem 2 yields a function $\delta: E^1 \rightarrow E$ such that $\lim_{x \rightarrow q} \delta(x)=\delta(q)=0$ and such that

$$\left(\forall y \in E^1\right) \quad f(y)-f(q)=\left[f^{\prime}(q)+\delta(y)\right] \Delta y, \Delta y=y-q. \tag{5.1.26}$$

Taking $y=g(x)$, we get

$$\left(\forall x \in E^1\right) \quad f(g(x))-f(q)=\left[f^{\prime}(q)+\delta(g(x))\right][g(x)-g(p)], \tag{5.1.27}$$

where

$$g(x)-g(p)=y-q=\Delta y \text{ and } f(g(x))-f(q)=\Delta h, \tag{5.1.28}$$

as noted above. Hence

$$\frac{\Delta h}{\Delta x}=\left[f^{\prime}(q)+\delta(g(x))\right] \cdot \frac{g(x)-g(p)}{x-p} \quad \text{ for all } x \neq p. \tag{5.1.29}$$

Let $x \rightarrow p$. Then we obtain $h^{\prime}(p)=f^{\prime}(q) g^{\prime}(p)$, for, by the continuity of $\delta \circ g$ at p (Chapter 4, §2, Theorem 3),

$$\lim_{x \rightarrow p} \delta(g(x))=\delta(g(p))=\delta(q)=0. \square \tag{5.1.30}$$

The proofs of the next two theorems are left to the reader.

Theorem 5.1.4

If f, g, and h are real or complex and are differentiable at p, so are

$$f \pm g, h f, \text{ and } \frac{f}{h} \tag{5.1.31}$$

(the latter if $h(p) \neq 0$), and at the point p we have

(i) $(f \pm g)^{\prime}=f^{\prime} \pm g^{\prime}$;

(ii) $(h f)^{\prime}=h f^{\prime}+h^{\prime} f$; and

(iii) $\left(\frac{f}{h}\right)^{\prime}=\frac{h f^{\prime}-h^{\prime} f}{h^2}$.

All this holds also if f and g are vector valued and h is scalar valued. It also applies to infinite (even one-sided) derivatives, except when the limits involved become indeterminate (Chapter 4, §4).

Note 3. By induction, if f, g, and h are n times differentiable at a point p, so are $f \pm g$ and $h f$, and, denoting by $\left(\begin{array}{l}n \\ k\end{array}\right)$ the binomial coefficients, we have

(i*) $(f \pm g)^{(n)}=f^{(n)} \pm g^{(n)}$; and (ii*) $(hf)^{(n)}=\sum_{k=0}^{n}\left(\begin{array}{l}n\\k\end{array}\right)h^{(k)}f^{(n-k)}$.

Formula (ii) is known as the Leibniz formula; its proof is analogous to that of the binomial theorem. It is symbolically written as $(hf)^{(n)}=(h+f)^{n}$, with the last term interpreted accordingly.

Theorem 5.1.5

(componentwise differentiation). A function $f:E^{1}\rightarrow E^{n}\left({}^{*}C^{n}\right)$ is differentiable at p iff each of its n components $(f_{1},\ldots,f_{n})$ is, and then

$$f^{\prime}(p)=\left(f_{1}^{\prime}(p),\ldots,f_{n}^{\prime}(p)\right)=\sum_{k=1}^{n}f_{k}^{\prime}(p)\bar{e}_{k}, \tag{5.1.32}$$

with $\bar{e}_{k}$ as in Theorem 2 of Chapter 3, §§1-3.

In particular, a complex function $f:E^{1}\rightarrow C$ is differentiable iff its real and imaginary parts are, and $f^{\prime}=f_{\mathrm{re}}^{\prime}+i\cdot f_{\mathrm{im}}^{\prime}$ Chapter 4, §3, Note 5).

Example 5.1.2

(b) Consider the complex exponential

$$f(x)=\cos x+i\cdot\sin x=e^{xi}(\text{ Chapter }4,\S3). \tag{5.1.33}$$

We assume the derivatives of $\cos x$ and $\sin x$ to be known (see Problem 8). By Theorem 5, we have

$$f^{\prime}(x)=-\sin x+i\cdot\cos x=\cos\left(x+\frac{1}{2}\pi\right)+i\cdot\sin\left(x+\frac{1}{2}\pi\right)=e^{\left(x+\frac{1}{2}\pi\right)i}. \tag{5.1.34}$$

Hence by induction,

$$f^{(n)}(x)=e^{\left(x+\frac{1}{2}n\pi\right)i},n=1,2,\ldots.(\text{ Verify! }) \tag{5.1.35}$$

(c) Define $f:E^{1}\rightarrow E^{3}$ by

$$f(x)=(1,\cos x,\sin x),\quad x\in E^{1}. \tag{5.1.36}$$

Here Theorem 5 yields

$$f^{\prime}(p)=(0,-\sin p,\cos p),\quad p\in E^{1}. \tag{5.1.37}$$

For a fixed $p=p_{0}$, we may consider the line

$$\bar{x}=\bar{a}+t\vec{u}, \tag{5.1.38}$$

where

$$\bar{a}=f\left(p_{0}\right)\text{ and }\vec{u}=f^{\prime}\left(p_{0}\right)=\left(0,-\sin p_{0},\cos p_{0}\right). \tag{5.1.39}$$

This is, by definition, the tangent vector at p_0 to the curve $f\left[E^{1}\right]$ in E^{3}.

More generally, if $f:E^{1}\rightarrow E$ is differentiable at p and continuous on some globe about p, we define the tangent at p to the curve $f\left[G_{p}\right]$ to be the line

$$\bar{x}=f(p)+t\cdot f^{\prime}(p); \tag{5.1.40}$$

$f^{\prime}(p)$ is its direction vector in E, while t is the variable real parameter. For real functions $f:E^{1}\rightarrow E^{1}$, we usually consider not $f\left[E^{1}\right]$ but the curve $y=f(x)$ in E^{2}, i.e., the set

$$\left\{(x,y)|y=f(x),x\in E^{1}\right\}. \tag{5.1.41}$$

The tangent to that curve at p is the line through $(p,f(p))$ with slope $f^{\prime}(p)$.

In conclusion, let us note that differentiation (i.e., taking derivatives) is a local limit process at some point p. Hence (cf. Chapter 4, §1, Note 4) the existence and the value of $f'(p)$ is not affected by restricting f to some globe G_p about p or by arbitrarily redefining f outside G_p. For one-sided derivatives, we may replace G_p by its corresponding "half."

5.1.E: Problems on Derived Functions in One Variable

Exercise $5.1.E.1$

Prove Theorems 4 and 5, including (i^*) and (ii^*). Do it for dot products as well.

Exercise $5.1.E.2$

Verify Note 2.

Exercise $5.1.E.3$

Verify Example (a).

Exercise $5.1.E.3'$

Verify Example (b).

Exercise $5.1.E.4$

Prove that if f has finite one-sided derivatives at p, it is continuous at p.

Exercise $5.1.E.5$

Restate and prove Theorems 2 and 3 for one-sided derivatives.

Exercise $5.1.E.6$

Prove that if the functions $f_i : E^1 \rightarrow E^*(C)$ are differentiable at p, so is their product, and

$$\left(f_1 f_2 \cdots f_m\right)^{\prime}=\sum_{i=1}^m\left(f_1 f_2 \cdots f_{i-1} f_i^{\prime} f_{i+1} \cdots f_m\right) \text { at } p. \tag{5.1.E.1}$$

Exercise $5.1.E.7$

A function $f : E^1 \rightarrow E$ is said to satisfy a Lipschitz condition (L) of order $\alpha(\alpha>0)$ at p iff

$$(\exists \delta>0)\left(\exists K \in E^1\right)\left(\forall x \in G_{\neg p}(\delta)\right) \quad|f(x)-f(p)| \leq K|x-p|^\alpha. \tag{5.1.E.2}$$

(i) This implies continuity at p but not conversely; take

$$f(x)=\frac{1}{\ln |x|}, \quad f(0)=0, \quad p=0. \tag{5.1.E.3}$$

[Hint: For the converse, start with Problem 14 (iii) of Chapter 4, §2.]
(ii) L of order $\alpha>1$ implies differentiability at p, with $f^{\prime}(p)=0$.
(iii) Differentiability implies L of order 1, but not conversely. (Take

$$f(x)=x \sin \frac{1}{x}, f(0)=0, p=0; \tag{5.1.E.4}$$

then even one-sided derivatives fail to exist.)

Exercise $5.1.E.8$

Let

$$f(x)=\sin x \text{ and } g(x)=\cos x. \tag{5.1.E.5}$$

Show that f and g are differentiable on E^1, with

$$f'(p)=\cos p \text{ and } g'(p)=-\sin p \text{ for each } p\in E^1. \tag{5.1.E.6}$$

Hence prove for $n=0,1,2,\ldots$ that

$$f^{(n)}(p)=\sin\left(p+\frac{n\pi}{2}\right) \text{ and } g^{(n)}(p)=\cos\left(p+\frac{n\pi}{2}\right). \tag{5.1.E.7}$$

[Hint: Evaluate Δf as in Example (d) of Chapter 4, §8. Then use the continuity of f and the formula

$$\lim_{z\to 0}\frac{\sin z}{z}=\lim_{z\to 0}\frac{z}{\sin z}=1. \tag{5.1.E.8}$$

To prove the latter, note that

$$|\sin z|\leq|z|\leq|\tan z|, \tag{5.1.E.9}$$

whence

$$1\leq\frac{z}{\sin z}\leq\frac{1}{|\cos z|}\to 1; \tag{5.1.E.10}$$

similarly for g.]

Exercise $5.1.E.9$

Prove that if f is differentiable at p then

$$\lim_{\substack{x\to p^+\\ y\to p^-}}\frac{f(x)-f(y)}{x-y} \text{ exists, is finite, and equals } f'(p); \tag{5.1.E.11}$$

i.e., $(\forall\varepsilon>0)(\exists\delta>0)(\forall x\in(p,p+\delta))(\forall y\in(p-\delta,p))$

$$\left|\frac{f(x)-f(y)}{x-y}-f'(p)\right|<\varepsilon. \tag{5.1.E.12}$$

Show, by redefining f at p, that even if the limit exists, f may not be differentiable (note that the above limit does not involve $f(p)$).
[Hint: If $y<p<x$ then

$$\begin{aligned}\left|\frac{f(x)-f(y)}{x-y}-f'(p)\right| &\leq\left|\frac{f(x)-f(p)}{x-y}-\frac{x-p}{x-y}f'(p)\right|+\left|\frac{f(p)-f(y)}{x-y}-\frac{p-y}{x-y}f'(p)\right|\\ &\leq\left|\frac{f(x)-f(p)}{x-p}-f'(p)\right|+\left|\frac{f(p)-f(y)}{p-y}-f'(p)\right|\to 0.\end{aligned}$$]

Exercise $5.1.E.10$

Prove that if f is twice differentiable at p, then

$$f''(x)=\lim_{h\rightarrow 0}\frac{f(p+h)-2f(p)+f(p-h)}{h^2}\neq\pm\infty. \tag{5.1.E.13}$$

Does the converse hold (cf. Problem 9)?

Exercise $5.1.E.11$

In Example (c), find the three coordinate equations of the tangent line at $p=\frac{1}{2}\pi$.

Exercise $5.1.E.12$

Judging from Figure 22 in §2, discuss the existence, finiteness, and sign of the derivatives (or one-sided derivatives) of f at the points p_i indicated.

Exercise $5.1.E.13$

Let $f: E^n\rightarrow E$ be linear, i.e., such that

$$\left(\forall\overline{x},\overline{y}\in E^n\right)\left(\forall a,b\in E^1\right)\quad f(a\overline{x}+b\overline{y})=af(\overline{x})+bf(\overline{y}). \tag{5.1.E.14}$$

Prove that if $g: E^1\rightarrow E^n$ is differentiable at p, so is $h=f\circ g$ and $h'(p)=f\left(g'(p)\right)$.
[Hint: f is continuous since $f(\overline{x})=\sum_{k=1}^n x_k f\left(\overline{e}_k\right)$. See Problem 5 in Chapter 3, §§4-6.]

5.2: Derivatives of Extended-Real Functions

For a while (in §§2 and 3), we limit ourselves to extended-real functions. Below, f and g are real or extended real (f, $g: E^1 \rightarrow E^*$). We assume, however, that they are not constantly infinite on any interval $(a, b), a < b$.

Lemma 5.2.1

If $f'(p) > 0$ at some $p \in E^1$, then

$$x < p < y \tag{5.2.1}$$

implies

$$f(x) < f(p) < f(y) \tag{5.2.2}$$

for all x, y in a sufficiently small globe $G_p(\delta) = (p - \delta, p + \delta)$.

Similarly, if $f'(p) < 0$, then $x < p < y$ implies $f(x) > f(p) > f(y)$ for x, y in some $G_p(\delta)$.

Proof

If $f'(p) > 0,$ the "0" case in Definition 1 of §1, is excluded, so

$$f'(p) = \lim_{x \rightarrow p} \frac{\Delta f}{\Delta x} > 0. \tag{5.2.3}$$

Hence we must also have $\Delta f / \Delta x > 0$ for x in some $G_p(\delta)$.

It follows that Δf and Δx have the same sign in $G_p(\delta)$; i.e.,

$$f(x) - f(p) > 0 \text{ if } x > p \text{ and } f(x) - f(p) < 0 \text{ if } x < p. \tag{5.2.4}$$

(This implies $f(p) \neq \pm\infty$. Why? Hence

$$x < p < y \Longrightarrow f(x) < f(p) < f(y), \tag{5.2.5}$$

as claimed; similarly in case $f'(p) < 0$. $\square$

Corollary 5.2.1

If $f(p)$ is the maximum or minimum value of $f(x)$ for x in some $G_p(\delta)$, then $f'(p) = 0$; i.e., f has a zero derivative, or none at all, at p.

For, by Lemma 1, $f'(p) \neq 0$ excludes a maximum or minimum at p. (Why?)

Note 1. Thus $f'(p) = 0$ is a necessary condition for a local maximum or minimum at p. It is insufficient, however. For example, if $f(x) = x^3$, f has no maxima or minima at all, yet $f'(0) = 0$. For sufficient conditions, see §6.

Figure 22 illustrates these facts at the points $p_2, p_3, \ldots, p_{11}$. Note that in Figure 22, the isolated points P, Q, R belong to the graph.

Geometrically, $f'(p) = 0$ means that the tangent at p is horizontal, or that a two-sided tangent does not exist at p.

Theorem 5.2.1

Let $f: E^1 \rightarrow E^*$ be relatively continuous on an interval $[a, b]$, with $f' \neq 0$ on (a, b). Then f is strictly monotone on $[a, b]$, and f' is signconstant there (possibly 0 at a and b), with $f' \geq 0$ if $f \uparrow$, and $f' \leq 0$ if $f \downarrow$.

Proof

By Theorem 2 of Chapter 4, §8, f attains a least value m, and a largest value M, at some points of $[a, b]$. However, neither can occur at an interior point $p \in (a, b),$ for, by Corollary 1, this would imply $f'(p) = 0,$ contrary to our assumption.

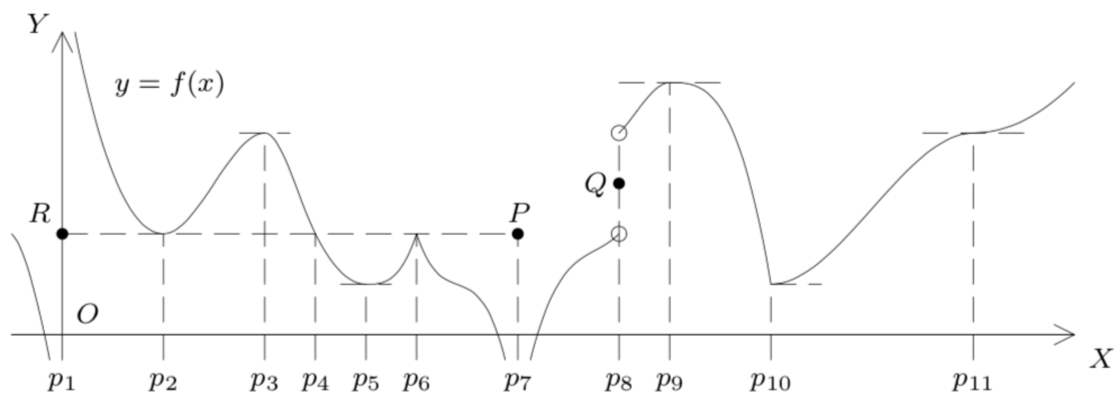

FIGURE 22

Thus $M=f(a)$ or $M=f(b)$; for the moment we assume $M=f(b)$ and $m=f(a)$. We must have $m<M$, for $m=M$ would make f constant on $[a,b]$, implying $f'=0$. Thus $m=f(a)<f(b)=M$.

Now let $a \leq x<y \leq b$. Applying the previous argument to each of the intervals $[a,x]$, $[a,y]$, $[x,y]$,and $[x,b]$ (now using that $m=f(a)<f(b)=M)$, we find that

$$f(a) \leq f(x)<f(y) \leq f(b). \quad \text{(Why?)} \tag{5.2.6}$$

Thus $a \leq x<y \leq b$ implies $f(x)<f(y)$; i.e., f increases on $[a,b]$. Hence f' cannot be negative at any $p \in[a,b]$, for, otherwise, by Lemma 1, f would decrease at p. Thus $f' \geq 0$ on $[a,b]$.

In the case $M=f(a)>f(b)=m$, we would obtain $f' \leq 0$. $\square$

Caution: The function f may increase or decrease at p even if $f'(p)=0$.
See Note 1.

Corollary 5.2.2 (Rolle's theorem)

If : $E^1 \rightarrow E^*$ is relatively continuous on $[a,b]$ and if $f(a)=f(b)$, then $f'(p)=0$ for at least one interior point $p \in(a,b)$.

For, if $f' \neq 0$ on all of (a,b), then by Theorem 1, f would be strictly monotone on $[a,b]$, so the equality $f(a)=f(b)$ would be impossible.

Figure 22 illustrates this on the intervals $[p_2, p_4]$ and $[p_4, p_6]$, with $f'(p_3)= f'(p_5)=0$. A discontinuity at 0 causes an apparent failure on $[0, p_2]$.

Note 2. Theorem 1 and Corollary 2 hold even if $f(a)$ and $f(b)$ are infinite, if continuity is interpreted in the sense of the metric ρ' of Problem 5 in Chapter 3, §11. (Weierstrass' Theorem 2 of Chapter 4, §8 applies to (E^*, ρ') , with the same proof.)

Theorem 5.2.2 (Cauchy's law of the mean)

Let the functions $f, g: E^1 \rightarrow E^*$ be relatively continuous and finite on $[a,b]$ and have derivatives on (a,b), with f' and g' never both infinite at the same point $p \in(a,b)$. Then

$$g'(q)[f(b)-f(a)]=f'(q)[g(b)-g(a)] \text{ for at least one } q \in(a,b). \tag{5.2.7}$$

Proof

Let $A=f(b)-f(a)$ and $B=g(b)-g(a)$. We must show that $Ag'(q)=Bf'(q)$ for some $q \in(a,b)$. For this purpose, consider the function $h=Ag-Bf$. It is relatively continuous and finite on $[a,b]$, as are g and f. Also,

$$h(a)=f(b)g(a)-g(b)f(a)=h(b). \quad \text{(Verify!)} \tag{5.2.8}$$

Thus by Corollary 2, $h'(q)=0$ for some $q \in (a,b)$. Here, by Theorem 4 of §1, $h'=(Ag-Bf)'=Ag'-Bf'$. (This is legitimate, for, by assumption, f' and g' never both become infinite, so no indeterminate limits occur.) Thus $h'(q)=Ag'(q)-Bf'(q)=0,$ and (1) follows. □

Corollary 5.2.3 (Lagrange's law of the mean)

If $f: E^1 \rightarrow E^1$ is relatively continuous on $[a,b]$ with a derivative on (a,b), then

$$f(b)-f(a)=f'(q)(b-a) \text{ for at least one } q \in (a,b). \tag{5.2.9}$$

Proof

Take $g(x)=x$ in Theorem 2, so $g'=1$ on E^1. □

Note 3. Geometrically,

$$\frac{f(b)-f(a)}{b-a} \tag{5.2.10}$$

is the slope of the secant through $(a, f(a))$ and $(b, f(b))$, and $f'(q)$ is the slope of the tangent line at q. Thus Corollary 3 states that the secant is parallel to the tangent at some intermediate point q; see Figure 23. Theorem 2 states the same for curves given parametrically: $x=f(t), y=g(t)$.

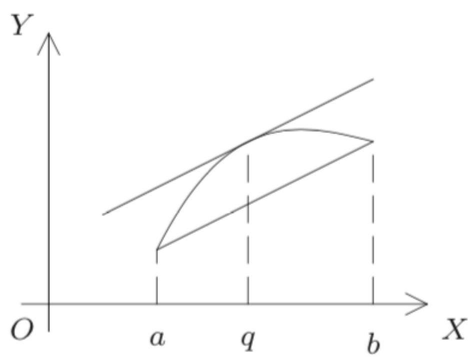

FIGURE 23

Corollary 5.2.4

Let f be as in Corollary 3. Then

(i) f is constant on $[a,b]$ iff $f'=0$ on (a,b);

(ii) $f \uparrow$ on $[a,b]$ iff $f' \geq 0$ on (a,b); and

(iii) $f \downarrow$ on $[a,b]$ iff $f' \leq 0$ on (a,b).

Proof

Let $f'=0$ on (a,b). If $a \leq x \leq y \leq b,$ apply Corollary 3 to the interval $[x,y]$ to obtain

$$f(y)-f(x)=f'(q)(y-x) \text{ for some } q \in (a,b) \text{ and } f'(q)=0. \tag{5.2.11}$$

Thus $f(y)-f(x)=0$ for $x, y \in [a,b]$, so f is constant.

The rest is left to the reader. □

Theorem 5.2.3 (inverse functions)

Let $f : E^1 \to E^1$ be relatively continuous and strictly monotone on an interval $I \subseteq E^1$. Let $f'(p) \neq 0$ at some interior point $p \in I$. Then the inverse function $g = f^{-1}$ (with f restricted to I) has a derivative at $q = f(p)$, and

$$g'(q) = \frac{1}{f'(p)}. \tag{5.2.12}$$

(If $f'(p) = \pm\infty$, then $g'(q) = 0$.)

Proof

By Theorem 3 of Chapter 4, §9, $g = f^{-1}$ is strictly monotone and relatively continuous on $f[I]$, itself an interval. If p is interior to I, then $q = f(p)$ is interior to $f[I]$. (Why?)

Now if $y \in f[I]$, we set

$$\Delta g = g(y) - g(q), \Delta y = y - q, x = f^{-1}(y) = g(y), \text{ and } f(x) = y \tag{5.2.13}$$

and obtain

$$\frac{\Delta g}{\Delta y} = \frac{g(y) - g(q)}{y - q} = \frac{x - p}{f(x) - f(p)} = \frac{\Delta x}{\Delta f} \text{ for } x \neq p. \tag{5.2.14}$$

Now if $y \to q$, the continuity of g at q yields $g(y) \to g(q)$; i.e., $x \to p$. Also, $x \neq p$ iff $y \neq q$, for f and g are one-to-one functions. Thus we may substitute $y = f(x)$ or $x = g(y)$ to get

$$g'(q) = \lim_{y \to q} \frac{\Delta g}{\Delta y} = \lim_{x \to p} \frac{\Delta x}{\Delta f} = \frac{1}{\lim_{x \to p} (\Delta f / \Delta x)} = \frac{1}{f'(p)}, \tag{5.2.15}$$

where we use the convention $\frac{1}{\infty} = 0$ if $f'(p) = \infty$. □

Examples

(A) Let

$$f(x) = \log_a |x| \text{ with } f(0) = 0. \tag{5.2.16}$$

Let $p > 0$. Then $(\forall x > 0)$

$$\begin{aligned} \Delta f &= f(x) - f(p) = \log_a x - \log_a p = \log_a (x/p) \\ &= \log_a \frac{p + (x - p)}{p} = \log_a \left(1 + \frac{\Delta x}{p}\right). \end{aligned}$$

Thus

$$\frac{\Delta f}{\Delta x} = \log_a \left(1 + \frac{\Delta x}{p}\right)^{1/\Delta x}. \tag{5.2.17}$$

Now let $z = \Delta x / p$. (Why is this substitution admissible?) Then using the formula

$$\lim_{z \to 0} (1 + z)^{1/z} = e \quad \text{(see Chapter 4, §2, Example (C))} \tag{5.2.18}$$

and the continuity of the log and power functions, we obtain

$$f'(p) = \lim_{x \to p} \frac{\Delta f}{\Delta x} = \lim_{z \to 0} \log_a \left[(1 + z)^{1/z}\right]^{1/p} = \log_a e^{1/p} = \frac{1}{p} \log_a e. \tag{5.2.19}$$

The same formula results also if $p < 0$, i.e., $|p| = -p$. At $p = 0$, f has one-sided derivatives $(\pm\infty)$ only (verify!), so $f'(0) = 0$ by Definition 1 in §1.

(B) The inverse of the $\log_a$ function is the exponential $g : E^1 \to E^1$, with

$$g(y)=a^y \quad (a>0, a\neq 1). \tag{5.2.20}$$

By Theorem 3, we have

$$\left(\forall q\in E^1\right) \quad g'(q)=\frac{1}{f'(p)}, p=g(q)=a^q. \tag{5.2.21}$$

Thus

$$g'(q)=\frac{1}{\frac{1}{p}\log_a e}=\frac{p}{\log_a e}=\frac{a^q}{\log_a e}. \tag{5.2.22}$$

Symbolically,

$$(\log_a|x|)'=\frac{1}{x}\log_a e(x\neq 0); \quad (a^x)'=\frac{a^x}{\log_a e}=a^x\ln a. \tag{5.2.23}$$

In particular, if $a=e$, we have $\log_e a=1$ and $\log_a x=\ln x$; hence

$$(\ln|x|)'=\frac{1}{x}(x\neq 0) \quad \text{and} \quad (e^x)'=e^x \quad \left(x\in E^1\right). \tag{5.2.24}$$

(C) The power function $g:(0,+\infty)\to E^1$ is given by

$$g(x)=x^a=\exp(a\cdot\ln x) \text{ for } x>0 \text{ and fixed } a\in E^1. \tag{5.2.25}$$

By the chain rule (§1, Theorem 3), we obtain

$$g'(x)=\exp(a\cdot\ln x)\cdot\frac{a}{x}=x^a\cdot\frac{a}{x}=a\cdot x^{a-1}. \tag{5.2.26}$$

Thus we have the symbolic formula

$$(x^a)'=a\cdot x^{a-1} \text{ for } x>0 \text{ and fixed } a\in E^1. \tag{5.2.27}$$

Theorem 5.2.4 (Darboux)

If $f:E^1\to E^*$ is relatively continuous and has a derivative on an interval I, then f' has the Darboux property (Chapter 4, §9) on I.

Proof

Let $p,q\in I$ and $f'(p)<c<f'(q)$. Put $g(x)=f(x)-cx$. Assume $g'\neq 0$ on (p,q) and find a contradiction to Theorem 1. Details are left to the reader. □

5.2.E: Problems on Derivatives of Extended-Real Functions

Exercise $5.2.E.1$

Complete the missing details in the proof of Theorems $1, 2$, and 4, Corollary 4, and Lemma 1.
$[$Hint for converse to Corollary 4(ii) : Use Lemma 1 for an indirect proof. $]$

Exercise $5.2.E.2$

Do cases $p \leq 0$ in Example (A).

Exercise $5.2.E.3$

Show that Theorems $1, 2$, and 4 and Corollaries 2 to 4 hold also if f is discontinuous at a and b but $f\left(a^{+}\right)$ and $f\left(b^{-}\right)$ exist and are finite. (In Corollary 2, assume also $f\left(a^{+}\right)=f\left(b^{-}\right)$; in Theorems 1 and 4 and Corollary 2, finiteness is unnecessary.)
$[$Hint: Redefine $f(a)$ and $f(b)$. $]$

Exercise $5.2.E.4$

Under the assumptions of Corollary 3, show that f' cannot stay infinite on any interval $(p, q), a \leq p<q \leq b$.
$[$Hint: Apply Corollary 3 to the interval $[p, q]$. $]$

Exercise $5.2.E.5$

Justify footnote 1.
[Hint: Let

$$f(x)=x+2 x^{2} \sin \frac{1}{x^{2}} \text { with } f(0)=0. \tag{5.2.E.1}$$

At 0, find f' from Definition 1 in §1. Use also Problem 8 of §1. Show that f is not monotone on any $G_{0}(\delta)$. $]$

Exercise $5.2.E.6$

Show that f' need not be continuous or bounded on $[a, b]$ (under the standard metric), even if f is differentiable there.
$[$Hint: Take f as in Problem 5. $]$

Exercise $5.2.E.7$

With f as in Corollaries 3 and 4, prove that if $f' \geq 0\left(f' \leq 0\right)$ on (a, b) and if f' is not constantly 0 on any subinterval $(p, q) \neq \emptyset$, then f is strictly monotone on $[a, b]$.

Exercise $5.2.E.8$

Let $x=f(t), y=g(t)$, where t varies over an open interval $I \subseteq E^{1}$, define a curve in E^{2} parametrically. Prove that if f and g have derivatives on I and $f' \neq 0$, then the function $h=f^{-1}$ has a derivative on $f[I]$, and the slope of the tangent to the curve at t_{0} equals $g'\left(t_{0}\right) / f'\left(t_{0}\right)$.
$[$Hint: The word "curve" implies that f and g are continuous on I (Chapter 4, §10), so Theorems 1 and 3 apply, and $h=f^{-1}$ is a function. Also, $y=g(h(x))$. Use Theorem 3 of §1. $]$

Exercise $5.2.E.9$

Prove that if f is continuous and has a derivative on (a, b) and if f' has a finite or infinite (even one-sided) limit at some $p \in (a, b)$, then this limit equals $f'(p)$. Deduce that f' is continuous at p if $f'\left(p^{-}\right)$ and $f'\left(p^{+}\right)$ exist.
[Hint: By Corollary 3, for each $x \in (a, b)$, there is some q_x between p and x such that

$$f'\left(q_x\right)=\frac{\Delta f}{\Delta x} \rightarrow f'(p) \text{ as } x \rightarrow p. \tag{5.2.E.2}$$

Set $y=q_x$, so $\lim_{y \rightarrow p} f'(y)=f'(p)$.]

Exercise $5.2.E.10$

From Theorem 3 and Problem 8 in §1, deduce the differentiation formulas

$$(\arcsin x)'=\frac{1}{\sqrt{1-x^2}};(\arccos x)'=\frac{-1}{\sqrt{1-x^2}};(\arctan x)'=\frac{1}{1+x^2}. \tag{5.2.E.3}$$

Exercise $5.2.E.11$

Prove that if f has a derivative at p, then $f(p)$ is finite, provided f is not constantly infinite on any interval (p, q) or $(q, p), p \neq q$.
[Hint: If $f(p)=\pm\infty$, each G_p has points at which $\frac{\Delta f}{\Delta x}=+\infty$, as well as those x with $\frac{\Delta f}{\Delta x}=-\infty$.]

5.3: L'Hôpital's Rule

We shall now prove a useful rule for resolving indeterminate limits. Below, $G_{\neg p}$ denotes a deleted globe $G_{\neg p}(\delta)$ in E^1, or one about $\pm\infty$ of the form $(a,+\infty)$ or $(-\infty, a)$. For one-sided limits, replace $G_{\neg p}$ by its appropriate "half."

Theorem 5.3.1 (L'Hôpital's rule)

Let $f, g : E^1 \rightarrow E^*$ be differentiable on $G_{\neg p}$, with $g' \neq 0$ there. If $|f(x)|$ and $|g(x)|$ tend both to $+\infty$,[1] or both to 0, as $x \rightarrow p$ and if

$$\lim_{x \rightarrow p} \frac{f'(x)}{g'(x)} = r \text{ exists in } E^*, \tag{5.3.1}$$

then also

$$\lim_{x \rightarrow p} \frac{f(x)}{g(x)} = r; \tag{5.3.2}$$

similarly for $x \rightarrow p^+$ or $x \rightarrow p^-$.

Proof

It suffices to consider left and right limits. Both combined then yield the two-sided limit.

First, let $-\infty \leq p < +\infty$,

$$\lim_{x \rightarrow p^+} |f(x)| = \lim_{x \rightarrow p^+} |g(x)| = +\infty \text{ and } \lim_{x \rightarrow p^+} \frac{f'(x)}{g'(x)} = r \text{ (finite)}. \tag{5.3.3}$$

Then given $\varepsilon > 0$, we can fix $a > p$ $(a \in G_{\neg p})$ such that

$$\left|\frac{f'(x)}{g'(x)} - r\right| < \varepsilon, \text{ for all } x \text{ in the interval } (p, a). \tag{5.3.4}$$

Now apply Cauchy's law of the mean (§2, Theorem 2) to each interval $[x, a]$, $p < x < a$. This yields, for each such x, some $q \in (x, a)$ with

$$g'(q)[f(x) - f(a)] = f'(q)[g(x) - g(a)]. \tag{5.3.5}$$

As $g' \neq 0$ (by assumption), $g(x) \neq g(a) \neq g(a)$ by Theorem 1, §2, so we may divide to obtain

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(q)}{g'(q)}, \quad \text{where } p < x < q < a. \tag{5.3.6}$$

This combined **with (1)** yields

$$\left|\frac{f(x) - f(a)}{g(x) - g(a)} - r\right| < \varepsilon, \tag{5.3.7}$$

or, setting

$$F(x) = \frac{1 - f(a)/f(x)}{1 - g(a)/g(x)}, \tag{5.3.8}$$

we have

$$\left|\frac{f(x)}{g(x)} \cdot F(x) - r\right| < \varepsilon \text{ for all } x \text{ inside } (p, a). \tag{5.3.9}$$

As $|f(x)|$ and $|g(x)| \rightarrow +\infty$ (by assumption), we have $F(x) \rightarrow 1$ as $x \rightarrow p^+$. Hence by rules for right limits, there is $b \in (p, a)$ such that for all $x \in (p, b)$, both $|F(x) - 1| < \varepsilon$ and $F(x) > \frac{1}{2}$. (Why?) For such x, **formula (2)** holds as well. Also,

$$\frac{1}{|F(x)|}<2 \text{ and } |r-rF(x)|=|r||1-F(x)|<|r|\varepsilon. \tag{5.3.10}$$

Combining this **with (2)**, we have for $x \in (p,b)$

$$\begin{aligned}\left|\frac{f(x)}{g(x)}-r\right| &= \frac{1}{|F(x)|}\left|\frac{f(x)}{g(x)}F(x)-rF(x)\right| \\ &< 2\left|\frac{f(x)}{g(x)}\cdot F(x)-r+r-rF(x)\right| \\ &< 2\varepsilon(1+|r|).\end{aligned}$$

Thus, given $\varepsilon>0,$ we found $b>p$ such that

$$\left|\frac{f(x)}{g(x)}-r\right|<2\varepsilon(1+|r|), \quad x\in(p,b). \tag{5.3.11}$$

As ε is arbitrary, we have $\lim_{x\to p^+}\frac{f(x)}{g(x)}=r,$ as claimed.

The case $\lim_{x\to p^+} f(x)=\lim_{x\to p^+} g(x)=0$ is simpler. As before, we obtain

$$\left|\frac{f(x)-f(a)}{g(x)-g(a)}-r\right|<\varepsilon. \tag{5.3.12}$$

Here we may as well replace " a'' by any $y\in(p,a).$ Keeping y fixed, let $x\to p^+$. Then $f(x)\to 0$ and $g(x)\to 0,$ so we get

$$\left|\frac{f(y)}{g(y)}-r\right|\leq\varepsilon \text{ for any } y\in(p,a). \tag{5.3.13}$$

As ε is arbitrary, this implies $\lim_{y\to p^+}\frac{f(y)}{g(y)}=r.$ Thus the case $x\to p^+$ is settled for a finite r.

The cases $r=\pm\infty$ and $x\to p^-$ are analogous, and we leave them to the reader. □

Note 1. $\lim\frac{f(x)}{g(x)}$ may exist even if $\lim\frac{f'(x)}{g'(x)}$ does not. For example, take

$$f(x)=x+\sin x \text{ and } g(x)=x. \tag{5.3.14}$$

Then

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}\left(1+\frac{\sin x}{x}\right)=1 \quad (\text{ why? }), \tag{5.3.15}$$

but

$$\frac{f'(x)}{g'(x)}=1+\cos x \tag{5.3.16}$$

does not tend to any limit as $x\to+\infty$.

Note 2. The rule fails if the required assumptions are not satisfied, e.g., if g' has zero values in each $G_{\neg p}$; see Problem 4 below.

Often it is useful to combine L'Hôpital's rule with some known limit formulas, such as

$$\lim_{z\to 0}(1+z)^{1/z}=e \text{ or } \lim_{x\to 0}\frac{x}{\sin x}=1 \text{ (see §1, Problem 8)}. \tag{5.3.17}$$

Examples

(a) $\lim_{x\to+\infty}\frac{\ln x}{x}=\lim_{x\to+\infty}\frac{(\ln x)'}{1}=\lim_{x\to+\infty}\frac{1}{x}=0.$

(b) $\lim_{x\to 0}\frac{\ln(1+x)}{x}=\lim_{x\to 0}\frac{1/(1+x)}{1}=1.$

(c) $\lim_{x\to 0}\frac{x-\sin x}{x^3}=\lim_{x\to 0}\frac{1-\cos x}{3x^2}=\lim_{x\to 0}\frac{\sin x}{6x}=\frac{1}{6}\lim_{x\to 0}\frac{\sin x}{x}=\frac{1}{6}$.

(Here we had to apply L'Hôpital's rule repeatedly.)

(d) Consider

$$\lim_{x\to 0^+}\frac{e^{-1/x}}{x}. \tag{5.3.18}$$

Taking derivatives (even n times), one gets

$$\lim_{x\to 0^+}\frac{e^{-1/x}}{n!x^{n+1}}, \quad n=1,2,3,\ldots \text{ (always indeterminate!)}. \tag{5.3.19}$$

Thus the rule gives no results. In this case, however, a simple device helps (see Problem 5 below).

(e) $\lim_{n\to\infty} n^{1/n}$ does not have the form $\frac{0}{0}$ or $\frac{\infty}{\infty}$, so the rule does not apply directly. Instead we compute

$$\lim_{n\to\infty}\ln n^{1/n}=\lim_{n\to\infty}\frac{\ln n}{n}=0 \text{ (Example (a))}. \tag{5.3.20}$$

Hence

$$n^{1/n}=\exp\left(\ln n^{1/n}\right)\to\exp(0)=e^0=1 \tag{5.3.21}$$

by the continuity of exponential functions. The answer is then 1.

5.3.E: Problems on L ′ L′ Hôpital's Rule

Elementary differentiation formulas are assumed known.

Exercise $5.3.E.1$

Complete the proof of L'Hôpital's rule. Verify that the differentiability assumption may be replaced by continuity plus existence of finite or infinite (but not both together infinite) derivatives f' and g' on $G_{\neg p}$ (same proof).

Exercise $5.3.E.2$

Show that the rule fails for complex functions. See, however, Problems 3, 7, and 8.
[Hint: Take $p=0$ with

$$f(x)=x \text{ and } g(x)=x+x^{2} e^{i / x^{2}}=x+x^{2}\left(\cos \frac{1}{x^{2}}+i \cdot \sin \frac{1}{x^{2}}\right). \tag{5.3.E.1}$$

Then

$$\lim _{x \rightarrow 0} \frac{f(x)}{g(x)}=1, \text { though } \lim _{x \rightarrow 0} \frac{f^{\prime}(x)}{g^{\prime}(x)}=\lim _{x \rightarrow 0} \frac{1}{g^{\prime}(x)}=0. \tag{5.3.E.2}$$

Indeed, $g^{\prime}(x)-1=(2 x-2 i / x) e^{i / x^{2}}$. (Verify!) Hence

$$\left|g^{\prime}(x)\right|+1 \geq|2 x-2 i / x| \quad\left(\text {for }\left|e^{i / x^{2}}\right|=1\right), \tag{5.3.E.3}$$

so

$$\left|g^{\prime}(x)\right| \geq-1+\frac{2}{x}. \quad(\text { Why? }) \tag{5.3.E.4}$$

Deduce that

$$\left.\left|\frac{1}{g^{\prime}(x)}\right| \leq\left|\frac{x}{2-x}\right| \rightarrow 0 .\right] \tag{5.3.E.5}$$

Exercise $5.3.E.3$

Prove the "simplified rule of L' Hôpital" for real or complex functions $(\text{also for vector-valued } f \text{ and scalar-valued } g):$ If f and g are differentiable at p, with $g^{\prime}(p) \neq 0$ and $f(p)=g(p)=0$, then

$$\lim _{x \rightarrow p} \frac{f(x)}{g(x)}=\frac{f^{\prime}(p)}{g^{\prime}(p)}. \tag{5.3.E.6}$$

[Hint:

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(p)}{g(x)-g(p)}=\frac{\Delta f}{\Delta x} / \frac{\Delta g}{\Delta x} \rightarrow \frac{f^{\prime}(p)}{g^{\prime}(p)}. \tag{5.3.E.7}$$

Exercise $5.3.E.4$

Why does $\lim_{x\to+\infty}\frac{f(x)}{g(x)}$ not exist, though $\lim_{x\to+\infty}\frac{f'(x)}{g'(x)}$ does, in the following example? Verify and explain.

$$f(x)=e^{-2x}(\cos x+2\sin x),\quad g(x)=e^{-x}(\cos x+\sin x). \tag{5.3.E.8}$$

[Hint: g' vanishes many times in each $G_{+\infty}$. Use the Darboux property for the proof.]

Exercise $5.3.E.5$

Find $\lim_{x\to 0^+}\frac{e^{-1/x}}{x}$.
$[\text{Hint: Substitute } z=\frac{1}{x}\to+\infty.\text{ Then use the rule.}]$

Exercise $5.3.E.6$

Verify that the assumptions of L'Hôpital's rule hold, and find the following limits.
(a) $\lim_{x\to 0}\frac{e^x-e^{-x}}{\ln(e-x)+x-1}$;
(b) $\lim_{x\to 0}\frac{e^x-e^{-x}-2x}{x-\sin x}$;
(c) $\lim_{x\to 0}\frac{(1+x)^{1/x}-e}{x}$;
(d) $\lim_{x\to 0^+}(x^q\ln x), q>0$;
(e) $\lim_{x\to+\infty}(x^{-q}\ln x), q>0$;
(f) $\lim_{x\to 0^+}x^x$;
(g) $\lim_{x\to+\infty}(x^q a^{-x}), a>1, q>0$;
(h) $\lim_{x\to 0}\left(\frac{1}{x^2}-\operatorname{cotan}^2 x\right)$;
(i) $\lim_{x\to+\infty}\left(\frac{\pi}{2}-\arctan x\right)^{1/\ln x}$;
(j) $\lim_{x\to 0}\left(\frac{\sin x}{x}\right)^{1/(1-\cos x)}$.

Exercise $5.3.E.7$

Prove L'Hôpital's rule for $f: E^1\to E^n(C)$ and $g: E^1\to E^1$, with

$$\lim_{k\to p}|f(x)|=0=\lim_{x\to p}|g(x)|, p\in E^* \text{ and } r\in E^n, \tag{5.3.E.9}$$

leaving the other assumptions unchanged.
$[\text{Hint: Apply the rule to the components of } \frac{f}{g} \text{ (respectively, to } \left(\frac{f}{g}\right)_{\text{re}} \text{ and } \left(\frac{f}{g}\right)_{\text{im}}\text{)}.]$

Exercise $5.3.E.8$

Let f and g be complex and differentiable on $G_{\neg p}, p\in E^1$. Let

$$\lim_{x\to p}f(x)=\lim_{x\to p}g(x)=0, \lim_{x\to p}f'(x)=q, \text{ and } \lim_{x\to p}g'(x)=r\neq 0. \tag{5.3.E.10}$$

Prove that $\lim_{x\to p}\frac{f(x)}{g(x)}=\frac{q}{r}$.
[Hint:

$$\frac{f(x)}{g(x)}=\frac{f(x)}{x-p}\Big/\frac{g(x)}{x-p}. \tag{5.3.E.11}$$

Apply Problem 7 to find

$$\lim_{x \to p} \frac{f(x)}{x-p} \text{ and } \lim_{x \to p} \frac{g(x)}{x-p}.] \tag{5.3.E.12}$$

Exercise $5.3.E.*9$

Do Problem 8 for $f : E^1 \to C^n$ and $g : E^1 \to C$.

5.4: Complex and Vector-Valued Functions on E 1 E1

The theorems of §§2-3 fail for complex and vector-valued functions (see Problem 3 below and Problem 2 in §3). However, some analogues hold. In a sense, they even are stronger, for, unlike the previous theorems, they do not require the existence of a derivative on an entire interval $I \subseteq E^1$, but only on $I - Q$, where Q is a countable set, one contained in the range of a sequence, $Q \subseteq \{p_m\}$. (We henceforth presuppose §9 of Chapter 1.)

In the following theorem, due to N. Bourbaki, $g: E^1 \rightarrow E^*$ is extended real while f may also be complex or vector valued. We call it the finite increments law since it deals with "finite increments" $f(b) - f(a)$ and $g(b) - g(a)$. Roughly, it states that $|f'| \leq g'$ implies a similar inequality for increments.

Theorem 5.4.1 (finite increments law)

Let $f: E^1 \rightarrow E$ and $g: E^1 \rightarrow E^*$ be relatively continuous and finite on a closed interval $I = [a, b] \subseteq E^1$, and have derivatives with $|f'| \leq g'$, on $I - Q$ where $Q \subseteq \{p_1, p_2, \ldots, p_m, \ldots\}$. Then

$$|f(b) - f(a)| \leq g(b) - g(a). \tag{5.4.1}$$

The proof is somewhat laborious, but worthwhile. (At a first reading, one may omit it, however.) We outline some preliminary ideas.

Given any $x \in I$, suppose first that $x > p_m$ for at least one $p_m \in Q$. In this case, we put

$$Q(x) = \sum_{p_m < x} 2^{-m}; \tag{5.4.2}$$

here the summation is only over those m for which $p_m < x$. If, however, there are no $p_m \in Q$ with $p_m < x$, we put $Q(x) = 0$. Thus $Q(x)$ is defined for all $x \in I$. It gives an idea as to "how many" p_m (at which f may have no derivative) precede x. Note that $x < y$ implies $Q(x) \leq Q(y)$. (Why?) Also,

$$Q(x) \leq \sum_{m=1}^{\infty} 2^{-m} = 1. \tag{5.4.3}$$

Our plan is as follows. To prove (1), it suffices to show that for some fixed $K \in E^1$, we have

$$(\forall \varepsilon > 0) \quad |f(b) - f(a)| \leq g(b) - g(a) + K\varepsilon, \tag{5.4.4}$$

for then, letting $\varepsilon \rightarrow 0$, we obtain (1). We choose

$$K = b - a + Q(b), \text{ with } Q(x) \text{ as above.} \tag{5.4.5}$$

Temporarily fixing $\varepsilon > 0$, let us call a point $r \in I$ "good" iff

$$|f(r) - f(a)| \leq g(r) - g(a) + [r - a + Q(r)]\varepsilon \tag{5.4.6}$$

and "bad" otherwise. We shall show that b is "good." First, we prove a lemma.

Lemma 5.4.1

Every "good" point $r \in I (r < b)$ is followed by a whole interval $(r, s), r < s \leq b$, consisting of "good" points only.

Proof

First let $r \notin Q$, so by assumption, f and g have derivatives at r, with

$$|f'(r)| \leq g'(r). \tag{5.4.7}$$

Suppose $g'(r) < +\infty$. Then (treating g' as a right derivative) we can find $s > r$ $(s \leq b)$ such that, for all x in the interval (r, s), ,

$$\left| \frac{g(x) - g(r)}{x - r} - g'(r) \right| < \frac{\varepsilon}{2} \quad \text{(why?)}; \tag{5.4.8}$$

similarly for f. Multiplying by $x-r$, we get

$$|f(x)-f(r)-f'(r)(x-r)|<(x-r)\frac{\varepsilon}{2} \text{ and }$$
$$|g(x)-g(r)-g'(r)(x-r)|<(x-r)\frac{\varepsilon}{2},$$

and hence by the triangle inequality (explain!),

$$|f(x)-f(r)|\leq|f'(r)|(x-r)+(x-r)\frac{\varepsilon}{2} \tag{5.4.9}$$

and

$$g'(r)(x-r)+(x-r)\frac{\varepsilon}{2}<g(x)-g(r)+(x-r)\varepsilon. \tag{5.4.10}$$

Combining this with $|f'(r)|\leq g'(r)$, we obtain

$$|f(x)-f(r)|\leq g(x)-g(r)+(x-r)\varepsilon \text{ whenever } r<x<s. \tag{5.4.11}$$

Now as r is "good," it satisfies (2); hence, certainly, as $Q(r)\leq Q(x)$,

$$|f(r)-f(a)|\leq g(r)-g(a)+(r-a)\varepsilon+Q(x)\varepsilon \text{ whenever } r<x<s. \tag{5.4.12}$$

Adding this to (3) and using the triangle inequality again, we have

$$|f(x)-f(a)|\leq g(x)-g(a)+[x-a+Q(x)]\varepsilon \text{ for all } x\in(r,s). \tag{5.4.13}$$

By definition, this shows that each $x\in(r,s)$ is "good," as claimed. Thus the lemma is proved for the case $r\in I-Q$, with $g'(r)<+\infty$.

The cases $g'(r)=+\infty$ and $r\in Q$ are left as Problems 1 and 2. □

We now return to Theorem 1.

Proof of Theorem 1. Seeking a contradiction, suppose b is "bad," and let $B\neq\emptyset$ be the set of all "bad" points in $[a,b]$. Let

$$r=\inf B, \quad r\in[a,b]. \tag{5.4.14}$$

Then the interval $[a,r)$ can contain only "good" points, i.e., points x such that

$$|f(x)-f(a)|\leq g(x)-g(a)+[x-a+Q(x)]\varepsilon. \tag{5.4.15}$$

As $x<r$ implies $Q(x)\leq Q(r)$, we have

$$|f(x)-f(a)|\leq g(x)-g(a)+[x-a+Q(r)]\varepsilon \text{ for all } x\in[a,r). \tag{5.4.16}$$

Note that $[a,r)\neq\emptyset$, for by (2), a is certainly "good' (why?), and so Lemma 1 yields a whole interval $[a,s)$ of "good" points contained in $[a,r)$.

Letting $x\rightarrow r$ in (4) and using the continuity of f at r, we obtain (2). Thus r is "good" itself. Then, however, Lemma 1 yields a new interval (r,q) of "good" points. Hence $[a,q)$ has no "bad" points, and so q is a lower bound of the set B of "bad" points in I, contrary to $q>r=\operatorname{glb} B$. This contradiction shows that b must be "good," i.e.,

$$|f(b)-f(a)|\leq g(b)-g(a)+[b-a+Q(b)]\varepsilon. \tag{5.4.17}$$

Now, letting $\varepsilon\rightarrow 0$, we obtain formula (1), and all is proved. □

Corollary 5.4.1

If $f:E^1\rightarrow E$ is relatively continuous and finite on $I=[a,b]\subseteq E^1$, and has a derivative on $I-Q$, then there is a real M such that

$$|f(b)-f(a)|\leq M(b-a) \text{ and } M\leq\sup_{t\in I-Q}|f'(t)|. \tag{5.4.18}$$

Proof

Let

$$M_0 = \sup_{t \in I-Q} |f'(t)|. \tag{5.4.19}$$

If $M_0 < +\infty$, put $M = M_0 \geq |f'|$ on $I - Q$, and take $g(x) = Mx$ in Theorem 1. Then $g' = M \geq |f'|$ on $I - Q$, so formula (1) yields (5) since

$$g(b) - g(a) = Mb - Ma = M(b-a). \tag{5.4.20}$$

If, however, $M_0 = +\infty$, let

$$M = \left| \frac{f(b) - f(a)}{b - a} \right| < M_0. \tag{5.4.21}$$

Then (5) clearly is true. Thus the required M exists always. $\square$

Corollary 5.4.2

Let f be as in Corollary 1. Then f is constant on I iff $f' = 0$ on $I - Q$.

Proof

If $f' = 0$ on $I - Q$, then $M = 0$ in Corollary 1, so Corollary 1 yields, for any subinterval $[a, x] (x \in I), |f(x) - f(a)| \leq 0;$ i.e., $f(x) = f(a)$ for all $x \in I$. Thus f is constant on I.

Conversely, if so, then $f' = 0,$ even on all of I. $\square$

Corollary 5.4.3

Let $f, g : E^1 \rightarrow E$ be relatively continuous and finite on $I = [a, b]$, and differentiable on $I - Q$. Then $f - g$ is constant on I iff $f' = g'$ on $I - Q$.

Proof

Apply Corollary 2 to the function $f - g$. $\square$

We can now also strengthen parts (ii) and (iii) of Corollary 4 in §2.

Theorem 5.4.2

Let f be real and have the properties stated in Corollary 1. Then

(i) $f \uparrow$ on $I = [a, b]$ iff $f' \geq 0$ on $I - Q$; and

(ii) $f \downarrow$ on I iff $f' \leq 0$ on $I - Q$.

Proof

Let $f' \geq 0$ on $I - Q$. Fix any $x, y \in I (x < y)$ and define $g(t) = 0$ on E^1. Then $|g'| = 0 \leq f'$ on $I - Q$. Thus g and f satisfy Theorem 1 (with their roles reversed on I, and certainly on the subinterval $[x, y]$. Thus we have

$$f(y) - f(x) \geq |g(y) - g(x)| = 0, \text{ i.e., } f(y) \geq f(x) \text{ whenever } y > x \text{ in } I, \tag{5.4.22}$$

so $f \uparrow$ on I.

Conversely, if $f \uparrow$ on I, then for every $p \in I$, we must have $f'(p) \geq 0,$ for otherwise, by Lemma 1 of §2, f would decrease at p. Thus $f' \geq 0$, even on all of I, and (i) is proved. Assertion (ii) is proved similarly. $\square$

5.4.E: Problems on Complex and Vector-Valued Functions on E 1 E1

Exercise $5.4.E.1$

Do the case $g^{\prime}(r)=+\infty$ in Lemma 1.
[Hint: Show that there is $s>r$ with

$$g(x)-g(r) \geq\left(\left|f^{\prime}(r)\right|+1\right)(x-r) \geq|f(x)-f(r)| \text { for } x \in(r, s). \tag{5.4.E.1}$$

Such x are "good."]

Exercise $5.4.E.2$

Do the case $r=p_n \in Q$ in Lemma 1.
[Hint: Show by continuity that there is $s>r$ such that $(\forall x \in(r, s))$

$$|f(x)-f(r)|<\frac{\varepsilon}{2^{n+1}} \text { and }|g(x)-g(r)|<\frac{\varepsilon}{2^{n+1}}. \tag{5.4.E.2}$$

Show that all such x are "good" since $x>r=p_n$ implies

$$\left.2^{-n}+Q(r) \leq Q(x). \quad(\text { Why? })\right] \tag{5.4.E.3}$$

Exercise $5.4.E.3$

Show that Corollary 3 in §2 (hence also Theorem 2 in §2) fails for complex functions.
$\left[\right.$Hint: Let $f(x)=e^{x i}=\cos x+i \cdot \sin x$. Verify that $\left|f^{\prime}\right|=1$ yet $f(2 \pi)-f(0)=0.\left.\right]$

Exercise $5.4.E.4$

(i) Verify that all propositions of §4 hold also if $f^{\prime}$ and $g^{\prime}$ are only right derivatives on $I-Q$.
(ii) Do the same for left derivatives. (See footnote 2.)

Exercise $5.4.E.5$

(i) Prove that if $f: E^1 \rightarrow E$ is continuous and finite on $I=(a, b)$ and differentiable on $I-Q$, and if

$$\sup _{t \in I-Q}\left|f^{\prime}(t)\right|<+\infty, \tag{5.4.E.4}$$

then f is uniformly continuous on I.
(ii) Moreover, if E is complete (e. g. , $E=E^n$) , then $f\left(a^{+}\right)$ and $f\left(b^{-}\right)$ exist and are finite.
[Hints: (i) Use Corollary 1. (ii) See the "hint" to Problem 11 (iii) of Chapter 4, §8.]

Exercise $5.4.E.6$

Prove that if f is as in Theorem 2, with $f^{\prime} \geq 0$ on $I-Q$ and $f^{\prime}>0$ at some $p \in I$, then $f(a)<f(b)$. Do it also with $f^{\prime}$ treated as a right derivative (see Problem 4).

Exercise $5.4.E.7$

Let $f, g : E^1 \to E^1$ be relatively continuous on $I=[a,b]$ and have right derivatives f'_+ and g'_+ (finite or infinite, but not both infinite) on $I-Q$.
(i) Prove that if

$$mg'_+ \leq f'_+ \leq Mg'_+ \text{ on } I-Q \tag{5.4.E.5}$$

for some fixed $m, M \in E^1$, then

$$m[g(b)-g(a)] \leq f(b)-f(a) \leq M[g(b)-g(a)]. \tag{5.4.E.6}$$

$[\text{Hint: Apply Theorem 2 and Problem 4 to each of } Mg-f \text{ and } f-mg.]$
(ii) Hence prove that

$$m_0(b-a) \leq f(b)-f(a) \leq M_0(b-a), \tag{5.4.E.7}$$

where

$$m_0=\inf f'_+[I-Q] \text{ and } M_0=\sup f'_+[I-Q] \text{ in } E^*. \tag{5.4.E.8}$$

$[\text{Hint: Take } g(x)=x \text{ if } m_0 \in E^1 \text{ or } M_0 \in E^1. \text{ The infinite case is simple.}]$

Exercise $5.4.E.8$

(i) Let $f:(a,b) \to E$ be finite, continuous, with a right derivative on (a,b). Prove that $q=\lim_{x\to a^+} f'_+(x)$ exists (finite) iff

$$q=\lim_{x,y\to a^+} \frac{f(x)-f(y)}{x-y}, \tag{5.4.E.9}$$

i.e., iff

$$(\forall \varepsilon>0)(\exists c>a)(\forall x,y \in (a,c) | x \neq y) \quad \left|\frac{f(x)-f(y)}{x-y}-q\right|<\varepsilon. \tag{5.4.E.10}$$

[Hints: If so, let $y \to x^+$ (keeping x fixed) to obtain

$$(\forall x \in (a,c)) \quad |f'_+(x)-q| \leq \varepsilon. \quad \text{(Why?)} \tag{5.4.E.11}$$

Conversely, if $\lim_{x\to a^+} f'_+(x)=q$, then

$$(\forall \varepsilon>0)(\exists c>a)(\forall t \in (a,c)) \quad |f'_+(t)-q|<\varepsilon. \tag{5.4.E.12}$$

Put

$$M=\sup_{a<t<c} |f'_+(t)-q| \leq \varepsilon \quad (\text{why } \leq \varepsilon?) \tag{5.4.E.13}$$

and
$h(t)=f(t)-tq, \quad t \in (a,b)$.
Apply Corollary 1 and Problem 4 to h on the interval $[x,y] \subseteq (a,c)$, to get

$$|f(y)-f(x)-(y-x)q| \leq M(y-x) \leq \varepsilon(y-x). \tag{5.4.E.14}$$

Proceed.]
(ii) Prove similar statements for the cases $q=\pm\infty$ and $x \rightarrow b^{-}$.
[Hint: In case $q=\pm\infty$, use Problem 7 (ii) instead of Corollary 1.]

Exercise $5.4.E.9$

From Problem 8 deduce that if f is as indicated and if $f_{+}^{\prime}$ is left continuous at some $p \in(a, b)$, then f also has a left derivative at p.
If $f_{+}^{\prime}$ is also right continuous at p, then $f_{+}^{\prime}(p)=f_{-}^{\prime}(p)=f^{\prime}(p)$.
[Hint: Apply Problem 8 to (a, p) and (p, b).]

Exercise $5.4.E.10$

In Problem 8, prove that if, in addition, E is complete and if

$$q=\lim_{x \rightarrow a^{+}} f_{+}^{\prime}(x) \neq \pm\infty \quad \text{(finite)}, \tag{5.4.E.15}$$

then $f\left(a^{+}\right) \neq \pm\infty$ exists, and

$$\lim_{x \rightarrow a^{+}} \frac{f(x)-f\left(a^{+}\right)}{x-a}=q; \tag{5.4.E.16}$$

similarly in case $\lim_{x \rightarrow b^{-}} f_{+}^{\prime}(x)=r$.
If both exist, set $f(a)=f\left(a^{+}\right)$ and $f(b)=f\left(b^{-}\right)$. Show that then f becomes relatively continuous on $[a, b]$, with $f_{+}^{\prime}(a)=q$ and $f_{-}^{\prime}(b)=r$.
[Hint: If

$$\lim_{x \rightarrow a^{+}} f_{+}^{\prime}(x)=q \neq \pm\infty, \tag{5.4.E.17}$$

then $f_{+}^{\prime}$ is bounded on some subinterval $(a, c), a<c \leq b($ why?$)$, so f is uniformly continuous on (a, c), by Problem 5, and $f\left(a^{+}\right)$ exists. Let $y \rightarrow a^{+}$, as in the hint to Problem 8.]

Exercise $5.4.E.11$

Do Problem 9 in §2 for complex and vector-valued functions.
[Hint: Use Corollary 1 of §4.]

Exercise $5.4.E.12$

Continuing Problem 7, show that the equalities

$$m=\frac{f(b)-f(a)}{b-a}=M \tag{5.4.E.18}$$

hold iff f is linear, i.e., $f(x)=c x+d$ for some $c, d \in E^{1}$, and then $c=m=M$.

Exercise $5.4.E.13$

Let $f: E^{1} \rightarrow C$ be as in Corollary 1, with $f \neq 0$ on I. Let g be the real part of $f^{\prime} / f$.
(i) Prove that $|f| \uparrow$ on I iff $g \geq 0$ on $I-Q$.
(ii) Extend Problem 4 to this result.

Exercise $5.4.E.14$

Define $f: E^1 \rightarrow C$ by

$$f(x)=\begin{cases} x^2 e^{i/x}=x^2\left(\cos \frac{1}{x}+i \cdot \sin \frac{1}{x}\right) & \text{if } x>0, \text{ and} \\ 0 & \text{if } x \leq 0. \end{cases} \tag{5.4.E.19}$$

Show that f is differentiable on $I=(-1,1)$, yet $f'[I]$ is not a convex set in $E^2=C$ (thus there is no analogue to Theorem 4 of §2).

5.5: Antiderivatives (Primitives, Integrals)

Given $f : E^1 \rightarrow E,$ we often have to find a function F such that $F' = f$ on I, or at least on $I - Q$. We also require F to be relatively continuous and finite on I. This process is called antidifferentiation or integration.

Definition 1

We call $F : E^1 \rightarrow E$ a primitive, or antiderivative, or an indefinite integral, of f on I iff

(i) F is relatively continuous and finite on I, and

(ii) F is differentiable, with $F' = f$, on $I - Q$ at least.

We then write

$$F = \int f, \text{ or } F(x) = \int f(x) dx, \text{ on } I. \tag{5.5.1}$$

(The latter is classical notation.)

If such an F exists (which is not always the case), we shall say that $\int f$ exists on I, or that f has a primitive (or antiderivative) on I, or that f is primitively integrable (briefly integrable) on I.

If $F' = f$ on a set $B \subseteq I$, we say that $\int f$ is exact on B and call F an exact primitive on B. Thus if $Q = \emptyset$, $\int f$ is exact on all of I.

Note 1. Clearly, if $F' = f$, then also $(F + c)' = f$ for a finite constant c. Thus the notation $F = \int f$ is rather incomplete; it means that F is one of many primitives. We now show that all of them have the form $F + c$ (or $\int f + c$).

Theorem 5.5.1

If F and G are primitive to f on I, then $G - F$ is constant on I.

Proof

By assumption, F and G are relatively continuous and finite on I; hence so is $G - F$. Also, $F' = f$ on $I - Q$ and $G' = f$ on $I - P$. (Q and P are countable, but possibly $Q \neq P$.)

Hence both F' and G' equal f on $I - S$, where $S = P \cup Q$, and S is countable itself by Theorem 2 of Chapter 1, §9.

Thus by Corollary 3 in §4, $F' = G'$ on $I - S$ implies $G - F = c$ (constant) on each $[x, y] \subseteq I$; hence $G - F = c$ (or $G = F + c$) on I. $\square$

Definition 2

If $F = \int f$ on I and if $a, b \in I$ (where $a \leq b$ or $b \leq a$), we define

$$\int_a^b f = \int_a^b f(x) dx = F(b) - F(a), \text{ also written } F(x)\Big|_a^b. \tag{5.5.2}$$

This expression is called the definite integral of f from a to b.

The definite integral of f from a to b is independent of the particular choice of the primitive F for f, and thus unambiguous, for if G is another primitive, Theorem 1 yields $G = F + c,$ so

$$G(b) - G(a) = F(b) + c - [F(a) + c] = F(b) - F(a), \tag{5.5.3}$$

and it does not matter whether we take F or G.

Note that $\int_a^b f(x) dx$, or $\int_a^b f$, is a constant in the range space E (a vector if f is vector valued). The "x" in $\int_a^b f(x) dx$ is a "dummy variable" only, and it may be replaced by any other letter. Thus

$$\int_a^b f(x)dx = \int_a^b f(y)dy = F(b) - F(a). \tag{5.5.4}$$

On the other hand, the indefinite integral is a function: $F : E^1 \to E$.

Note 2. We may, however, vary a or b (or both) in (1). Thus, keeping a fixed and varying b, we can define a function

$$G(t) = \int_a^t f = F(t) - F(a), \quad t \in I. \tag{5.5.5}$$

Then $G' = F' = f$ on I, and $G(a) = F(a) - F(a) = 0$. Thus if $\int f$ exists on I, f has a (unique) primitive G on I such that $G(a) = 0$. (It is unique by Theorem 1. Why?)

Examples

(a) Let

$$f(x) = \frac{1}{x} \text{ and } F(x) = \ln|x|, \text{ with } F(0) = f(0) = 0. \tag{5.5.6}$$

Then $F' = f$ and $F = \int f$ on $(-\infty, 0)$ and on $(0, +\infty)$ but not on E^1, since F is discontinuous at 0, contrary to Definition 1. We compute

$$\int_1^2 f = \ln 2 - \ln 1 = \ln 2. \tag{5.5.7}$$

(b) On E^1, let

$$f(x) = \frac{|x|}{x} \text{ and } F(x) = |x|, \text{ with } f(0) = 1. \tag{5.5.8}$$

Here F is continuous and $F' = f$ on $E^1 - \{0\}$. Thus $F = \int f$ on E^1, exact on $E^1 - \{0\}$. Here $I = E^1, Q = \{0\}$.

We compute

$$\int_{-2}^2 f = F(2) - F(-2) = 2 - 2 = 0 \tag{5.5.9}$$

(even though f never vanishes on E^1).

Basic properties of integrals follow from those of derivatives. Thus we have the following.

Corollary 5.5.1 (linearity)

If $\int f$ and $\int g$ exist on I, so does $\int (pf + qg)$ for any scalars p, q (in the scalar field of E). Moreover, for any $a, b \in I$, we obtain

(i) $\int_a^b (pf + qg) = p\int_a^b f + q\int_a^b g$;

(ii) $\int_a^b (f \pm g) = \int_a^b f \pm \int_a^b g$; and

(iii) $\int_a^b pf = p\int_a^b f$.

Proof

By assumption, there are F and G such that

$$F' = f \text{ on } I - Q \text{ and } G' = g \text{ on } I - P. \tag{5.5.10}$$

Thus, setting $S = P \cup Q$ and $H = pF + qG$, we have

$$H' = pF' + qG' = pf + qg \text{ on } I - S, \tag{5.5.11}$$

with P, Q, and S countable. Also, $H = pF + qG$ is relatively continuous and finite on I, as are F and G.

Thus by definition, $H=\int(pf+qg)$ exists on I, and by (1),

$$\int_a^b (pf+qg)=H(b)-H(a)=pF(b)+qG(b)-pF(a)-qG(a)=p\int_a^b f+q\int_a^b g, \tag{5.5.12}$$

proving (i*).

With $p=1$ and $q=\pm 1,$ we obtain (ii*).

Taking $q=0,$ we get (iii*). $\square$

Corollary 5.5.2

If both $\int f$ and $\int|f|$ exist on $I=[a,b]$, then

$$\left|\int_a^b f\right| \leq \int_a^b |f|. \tag{5.5.13}$$

Proof

As before, let

$$F'=f \text{ and } G'=|f| \text{ on } I-S(S=Q\cup P, \text{ all countable}), \tag{5.5.14}$$

where F and G are relatively continuous and finite on I and $G=\int|f|$ is real. Also, $|F'|=|f|=G'$ on $I-S.$ Thus by Theorem 1 of §4,

$$|F(b)-F(a)| \leq G(b)-G(a)=\int_a^b |f|. \quad \square \tag{5.5.15}$$

Corollary 5.5.3

If $\int f$ exists on $I=[a,b]$, exact on $I-Q$, then

$$\left|\int_a^b f\right| \leq M(b-a) \tag{5.5.16}$$

for some real

$$M \leq \sup_{t\in I-Q} |f(t)|. \tag{5.5.17}$$

This is simply Corollary 1 of §4, when applied to a primitive, $F=\int f$

Corollary 5.5.4

If $F=\int f$ on I and $f=g$ on $I-Q$, then F is also a primitive of g, and

$$\int_a^b f=\int_a^b g \quad \text{for } a,b\in I. \tag{5.5.18}$$

(Thus we may arbitrarily redefine f on a countable Q.)

Proof

Let $F'=f$ on $I-P.$ Then $F'=g$ on $I-(P\cup Q).$ The rest is clear. $\square$

Corollary 5.5.5 (integration by parts)

Let f and g be real or complex (or let f be scalar valued and g vector valued), both relatively continuous on I and differentiable on $I-Q$. Then if $\int f'g$ exists on I, so does $\int fg'$, and we have

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g \quad \text{for any } a, b \in I. \tag{5.5.19}$$

Proof

By assumption, fg is relatively continuous and finite on I, and

$$(fg)' = fg' + f'g \text{ on } I - Q. \tag{5.5.20}$$

Thus, setting $H = fg$, we have $H = \int(fg' + f'g)$ on I. Hence by Corollary 1 if $\int f'g$ exists on I, so does $\int((fg'+f'g)-f'g) = \int fg'$, and

$$\int_a^b fg' + \int_a^b f'g = \int_a^b (fg' + f'g) = H(b) - H(a) = f(b)g(b) - f(a)g(a). \tag{5.5.21}$$

Thus (2) follows. □

The proof of the next three corollaries is left to the reader.

Corollary 5.5.6 (additivity of the integral)

If $\int f$ exists on I then, for $a, b, c \in I$, we have

(i) $\int_a^b f = \int_a^c f + \int_c^b f$;

(ii) $\int_a^a f = 0$; and

(iii) $\int_b^a f = -\int_a^b f$.

Corollary 5.5.7 (componentwise integration)

A function $f : E^1 \rightarrow E^n\ ({}^*C^n)$ is integrable on I iff all its components $(f_1, f_2, \ldots, f_n)$ are, and then by Theorem 5 in §1)

$$\int_a^b f = \left(\int_a^b f_1, \ldots, \int_a^b f_n\right) = \sum_{k=1}^n \vec{e}_k \int_a^b f_k \text{ for any } a, b \in I. \tag{5.5.22}$$

Hence if f is complex,

$$\int_a^b f = \int_a^b f_{\text{re}} + i \cdot \int_a^b f_{\text{im}} \tag{5.5.23}$$

(see Chapter 4, §3, Note 5).

Examples (continued)

(c) Define $f : E^1 \rightarrow E^3$ by

$$f(x) = (a \cdot \cos x, a \cdot \sin x, 2cx), \quad a, c \in E^1. \tag{5.5.24}$$

Verify that

$$\int_0^\pi f(x)dx = \left(a \cdot \sin x, -a \cdot \cos x, cx^2\right)\Big|_0^\pi = \left(0, 2a, c\pi^2\right) = 2a\vec{j} + c\pi^2\vec{k}. \tag{5.5.25}$$

(d) $\int_0^\pi e^{ix}dx = \int_0^\pi (\cos x + i \cdot \sin x)dx = (\sin x - i \cdot \cos x)\big|_0^\pi = 2i$.

Corollary 5.5.8

If $f=0$ on $I-Q$, then $\int f$ exists on I, and

$$\left|\int_{a}^{b} f\right|=\int_{a}^{b}|f|=0 \quad \text{for } a, b \in I. \tag{5.5.26}$$

Theorem 5.5.2 (change of variables)

Suppose $g: E^{1} \rightarrow E^{1}$ (real) is differentiable on I, while $f: E^{1} \rightarrow E$ has a primitive on $g[I]$, exact on $g[I-Q]$. Then

$$\int f(g(x)) g^{\prime}(x) d x \quad \left(i.e., \int(f \circ g) g^{\prime}\right) \tag{5.5.27}$$

exists on I, and for any $a, b \in I$, we have

$$\int_{a}^{b} f(g(x)) g^{\prime}(x) d x=\int_{p}^{q} f(y) d y, \text{ where } p=g(a) \text{ and } q=g(b). \tag{5.5.28}$$

Thus, using classical notation, we may substitute $y=g(x)$, provided that we also substitute $d y=g^{\prime}(x) d x$ and change the bounds of **integrals (3)**. Here we treat the expressions dy and $g^{\prime}(x) d x$ purely formally, without assigning them any separate meaning outside the context of the integrals.

Proof

Let $F=\int f$ on $g[I]$, and $F^{\prime}=f$ on $g[I-Q]$. Then the composite function $H=F \circ g$ is relatively continuous and finite on I. (Why?) By Theorem 3 of §1,

$$H^{\prime}(x)=F^{\prime}(g(x)) g^{\prime}(x) \text{ for } x \in I-Q; \tag{5.5.29}$$

i.e.,

$$H^{\prime}=\left(F^{\prime} \circ g\right) g^{\prime} \text{ on } I-Q. \tag{5.5.30}$$

Thus $H=\int(f \circ g) g^{\prime}$ exists on I, and

$$\int_{a}^{b}(f \circ g) g^{\prime}=H(b)-H(a)=F(g(b))-F(g(a))=F(q)-F(p)=\int_{p}^{q} f. \quad \square \tag{5.5.31}$$

Note 3. The theorem does not require that g be one to one on I, but if it is, then one can drop the assumption that $\int f$ is exact on $g[I-Q]$. (See Problem 4.)

Examples (continued)

(e) Find $\int_{0}^{\pi / 2} \sin ^{2} x \cdot \cos x d x$.

Here $f(y)=y^{2}, y=g(x)=\sin x, d y=\cos x d x, F(y)=y^{3} / 3, a=0, \quad b=\pi / 2, p=\sin 0=0,$ and $q=\sin (\pi / 2)=1,$ so **(3)** yields

$$\int_{0}^{\pi / 2} \sin ^{2} x \cdot \cos x d x=\int_{0}^{1} y^{2} d y=\left.\frac{y^{3}}{3}\right|_{0}^{1}=\frac{1}{3}-0=\frac{1}{3}. \tag{5.5.32}$$

For real functions, we obtain some inferences dealing with inequalities.

Theorem 5.5.3

If $f, g: E^1 \rightarrow E^1$ are integrable on $I=[a, b]$, then we have the following:

(i) $f \geq 0$ on $I-Q$ implies $\int_a^b f \geq 0$.

(i') $f \leq 0$ on $I-Q$ implies $\int_a^b f \leq 0$.

(ii) $f \geq g$ on $I-Q$ implies

$$\int_a^b f \geq \int_a^b g \text{ (dominance law). } \tag{5.5.33}$$

(iii) If $f \geq 0$ on $I-Q$ and $a \leq c \leq d \leq b$, then

$$\int_a^b f \geq \int_c^d f \text{ (monotonicity law). } \tag{5.5.34}$$

(iv) If $\int_a^b f=0$, and $f \geq 0$ on $I-Q$, then $f=0$ on some $I-P, P$ countable.

Proof

By Corollary 4, we may redefine f on Q so that our assumptions in (i)-(iv) hold on all of I. Thus we write "I" for "$I-Q$."

By assumption, $F=\int f$ and $G=\int g$ exist on I. Here F and G are relatively continuous and finite on $I=[a, b]$, with $F^{\prime}=f$ and $I-P$, for another countable set P (this P cannot be omitted). Now consider the cases (i)-(iv). (P is fixed henceforth.)

(i) Let $f \geq 0$ on I; i.e., $F^{\prime}=f \geq 0$ on $I-P$. Then by Theorem 2 in §4, $F \uparrow$ on $I=[a, b]$. Hence $F(a) \leq F(b)$, and so

$$\int_a^b f=F(b)-F(a) \geq 0. \tag{5.5.35}$$

One proves (i') similarly.

(ii) If $f-g \geq 0$, then by (i),

$$\int_a^b(f-g)=\int_a^b f-\int_a^b g \geq 0, \tag{5.5.36}$$

so $\int_a^b f \geq \int_a^b g$, as claimed.

(iii) Let $f \geq 0$ on I and $a \leq c \leq d \leq b$. Then by (i),

$$\int_a^c f \geq 0 \text{ and } \int_d^b f \geq 0. \tag{5.5.37}$$

Thus by Corollary 6,

$$\int_a^b f=\int_a^c f+\int_c^d f+\int_d^b f \geq \int_c^d f, \tag{5.5.38}$$

as asserted.

(iv) Seeking a contradiction, suppose $\int_a^b f=0, f \geq 0$ on I, yet $f(p)>0$ for some $p \in I-P$ (P as above), so $F^{\prime}(p)=f(p)>0$.

Now if $a \leq p<b$, Lemma 1 of §2 yields $F(c)>F(p)$ for some $c \in(p, b]$. Then by (iii),

$$\int_a^b f \geq \int_p^c f=F(c)-F(p)>0, \tag{5.5.39}$$

contrary to $\int_a^b f=0$; similarly in case $a<p \leq b$. $\square$

Note 4. Hence

$$\int_a^b |f| = 0 \text{ implies } f = 0 \text{ on } [a,b] - P \tag{5.5.40}$$

(P countable), even for vector-valued functions (for $|f|$ is always real, and so Theorem 3 applies).

However, $\int_a^b f = 0$ does not suffice, even for real functions (unless f is signconstant). For example,

$$\int_0^{2\pi} \sin x\, dx = 0, \text{ yet } \sin x \not\equiv 0 \text{ on any } I - P. \tag{5.5.41}$$

See also Example (b).

Corollary 5.5.9 (first law of the mean)

If f is real and $\int f$ exists on $[a,b]$, exact on (a,b), then

$$\int_a^b f = f(q)(b-a) \text{ for some } q \in (a,b). \tag{5.5.42}$$

Proof

Apply Corollary 3 in §2 to the function $F = \int f$. $\square$

Caution: Corollary 9 may fail if $\int f$ is inexact at some $p \in (a,b)$. (Exactness on $[a,b] - Q$ does not suffice, as it does not in Corollary 3 of §2, used here.) Thus in Example (b) above, $\int_{-2}^{2} f = 0$. Yet for no q is $f(q)(2+2) = 0$, since $f(q) = \pm 1$. The reason is that $\int f$ is inexact just at 0, an interior point of $[-2,2]$.

5.5.E: Problems on Antiderivatives

Exercise $5.5.E.1$

Prove in detail Corollaries $3, 4, 6, 7, 8,$and 9 and Theorem 3(i$'$) and (iv).

Exercise $5.5.E.2$

In Examples (a) and (b) discuss continuity and differentiability of f and F at 0. In (a) show that $\int f$ does not exist on any interval $(-a, a)$.
[Hint: Use Theorem 1.]

Exercise $5.5.E.3$

Show that Theorem 2 holds also if g is relatively continuous on I and differentiable on $I-Q$.

Exercise $5.5.E.4$

Under the assumptions of Theorem 2, show that if g is one to one on I, then automatically $\int f$ is exact on $g[I-Q] (Q \text{ countable})$.
[Hint: If $F=\int f$ on $g[I]$, then

$$F^{\prime}=f \text{ on } g[I]-P, P \text{ countable.} \tag{5.5.E.1}$$

Let $Q=g^{-1}[P]$. Use Problem 6 of Chapter $1, §§4-7$ and Problem 2 of Chapter 1 §9 to show that Q is countable and $g[I]-P=g[I-Q]$.

Exercise $5.5.E.5$

Prove Corollary 5 for dot products $f \cdot g$ of vector-valued functions.

Exercise $5.5.E.6$

Prove that if $\int f$ exists on $[a, p]$ and $[p, b]$, then it exists on $[a, b]$. By induction, extend this to unions of n adjacent intervals.
[Hint: Choose $F=\int f$ on $[a, p]$ and $G=\int f$ on $[p, b]$ such that $F(p)=G(p)$. (Why do such F, G exist?) Then construct a primitive $H=\int f$ that is relatively continuous on *all* of $[a, b]$.]

Exercise $5.5.E.7$

Prove the weighted law of the mean: If g is real and nonnegative on $I=[a, b]$, and if $\int g$ and $\int g f$ exist on I for some $f: E^{1} \rightarrow E,$ then there is a finite $c \in E$ with

$$\int_{a}^{b} g f=c \int_{a}^{b} g. \tag{5.5.E.2}$$

(The value c is called a g-weighted mean of f.)
[Hint: If $\int_{a}^{b} g>0,$ put

$$c=\int_{a}^{b} g f / \int_{a}^{b} g. \tag{5.5.E.3}$$

If $\int_a^b g=0,$ use Theorem 3(i) and (iv) to show that also $\int_a^b gf=0,$ so any c will do.]

? Exercise 5.5.E.8

In Problem 7, prove that if, in addition, f is real and has the Darboux property on I, then $c=f(q)$ for some $q \in I$ (the second law of the mean).

[Hint: Choose c as in Problem 7. If $\int_a^b g>0,$ put

$$m=\inf f[I] \text{ and } M=\sup f[I], \text{ in } E^{*}, \tag{5.5.E.4}$$

so $m \leq f \leq M$ on I. Deduce that

$$m \int_a^b g \leq \int_a^b gf \leq M \int_a^b g, \tag{5.5.E.5}$$

whence $m \leq c \leq M$.

If $m<c<M,$ then $f(x)<c<f(y)$ for some $x, y \in I$ (why?), so the Darboux property applies.

If $c=m,$ then $g \cdot (f-c) \geq 0$ and Theorem 3(iv) yields $gf=gc$ on $I-P$. (Why?) Deduce that $f(q)=c$ if $g(q) \neq 0$ and $q \in I-P$. (Why does such a q exist?)

What if $c=M$?]

? Exercise 5.5.E.9

Taking $g(x) \equiv 1$ in Problem 8, obtain a new version of Corollary 9. State it precisely!

? Exercise 5.5.E.10

$\Rightarrow$ 10. Prove that if $F=\int f$ on $I=(a, b)$ and f is right (left) continuous and finite at $p \in I$, then

$$f(p)=F_{+}^{\prime}(p) \text{ (respectively, } F_{-}^{\prime}(p)). \tag{5.5.E.6}$$

Deduce that if f is continuous and finite on I, all its primitives on I are exact on I.

[Hint: Fix $\varepsilon>0$. If f is right continuous at p, there is $c \in I(c>p)$, with

$$|f(x)-f(p)|<\varepsilon \text{ for } x \in[p, c). \tag{5.5.E.7}$$

Fix such an x. Let

$$G(t)=F(t)-t f(p), \quad t \in E^{1}. \tag{5.5.E.8}$$

Deduce that $G^{\prime}(t)=f(t)-f(p)$ for $t \in I-Q$.

By Corollary 1 of §4,

$$|G(x)-G(p)|=|F(x)-F(p)-(x-p) f(p)| \leq M(x-p), \tag{5.5.E.9}$$

with $M \leq \varepsilon$. (Why?) Hence

$$\left|\frac{\Delta F}{\Delta x}-f(p)\right| \leq \varepsilon \text{ for } x \in[p, c), \tag{5.5.E.10}$$

and so

$$\lim_{x \to p^+} \frac{\Delta F}{\Delta x} = f(p) \quad (\text{ why? }); \tag{5.5.E.11}$$

similarly for a left-continuous f.]

Exercise 5.5.E.11

State and solve Problem 10 for the case $I = [a, b]$.

Exercise 5.5.E.12

(i) Prove that if f is constant $(f = c \neq \pm\infty)$ on $I - Q$, then

$$\int_a^b f = (b - a)c \quad \text{for } a, b \in I. \tag{5.5.E.12}$$

(ii) Hence prove that if $f = c_k \neq \pm\infty$ on

$$I_k = [a_k, a_{k+1}), \quad a = a_0 < a_1 < \cdots < a_n = b, \tag{5.5.E.13}$$

then $\int f$ exists on $[a, b]$, and

$$\int_a^b f = \sum_{k=0}^{n-1} (a_{k+1} - a_k) c_k. \tag{5.5.E.14}$$

Show that this is true also if $f = c_k \neq \pm\infty$ on $I_k - Q_k$.
[Hint: Use Problem 6.]

Exercise 5.5.E.13

Prove that if $\int f$ exists on each $I_n = [a_n, b_n]$, where

$$a_{n+1} \leq a_n \leq b_n \leq b_{n+1}, \quad n = 1, 2, \ldots, \tag{5.5.E.15}$$

then $\int f$ exists on

$$I = \bigcup_{n=1}^{\infty} [a_n, b_n], \tag{5.5.E.16}$$

itself an interval with endpoints $a = \inf a_n$ and $b = \sup b_n$, $a, b \in E^*$.
[Hint: Fix some $c \in I_1$. Define

$$H_n(t) = \int_c^t f \text{ on } I_n, n = 1, 2, \ldots. \tag{5.5.E.17}$$

Prove that

$$(\forall n \leq m) \quad H_n = H_m \text{ on } I_n \,(\text{ since } \{I_n\} \uparrow). \tag{5.5.E.18}$$

Thus $H_n(t)$ is the same for all n such that $t \in I_n$, so we may simply write H for H_n on $I = \bigcup_{n=1}^{\infty} I_n$. Show that $H = \int f$ on all of I; verify that I is, indeed, an interval.]

Exercise $5.5.E.14$

Continuing Problem 13, prove that $\int f$ exists on an interval I iff it exists on each closed subinterval $[a, b] \subseteq I$.
[Hint: Show that each I is the union of an expanding sequence $I_n = [a_n, b_n]$. For example, if $I = (a, b), a, b \in E^1,$ put

$$a_n = a + \frac{1}{n} \text{ and } b_n = b - \frac{1}{n} \text{ for large } n \text{ (how large?)}, \tag{5.5.E.19}$$

and show that

$$\left. I = \bigcup_n [a_n, b_n] \text{ over such } n. \right] \tag{5.5.E.20}$$

5.6: Differentials. Taylor's Theorem and Taylor's Series

Recall (Theorem 2 of §1) that a function f is differentiable at p iff

$$\Delta f=f^{\prime}(p) \Delta x+\delta(x) \Delta x, \tag{5.6.1}$$

with $\lim_{x \rightarrow p} \delta(x)=\delta(p)=0$. It is customary to write $d f$ for $f^{\prime}(p) \Delta x$ and $o(\Delta x)$ for $\delta(x) \Delta x ; d f$ is called the differential of f (at p and x). Thus

$$\Delta f=d f+o(\Delta x); \tag{5.6.2}$$

i.e., df approximates Δf to within $o(\Delta x)$.

More generally, given any function $f: E^{1} \rightarrow E$ and $p, x \in E^{1}$, we define

$$d^{n} f=d^{n} f(p, x)=f^{(n)}(p)(x-p)^{n}, \quad n=0,1,2, \ldots, \tag{5.6.3}$$

where $f^{(n)}$ is the n th derived function (Definition 2 in §1); $d^{n} f$ is called the nth differential, or differential of order n, of f (at p and x). In particular, $d^{1} f=f^{\prime}(p) \Delta x=d f$. By our conventions, $d^{n} f$ is always defined, as is $f^{(n)}$.

As we shall see, good approximations of Δf (suggested by Taylor) can often be obtained by using higher differentials (1), as follows:

$$\Delta f=d f+\frac{d^{2} f}{2 !}+\frac{d^{3} f}{3 !}+\cdots+\frac{d^{n} f}{n !}+R_{n}, \quad n=1,2,3, \ldots, \tag{5.6.4}$$

where

$$R_{n}=\Delta f-\sum_{k=1}^{n} \frac{d^{k} f}{k !} \quad (\text{the "remainder term"}) \tag{5.6.5}$$

is the error of the approximation. Substituting the values of Δf and $d^{k} f$ and transposing $f(p)$, we have

$$f(x)=f(p)+\frac{f^{\prime}(p)}{1 !}(x-p)+\frac{f^{\prime \prime}(p)}{2 !}(x-p)^{2}+\cdots+\frac{f^{(n)}(p)}{n !}(x-p)^{n}+R_{n}. \tag{5.6.6}$$

Formula (3) is known as the nth Taylor expansion of f about p (with remainder term R_{n} to be estimated). Usually we treat p as fixed and x as variable. Writing $R_{n}(x)$ for R_{n} and setting

$$P_{n}(x)=\sum_{k=0}^{n} \frac{f^{(k)}(p)}{k !}(x-p)^{k}, \tag{5.6.7}$$

we have

$$f(x)=P_{n}(x)+R_{n}(x). \tag{5.6.8}$$

The function $P_{n}: E^{1} \rightarrow E$ so defined is called the nth Taylor polynomial for f about p. Thus (3) yields approximations of f by polynomials $P_{n}, n=1,2,3, \ldots$. This is one way of interpreting it. The other (easy to remember) one is (2), which gives approximations of Δf by the $d^{k} f$. It remains, however, to find a good estimate for R_{n}. We do it next.

Theorem 5.6.1 (taylor)

Let the function $f: E^{1} \rightarrow E$ and its first n derived functions be relatively continuous and finite on an interval I and differentiable on $I-Q$ (Q countable). Let $p, x \in I$. Then formulas (2) and (3) hold, with

$$R_{n}=\frac{1}{n !} \int_{p}^{x} f^{(n+1)}(t) \cdot(x-t)^{n} d t \quad (\text{"integral form of } R_{n}\text{"}) \tag{5.6.9}$$

and

$$\left|R_{n}\right| \leq M_{n} \frac{|x-p|^{n+1}}{(n+1) !} \text { for some real } M_{n} \leq \sup_{t \in I-Q}\left|f^{(n+1)}(t)\right|. \tag{5.6.10}$$

Proof

By definition, $R_n = f - P_n,$ or

$$R_n = f(x) - f(p) - \sum_{k=1}^{n} f^{(k)}(p) \frac{(x-p)^k}{k!}. \tag{5.6.11}$$

We use the right side as a "pattern" to define a function $h: E^1 \rightarrow E$. This time, we keep x fixed $(\text{ say }, x = a \in I)$ and replace p by a variable t. Thus we set

$$h(t) = f(a) - f(t) - \frac{f'(t)}{1!}(a-t) - \cdots - \frac{f^{(n)}(t)}{n!}(a-t)^n \text{ for all } t \in E^1. \tag{5.6.12}$$

Then $h(p) = R_n$ and $h(a) = 0$. Our assumptions imply that h is relatively continuous and finite on I, and differentiable on $I - Q$. Differentiating (4), we see that all cancels out except for one term

$$h'(t) = -f^{(n+1)}(t) \frac{(a-t)^n}{n!}, \quad t \in I - Q. \quad (\text{Verify!}) \tag{5.6.13}$$

Hence by Definitions 1 and 2 of §5,

$$-h(t) = \frac{1}{n!} \int_t^a f^{(n+1)}(s)(a-s)^n ds \quad \text{ on } I \tag{5.6.14}$$

and

$$\frac{1}{n!} \int_p^a f^{(n+1)}(t)(a-t)^n dt = -h(a) + h(p) = 0 + R_n = R_n \quad (\text{for } h(p) = R_n). \tag{5.6.15}$$

As $x = a,$ (3') is proved.

Next, let

$$M = \sup_{t \in I-Q} \left|f^{(n+1)}(t)\right|. \tag{5.6.16}$$

If $M < +\infty,$ define

$$g(t) = M \frac{(t-a)^{n+1}}{(n+1)!} \text{ for } t \geq a \text{ and } g(t) = -M \frac{(a-t)^{n+1}}{(n+1)!} \text{ for } t \leq a. \tag{5.6.17}$$

In both cases,

$$g'(t) = M \frac{|a-t|^n}{n!} \geq |h'(t)| \text{ on } I - Q \text{ by } (5). \tag{5.6.18}$$

Hence, applying Theorem 1 in §4 to the functions h and g on the interval $[a, p]$ (or $[p, a]$), we get

$$|h(p) - h(a)| \leq |g(p) - g(a)|, \tag{5.6.19}$$

or

$$|R_n - 0| \leq M \frac{|a-p|^{n+1}}{(n+1)!}. \tag{5.6.20}$$

Thus (3") follows, with $M_n = M$.

Finally, if $M = +\infty,$ we put

$$M_n = |R_n| \frac{(n+1)!}{|a-p|^{n+1}} < M. \quad \square \tag{5.6.21}$$

For real functions, we obtain some additional estimates of R_n.

Theorem 5.6.1′

If f is real and $n+1$ times differentiable on I, then for $p \neq x$ $(p, x \in I)$, there are q_n, q_n' in the interval (p, x) (respectively, (x, p)) such that

$$R_n = \frac{f^{(n+1)}(q_n)}{(n+1)!}(x-p)^{n+1} \tag{5.6.22}$$

and

$$R_n = \frac{f^{(n+1)}(q_n')}{n!}(x-p)(x-q_n')^n. \tag{5.6.23}$$

(Formulas (5') and (5") are known as the Lagrange and Cauchy forms of R_n, respectively.)

Proof

Exactly as in the proof of Theorem 1, we obtain the function h and formula (5). By our present assumptions, h is differentiable (hence continuous) on I, so we may apply to it Cauchy's law of the mean (Theorem 2 of §2) on the interval $[a, p]$ (or $[p, a]$ if $p < a$), where $a = x \in I$.

For this purpose, we shall associate h with another suitable function g (to be specified later). Then by Theorem 2 of §2, there is a real $q \in (a, p)$ (respectively, $q \in (p, a)$) such that

$$g'(q)[h(a)-h(p)] = h'(q)[g(a)-g(p)]. \tag{5.6.24}$$

Here by the previous proof, $h(a)=0, h(p)=R_n,$ and

$$h'(q) = -\frac{f^{(n+1)}}{n!}(a-q)^n. \tag{5.6.25}$$

Thus

$$g'(q)\cdot R_n = \frac{f^{(n+1)}(q)}{n!}(a-q)^n[g(a)-g(p)]. \tag{5.6.26}$$

Now define g by

$$g(t) = a-t, \quad t \in E^1. \tag{5.6.27}$$

Then

$$g(a)-g(p) = -(a-p) \text{ and } g'(q) = -1, \tag{5.6.28}$$

so (6) yields (5") (with $q_n' = q$ and $a = x$) .

Similarly, setting $g(t) = (a-t)^{n+1}$, we obtain (5'). (Verify!) Thus all is proved. □

Note 1. In (5') and (5"), the numbers q_n and q_n' depend on n and are different in general $(q_n \neq q_n')$, for they depend on the choice of the function g. Since they are between p and x, they may be written as

$$q_n = p+\theta_n(x-p) \text{ and } q_n' = p+\theta_n'(x-p), \tag{5.6.29}$$

where $0 < \theta_n < 1$ and $0 < \theta_n' < 1$. (Explain!)

Note 2. For any function $f: E^1 \rightarrow E,$ the Taylor polynomials P_n are partial sums of a power series, called the Taylor series for f (about p). We say that f admits such a series on a set B iff the series converges to f on B; i.e.,

$$f(x) = \lim_{n\to\infty} P_n(x) = \sum_{n=1}^{\infty} \frac{f^{(n)}(p)}{n!}(x-p)^n \neq \pm\infty \text{ for } x \in B. \tag{5.6.30}$$

This is clearly the case iff

$$\lim_{n\to\infty} R_n(x) = \lim_{n\to\infty}[f(x)-P_n(x)] = 0 \text{ for } x \in B; \tag{5.6.31}$$

briefly, $R_n \rightarrow 0$. Thus

$$f \text{ admits a Taylor series (about p) iff } R_n \rightarrow 0. \tag{5.6.32}$$

Caution: The convergence of the series alone (be it pointwise or uniform) does not suffice. Sometimes the series converges to a sum other than $f(x)$; then (7) fails. Thus all depends on the necessary and sufficient condition: $R_n \rightarrow 0$.

Definition 1

We say that f is of class CD^n, or continuously differentiable n times, on a set B iff f is n times differentiable on B, and $f^{(n)}$ is relatively continuous on B. Notation: $f \in \mathrm{CD}^n$ (on B).

If this holds for each $n \in N$, we say that f is infinitely differentiable on B and write $f \in \mathrm{CD}^\infty$ (on B).

The notation $f \in \mathrm{CD}^0$ means that f is finite and relatively continuous (all on B).

Examples

(a) Let

$$f(x) = e^x \text{ on } E^1. \tag{5.6.33}$$

Then

$$(\forall n) \quad f^{(n)}(x) = e^x, \tag{5.6.34}$$

so $f \in \mathrm{CD}^\infty$ on E^1. At $p = 0$, $f^{(n)}(p) = 1$, so we obtain by Theorem 1' (using (5') and Note 1)

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{e^{\theta_n x}}{(n+1)!} x^{n+1}, \quad 0 < \theta_n < 1. \tag{5.6.35}$$

Thus on an interval $[-a, a]$,

$$e^x \approx 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} \tag{5.6.36}$$

to within an error $R_n (> 0 \text{ if } x > 0)$ with

$$|R_n| < e^a \frac{a^{n+1}}{(n+1)!}, \tag{5.6.37}$$

which tends to 0 as $n \rightarrow +\infty$. For $a = 1 = x$, we get

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + R_n \text{ with } 0 < R_n < \frac{e^1}{(n+1)!}. \tag{5.6.38}$$

Taking $n = 10$, we have

$$e \approx 2.7182818|011463845\ldots \tag{5.6.39}$$

with a nonnegative error of no more than

$$\frac{e}{11!} = 0.00000006809869\ldots; \tag{5.6.40}$$

all digits are correct before the vertical bar.

(b) Let

$$f(x) = e^{-1/x^2} \text{ with } f(0) = 0. \tag{5.6.41}$$

As $\lim_{x \rightarrow 0} f(x) = 0 = f(0)$, f is continuous at 0. We now show that $f \in \mathrm{CD}^\infty$ on E^1.

For $x \neq 0$, this is clear; moreover, induction yields

$$f^{(n)}(x) = e^{-1/x^2} x^{-3n} S_n(x), \tag{5.6.42}$$

where S_n is a polynomial in x of degree $2(n-1)$ (this is all we need know about S_n). A repeated application of L'Hôpital's rule then shows that

$$\lim_{x \rightarrow 0} f^{(n)}(x)=0 \text { for each } n. \tag{5.6.43}$$

To find $f'(0)$, we have to use the definition of a derivative:

$$f'(0)=\lim_{x \rightarrow 0} \frac{f(x)-f(0)}{x-0}, \tag{5.6.44}$$

or by L'Hôpital's rule,

$$f'(0)=\lim_{x \rightarrow 0} \frac{f'(x)}{1}=0. \tag{5.6.45}$$

Using induction again, we get

$$f^{(n)}(0)=0, \quad n=1,2, \ldots. \tag{5.6.46}$$

Thus, indeed, f has finite derivatives of all orders at each $x \in E^1$, including $x=0$, so $f \in \mathrm{CD}^{\infty}$ on E^1, as claimed.

Nevertheless, any attempt to use formula (3) at $p=0$ yields nothing. As all $f^{(n)}$ vanish at 0, so do all terms except R_n. Thus no approximation by polynomials results - we only get $P_n=0$ on E^1 and $R_n(x)=e^{-1/x^2}$. R_n does not tend to 0 except at $x=0$, so f admits no Taylor series about 0 (except on $E=\{0\}$).

Taylor's theorem also yields sufficient conditions for maxima and minima, as we see in the following theorem.

Theorem 5.6.2

Let $f: E^1 \rightarrow E^*$ be of class CD^n on $G_p(\delta)$ for an even number $n \geq 2$, and let

$$f^{(k)}(p)=0 \text { for } k=1,2, \ldots, n-1, \tag{5.6.47}$$

while

$$f^{(n)}(p)<0 \text { (respectively, } f^{(n)}(p)>0). \tag{5.6.48}$$

Then $f(p)$ is the maximum (respectively, minimum) value of f on some $G_p(\varepsilon)$ $\varepsilon \leq \delta$.

If, however, these conditions hold for some odd $n \geq 1$ (i.e., the first nonvanishing derivative at p is of odd order), f has no maximum or minimum at p.

Proof

As

$$f^{(k)}(p)=0, \quad k=1,2, \ldots, n-1, \tag{5.6.49}$$

Theorem 1' (with n replaced by $n-1$) yields

$$f(x)=f(p)+f^{(n)}\left(q_n\right) \frac{(x-p)^n}{n!} \quad \text { for all } x \in G_p(\delta), \tag{5.6.50}$$

with q_n between x and p.

Also, as $f \in \mathrm{CD}^n, f^{(n)}$ is continuous at p. Thus if $f^{(n)}(p)<0$, then $f^{(n)}<0$ on some $G_p(\varepsilon), 0<\varepsilon \leq \delta$. However, $x \in G_p(\varepsilon)$ implies $q_n \in G_p(\varepsilon)$, so

$$f^{(n)}\left(q_n\right)<0, \tag{5.6.51}$$

while

$$(x-p)^n \geq 0 \text { if } n \text { is even. } \tag{5.6.52}$$

It follows that

$$f^{(n)}(q_n)\frac{(x-p)^n}{n!} \leq 0, \tag{5.6.53}$$

and so

$$f(x)=f(p)+f^{(n)}(q_n)\frac{(x-p)^n}{n!} \leq f(p) \quad \text{for } x \in G_p(\varepsilon), \tag{5.6.54}$$

i.e., $f(p)$ is the maximum value of f on $G_p(\varepsilon)$, as claimed.

Similarly, in the case $f^{(n)}(p)>0$, a minimum would result.

If, however, n is odd, then $(x-p)^n$ is negative for $x<p$ but positive for $x>p$. The same argument then shows that $f(x)<f(p)$ on one side of p and $f(x)>f(p)$ on the other side; thus no local maximum or minimum can exist at p. This completes the proof. □

Examples

(a') Let

$$f(x)=x^2 \text{ on } E^1 \text{ and } p=0. \tag{5.6.55}$$

Then

$$f'(x)=2x \text{ and } f''(x)=2>0, \tag{5.6.56}$$

so

$$f'(0)=0 \text{ and } f''(0)=2>0. \tag{5.6.57}$$

By Theorem 2, $f(p)=0^2=0$ is a minimum value.

It turns out to be a minimum on all of E^1. Indeed, $f'(x)>0$ for $x>0$, and $f'<0$ for $x<0$, so f strictly decreases on $(-\infty,0)$ and increases on $(0,+\infty)$.

Actually, even without using Theorem 2, the last argument yields the answer.

(b') Let

$$f(x)=\ln x \text{ on } (0,+\infty). \tag{5.6.58}$$

Then

$$f'(x)=\frac{1}{x}>0 \text{ on all of } (0,+\infty). \tag{5.6.59}$$

This shows that f strictly increases everywhere and hence can have no maximum or minimum anywhere. The same follows by the second part of Theorem 2, with $n=1$.

(b'') In Example (b'), consider also

$$f''(x)=-\frac{1}{x^2}<0. \tag{5.6.60}$$

In this case, f'' has no bearing on the existence of a maximum or minimum because $f'\neq 0$. Still, the formula $f''<0$ does have a certain meaning. In fact, if $f''(p)<0$ and $f\in \mathrm{CD}^2$ on $G_p(\delta)$, then (using the same argument as in Theorem 2) the reader will easily find that

$$f(x) \leq f(p)+f'(p)(x-p) \quad \text{for } x \text{ in some } G_p(\varepsilon), 0<\varepsilon\leq\delta. \tag{5.6.61}$$

since $y=f(p)+f'(p)(x-p)$ is the equation of the tangent at p, it follows that $f(x)\leq y$; i.e., near p the curve lies below the tangent at p.

Similarly, $f''(p)>0$ and $f\in \mathrm{CD}^2$ on $G_p(\delta)$ implies that the curve near p lies above the tangent.

5.6.E: Problems on Tayior's Theorem

Exercise $5.6.E.1$

Complete the proofs of Theorems $1, 1'$, and 2.

Exercise $5.6.E.2$

Verify Note 1 and Examples (b) and $\left(\mathrm{b}''\right)$.

Exercise $5.6.E.3$

Taking $g(t)=(a-t)^s, s>0$, in (6), find

$$R_n=\frac{f^{(n+1)}(q)}{n!s}(x-p)^s(x-q)^{n+1-s} \quad (\text{ Schloemilch-Roche remainder }). \tag{5.6.E.1}$$

Obtain $(5')$ and $(5'')$ from it.

Exercise $5.6.E.4$

Prove that P_n (as defined) is the only polynomial of degree n such that

$$f^{(k)}(p)=P_n^{(k)}(p), \quad k=0,1,\ldots,n. \tag{5.6.E.2}$$

[Hint: Differentiate $P_n n$ times to verify that it satisfies this property.
For uniqueness, suppose this also holds for

$$P(x)=\sum_{k=0}^n a_k(x-p)^k. \tag{5.6.E.3}$$

Differentiate Pn times to show that

$$P^{(k)}(p)=f^{(k)}(p)=a_k k!, \tag{5.6.E.4}$$

so $P=P_n$. (Why?)]

Exercise $5.6.E.5$

With P_n as defined, prove that if f is n times differentiable at p, then

$$f(x)-P_n(x)=o\left((x-p)^n\right) \text{ as } x \rightarrow p \tag{5.6.E.5}$$

(Taylor's theorem with Peano remainder term).
[Hint: Let $R(x)=f(x)-P_n(x)$ and

$$\delta(x)=\frac{R(x)}{(x-p)^n} \text{ with } \delta(p)=0. \tag{5.6.E.6}$$

Using the "simplified" L'Hôpital rule (Problem 3 in §3) repeatedly n times, prove that $\lim_{x \rightarrow p} \delta(x)=0.$]

Exercise 5.6.E.6

Use Theorem $1'$ with $p=0$ to verify the following expansions, and prove that $\lim_{n\to\infty} R_n = 0$.

(a) $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots - \frac{(-1)^m x^{2m-1}}{(2m-1)!} + \frac{(-1)^m x^{2m+1}}{(2m+1)!} \cos\theta_m x$
for all $x \in E^1$;

(b) $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + \frac{(-1)^m x^{2m}}{(2m)!} - \frac{(-1)^m x^{2m+2}}{(2m+2)!} \sin\theta_m x$ for
all $x \in E^1$.

[Hints: Let $f(x) = \sin x$ and $g(x) = \cos x$. Induction shows that

$$f^{(n)}(x) = \sin\left(x + \frac{n\pi}{2}\right) \text{ and } g^{(n)}(x) = \cos\left(x + \frac{n\pi}{2}\right). \tag{5.6.E.7}$$

Using formula $(5')$, prove that

$$|R_n(x)| \leq \left|\frac{x^{n+1}}{(n+1)!}\right| \to 0. \tag{5.6.E.8}$$

Indeed, $x^n/n!$ is the general term of a convergent series

$$\sum \frac{x^n}{n!} \quad \left(\text{see Chapter } 4, \S 13, \text{ Example (d)}\right). \tag{5.6.E.9}$$

Thus $x^n/n! \to 0$ by Theorem 4 of the same section.]

Exercise 5.6.E.7

For any $s \in E^1$ and $n \in N$, define

$$\binom{s}{n} = \frac{s(s-1)\cdots(s-n+1)}{n!} \text{ with } \binom{s}{0} = 1. \tag{5.6.E.10}$$

Then prove the following.

(i) $\lim_{n\to\infty} n\binom{s}{n} = 0$ if $s > 0$,

(ii) $\lim_{n\to\infty} \binom{s}{n} = 0$ if $s > -1$,

(iii) For any fixed $s \in E^1$ and $x \in (-1,1)$.

$$\lim_{n\to\infty} \binom{s}{n} n x^n = 0; \tag{5.6.E.11}$$

hence

$$\lim_{n\to\infty} \binom{s}{n} x^n = 0. \tag{5.6.E.12}$$

$\Big[$Hints: (i) Let $a_n = \left|n\binom{s}{n}\right|$. Verify that

$$a_n = |s|\left|1 - \frac{s}{1}\right|\left|1 - \frac{s}{2}\right|\cdots\left|1 - \frac{s}{n-1}\right|. \tag{5.6.E.13}$$

If $s>0, \{a_n\} \downarrow$ for $n>s+1,$ so we may put $L=\lim a_n=\lim a_{2n} \geq 0.$ (Explain!)
Prove that

$$\frac{a_{2n}}{a_n}<\left|1-\frac{s}{2n}\right|^n \rightarrow e^{-\frac{1}{2}s} \text{ as } n \rightarrow \infty, \tag{5.6.E.14}$$

so for large n,

$$\frac{a_{2n}}{a_n}<e^{-\frac{1}{2}s}+\varepsilon; \text{ i.e., } a_{2n}<\left(e^{-\frac{1}{2}s}+\varepsilon\right) a_n. \tag{5.6.E.15}$$

With ε fixed, let $n \rightarrow \infty$ to get $L \leq\left(e^{-\frac{1}{2}s}+\varepsilon\right) L.$ Then with $\varepsilon \rightarrow 0,$ obtain $L e^{\frac{1}{2}s} \leq L.$
As $e^{\frac{1}{2}s}>1$ (for $s>0$), this implies $L=0,$ as claimed.
(ii) For $s>-1, s+1>0,$ so by (i),

$$(n+1)\binom{s+1}{n+1} \rightarrow 0; \text{ i.e., } (s+1)\binom{s}{n} \rightarrow 0. \quad (\text{ Why? }) \tag{5.6.E.16}$$

(iii) Use the ratio test to show that the series $\sum\binom{s}{n} n x^n$ converges when $|x|<1.$
Then apply Theorem 4 of Chapter 4, §13.]

? Exercise 5.6.E.8

Continuing Problems 6 and 7, prove that

$$(1+x)^s=\sum_{k=0}^n\binom{s}{k} x^k+R_n(x), \tag{5.6.E.17}$$

where $R_n(x) \rightarrow 0$ if either $|x|<1,$ or $x=1$ and $s>-1,$ or $x=-1$ and $s>0.$
[Hints: (a) If $0 \leq x \leq 1,$ use $(5')$ for

$$R_{n-1}(x)=\binom{s}{n} x^n\left(1+\theta_n x\right)^{s-n}, \quad 0<\theta_n<1. \text{ (Verify!)} \tag{5.6.E.18}$$

Deduce that $\left|R_{n-1}(x)\right| \leq\left|\binom{s}{n} x^n\right| \rightarrow 0.$ Use Problem 7(iii) if $|x|<1$ or Problem 7(ii) if $x=1.$
(b) If $-1 \leq x<0,$ write $(5'')$ as

$$R_{n-1}(x)=\binom{s}{n} n x^n\left(1+\theta_n' x\right) s^{-1}\left(\frac{1-\theta_n'}{1+\theta_n' x}\right)^{n-1}. \text{ (Check!)} \tag{5.6.E.19}$$

As $-1 \leq x<0,$ the last fraction is $\leq 1.$ (Why?) Also,

$$\left(1+\theta_n' x\right)^{s-1} \leq 1 \text{ if } s>1, \text{ and } \leq(1+x)^{s-1} \text{ if } s \leq 1. \tag{5.6.E.20}$$

Thus, with x fixed, these expressions are bounded, while $\binom{s}{n} n x^n \rightarrow 0$ by Problem 7(i) or (iii). Deduce that $R_{n-1} \rightarrow 0,$ hence $R_n \rightarrow 0.$]

Exercise 5.6.E.9

Prove that

$$\ln(1+x)=\sum_{k=1}^{n}(-1)^{k+1}\frac{x^{k}}{k}+R_{n}(x), \tag{5.6.E.21}$$

where $\lim_{n\rightarrow\infty} R_{n}(x)=0$ if $-1<x\leq 1$.
[Hints: If $0\leq x\leq 1,$ use formula $(5')$.
If $-1<x<0,$ use formula (6) with $g(t)=\ln(1+t)$ to obtain

$$R_{n}(x)=\frac{\ln(1+x)}{(-1)^{n}}\left(\frac{1-\theta_{n}}{1+\theta_{n}x}\cdot x\right)^{n}. \tag{5.6.E.22}$$

Proceed as in Problem 8.]

Exercise 5.6.E.10

Prove that if $f: E^{1}\rightarrow E^{*}$ is of class CD^{1} on $[a,b]$ and if $-\infty<f''<0$ on (a,b), then for each $x_{0}\in(a,b)$,

$$f\left(x_{0}\right)>\frac{f(b)-f(a)}{b-a}\left(x_{0}-a\right)+f(a); \tag{5.6.E.23}$$

i.e., the curve $y=f(x)$ lies above the secant through $(a, f(a))$ and $(b, f(b))$.
[Hint: This formula is equivalent to

$$\frac{f\left(x_{0}\right)-f(a)}{x_{0}-a}>\frac{f(b)-f(a)}{b-a}, \tag{5.6.E.24}$$

i.e., the average of f' on $\left[a, x_{0}\right]$ is strictly greater than the average of f' on $[a,b]$, which follows because f' decreases on (a,b). (Explain!)]

Exercise 5.6.E.11

Prove that if $a, b, r,$ and s are positive reals and $r+s=1,$ then

$$a^{r} b^{s}\leq ra+sb. \tag{5.6.E.25}$$

(This inequality is important for the theory of so-called L^{p} -spaces.)
[Hints: If $a=b$, all is trivial.
Therefore, assume $a<b$. Then

$$a=(r+s)a<ra+sb<b. \tag{5.6.E.26}$$

Use Problem 10 with $x_{0}=ra+sb\in(a,b)$ and

$$f(x)=\ln x, f''(x)=-\frac{1}{x^{2}}<0. \tag{5.6.E.27}$$

Verify that

$$x_{0}-a=x_{0}-(r+s)a=s(b-a) \tag{5.6.E.28}$$

and $r \cdot \ln a = (1-s) \ln a;$ hence deduce that

$$r \cdot \ln a + s \cdot \ln b - \ln a = s(\ln b - \ln a) = s(f(b) - f(a)). \tag{5.6.E.29}$$

After substitutions, obtain

$$f\left(x_0\right) = \ln(ra + sb) > r \cdot \ln a + s \cdot \ln b = \ln\left(a^r b^s\right).] \tag{5.6.E.30}$$

Exercise $5.6.E.12$

Use Taylor's theorem (Theorem 1') to prove the following inequalities:
(a) $\sqrt[3]{1+x} < 1 + \frac{x}{3}$ if $x > -1, x \neq 0.$
(b) $\cos x > 1 - \frac{1}{2} x^2$ if $x \neq 0.$
(c) $\frac{x}{1+x^2} < \arctan x < x$ if $x > 0.$
(d) $x > \sin x > x - \frac{1}{6} x^3$ if $x > 0.$

5.7: The Total Variation (Length) of a Function f - E1 → E

This page is a draft and is under active development.

The question that we shall consider now is how to define reasonably (and precisely) the notion of the length of a curve (Chapter 4, §10) described by a function $f: E^{1} \rightarrow E$ over an interval $I=[a, b]$, i.e., $f[I]$.

We proceed as follows (see Figure 24).

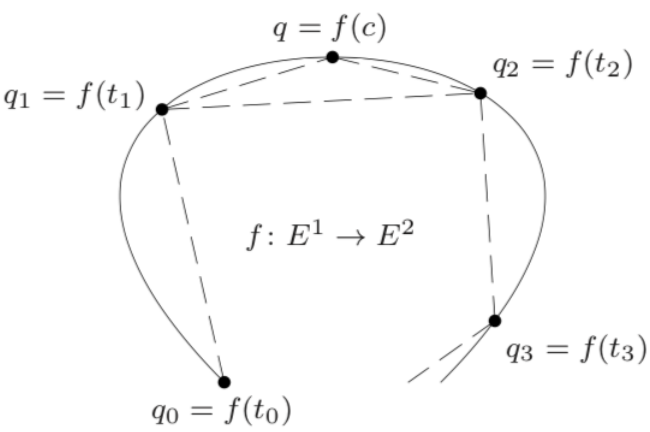

FIGURE 24

Subdivide $[a, b]$ by a finite set of points $P=\left\{t_{0}, t_{1}, \ldots, t_{m}\right\}$, with

$$a=t_{0} \leq t_{1} \leq \cdots \leq t_{m}=b; \tag{5.7.1}$$

P is called a partition of $[a, b]$. Let

$$q_{i}=f\left(t_{i}\right), \quad i=1,2, \ldots, m, \tag{5.7.2}$$

and, for $i=1,2, \ldots, m$,

$$\begin{aligned}
\Delta_{i} f &=q_{i}-q_{i-1} \\
&=f\left(t_{i}\right)-f\left(t_{i-1}\right).
\end{aligned}$$

We also define

$$S(f, P)=\sum_{i=1}^{m}\left|\Delta_{i} f\right|=\sum_{i=1}^{m}\left|q_{i}-q_{i-1}\right|. \tag{5.7.3}$$

Geometrically, $\left|\Delta_{i} f\right|=\left|q_{i}-q_{i-1}\right|$ is the length of the line segment $L\left[q_{i-1}, q_{i}\right]$ in E, and $S(f, P)$ is the sum of such lengths, i.e., the length of the polygon

$$W=\bigcup_{i=1}^{m} L\left[q_{i-1}, q_{i}\right] \tag{5.7.4}$$

inscribed into $f[I]$; we denote it by

$$\ell W=S(f, P). \tag{5.7.5}$$

Now suppose we add a new partition point c, with

$$t_{i-1} \leq c \leq t_{i}. \tag{5.7.6}$$

Then we obtain a new partition

$$P_c = \{t_0, \ldots, t_{i-1}, c, t_i, \ldots, t_m\}, \tag{5.7.7}$$

called a refinement of P, and a new inscribed polygon W_c in which $L[q_{i-1}, q_i]$ is replaced by two segments, $L[q_{i-1}, q]$ and $L[q, q_i]$, where $q = f(c)$; see Figure 24. Accordingly, the term $|\Delta_i f| = |q_i - q_{i-1}|$ in $S(f, P)$ is replaced by

$$|q_i - q| + |q - q_{i-1}| \geq |q_i - q_{i-1}| \quad \text{(triangle law)}. \tag{5.7.8}$$

It follows that

$$S(f, P) \leq S(f, P_c); \text{ i.e., } \ell W \leq \ell W_c. \tag{5.7.9}$$

Hence we obtain the following result.

Corollary 5.7.1

The sum $S(f, P) = \ell W$ cannot decrease when P is refined.

Thus when new partition points are added, $S(f, P)$ grows in general; i.e., it approaches some supremum value (finite or not). Roughly speaking, the inscribed polygon W gets "closer" to the curve. It is natural to define the desired length of the curve to be the lub of all lengths ℓW, i.e., of all sums $S(f, P)$ resulting from the various partitions P. This supremum is also called the total variation of f over $[a, b]$, denoted $V_f[a, b]$.

Definition 1

Given any function $f : E^1 \rightarrow E$, and $I = [a, b] \subset E^1$, we set

$$V_f[I] = V_f[a, b] = \sup_P S(f, P) = \sup_P \sum_{i=1}^{m} |f(t_i) - f(t_{i-1})| \geq 0 \text{ in } E^*, \tag{5.7.10}$$

where the supremum is over all partitions $P = \{t_0, \ldots, t_m\}$ of I. We call $V_f[I]$ the total variation, or length, of f on I. Briefly, we denote it by V_f.

Note 1. If f is continuous on $[a, b]$, the image set $A = f[I]$ is an arc (Chapter 4, §10). It is customary to call $V_f[I]$ the length of that arc, denoted $\ell_f A$ or briefly ℓA. Note, however, that there may well be another function g, continuous on an interval J, such that $g[J] = A$ but $V_f[I] \neq V_g[J]$, and so $\ell_f A \neq \ell_g A$. Thus it is safer to say "the length of A as described by f on I." Only for simple arcs (where f is one to one), is "ℓA" unambiguous. (See Problems 6-8.)

In the propositions below, f is an arbitrary function, $f : E^1 \rightarrow E$.

Theorem 5.7.1 (additivity of V_f)

If $a \leq c \leq b$, then

$$V_f[a, b] = V_f[a, c] + V_f[c, b]; \tag{5.7.11}$$

i.e., the length of the whole equals the sum of the lengths of the parts.

Proof

Take any partition $P = \{t_0, \ldots, t_m\}$ of $[a, b]$. If $c \notin P$, refine P to

$$P_c = \{t_0, \ldots, t_i, c, t_i, \ldots, t_m\}. \tag{5.7.12}$$

Then by Corollary 1, $S(f, P) \leq S(f, P_c)$.

Now P_c splits into partitions of $[a, c]$ and $[c, b]$, namely,

$$P' = \{t_0, \ldots, t_{i-1}, c\} \text{ and } P'' = \{c, t_i, \ldots, t_m\}, \tag{5.7.13}$$

so that

$$S(f, P') + S(f, P'') = S(f, P_c). \text{ (Verify!)} \tag{5.7.14}$$

Hence (as V_f is the lub of the corresponding sums) ,

$$V_f[a,c]+V_f[c,d]\geq S(f,P_c)\geq S(f,P). \tag{5.7.15}$$

As P is an arbitrary partition of $[a,b]$, we also have

$$V_f[a,c]+V_f[c,b]\geq \sup S(f,P)=V_f[a,b]. \tag{5.7.16}$$

Thus it remains to show that, conversely,

$$V_f[a,b]\geq V_f[a,c]+V_f[c,b]. \tag{5.7.17}$$

The latter is trivial if $V_f[a,b]=+\infty$. Thus assume $V_f[a,b]=K<+\infty$. Let P' and P'' be any partitions of $[a,c]$ and $[c,b]$, respectively. Then $P^*=P'\cup P''$ is a partition of $[a,b]$, and

$$S(f,P')+S(f,P'')=S(f,P^*)\leq V_f[a,b]=K, \tag{5.7.18}$$

whence

$$S(f,P')\leq K-S(f,P''). \tag{5.7.19}$$

Keeping P'' fixed and varying P', we see that $K-S(f,P'')$ is an upper bound of all $S(f,P')$ over $[a,c]$. Hence

$$V_f[a,c]\leq K-S(f,P'') \tag{5.7.20}$$

or

$$S(f,P'')\leq K-V_f[a,c]. \tag{5.7.21}$$

Similarly, varying P'', we now obtain

$$V_f[c,b]\leq K-V_f[a,c] \tag{5.7.22}$$

or

$$V_f[a,c]+V_f[c,b]\leq K=V_f[a,b], \tag{5.7.23}$$

as required. Thus all is proved. □

Corollary 5.7.2 (monotonicity of V_f)

If $a\leq c\leq d\leq b$, then

$$V_f[c,d]\leq V_f[a,b]. \tag{5.7.24}$$

Proof

By Theorem 1,

$$V_f[a,b]=V_f[a,c]+V_f[c,d]+V_f[d,b]\geq V_f[c,d]. \quad \square \tag{5.7.25}$$

Definition 2

If $V_f[a,b]<+\infty$, we say that f is of bounded variation on $I=[a,b]$, and that the set $f[I]$ is rectifiable (by f on I).

Corollary 5.7.3

For each $t\in[a,b]$,

$$|f(t)-f(a)|\leq V_f[a,b]. \tag{5.7.26}$$

Hence if f is of bounded variation on $[a,b]$, it is bounded on $[a,b]$.

Proof

If $t \in [a,b]$, let $P = \{a, t, b\}$, so

$$|f(t) - f(a)| \leq |f(t) - f(a)| + |f(b) - f(t)| = S(f,P) \leq V_f[a,b], \tag{5.7.27}$$

proving our first assertion. Hence

$$(\forall t \in [a,b]) \quad |f(t)| \leq |f(t) - f(a)| + |f(a)| \leq V_f[a,b] + |f(a)|. \tag{5.7.28}$$

This proves the second assertion. □

Note 2. Neither boundedness, nor continuity, nor differentiability of f on $[a,b]$ implies $V_f[a,b] < +\infty$, but boundedness of f' does. (See Problems 1 and 3.)

Corollary 5.7.4

A function f is finite and constant on $[a,b]$ iff $V_f[a,b] = 0$.

The proof is left to the reader. (Use Corollary 3 and the definitions.

Theorem 5.7.2

Let f, g, h be real or complex (or let f and g be vector valued and h scalar valued). Then on any interval $I = [a,b]$, we have

(i) $V_{|f|} \leq V_f$;

(ii) $V_{f \pm g} \leq V_f + V_g$; and

(iii) $V_{hf} \leq sV_f + rV_h$, with $r = \sup_{t \in I} |f(t)|$ and $s = \sup_{t \in I} |h(t)|$.

Hence if f, g, and h are of bounded variation on I, so are $f \pm g$, hf, and $|f|$.

Proof

We first prove (iii).

Take any partition $P = \{t_0, \ldots, t_m\}$ of I. Then

$$\begin{aligned} \left|\Delta_i hf\right| &= \left|h\left(t_i\right) f\left(t_i\right) - h\left(t_{i-1}\right) f\left(t_{i-1}\right)\right| \\ &\leq \left|h\left(t_i\right) f\left(t_i\right) - h\left(t_{i-1}\right) f\left(t_i\right)\right| + \left|h\left(t_{i-1}\right) f\left(t_i\right) - h\left(t_{i-1}\right) f\left(t_{i-1}\right)\right| \\ &= \left|f\left(t_i\right)\right|\left|\Delta_i h\right| + \left|h\left(t_{i-1}\right)\right|\left|\Delta_i f\right| \\ &\leq r\left|\Delta_i h\right| + s\left|\Delta_i f\right|. \end{aligned}$$

Adding these inequalities, we obtain

$$S(hf, P) \leq r \cdot S(h,P) + s \cdot S(f,P) \leq rV_h + sV_f. \tag{5.7.29}$$

As this holds for all sums $S(hf, P)$, it holds for their supremum, so

$$V_{hf} = \sup S(hf, P) \leq rV_h + sV_f, \tag{5.7.30}$$

as claimed.

Similarly, (i) follows from

$$\left|\left|f\left(t_i\right)\right| - \left|f\left(t_{i-1}\right)\right|\right| \leq \left|f\left(t_i\right) - f\left(t_{i-1}\right)\right|. \tag{5.7.31}$$

The analogous proof of (ii) is left to the reader.

Finally, (i)-(iii) imply that V_f, $V_{f \pm g}$, and V_{hf} are finite if V_f, V_g, and V_h are. This proves our last assertion. □

Note 3. Also f/h is of bounded variation on I if f and h are, provided h is bounded away from 0 on I; i.e.,

$$(\exists \varepsilon > 0) \quad |h| \geq \varepsilon \text{ on } I. \tag{5.7.32}$$

(See Problem 5.)

Special theorems apply in case the range space E is E^1 or E^n (* or C^n).

Theorem 5.7.3

(i) A real function f is of bounded variation on $I=[a,b]$ iff $f=g-h$ for some nondecreasing real functions g and h on I.

(ii) If f is real and monotone on I, it is of bounded variation there.

Proof

We prove (ii) first.

Let $f \uparrow$ on I. If $P=\{t_0,\ldots,t_m\}$, then

$$t_i \geq t_{i-1} \text{ implies } f(t_i) \geq f(t_{i-1}). \tag{5.7.33}$$

Hence $\Delta_i f \geq 0$; i.e., $|\Delta_i f|=\Delta_i f$. Thus

$$\begin{aligned} S(f,P) &= \sum_{i=1}^{m}|\Delta_i f| = \sum_{i=1}^{m}\Delta_i f = \sum_{i=1}^{m}[f(t_i)-f(t_{i-1})] \\ &= f(t_m)-f(t_0)=f(b)-f(a) \end{aligned}$$

for any P. (Verify!) This implies that also

$$V_f[I]=\sup S(f,P)=f(b)-f(a)<+\infty. \tag{5.7.34}$$

Thus (ii) is proved.

Now for (i), let $f=g-h$ with $g\uparrow$ and $h\uparrow$ on I. By (ii), g and h are of bounded variation on I. Hence so is $f=g-h$ by Theorem 2 (last clause).

Conversely, suppose $V_f[I]<+\infty$. Then define

$$g(x)=V_f[a,x], x\in I, \text{ and } h=g-f \text{ on } I, \tag{5.7.35}$$

so $f=g-h$, and it only remains to show that $g\uparrow$ and $h\uparrow$.

To prove it, let $a\leq x\leq y\leq b$. Then Theorem 1 yields

$$V_f[a,y]-V_f[a,x]=V_f[x,y]; \tag{5.7.36}$$

i.e.,

$$g(y)-g(x)=V_f[x,y]\geq|f(y)-f(x)|\geq 0 \quad \text{(by Corollary 3)}. \tag{5.7.37}$$

Hence $g(y)\geq g(x)$. Also, as $h=g-f$, we have

$$\begin{aligned} h(y)-h(x) &= g(y)-f(y)-[g(x)-f(x)] \\ &= g(y)-g(x)-[f(y)-f(x)] \\ &\geq 0 \quad \text{by (2)}. \end{aligned}$$

Thus $h(y)\geq h(x)$. We see that $a\leq x\leq y\leq b$ implies $g(x)\leq g(y)$ and $h(x)\leq h(y)$, so $h\uparrow$ and $g\uparrow$, indeed. □

Theorem 5.7.4

(i) A function $f: E^1\rightarrow E^n\ ({}^*C^n)$ is of bounded variation on $I=[a,b]$ iff all of its components $(f_1,f_2,\ldots,f_n)$ are.

(ii) If this is the case, then finite limits $f(p^+)$ and $f(q^-)$ exist for every $p\in[a,b)$ and $q\in(a,b]$.

Proof

(i) Take any partition $P=\{t_0,\ldots,t_m\}$ of I. Then

$$|f_k(t_i)-f_k(t_{i-1})|^2 \leq \sum_{j=1}^{n}|f_j(t_i)-f_j(t_{i-1})|^2=|f(t_i)-f(t_{i-1})|^2; \tag{5.7.38}$$

i.e., $|\Delta_i f_k| \leq |\Delta_i f|, i=1,2,\ldots, m$. Thus

$$(\forall P) \quad S(f_k, P) \leq S(f, P) \leq V_f, \tag{5.7.39}$$

and $V_{f_k} \leq V_f$ follows. Thus

$$V_f < +\infty \text{ implies } V_{f_k} < +\infty, \quad k=1,2,\ldots, n. \tag{5.7.40}$$

The converse follows by Theorem 2 since $f=\sum_{k=1}^n f_k \vec{e}_k$. (Explain!)

(ii) For real monotone functions, $f(p^+)$ and $f(q^-)$ exist by Theorem 1 of Chapter 4, §5. This also applies if f is real and of bounded variation, for by Theorem 3,

$$f=g-h \text{ with } g \uparrow \text{ and } h \uparrow \text{ on } I, \tag{5.7.41}$$

and so

$$f(p^+)=g(p^+)-h(p^+) \text{ and } f(q^-)=g(q^-)-h(q^-) \text{ exist.} \tag{5.7.42}$$

The limits are finite since f is bounded on I by Corollary 3.

Via components (Theorem 2 of Chapter 4, §3), this also applies to functions $f: E^1 \rightarrow E^n$. (Why?) In particular, (ii) applies to complex functions (treat C as E^2 (*and so it extends to functions $f: E^1 \rightarrow C^n$. as well). □

We also have proved the following corollary.

Corollary 5.7.5

A complex function $f: E^1 \rightarrow C$ is of bounded variation on $[a, b]$ iff its real and imaginary parts are. (See Chapter 4, §3, Note 5).

5.7.E: Problems on Total Variation and Graph Length

Exercise $5.7.E.1$

In the following cases show that $V_f[I]=+\infty$, though f is bounded on I. (In case (iii), f is continuous, and in case (iv), it is even differentiable on I.)

(i) For $I=[a, b](a<b)$, $f(x)=\begin{cases}1 & \text{if } x \in R(\text{ rational }), \text{ and } \\ 0 & \text{if } x \in E^1-R.\end{cases}$

(ii) $f(x)=\sin \frac{1}{x} ; f(0)=0 ; I=[a, b], a \leq 0 \leq b, a<b$.

(iii) $f(x)=x \cdot \sin \frac{\pi}{2 x} ; f(0)=0 ; I=[0,1]$.

(iv) $f(x)=x^2 \sin \frac{1}{x^2} ; f(0)=0 ; I=[0,1]$.

[Hints: (i) For any m there is P, with

$$\left|\Delta_i f\right|=1, \quad i=1,2, \ldots, m, \tag{5.7.E.1}$$

so $S(f, P)=m \rightarrow+\infty$.

(iii) Let

$$P_m=\left\{0, \frac{1}{m}, \frac{1}{m-1}, \ldots, \frac{1}{2}, 1\right\}. \tag{5.7.E.2}$$

Prove that $S\left(f, P_m\right) \geq \sum_{k=1}^m \frac{1}{k} \rightarrow+\infty$.]

Exercise $5.7.E.2$

Let $f: E^1 \rightarrow E^1$ be monotone on each of the intervals

$$\left[a_{k-1}, a_k\right], \quad k=1, \ldots, n \quad (\text{"piecewise monotone"}). \tag{5.7.E.3}$$

Prove that

$$V_f\left[a_0, a_n\right]=\sum_{k=1}^n\left|f\left(a_k\right)-f\left(a_{k-1}\right)\right|. \tag{5.7.E.4}$$

In particular, show that this applies if $f(x)=\sum_{i=1}^n c_i x^i$ (polynomial), with $c_i \in E^1$.

[Hint: It is known that a polynomial of degree n has at most n real roots. Thus it is piecewise monotone, for its derivative vanishes at finitely many points (being of degree $n-1$). Use Theorem 1 in §2.]

Exercise $5.7.E.3$

$\Rightarrow$ Prove that if f is finite and relatively continuous on $I=[a, b]$, with a bounded derivative, $\left|f^{\prime}\right| \leq M$, on $I-Q$(see §4), then

$$V_f[a, b] \leq M(b-a). \tag{5.7.E.5}$$

However, we may have $V_f[I]<+\infty$, and yet $\left|f^{\prime}\right|=+\infty$ at some $p \in I$.

[Hint: Take $f(x)=\sqrt[3]{x}$ on $[-1,1]$.]

Exercise 5.7.E.4

Complete the proofs of Corollary 4 and Theorems 2 and 4.

Exercise 5.7.E.5

Prove Note 3.
[Hint: If $|h| \geq \varepsilon$ on I, show that

$$\left|\frac{1}{h\left(t_{i}\right)}-\frac{1}{h\left(t_{i-1}\right)}\right| \leq \frac{\left|\Delta_{i} h\right|}{\varepsilon^{2}} \tag{5.7.E.6}$$

and hence

$$S\left(\frac{1}{h}, P\right) \leq \frac{S(h, P)}{\varepsilon^{2}} \leq \frac{V_{h}}{\varepsilon^{2}}. \tag{5.7.E.7}$$

Deduce that $\frac{1}{h}$ is of bounded variation on I if h is. Then apply Theorem 2(iii) to $\frac{1}{h} \cdot f.$]

Exercise 5.7.E.6

Let $g: E^{1} \rightarrow E^{1}$ (real) and $f: E^{1} \rightarrow E$ be relatively continuous on $J=[c, d]$ and $I=[a, b]$, respectively, with $a=g(c)$ and $b=g(d)$. Let

$$h=f \circ g. \tag{5.7.E.8}$$

Prove that if g is one to one on J, then
(i) $g[J]=I$, so f and h describe one and the same arc $A=f[I]=h[J]$;
(ii) $V_{f}[I]=V_{h}[J]$; i.e., $\ell_{f} A=\ell_{h} A$.
[Hint for (ii): Given $P=\left\{a=t_{0}, \ldots, t_{m}=b\right\}$, show that the points $s_{i}=g^{-1}\left(t_{i}\right)$ form a partition $P^{\prime}$ of $J=[c, d]$, with $S\left(h, P^{\prime}\right)=S(f, P)$. Hence deduce $V_{f}[I] \leq V_{h}[J]$.
Then prove that $V_{h}[J] \leq V_{f}[I]$, taking an arbitrary $P^{\prime}=\left\{c=s_{0}, \ldots, s_{m}=d\right\}$, and defining $P=\left\{t_{0}, \ldots, t_{m}\right\}$, with $t_{i}=g\left(s_{i}\right)$. What if $g(c)=b, g(d)=a$?]

Exercise 5.7.E.7

Prove that if $f, h: E^{1} \rightarrow E$ are relatively continuous and one to one on $I=[a, b]$ and $J=[c, d]$, respectively, and if

$$f[I]=h[J]=A \tag{5.7.E.9}$$

(i.e., f and h describe the same simple arc A), then

$$\ell_{f} A=\ell_{h} A. \tag{5.7.E.10}$$

Thus for simple arcs, $\ell_{f} A$ is independent of f.
[Hint: Define $g: J \rightarrow E^{1}$ by $g=f^{-1} \circ h$. Use Problem 6 and Chapter 4, §9, Theorem 3. First check that Problem 6 works also if $g(c)=b$ and $g(d)=a$, i.e., $g \downarrow$ on J.]

Exercise 5.7.E.8

Let $I=[0,2 \pi]$ and define $f, g, h: E^{1} \rightarrow E^{2}(C)$ by

$$\begin{aligned} f(x) &= (\sin x, \cos x), \\ g(x) &= (\sin 3x, \cos 3x), \\ h(x) &= \left(\sin \tfrac{1}{x}, \cos \tfrac{1}{x}\right) \text{ with } h(0) = (0, 1). \end{aligned} \tag{5.7.E.11}$$

Show that $f[I] = g[I] = h[I]$ (the unit circle; call it A), yet $\ell_f A = 2\pi$, $\ell_g A = 6\pi$, while $V_h[I] = +\infty$. (Thus the result of Problem 7 fails for closed curves and nonsimple arcs.)

Exercise $5.7.E.9$

In Theorem 3, define two functions $G, H : E^1 \rightarrow E^1$ by

$$G(x) = \frac{1}{2}[V_f[a, x] + f(x) - f(a)] \tag{5.7.E.12}$$

and

$$H(x) = G(x) - f(x) + f(a). \tag{5.7.E.13}$$

$(G$ and H are called, respectively, the positive and negative variation functions for $f.)$ Prove that
(i) $G \uparrow$ and $H \uparrow$ on $[a, b]$;
(ii) $f(x) = G(x) - [H(x) - f(a)]$ (thus the functions f and g of Theorem 3 are not unique);
(iii) $V_f[a, x] = G(x) + H(x)$;
(iv) if $f = g - h$, with $g \uparrow$ and $h \uparrow$ on $[a, b]$, then

$$V_G[a, b] \leq V_g[a, b], \text{ and } V_H[a, b] \leq V_h[a, b]; \tag{5.7.E.14}$$

(v) $G(a) = H(a) = 0$.

Exercise $5.7.E.10*$

Prove that if $f : E^1 \rightarrow E^n\ (^*C^n)$ is of bounded variation on $I = [a, b]$, then f has at most countably many discontinuities in I.
[Hint: Apply Problem 5 of Chapter 4, §5. Proceed as in the proof of Theorem 4 in §7. Finally, use Theorem 2 of Chapter 1, §9.]

5.8: Rectifiable Arcs. Absolute Continuity

This page is a draft and is under active development.

If a function $f : E^1 \rightarrow E$ is of bounded variation (§7) on an interval $I = [a, b]$, we can define a real function v_f on I by

$$v_f(x) = V_f[a, x] (= \text{ total variation of } f \text{ on } [a, x]) \text{ for } x \in I; \tag{5.8.1}$$

v_f is called the total variation function, or length function, generated by f on I. Note that $v_f \uparrow$ on I. (Why?) We now consider the case where f is also relatively continuous on I, so that the set $A = f[I]$ is a rectifiable arc (see §7, Note 1 and Definition 2).

Definition 1

A function $f : E^1 \rightarrow E$ is (weakly) absolutely continuous on $I = [a, b]$ iff $V_f[I] < +\infty$ and f is relatively continuous on I.

Theorem 5.8.1

The following are equivalent:

(i) f is (weakly) absolutely continuous on $I = [a, b]$;

(ii) v_f is finite and relatively continuous on I; and

(iii) $(\forall \varepsilon > 0)\ (\exists \delta > 0)\ (\forall x, y \in I | 0 \leq y - x < \delta)\ V_f[x, y] < \varepsilon$.

Proof

We shall show that (ii) ⇒ (iii) ⇒ (i) ⇒ (ii).

(ii) ⇒ (iii). As $I = [a, b]$ is compact, (ii) implies that v_f is uniformly continuous on I (Theorem 4 of Chapter 4, §8). Thus

$$(\forall \varepsilon > 0)\ (\exists \delta > 0)\ (\forall x, y \in I | 0 \leq y - x < \delta) \quad v_f(y) - v_f(x) < \varepsilon. \tag{5.8.2}$$

However,

$$v_f(y) - v_f(x) = V_f[a, y] - V_f[a, x] = V_f[x, y] \tag{5.8.3}$$

by additivity (Theorem 1 in §7). Thus (iii) follows.

(iii) ⇒ (i). By Corollary 3 of §7, $|f(x) - f(y)| \leq V_f[x, y]$. Therefore, (iii) implies that

$$(\forall \varepsilon > 0)\ (\exists \delta > 0)\ (\forall x, y \in I | |x - y| < \delta) \quad |f(x) - f(y)| < \varepsilon, \tag{5.8.4}$$

and so f is relatively (even uniformly) continuous on I.

Now with ε and δ as in (iii), take a partition $P = \{t_0, \ldots, t_m\}$ of I so fine that

$$t_i - t_{i-1} < \delta, \quad i = 1, 2, \ldots, m. \tag{5.8.5}$$

Then $(\forall i) V_f\left[t_{i-1}, t_i\right] < \varepsilon$. Adding up these m inequalities and using the additivity of V_f, we obtain

$$V_f[I] = \sum_{i=1}^{m} V_f\left[t_{i-1}, t_i\right] < m\varepsilon < +\infty. \tag{5.8.6}$$

Thus (i) follows, by definition.

That (i) ⇒ (ii) is given as the next theorem. □

Theorem 5.8.2

If $V_f[I] < +\infty$ and if f is relatively continuous at some $p \in I$ (over $I = [a, b]$), then the same applies to the length function v_f.

Proof

We consider left continuity first, with $a < p \leq b$.

Let $\varepsilon > 0$. By assumption, there is $\delta > 0$ such that

$$|f(x) - f(p)| < \frac{\varepsilon}{2} \text{ when } |x - p| < \delta \text{ and } x \in [a, p]. \tag{5.8.7}$$

Fix any such x. Also, $V_f[a, p] = \sup_P S(f, P)$ over $[a, p]$. Thus

$$V_f[a, p] - \frac{\varepsilon}{2} < \sum_{i=1}^{k} |\Delta_i f| \tag{5.8.8}$$

for some partition

$$P = \{t_0 = a, \ldots, t_{k-1}, t_k = p\} \text{ of } [a, p]. \text{ (Why?)} \tag{5.8.9}$$

We may assume $t_{k-1} = x, x$ as above. (If $t_{k-1} \neq x$, add x to P.) Then

$$|\Delta_k f| = |f(p) - f(x)| < \frac{\varepsilon}{2}, \tag{5.8.10}$$

and hence

$$V_f[a, p] - \frac{\varepsilon}{2} < \sum_{i=1}^{k-1} |\Delta_i f| + |\Delta_k f| < \sum_{i=1}^{k-1} |\Delta_i f| + \frac{\varepsilon}{2} \leq V_f\left[a, t_{k-1}\right] + \frac{\varepsilon}{2}. \tag{5.8.11}$$

However,

$$V_f[a, p] = v_f(p) \tag{5.8.12}$$

and

$$V_f\left[a, t_{k-1}\right] = V_f[a, x] = v_f(x). \tag{5.8.13}$$

Thus (1) yields

$$|v_f(p) - v_f(x)| = V_f[a, p] - V_f[a, x] < \varepsilon \text{ for } x \in [a, p] \text{ with } |x - p| < \delta. \tag{5.8.14}$$

This shows that v_f is left continuous at p.

Right continuity is proved similarly on noting that

$$v_f(x) - v_f(p) = V_f[p, b] - V_f[x, b] \text{ for } p \leq x < b. \text{ (Why?)} \tag{5.8.15}$$

Thus v_f is, indeed, relatively continuous at p. Observe that v_f is also of bounded variation on I, being monotone and finite (see Theorem 3(ii) of §7).

This completes the proof of both Theorem 2 and Theorem 1. □

We also have the following.

Corollary 5.8.1

If f is real and absolutely continuous on $I = [a, b]$ (weakly), so are the nondecreasing functions g and $h (f = g - h)$ defined in Theorem 3 of §7.

Indeed, the function g as defined there is simply v_f. Thus it is relatively continuous and finite on I by Theorem 1. Hence so also is $h = f - g$. Both are of bounded variation (monotone!) and hence absolutely continuous (weakly).

Note 1. The proof of Theorem 1 shows that (weak) absolute continuity implies uniform continuity. The converse fails, however (see Problem 1(iv) in §7).

We now apply our theory to antiderivatives (integrals).

Corollary 5.8.2

If $F=\int f$ on $I=[a, b]$ and if f is bounded $\left(|f| \leq K \in E^1\right)$ on $I-Q$ (Q countable), then F is weakly absolutely continuous on I.

(Actually, even the stronger variety of absolute continuity follows. See Chapter 7, §11, Problem 17).

Proof

By definition, $F=\int f$ is finite and relatively continuous on I, so we only have to show that $V_F[I]<+\infty$. This, however, easily follows by Problem 3 of §7 on noting that $F^{\prime}=f$ on $I-S$ (S countable). Details are left to the reader. □

Our next theorem expresses arc length in the form of an integral.

Theorem 5.8.3

If $f: E^1 \rightarrow E$ is continuously differentiable on $I=[a, b]$ (§6), then $v_f=\int\left|f^{\prime}\right|$ on I and

$$V_f[a, b]=\int_a^b\left|f^{\prime}\right| . \tag{5.8.16}$$

Proof

Let $a<p<x \leq b, \Delta x=x-p$, and

$$\Delta v_f=v_f(x)-v_f(p)=V_f[p, x]. \quad \text{(Why?)} \tag{5.8.17}$$

As a first step, we shall show that

$$\frac{\Delta v_f}{\Delta x} \leq \sup _{[p, x]}\left|f^{\prime}\right| . \tag{5.8.18}$$

For any partition $P=\left\{p=t_0, \ldots, t_m=x\right\}$ of $[p, x]$, we have

$$S(f, P)=\sum_{i=1}^m\left|\Delta_i f\right| \leq \sum_{i=1}^m \sup _{\left[t_{i-1}, t_i\right]}\left|f^{\prime}\right|\left(t_i-t_{i-1}\right) \leq \sup _{[p, x]}\left|f^{\prime}\right| \Delta x . \tag{5.8.19}$$

Since this holds for any partition P, we have

$$V_f[p, x] \leq \sup _{[p, x]}\left|f^{\prime}\right| \Delta x, \tag{5.8.20}$$

which implies (2).

On the other hand,

$$\Delta v_f=V_f[p, x] \geq|f(x)-f(p)|=|\Delta f| . \tag{5.8.21}$$

Combining, we get

$$\left|\frac{\Delta f}{\Delta x}\right| \leq \frac{\Delta v_f}{\Delta x} \leq \sup _{[p, x]}\left|f^{\prime}\right|<+\infty \tag{5.8.22}$$

since $f^{\prime}$ is relatively continuous on $[a, b]$, hence also uniformly continuous and bounded. (Here we assumed $a<p<x \leq b$. However, (3) holds also if $a \leq x<p<b$, with $\Delta v_f=-V[x, p]$ and $\Delta x<0$. Verify!)

Now

$$||f'(p)|-|f'(x)|| \leq |f'(p)-f'(x)| \rightarrow 0 \quad \text{as } x \rightarrow p, \tag{5.8.23}$$

so, taking limits as $x \rightarrow p$, we obtain

$$\lim_{x \rightarrow p} \frac{\Delta v_f}{\Delta x} = |f'(p)|. \tag{5.8.24}$$

Thus v_f is differentiable at each p in (a, b), with $v_f'(p) = |f'(p)|$. Also, v_f is relatively continuous and finite on $[a, b]$ (by Theorem 1). Hence $v_f = \int |f'|$ on $[a, b]$, and we obtain

$$\int_a^b |f'| = v_f(b) - v_f(a) = V_f[a, b], \text{ as asserted. } \square \tag{5.8.25}$$

Note 2. If the range space E is E^n (*or C^n), f has n components

$$f_1, f_2, \ldots, f_n. \tag{5.8.26}$$

By Theorem 5 in §1, $f' = (f_1', f_2', \ldots, f_n')$, so

$$|f'| = \sqrt{\sum_{k=1}^n |f_k'|^2}, \tag{5.8.27}$$

and we get

$$V_f[a, b] = \int_a^b \sqrt{\sum_{k=1}^n |f_k'|^2} = \int_a^b \sqrt{\sum_{k=1}^n |f_k'(t)|^2} dt \quad \text{(classical notation)}. \tag{5.8.28}$$

In particular, for complex functions, we have (see Chapter 4, §3, Note 5)

$$V_f[a, b] = \int_a^b \sqrt{f_{\text{re}}'(t)^2 + f_{\text{im}}'(t)^2} dt. \tag{5.8.29}$$

In practice, formula (5) is used when a curve is given parametrically by

$$x_k = f_k(t), \quad k = 1, 2, \ldots, n, \tag{5.8.30}$$

with the f_k differentiable on $[a, b]$. Curves in E^2 are often given in nonparametric form as

$$y = F(x), \quad F : E^1 \rightarrow E^1. \tag{5.8.31}$$

Here $F[I]$ is *not* the desired curve but simply a set in E^1. To apply (5) here, we first replace "$y = F(x)$" by suitable parametric equations,

$$x = f_1(t) \text{ and } y = f_2(t); \tag{5.8.32}$$

i.e., we introduce a function $f : E^1 \rightarrow E$, with $f = (f_1, f_2)$. An obvious (but not the only) way of achieving it is to set

$$x = f_1(t) = t \text{ and } y = f_2(t) = F(t) \tag{5.8.33}$$

so that $f_1' = 1$ and $f_2' = F'$. Then formula (5) may be written as

$$V_f[a, b] = \int_a^b \sqrt{1 + F'(x)^2} dx, \quad f(x) = (x, F(x)). \tag{5.8.34}$$

Example

Find the length of the circle

$$x^2 + y^2 = r^2. \tag{5.8.35}$$

Here it is convenient to use the parametric equations

$$x = r \cos t \text{ and } y = r \sin t, \tag{5.8.36}$$

i.e., to define $f : E^1 \rightarrow E^2$ by

$$f(t)=(r \cos t, r \sin t), \tag{5.8.37}$$

or, in complex notation,

$$f(t)=r e^{t i}. \tag{5.8.38}$$

Then the circle is obtained by letting t vary through $[0, 2\pi]$. Thus (5) yields

$$V_f[0, 2\pi]=\int_a^b r \sqrt{\cos^2 t+\sin^2 t}\, dt=r \int_a^b 1 dt=\left.r t\right|_0^{2\pi}=2 r \pi. \tag{5.8.39}$$

Note that f describes the same circle $A=f[I]$ over $I=[0, 4\pi]$. More generally, we could let t vary through any interval $[a, b]$ with $b-a \geq 2\pi$. However, the length, $V_f[a, b]$, would change (depending on $b-a$). This is because the circle $A=f[I]$ is not a simple arc (see §7, Note 1), so ℓA depends on f and I, and one must be careful in selecting both appropriately.

5.8.E: Problems on Absolute Continuity and Rectifiable Arcs

Exercise 5.8.E.1

Complete the proofs of Theorems 2 and 3, giving all missing details.

Exercise 5.8.E.2

$\Rightarrow$ Show that f is absolutely continuous (in the weaker sense) on $[a, b]$ if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$\sum_{i=1}^{m} |f(t_i) - f(s_i)| < \varepsilon \text{ whenever } \sum_{i=1}^{m} (t_i - s_i) < \delta \text{ and}$$
$$a \leq s_1 \leq t_1 \leq s_2 \leq t_2 \leq \cdots \leq s_m \leq t_m \leq b.$$

(This is absolute continuity in the stronger sense.)

Exercise 5.8.E.3

Prove that v_f is strictly monotone on $[a, b]$ iff f is not constant on any nondegenerate subinterval of $[a, b]$.
$[\text{Hint: If } x < y, V_f[x, y] > 0, \text{ by Corollary 4 of } \S 7]$.

Exercise 5.8.E.4

With f, g, h as in Theorem 2 of §7, prove that if f, g, h are absolutely continuous (in the weaker sense) on I, so are $f \pm g, hf$, and $|f|$; so also is f/h if $(\exists \varepsilon > 0)|h| \geq \varepsilon$ on I.

Exercise 5.8.E.5

Prove that
(i) If f' is finite and $\neq 0$ on $I = [a, b]$, so also is v_f'(with one-sided derivatives at the endpoints of the interval); moreover,

$$\left|\frac{f'}{v_f'}\right| = 1 \text{ on } I. \tag{5.8.E.1}$$

$\text{Thus show that } f'/v_f' \text{ is the tangent unit vector (see } \S 1\text{)}$.
(ii) Under the same assumptions, $F = f \circ v_f^{-1}$ is differentiable on $J = [0, v_f(b)]$ (with one-sided derivatives at the endpoints of the interval) and $F[J] = f[I]$; i.e., F and f describe the same simple arc, with $V_F[I] = V_f[I]$.
Note that this is tantamount to a change of parameter. Setting $s = v_f(t)$, i.e., $t = v_f^{-1}(s)$, we have $f(t) = f\left(v_f^{-1}(s)\right) = F(s)$, with the arclength s as parameter.

5.9: Convergence Theorems in Differentiation and Integration

This page is a draft and is under active development.

Given

$$F_n = \int f_n \text{ or } F_n' = f_n, \quad n = 1, 2, \ldots, \tag{5.9.1}$$

what can one say about $\int \lim f_n$ or $(\lim F_n)'$ if the limits exist? Below we give some answers, for complete range spaces E (such as E^n). Roughly, we have $\lim F_n' = (\lim F_n)'$ on $I - Q$ if

(a) $\{F_n(p)\}$ converges for at least one $p \in I$ and

(b) $\{F_n'\}$ converges uniformly.

Here I is a finite or infinite interval in E^1 and Q is countable. We include in Q the endpoints of I (if any), so $I - Q \subseteq I^0 (= \text{ interior of } I)$.

Theorem 5.9.1

Let $F_n : E^1 \rightarrow E\ (n = 1, 2, \ldots)$ be finite and relatively continuous on I and differentiable on $I - Q$. Suppose that

(a) $\lim_{n \rightarrow \infty} F_n(p)$ exists for some $p \in I$;

(b) $F_n' \rightarrow f \neq \pm\infty (\text{uniformly})$ on $J - Q$ for each finite subinterval $J \subseteq I$;

(c) E is complete.

Then

(i) $\lim_{n \rightarrow \infty} F_n = F$ exists uniformly on each finite subinterval $J \subseteq I$;

(ii) $F = \int f$ on I; and

(iii) $(\lim F_n)' = F' = f = \lim_{n \rightarrow \infty} F_n'$ on $I - Q$.

Proof

Fix $\varepsilon > 0$ and any subinterval $J \subseteq I$ of length $\delta < \infty$, with $p \in J$ (p as in (a)). By (b), $F_n' \rightarrow f$ (uniformly) on $J - Q$, so there is a k such that for $m, n > k$,

$$|F_n'(t) - f(t)| < \frac{\varepsilon}{2}, \quad t \in J - Q; \tag{5.9.2}$$

hence

$$\sup_{t \in J - Q} |F_m'(t) - F_n'(t)| \leq \varepsilon. \quad \text{(Why?)} \tag{5.9.3}$$

Now apply Corollary 1 in §4 to the function $h = F_m - F_n$ on J. Then for each $x \in J, |h(x) - h(p)| \leq M|x - p|$, where

$$M \leq \sup_{t \in J - Q} |h'(t)| \leq \varepsilon \tag{5.9.4}$$

by (2). Hence for $m, n > k, x \in J$ and

$$|F_m(x) - F_n(x) - F_m(p) + F_n(p)| \leq \varepsilon|x - p| \leq \varepsilon\delta. \tag{5.9.5}$$

As ε is arbitrary, this shows that the sequence

$$\{F_n - F_n(p)\} \tag{5.9.6}$$

satisfies the uniform Cauchy criterion (Chapter 4, §12, Theorem 3). Thus as E is complete, $\{F_n - F_n(p)\}$ converges uniformly on J. So does $\{F_n\}$, for $\{F_n(p)\}$ converges, by (a). Thus we may set

$$F=\lim F_n \text{ (uniformly) on } J, \tag{5.9.7}$$

proving assertion (i).

Here by Theorem 2 of Chapter 4, §12, F is relatively continuous on each such $J \subseteq I$, hence on all of I. Also, letting $m \rightarrow+\infty$ (with n fixed), we have $F_m \rightarrow F$ in (3), and it follows that for $n>k$ and $x \in G_p(\delta) \cap I$.

$$\left|F(x)-F_n(x)-F(p)+F_n(p)\right| \leq \varepsilon|x-p| \leq \varepsilon \delta. \tag{5.9.8}$$

Having proved (i), we may now treat p as just any point in I. Thus formula (4) holds for any globe $G_p(\delta), p \in I$. We now show that

$$F^{\prime}=f \text{ on } I-Q; \text{ i.e., } F=\int f \text{ on } I. \tag{5.9.9}$$

Indeed, if $p \in I-Q$, each F_n is differentiable at p (by assumption), and $p \in I^0$ (since $I-Q \subseteq I^0$ by our convention). Thus for each n, there is $\delta_n>0$ such that

$$\left|\frac{\Delta F_n}{\Delta x}-F_n^{\prime}(p)\right|=\left|\frac{F_n(x)-F_n(p)}{x-p}-F_n^{\prime}(p)\right|<\frac{\varepsilon}{2} \tag{5.9.10}$$

for all $x \in G_{\neg p}\left(\delta_n\right) ; G_p\left(\delta_n\right) \subseteq I$.

By assumption (b) and by (4), we can fix n so large that

$$\left|F_n^{\prime}(p)-f(p)\right|<\frac{\varepsilon}{2} \tag{5.9.11}$$

and so that (4) holds for $\delta=\delta_n$. Then, dividing by $|\Delta x|=|x-p|$, we have

$$\left|\frac{\Delta F}{\Delta x}-\frac{\Delta F_n}{\Delta x}\right| \leq \varepsilon. \tag{5.9.12}$$

Combining with (5), we infer that for each $\varepsilon>0$, there is $\delta>0$ such that

$$\left|\frac{\Delta F}{\Delta x}-f(p)\right| \leq\left|\frac{\Delta F}{\Delta x}-\frac{\Delta F_n}{\Delta x}\right|+\left|\frac{\Delta F_n}{\Delta x}-F_n^{\prime}(p)\right|+\left|F_n^{\prime}(p)-f(p)\right|<2 \varepsilon, x \in G_p(\delta). \tag{5.9.13}$$

This shows that

$$\lim _{x \rightarrow p} \frac{\Delta F}{\Delta x}=f(p) \text{ for } p \in I-Q, \tag{5.9.14}$$

i.e., $F^{\prime}=f$ on $I-Q$, with f finite by assumption, and F finite by (4). As F is also relatively continuous on I, assertion (ii) is proved, and (iii) follows since $F=\lim F_n$ and $f=\lim F_n^{\prime}$. $\square$

Note 1. The same proof also shows that $F_n \rightarrow F$ (uniformly) on each closed subinterval $J \subseteq I$ if $F_n^{\prime} \rightarrow f$ (uniformly) on $J-Q$ for all such J (with the other assumptions unchanged). In any case, we then have $F_n \rightarrow F$ (pointwise) on all of I.

We now apply Theorem 1 to antiderivatives.

Theorem 5.9.2

Let the functions $f_n: E^1 \rightarrow E, n=1,2, \ldots,$ have antiderivatives, $F_n=\int f_n$, on I. Suppose E is complete and $f_n \rightarrow f$ (uniformly) on each finite subinterval $J \subseteq I$, with f finite there. Then $\int f$ exists on I, and

$$\int_p^x f=\int_p^x \lim _{n \rightarrow \infty} f_n=\lim _{n \rightarrow \infty} \int_p^x f_n \text{ for any } p, x \in I. \tag{5.9.15}$$

Proof

Fix any $p \in I$. By Note 2 in §5, we may choose

$$F_n(x) = \int_p^x f_n \text{ for } x \in I. \tag{5.9.16}$$

Then $F_n(p) = \int_p^p f_n = 0,$ and so $\lim_{n\to\infty} F_n(p) = 0$ exists, as required in Theorem 1(a).

Also, by definition, each F_n is relatively continuous and finite on I and differentiable, with $F_n' = f_n,$ on $I - Q_n.$ The countable sets Q_n need not be the same, so we replace them by

$$Q = \bigcup_{n=1}^{\infty} Q_n \tag{5.9.17}$$

(including in Q also the endpoints of I, if any. Then Q is countable (see Chapter 1, §9, Theorem 2), and $I - Q \subseteq I - Q_n,$ so all F_n are differentiable on $I - Q,$ with $F_n' = f_n$ there.

Additionally, by assumption,

$$f_n \rightarrow f \text{ (uniformly)} \tag{5.9.18}$$

on finite subintervals $J \subseteq I.$ Hence

$$F_n' \rightarrow f \text{ (uniformly) on } J - Q \tag{5.9.19}$$

for all such J, and so the conditions of Theorem 1 are satisfied.

By that theorem, then,

$$F = \int f = \lim F_n \text{ exists on } I \tag{5.9.20}$$

and, recalling that

$$F_n(x) = \int_p^x f_n, \tag{5.9.21}$$

we obtain for $x \in I$

$$\int_p^x f = F(x) - F(p) = \lim F_n(x) - \lim F_n(p) = \lim F_n(x) - 0 = \lim \int_p^x f_n. \tag{5.9.22}$$

As $p \in I$ was arbitrary, and $f = \lim f_n$ (by assumption), all is proved. $\square$

Note 2. By Theorem 1, the convergence

$$\int_p^x f_n \rightarrow \int_p^x f \quad (\text{ i.e. }, F_n \rightarrow F) \tag{5.9.23}$$

is uniform in x (with p fixed), on each finite subinterval $J \subseteq I$.

We now "translate" Theorems 1 and 2 into the language of series.

Corollary 5.9.1

Let E and the functions $F_n : E^1 \rightarrow E$ be as in Theorem 1. Suppose the series

$$\sum F_n(p) \tag{5.9.24}$$

converges for some $p \in I$, and

$$\sum F_n' \tag{5.9.25}$$

converges uniformly on $J - Q,$ for each finite subinterval $J \subseteq I$.

Then $\sum F_n$ converges uniformly on each such J, and

$$F=\sum_{n=1}^{\infty} F_n \tag{5.9.26}$$

is differentiable on $I-Q$, with

$$F'=\left(\sum_{n=1}^{\infty} F_n\right)'=\sum_{n-1}^{\infty} F_n' \text{ there.} \tag{5.9.27}$$

In other words, the series can be differentiated termwise.

Proof

Let

$$s_n=\sum_{k=1}^{n} F_k, \quad n=1,2,\ldots, \tag{5.9.28}$$

be the partial sums of $\sum F_n$. From our assumptions, it then follows that the s_n satisfy all conditions of Theorem 1. (Verify!) Thus the conclusions of Theorem 1 hold, with F_n replaced by s_n.

We have $F=\lim s_n$ and $F'=(\lim s_n)'=\lim s_n'$, whence (7) follows. □

Corollary 5.9.2

If E and the f_n are as in Theorem 2 and if $\sum f_n$ converges uniformly to f on each finite interval $J\subseteq I$, then $\int f$ exists on I, and

$$\int_p^x f=\int_p^x \sum_{n=1}^{\infty} f_n=\sum_{n=1}^{\infty}\int_p^x f_n \text{ for any } p,x\in I. \tag{5.9.29}$$

Briefly, a uniformly convergent series can be integrated termwise.

Theorem 5.9.3 (Power Series)

Let r be the convergence radius of

$$\sum a_n(x-p)^n, \quad a_n\in E, p\in E^1. \tag{5.9.30}$$

Suppose E is complete. Set

$$f(x)=\sum_{n=0}^{\infty} a_n(x-p)^n \quad \text{on } I=(p-r,p+r). \tag{5.9.31}$$

Then the following are true:

(i) f is differentiable and has an exact antiderivative on I.

(ii) $f'(x)=\sum_{n=1}^{\infty} na_n(x-p)^{n-1}$ and $\int_p^x f=\sum_{n=0}^{\infty}\frac{a_n}{n+1}(x-p)^{n+1}, x\in I$.

(iii) r is also the convergence radius of the two series in (ii).

(iv) $\sum_{n=0}^{\infty} a_n(x-p)^n$ is exactly the Taylor series for $f(x)$ on I about p.

Proof

We prove (iii) first.

By Theorem 6 of Chapter 4, §13, $r=1/d$, where

$$d=\overline{\lim}\sqrt[n]{a_n}. \tag{5.9.32}$$

Let r' be the convergence radius of $\sum na_n(x-p)^{n-1}$, so

$$r'=\frac{1}{d'} \text{ with } d'=\overline{\lim}\sqrt[n]{na_n}. \tag{5.9.33}$$

However, $\lim \sqrt[n]{n}=1$ (see §3, Example (e)). It easily follows that

$$d'=\overline{\lim}\sqrt[n]{na_n}=1\cdot\overline{\lim}\sqrt[n]{a_n}=d^2. \tag{5.9.34}$$

Hence $r'=1/d'=1/d=r$.

The convergence radius of $\sum \frac{a_n}{n+1}(x-p)^{n+1}$ is determined similarly. Thus (iii) is proved.

Next, let

$$f_n(x)=a_n(x-p)^n \text{ and } F_n(x)=\frac{a_n}{n+1}(x-p)^{n+1}, n=0,1,2,\ldots. \tag{5.9.35}$$

Then the F_n are differentiable on I, with $F_n'=f_n$ there. Also, by Theorems 6 and 7 of Chapter 4, §13, the series

$$\sum F_n'=\sum a_n(x-p)^n \tag{5.9.36}$$

converges uniformly on each closed subinterval $J\subseteq I=(p-r,p+r)$. Thus the functions F_n satisfy all conditions of Corollary 1, with $Q=\emptyset$, and the f_n satisfy Corollary 2. By Corollary 1, then,

$$F=\sum_{n=1}^{\infty}F_n \tag{5.9.37}$$

is differentiable on I, with

$$F'(x)=\sum_{n=1}^{\infty}F_n'(x)=\sum_{n=1}^{\infty}a_n(x-p)^n=f(x) \tag{5.9.38}$$

for all $x\in I$. Hence F is an exact antiderivative of f on I, and (8) yields the second formula in (ii).

Quite similarly, replacing F_n and F by f_n and f, one shows that f is differentiable on I, and the first formula in (ii) follows. This proves (i) as well.

Finally, to prove (iv), we apply (i)-(iii) to the consecutive derivatives of f and obtain

$$f^{(k)}(x)=\sum_{n=k}^{\infty}n(n-1)\cdots(n-k+1)a_n(x-p)^{n-k} \tag{5.9.39}$$

for each $x\in I$ and $k\in N$.

If $x=p$, all terms vanish except the first term $(n=k)$, i.e., $k!a_k$. Thus $f^{(k)}(p)=k!a_k, k\in N$. We may rewrite it as

$$a_n=\frac{f^{(n)}(p)}{n!}, \quad n=0,1,2,\ldots, \tag{5.9.40}$$

since $f^{(0)}(p)=f(p)=a_0$. Assertion (iv) now follows since

$$f(x)=\sum_{n=0}^{\infty}a_n(x-p)^n=\sum_{n=0}^{\infty}\frac{f^{(n)}(p)}{n!}(x-p)^n, \quad x\in I, \text{ as required.} \quad \square \tag{5.9.41}$$

Note 3. If $\sum a_n(x-p)^n$ converges also for $x=p-r$ or $x=p+r$, so does the integrated series

$$\sum a_n\frac{(x-p)^{n+1}}{n+1} \tag{5.9.42}$$

since we may include such x in I. However, the derived series $\sum na_n(x-p)^{n-1}$ need not converge at such x. (Why?) For example (see §6, Problem 9), the expansion

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots \tag{5.9.43}$$

converges for $x=1$ but the derived series

$$1-x+x^2-\cdots \tag{5.9.44}$$

does not.

In this respect, there is the following useful rule for functions $f: E^1 \rightarrow E^m\ \left({}^*C^m\right)$.

Corollary 5.9.3

Let a function $f: E^1 \rightarrow E^m\left({}^*C^m\right)$ be relatively continuous on $\left[p, x_0\right]$ (or $\left[x_0, p\right]$), $x_0 \neq p$. If

$$f(x)=\sum_{n=0}^{\infty} a_n(x-p)^n \text{ for } p \leq x<x_0 \text{ (respectively, } x_0<x \leq p), \tag{5.9.45}$$

and if $\sum a_n\left(x_0-p\right)^n$ converges, then necessarily

$$f\left(x_0\right)=\sum_{n=0}^{\infty} a_n\left(x_0-p\right)^n. \tag{5.9.46}$$

Proof

The proof is sketched in Problems 4 and 5.

Thus in the above example, $f(x)=\ln(1+x)$ defines a continuous function on $[0,1]$, with

$$f(x)=\sum_{n=1}^{\infty}(-1)^{n-1} \frac{x^n}{n} \text{ on } [0,1]. \tag{5.9.47}$$

We gave a direct proof in §6, Problem 9. However, by Corollary 3, it suffices to prove this for $[0,1)$, which is much easier. Then the convergence of

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \tag{5.9.48}$$

yields the result for $x=1$ as well.

5.9.E: Problems on Convergence in Differentiation and Integration

Exercise $5.9.E.1$

Complete all proof details in Theorems 1 and 3, Corollaries 1 and 2, and Note 3.

Exercise $5.9.E.2$

Show that assumptions (a) and (c) in Theorem 1 can be replaced by $F_n \rightarrow F$ (pointwise) on I. (In this form, the theorem applies to incomplete spaces E as well.)
$[\text{Hint: } F_n \rightarrow F(\text{ pointwise } e), \text{ together with formula } (3), \text{ implies } F_n \rightarrow F \text{ (uniformly) on } I.]$

Exercise $5.9.E.3$

Show that Theorem 1 fails without assumption (b), even if $F_n \rightarrow F$ (uniformly) and if F is differentiable on I.
[Hint: For a counterexample, try $F_n(x)=\frac{1}{n} \sin n x$, on any nondegenerate I. Verify that $F_n \rightarrow 0$ (uniformly), yet (b) and assertion (iii) fail.

Exercise $5.9.E.4$

Prove Abel's theorem (Chapter 4, $\S 13$, Problem 15) for series

$$\sum a_n(x-p)^n, \tag{5.9.E.1}$$

with all a_n in E^m (* or in C^m) but with $x, p \in E^1$.
[Hint: Split $a_n(x-p)^n$ into components.]

Exercise $5.9.E.5$

Prove Corollary 3.
$[\text{Hint: By Abel's theorem (see Problem 4), we may put}$

$$\sum_{n=0}^{\infty} a_n(x-p)^n=F(x) \tag{5.9.E.2}$$

uniformly on $[p, x_0]$ (respectively, $[x_0, p]$). This implies that F is relatively continuous at x_0. (Why?) So is f, by assumption. Also $f=F$ on $[p, x_0)\ ((x_0, p])$. Show that

$$f(x_0)=\lim f(x)=\lim F(x)=F(x_0) \tag{5.9.E.3}$$

as $x \rightarrow x_0$ from the left (right).]

Exercise $5.9.E.6$

In the following cases, find the Taylor series of F about 0 by integrating the series of F'. Use Theorem 3 and Corollary 3 to find the convergence radius r and to investigate convergence at $-r$ and r. Use (b) to find a formula for π.
(a) $F(x)=\ln(1+x)$;
(b) $F(x)=\arctan x$;
(c) $F(x)=\arcsin x$.

Exercise $5.9.E.7$

Prove that

$$\int_0^x \frac{\ln(1-t)}{t} dt = \sum_{n=1}^{\infty} \frac{x^n}{n^2} \quad \text{for } x \in [-1,1]. \tag{5.9.E.4}$$

[Hint: Use Theorem 3 and Corollary 3. Take derivatives of both sides.]

5.10: Sufficient Condition of Integrability. Regulated Functions

This page is a draft and is under active development.

In this section, we shall determine a large family of functions that do have antiderivatives. First, we give a general definition, valid for any range space (T, p) (not necessarily E). The domain space remains E^1.

Definition 1

A function $f : E^1 \rightarrow (T, p)$ is said to be regulated on an interval $I \subseteq E^1$, with endpoints $a < b$, iff the limits $f\left(p^{-}\right)$ and $f\left(p^{+}\right)$, other than $\pm\infty$, exist at each $p \in I$. However, at the endpoints a, b, if in I, we only require $f\left(a^{+}\right)$ and $f\left(b^{-}\right)$ to exist.

Examples

(a) If f is relatively continuous and finite on I, it is regulated.

(b) Every real monotone function is regulated (see Chapter 4, §5, Theorem 1).

(c) If $f : E^1 \rightarrow E^n\left({}^*C^n\right)$ has bounded variation on I, it is regulated (§7, Theorem 4).

(d) The characteristic function of a set B, denoted C_B, is defined by

$$C_B(x) = 1 \text{ if } x \in B \text{ and } C_B = 0 \text{ on } -B. \tag{5.10.1}$$

For any interval $J \subseteq E^1, C_J$ is regulated on E^1.

(e) A function f is called a step function on I iff I can be represented as the union, $I = \bigcup_k I_k$, of a sequence of disjoint intervals I_k such that f is constant and $\neq \pm\infty$ on each I_k. Note that some I_k may be singletons, $\{p\}$.

If the number of the I_k is finite, we call f a simple step function.

When the range space T is E, we can give the following convenient alternative definition. If, say, $f = a_k \neq \pm\infty$ on I_k, then

$$f = \sum_k a_k C_{I_k} \quad \text{ on } I, \tag{5.10.2}$$

where C_{I_k} is as in (d). Note that $\sum_k a_k C_{I_k}(x)$ always exists for disjoint I_k. (Why?)

Each simple step function is regulated. (Why?)

Theorem 5.10.1

Let the functions f, g, h be real or complex (or let f, g be vector valued and h scalar valued).

If they are regulated on I, so are $f \pm g, fh$, and $|f|$; so also is f / h if h is bounded away from 0 on I, i.e., $(\exists \varepsilon > 0)|h| \geq \varepsilon$ on I.

Proof

The proof, based on the usual limit properties, is left to the reader.

We shall need two lemmas. One is the famous Heine-Borel lemma.

lemma 5.10.1 (Heine-Borel)

If a closed interval $A = [a, b]$ in E^1 (or E^n) is covered by open sets $G_i (i \in I)$, i.e.,

$$A \subseteq \bigcup_{i \in I} G_i, \tag{5.10.3}$$

then A can be covered by a finite number of these G_i.

Proof

The proof was sketched in Problem 10 of Chapter 4, §6.

Note 1. This fails for nonclosed intervals A. For example, let

$$A=(0,1) \subseteq E^1 \text{ and } G_n=\left(\frac{1}{n}, 1\right). \tag{5.10.4}$$

Then

$$A=\bigcup_{n=1}^{\infty} G_n(\text{ verify! }), \text{ but not } A \subseteq \bigcup_{n=1}^m G_n \tag{5.10.5}$$

for any finite m. (Why?)

The lemma also fails for nonopen sets G_i. For example, cover A by singletons $\{x\}, x \in A$. Then none of the $\{x\}$ can be dropped!

lemma 5.10.2

If a function $f: E^1 \rightarrow T$ is regulated on $I=[a, b]$, then f can be uniformly approximated by simple step functions on I.

That is, for any $\varepsilon>0$, there is a simple step function g, with $\rho(f, g) \leq \varepsilon$ on I; hence

$$\sup_{x \in I} \rho(f(x), g(x)) \leq \varepsilon. \tag{5.10.6}$$

Proof

By assumption, $f\left(p^{-}\right)$ exists for each $p \in(a, b]$, and $f\left(p^{+}\right)$ exists for $p \in[a, b)$, all finite.

Thus, given $\varepsilon>0$ and any $p \in I$, there is $G_p(\delta)$ (δ depending on p) such that $\rho(f(x), r)<\varepsilon$ whenever $r=f\left(p^{-}\right)$ and $x \in(p-\delta, p)$, and $\rho(f(x), s)<\varepsilon$ whenever $s=f\left(p^{+}\right)$ and $x \in(p, p+\delta) ; x \in I$

We choose such a $G_p(\delta)$ for every $p \in I$. Then the open globes $G_p=G_p(\delta)$ cover the closed interval $I=[a, b]$, so by Lemma 1, I is covered by a finite number of such globes, say,

$$I \subseteq \bigcup_{k=1}^n G_{p_k}\left(\delta_k\right), \quad a \in G_{p_1}, a \leq p_1<p_2<\cdots<p_n \leq b. \tag{5.10.7}$$

We now define the step function g on I as follows.

If $x=p_k$, we put

$$g(x)=f\left(p_k\right), \quad k=1,2, \ldots, n. \tag{5.10.8}$$

If $x \in\left[a, p_1\right)$, then

$$g(x)=f\left(p_1^{-}\right). \tag{5.10.9}$$

If $x \in\left(p_1, p_1+\delta_1\right)$, then

$$g(x)=f\left(p_1^{+}\right). \tag{5.10.10}$$

More generally, if x is in $G_{\neg p_k}\left(\delta_k\right)$ but in none of the $G_{p_i}\left(\delta_i\right), i<k$, we put

$$g(x)=f\left(p_k^{-}\right) \quad \text{if } x<p_k \tag{5.10.11}$$

and

$$g(x)=f\left(p_k^{+}\right) \quad \text{if } x>p_k. \tag{5.10.12}$$

Then by construction, $\rho(f, g)<\varepsilon$ on each G_{p_k}, hence on I. $\square$

***Note 2.** If T is complete, we can say more: f is regulated on $I=[a,b]$ iff f is uniformly approximated by simple step functions on I. (See Problem 2.)

Theorem 5.10.2

If $f: E^{1} \rightarrow E$ is regulated on an interval $I \subseteq E^{1}$ and if E is complete, then $\int f$ exists on I, exact at every continuity point of f in I^{0}.

In particular, all continuous maps $f: E^{1} \rightarrow E^{n}\left({ }^{*} C^{n}\right)$ have exact primitives.

Proof

In view of Problem 14 of §5, it suffices to consider closed intervals.

Thus let $I=[a, b], a<b$, in E^{1}. Suppose first that f is the characteristic function C_{J} of a subinterval $J \subseteq I$ with endpoints c and d $(a \leq c \leq d \leq b)$, so $f=1$ on J and $f=0$ on $I-J$. We then define $F(x)=x$ on $J, F=c$ on $[a, c]$, and $F=d$ on $[d, b]$ (see Figure 25). Thus F is continuous (why?), and $F^{\prime}=f$ on $I-\{a, b, c, d\}$ (why?). Hence $F=\int f$ on I; i.e., characteristic functions are integrable.

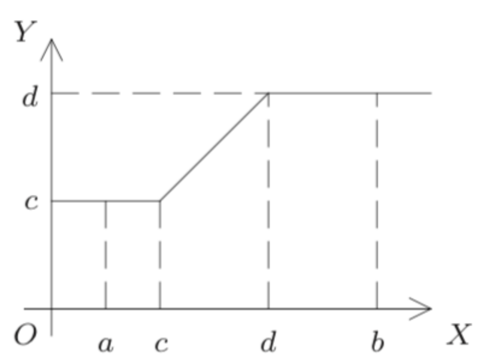

FIGURE 25

Then, however, so is any simple step function

$$f=\sum_{k=1}^{m} a_{k} C_{I_{k}}, \tag{5.10.13}$$

by repeated use of Corollary 1 in §5.

Finally, let f be any regulated function on I. Then by Lemma 2, for any $\varepsilon_{n}=\frac{1}{n}$, there is a simple step function g_{n} such that

$$\sup _{x \in I}\left|g_{n}(x)-f(x)\right| \leq \frac{1}{n}, \quad n=1,2, \ldots . \tag{5.10.14}$$

As $\frac{1}{n} \rightarrow 0$, this implies that $g_{n} \rightarrow f$ (uniformly) on I (see Chapter 4, §12, Theorem 1). Also, by what was proved above, the step functions g_{n} have antiderivatives, hence so has f (Theorem 2 in §9); ; i.e., $F=\int f$ exists on I, as claimed. Moreover, $\int f$ is exact at continuity points of f in I^{0} (Problem 10 in §5). $\square$

In view of the sufficient condition expressed in Theorem 2, we can now replace the assumption "$\int f$ exists" in our previous theorems by "f is regulated" (provided E is complete). For example, let us now review Problems 7 and 8 in §5.

Theorem 5.10.3 (weighted law of the mean)

Let $f: E^1 \to E$ (E complete) and $g: E^1 \to E^1$ be regulated on $I=[a,b]$, with $g \geq 0$ on I. Then the following are true:

(i) There is a finite $c \in E$ (called the "g-weighted mean of f on I") such that $\int_a^b gf = c\int_a^b g$.

(ii) If f, too, is real and has the Darboux property on I, then $c=f(q)$ for some $q \in I$.

Proof

Indeed, as f and g are regulated, so is gf by Theorem 1. Hence by Theorem 2, $\int f$ and $\int gf$ exist on I. The rest follows as in Problems 7 and 8 of §5. □

Theorem 5.10.4 (second law of the mean)

Suppose f and g are real, f is monotone with $f=\int f'$ on I, and g is regulated on I; $I=[a,b]$. Then

$$\int_a^b fg = f(a)\int_a^q g + f(b)\int_q^b g \text{ for some } q \in I. \tag{5.10.15}$$

Proof

To fix ideas, let $f \uparrow$; i.e., $f' \geq 0$ on I.

The formula $f=\int f'$ means that f is relatively continuous (hence regulated) on I and differentiable on $I-Q$ (Q countable). As g is regulated,

$$\int_a^x g = G(x) \tag{5.10.16}$$

does exist on I, so G has similar properties, with $G(a)=\int_a^a g=0$.

By Theorems 1 and 2, $\int fG' = \int fg$ exists on I. (Why?) Hence by Corollary 5 in §5, so does $\int Gf'$, and we have

$$\int_a^b fg = \int_a^b fG' = f(x)G(x)\Big|_a^b - \int_a^b Gf' = f(b)G(b) - \int_a^b Gf'. \tag{5.10.17}$$

Now G has the Darboux property on I (being relatively continuous), and $f' \geq 0$. Also, $\int G$ and $\int Gf'$ exist on I. Thus by Problems 7 and 8 in §5,

$$\int_a^b Gf' = G(q)\int_a^b f' = G(q)f(x)\Big|_a^b, \quad q \in I. \tag{5.10.18}$$

Combining all, we obtain the required result (1) since

$$\begin{aligned}\int fg &= f(b)G(b) - \int_a^b Gf' \\ &= f(b)G(b) - f(b)G(q) + f(a)G(q) \\ &= f(b)\int_q^b g + f(a)\int_a^q g. \quad \square\end{aligned}$$

We conclude with an application to infinite series. Given $f: E^1 \to E$, we define

$$\int_a^\infty f = \lim_{x\to+\infty}\int_a^x f \text{ and } \int_{-\infty}^a f = \lim_{x\to-\infty}\int_x^a f \tag{5.10.19}$$

if these integrals and limits exist.

We say that $\int_a^\infty f$ and $\int_{-\infty}^a f$ converge iff they exist and are finite.

Theorem 5.10.5 (integral test of convergence)

If $f: E^1 \rightarrow E^1$ is nonnegative and nonincreasing on $I=[a,+\infty)$, then

$$\int_a^\infty f \text{ converges iff } \sum_{n=1}^\infty f(n) \text{ does.} \tag{5.10.20}$$

Proof

As $f \downarrow$, f is regulated, so $\int f$ exists on $I=[a,+\infty)$. We fix some natural $k \geq a$ and define

$$F(x)=\int_k^x f \text{ for } x \in I. \tag{5.10.21}$$

By Theorem 3(iii) in §5, $F \uparrow$ on I. Thus by monotonicity,

$$\lim_{x \rightarrow+\infty} F(x)=\lim_{x \rightarrow+\infty} \int_k^x f=\int_k^\infty f \tag{5.10.22}$$

exists in E^*; so does $\int_a^k f$. Since

$$\int_a^x f=\int_a^k f+\int_k^x f, \tag{5.10.23}$$

where $\int_a^k f$ is finite by definition, we have

$$\int_a^\infty f<+\infty \text{ iff } \int_k^\infty f<+\infty. \tag{5.10.24}$$

Similarly,

$$\sum_{n=1}^\infty f(n)<+\infty \quad \text{ iff } \sum_{n=k}^\infty f(n)<+\infty. \tag{5.10.25}$$

Thus we may replace "a" by "k."

Let

$$I_n=[n, n+1), \quad n=k, k+1, \ldots, \tag{5.10.26}$$

and define two step functions, g and h, constant on each I_n, by

$$h=f(n) \text{ and } g=f(n+1) \text{ on } I_n, n \geq k. \tag{5.10.27}$$

Since $f \downarrow$, we have $g \leq f \leq h$ on all I_n, hence on $J=[k,+\infty)$. Therefore,

$$\int_k^x g \leq \int_k^x f \leq \int_k^x h \text{ for } x \in J. \tag{5.10.28}$$

Also,

$$\int_k^m h=\sum_{n=k}^{m-1} \int_n^{n+1} h=\sum_{n=k}^{m-1} f(n), \tag{5.10.29}$$

since $h=f(n)$ (constant) on $[n, n+1)$, and so

$$\int_n^{n+1} h(x) d x=f(n) \int_n^{n+1} 1 d x=\left.f(n) \cdot x\right|_n^{n+1}=f(n)(n+1-n)=f(n). \tag{5.10.30}$$

Similarly,

$$\int_k^m g=\sum_{n=k}^{m-1} f(n+1)=\sum_{n=k+1}^m f(n). \tag{5.10.31}$$

Thus we obtain

$$\sum_{n=k+1}^{m} f(n)=\int_{k}^{m} g \leq \int_{k}^{m} f \leq \int_{k}^{m} h=\sum_{n=k}^{m-1} f(n), \tag{5.10.32}$$

or, letting $m \rightarrow \infty$,

$$\sum_{n=k+1}^{\infty} f(n) \leq \int_{k}^{\infty} f \leq \sum_{n=k}^{\infty} f(n). \tag{5.10.33}$$

Hence $\int_{k}^{\infty} f$ is finite iff $\sum_{n=1}^{\infty} f(n)$ is, and all is proved. $\square$

Examples (continued)

(f) Consider the hyperharmonic series

$$\sum \frac{1}{n^{p}} \quad \text{(Problem 2 of Chapter 4, §13).} \tag{5.10.34}$$

Let

$$f(x)=\frac{1}{x^{p}}, \quad x \geq 1. \tag{5.10.35}$$

If $p=1$, then $f(x)=1 / x$, so $\int_{1}^{x} f=\ln x \rightarrow+\infty$ as $x \rightarrow+\infty$. Hence $\sum 1 / n$ diverges.

If $p \neq 1$, then

$$\int_{1}^{\infty} f=\lim_{x \rightarrow+\infty} \int_{1}^{x} f=\lim_{x \rightarrow+\infty} \left.\frac{x^{1-p}}{1-p}\right|_{1}^{x}, \tag{5.10.36}$$

so $\int_{1}^{\infty} f$ converges or diverges according as $p>1$ or $p<1$, and the same applies to the series $\sum 1 / n^{p}$.

(g) Even nonregulated functions may be integrable. Such is Dirichlet's function (Example (c) in Chapter 4, §1). Explain, using the countability of the rationals.

5.10.E: Problems on Regulated Functions

Exercise $5.10.E.1$

Complete all details in the proof of Theorems $1-3$.

Exercise $5.10.E.1'$

Explain Examples $(a)-(g)$.

Exercise $5.10.E.2*$

Prove Note 2. More generally, assuming T to be complete, prove that if

$$g_n \rightarrow f(\text{ uniformly }) \text{ on } I=[a, b] \tag{5.10.E.1}$$

and if the g_n are regulated on I, so is f.
[Hint: Fix $p \in(a, b]$. Use Theorem 2 of Chapter 4, §11 with

$$X=[a, p], Y=N \cup\{+\infty\}, q=+\infty, \text { and } F(x, n)=g_n(x). \tag{5.10.E.2}$$

Then show that

$$f\left(p^{-}\right)=\lim _{x \rightarrow p^{-}} \lim _{n \rightarrow \infty} g_n(x) \text { exists; } \tag{5.10.E.3}$$

similarly for $f\left(p^{+}\right)$.]

Exercise $5.10.E.3$

Given $f, g: E^1 \rightarrow E^1$, define $f \vee g$ and $f \wedge g$ as in Problem 12 of Chapter 4, §8. Using the hint given there, show that $f \vee g$ and $f \wedge g$ are regulated if f and g are.

Exercise $5.10.E.4$

Show that the function $g \circ f$ need not be regulated even if g and f are.
[Hint: Let

$$f(x)=x \cdot \sin \frac{1}{x}, g(x)=\frac{x}{|x|}, \text { and } f(0)=g(0)=0 \text { with } I=[0,1]. \tag{5.10.E.4}$$

Proceed.]

Exercise $5.10.E.5$

$\Rightarrow$ Given $f: E^1 \rightarrow(T, \rho)$, regulated on I, put

$$j(p)=\max \left\{\rho\left(f(p), f\left(p^{-}\right)\right), \rho\left(f(p), f\left(p^{+}\right)\right), \rho\left(f\left(p^{-}\right), f\left(p^{+}\right)\right)\right\}; \tag{5.10.E.5}$$

call it the *jump* at p.
(i) Prove that f is discontinuous at $p \in I^0$ iff $j(p)>0$, i.e., iff

$$(\exists n \in N) \quad j(p) > \frac{1}{n}. \tag{5.10.E.6}$$

(ii) For a fixed $n \in N$, prove that a closed subinterval $J \subseteq I$ contains at most finitely many x with $j(x) > 1/n$.
[Hint: Otherwise, there is a sequence of distinct points $x_m \in J, j(x_m) > \frac{1}{n}$, hence a subsequence $x_{m_k} \to p \in J$. (Why?) Use Theorem 1 of Chapter 4, §2, to show that $f(p^-)$ or $f(p^+)$ fails to exist.]

Exercise 5.10.E.6

⇒ Show that if $f : E^1 \to (T, \rho)$ is regulated on I, then it has at most countably many discontinuities in I; all are of the "jump" type (Problem 5).
[Hint: By Problem 5, any closed subinterval $J \subseteq I$ contains, for each n, at most finitely many discontinuities x with $j(x) > 1/n$. Thus for $n = 1, 2, \ldots,$ obtain countably many such x.]

Exercise 5.10.E.7

Prove that if E is complete, all maps $f : E^1 \to E$, with $V_f[I] < +\infty$ on $I = [a, b]$, are regulated on I.
[Hint: Use Corollary 1, Chapter 4, §2, to show that $f(p^-)$ and $f(p^+)$ exist.
Say,

$$x_n \to p \text{ with } x_n < p \quad (x_n, p \in I), \tag{5.10.E.7}$$

but $\{f(x_n)\}$ is not Cauchy. Then find a subsequence, $\{x_{n_k}\} \uparrow$, and $\varepsilon > 0$ such that

$$|f(x_{n_{k+1}}) - f(x_{n_k})| \geq \varepsilon, \quad k = 1, 3, 5, \ldots \tag{5.10.E.8}$$

Deduce a contradiction to $V_f[I] < +\infty$.
Provide a similar argument for the case $x_n > p$.]

Exercise 5.10.E.8

Prove that if $f : E^1 \to (T, \rho)$ is regulated on I, then $\overline{f[B]}$ (the closure of $f[B]$) is compact in (T, ρ) whenever B is a compact subset of I.
[Hint: Given $\{z_m\}$ in $\overline{f[B]}$, find $\{y_m\} \subseteq f[B]$ such that $\rho(z_m, y_m) \to 0$ (use Theorem 3 of Chapter 3, §16). Then "imitate" the proof of Theorem 1 in Chap ter 4, §8 suitably. Distinguish the cases:
(i) all but finitely many x_m are $< p$;
(ii) infinitely many x_m exceed p; or
(iii) infinitely many x_m equal p.]

5.11: Integral Definitions of Some Functions

This page is a draft and is under active development.

By Theorem 2 in §10, $\int f$ exists on I whenever the function $f: E^{1} \rightarrow E$ is regulated on I, and E is complete. Hence whenever such an f is given, we can define a new function F by setting

$$F=\int_{a}^{x} f \tag{5.11.1}$$

on I for some $a \in I$. This is a convenient method of obtaining new continuous functions, differentiable on $I-Q$ (Q countable). We shall now apply it to obtain new definitions of some functions previously defined in a rather strenuous step-by-step manner.

I. Logarithmic and Exponential Functions. From our former definitions, we proved that

$$\ln x=\int_{1}^{x} \frac{1}{t} d t, \quad x>0. \tag{5.11.2}$$

Now we want to treat this as a definition of logarithms. We start by setting

$$f(t)=\frac{1}{t}, \quad t \in E^{1}, t \neq 0, \tag{5.11.3}$$

and $f(0)=0$.

Then f is continuous on $I=(0,+\infty)$ and $J=(-\infty, 0),$ so it has an exact primitive on I and J separately (not on E^{1}). Thus we can now define the log function on I by

$$\int_{1}^{x} \frac{1}{t} d t=\log x \text{ (also written } \ln x) \text{ for } x>0. \tag{5.11.4}$$

By the very definition of an exact primitive, the log function is continuous and differentiable on $I=(0,+\infty)$; its derivative on I is f. Thus we again have the symbolic formula

$$(\log x)^{\prime}=\frac{1}{x}, \quad x>0. \tag{5.11.5}$$

If $x<0$, we can consider $\log (-x)$. Then the chain rule (Theorem 3 of §1) yields

$$(\log (-x))^{\prime}=\frac{1}{x}. \quad \text{(Verify!)} \tag{5.11.6}$$

Hence

$$(\log |x|)^{\prime}=\frac{1}{x} \quad \text{for } x \neq 0. \tag{5.11.7}$$

Other properties of logarithms easily follow from (1). We summarize them now.

Theorem 5.11.1

(i) $\log 1=\int_{1}^{1} \frac{1}{t} d t=0$.

(ii) $\log x<\log y$ whenever $0<x<y$.

(iii) $\lim _{x \rightarrow+\infty} \log x=+\infty$ and $\lim _{x \rightarrow 0^{+}} \log x=-\infty$.

(iv) The range of log is all of E^{1}.

(v) For any positive $x, y \in E^{1}$,

$$\log (x y)=\log x+\log y \text{ and } \log \left(\frac{x}{y}\right)=\log x-\log y. \tag{5.11.8}$$

(vi) $\log a^r = r \cdot \log a, a > 0, r \in N$.

(vii) $\log e = 1$, where $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$.

Proof

(ii) By (2), $(\log x)' > 0$ on $I = (0, +\infty)$, so $\log x$ is increasing on I.

(iii) By Theorem 5 in §10,

$$\lim_{x\to+\infty} \log x = \int_1^\infty \frac{1}{t} dt = +\infty \tag{5.11.9}$$

since

$$\sum_{n=1}^{\infty} \frac{1}{n} = +\infty \quad \text{(Chapter 4, §13, Example (b)).} \tag{5.11.10}$$

Hence, substituting $y = 1/x$, we obtain

$$\lim_{y\to 0^+} \log y = \lim_{x\to+\infty} \log \frac{1}{x}. \tag{5.11.11}$$

However, by Theorem 2 in §5 (substituting $s = 1/t$),

$$\log \frac{1}{x} = \int_1^{1/x} \frac{1}{t} dt = -\int_1^x \frac{1}{s} ds = -\log x. \tag{5.11.12}$$

Thus

$$\lim_{y\to 0^+} \log y = \lim_{x\to+\infty} \log \frac{1}{x} = -\lim_{x\to+\infty} \log x = -\infty \tag{5.11.13}$$

as claimed. (We also proved that $\log \frac{1}{x} = -\log x$.)

(iv) Assertion (iv) now follows by the Darboux property (as in Chapter 4, §9, Example (b)).

(v) With x, y fixed, we substitute $t = xs$ in

$$\int_1^{xy} \frac{1}{t} dt = \log xy \tag{5.11.14}$$

and obtain

$$\begin{aligned} \log xy &= \int_1^{xy} \frac{1}{t} dt = \int_{1/x}^{y} \frac{1}{s} ds \\ &= \int_{1/x}^{1} \frac{1}{s} ds + \int_1^y \frac{1}{s} ds \\ &= -\log \frac{1}{x} + \log y \\ &= \log x + \log y. \end{aligned}$$

Replacing y by $1/y$ here, we have

$$\log \frac{x}{y} = \log x + \log \frac{1}{y} = \log x - \log y. \tag{5.11.15}$$

Thus (v) is proved, and (vi) follows by induction over r.

(vii) By continuity,

$$\log e = \lim_{x\to e} \log x = \lim_{n\to\infty} \log\left(1+\frac{1}{n}\right)^n = \lim_{n\to\infty} \frac{\log(1+1/n)}{1/n}, \tag{5.11.16}$$

where the last equality follows by (vi). Now, L'Hôpital's rule yields

$$\lim_{x\to 0}\frac{\log(1+x)}{x}=\lim_{x\to 0}\frac{1}{1+x}=1. \tag{5.11.17}$$

Letting x run over $\frac{1}{n}\to 0$, we get (vii). $\square$

Note 1. Actually, (vi) holds for any $r\in E^1$, with a^r as in Chapter 2, §§11-12. One uses the techniques from that section to prove it first for rational r, and then it follows for all real r by continuity. However, we prefer not to use this now.

Next, we define the exponential function ("exp") to be the inverse of the log function. This inverse function exists; it is continuous (even differentiable) and strictly increasing on its domain (by Theorem 3 of Chapter 4, §9 and Theorem 3 of Chapter 5, §2) since the log function has these properties. From $(\log x)'=1/x$ we get, as in 2,

$$(\exp x)'=\exp x \quad (\text{cf. §2, Example (B)}). \tag{5.11.18}$$

The domain of the exponential is the range of its inverse, i.e., E^1 (cf. Theorem 1(iv)). Thus $\exp x$ is defined for all $x\in E^1$. The range of exp is the domain of log, i.e., $(0,+\infty)$. Hence exp $x>0$ for all $x\in E^1$. Also, by definition,

$$\begin{aligned}\exp(\log x)&=x \text{ for } x>0\\ \exp 0&=1 \text{ (cf. Theorem 1(i)), and}\\ \exp r&=e^r \text{ for } r\in N.\end{aligned}$$

Indeed, by Theorem 1(vi) and (vii), log $e^r=r\cdot\log e=r$. Hence (6) follows. If the definitions and rules of Chapter 2, §§11-12 are used, this proof even works for any r by Note 1. Thus our new definition of exp agrees with the old one.

Our next step is to give a new definition of a^r, for any $a,r\in E^1(a>0)$. We set

$$\begin{aligned}a^r&=\exp(r\cdot\log a) \text{ or}\\ \log a^r&=r\cdot\log a \quad \left(r\in E^1\right).\end{aligned}$$

In case $r\in N$, (8) becomes Theorem 1(vi). Thus for natural r, our new definition of a^r is consistent with the previous one. We also obtain, for $a,b>0$,

$$(ab)^r=a^rb^r;\quad a^{rs}=(a^r)^s;\quad a^{r+s}=a^ra^s;\quad \left(r,s\in E^1\right). \tag{5.11.19}$$

The proof is by taking logarithms. For example,

$$\begin{aligned}\log(ab)^r&=r\log ab=r(\log a+\log b)=r\cdot\log a+r\cdot\log b\\ &=\log a^r+\log b^r=\log(a^rb^r).\end{aligned}$$

Thus $(ab)^r=a^rb^r$. Similar arguments can be given for the rest of (9) and other laws stated in Chapter 2, §§11-12.

We can now define the exponential to the base $a(a>0)$ and its inverse, $\log_a$, as before (see the example in Chapter 4, §5 and Example (b) in Chapter 4, §9). The differentiability of the former is now immediate from (7), and the rest follows as before.

II. Trigonometric Functions. These shall now be defined in a precise analytic manner (not based on geometry).

We start with an integral definition of what is usually called the principal value of the arcsine function,

$$\arcsin x=\int_0^x\frac{1}{\sqrt{1-t^2}}dt. \tag{5.11.20}$$

We shall denote it by $F(x)$ and set

$$f(x)=\frac{1}{\sqrt{1-x^2}} \text{ on } I=(-1,1). \tag{5.11.21}$$

($F=f=0$ on E^1-I.) Thus by definition, $F=\int f$ on I.

Note that $\int f$ exists and is exact on I since f is continuous on I. Thus

$$F'(x)=f(x)=\frac{1}{\sqrt{1-x^2}}>0 \quad \text{for } x\in I, \tag{5.11.22}$$

and so F is strictly increasing on I. Also, $F(0)=\int_0^0 f=0$.

We also define the number π by setting

$$\frac{\pi}{2}=2\arcsin\sqrt{\frac{1}{2}}=2F(c)=2\int_0^c f,\quad c=\sqrt{\frac{1}{2}}. \tag{5.11.23}$$

Then we obtain the following theorem.

Theorem 5.11.2

F has the limits

$$F\left(1^{-}\right)=\frac{\pi}{2} \text{ and } F\left(-1^{+}\right)=-\frac{\pi}{2}. \tag{5.11.24}$$

Thus F becomes relatively continuous on $\bar{I}=[-1,1]$ if one sets

$$F(1)=\frac{\pi}{2} \text{ and } F(-1)=-\frac{\pi}{2}, \tag{5.11.25}$$

i.e.,

$$\arcsin 1=\frac{\pi}{2} \text{ and } \arcsin(-1)=-\frac{\pi}{2}. \tag{5.11.26}$$

Proof

We have

$$F(x)=\int_0^x f=\int_0^c f+\int_c^x f,\quad c=\sqrt{\frac{1}{2}}. \tag{5.11.27}$$

By substituting $s=\sqrt{1-t^2}$ in the last integral and setting, for brevity, $y=\sqrt{1-x^2}$, we obtain

$$\int_c^x f=\int_c^x \frac{1}{\sqrt{1-t^2}}dt=\int_y^c \frac{1}{\sqrt{1-s^2}}ds=F(c)-F(y).\quad \text{(Verify!)} \tag{5.11.28}$$

Now as $x\rightarrow 1^{-}$, we have $y=\sqrt{1-x^2}\rightarrow 0$, and hence $F(y)\rightarrow F(0)=0$ (for F is continuous at 0). Thus

$$F\left(1^{-}\right)=\lim_{x\rightarrow 1^{-}} F(x)=\lim_{y\rightarrow 0}\left(\int_0^c f+\int_y^c f\right)=\int_0^c f+F(c)=2\int_0^c f=\frac{\pi}{2}. \tag{5.11.29}$$

Similarly, one gets $F\left(-1^{+}\right)=-\pi/2$. $\square$

The function F as redefined in Theorem 2 will be denoted by F_0. It is a primitive of f on the closed interval $\bar{I}$ (exact on I). Thus $F_0(x)=\int_0^x f,\ -1\leq x\leq 1$, and we may now write

$$\frac{\pi}{2}=\int_0^1 f \text{ and } \pi=\int_{-1}^0 f+\int_0^1 f=\int_{-1}^1 f. \tag{5.11.30}$$

Note 2. In classical analysis, the last integrals are regarded as so-called improper integrals, i.e., limits of integrals rather than integrals proper. In our theory, this is unnecessary since F_0 is a genuine primitive of f on $\bar{I}$.

For each integer n (negatives included), we now define $F_n: E^1\rightarrow E^1$ by

$$\begin{aligned} F_n(x)&=n\pi+(-1)^n F_0(x) \text{ for } x\in\bar{I}=[-1,1] \\ F_n&=0 \qquad\qquad \text{on } -\bar{I}. \end{aligned}$$

F_n is called the n th branch of the arcsine. Figure 26 shows the graphs of F_0 and F_1 (that of F_1 is dotted). We now obtain the following theorem.

Theorem 5.11.3

(i) Each F_n is differentiable on $I=(-1,1)$ and relatively continuous on $\bar{I}=[-1,1]$.

(ii) F_n is increasing on $\bar{I}$ if n is even, and decreasing if n is odd.

(iii) $F_n'(x)=\frac{(-1)^n}{\sqrt{1-x^2}}$ on I.

(iv) $F_n(-1)=F_{n-1}(-1)=n\pi-(-1)^n\frac{\pi}{2}; F_n(1)=F_{n-1}(1)=n\pi+(-1)^n\frac{\pi}{2}$.

Proof

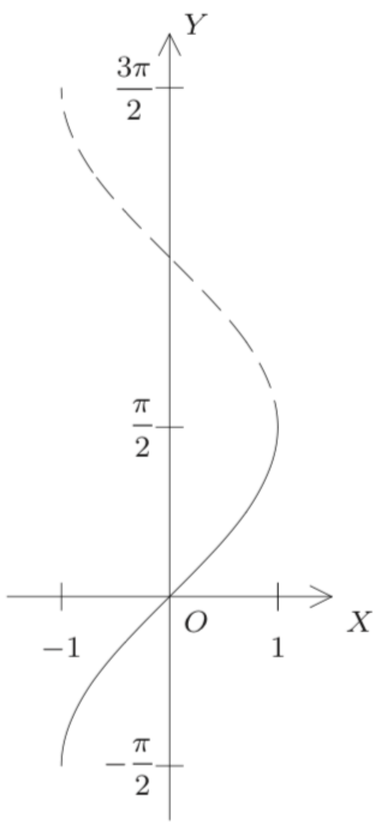

FIGURE 26

The proof is obvious from (12) and the properties of F_0. Assertion (iv) ensures that the graphs of the F_n add up to one curve. By (ii), each F_n is one to one (strictly monotone) on $\bar{I}$. Thus it has a strictly monotone inverse on the interval $\overline{J_n}=F_n\ [\![-1,1]\!]$, i.e., on the F_n -image of $\bar{I}$. For simplicity, we consider only

$$\overline{J_0}=\left[-\frac{\pi}{2},\frac{\pi}{2}\right] \text{ and } J_1=\left[\frac{\pi}{2},\frac{3\pi}{2}\right], \tag{5.11.31}$$

as shown on the Y -axis in Figure 26. On these, we define for $x\in\overline{J_0}$

$$\sin x=F_0^{-1}(x) \tag{5.11.32}$$

and

$$\cos x=\sqrt{1-\sin^2 x}, \tag{5.11.33}$$

and for $x \in \overline{J_1}$

$$\sin x = F_1^{-1}(x) \tag{5.11.34}$$

and

$$\cos x = -\sqrt{1-\sin^2 x}. \tag{5.11.35}$$

On the rest of E^1, we define $\sin x$ and $\cos x$ periodically by setting

$$\sin(x+2n\pi) = \sin x \text{ and } \cos(x+2n\pi) = \cos x, \quad n = 0, \pm 1, \pm 2, \ldots. \tag{5.11.36}$$

Note that by Theorem 3(iv),

$$F_0^{-1}\left(\frac{\pi}{2}\right) = F_1^{-1}\left(\frac{\pi}{2}\right) = 1. \tag{5.11.37}$$

Thus (13) and (14) both yield $\sin \pi/2 = 1$ for the common endpoint $\pi/2$ of $\overline{J_0}$ and $\overline{J_1}$, so the two formulas are consistent. We also have

$$\sin\left(-\frac{\pi}{2}\right) = \sin\left(\frac{3\pi}{2}\right) = -1, \tag{5.11.38}$$

in agreement with (15). Thus the sine and cosine functions (briefly, s and c) are well defined on E^1.

Theorem 5.11.4

The sine and cosine functions (s and c) are differentiable, hence continuous, on all of E^1, with derivatives $s' = c$ and $c' = -s$; that is,

$$(\sin x)' = \cos x \text{ and } (\cos x)' = -\sin x. \tag{5.11.39}$$

Proof

It suffices to consider the intervals $\overline{J_0}$ and $\overline{J_1}$, for by (15), all properties of s and c repeat themselves, with period 2π, on the rest of E^1.

By (13),

$$s = F_0^{-1} \text{ on } \overline{J_0} = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \tag{5.11.40}$$

where F_0 is differentiable on $I = (-1, 1)$. Thus Theorem 3 of §2 shows that s is differentiable on $J_0 = (-\pi/2, \pi/2)$ and that

$$s'(q) = \frac{1}{F_0'(p)} \text{ whenever } p \in I \text{ and } q = F_0(p); \tag{5.11.41}$$

i.e., $q \in J$ and $p = s(q)$. However, by Theorem 3(iii),

$$F_0'(p) = \frac{1}{\sqrt{1-p^2}}. \tag{5.11.42}$$

Hence,

$$s'(q) = \sqrt{1-\sin^2 q} = \cos q = c(q), \quad q \in J. \tag{5.11.43}$$

This proves the theorem for interior points of $\overline{J_0}$ as far as s is concerned.

As

$$c = \sqrt{1-s^2} = \left(1-s^2\right)^{\frac{1}{2}} \text{ on } J_0 \text{ (by (13))}, \tag{5.11.44}$$

we can use the chain rule (Theorem 3 in §1) to obtain

$$c' = \frac{1}{2}\left(1-s^2\right)^{-\frac{1}{2}}(-2s)s' = -s \tag{5.11.45}$$

on noting that $s' = c = \left(1-s^2\right)^{\frac{1}{2}}$ on J_0. Similarly, using (14), one proves that $s' = c$ and $c' = -s$ on J_1 (interior of $\overline{J_1}$). Next, let q be an endpoint, say, $q = \pi/2$. We take the left derivative

$$s'_-(q) = \lim_{x\to q^-} \frac{s(x)-s(q)}{x-q}, \quad x \in J_0. \tag{5.11.46}$$

By L'Hôpital's rule, we get

$$s'_-(q) = \lim_{x\to q^-} \frac{s'(x)}{1} = \lim_{x\to q^-} c(x) \tag{5.11.47}$$

since $s' = c$ on J_0. However, $s = F_0^{-1}$ is left continuous at q (why?); hence so is $c = \sqrt{1-s^2}$. (Why?) Therefore,

$$s'_-(q) = \lim_{x\to q^-} c(x) = c(q), \quad \text{as required.} \tag{5.11.48}$$

Similarly, one shows that $s'_+(q) = c(q)$. Hence $s'(q) = c(q)$ and $c'(q) = -s(q)$ as before. □

The other trigonometric functions reduce to s and c by their defining formulas

$$\tan x = \frac{\sin x}{\cos x}, \cot x = \frac{\cos x}{\sin x}, \sec x = \frac{1}{\cos x}, \text{ and } \csc x = \frac{1}{\sin x}, \tag{5.11.49}$$

so we shall not dwell on them in detail. The various trigonometric laws easily follow from our present definitions; for hints, see the problems below.

5.11.E: Problems on Exponential and Trigonometric Functions

Exercise $5.11.E.1$

Verify formula (2).

Exercise $5.11.E.2$

Prove Note 1, as suggested (using Chapter 2, §§$11-12$).

Exercise $5.11.E.3$

Prove formulas (1) of Chapter 2, §§$11-12$ from our new definitions.

Exercise $5.11.E.4$

Complete the missing details in the proofs of Theorems $2-4$.

Exercise $5.11.E.5$

Prove that
(i) $\sin 0=\sin (n \pi)=0$;
(ii) $\cos 0=\cos (2 n \pi)=1$;
(iii) $\sin \frac{\pi}{2}=1$;
(iv) $\sin \left(-\frac{\pi}{2}\right)=-1$;
(v) $\cos \left(\pm \frac{\pi}{2}\right)=0$;
(vi) $|\sin x| \leq 1$ and $|\cos x| \leq 1$ for $x \in E^{1}$.

Exercise $5.11.E.6$

Prove that
(i) $\sin (-x)=-\sin x$ and
(ii) $\cos (-x)=\cos x$ for $x \in E^{1}$.
[Hint: For (i), let $h(x)=\sin x+\sin (-x)$. Show that $h^{\prime}=0$; hence h is constant, say, $h=q$ on E^{1}. Substitute $x=0$ to find q. For (ii), use $(13)-(15)$.]

Exercise $5.11.E.7$

Prove the following for $x, y \in E^{1}$:
(i) $\sin (x+y)=\sin x \cos y+\cos x \sin y$; hence $\sin \left(x+\frac{\pi}{2}\right)=\cos x$.
(ii) $\cos (x+y)=\cos x \cos y-\sin x \sin y$; hence $\cos \left(x+\frac{\pi}{2}\right)=-\sin x$.
[Hint for (i): Fix x, y and let $p=x+y$. Define $h: E^{1} \rightarrow E^{1}$ by

$$h(t)=\sin t \cos (p-t)+\cos t \sin (p-t), \quad t \in E^{1}. \tag{5.11.E.1}$$

Proceed as in Problem 6. Then let $t=x$.]

Exercise $5.11.E.8$

With $\overline{J_{n}}$ as in the text, show that the sine increases on $\overline{J_{n}}$ if n is even and decreases if n is odd. How about the cosine? Find the endpoints of $\overline{J_{n}}$.

CHAPTER OVERVIEW

6: Differentiation on E^n and Other Normed Linear Spaces

6.1: Directional and Partial Derivatives

In Chapter 5 we considered functions $f : E^1 \rightarrow E$ of one real variable.

Now we take up functions $f : E' \rightarrow E$ where both E' and E are normed spaces.

The scalar field of both is always assumed the same: E^1 or C (the complex field). The case $E = E^*$ is excluded here; thus all is assumed finite.

We mostly use arrowed letters $\vec{p}, \vec{q}, \ldots, \vec{x}, \vec{y}, \vec{z}$ for vectors in the domain space E', and nonarrowed letters for those in E and for scalars.

As before, we adopt the convention that f is defined on all of E', with $f(\vec{x}) = 0$ if not defined otherwise.

Note that, if $\vec{p} \in E'$, one can express any point $\vec{x} \in E'$ as

$$\vec{x} = \vec{p} + t\vec{u}, \tag{6.1.1}$$

with $t \in E^1$ and $\vec{u}$ a unit vector. For if $\vec{x} \neq \vec{p}$, set

$$t = |\vec{x} - \vec{p}| \text{ and } \vec{u} = \frac{1}{t}(\vec{x} - \vec{p}); \tag{6.1.2}$$

and if $\vec{x} = \vec{p}$, set $t = 0$, and any$\vec{u}$ will do. We often use the notation

$$\vec{t} = \Delta\vec{x} = \vec{x} - \vec{p} = t\vec{u} \quad \left(t \in E^1, \vec{t}, \vec{u} \in E'\right). \tag{6.1.3}$$

First of all, we generalize Definition 1 in Chapter 5, §1.

Definition 1

Given $f : E' \rightarrow E$ and $\vec{p}, \vec{u} \in E'(\vec{u} \neq \overrightarrow{0})$, we define the directional derivative of f along $\vec{u}$ (or $\vec{u}$ -directed derivative of f) at $\vec{p}$ by

$$D_{\vec{u}} f(\vec{p}) = \lim_{t \rightarrow 0} \frac{1}{t}[f(\vec{p} + t\vec{u}) - f(\vec{p})], \tag{6.1.4}$$

if this limit exists in E (finite).

We also define the $\vec{u}$ -directed derived function,

$$D_{\vec{u}} f : E' \rightarrow E, \tag{6.1.5}$$

as follows. For any $\vec{p} \in E'$,

$$D_{\vec{u} f(\vec{p})} = \begin{cases} \lim_{t \rightarrow 0} \frac{1}{t}[f(\vec{p} + t\vec{u}) - f(\vec{p})] & \text{if this limit exists,} \\ 0 & \text{otherwise.} \end{cases} \tag{6.1.6}$$

Thus $D_{\vec{u}} f$ is always defined, but the name derivative is used only if the limit (1) exists (finite). If it exists for each $\vec{p}$ in a set $B \subseteq E'$, we call $D_{\vec{u}} f$ (in classical notation $\partial f / \partial \vec{u}$) the $\vec{u}$ -directed derivative of f on B.

Note that, as $t \rightarrow 0, \vec{x}$ tends to $\vec{p}$ over the line $\vec{x} = \vec{p} + t\vec{u}$. Thus $D_{\vec{u}} f(\vec{p})$ can be treated as a relative limit over that line. Observe that it depends on both the direction and the length of $\vec{u}$. Indeed, we have the following result.

Corollary 6.1.1

Given $f : E' \rightarrow E, \vec{u} \neq \overrightarrow{0}$, and a scalar $s \neq 0$, we have

$$D_{s\vec{u}} f = s D_{\vec{u}} f. \tag{6.1.7}$$

Moreover, $D_{s\vec{u}} f(\vec{p})$ is a genuine derivative iff $D_{\vec{u}} f(\vec{p})$ is.

Proof

Set $t=s\theta$ in (1) to get

$$sD_{\vec{u}}f(\vec{p})=\lim_{\theta\to 0}\frac{1}{\theta}[f(\vec{p}+\theta s\vec{u})-f(\vec{p})]=D_{s\vec{u}}f(\vec{p}).\quad\square \tag{6.1.8}$$

In particular, taking $s=1/|\vec{u}|$, we have

$$|s\vec{u}|=\frac{|\vec{u}|}{|\vec{u}|}=1 \text{ and } D_{\vec{u}}f=\frac{1}{s}D_{s\vec{u}}f. \tag{6.1.9}$$

Thus all reduces to the case $D_{\vec{v}}f$, where $\vec{v}=s\vec{u}$ is a unit vector. This device, called normalization, is often used, but actually it does not simplify matters.

If $E'=E^n\left(C^n\right)$, then f is a function of n scalar variables $x_k(k=1,\ldots,n)$ and E' has the n basic unit vectors $\vec{e}_k$. This example leads us to the following definition.

Definition 2

If in formula (1), $E'=E^n\left(C^n\right)$ and $\vec{u}=\vec{e}_k$ for a fixed $k\leq n$, we call $D_{\vec{u}}f$ the partially derived function for f, with respect to x_k, denoted

$$D_kf \text{ or } \frac{\partial f}{\partial x_k}, \tag{6.1.10}$$

and the limit (1) is called the partial derivative of f at $\vec{p}$, with respect to x_k, denoted

$$D_kf(\vec{p}), \text{ or } \frac{\partial}{\partial x_k}f(\vec{p}), \text{ or } \left.\frac{\partial f}{\partial x_k}\right|_{\vec{x}=\vec{p}}. \tag{6.1.11}$$

If it exists for all $\vec{p}\in B$, we call D_kf the partial derivative (briefly, partial) of f on B, with respect to x_k.

In any case, the derived functions $D_kf(k=1,\ldots,n)$ are always defined on all of $E^n\left(C^n\right)$.

If $E'=E^3\left(C^3\right)$, we often write x,y,z for x_1,x_2,x_3, and

$$\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z} \text{ for } D_kf \quad (k=1,2,3). \tag{6.1.12}$$

Note 1. If $E'=E^1$, scalars are also "vectors," and D_1f coincides with f' as defined in Chapter 5, §1 (except where $f'=\pm\infty$). Explain!

Note 2. As we have observed, the $\vec{u}$ -directed derivative (1) is obtained by keeping $\vec{x}$ on the line $\vec{x}=\vec{p}+t\vec{u}$.

If $\vec{u}=\vec{e}_k$, the line is parallel to the kth axis; so all coordinates of $\vec{x}$, except x_k, remain fixed $\left(x_i=p_i, i\neq k\right)$, and f behaves like a function of one variable, x_k. Thus we can compute D_kf by the usual rules of differentiation, treating all $x_i(i\neq k)$ as constants and x_k as the only variable.

For example, let $f(x,y)=x^2y$. Then

$$\frac{\partial f}{\partial x}=D_1f(x,y)=2xy \text{ and } \frac{\partial f}{\partial y}=D_2f(x,y)=x^2. \tag{6.1.13}$$

Note 3. More generally, given $\vec{p}$ and $\vec{u}\neq\overrightarrow{0}$, set

$$h(t)=f(\vec{p}+t\vec{u}),\quad t\in E^1. \tag{6.1.14}$$

Then $h(0)=f(\vec{p})$; so

$$\begin{aligned} D_{\vec{u}} f(\vec{p}) &= \lim_{t \to 0} \frac{1}{t}[f(\vec{p}+t\vec{u}) - f(\vec{p})] \\ &= \lim_{t \to 0} \frac{h(t)-h(0)}{t-0} \\ &= h'(0) \end{aligned}$$

if the limit exists. Thus all reduces to a function h of one real variable.

For functions $f : E^1 \to E$, the existence of a finite derivative ("differentiability") at p implies continuity at p (Theorem 1 of Chapter 5, §1). But in the general case, $f : E' \to E$, this may fail even if $D_{\vec{u}} f(\vec{p})$ exists for all $\vec{u} \neq \overrightarrow{0}$.

Examples

(a) Define $f : E^2 \to E^1$ by

$$f(x, y) = \frac{x^2 y}{x^4 + y^2}, \quad f(0,0) = 0. \tag{6.1.15}$$

Fix a unit vector $\vec{u} = (u_1, u_2)$ in E^2. Let $\vec{p} = (0,0)$. To find $D_{\vec{u}} f(p)$, use the h of Note 3 :

$$h(t) = f(\vec{p}+t\vec{u}) = f(t\vec{u}) = f(tu_1, tu_2) = \frac{tu_1^2 u_2}{t^2 u_1^4 + u_2^2} \text{ if } u_2 \neq 0, \tag{6.1.16}$$

and $h = 0$ if $u_2 = 0$. Hence

$$D_{\vec{u}} f(\vec{p}) = h'(0) = \frac{u_1^2}{u_2} \text{ if } u_2 \neq 0, \tag{6.1.17}$$

and $h'(0) = 0$ if $u_2 = 0$. Thus $D_{\vec{u}}(\overrightarrow{0})$ exists for all $\vec{u}$. Yet f is discontinuous at $\overrightarrow{0}$ (see Problem 9 in Chapter 4, §3).

(b) Let

$$f(x, y) = \begin{cases} x + y & \text{if } xy = 0, \\ 1 & \text{otherwise.} \end{cases} \tag{6.1.18}$$

Then $f(x, y) = x$ on the x-axis; so $D_1 f(0,0) = 1$.

Yet f is discontinuous at $\overrightarrow{0}$ (even relatively so) over any line $y = ax$ $(a \neq 0)$. For on that line, $f(x, y) = 1$ if $(x, y) \neq (0,0)$; so $f(x, y) \to 1$ but $f(0,0) = 0 + 0 = 0$.

Thus continuity at $\overrightarrow{0}$ fails. (But see Theorem 1 below!)

Hence, if differentiability is to imply continuity, it must be defined in a stronger manner. We do it in §3. For now, we prove only some theorems on partial and directional derivatives, based on those of Chapter 5.

Theorem 6.1.1

If $f : E' \to E$ has a $\vec{u}$-directed derivative at $\vec{p} \in E'$, then f is relatively continuous at $\vec{p}$ over the line

$$\vec{x} = \vec{p} + t\vec{u} \quad \left(\overrightarrow{0} \neq \vec{u} \in E'\right). \tag{6.1.19}$$

Proof

Set $h(t) = f(\vec{p} + t\vec{u}), t \in E^1$.

By Note 3, our assumption implies that h (a function on E^1) is differentiable at 0.

By Theorem 1 in Chapter 5, §1, then, h is continuous at 0; so

$$\lim_{t \to 0} h(t) = h(0) = f(\vec{p}), \tag{6.1.20}$$

i.e.,

$$\lim_{t\to 0} f(\vec{p}+t\vec{u})=f(\vec{p}). \tag{6.1.21}$$

But this means that $f(\vec{x})\to f(\vec{p})$ as $\vec{x}\to\vec{p}$ over the line $\vec{x}=\vec{p}+t\vec{u},$ for, on that line, $\vec{x}=\vec{p}+t\vec{u}$.

Thus, indeed, f is relatively continuous at $\vec{p}$, as stated. □

Note that we actually used the substitution $\vec{x}=\vec{p}+t\vec{u}$. This is admissible since the dependence between x and t is one-to-tone (Corollary 2(iii) of Chapter 4, §2). Why?

Theorem 6.1.2

Let $E'\ni\vec{u}=\vec{q}-\vec{p},\vec{u}\neq\overrightarrow{0}$.

If $f: E'\to E$ is relatively continuous on the segment $I=L[\vec{p},\vec{q}]$ and has a $\vec{u}$-directed derivative on $I-Q$ (Q countable), then

$$|f(\vec{q})-f(\vec{p})|\leq\sup|D_{\vec{u}}f(\vec{x})|, \quad \vec{x}\in I-Q. \tag{6.1.22}$$

Proof

Set again $h(t)=f(\vec{p}+t\vec{u})$ and $g(t)=\vec{p}+t\vec{u}$.

Then $h=f\circ g,$ and g is continuous on E^1. (Why?)

As f is relatively continuous on $I=L[\vec{p},\vec{q}]$, so is $h=f\circ g$ on the interval $J=[0,1]\subset E^1$ (cf. Chapter 4, §8, Example (1)).

Now fix $t_0\in J$. If $\vec{x}_0=\vec{p}+t_0\vec{u}\in I-Q,$ our assumptions imply the existence of

$$\begin{aligned} D_{\vec{u}}f\left(\vec{x}_0\right) &=\lim_{t\to 0}\frac{1}{t}\left[f\left(\vec{x}_0+t\vec{u}\right)-f\left(\vec{x}_0\right)\right] \\ &=\lim_{t\to 0}\frac{1}{t}\left[f\left(\vec{p}+t_0\vec{u}+t\vec{u}\right)-f\left(\vec{p}+t_0\vec{u}\right)\right] \\ &=\lim_{t\to 0}\frac{1}{t}\left[h\left(t_0+t\right)-h\left(t_0\right)\right] \\ &=h'\left(t_0\right). \quad \text{(Explain!)} \end{aligned}$$

This can fail for at most a countable set Q' of points $t_0\in J$ (those for which $\vec{x}_0\in Q$).

Thus h is differentiable on $J-Q'$; and so, by Corollary 1 in Chapter 5, §4,

$$|h(1)-h(0)|\leq\sup_{t\in J-Q'}|h'(t)|=\sup_{\vec{x}\in I-Q}|D_{\vec{u}}f(\vec{x})|. \tag{6.1.23}$$

Now as $h(1)=f(\vec{p}+\vec{u})=f(\vec{q})$ and $h(\overrightarrow{0})=f(\vec{p}),$ formula (2) follows. □

Theorem 6.1.3

If in Theorem 2, $E=E^1$ and if f has a $\vec{u}$-directed derivative at least on the open line segment $L(\vec{p},\vec{q})$, then

$$f(\vec{q})-f(\vec{p})=D_{\vec{u}}f\left(\vec{x}_0\right) \tag{6.1.24}$$

for some $\vec{x}_0\in L(\vec{p},\vec{q})$.

Proof

The proof is as in Theorem 2, based on Corollary 3 in Chapter 5, §2 (instead of Corollary 1 in Chapter 5, §4).

Theorems 2 and 3 are often used in "normalized" form, as follows.

Corollary 6.1.2

If in Theorems 2 and 3, we set

$$r=|\vec{u}|=|\vec{q}-\vec{p}| \neq 0 \text{ and } \vec{v}=\frac{1}{r} \vec{u}, \tag{6.1.25}$$

then formulas (2) and (3) can be written as

$$|f(\vec{q})-f(\vec{p})| \leq|\vec{q}-\vec{p}| \sup \left|D_{\vec{v}} f(\vec{x})\right|, \quad \vec{x} \in I-Q, \tag{6.1.26}$$

and

$$f(\vec{q})-f(\vec{p})=|\vec{q}-\vec{p}| D_{\vec{v}} f\left(\vec{x}_{0}\right) \tag{6.1.27}$$

for some $\vec{x}_{0} \in L(\vec{p}, \vec{q})$.

For by Corollary 1,

$$D_{\vec{u}} f=r D_{\vec{v}} f=|\vec{q}-\vec{p}| D_{\vec{v}} f; \tag{6.1.28}$$

so (2') and (3') follow.

6.1.E: Problems on Directional and Partial Derivatives

Exercise $6.1.E.1$

Complete all missing details in the proof of Theorems 1 to 3 and Corollaries 1 and 2.

Exercise $6.1.E.2$

Complete all details in Examples (a) and (b). Find $D_1 f(\vec{p})$ and $D_2 f(\vec{p})$ also for $\vec{p} \neq 0$. Do Example (b) in two ways: (i) use Note 3; (ii) use Definition 2 only.

Exercise $6.1.E.3$

In Examples (a) and (b) describe $D_{\vec{u}} f : E^2 \rightarrow E^1$. Compute it for \(\vec{u =(1,1)=\vec{p}.\)
In (b), show that f has no directional derivatives $D_{\vec{u}} f(\vec{p})$ except if $\vec{u} \| \vec{e}_1$ or $\vec{u} \| \vec{e}_2$. Give two proofs: (i) use Theorem 1; (ii) use definitions only.

Exercise $6.1.E.4$

Prove that if $f : E^n\left(C^n\right) \rightarrow E$ has a zero partial derivative, $D_k f = 0$, on a convex set A, then $f(\vec{x})$ does not depend on x_k, for $\vec{x} \in A$. (Use Theorems 1 and 2.)

Exercise $6.1.E.5$

Describe $D_1 f$ and $D_2 f$ on the various parts of E^2, and discuss the relative continuity of f over lines through $\overrightarrow{0}$, given that $f(x, y)$ equals:

$$\begin{array}{ll} \text{(i) } \frac{xy}{x^2+y^2}; & \text{(ii) the integral part of } x+y; \\ \text{(iii) } \frac{xy}{|x|}+x\sin\frac{1}{y}; & \text{(iv) } xy\frac{x^2-y^2}{x^2+y^2}; \\ \text{(v) } \sin(y\cos x); & \text{(vi) } x^y. \end{array} \tag{6.1.E.1}$$

(Set $f = 0$ wherever the formula makes no sense.)

Exercise $6.1.E.6$

$\Rightarrow$ Prove that if $f : E' \rightarrow E^1$ has a local maximum or minimum at $\vec{p} \in E'$, then $D_{\vec{u}} f(\vec{p}) = 0$ for every vector $\vec{u} \neq \overrightarrow{0}$ in E'.
[Hint: Use Note 3, then Corollary 1 in Chapter 5, §2.]

Exercise $6.1.E.7$

State and prove the Finite Increments Law (Theorem 1 of Chapter 5, §4) for directional derivatives.
[Hint: Imitate Theorem 2 using two auxiliary functions, h and k.]

Exercise $6.1.E.8$

State and prove Theorems 4 and 5 of Chapter 5, §1, for directional derivatives.

6.2: Linear Maps and Functionals. Matrices

For an adequate definition of differentiability, we need the notion of a linear map. Below, E', E'', and E denote normed spaces over the same scalar field, E^1 or C.

Definition 1

A function $f : E' \rightarrow E$ is a linear map if and only if for all $\vec{x}, \vec{y} \in E'$ and scalars a, b

$$f(a\vec{x}+b\vec{y})=af(\vec{x})+bf(\vec{y}); \tag{6.2.1}$$

equivalently, iff for all such $\vec{x}, \vec{y}$, and a

$$f(\vec{x}+\vec{y})=f(x)+f(y) \text{ and } f(a\vec{x})=af(\vec{x}). \text{ (Verify!)} \tag{6.2.2}$$

If $E=E'$, such a map is also called a linear operator.

If the range space E is the scalar field of E', (i.e., E^1 or C,) the linear f is also called a (real or complex) linear functional on E'.

Note 1. Induction extends formula (1) to any "linear combinations":

$$f\left(\sum_{i=1}^{m} a_i \vec{x}_i\right)=\sum_{i=1}^{m} a_i f\left(\vec{x}_i\right) \tag{6.2.3}$$

for all $\vec{x}_i \in E'$ and scalars a_i.

Briefly: A linear map f preserves linear combinations.

Note 2. Taking $a=b=0$ in (1), we obtain $f(\overrightarrow{0})=0$ if f is linear.

Examples

(a) Let $E'=E^n\left(C^n\right)$. Fix a vector $\vec{v}=(v_1, \ldots, v_n)$ in E' and set

$$(\forall \vec{x} \in E') \quad f(\vec{x})=\vec{x} \cdot \vec{v} \tag{6.2.4}$$

(inner product; see Chapter 3, §§1-3 and §9).

Then

$$\begin{aligned} f(a\vec{x}+b\vec{y}) &=(a\vec{x}) \cdot \vec{v}+(b\vec{y}) \cdot \vec{v} \\ &=a(\vec{x} \cdot \vec{v})+b(\vec{y} \cdot \vec{v}) \\ &=af(\vec{x})+bf(\vec{y}); \end{aligned}$$

so f is linear. Note that if $E'=E^n$, then by definition,

$$f(\vec{x})=\vec{x} \cdot \vec{v}=\sum_{k=1}^{n} x_k v_k=\sum_{k=1}^{n} v_k x_k. \tag{6.2.5}$$

If, however, $E'=C^n$, then

$$f(\vec{x})=\vec{x} \cdot \vec{v}=\sum_{k=1}^{n} x_k \bar{v}_k=\sum_{k=1}^{n} \bar{v}_k x_k, \tag{6.2.6}$$

where $\bar{v}_k$ is the conjugate of the complex number v_k.

By Theorem 3 in Chapter 4, §3, f is continuous (a polynomial!).

Moreover, $f(\vec{x})=\vec{x} \cdot \vec{v}$ is a scalar (in E^1 or C). Thus the range of f lies in the scalar field of E'; so f is a linear functional on E'.

(b) Let $I=[0,1]$. Let E' be the set of all functions $u: I \rightarrow E$ that are of class CD^{∞} (Chapter 5, §6) on I, hence bounded there (Theorem 2 of Chapter 4, §8).

As in Example (C) in Chapter 3, §10, E' is a normed linear space, with norm

$$\|u\|=\sup_{x \in I}|u(x)|. \tag{6.2.7}$$

Here each function $u \in E'$ is treated as a single "point" in E'. The distance between two such points, u and v, equals $\|u-v\|$, by definition.

Now define a map D on E' by setting $D(u)=u'$ (derivative of u on I). As every $u \in E'$ is of class CD^{∞}, so is u'.

Thus $D(u)=u' \in E'$, and so $D: E' \rightarrow E'$ is a linear operator. (Its linearity follows from Theorem 4 in Chapter 5, §1.)

(c) Let again $I=[0,1]$. Let E' be the set of all functions $u: I \rightarrow E$ that are bounded and have antiderivatives (Chapter 5, §5) on I. With norm $\|u\|$ as in Example (b), E' is a normed linear space.

Now define $\phi: E' \rightarrow E$ by

$$\phi(u)=\int_{0}^{1} u, \tag{6.2.8}$$

with $\int u$ as in Chapter 5, §5. (Recall that $\int_{0}^{1} u$ is an element of E if $u: I \rightarrow E$.) By Corollary 1 in Chapter 5, §5, ϕ is a linear map of E' into E. (Why?)

(d) The zero map $f=0$ on E' is always linear. (Why?)

Theorem 6.2.1

A linear map $f: E' \rightarrow E$ is continuous (even uniformly so) on all of E' iff it is continuous at $\overrightarrow{0}$; equivalently, iff there is a real $c>0$ such that

$$(\forall \vec{x} \in E') \quad |f(\vec{x})| \leq c|\vec{x}|. \tag{6.2.9}$$

(We call this property linear boundedness.)

Proof

Assume that f is continuous at $\overrightarrow{0}$. Then, given $\varepsilon>0$, there is $\delta>0$ such that

$$|f(\vec{x})-f(\overrightarrow{0})|=|f(\vec{x})| \leq \varepsilon \tag{6.2.10}$$

whenever $|\vec{x}-\overrightarrow{0}|=|\vec{x}|<\delta$.

Now, for any $\vec{x} \neq \overrightarrow{0}$, we surely have

$$\left|\frac{\delta \vec{x}}{|\vec{x}|}\right|=\frac{\delta}{2}<\delta. \tag{6.2.11}$$

Hence

$$(\forall \vec{x} \neq \overrightarrow{0}) \quad\left|f\left(\frac{\delta \vec{x}}{2|\vec{x}|}\right)\right| \leq \varepsilon, \tag{6.2.12}$$

or, by linearity,

$$\frac{\delta}{2|\vec{x}|}|f(\vec{x})| \leq \varepsilon, \tag{6.2.13}$$

i.e.,

$$|f(\vec{x})| \leq \frac{2\varepsilon}{\delta}|\vec{x}|. \tag{6.2.14}$$

By Note 2, this also holds if $\vec{x} = \overrightarrow{0}$.

Thus, taking $c = 2\varepsilon/\delta$, we obtain

$$(\forall \vec{x} \in E') \quad f(\vec{x}) \leq c|\vec{x}| \quad \text{(linear boundedness)}. \tag{6.2.15}$$

Now assume (3). Then

$$(\forall \vec{x}, \vec{y} \in E') \quad |f(\vec{x} - \vec{y})| \leq c|\vec{x} - \vec{y}|; \tag{6.2.16}$$

or, by linearity,

$$(\forall \vec{x}, \vec{y} \in E') \quad |f(\vec{x}) - f(\vec{y})| \leq c|\vec{x} - \vec{y}|. \tag{6.2.17}$$

Hence f is uniformly continuous (given $\varepsilon > 0$, take $\delta = \varepsilon/c$). This, in turn, implies continuity at $\overrightarrow{0}$; so all conditions are equivalent, as claimed. $\square$

A linear map need not be continuous. But, for E^n and C^n, we have the following result.

Theorem 6.2.2

(i) Any linear map on E^n or C^n is uniformly continuous.

(ii) Every linear functional on E^n (C^n) has the form

$$f(\vec{x}) = \vec{x} \cdot \vec{v} \quad \text{(dot product)} \tag{6.2.18}$$

for some unique vector $\vec{v} \in E^n$ (C^n), dependent on f only.

Proof

Suppose $f : E^n \rightarrow E$ is linear; so f preserves linear combinations.

But every $\vec{x} \in E^n$ is such a combination,

$$\vec{x} = \sum_{k=1}^{n} x_k \vec{e}_k \quad \text{(Theorem 2 in Chapter 3, §§1-3)}. \tag{6.2.19}$$

Thus, by Note 1,

$$f(\vec{x}) = f\left(\sum_{k=1}^{n} x_k \vec{e}_k\right) = \sum_{k=1}^{n} x_k f\left(\vec{e}_k\right). \tag{6.2.20}$$

Here the function values $f\left(\vec{e}_k\right)$ are fixed vectors in the range space E, say,

$$f\left(\vec{e}_k\right) = v_k \in E, \tag{6.2.21}$$

so that

$$f(\vec{x}) = \sum_{k=1}^{n} x_k f\left(\vec{e}_k\right) = \sum_{k=1}^{n} x_k v_k, \quad v_k \in E. \tag{6.2.22}$$

Thus f is a polynomial in n real variables x_k, hence continuous (even uniformly so, by Theorem 1).

In particular, if $E = E^1$ (i.e., f is a linear functional) then all v_k in (5) are real numbers; so they form a vector

$$\vec{v} = (v_1, \ldots, v_k) \text{ in } E^n, \tag{6.2.23}$$

and (5) can be written as

$$f(\vec{x}) = \vec{x} \cdot \vec{v}. \tag{6.2.24}$$

The vector $\vec{v}$ is unique. For suppose there are two vectors, $\vec{u}$ and $\vec{v}$, such that

$$(\forall \vec{x} \in E^n) \quad f(\vec{x}) = \vec{x} \cdot \vec{v} = \vec{x} \cdot \vec{u}. \tag{6.2.25}$$

Then

$$(\forall \vec{x} \in E^n) \quad \vec{x} \cdot (\vec{v} - \vec{u}) = 0. \tag{6.2.26}$$

By Problem 10 of Chapter 3, §§1-3, this yields $\vec{v} - \vec{u} = \overrightarrow{0}$, or $\vec{v} = \vec{u}$. This completes the proof for $E = E^n$.

It is analogous for C^n; only in (ii) the v_k are complex and one has to replace them by their conjugates $\bar{v}_k$ when forming the vector $\vec{v}$ to obtain $f(\vec{x}) = \vec{x} \cdot \vec{v}$. Thus all is proved. □

Note 3. Formula (5) shows that a linear map $f : E^n\left(C^n\right) \rightarrow E$ is uniquely determined by the n function values $v_k = f\left(\vec{e}_k\right)$.

If further $E = E^m\left(C^m\right)$, the vectors v_k are m -tuples of scalars,

$$v_k = \left(v_{1k}, \ldots, v_{mk}\right). \tag{6.2.27}$$

We often write such vectors vertically, as the n "columns" in an array of m "rows" and n "columns":

$$\left(\begin{array}{cccc} v_{11} & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} & \ldots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{m1} & v_{m2} & \ldots & v_{mn} \end{array}\right). \tag{6.2.28}$$

Formally, (6) is a double sequence of mn terms, called an $m \times n$ matrix. We denote it by $[f] = \left(v_{ik}\right)$, where for $k = 1, 2, \ldots, n$,

$$f\left(\vec{e}_k\right) = v_k = \left(v_{1k}, \ldots, v_{mk}\right). \tag{6.2.29}$$

Thus linear maps $f : E^n \rightarrow E^m$ (or $f : C^n \rightarrow C^m$) correspond one-to-one to their matrices $[f]$.

The easy proof of Corollaries 1 to 3 below is left to the reader.

Corollary 6.2.1

If $f, g : E' \rightarrow E$ are linear, so is

$$h = af + bg \tag{6.2.30}$$

for any scalars a, b.

If further $E' = E^n\left(C^n\right)$ and $E = E^m\left(C^m\right)$, with $[f] = \left(v_{ik}\right)$ and $[g] = \left(w_{ik}\right)$, then

$$[h] = \left(av_{ik} + bw_{ik}\right). \tag{6.2.31}$$

Corollary 6.2.2

A map $f : E^n\left(C^n\right) \rightarrow E$ is linear iff

$$f(\vec{x}) = \sum_{k=1}^{n} v_k x_k, \tag{6.2.32}$$

where $v_k = f\left(\vec{e}_k\right)$.

Hint: For the "if," use Corollary 1. For the "only if," use formula (5) above.

Corollary 6.2.3

If $f : E' \rightarrow E''$ and $g : E'' \rightarrow E$ are linear, so is the composite $h = g \circ f$.

Our next theorem deals with the matrix of the composite linear map $g \circ f$

Theorem 6.2.3

Let $f : E' \rightarrow E''$ and $g : E'' \rightarrow E$ be linear, with

$$E' = E^n\left(C^n\right), E'' = E^m\left(C^m\right), \text{ and } E = E^r\left(C^r\right). \tag{6.2.33}$$

If $[f] = \left(v_{ik}\right)$ and $[g] = \left(w_{ji}\right)$, then

$$[h] = [g \circ f] = \left(z_{jk}\right), \tag{6.2.34}$$

where

$$z_{jk} = \sum_{i=1}^{m} w_{ji} v_{ik}, \quad j = 1, 2, \ldots, r, k = 1, 2, \ldots, n. \tag{6.2.35}$$

Proof

Denote the basic unit vectors in E' by

$$e_1', \ldots, e_n', \tag{6.2.36}$$

those in E'' by

$$e_1'', \ldots, e_m'', \tag{6.2.37}$$

and those in E by

$$e_1, \ldots, e_r. \tag{6.2.38}$$

Then for $k = 1, 2, \ldots, n$,

$$f\left(e_k'\right) = v_k = \sum_{i=1}^{m} v_{ik} e_i'' \text{ and } h\left(e_k'\right) = \sum_{j=1}^{r} z_{jk} e_j, \tag{6.2.39}$$

and for $i = 1, \ldots m$,

$$g\left(e_i''\right) = \sum_{j=1}^{r} w_{ji} e_j. \tag{6.2.40}$$

Also,

$$h\left(e_k'\right) = g\left(f\left(e_k'\right)\right) = g\left(\sum_{i=1}^{m} v_{ik} e_i''\right) = \sum_{i=1}^{m} v_{ik} g\left(e_i''\right) = \sum_{i=1}^{m} v_{ik}\left(\sum_{j=1}^{r} w_{ji} e_j\right). \tag{6.2.41}$$

Thus

$$h\left(e_k'\right) = \sum_{j=1}^{r} z_{jk} e_j = \sum_{j=1}^{r}\left(\sum_{i=1}^{m} w_{ji} v_{ik}\right) e_j. \tag{6.2.42}$$

But the representation in terms of the e_j is unique (Theorem 2 in Chapter 3, §§1-3), so, equating coefficients, we get (7). □

Note 4. Observe that z_{jk} is obtained, so to say, by "dot-multiplying" the jth row of $[g]$ (an $r \times m$ matrix) by the kth column of $[f]$ (an $m \times n$ matrix).

It is natural to set

$$[g][f] = [g \circ f], \tag{6.2.43}$$

or

$$\left(w_{ji}\right)\left(v_{ik}\right) = \left(z_{jk}\right), \tag{6.2.44}$$

with z_{jk} as in (7).

Caution. Matrix multiplication, so defined, is not commutative.

Definition 2

The set of all continuous linear maps $f: E' \rightarrow E$ (for fixed E' and E) is denoted $L(E', E)$.

If $E = E'$, we write $L(E)$ instead.

For each f in $L\left(E', E\right)$, we define its norm by

$$\|f\|=\sup _{|\vec{x}| \leq 1}|f(\vec{x})|. \tag{6.2.45}$$

Note that $\|f\|<+\infty$, by Theorem 1.

Theorem 6.2.4

$L(E', E)$ is a normed linear space under the norm defined above and under the usual operations on functions, as in Corollary 1.

Proof

Corollary 1 easily implies that $L(E', E)$ is a vector space. We now show that $\|\cdot\|$ is a genuine norm.

The triangle law,

$$\|f+g\| \leq\|f\|+\|g\|, \tag{6.2.46}$$

follows exactly as in Example (C) of Chapter 3, §10. (Verify!)

Also, by Problem 5 in Chapter 2, §§8-9, $\sup |a f(\vec{x})|=|a| \sup |f(\vec{x})|$. Hence $\|a f\|=|a|\|f\|$for any scalar a.

As noted above, $0 \leq\|f\|<+\infty$.

It remains to show that $\|f\|=0$ iff f is the zero map. If

$$\|f\|=\sup _{|\vec{x}| \leq 1}|f(\vec{x})|=0, \tag{6.2.47}$$

then $|f(\vec{x})|=0$ when $|\vec{x}| \leq 1$. Hence, if $\vec{x} \neq \overrightarrow{0}$,

$$f\left(\frac{\vec{x}}{|\vec{x}|}\right)=\frac{1}{|\vec{x}|} f(\vec{x})=0. \tag{6.2.48}$$

As $f(\overrightarrow{0})=0$, we have $f(\vec{x})=0$ for all $\vec{x} \in E'$.

Thus $\|f\|=0$ implies $f=0$, and the converse is clear. Thus all is proved. $\square$

Note 5. A similar proof, via $f\left(\frac{\vec{x}}{|\vec{x}|}\right)$ and properties of lub, shows that

$$\|f\|=\sup _{\vec{x} \neq 0}\left|\frac{f(\vec{x})}{|\vec{x}|}\right| \tag{6.2.49}$$

and

$$\left(\forall \vec{x} \in E'\right) \quad|f(\vec{x})| \leq\|f\||\vec{x}|. \tag{6.2.50}$$

It also follows that $\|f\|$ is the least real c such that

$$\left(\forall \vec{x} \in E'\right) \quad|f(\vec{x})| \leq c|\vec{x}|. \tag{6.2.51}$$

Verify. (See Problem 3'.)

As in any normed space, we define distances in $L(E', E)$ by

$$\rho(f, g) = \|f - g\|, \tag{6.2.52}$$

making it a metric space; so we may speak of convergence, limits, etc. in it.

Corollary 6.2.4

If $f \in L(E', E'')$ and $g \in L(E'', E)$, then

$$\|g \circ f\| \leq \|g\|\|f\|. \tag{6.2.53}$$

Proof

By Note 5,

$$(\forall \vec{x} \in E') \quad |g(f(\vec{x}))| \leq \|g\||f(\vec{x})| \leq \|g\|\|f\||\vec{x}|. \tag{6.2.54}$$

Hence

$$(\forall \vec{x} \neq \overrightarrow{0}) \quad \left|\frac{(g \circ f)(\vec{x})}{|\vec{x}|}\right| \leq \|g\|\|f\|, \tag{6.2.55}$$

and so

$$\|g\|\|f\| \geq \sup_{\vec{x} \neq \overrightarrow{0}} \frac{|(g \circ f)(\vec{x})|}{|\vec{x}|} = \|g \circ f\|. \quad \square \tag{6.2.56}$$

6.2.E: Problems on Linear Maps and Matrices

Exercise $6.2.E.1$

Verify Note 1 and the equivalence of the two statements in Definition 1.

Exercise $6.2.E.2$

In Examples (b) and (c) show that

$$f_n \rightarrow f(\text{uniformly}) \text{ on } I \text{ iff } \|f_n - f\| \rightarrow 0, \tag{6.2.E.1}$$

i.e., $f_n \rightarrow f$ in E'.
[Hint: Use Theorem 1 in Chapter 4, §2.]
Hence deduce the following.
(i) If E is complete, then the map ϕ in Example (c) is continuous.
[Hint: Use Theorem 2 of Chapter 5, §9, and Theorem 1 in Chapter 4, §12.]
(ii) The map D of Example (b}) is not continuous.
[Hint: Use Problem 3 in Chapter 5, §9.]

Exercise $6.2.E.3$

Prove Corollaries 1 to 3.

Exercise $6.2.E.3'$

Show that

$$\|f\| = \sup_{|\vec{x}| \leq 1} |f(\vec{x})| = \sup_{|\vec{x}|=1} |f(\vec{x})| = \sup_{\vec{x} \neq \overrightarrow{0}} \frac{|f(\vec{x})|}{|\vec{x}|}. \tag{6.2.E.2}$$

[Hint: From linearity of f deduce that $|f(\vec{x})| \geq |f(c x)|$ if $|c|<1$. Hence one may disregard vectors of length <1 when computing sup $|f(\vec{x})|$. Why?]

Exercise $6.2.E.4$

Find the matrices $[f],[g],[h],[k]$, and the defining formulas for the linear maps $f: E^2 \rightarrow E^1, g: E^3 \rightarrow E^4, h: E^4 \rightarrow E^2, k: E^1 \rightarrow E^3$ if
(i) $f\left(\vec{e}_1\right)=3, f\left(\vec{e}_2\right)=-2$;
(ii) $g\left(\vec{e}_1\right)=(1,0,-2,4), g\left(\vec{e}_2\right)=(0,2,-1,1), g\left(\vec{e}_3\right)=(0,1,0,-1)$;
(iii) $h\left(\vec{e}_1\right)=(2,2), h\left(\vec{e}_2\right)=(0,-2), h\left(\vec{e}_3\right)=(1,0), h\left(\vec{e}_4\right)=(-1,1)$;
(iv) $k(1)=(0,1,-1)$.

Exercise $6.2.E.5$

In Problem 4, use Note 4 to find the product matrices $[k][f],[g][k],[f][h]$,and $[h][g]$. Hence obtain the defining formulas for $k \circ f, g \circ k, f \circ h,$ and $h \circ g$.

Exercise $6.2.E.6$

For $m \times n$ matrices (with m and n fixed) define addition and multiplication by scalars as follows:

$$a[f]+b[g]=[af+bg] \text{ if } f, g \in L\left(E^{n}, E^{m}\right)\left(\text{ or } L\left(C^{n}, C^{m}\right)\right). \tag{6.2.E.3}$$

Show that these matrices form a vector space over E^{1} (or C).

Exercise $6.2.E.7$

With matrix addition as in Problem 6, and multiplication as in Note 4, show that all $n \times n$ matrices form a noncommutative ring with unity, i.e., satisfy the field axioms (Chapter 2, §§1-4) except the commutativity of multiplication and existence of multiplicative inverses (give counterex-amplest!).
Which is the "unity" matrix?

Exercise $6.2.E.8$

Let $f: E^{\prime} \rightarrow E$ be linear. Prove the following statements.

(i) The derivative $D_{\vec{u}} f(\vec{p})$ exists and equals $f(\vec{u})$ for every $\vec{p}, \vec{u} \in E^{\prime}(\vec{u} \neq \overrightarrow{0})$;
(ii) f is relatively continuous on any line in $E^{\prime}$ (use Theorem 1 in §1);
(iii) f carries any such line into a line in E.

Exercise $6.2.E.9$

Let $g: E^{\prime \prime} \rightarrow E$ be linear. Prove that if some $f: E^{\prime} \rightarrow E^{\prime \prime}$ has a $\vec{u}$-directed derivative at $\vec{p} \in E^{\prime}$, so has $h=g \circ f$, and $D_{\vec{u}} h(\vec{p})=g\left(D_{\vec{u}} f(\vec{p})\right)$.
[Hint: Use Problem 8.]

Exercise $6.2.E.10$

A set A in a vector space $V(A \subseteq V)$ is said to be linear (or a linear subspace of V) iff $a \vec{x}+b \vec{y} \in A$ for any $\vec{x}, \vec{y} \in A$ and any scalars a, b. Prove the following.
(i) Any such A is itself a vector space.
(ii) If $f: E^{\prime} \rightarrow E$ is a linear map and A is linear in $E^{\prime}$ (respectively, in E), so is $f[A]$ in E (respectively, so is $f^{-1}[A]$ in $E^{\prime}$).

Exercise $6.2.E.11$

A set A in a vector space V is called the span of a set $B \subseteq A(A=\operatorname{sp}(B))$ iff A consists of all linear combinations of vectors from B. We then also say that B spans A.
Prove the following:
(i) $A=\operatorname{sp}(B)$ is the smallest linear subspace of V that contains B.
(ii) If $f: V \rightarrow E$ is linear and $A=\operatorname{sp}(B)$, then $f[A]=\operatorname{sp}(f[B])$ in E.

Exercise $6.2.E.12$

A set $B=\left\{\vec{x}_{1}, \vec{x}_{2}, \ldots, \vec{x}_{n}\right\}$ in a vector space V is called a basis iff each $\vec{v} \in V$ has a unique representation as

$$\vec{v}=\sum_{i=1}^{n} a_{i} \vec{x}_{i} \tag{6.2.E.4}$$

for some scalars a_{i}. If so, the number n of the vectors in B is called the dimension of V, and V is said to be n- dimensional.

Examples of such spaces are E^n and C^n (the $\vec{e}_k$ form a basis!).
(i) Show that B is a basis iff it spans V (see Problem 11) and its elements $\vec{x}_i$ are linearly independent, i.e.,

$$\sum_{i=1}^{n} a_i \vec{x}_i = \overrightarrow{0} \text{ iff all } a_i \text{ vanish.} \tag{6.2.E.5}$$

(ii) If E' is finite-dimensional, all linear maps on E' are uniformly continuous. (See also Problems 3 and 4 of §6.)

Exercise $6.2.E.13$

Prove that if $f: E^1 \rightarrow E$ is continuous and $\left(\forall x, y \in E^1\right)$

$$f(x+y) = f(x) + f(y), \tag{6.2.E.6}$$

then f is linear; so, by Corollary 2, $f(x) = vx$ where $v = f(1)$.
[Hint: Show that $f(ax) = af(x)$; first for $a = 1, 2, \ldots$ (note: $nx = x + x + \cdots + x, n$ terms); then for rational $a = m/n$; then for $a = 0$ and $a = -1$. Any $a \in E^1$ is a limit of rationals; so use continuity and Theorem 1 in Chapter 4, §2.]

6.3: Differentiable Functions

As we know, a function $f: E^{1} \rightarrow E\left(\text{on } E^{1}\right)$ is differentiable at $p \in E^{1}$ iff, with $\Delta f=f(x)-f(p)$ and $\Delta x=x-p$,

$$f^{\prime}(p)=\lim_{x \rightarrow p} \frac{\Delta f}{\Delta x} \text{ exists } \quad (\text{ finite }). \tag{6.3.1}$$

Setting $\Delta x=x-p=t, \Delta f=f(p+t)-f(p), \quad$ and $f^{\prime}(p)=v$, we may write this equation as

$$\lim_{t \rightarrow 0}\left|\frac{\Delta f}{t}-v\right|=0, \tag{6.3.2}$$

or

$$\lim_{t \rightarrow 0} \frac{1}{|t|}|f(p+t)-f(p)-v t|=0 \tag{6.3.3}$$

Now define a map $\phi: E^{1} \rightarrow E$ by $\phi(t)=t v, v=f^{\prime}(p) \in E$.

Then ϕ is linear and continuous, i.e., $\phi \in L\left(E^{1}, E\right)$; so by Corollary 2 in §2, we may express (1) as follows: there is a map $\phi \in L\left(E^{1}, E\right)$ such that

$$\lim_{t \rightarrow 0} \frac{1}{|t|}|\Delta f-\phi(t)|=0. \tag{6.3.4}$$

We adopt this as a definition in the general case, $f: E^{\prime} \rightarrow E,$ as well.

Definition: Differentiable at a Point

A function $f: E^{\prime} \rightarrow E$ where $E^{\prime}$ and E are normed spaces over the same scalar field) is said to be **differentiable** at a point $\vec{p} \in E^{\prime}$ iff there is a map

$$\phi \in L\left(E^{\prime}, E\right) \tag{6.3.5}$$

such that

$$\lim_{\vec{t} \rightarrow \overrightarrow{0}} \frac{1}{|\vec{t}|}|\Delta f-\phi(\vec{t})|=0; \tag{6.3.6}$$

that is,

$$\lim_{\vec{t} \rightarrow \overrightarrow{0}} \frac{1}{|\vec{t}|}[f(\vec{p}+\vec{t})-f(\vec{p})-\phi(\vec{t})]=0. \tag{6.3.7}$$

As we show below, ϕ is unique (for a fixed $\vec{p}$), if it exists.

We call ϕ the differential of f at $\vec{p}$, briefly denoted df. As it depends on $\vec{p}$, we also write $df(\vec{p}; \vec{t})$ for $df(\vec{t})$ and $df(\vec{p}; \cdot)$ for df.

Some authors write $f^{\prime}(\vec{p})$ for $df(\vec{p}; \cdot)$ and call it the derivative at $\vec{p}$, but we shall not do this (see Preface). Following M. Spivak, however, we shall use " $\left[f^{\prime}(\vec{p})\right]$ " for its matrix, as follows.

Definition: Jacobian matrix

If $E^{\prime}=E^{n}\left(C^{n}\right)$ and $E=E^{m}\left(C^{m}\right)$, and $f: E^{\prime} \rightarrow E$ is differentiable at $\vec{p}$, we set

$$\left[f^{\prime}(\vec{p})\right]=[df(\vec{p}; \cdot)] \tag{6.3.8}$$

and call it the **Jacobian matrix** of f at $\vec{p}$.

Note 1. In Chapter 5, §6, we did not define df as a mapping. However, if $E^{\prime}=E^{1}$, the function value

$$df(p; t)=v t=f^{\prime}(p) \Delta x \tag{6.3.9}$$

is as in Chapter 5, §6.

Also, $[f'(p)]$ is a 1×1 matrix with single term $f'(p)$. (Why?) This motivated Definition 2.

Theorem 6.3.1

(uniqueness of df). If $f:E'\to E$ is differentiable at $\vec{p}$, then the map ϕ described in Definition 1 is unique (dependent on f and $\vec{p}$ only).

Proof

Suppose there is another linear map $g:E'\to E$ such that

$$\lim_{\vec{t}\to\vec{0}}\frac{1}{|\vec{t}|}[f(\vec{p}+\vec{t})-f(\vec{p})-g(\vec{t})]=\lim_{\vec{t}\to\vec{0}}\frac{1}{|\vec{t}|}[\Delta f-g(\vec{t})]=0. \tag{6.3.10}$$

Let $h=\phi-g$. By Corollary 1 in §2, h is linear.

Also, by the triangle law,

$$|h(\vec{t})|=|\phi(\vec{t})-g(\vec{t})|\leq|\Delta f-\phi(\vec{t})|+|\Delta f-g(\vec{t})|. \tag{6.3.11}$$

Hence, dividing by $|\vec{t}|$,

$$\left|h\left(\frac{\vec{t}}{|\vec{t}|}\right)\right|=\frac{1}{|\vec{t}|}|h(\vec{t})|\leq\frac{1}{|\vec{t}|}|\Delta f-\phi(\vec{t})|+\frac{1}{|\vec{t}|}|\Delta f-g(\vec{t})|. \tag{6.3.12}$$

By (3) and (2), the right side expressions tend to 0 as $\vec{t}\to\overrightarrow{0}$. Thus

$$\lim_{\vec{t}\to\vec{0}}h\left(\frac{\vec{t}}{|\vec{t}|}\right)=0. \tag{6.3.13}$$

This remains valid also if $\vec{t}\to\overrightarrow{0}$ over any line through $\overrightarrow{0}$, so that $\vec{t}/|\vec{t}|$ remains constant, say $\vec{t}/|\vec{t}|=\vec{u}$, where $\vec{u}$ is an arbitrary (but fixed) unit vector.

Then

$$h\left(\frac{\vec{t}}{|\vec{t}|}\right)=h(\vec{u}) \tag{6.3.14}$$

is constant; so it can tend to 0 only if it equals 0, so $h(\vec{u})=0$ for any unit vector $\vec{u}$.

Since any $\vec{x}\in E'$ can be written as $\vec{x}=|\vec{x}|\vec{u}$, linearity yields

$$h(\vec{x})=|\vec{x}|h(\vec{u})=0. \tag{6.3.15}$$

Thus $h=\phi-g=0$ on E', and so $\phi=g$ after all, proving the uniqueness of ϕ. □

Theorem 6.3.2

If f is differentiable at $\vec{p}$, then

(i) f is continuous at $\vec{p}$;

(ii) for any $\vec{u}\neq\overrightarrow{0}$, has the $\vec{u}$-directed derivative

$$D_{\vec{u}}f(\vec{p})=df(\vec{p};\vec{u}). \tag{6.3.16}$$

Proof

By assumption, formula (2) holds for $\phi=df(\vec{p};\cdot)$.

Thus, given $\varepsilon>0,$ there is $\delta>0$ such that, setting $\Delta f=f(\vec{p}+\vec{t})-f(\vec{p})$ we have

$$\frac{1}{|\vec{t}|}|\Delta f-\phi(\vec{t})|<\varepsilon \text{ whenever } 0<|\vec{t}|<\delta; \tag{6.3.17}$$

or, by the triangle law,

$$|\Delta f| \leq|\Delta f-\phi(\vec{t})|+|\phi(\vec{t})| \leq \varepsilon|\vec{t}|+|\phi(\vec{t})|, \quad 0<|\vec{t}|<\delta. \tag{6.3.18}$$

Now, by Definition 1, ϕ is linear and continuous; so

$$\lim_{\vec{t} \rightarrow \overrightarrow{0}}|\phi(\vec{t})|=|\phi(\overrightarrow{0})|=0. \tag{6.3.19}$$

Thus, making $\vec{t} \rightarrow \overrightarrow{0}$ in (5), with ε fixed, we get

$$\lim_{\vec{t} \rightarrow \overrightarrow{0}}|\Delta f|=0. \tag{6.3.20}$$

As $\vec{t}$ is just another notation for $\Delta \vec{x}=\vec{x}-\vec{p},$ this proves assertion (i).

Next, fix any $\vec{u} \neq \overrightarrow{0}$ in E', and substitute $t \vec{u}$ for $\vec{t}$ in (4).

In other words, t is a real variable, $0<t<\delta /|\vec{u}|,$ so that $\vec{t}=t \vec{u}$ satisfies $0<|\vec{t}|<\delta$.

Multiplying by $|\vec{u}|$, we use the linearity of ϕ to get

$$\varepsilon|\vec{u}|>\left|\frac{\Delta f}{t}-\frac{\phi(t \vec{u})}{t}\right|=\left|\frac{\Delta f}{t}-\phi(\vec{u})\right|=\left|\frac{f(\vec{p}+t \vec{u})-f(\vec{p})}{t}-\phi(\vec{u})\right|. \tag{6.3.21}$$

As ε is arbitrary, we have

$$\phi(\vec{u})=\lim_{t \rightarrow 0} \frac{1}{t}[f(\vec{p}+t \vec{u})-f(\vec{p})]. \tag{6.3.22}$$

But this is simply $D_{\vec{u}} f(\vec{p})$, by Definition 1 in §1.

Thus $D_{\vec{u}} f(\vec{p})=\phi(\vec{u})=d f(\vec{p} ; \vec{u}),$ proving (ii). $\square$

Note 2. If $E'=E^{n}\left(C^{n}\right),$ Theorem 2(iii) shows that if f is differentiable at $\vec{p}$, it has the n partials

$$D_{k} f(\vec{p})=d f\left(\vec{p} ; \vec{e}_{k}\right), \quad k=1, \ldots, n. \tag{6.3.23}$$

But the converse fails: the existence of the $D_{k} f(\vec{p})$ does not even imply continuity, let alone differentiability (see §1). Moreover, we have the following result.

Corollary 6.3.1

If $E'=E^{n}\left(C^{n}\right)$ and if $f: E' \rightarrow E$ is differentiable at $\vec{p}$, then

$$d f(\vec{p} ; \vec{t})=\sum_{k=1}^{n} t_{k} D_{k} f(\vec{p})=\sum_{k=1}^{n} t_{k} \frac{\partial}{\partial x_{k}} f(\vec{p}), \tag{6.3.24}$$

where $\vec{t}=\left(t_{1}, \ldots, t_{n}\right)$.

Proof

By definition, $\phi=d f(\vec{p} ; \cdot)$ is a linear map for a fixed $\vec{p}$.

If $E'=E^{n}$ or C^{n}, we may use formula (3) of §2, replacing f and $\vec{x}$ by ϕ and $\vec{t}$, and get

$$\phi(\vec{t})=d f(\vec{p} ; \vec{t})=\sum_{k=1}^{n} t_{k} d f\left(\vec{p} ; \vec{e}_{k}\right)=\sum_{k=1}^{n} t_{k} D_{k} f(\vec{p}) \tag{6.3.25}$$

by Note 2. □

Note 3. In classical notation, one writes Δx_k or dx_k for t_k in (6). Thus, omitting $\vec{p}$ and $\vec{t}$, formula (6) is often written as

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n. \tag{6.3.26}$$

In particular, if $n=3$, we write x, y, z for x_1, x_2, x_3. This yields

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \tag{6.3.27}$$

(a familiar calculus formula).

Note 4. If the range space E in Corollary 1 is $E^1(C)$, then the $D_k f(\vec{p})$ form an n-tuple of scalars, i.e., a vector in $E^n(C^n)$.

In case $f : E^n \rightarrow E^1$, we denote it by

$$\nabla f(\vec{p}) = (D_1 f(\vec{p}), \ldots, D_n f(\vec{p})) = \sum_{k=1}^{n} \vec{e}_k D_k f(\vec{p}). \tag{6.3.28}$$

In case $f : C^n \rightarrow C$, we replace the $D_k f(\vec{p})$ by their conjugates $\overline{D_k f(\vec{p})}$ and set

$$\nabla f(\vec{p}) = \sum_{k=1}^{n} \vec{e}_k \overline{D_k f(\vec{p})}. \tag{6.3.29}$$

The vector $\nabla f(\vec{p})$ is called the gradient of f ("grad f") at $\vec{p}$.

From (6) we obtain

$$df(\vec{p}; \vec{t}) = \sum_{k=1}^{n} t_k D_k f(\vec{p}) = \vec{t} \cdot \nabla f(\vec{p}) \tag{6.3.30}$$

(dot product of $\vec{t}$ by $\nabla f(\vec{p})$), provided $f : E^n \rightarrow E^1$ (or $f : C^n \rightarrow C$) is differentiable at $\vec{p}$.

This leads us to the following result.

Corollary 6.3.2

A function $f : E^n \rightarrow E^1$ (or $f : C^n \rightarrow C$) is differentiable at $\vec{p}$ iff

$$\lim_{\vec{t} \rightarrow \overrightarrow{0}} \frac{1}{|\vec{t}|} |f(\vec{p} + \vec{t}) - f(\vec{p}) - \vec{t} \cdot \vec{v}| = 0 \tag{6.3.31}$$

for some $\vec{v} \in E^n(C^n)$.

In this case, necessarily $\vec{v} = \nabla f(\vec{p})$ and $\vec{t} \cdot \vec{v} = df(\vec{p}; \vec{t}), \vec{t} \in E^n(C^n)$.

Proof

If f is differentiable at $\vec{p}$, we may set $\phi = df(\vec{p}; \cdot)$ and $\vec{v} = \nabla f(\vec{p})$

Then by (7),

$$\phi(\vec{t}) = df(\vec{p}; \vec{t}) = \vec{t} \cdot \vec{v}; \tag{6.3.32}$$

so by Definition 1, (8) results.

Conversely, if some $\vec{v}$ satisfies (8), set $\phi(\vec{t}) = \vec{t} \cdot \vec{v}$. Then (8) implies (2), and ϕ is linear and continuous.

Thus by definition, f is differentiable at $\vec{p}$; so (7) holds.

Also, ϕ is a linear functional on $E^n(C^n)$. By Theorem 2(ii) in §2, the $\vec{v}$ in $\phi(\vec{t}) = \vec{t} \cdot \vec{v}$ is unique, as is ϕ.

Thus by (7), $\vec{v} = \nabla f(\vec{p})$ necessarily. □

Corollary 6.3.3 (law of the mean)

If $f: E^{n} \rightarrow E^{1}$ (real) is relatively continuous on a closed segment $L[\vec{p}, \vec{q}], \vec{p} \neq \vec{q}$, and differentiable on $L(\vec{p}, \vec{q})$, then

$$f(\vec{q})-f(\vec{p})=(\vec{q}-\vec{p}) \cdot \nabla f\left(\vec{x}_{0}\right) \tag{6.3.33}$$

for some $\vec{x}_{0} \in L(\vec{p}, \vec{q})$.

Proof

Let

$$r=|\vec{q}-\vec{p}|, \quad \vec{v}=\frac{1}{r}(\vec{q}-\vec{p}), \text{ and } r \vec{v}=(\vec{q}-\vec{p}). \tag{6.3.34}$$

By (7) and Theorem 2(ii),

$$D_{\vec{v}} f(\vec{x})=d f(\vec{x} ; \vec{v})=\vec{v} \cdot \nabla f(\vec{x}) \tag{6.3.35}$$

for $\vec{x} \in L(\vec{p}, \vec{q})$. Thus by formula (3') of Corollary 2 in §1,

$$f(\vec{q})-f(\vec{p})=r D_{\vec{v}} f\left(\vec{x}_{0}\right)=r \vec{v} \cdot \nabla f\left(\vec{x}_{0}\right)=(\vec{q}-\vec{p}) \cdot \nabla f\left(\vec{x}_{0}\right) \tag{6.3.36}$$

for some $\vec{x}_{0} \in L(\vec{p}, \vec{q})$. $\square$

As we know, the mere existence of partials does not imply differentiability. But the existence of continuous partials does. Indeed, we have the following theorem.

Theorem 6.3.3

Let $E^{\prime}=E^{n}\left(C^{n}\right)$.

If $f: E^{\prime} \rightarrow E$ has the partial derivatives $D_{k} f(k=1, \ldots, n)$ on all of an open set $A \subseteq E^{\prime}$, and if the $D_{k} f$ are continuous at some $\vec{p} \in A$, then f is differentiable at $\vec{p}$.

Proof

With $\vec{p}$ as above, let

$$\phi(\vec{t})=\sum_{k=1}^{n} t_{k} D_{k} f(\vec{p}) \text{ with } \vec{t}=\sum_{k=1}^{n} t_{k} \vec{e}_{k} \in E^{\prime}. \tag{6.3.37}$$

Then ϕ is continuous (a polynomial!) and linear (Corollary 2 in §2).

Thus by Definition 1, it remains to show that

$$\lim _{\vec{t} \rightarrow \overrightarrow{0}} \frac{1}{|\vec{t}|}|\Delta f-\phi(\vec{t})|=0; \tag{6.3.38}$$

that is;

$$\lim _{\vec{t} \in \overrightarrow{0}} \frac{1}{|\vec{t}|}\left|f(\vec{p}+\vec{t})-f(\vec{p})-\sum_{k=1}^{n} t_{k} D_{k} f(\vec{p})\right|=0. \tag{6.3.39}$$

To do this, fix $\varepsilon>0$. As A is open and the $D_{k} f$ are continuous at $\vec{p} \in A$ there is a $\delta>0$ such that $G_{\vec{p}}(\delta) \subseteq A$ and simultaneously (explain this!)

$$\left(\forall \vec{x} \in G_{\vec{p}}(\delta)\right) \quad\left|D_{k} f(\vec{x})-D_{k} f(\vec{p})\right|<\frac{\varepsilon}{n}, k=1, \ldots, n. \tag{6.3.40}$$

Hence for any set $I \subseteq G_{\vec{p}}(\delta)$

$$\sup_{\vec{x} \in I} |D_k f(\vec{x}) - D_k f(\vec{p})| \leq \frac{\varepsilon}{n}. \quad (Why?) \tag{6.3.41}$$

Now fix any $\vec{t} \in E', 0 < |\vec{t}| < \delta,$ and let $\vec{p}_0 = \vec{p}$,

$$\vec{p}_k = \vec{p} + \sum_{i=1}^{k} t_i e_i, \quad k = 1, \ldots, n. \tag{6.3.42}$$

Then

$$\vec{p}_n = \vec{p} + \sum_{i=1}^{n} t_i \vec{e}_i = \vec{p} + \vec{t}, \tag{6.3.43}$$

$|\vec{p}_k - \vec{p}_{k-1}| = |t_k|$, and all $\vec{p}_k$ lie in $G_{\vec{p}}(\delta)$, for

$$|\vec{p}_k - \vec{p}| = \left|\sum_{i=1}^{k} t_i e_i\right| = \sqrt{\sum_{i=1}^{k} |t_i|^2} \leq \sqrt{\sum_{i=1}^{n} |t_i|^2} = |\vec{t}| < \delta, \tag{6.3.44}$$

as required.

As $G_p(\delta)$ is convex (Chapter 4, §9), the segments $I_k = L[\vec{p}_{k-1}, \vec{p}_k]$ all lie in $G_{\vec{p}}(\delta) \subseteq A$; and by assumption, f has all partials there.

Hence by Theorem 1 in §1, f is relatively continuous on all I_k.

All this also applies to the functions g_k, defined by

$$(\forall \vec{x} \in E') \quad g_k(\vec{x}) = f(\vec{x}) - x_k D_k f(\vec{p}), \quad k = 1, \ldots, n. \tag{6.3.45}$$

(Why?) Here

$$D_k g_k(\vec{x}) = D_k f(\vec{x}) - D_k f(\vec{p}). \tag{6.3.46}$$

(Why?)

Thus by Corollary 2 in §1, and (11) above,

$$\begin{aligned} |g_k(\vec{p}_k) - g_k(\vec{p}_{k-1})| &\leq |\vec{p}_k - \vec{p}_{k-1}| \sup_{x \in I_k} |D_k f(\vec{x}) - D_k f(\vec{p})| \\ &\leq \frac{\varepsilon}{n} |t_k| \leq \frac{\varepsilon}{n} |\vec{t}|, \end{aligned}$$

since

$$|\vec{p}_k - \vec{p}_{k-1}| = |t_k \vec{e}_k| \leq |\vec{t}|, \tag{6.3.47}$$

by construction.

Combine with (12), recalling that the kth coordinates x_k, for $\vec{p}_k$ and $\vec{p}_{k-1}$ differ by t_k; so we obtain

$$\begin{aligned} |g_k(\vec{p}_k) - g_k(\vec{p}_{k-1})| &= |f(\vec{p}_k) - f(\vec{p}_{k-1}) - t_k D_k f(\vec{p})| \\ &\leq \frac{\varepsilon}{n} |\vec{t}|. \end{aligned}$$

Also,

$$\begin{aligned} \sum_{k=1}^{n} [f(\vec{p}_k) - f(\vec{p}_{k-1})] &= f(\vec{p}_n) - f(\vec{p}_0) \\ &= f(\vec{p} + \vec{t}) - f(\vec{p}) = \Delta f (\text{see above}). \end{aligned}$$

Thus,

$$\left|\Delta f-\sum_{k=1}^{n} t_{k} D_{k} f(\vec{p})\right| =\left|\sum_{k=1}^{n}\left[f\left(\vec{p}_{k}\right)-f\left(\vec{p}_{k-1}\right)-t_{k} D_{k} f(\vec{p})\right]\right| \leq n \cdot \frac{\varepsilon}{n}|\vec{t}|=\varepsilon|\vec{t}|.$$

As ε is arbitrary, (10) follows, and all is proved. □

Theorem 6.3.4

If $f: E^{n} \rightarrow E^{m}$ (or $f: C^{n} \rightarrow C^{m}$) is differentiable at $\vec{p}$, with $f=\left(f_{1}, \ldots, f_{m}\right)$, then $\left[f^{\prime}(\vec{p})\right]$ is an $m \times n$ matrix,

$$\left[f^{\prime}(\vec{p})\right]=\left[D_{k} f_{i}(\vec{p})\right], \quad i=1, \ldots, m, k=1, \ldots, n. \tag{6.3.48}$$

Proof

By definition, $\left[f^{\prime}(\vec{p})\right]$ is the matrix of the linear map $\phi=d f(\vec{p} ; \cdot)$, $\phi=\left(\phi_{1}, \ldots, \phi_{m}\right)$. Here

$$\phi(\vec{t})=\sum_{k=1}^{n} t_{k} D_{k} f(\vec{p}) \tag{6.3.49}$$

by Corollary 1.

As $f=\left(f_{1}, \ldots, f_{m}\right)$, we can compute $D_{k} f(\vec{p})$ componentwise by Theorem 5 of Chapter 5, §1, and Note 2 in §1 to get

$$\begin{aligned} D_{k} f(\vec{p}) &=\left(D_{k} f_{1}(\vec{p}), \ldots, D_{k} f_{m}(\vec{p})\right) \\ &=\sum_{i=1}^{m} e_{i}^{\prime} D_{k} f_{i}(\vec{p}), \quad k=1,2, \ldots, n, \end{aligned}$$

where the $e_{i}^{\prime}$ are the basic vectors in $E^{m}\left(C^{m}\right)$. (Recall that the $\vec{e}_{k}$ are the basic vectors in $E^{n}\left(C^{n}\right)$.)

Thus

$$\phi(\vec{t})=\sum_{i=1}^{m} e_{i}^{\prime} \phi_{i}(\vec{t}). \tag{6.3.50}$$

Also,

$$\phi(\vec{t})=\sum_{k=1}^{n} t_{k} \sum_{i=1}^{m} e_{i}^{\prime} D_{k} f_{i}(\vec{p})=\sum_{i=1}^{m} e_{i}^{\prime} \sum_{k=1}^{n} t_{k} D_{k} f_{i}(\vec{p}). \tag{6.3.51}$$

The uniqueness of the decomposition (Theorem 2 in Chapter 3, §§1-3) now yields

$$\phi_{i}(\vec{t})=\sum_{k=1}^{n} t_{k} D_{k} f_{i}(\vec{p}), \quad i=1, \ldots, m, \quad \vec{t} \in E^{n}\left(C^{n}\right). \tag{6.3.52}$$

If here $\vec{t}=\vec{e}_{k}$, then $t_{k}=1$, while $t_{j}=0$ for $j \neq k$. Thus we obtain

$$\phi_{i}\left(\vec{e}_{k}\right)=D_{k} f_{i}(\vec{p}), \quad i=1, \ldots, m, k=1, \ldots, n. \tag{6.3.53}$$

Hence,

$$\phi\left(\vec{e}_{k}\right)=\left(v_{1 k}, v_{2 k}, \ldots, v_{m k}\right), \tag{6.3.54}$$

where

$$v_{i k}=\phi_{i}\left(\vec{e}_{k}\right)=D_{k} f_{i}(\vec{p}). \tag{6.3.55}$$

But by Note 3 of §2, $v_{1 k}, \ldots, v_{m k}$ (written vertically) is the kth column of the $m \times n$ matrix $[\phi]=\left[f^{\prime}(\vec{p})\right]$. Thus formula (14) results indeed. □

In conclusion, let us stress again that while $D_{\vec{u}} f(\vec{p})$ is a constant, for a fixed $\vec{p}$, $d f(\vec{p} ; \cdot)$ is a mapping

$$\phi \in L\left(E^{\prime}, E\right), \tag{6.3.56}$$

especially "tailored" for $\vec{p}$.

The reader should carefully study at least the "arrowed" problems below.

6.3.E: Problems on Differentiable Functions

Exercise $6.3.E.1$

Complete the missing details in the proofs of this section.

Exercise $6.3.E.2$

Verify Note 1. Describe $[f'(\vec{p})]$ for $f : E^1 \rightarrow E^m,$ too. Give examples.

Exercise $6.3.E.3$

$\Rightarrow A$ map $f : E' \rightarrow E$ is said to satisfy a Lipschitz condition (L) of order $\alpha > 0$ at $\vec{p}$ iff

$$(\exists \delta>0)\left(\exists K \in E^{1}\right)\left(\forall \vec{x} \in G_{\neg \vec{p}}(\delta)\right) \quad|f(\vec{x})-f(\vec{p})| \leq K|\vec{x}-\vec{p}|^{\alpha}. \tag{6.3.E.1}$$

Prove the following.
(i) This implies continuity at $\vec{p}$ (but not conversely; see Problem 7 in Chapter 5, §1).
(ii) L of order >1 implies differentiability at $\vec{p}$, with $df(\vec{p};\cdot) = 0$ on E'.
(iii) Differentiability at $\vec{p}$ implies L of order 1 (apply Theorem 1 in §2 to $\phi = df$).
(iv) If f and g are differentiable at $\vec{p}$, then

$$\lim_{\vec{x} \rightarrow \vec{p}} \frac{1}{|\Delta \vec{x}|}|\Delta f||\Delta g|=0. \tag{6.3.E.2}$$

Exercise $6.3.E.4$

For the functions of Problem 5 in §1, find those $\vec{p}$ at which f is differentiable. Find

$$\nabla f(\vec{p}), df(\vec{p};\cdot), \text{ and } \left[f'(\vec{p})\right]. \tag{6.3.E.3}$$

[Hint: Use Theorem 3 and Corollary 1.]

Exercise $6.3.E.5$

$\Rightarrow$ Prove the following statements.
(i) If $f : E' \rightarrow E$ is constant on an open globe $G \subset E',$ it is differentiable at each $\vec{p} \in G,$ and $df(\vec{p},\cdot) = 0$ on $E'.$
(ii) If the latter holds for each $\vec{p} \in G - Q$ (Q countable), then f is constant on G (even on $\overline{G}$) provided f is relatively continuous there.
[Hint: Given $\vec{p}, \vec{q} \in G,$ use Theorem 2 in §1 to get $f(\vec{p}) = f(\vec{q})$.]

Exercise $6.3.E.6$

Do Problem 5 in case G is any open polygon-connected set in E'. (See Chapter 4, §9.)

Exercise $6.3.E.7$

$\Rightarrow$ Prove the following.
(i) If $f, g : E' \rightarrow E$ are differentiable at $\vec{p}$, so is

$$h = af + bg, \tag{6.3.E.4}$$

for any scalars a, b (if f and g are scalar valued, a and b may be vectors; moreover,

$$d(af+bg) = adf + bdg, \tag{6.3.E.5}$$

i.e.,

$$dh(\vec{p};\vec{t}) = adf(\vec{p};\vec{t}) + bdg(\vec{p};\vec{t}), \quad \vec{t} \in E'. \tag{6.3.E.6}$$

(ii) In case $f, g : E^m \rightarrow E^1$ or $C^m \rightarrow C$, deduce also that

$$\nabla h(\vec{p}) = a\nabla f(\vec{p}) + b\nabla g(\vec{p}). \tag{6.3.E.7}$$

Exercise $6.3.E.8$

⇒ Prove that if $f, g : E' \rightarrow E^1(C)$ are differentiable at $\vec{p}$, then so are

$$h = gf \text{ and } k = \frac{g}{f}. \tag{6.3.E.8}$$

(the latter, if $f(\vec{p}) \neq 0$). Moreover, with $a = f(\vec{p})$ and $b = g(\vec{p})$, show that
(i) $dh = adg + bdf$ and
(ii) $dk = (adg - bdf)/a^2$.
If further $E' = E^n\,(C^n)$, verify that
(iii) $\nabla h(\vec{p}) = a\nabla g(\vec{p}) + b\nabla f(\vec{p})$ and
(iv) $\nabla k(\vec{p}) = (a\nabla g(\vec{p}) - b\nabla f(\vec{p}))/a^2$.
Prove (i) and (ii) for vector-valued g, too.
[Hints: (i) Set $\phi = adg + bdf$, with a and b as above. Verify that

$$\Delta h - \phi(\vec{t}) = g(\vec{p})(\Delta f - df(\vec{t})) + f(\vec{p})(\Delta g - dg(\vec{t})) + (\Delta f)(\Delta g). \tag{6.3.E.9}$$

Use Problem 3(iv) and Definition 1.
(ii) Let $F(\vec{t}) = 1/f(\vec{t})$. Show that $dF = -df/a^2$. Then apply (i) to gF.]

Exercise $6.3.E.9$

⇒ Let $f : E' \rightarrow E^m\,(C^m), f = (f_1, \ldots, f_m)$. Prove that
(i) f is linear iff all its m components f_k are;
(ii) f is differentiable at $\vec{p}$ iff all f_k are, and then $df = (df_1, \ldots, df_m)$. Hence if f is complex, $df = df_{re} + i \cdot df_{im}$.

Exercise $6.3.E.10$

Prove the following statements.
(i) If $f \in L(E', E)$ then f is differentiable on E', and $df(\vec{p};\cdot) = f, \vec{p} \in E'$.
(ii) Such is any first-degree monomial, hence any sum of such monomials.

Exercise $6.3.E.11$

Any rational function is differentiable in its domain.
[Hint: Use Problems 10(i), 7, and 8. Proceed as in Theorem 3 in Chapter 4, §3.]

Exercise 6.3.E.12

Do Problem 8(i) in case g is only continuous at $\vec{p}$, and $f(\vec{p})=0$. Find dh.

Exercise 6.3.E.13

Do Problem 8(i) for dot products $h=f\cdot g$ of functions $f, g: E^{\prime} \rightarrow E^{m} \quad\left(C^{m}\right)$.

Exercise 6.3.E.14

Prove the following.
(i) If $\phi \in L\left(E^{n}, E^{1}\right)$ or $\phi \in L\left(C^{n}, C\right)$, then $\|\phi\|=|\vec{v}|$, with $\vec{v}$ as in §2, Theorem 2(ii).
(ii) If $f: E^{n} \rightarrow E^{1}\left(f: C^{n} \rightarrow C^{1}\right)$ is differentiable at $\vec{p}$, then

$$\|d f(\vec{p} ; \cdot)\|=|\nabla f(\vec{p})|. \tag{6.3.E.10}$$

Moreover, in case $f: E^{n} \rightarrow E^{1}$,

$$|\nabla f(\vec{p})| \geq D_{\vec{u}} f(\vec{p}) \quad \text{if } |\vec{u}|=1 \tag{6.3.E.11}$$

and

$$|\nabla f(\vec{p})|=D_{\vec{u}} f(\vec{p}) \quad \text{when } \vec{u}=\frac{\nabla f(\vec{p})}{|\nabla f(\vec{p})|}; \tag{6.3.E.12}$$

thus

$$|\nabla f(\vec{p})|=\max_{|\vec{u}|=1} D_{\vec{u}} f(\vec{p}). \tag{6.3.E.13}$$

[Hints: Use the equality case in Theorem 4(c') of Chapter 3, §§1-3. Use formula (7), Corollary 2, and Theorem 2(ii).]

Exercise 6.3.E.15

Show that Theorem 3 holds even if
(i) $D_{1} f$ is discontinuous at $\vec{p}$, and
(ii) f has partials on $A-Q$ only (Q countable, $\vec{p} \notin Q$), provided f is continuous on A in each of the last $n-1$ variables.
[Hint: For $k=1,$ formula (13) still results by definition of $D_{1} f$, if a suitable δ has been chosen.]

Exercise 6.3.E.16*

Show that Theorem 3 and Problem 15 apply also to any $f: E^{\prime} \rightarrow E$ where $E^{\prime}$ is n-dimensional with basis $\left\{\vec{u}_{1}, \ldots, \vec{u}_{n}\right\}$ (see Problem 12 in §2) if we write $D_{k} f$ for $D_{\vec{u}_{k}} f$.
[Hints: Assume $\left|\vec{u}_{k}\right|=1, 1 \leq k \leq n$ (if not, replace $\vec{u}_{k}$ by $\vec{u}_{k} /\left|\vec{u}_{k}\right|$; show that this yields another basis). Modify the proof so that the $\vec{p}_{k}$ are still in $G_{\vec{p}}(\delta)$. Caution: The standard norm of E^{n} does not apply here.]

Exercise 6.3.E.17

Let $f_{k}: E^{1} \rightarrow E^{1}$ be differentiable at $p_{k}(k=1, \ldots, n)$. For $\vec{x}=\left(x_{1}, \ldots, x_{n}\right) \in E^{n},$ set

$$F(\vec{x})=\sum_{k=1}^{n} f_{k}\left(x_{k}\right) \text{ and } G(\vec{x})=\prod_{k=1}^{n} f_{k}\left(x_{k}\right). \tag{6.3.E.14}$$

Show that F and G are differentiable at $\vec{p}=(p_1,\ldots,p_n)$. Express $\nabla F(\vec{p})$ and $\nabla G(\vec{p})$ in terms of the $f_k^{\prime}(p_k)$.
[Hint: In order to use Problems 7 and 8, replace the f_k by suitable functions defined on E^n. For $\nabla G(\vec{p})$, "imitate" Problem 6 in Chapter 5, §1.]

6.4: The Chain Rule. The Cauchy Invariant Rule

To generalize the chain rule (Chapter 5, §1), we consider the composite $h=g \circ f$ of two functions, $f: E^{\prime} \rightarrow E^{\prime \prime}$ and $g: E^{\prime \prime} \rightarrow E$, with $E^{\prime}, E^{\prime \prime}$, and E as before.

Theorem 6.4.1 (chain rule)

If

$$f: E^{\prime} \rightarrow E^{\prime \prime} \text { and } g: E^{\prime \prime} \rightarrow E \tag{6.4.1}$$

are differentiable at $\vec{p}$ and $\vec{q}=f(\vec{p})$, respectively, then

$$h=g \circ f \tag{6.4.2}$$

is differentiable at $\vec{p}$, and

$$d h(\vec{p} ; \cdot)=d g(\vec{q} ; \cdot) \circ d f(\vec{p} ; \cdot). \tag{6.4.3}$$

Briefly: "The differential of the composite is the composite of differentials."

Proof

Let $U=d f(\vec{p} ; \cdot), V=d g(\vec{q} ; \cdot)$, and $\phi=V \circ U$.

As U and V are linear continuous maps, so is ϕ. We must show that $\phi=d h(\vec{p} ; \cdot)$.

Here it is more convenient to write $\Delta \vec{x}$ or $\vec{x}-\vec{p}$ for the "$\vec{t}$" of Definition 1 in §3. For brevity, we set (with $\vec{q}=f(\vec{p})$)

$$\begin{aligned} w(\vec{x}) &=\Delta h-\phi(\Delta \vec{x})=h(\vec{x})-h(\vec{p})-\phi(\vec{x}-\vec{p}), \quad \vec{x} \in E^{\prime}, \\ u(\vec{x}) &=\Delta f-U(\Delta \vec{x})=f(\vec{x})-f(\vec{p})-U(\vec{x}-\vec{p}), \quad \vec{x} \in E^{\prime}, \\ v(\vec{y}) &=\Delta g-V(\Delta \vec{y})=g(\vec{y})-g(\vec{q})-V(\vec{y}-\vec{q}), \quad \vec{y} \in E^{\prime \prime}. \end{aligned}$$

Then what we have to prove (see Definition 1 in §3) reduces to

$$\lim _{\vec{x} \rightarrow \vec{p}} \frac{w(\vec{x})}{|\vec{x}-\vec{p}|}=0, \tag{6.4.4}$$

while the assumed existence of $d f(\vec{p} ; \cdot)=U$ and $d g(\vec{q} ; \cdot)=V$ can be expressed as

$$\lim _{\vec{x} \rightarrow \vec{p}} \frac{u(\vec{x})}{|\vec{x}-\vec{p}|}=0, \tag{6.4.5}$$

and

$$\lim _{\vec{y} \rightarrow \vec{q}} \frac{v(\vec{y})}{|\vec{y}-\vec{q}|}=0, \quad \vec{q}=f(\vec{p}). \tag{6.4.6}$$

From (2) and (3), recalling that $h=g \circ f$ and $\phi=V \circ U$, we obtain

$$\begin{aligned} w(\vec{x}) &=g(f(\vec{x}))-g(\vec{q})-V(U(\vec{x}-\vec{p})) \\ &=g(f(\vec{x}))-g(\vec{q})-V(f(\vec{x})-f(\vec{p})-u(\vec{x})). \end{aligned}$$

Using (4), with $\vec{y}=f(\vec{x})$, and the linearity of V, we rewrite (6) as

$$\begin{aligned} w(\vec{x}) &=g(f(\vec{x}))-g(\vec{q})-V(f(\vec{x})-f(\vec{p}))-V(u(\vec{x})) \\ &=v(f(\vec{x}))+V(u(\vec{x})). \end{aligned}$$

(Verify!) Thus the desired formula (5) will be proved if we show that

$$\lim _{\vec{x} \rightarrow \vec{p}} \frac{V(u(\vec{x}))}{|\vec{x}-\vec{p}|}=0 \tag{6.4.7}$$

and

$$\lim_{\vec{x} \rightarrow \vec{p}} \frac{v(f(\vec{x}))}{|\vec{x}-\vec{p}|}=0. \tag{6.4.8}$$

Now, as V is linear and continuous, formula (5') yields (6'). Indeed,

$$\lim_{\vec{x} \rightarrow \vec{p}} \frac{V(u(\vec{x}))}{|\vec{x}-\vec{p}|}=\lim_{\vec{x} \rightarrow \vec{p}} V\left(\frac{u(\vec{x})}{|\vec{x}-\vec{p}|}\right)=V(0)=0 \tag{6.4.9}$$

by Corollary 2 in Chapter 4, §2. (Why?)

Similarly, (5") implies (6") by substituting $\vec{y}=f(\vec{x}),$ since

$$|f(\vec{x})-f(\vec{p})| \leq K|\vec{x}-\vec{p}| \tag{6.4.10}$$

by Problem 3(iii) in §3. (Explain!) Thus all is proved. □

Note 1 (Cauchy invariant rule). Under the same assumptions, we also have

$$d h(\vec{p} ; \vec{t})=d g(\vec{q} ; \vec{s}) \tag{6.4.11}$$

if $\vec{s}=d f(\vec{p} ; \vec{t}), \vec{t} \in E^{\prime}$.

For with U and V as above,

$$d h(\vec{p} ; \cdot)=\phi=V \circ U. \tag{6.4.12}$$

Thus if

$$\vec{s}=d f(\vec{p} ; \vec{t})=U(\vec{t}), \tag{6.4.13}$$

we have

$$d h(\vec{p} ; \vec{t})=\phi(\vec{t})=V(U(\vec{t}))=V(\vec{s})=d g(\vec{q} ; \vec{s}), \tag{6.4.14}$$

proving (7).

Note 2. If

$$E^{\prime}=E^{n}\left(C^{n}\right), E^{\prime \prime}=E^{m}\left(C^{m}\right), \text { and } E=E^{r}\left(C^{r}\right) \tag{6.4.15}$$

then by Theorem 3 of §2 and Definition 2 in §3, we can write (1) in matrix form,

$$\left[h^{\prime}(\vec{p})\right]=\left[g^{\prime}(\vec{q})\right]\left[f^{\prime}(\vec{p})\right], \tag{6.4.16}$$

resembling Theorem 3 in Chapter 5, §1 (with f and g interchanged). Moreover, we have the following theorem.

Theorem 6.4.2

With all as in Theorem 1, let

$$E^{\prime}=E^{n}\left(C^{n}\right), E^{\prime \prime}=E^{m}\left(C^{m}\right), \tag{6.4.17}$$

and

$$f=\left(f_{1}, \ldots, f_{m}\right). \tag{6.4.18}$$

Then

$$D_{k} h(\vec{p})=\sum_{i=1}^{m} D_{i} g(\vec{q}) D_{k} f_{i}(\vec{p}); \tag{6.4.19}$$

or, in classical notation,

$$\frac{\partial}{\partial x_{k}} h(\vec{p})=\sum_{i=1}^{m} \frac{\partial}{\partial y_{i}} g(\vec{q}) \cdot \frac{\partial}{\partial x_{k}} f_{i}(\vec{p}), \quad k=1,2, \ldots, n. \tag{6.4.20}$$

Proof

Fix any basic vector $\vec{e}_k$ in E' and set

$$\vec{s}=df\left(\vec{p};\vec{e}_k\right), \quad \vec{s}=\left(s_1,\ldots,s_m\right)\in E''. \tag{6.4.21}$$

As f is differentiable at $\vec{p}$, so are its components f_i (Problem 9 in §3), and

$$s_i=df_i\left(\vec{p};\vec{e}_k\right)=D_k f_i(\vec{p}) \tag{6.4.22}$$

by Theorem 2(ii) in §3. Using also Corollary 1 in §3, we get

$$dg(\vec{q};\vec{s})=\sum_{i=1}^{m} s_i D_i g(\vec{q})=\sum_{i=1}^{m} D_k f_i(\vec{p}) D_i g(\vec{q}). \tag{6.4.23}$$

But as $\vec{s}=df\left(\vec{p};\vec{e}_k\right),$ formula (7) yields

$$dg(\vec{q};\vec{s})=dh\left(\vec{p};\vec{e}_k\right)=D_k h(\vec{p}) \tag{6.4.24}$$

by Theorem 2(ii) in §3. Thus the result follows. $\square$

Note 3. Theorem 2 is often called the chain rule for functions of several variables. It yields Theorem 3 in Chapter 5, §1, if $m=n=1$.

In classical calculus one often speaks of derivatives and differentials of variables $y=f\left(x_1,\ldots,x_n\right)$ rather than those of mappings. Thus Theorem 2 is stated as follows.

Let $u=g\left(y_1,\ldots,y_m\right)$ be differentiable. If, in turn, each

$$y_i=f_i\left(x_1,\ldots,x_n\right) \tag{6.4.25}$$

is differentiable for $i=1,\ldots,m$, then u is also differentiable as a composite function of the n variables x_k, and ("simplifying" formula (8)) we have

$$\frac{\partial u}{\partial x_k}=\sum_{i=1}^{m} \frac{\partial u}{\partial y_i} \frac{\partial y_i}{\partial x_k}, \quad k=1,2,\ldots,n. \tag{6.4.26}$$

It is understood that the partials

$$\frac{\partial u}{\partial x_k} \text{ and } \frac{\partial y_i}{\partial x_k} \text{ are taken at some } \vec{p}\in E', \tag{6.4.27}$$

while the $\partial u/\partial y_i$ are at $\vec{q}=f(\vec{p})$, where $f=\left(f_1,\ldots,f_m\right)$. This "variable" notation is convenient in computations, but may cause ambiguities (see the next example).

Example

Let $u=g(x,y,z)$, where z depends on x and y :

$$z=f_3(x,y). \tag{6.4.28}$$

Set $f_1(x,y)=x$, $f_2(x,y)=y$, $f=\left(f_1,f_2,f_3\right)$, and $h=g\circ f;$ so

$$h(x,y)=g(x,y,z). \tag{6.4.29}$$

By (8'),

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial x}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial x}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial x}. \tag{6.4.30}$$

Here

$$\frac{\partial x}{\partial x}=\frac{\partial f_1}{\partial x}=1 \text{ and } \frac{\partial y}{\partial x}=0, \tag{6.4.31}$$

for f_2 does not depend on x. Thus we obtain

$$\frac{\partial u}{\partial x}=\frac{\partial u}{\partial x}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial x}. \tag{6.4.32}$$

(Question: Is $(\partial u/\partial z)(\partial z/\partial x)=0$?)

The trouble with (9) is that the variable u "poses" as both g and h. On the left, it is h; on the right, it is g.

To avoid this, our method is to differentiate well-defined mappings, not "variables." Thus in (9), we have the maps

$$g: E^3 \rightarrow E \text{ and } f: E^2 \rightarrow E^3, \tag{6.4.33}$$

with f_1, f_2, f_3 as indicated. Then if $h=g\circ f$, Theorem 2 states (9) unambiguously as

$$D_1 h(\vec{p})=D_1 g(\vec{q})+D_3 g(\vec{q})\cdot D_1 f(\vec{p}), \tag{6.4.34}$$

where $\vec{p}\in E^2$ and

$$\vec{q}=f(\vec{p})=(p_1, p_2, f_3(\vec{p})). \tag{6.4.35}$$

(Why?) In classical notation,

$$\frac{\partial h}{\partial x}=\frac{\partial g}{\partial x}+\frac{\partial g}{\partial z}\frac{\partial f_3}{\partial x} \tag{6.4.36}$$

(avoiding the "paradox" of (9)).

Nonetheless, with due caution, one may use the "variable" notation where convenient. The reader should practice both (see the Problems).

Note 4. The Cauchy rule (7), in "variable" notation, turns into

$$du=\sum_{i=1}^{m}\frac{\partial u}{\partial y_i}dy_i=\sum_{k=1}^{n}\frac{\partial u}{\partial x_k}dx_k, \tag{6.4.37}$$

where $dx_k=t_k$ and $dy_i=df_i(\vec{p};\vec{t})$.

Indeed, by Corollary 1 in §3,

$$dh(\vec{p};\vec{t})=\sum_{k=1}^{n}D_k h(\vec{p})\cdot t_k \text{ and } dg(\vec{q};\vec{s})=\sum_{i=1}^{m}D_i g(\vec{q})\cdot s_i. \tag{6.4.38}$$

Now, in (7),

$$\vec{s}=(s_1,\ldots,s_m)=df(\vec{p};\vec{t}); \tag{6.4.39}$$

so by Problem 9 in §3,

$$df_i(\vec{p};\vec{t})=s_i, \quad i=1,\ldots,m. \tag{6.4.40}$$

Rewriting all in the "variable" notation, we obtain (10).

The "advantage" of (10) is that du has the same form, independently of whether u is treated as a function of the x_k or of the y_i (hence the name "invariant" rule). However, one must remember the meaning of dx_k and dy_i, which are quite different.

The "invariance" also fails completely for differentials of higher order (§5).

The advantages of the "variable" notation vanish unless one is able to "translate" it into precise formulas.

6.4.E: Further Problems on Differentiable Functions

Exercise $6.4.E.1$

For $E=E^r\left(C^r\right)$ prove Theorem 2 directly.
[Hint: Find

$$D_k h_j(\vec{p}), \quad j=1,\ldots,r, \tag{6.4.E.1}$$

from Theorem 4 of §3, and Theorem 3 of §2. Verify that

$$D_k h(\vec{p})=\sum_{j=1}^r e_j D_k h_j(\vec{p}) \text{ and } D_i g(\vec{q})=\sum_{j=1}^r e_j D_i g_j(\vec{q}), \tag{6.4.E.2}$$

where the e_j are the basic unit vectors in E^r. Proceed.]

Exercise $6.4.E.2$

Let $g(x,y,z)=u, x=f_1(r,\theta), y=f_2(r,\theta), z=f_3(r,\theta)$, and

$$f=\left(f_1,f_2,f_3\right): E^2\rightarrow E^3. \tag{6.4.E.3}$$

Assuming differentiability, verify (using "variables") that

$$du=\frac{\partial u}{\partial x}dx+\frac{\partial u}{\partial y}dy+\frac{\partial u}{\partial z}dz=\frac{\partial u}{\partial r}dr+\frac{\partial u}{\partial \theta}d\theta \tag{6.4.E.4}$$

by computing derivatives from (8'). Then do all in the mapping notation for $H=g\circ f, dH(\vec{p};\vec{t})$.

Exercise $6.4.E.3$

For the specific functions f, g, h, and k of Problems 4 and 5 of §2, set up and solve problems analogous to Problem 2, using

$$\text{(a) } k\circ f; \quad \text{(b) } g\circ k; \quad \text{(c) } f\circ h; \quad \text{(d) } h\circ g. \tag{6.4.E.5}$$

Exercise $6.4.E.4$

For the functions of Problem 5 in §1, find the formulas for $df(\vec{p};\vec{t})$. At which $\vec{p}$ does $df(\vec{p};\cdot)$ exist in each given case? Describe it for a chosen $\vec{p}$.

Exercise $6.4.E.5$

From Theorem 2, with $E=E^1(C)$, find

$$\nabla h(\vec{p})=\sum_{k=1}^n D_k g(\vec{q})\nabla f_k(\vec{p}). \tag{6.4.E.6}$$

Exercise $6.4.E.6$

Use Theorem 1 for a new solution of Problem 7 in §3 with $E=E^1(C)$.
[Hint: Define F on E' and G on $E^2\left(C^2\right)$ by

$$F(\vec{x})=(f(\vec{x}),g(\vec{x})) \text{ and } G(\vec{y})=ay_1+by_2. \tag{6.4.E.7}$$

Then $h=af+bg=G\circ F$. (Why?) Use Problems 9 and 10(ii) of §3. Do all in "variable" notation, too.]

Exercise $6.4.E.7$

Use Theorem 1 for a new proof of the "only if " in Problem 9 in §3.
[Hint: Set $f_i=g\circ f$, where $g(\vec{x})=x_i$ (the ith "projection map") is a monomial. Verify!]

Exercise $6.4.E.8$

Do Problem 8(I) in §3 for the case $E'=E^2\left(C^2\right)$, with

$$f(\vec{x})=x_1 \text{ and } g(\vec{x})=x_2. \tag{6.4.E.8}$$

(Simplify!) Then do the general case as in Problem 6 above, with

$$G(\vec{y})=y_1y_2. \tag{6.4.E.9}$$

Exercise $6.4.E.9$

Use Theorem 2 for a new proof of Theorem 4 in Chapter 5, §1. (Proceed as in Problems 6 and 8, with $E'=E^1$, so that $D_1h=h'$). Do it in the "variable" notation, too.

Exercise $6.4.E.10$

Under proper differentiability assumptions, use formula (8') to express the partials of u if
(i) $u=g(x,y), x=f(r)h(\theta), y=r+h(\theta)+\theta f(r)$;
(ii) $u=g(r,\theta), r=f(x+f(y)), \theta=f(xf(y))$;
(iii) $u=g\left(x^y,y^z,z^{x+y}\right)$.
Then redo all in the "mapping" terminology, too.

Exercise $6.4.E.11$

Let the map $g:E^1\rightarrow E^1$ be differentiable on E^1. Find $|\nabla h(\vec{p})|$ if
$h=g\circ f$ and
(i) $f(\vec{x})=\sum_{k=1}^n x_k, \vec{x}\in E^n$;
(ii) $f(\vec{x})=|\vec{x}|^2, \vec{x}\in E^n$.

Exercise $6.4.E.12$

(Euler's theorem.) A map $f:E^n\rightarrow E^1$ (or $C^n\rightarrow C$) is called homogeneous of degree m on G iff

$$\left(\forall t\in E^1(C)\right) \quad f(t\vec{x})=t^m f(\vec{x}) \tag{6.4.E.10}$$

when $\vec{x}, t\vec{x}\in G$. Prove the following statements.
(i) If so, and f is differentiable at $\vec{p}\in G$ (an open globe), then

$$\vec{p}\cdot\nabla f(\vec{p})=mf(\vec{p}). \tag{6.4.E.11}$$

*(ii) Conversely, if the latter holds for all $\vec{p}\in G$ and if $\overrightarrow{0}\notin G$, then f is homogeneous of degree m on G.
(iii) What if $\overrightarrow{0}\in G$?
[Hints: (i) Let $g(t)=f(t\vec{p})$. Find $g'(1)$. (iii) Take $f(x,y)=x^2y^2$ if $x\leq 0, f=0$ if $x>0, G=G_0(1)$.]

Exercise $6.4.E.13$

Try Problem 12 for $f: E' \rightarrow E,$ replacing $\vec{p} \cdot \nabla f(\vec{p})$ by $df(\vec{p}; \vec{p})$.

Exercise $6.4.E.14$

With all as in Theorem 1, prove the following.

(i) If $E' = E^1$ and $\vec{s} = f'(p) \neq \overrightarrow{0}$, then $h'(p) = D_{\vec{s}}\, g(\vec{q})$.

(ii) If $\vec{u}$ and $\vec{v}$ are nonzero in E' and $a\vec{u} + b\vec{v} \neq \overrightarrow{0}$ for some scalars a, b, then

$$D_{a\vec{u}+b\vec{v}} f(\vec{p}) = a D_{\vec{u}} f(\vec{p}) + b D_{\vec{v}} f(\vec{p}). \tag{6.4.E.12}$$

(iii) If f is differentiable on a globe $G_{\vec{p}}$, and $\vec{u} \neq \overrightarrow{0}$ in E', then

$$D_{\vec{u}} f(\vec{p}) = \lim_{\vec{x} \rightarrow \vec{u}} D_{\vec{x}}(\vec{p}). \tag{6.4.E.13}$$

[Hints: Use Theorem 2(ii) from §3 and Note 1.]

Exercise $6.4.E.15$

Use Theorem 2 to find the partially derived functions of f, if
(i) $f(x, y, z) = (\sin(xy/z))^x$;
(ii) $f(x, y) = \log_x |\tan(y/x)|$.
(Set $f = 0$ wherever undefined.)

6.5: Repeated Differentiation. Taylor's Theorem

In §1 we defined $\vec{u}$-directed derived functions, $D_{\vec{u}} f$ for any $f: E' \rightarrow E$ and any $\vec{u} \neq \overrightarrow{0}$ in E'.

Thus given a sequence $\{\vec{u}_i\} \subseteq E' - \{\overrightarrow{0}\}$, we can first form $D_{\vec{u}_1} f$, then $D_{\vec{u}_2}\left(D_{\vec{u}_1} f\right)$ (the $\vec{u}_2$-directed derived function of $D_{\vec{u}_1} f$), then the $\vec{u}_3$-directed derived function of $D_{\vec{u}_2}\left(D_{\vec{u}_1} f\right)$, and so on. We call all functions so formed the higher-order directional derived functions of f.

If at each step the limit postulated in Definition 1 of §1 exists for all $\vec{p}$ in a set $B \subseteq E'$, we call them the higher-order directional derivatives of f (on B).

If all $\vec{u}_i$ are basic unit vectors in $E^n\left(C^n\right)$, we say "partial' instead of "directional."

We also define $D_{\vec{u}}^1 f = D_{\vec{u}} f$ and

$$D_{\vec{u}_1 \vec{u}_2 \ldots \vec{u}_k}^k f = D_{\vec{u}_k}\left(D_{\vec{u}_1 \vec{u}_2 \ldots \vec{u}_{k-1}}^{k-1} f\right), \quad k = 2, 3, \ldots, \tag{6.5.1}$$

and call $D_{\vec{u}_1 \vec{u}_2 \ldots \vec{u}_k}^k f$ a directional derived function of order k. (Some authors denote it by $D_{\vec{u}_k \vec{u}_{k-1} \ldots \vec{u}_1} f$.)

If all $\vec{u}_i$ equal $\vec{u}$, we write $D_{\vec{u}}^k f$ instead.

For partially derived functions, we simplify this notation, writing $12\ldots$ for $\vec{e}_1 \vec{e}_2 \ldots$ and omitting the "k" in D^k (except in classical notation):

$$D_{12} f = D_{\vec{e}_1 \vec{e}_2}^2 f = \frac{\partial^2 f}{\partial x_1 \partial x_2}, \quad D_{11} f = D_{\vec{e}_1 \vec{e}_1}^2 f = \frac{\partial^2 f}{\partial x_1^2}, \text{ etc.} \tag{6.5.2}$$

We also set $D_{\vec{u}}^0 f = f$ for any vector $\vec{u}$.

Example

(A) Define $f: E^2 \rightarrow E^1$ by

$$f(0,0) = 0, \quad f(x, y) = \frac{xy\left(x^2 - y^2\right)}{x^2 + y^2}. \tag{6.5.3}$$

Then

$$\frac{\partial f}{\partial x} = D_1 f(x, y) = \frac{y\left(x^4 + 4x^2 y^2 - y^4\right)}{\left(x^2 + y^2\right)^2}, \tag{6.5.4}$$

whence $D_1 f(0, y) = -y$ if $y \neq 0$; and also

$$D_1 f(0,0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0,0)}{x} = 0. \quad \text{(Verify!)} \tag{6.5.5}$$

Thus $D_1 f(0, y) = -y$ always, and so $D_{12} f(0, y) = -1$; $D_{12} f(0,0) = -1$ Similarly,

$$D_2 f(x, y) = \frac{x\left(x^4 - 4x^2 y^2 - y^4\right)}{\left(x^2 + y^2\right)^2} \tag{6.5.6}$$

if $x \neq 0$ and $D_2 f(0,0) = 0$. Thus $(\forall x) D_2 f(x, 0) = x$ and so

$$D_{21} f(x, 0) = 1 \text{ and } D_{21} f(0,0) = 1 \neq D_{12} f(0,0) = -1. \tag{6.5.7}$$

The previous example shows that we may well have $D_{12} f \neq D_{21} f$, or more generally, $D_{\vec{u}\vec{v}}^2 f \neq D_{\vec{v}\vec{u}}^2 f$. However, we obtain the following theorem.

Theorem 6.5.1

Given nonzero vectors $\vec{u}$ and $\vec{v}$ in E', suppose $f : E' \rightarrow E$ has the derivatives

$$D_{\vec{u}} f, D_{\vec{v}} f, \text{ and } D_{\vec{u}\vec{v}}^2 f \tag{6.5.8}$$

on an open set $A \subseteq E'$.

If $D_{\vec{u}\vec{v}}^2 f$ is continuous at some $\vec{p} \in A$, then the derivative $D_{\vec{v}\vec{u}}^2 f(\vec{p})$ also exists and equals $D_{\vec{u}\vec{v}}^2 f(\vec{p})$.

Proof

By Corollary 1 in §1, all reduces to the case $|\vec{u}| = 1 = |\vec{v}|$. (Why?)

Given $\varepsilon > 0$, fix $\delta > 0$ so small that $G = G_{\vec{p}}(\delta) \subseteq A$ and simultaneously

$$\sup_{\vec{x} \in G} \left| D_{\vec{u}\vec{v}}^2 f(\vec{x}) - D_{\vec{u}\vec{v}}^2 f(\vec{p}) \right| \leq \varepsilon \tag{6.5.9}$$

(by the continuity of $D_{\vec{u}\vec{v}}^2 f$ at $\vec{p}$).

Now $\left(\forall s, t \in E^1\right)$ define $H_t : E^1 \rightarrow E$ by

$$H_t(s) = D_{\vec{u}} f(\vec{p} + t\vec{u} + s\vec{v}). \tag{6.5.10}$$

Let

$$I = \left(-\frac{\delta}{2}, \frac{\delta}{2}\right). \tag{6.5.11}$$

If $s, t \in I$, the point $\vec{x} = \vec{p} + t\vec{u} + s\vec{v}$ is in $G_{\vec{p}}(\delta) \subseteq A$, since

$$|\vec{x} - \vec{p}| = |t\vec{u} + s\vec{v}| < \frac{\delta}{2} + \frac{\delta}{2} = \delta. \tag{6.5.12}$$

Thus by assumption, the derivative $D_{\vec{u}\vec{v}}^2 f(\vec{p})$ exists. Also,

$$\begin{aligned} H_t'(s) &= \lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s}\left[H_t(s + \Delta s) - H_t(s)\right] \\ &= \lim_{\Delta s \rightarrow 0} \frac{1}{\Delta s}\left[D_{\vec{u}} f(\vec{x} + \Delta s \cdot \vec{v}) - D_{\vec{u}} f(\vec{x})\right]. \end{aligned}$$

But the last limit is $D_{\vec{u}\vec{v}}^2 f(\vec{x})$, by definition. Thus, setting

$$h_t(s) = H_t(s) - s D_{\vec{u}\vec{v}}^2 f(\vec{p}), \tag{6.5.13}$$

we get

$$\begin{aligned} h_t'(s) &= H_t'(s) - D_{\vec{u}\vec{v}}^2 f(\vec{p}) \\ &= D_{\vec{u}\vec{v}}^2 f(\vec{x}) - D_{\vec{u}\vec{v}}^2 f(\vec{p}). \end{aligned}$$

We see that h_t is differentiable on I, and by (2),

$$\sup_{s \in I} \left|h_t'(s)\right| \leq \sup_{\vec{x} \in G} \left| D_{\vec{u}\vec{v}}^2 f(\vec{x}) - D_{\vec{u}\vec{v}}^2 f(\vec{p}) \right| \leq \varepsilon \tag{6.5.14}$$

for all $t \in I$. Hence by Corollary 1 of Chapter 5, §4,

$$\left|h_t(s) - h_t(0)\right| \leq \left| s \right| \sup_{\sigma \in I} \left| h_t'(\sigma)\right| \leq |s| \varepsilon. \tag{6.5.15}$$

But by definition,

$$h_t(s) = D_{\vec{u}} f(\vec{p} + t\vec{u} + s\vec{v}) - s D_{\vec{u}\vec{v}}^2 f(\vec{p}) \tag{6.5.16}$$

and

$$h_t(0) = D_{\vec{u}} f(\vec{p} + t\vec{u}). \tag{6.5.17}$$

Thus

$$\left| D_{\vec{u}} f(\vec{p} + t\vec{u} + s\vec{v}) - D_{\vec{u}} f(\vec{p} + t\vec{u}) - s D^2_{\vec{u}\vec{v}} f(\vec{p}) \right| \leq |s|\varepsilon \tag{6.5.18}$$

for all $s, t \in I$.

Next, set

$$G_s(t) = f(\vec{p} + t\vec{u} + s\vec{v}) - f(\vec{p} + t\vec{u}) \tag{6.5.19}$$

and

$$g_s(t) = G_s(t) - st \cdot D^2_{\vec{u}\vec{v}} f(\vec{p}). \tag{6.5.20}$$

As before, one finds that $(\forall s \in I) g_s$ is differentiable on I and that

$$g_s'(t) = D_{\vec{u}} f(\vec{p} + t\vec{u} + s\vec{v}) - D_{\vec{u}} f(\vec{p} + t\vec{u}) - s D^2_{\vec{u}\vec{v}} f(\vec{p}) \tag{6.5.21}$$

for $s, t \in I$. (Verify!)

Hence by (3),

$$\sup_{t \in I} |g_s'(t)| \leq |s|\varepsilon. \tag{6.5.22}$$

Again, by Corollary 1 of Chapter 5, §4,

$$|g_s(t) - g_s(0)| \leq |st|\varepsilon, \tag{6.5.23}$$

or by the definition of g_s (assuming $s, t \in I - \{0\}$ and dividing by st),

$$\left| \frac{1}{st}[f(\vec{p} + t\vec{u} + s\vec{v}) - f(\vec{p} + t\vec{u})] - D^2_{\vec{u}\vec{v}} f(\vec{p}) - \frac{1}{st}[f(\vec{p} + s\vec{v}) - f(\vec{p})] \right| \leq \varepsilon. \tag{6.5.24}$$

(Verify!) Making $s \rightarrow 0$ (with t fixed), we get, by the definition of $D_{\vec{v}} f$,

$$\left| \frac{1}{t} D_{\vec{v}} f(\vec{p} + t\vec{u}) - \frac{1}{t} D_{\vec{v}} f(\vec{p}) - D^2_{\vec{u}\vec{v}} f(\vec{p}) \right| \leq \varepsilon \tag{6.5.25}$$

whenever $0 < |t| < \delta/2$.

As ε is arbitrary, we have

$$D^2_{\vec{u}\vec{v}} f(\vec{p}) = \lim_{t \rightarrow 0} \frac{1}{t}[D_v f(\vec{p} + t\vec{u}) - D_{\vec{v}} f(\vec{p})]. \tag{6.5.26}$$

But by definition, this limit is the derivative $D^2_{\vec{v}\vec{u}} f(\vec{p})$. Thus all is proved. □

Note 1. By induction, the theorem extends to derivatives of order > 2. Thus the derivative $D_{\vec{u}_1 \vec{u}_2 \ldots \vec{u}_k} f$ is independent of the order in which the $\vec{u}_i$ follow each other if it exists and is continuous on an open set $A \subseteq E'$, along with appropriate derivatives of order $< k$.

If $E' = E^n\left(C^n\right)$, this applies to partials as a special case.

For E^n and C^n only, we also formulate the following definition.

Definition 1

Let $E' = E^n\left(C^n\right)$. We say that $f : E' \rightarrow E$ is m times differentiable at $\vec{p} \in E'$ iff f and all its partials of order $< m$ are differentiable at $\vec{p}$.

If this holds for all $\vec{p}$ in a set $B \subseteq E'$, we say that f is m times differentiable on B.

If, in addition, all partials of order m are continuous at $\vec{p}$ (on B), we say that f is of class CD^m, or continuously differentiable m times there, and write $f \in CD^m$ at $\vec{p}$ (on B).

Finally, if this holds for all natural m, we write $f \in CD^\infty$ at $\vec{p}$ (on B, respectively).

Definition 2

Given the space $E' = E^n\left(C^n\right)$, the function $f : E' \rightarrow E$, and a point $\vec{p} \in E'$, we define the mappings

$$d^m f(\vec{p}; \cdot), \quad m = 1, 2, \ldots, \tag{6.5.27}$$

from E' to E by setting for every $\vec{t} = (t_1, \ldots, t_n)$

$$\begin{aligned} d^1 f(\vec{p}; \vec{t}) &= \sum_{i=1}^n D_i f(\vec{p}) \cdot t_i, \\ d^2 f(\vec{p}; \vec{t}) &= \sum_{j=1}^n \sum_{i=1}^n D_{ij} f(\vec{p}) \cdot t_i t_j, \\ d^3 f(\vec{p}; \vec{t}) &= \sum_{k=1}^n \sum_{j=1}^n \sum_{i=1}^n D_{ijk} f(\vec{p}) \cdot t_i t_j t_k, \quad \text{and so on.} \end{aligned}$$

We call $d^m f(\vec{p}; \cdot)$ the mth differential (or differential of order m) of f at $\vec{p}$. By our conventions, it is always defined on $E^n\left(C^n\right)$ as are the partially derived functions involved.

If f is differentiable at $\vec{p}$ (but not otherwise), then $d^1 f(\vec{p}; \vec{t}) = df(\vec{p}; \vec{t})$ by Corollary 1 in §3; $d^1 f(\vec{p}; \cdot)$ is linear and continuous (why?) but need not satisfy Definition 1 in §3.

In classical notation, we write dx_i for t_i; e.g.,

$$d^2 f = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j. \tag{6.5.28}$$

Note 2. Classical analysis tends to define differentials as above in terms of partials. Formula (4) for $d^m f$ is often written symbolically:

$$d^m f = \left(\frac{\partial}{\partial x_1} dx_1 + \frac{\partial}{\partial x_2} dx_2 + \cdots + \frac{\partial}{\partial x_n} dx_n\right)^m f, \quad m = 1, 2, \ldots \tag{6.5.29}$$

Indeed, raising the bracketed expression to the mth "power" as in algebra (removing brackets, without collecting "similar" terms) and then "multiplying" by f, we obtain sums that agree with (4). (Of course, this is not genuine multiplication but only a convenient memorizing device.)

Example

(B) Define $f : E^2 \rightarrow E^1$ by

$$f(x, y) = x \sin y. \tag{6.5.30}$$

Take any $\vec{p} = (x, y) \in E^2$. Then

$$D_1 f(x, y) = \sin y \text{ and } D_2 f(x, y) = x \cos y; \tag{6.5.31}$$

$$D_{12} f(x, y) = D_{21} f(x, y) = \cos y, \tag{6.5.32}$$

$$D_{11} f(x, y) = 0, \text{ and } D_{22} f(x, y) = -x \sin y; \tag{6.5.33}$$

$$D_{111} f(x, y) = D_{112} f(x, y) = D_{121} f(x, y) = D_{211} f(x, y) = 0, \tag{6.5.34}$$

$$D_{221} f(x, y) = D_{212} f(x, y) = D_{122} f(x, y) = -\sin y, \text{ and} \tag{6.5.35}$$

$$D_{222} f(x, y) = -x \cos y; \text{ etc.} \tag{6.5.36}$$

As is easily seen, f has continuous partials of all orders; so $f \in CD^{\infty}$ on all of E^2. Also,

$$\begin{aligned} df(\vec{p}; \vec{t}) &= t_1 D_1 f(\vec{p}) + t_2 D_2 f(\vec{p}) \\ &= t_1 \sin y + t_2 x \cos y. \end{aligned}$$

In classical notation,

$$\begin{aligned} df &= d^1 f = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \\ &= \sin y dx + x \cos y dy; \\ d^2 f &= \frac{\partial^2 f}{\partial x^2} dx^2 + 2 \frac{\partial^2 f}{\partial x \partial y} dx dy + \frac{\partial^2 f}{\partial y^2} dy^2 \\ &= 2 \cos y dx dy - x \sin y dy^2; \\ d^3 f &= -3 \sin y dx dy^2 - x \cos y dy^3; \end{aligned}$$

and so on. (Verify!)

We can now extend Taylor's theorem (Theorem 1 in Chapter 5, §6) to the case $E' = E^n\left(C^n\right)$.

Theorem 6.5.2 (Taylor)

Let $\vec{u} = \vec{x} - \vec{p} \neq \overrightarrow{0}$ in $E' = E^n\left(C^n\right)$.

If $f : E' \rightarrow E$ is $m+1$ times differentiable on the line segment

$$I = L[\vec{p}, \vec{x}] \subset E' \tag{6.5.37}$$

then

$$f(\vec{x}) - f(\vec{p}) = \sum_{i=1}^{m} \frac{1}{i!} d^i f(\vec{p}; \vec{u}) + R_m, \tag{6.5.38}$$

with

$$\left|R_m\right| \leq \frac{K_m}{(m+1)!}, K_m \in E^1, \tag{6.5.39}$$

and

$$0 \leq K_m \leq \sup_{\vec{s} \in I} \left|d^{m+1} f(\vec{s}; \vec{u})\right|. \tag{6.5.40}$$

Proof

Define $g : E^1 \rightarrow E'$ and $h : E^1 \rightarrow E$ by $g(t) = \vec{p} + t\vec{u}$ and $h = f \circ g$.

As $E' = E^n\left(C^n\right)$, we may consider the components of g,

$$g_k(t) = p_k + t u_k, \quad k \leq n. \tag{6.5.41}$$

Clearly, g_k is differentiable, $g_k'(t) = u_k$.

By assumption, so is f on $I = L[\vec{p}, \vec{x}]$. Thus, by the chain rule, $h = f \circ g$ is differentiable on the interval $J = [0, 1] \subset E^1$; for, by definition,

$$\vec{p} + t\vec{u} \in L[\vec{p}, \vec{x}] \text{ iff } t \in [0, 1]. \tag{6.5.42}$$

By Theorem 2 in §4,

$$h'(t) = \sum_{k=1}^{n} D_k f(\vec{p} + t\vec{u}) \cdot u_k = df(\vec{p} + t\vec{u}; \vec{u}), \quad t \in J. \tag{6.5.43}$$

(Explain!)

By assumption (and Definition 1), the $D_k f$ are differentiable on I. Hence, by (7), h' is differentiable on J. Reapplying Theorem 2 in §4, we obtain

$$\begin{aligned} h''(t) &= \sum_{j=1}^{n} \sum_{k=1}^{n} D_{kj} f(\vec{p}+t\vec{u}) \cdot u_k u_j \\ &= d^2 f(\vec{p}+t\vec{u}; \vec{u}), \quad t \in J. \end{aligned}$$

By induction, h is $m+1$ times differentiable on J, and

$$h^{(i)}(t) = d^i f(\vec{p}+t\vec{u}; \vec{u}), \quad t \in J, i = 1, 2, \ldots, m+1. \tag{6.5.44}$$

The differentiability of $h^{(i)} (i \leq m)$ implies its continuity on $J = [0, 1]$.

Thus h satisfies Theorem 1 of Chapter 5, §6 (with $x = 1, p = 0,$ and $Q = \emptyset$); hence

$$\begin{aligned} h(1) - h(0) &= \sum_{i=1}^{m} \frac{h^{(i)}(0)}{i!} + R_m, \\ |R_m| &\leq \frac{K_m}{(m+1)!}, \quad K_m \in E^1, \\ K_m &\leq \sup_{t \in J} \left| h^{(m+1)}(t) \right|. \end{aligned}$$

By construction,

$$h(t) = f(g(t)) = f(\vec{p}+t\vec{u}); \tag{6.5.45}$$

so

$$h(1) = f(\vec{p}+\vec{u}) = f(\vec{x}) \text{ and } h(0) = f(\vec{p}). \tag{6.5.46}$$

Thus using (8) also, we see that (9) implies (6), indeed. □

Note 3. Formula (3') of Chapter 5, §6, combined with (8), also yields

$$\begin{aligned} R_m &= \frac{1}{m!} \int_0^1 h^{(m+1)}(t) \cdot (1-t)^m dt \\ &= \frac{1}{m!} \int_0^1 d^{m+1} f(\vec{p}+t\vec{u}; \vec{u}) \cdot (1-t)^m dt. \end{aligned}$$

Corollary 6.5.1 (the Lagrange form of R_m)

If $E = E^1$ in Theorem 2, then

$$R_m = \frac{1}{(m+1)!} d^{m+1} f(\vec{s}; \vec{u}) \tag{6.5.47}$$

for some $\vec{s} \in L(\vec{p}, \vec{x})$.

Proof

Here the function h defined in the proof of Theorem 2 is real; so Theorem 1' and formula (3') of Chapter 5, §6 apply. This yields (10). Explain! □

Corollary 6.5.2

If $f : E^n\ (C^n) \rightarrow E$ is m times differentiable at $\vec{p}$ and if $\vec{u} \neq \vec{0}$ $(\vec{p}, \vec{u} \in E^n\ (C^n))$, then the derivative $D_{\vec{u}}^m f(\vec{p})$ exists and equals $d^m f(\vec{p}; \vec{u})$.

This follows as in the proof of Theorem 2 (with $t = 0$). For by definition,

$$\begin{aligned} D_{\vec{u}} f(\vec{p}) &= \lim_{s \to 0} \frac{1}{s}[f(\vec{p}+s\vec{u}) - f(\vec{p})] \\ &= \lim \frac{1}{s}[h(s) - h(0)] \\ &= h'(0) = df(\vec{p};\vec{u}) \end{aligned}$$

by (7). Induction yields

$$D_{\vec{u}}^{m} f(\vec{p}) = h^{(m)}(0) = d^{m}(\vec{p};\vec{u}) \tag{6.5.48}$$

by (8). (See Problem 3.

Example

(C) Continuing Example (B), fix

$$\vec{p} = (1,0); \tag{6.5.49}$$

thus replace (x,y) by $(1,0)$ there. Instead, write (x,y) for $\vec{x}$ in Theorem 2. Then

$$\vec{u} = \vec{x} - \vec{p} = (x-1, y); \tag{6.5.50}$$

so

$$u_1 = x - 1 = dx \text{ and } u_2 = y = dy, \tag{6.5.51}$$

and we obtain

$$\begin{aligned} df(\vec{p};\vec{u}) &= D_1 f(1,0)\cdot(x-1) + D_2 f(1,0)\cdot y \\ &= (\sin 0)\cdot(x-1) + (1\cdot\cos 0)\cdot y \\ &= y; \\ d^2 f(\vec{p};\vec{u}) &= D_{11} f(1,0)\cdot(x-1)^2 + 2D_{12} f(1,0)\cdot(x-1)y \\ &\quad + D_{22} f(1,0)\cdot y^2 \\ &= (0)\cdot(x-1)^2 + 2(\cos 0)\cdot(x-1)y - (1\cdot\sin 0)\cdot y^2 \\ &= 2(x-1)y; \end{aligned}$$

and for all $\vec{s} = (s_1, s_2) \in I$,

$$\begin{aligned} d^3 f(\vec{s};\vec{u}) = &\ D_{111} f(s_1,s_2)\cdot(x-1)^3 + 3D_{112} f(s_1,s_2)\cdot(x-1)^2 y \\ &+ 3D_{122} f(s_1,s_2)\cdot(x-1)y^2 + D_{222} f(s_1,s_2)\cdot y^3 \\ = &\ -3\sin s_2\cdot(x-1)y^2 - s_1\cos s_2\cdot y^3. \end{aligned}$$

Hence by (6) and Corollary 1 (with $m=2$), noting that $f(\vec{p}) = f(1,0) = 0$, we get

$$\begin{aligned} f(x,y) &= x\cdot\sin y \\ &= y + (x-1)y + R_2, \end{aligned}$$

where for some $\vec{s} \in I$,

$$R_2 = \frac{1}{3!} d^3 f(\vec{s};\vec{u}) = \frac{1}{6}\left[-3\sin s_2\cdot(x-1)y^2 - s_1\cos s_2\cdot y^3\right]. \tag{6.5.52}$$

As $\vec{s} \in L(\vec{p},\vec{x})$, where $\vec{p} = (1,0)$ and $\vec{x} = (x,y)$, s_1 is between 1 and x; so

$$|s_1| \le \max(|x|,1) \le |x| + 1. \tag{6.5.53}$$

Finally, since $|\sin s_2| \le 1$ and $|\cos s_2| \le 1$, we obtain

$$|R_2| \le \frac{1}{6}[3|x-1| + (|x|+1)|y|]y^2. \tag{6.5.54}$$

This bounds the maximum error that arises if we use (11) to express $x\sin y$ as a second-degree polynomial in $(x-1)$ and y. (See also Problem 4 and Note 4 below.)

Note 4. Formula (6), briefly

$$\Delta f=\sum_{i=1}^{m}\frac{d^{i}f}{i!}+R_{2}, \tag{6.5.55}$$

generalizes formula (2) in Chapter 5, §6.

As in Chapter 5, §6, we set

$$P_{m}(\vec{x})=f(\vec{p})+\sum_{i=1}^{m}\frac{1}{i!}d^{i}f(\vec{p};\vec{x}-\vec{p}) \tag{6.5.56}$$

and call P_m the m th Taylor polynomial for f about $\vec{p}$, treating it as a function of n variables x_k, with $\vec{x}=(x_1,\ldots,x_n)$.

When expanded as in Example (C), formula (6) expresses $f(\vec{x})$ in powers of

$$u_{k}=x_{k}-p_{k},\quad k=1,\ldots,n, \tag{6.5.57}$$

plus the remainder term R_m.

If $f\in CD^{\infty}$ on some $G_{\vec{p}}$ and if $R_m\rightarrow 0$ as $m\rightarrow\infty$, we can express $f(\vec{x})$ as a convergent power series

$$f(\vec{x})=\lim_{m\rightarrow\infty}P_{m}(\vec{x})=f(\vec{p})+\sum_{i=1}^{\infty}\frac{1}{i!}d^{i}f(\vec{p};\vec{x}-\vec{p}). \tag{6.5.58}$$

We then say that f admits a Taylor series about $\vec{p}$, on $G_{\vec{p}}$.

6.5.E: Problems on Repeated Differentiation and Taylor Expansions

Exercise $6.5.E.1$

Complete all details in the proof of Theorem 1. What is the motivation for introducing the auxiliary functions h_t and g_s in this particular way?

Exercise $6.5.E.2$

Is symbolic "multiplication" in Note 2 always commutative? (See Example (A).) Why was it possible to collect "similar" terms

$$\frac{\partial^2 f}{\partial x \partial y} dx dy \text{ and } \frac{\partial^2 f}{\partial y \partial x} dy dx \tag{6.5.E.1}$$

in Example (B)? Using (5), find the general formula for $d^3 f$. Expand it!

Exercise $6.5.E.3$

Carry out the induction in Theorem 2 and Corollary 2. (Use a suitable notation for subscripts: $k_1 k_2 \ldots$ instead of $jk \ldots$.)

Exercise $6.5.E.4$

Do Example(C) with $m=3$ (instead of $m=2$) and with $\vec{p}=(0,0)$. Show that $R_m \rightarrow 0$, i.e., f admits a Taylor series about $\vec{p}$.
Do it in the following two ways.
(i) Use Theorem 2.
(ii) Expand $\sin y$ as in Problem 6(a) in Chapter 5, §6, and then multiply termwise by x.
Give an estimate for R_3.

Exercise $6.5.E.5$

Use Theorem 2 to expand the following functions in powers of $x-3$ and $y+2$ exactly (choosing m so that $R_m=0$).
(i) $f(x, y)=2 x y^2-3 y^3+y x^2-x^3$;
(ii) $f(x, y)=x^4-x^3 y^2+2 x y-1$;
(iii) $f(x, y)=x^5 y-a x y^5-x^3$.

Exercise $6.5.E.6$

For the functions of Problem 15 in §4, give their Taylor expansions up to R_2, with

$$\vec{p}=\left(1, \frac{\pi}{4}, 1\right) \tag{6.5.E.2}$$

in case (i) and

$$\vec{p}=\left(e, \frac{\pi}{4} e\right) \tag{6.5.E.3}$$

in (ii). Bound R_2.

Exercise $6.5.E.7$

(Generalized Taylor theorem.) Let $\vec{u}=\vec{x}-\vec{p} \neq \overrightarrow{0}$ in $E^{\prime}(E^{\prime}$ need not be E^{n} or C^{n}); let $I=L[\vec{p}, \vec{x}]$. Prove the following statement:
If $f: E^{\prime} \rightarrow E$ and the derived functions $D_{\vec{u}}^{i} f(i \leq m)$ are relatively continuous on I and have $\vec{u}$-directed derivatives on $I-Q$ (Q countable), then formula (6) and Note 3 hold, with $d^{i} f(\vec{p} ; \vec{u})$ replaced by $D_{\vec{u}}^{i} f(\vec{p})$.
[Hint: Proceed as in Theorem 2 without using the chain rule or any partials or components. Instead of (8), prove that $h^{(i)}(t)=D_{\vec{u}}^{i} f(\vec{p}+t \vec{u})$ on $J-Q^{\prime}, Q^{\prime}=g^{-1}[Q]$.]

Exercise $6.5.E.8$

(i) Modify Problem 7 by setting

$$\vec{u}=\frac{\vec{x}-\vec{p}}{|\vec{x}-\vec{p}|}. \tag{6.5.E.4}$$

Thus expand $f(\vec{x})$ in powers of $|\vec{x}-\vec{p}|$.
(ii) Deduce Theorem 2 from Problem 7, using Corollary 2.

Exercise $6.5.E.9$

Given $f: E^{2}\left(C^{2}\right) \rightarrow E, f \in C D^{m}$ on an open set A, and $\vec{s} \in A,$ prove that $\left(\forall \vec{u} \in E^{2}\left(C^{2}\right)\right)$

$$d^{i} f(\vec{s} ; \vec{u})=\sum_{j=0}^{i}\binom{i}{j} u_{1}^{j} u_{2}^{i-j} D_{k_{1} \ldots k_{i}} f(\vec{s}), \quad 1 \leq i \leq m, \tag{6.5.E.5}$$

where the $\binom{i}{j}$ are binomial coefficients, and in the jth term,

$$k_{1}=k_{2}=\cdots=k_{j}=2 \tag{6.5.E.6}$$

and

$$k_{j+1}=\cdots=k_{i}=1. \tag{6.5.E.7}$$

Then restate formula (6) for $n=2$.
[Hint: Use induction, as in the binomial theorem.]

Exercise $6.5.E.10$

$\Rightarrow$ Given $\vec{p} \in E^{\prime}=E^{n}\left(C^{n}\right)$ and $f: E^{\prime} \rightarrow E,$ prove that $f \in C D^{1}$ at $\vec{p}$ iff f is differentiable at $\vec{p}$ and

$$(\forall \varepsilon>0)(\exists \delta>0)\left(\forall \vec{x} \in G_{\vec{p}}(\delta)\right) \quad\left\|d^{1} f(\vec{p} ; \cdot)-d^{1} f(\vec{x} ; \cdot)\right\|<\varepsilon, \tag{6.5.E.8}$$

with norm $\|\|$ as in Definition 2 in §2. (Does it apply?)
[Hint: If $f \in C D^{1}$, use Theorem 2 in §3. For the converse, verify that

$$\varepsilon \geq\left|d^{1} f(\vec{p} ; \vec{t})-d^{1} f(\vec{x} ; \vec{t})\right|=\left|\sum_{k=1}^{n}\left[D_{k} f(\vec{p})-D_{k} f(\vec{x})\right] t_{k}\right| \tag{6.5.E.9}$$

if $\vec{x} \in G_{\vec{p}}(\delta)$ and $|\vec{t}| \leq 1$. Take $\vec{t}=\vec{e}_{k}$, to prove continuity of $D_{k} f$ at $\vec{p}$.]

Exercise $6.5.E.11$

Prove the following.
(i) If $\phi : E^n \rightarrow E^m$ is linear and $[\phi] = (v_{ik})$, then

$$\|\phi\|^2 \leq \sum_{i,k} |v_{ik}|^2. \tag{6.5.E.10}$$

(ii) If $f : E^n \rightarrow E^m$ is differentiable at $\vec{p}$, then

$$\|df(\vec{p}; \cdot)\|^2 \leq \sum_{i,k} |D_k f_i(\vec{p})|^2. \tag{6.5.E.11}$$

(iii) Hence find a new converse proof in Problem 10 for $f : E^n \rightarrow E^m$.
Consider $f : C^n \rightarrow C^m$, too.
[Hints: (i) By the Cauchy-Schwarz inequality, $|\phi(\vec{x})|^2 \leq |\vec{x}|^2 \sum_{i,k} |v_{ik}|^2$. (Why?) (ii) Use part (i) and Theorem 4 in §3.]

Exercise $6.5.E.12$

(i) Find d^2u for the functions of Problem 10 in §4, in the "variable" and "mapping" notations.
(ii) Do it also for

$$u = f(x, y, z) = \left(x^2 + y^2 + z^2\right)^{-\frac{1}{2}} \tag{6.5.E.12}$$

and show that $D_{11}f + D_{22}f + D_{33}f = 0$.
(iii) Does the latter hold for $u = \arctan \frac{y}{x}$?

Exercise $6.5.E.13$

Let $u = g(x, y), x = r\cos\theta, y = r\sin\theta$ (passage to polars).
Using "variables" and then the "mappings" notation, prove that if g is differentiable, then
(i) $\frac{\partial u}{\partial r} = \cos\theta \frac{\partial u}{\partial x} + \sin\theta \frac{\partial u}{\partial y}$ and
(ii) $|\nabla g(x, y)|^2 = \left(\frac{\partial u}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial u}{\partial \theta}\right)^2$.
(iii) Assuming $g \in CD^2$, express $\frac{\partial^2 u}{\partial r \partial \theta}$, $\frac{\partial^2 u}{\partial r^2}$, and $\frac{\partial^2 u}{\partial \theta^2}$ as in (i).

Exercise $6.5.E.14$

Let $f, g : E^1 \rightarrow E^1$ be of class CD^2 on E^1. Verify (in "variable" notation, too) the following statements.
(i) $D_{11}h = a^2 D_{22}h$ if $a \in E^1$ (fixed) and

$$h(x, y) = f(ax + y) + g(y - ax). \tag{6.5.E.13}$$

(ii) $x^2 D_{11}h(x, y) + 2xy D_{12}h(x, y) + y^2 D_{22}h(x, y) = 0$ if

$$h(x, y) = xf\left(\frac{y}{x}\right) + g\left(\frac{y}{x}\right). \tag{6.5.E.14}$$

(iii) $D_1 h \cdot D_{21}h = D_2 h \cdot D_{11}h$ if

$$h(x, y) = g(f(x) + y) \tag{6.5.E.15}$$

Find $D_{12}h$, too.

Exercise 6.5.E. 15

Assume $E' = E^n\,(C^n)$ and $E'' = E^m\,(C^m)$. Let $f: E' \rightarrow E''$ and $g: E'' \rightarrow E$ be twice differentiable at $\vec{p} \in E'$ and $\vec{q} = f(\vec{p}) \in E''$, respectively, and set $h = g \circ f$.
Show that h is twice differentiable at $\vec{p}$, and

$$d^2 h(\vec{p}; \vec{t}) = d^2 g(\vec{q}; \vec{s}) + dg(\vec{q}; \vec{v}), \tag{6.5.E.16}$$

where $\vec{t} \in E', \vec{s} = df(\vec{p}; \vec{t}),$ and $\vec{v} = (v_1, \ldots, v_m) \in E''$ satisfies

$$v_i = d^2 f_i(\vec{p}; \vec{t}), \quad i = 1, \ldots, m. \tag{6.5.E.17}$$

Thus the second differential is not invariant in the sense of Note 4 in §4.
[Hint: Show that

$$D_{kl} h(\vec{p}) = \sum_{j=1}^{m} \sum_{i=1}^{m} D_{ij} g(\vec{q}) D_k f_i(\vec{p}) D_l f_j(\vec{p}) + \sum_{i=1}^{m} D_i g(\vec{q}) D_{kl} f_i(\vec{p}). \tag{6.5.E.18}$$

Proceed.]

Exercise 6.5.E. 16

Continuing Problem 15, prove the invariant rule:

$$d^r h(\vec{p}; \vec{t}) = d^r g(\vec{q}; \vec{s}), \tag{6.5.E.19}$$

if f is a first-degree polynomial and g is r times differentiable at $\vec{q}$.
[Hint: Here all higher-order partials of f vanish. Use induction.]

6.6: Determinants. Jacobians. Bijective Linear Operators

We assume the reader to be familiar with elements of linear algebra. Thus we only briefly recall some definitions and well-known rules.

Definition

Given a linear operator $\phi: E^{n} \rightarrow E^{n}\left(\text{ or } \phi: C^{n} \rightarrow C^{n}\right)$, with matrix

$$[\phi]=\left(v_{ik}\right), \quad i, k=1, \ldots, n, \tag{6.6.1}$$

we define the determinant of $[\phi]$ by

$$\begin{aligned}\operatorname{det}[\phi]=\operatorname{det}\left(v_{ik}\right) &=\left|\begin{array}{cccc}v_{11} & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} & \ldots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \ldots & v_{nn}\end{array}\right| \\ &=\sum(-1)^{\lambda} v_{1k_{1}} v_{2k_{2}} \ldots v_{nk_{n}}\end{aligned}$$

where the sum is over all ordered n-tuples $\left(k_{1}, \ldots, k_{n}\right)$ of distinct integers $k_{j}\left(1 \leq k_{j} \leq n\right)$, and

$$\lambda=\left\{\begin{array}{ll}0 & \text{if } \prod_{j<m}\left(k_{m}-k_{j}\right)>0 \text{ and} \\ 1 & \text{if } \prod_{j<m}\left(k_{m}-k_{j}\right)<0\end{array}\right. \tag{6.6.2}$$

Recall (Problem 12 in §2) that a set $B=\left\{\vec{v}_{1}, \vec{v}_{2}, \ldots, \vec{v}_{n}\right\}$ in a vector space E is a basis iff
(i) B spans E, i.e., each $\vec{v} \in E$ has the form

$$\vec{v}=\sum_{i=1}^{n} a_{i} \vec{v}_{i} \tag{6.6.3}$$

for some scalars a_{i}, and
(ii) this representation is unique.
The latter is true iff the $\vec{v}_{i}$ are independent, i.e.,

$$\sum_{i=1}^{n} a_{i} \vec{v}_{i}=\overrightarrow{0} \Longleftrightarrow a_{i}=0, i=1, \ldots, n. \tag{6.6.4}$$

If E has a basis of n vectors, we call E n-dimensional (e.g., E^{n} and C^{n}).
Determinants and bases satisfy the following rules.
(a) Multiplication rule. If $\phi, g: E^{n} \rightarrow E^{n}\left(\text{ or } C^{n} \rightarrow C^{n}\right)$ are linear, then

$$\operatorname{det}[g] \cdot \operatorname{det}[\phi]=\operatorname{det}([g][\phi])=\operatorname{det}[g \circ \phi] \tag{6.6.5}$$

(see §2, Theorem 3 and Note 4).
(b) If $\phi(\vec{x})=\vec{x}$ (identity map), then $[\phi]=\left(v_{ik}\right)$, where

$$v_{ik}=\left\{\begin{array}{ll}0 & \text{if } i \neq k \text{ and} \\ 1 & \text{if } i=k\end{array}\right. \tag{6.6.6}$$

hence det $[\phi]=1 .($ Why $?)$ See also the Problems.

(c) An n -dimensional space E is spanned by a set of n vectors iff they are independent. If so, each basis consists of exactly n vectors.

Definition

For any function $f: E^n \rightarrow E^n$ (or $f: C^n \rightarrow C^n$), we define the f-induced Jacobian map $J_f: E^n \rightarrow E^1\left(J_f: C^n \rightarrow C\right)$ by setting

$$J_f(\vec{x})=\operatorname{det}\left(v_{ik}\right), \tag{6.6.7}$$

where $v_{ik}=D_k f_i(\vec{x}), \vec{x} \in E^n\left(C^n\right)$, and $f=\left(f_1, \ldots, f_n\right)$.
The determinant

$$J_f(\vec{p})=\operatorname{det}\left(D_k f_i(\vec{p})\right) \tag{6.6.8}$$

is called the Jacobian of f at $\vec{p}$.
By our conventions, it is always defined, as are the functions $D_k f_i$.

Explicitly, $J_f(\vec{p})$ is the determinant of the right-side matrix in formula (14) in §3. Briefly,

$$J_f=\operatorname{det}\left(D_k f_i\right). \tag{6.6.9}$$

By Definition 2 and Note 2 in §5,

$$J_f(\vec{p})=\operatorname{det}\left[d^1 f(\vec{p} ; \cdot)\right]. \tag{6.6.10}$$

If f is differentiable at $\vec{p}$,

$$J_f(\vec{p})=\operatorname{det}\left[f^{\prime}(\vec{p})\right]. \tag{6.6.11}$$

Note 1. More generally, given any functions $v_{ik}: E^{\prime} \rightarrow E^1(C)$, we can define a map $f: E^{\prime} \rightarrow E^1(C)$ by

$$f(\vec{x})=\operatorname{det}\left(v_{ik}(\vec{x})\right); \tag{6.6.12}$$

briefly $f=\operatorname{det}\left(v_{ik}\right), i, k=1, \ldots, n$.

We then call f a functional determinant.

If $E^{\prime}=E^n\left(C^n\right)$ then f is a function of n variables, since $\vec{x}=\left(x_1, x_2, \ldots, x_n\right)$. If all v_{ik} are continuous or differentiable at some $\vec{p} \in E^{\prime}$, so is f; for by (1), f is a finite sum of functions of the form

$$(-1)^\lambda v_{ik_1} v_{ik_2} \ldots v_{ik_n}, \tag{6.6.13}$$

and each of these is continuous or differentiable if the v_{ik_i} are (see Problems 7 and 8 in §3).

Note 2. Hence the Jacobian map J_f is continuous or differentiable at $\vec{p}$ if all the partially derived functions $D_k f_i(i, k \leq n)$ are.

If, in addition, $J_f(\vec{p}) \neq 0$, then $J_f \neq 0$ on some globe about $\vec{p}$. (Apply Problem 7 in Chapter 4, §2, to $\left|J_f\right|$.)

In classical notation, one writes

$$\frac{\partial\left(f_1, \ldots, f_n\right)}{\partial\left(x_1, \ldots, x_n\right)} \text { or } \frac{\partial\left(y_1, \ldots, y_n\right)}{\partial\left(x_1, \ldots, x_n\right)} \tag{6.6.14}$$

for $J_f(\vec{x})$. Here $\left(y_1, \ldots, y_n\right)=f\left(x_1, \ldots, x_n\right)$.

The remarks made in §4 apply to this "variable" notation too. The chain rule easily yields the following corollary.

Corollary 6.6.1

If $f: E^n \rightarrow E^n$ and $g: E^n \rightarrow E^n$ (or $f, g: C^n \rightarrow C^n$) are differentiable at $\vec{p}$ and $\vec{q}=f(\vec{p})$, respectively, and if

$$h=g \circ f, \tag{6.6.15}$$

then

$$J_h(\vec{p}) = J_g(\vec{q}) \cdot J_f(\vec{p}) = \det(z_{ik}), \tag{6.6.16}$$

where

$$z_{ik} = D_k h_i(\vec{p}), \quad i, k = 1, \ldots, n; \tag{6.6.17}$$

or, setting

$$\begin{aligned}(u_1, \ldots, u_n) &= g(y_1, \ldots, y_n) \text{ and} \\ (y_1, \ldots, y_n) &= f(x_1, \ldots, x_n) \ (\text{"variables"}),\end{aligned}$$

we have

$$\frac{\partial(u_1, \ldots, u_n)}{\partial(x_1, \ldots, x_n)} = \frac{\partial(u_1, \ldots, u_n)}{\partial(y_1, \ldots, y_n)} \cdot \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)} = \det(z_{ik}), \tag{6.6.18}$$

where

$$z_{ik} = \frac{\partial u_i}{\partial x_k}, \quad i, k = 1, \ldots, n. \tag{6.6.19}$$

Proof

By Note 2 in §4,

$$[h'(\vec{p})] = [g'(\vec{q})] \cdot [f'(\vec{p})]. \tag{6.6.20}$$

Thus by rule (a) above,

$$\det[h'(\vec{p})] = \det[g'(\vec{q})] \cdot \det[f'(\vec{p})], \tag{6.6.21}$$

i.e.,

$$J_h(\vec{p}) = J_g(\vec{q}) \cdot J_f(\vec{p}). \tag{6.6.22}$$

Also, if $[h'(\vec{p})] = (z_{ik})$, Definition 2 yields $z_{ik} = D_k h_i(\vec{p})$.

This proves (i), hence (ii) also. $\square$

In practice, Jacobians mostly occur when a change of variables is made. For instance, in E^2, we may pass from Cartesian coordinates (x, y) to another system (u, v) such that

$$x = f_1(u, v) \text{ and } y = f_2(u, v). \tag{6.6.23}$$

We then set $f = (f_1, f_2)$ and obtain $f : E^2 \to E^2$,

$$J_f = \det(D_k f_i), \quad k, i = 1, 2. \tag{6.6.24}$$

✔ Example (passage to polar coordinates)

Let $x = f_1(r, \theta) = r\cos\theta$ and $y = f_2(r, \theta) = r\sin\theta$.

Then using the "variable" notation, we obtain $J_f(r, \theta)$ as

$$\begin{aligned}\frac{\partial(x, y)}{\partial(r, \theta)} = \left|\begin{array}{cc}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{array}\right| &= \left|\begin{array}{cc}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{array}\right| \\ &= r\cos^2\theta + r\sin^2\theta = r.\end{aligned}$$

Thus here $J_f(r, \theta) = r$ for all $r, \theta \in E^1$; J_f is independent of θ.

We now concentrate on one-to-one (invertible) functions.

Theorem 6.6.1

For a linear map $\phi : E^n \rightarrow E^n\ (\text{or} \phi : C^n \rightarrow C^n)$, the following are equivalent:

(i) ϕ is one-to-one;

(ii) the column vectors $\vec{v}_1, \ldots, \vec{v}_n$ of the matrix $[\phi]$ are independent;

(iii) ϕ is onto $E^n\ (C^n)$;

(iv) $\det[\phi] \neq 0$.

Proof

Assume (i) and let

$$\sum_{k=1}^{n} c_k \vec{v}_k = \overrightarrow{0}. \tag{6.6.25}$$

To deduce (ii), we must show that all c_k vanish.

Now, by Note 3 in §2, $\vec{v}_k = \phi\left(\vec{e}_k\right)$; so by linearity,

$$\sum_{k=1}^{n} c_k \vec{v}_k = \overrightarrow{0} \tag{6.6.26}$$

implies

$$\phi\left(\sum_{k=1}^{n} c_k \vec{e}_k\right) = \overrightarrow{0}. \tag{6.6.27}$$

As ϕ is one-to-one, it can vanish at $\overrightarrow{0}$ only. Thus

$$\sum_{k=1}^{n} c_k \vec{e}_k = \overrightarrow{0}. \tag{6.6.28}$$

Hence by Theorem 2 in Chapter 3, §§1-3, $c_k = 0, k = 1, \ldots, n,$ and (ii) follows.

Next, assume (ii); so, by rule (c) above, $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis.

Thus each $\vec{y} \in E^n\ (C^n)$ has the form

$$\vec{y} = \sum_{k=1}^{n} a_k \vec{v}_k = \sum_{k=1}^{n} a_k \phi\left(\vec{e}_k\right) = \phi\left(\sum_{k=1}^{n} a_k \vec{e}_k\right) = \phi(\vec{x}), \tag{6.6.29}$$

where

$$\vec{x} = \sum_{k=1}^{n} a_k \vec{e}_k \text{ (uniquely)}. \tag{6.6.30}$$

Hence (ii) implies both (iii) and (i). (Why?)

Now assume (iii). Then each $\vec{y} \in E^n\ (C^n)$ has the form $\vec{y} = \phi(\vec{x}),$ where

$$\vec{x} = \sum_{k=1}^{n} x_k \vec{e}_k, \tag{6.6.31}$$

by Theorem 2 in Chapter 3, §§1-3. Hence again

$$\vec{y} = \sum_{k=1}^{n} x_k \phi\left(\vec{e}_k\right) = \sum_{k=1}^{n} x_k \vec{v}_k; \tag{6.6.32}$$

so the $\vec{v}_k$ span all of $E^n\ (C^n)$. By rule (c) above, this implies (ii), hence (i), too. Thus (i), (ii), and (iii) are equivalent.

Also, by rules (a) and (b), we have

$$\det[\phi]\cdot\det\left[\phi^{-1}\right]=\det\left[\phi\circ\phi^{-1}\right]=1 \tag{6.6.33}$$

if ϕ is one-to-one (for $\phi\circ\phi^{-1}$ is the identity map). Hence $\det[\phi]\neq 0$ if (i) holds.

For the converse, suppose ϕ is not one-to-one. Then by (ii), the $\vec{v}_k$ are not independent. Thus one of them is a linear combination of the others, say,

$$\vec{v}_1=\sum_{k=2}^{n}a_k\vec{v}_k. \tag{6.6.34}$$

But by linear algebra (Problem 13(iii)), $\det[\phi]$ does not change if $\vec{v}_1$ is replaced by

$$\vec{v}_1-\sum_{k=2}^{n}a_k\vec{v}_k=\overrightarrow{0}. \tag{6.6.35}$$

Thus $\det[\phi]=0$ (one column turning to $\overrightarrow{0}$). This completes the proof. □

Note 3. Maps that are both onto and one-to-one are called bijective. Such is ϕ in Theorem 1. This means that the equation

$$\phi(\vec{x})=\vec{y} \tag{6.6.36}$$

has a unique solution

$$\vec{x}=\phi^{-1}(\vec{y}) \tag{6.6.37}$$

for each $\vec{y}$. Componentwise, by Theorem 1, the equations

$$\sum_{k=1}^{n}x_kv_{ik}=y_i,\quad i=1,\ldots,n, \tag{6.6.38}$$

have a unique solution for the x_k iff $\det(v_{ik})\neq 0$.

Corollary 6.6.2

If $\phi\in L\left(E',E\right)$ is bijective, with E' and E complete, then $\phi^{-1}\in L\left(E,E'\right)$.

Proof for $E=E^n\left(C^n\right)$

The notation $\phi\in L\left(E',E\right)$ means that $\phi:E'\rightarrow E$ is linear and continuous.

As ϕ is bijective, $\phi^{-1}:E\rightarrow E'$ is linear (Problem 12).

If $E=E^n\left(C^n\right)$, it is continuous, too (Theorem 2 in §2).

Thus $\phi^{-1}\in L\left(E,E'\right)$. □

Note. The case $E=E^n\left(C^n\right)$ suffices for an undergraduate course. (The beginner is advised to omit the "starred" §8.) Corollary 2 and Theorem 2 below, however, are valid in the general case. So is Theorem 1 in §7.

Theorem 6.6.2

Let E, E' and ϕ be as in Corollary 2. Set

$$\left\|\phi^{-1}\right\|=\frac{1}{\varepsilon}. \tag{6.6.39}$$

Then any map $\theta\in L\left(E',E\right)$ with $\|\theta-\phi\|<\varepsilon$ is one-to-one, and θ^{-1} is uniformly continuous.

Proof

Proof. By Corollary 2, $\phi^{-1}\in L\left(E,E'\right)$, so $\left\|\phi^{-1}\right\|$ is defined and >0 (for ϕ^{-1} is not the zero map, being one-to-one).

Thus we may set

$$\varepsilon=\frac{1}{\left\|\phi^{-1}\right\|}, \quad\left\|\phi^{-1}\right\|=\frac{1}{\varepsilon}. \tag{6.6.40}$$

Clearly $\vec{x}=\phi^{-1}(\vec{y})$ if $\vec{y}=\phi(\vec{x})$. Also,

$$\left|\phi^{-1}(\vec{y})\right| \leq \frac{1}{\varepsilon}|\vec{y}| \tag{6.6.41}$$

by Note 5 in §2, Hence

$$|\vec{y}| \geq \varepsilon\left|\phi^{-1}(\vec{y})\right|, \tag{6.6.42}$$

i.e.,

$$|\phi(\vec{x})| \geq \varepsilon|\vec{x}| \tag{6.6.43}$$

for all $\vec{x} \in E^{\prime}$ and $\vec{y} \in E$.

Now suppose $\phi \in L\left(E^{\prime}, E\right)$ and $\|\theta-\phi\|=\sigma<\varepsilon$.

Obviously, $\theta=\phi-(\phi-\theta)$, and by Note 5 in §2,

$$|(\phi-\theta)(\vec{x})| \leq\|\phi-\theta\||\vec{x}|=\sigma|\vec{x}|. \tag{6.6.44}$$

Thus for every $\vec{x} \in E^{\prime}$,

$$\begin{aligned} |\theta(\vec{x})| & \geq|\phi(\vec{x})|-|(\phi-\theta)(\vec{x})| \\ & \geq|\phi(\vec{x})|-\sigma|\vec{x}| \\ & \geq(\varepsilon-\sigma)|\vec{x}| \end{aligned}$$

by (2). Therefore, given $\vec{p} \neq \vec{r}$ in $E^{\prime}$ and setting $\vec{x}=\vec{p}-\vec{r} \neq \overrightarrow{0}$, we obtain

$$|\theta(\vec{p})-\theta(\vec{r})|=|\theta(\vec{p}-\vec{r})|=|\theta(\vec{x})| \geq(\varepsilon-\sigma)|\vec{x}|>0 \tag{6.6.45}$$

(since $\sigma<\varepsilon$) .

We see that $\vec{p} \neq \vec{r}$ implies $\theta(\vec{p}) \neq \theta(\vec{r})$; so θ is one-to-one, indeed.

Also, setting $\theta(\vec{x})=\vec{z}$ and $\vec{x}=\theta^{-1}(\vec{z})$ in (3), we get

$$|\vec{z}| \geq(\varepsilon-\sigma)\left|\theta^{-1}(\vec{z})\right|; \tag{6.6.46}$$

that is,

$$\left|\theta^{-1}(\vec{z})\right| \leq(\varepsilon-\sigma)^{-1}|\vec{z}| \tag{6.6.47}$$

for all $\vec{z}$ in the range of θ (domain of θ^{-1}).

Thus θ^{-1} is linearly bounded (by Theorem 1 in §2), hence uniformly continuous, as claimed. □

Corollary 6.6.3

If $E^{\prime}=E=E^{n}\left(C^{n}\right)$ in Theorem 2 above, then for given ϕ and $\delta>0$, there always is $\delta^{\prime}>0$ such that

$$\|\theta-\phi\|<\delta^{\prime} \text{ implies }\left\|\theta^{-1}-\phi^{-1}\right\|<\delta. \tag{6.6.48}$$

In other words, the transformation $\phi \rightarrow \phi^{-1}$ is continuous on $L(E), E=E^{n}\left(C^{n}\right)$.

Proof

First, since $E^{\prime}=E=E^{n}\left(C^{n}\right), \theta$ is bijective by Theorem 1(iii), so $\theta^{-1} \in L(E)$.

As before, set $\|\theta-\phi\|=\sigma<\varepsilon$.

By Note 5 in §2, formula (5) above implies that

$$\left\|\theta^{-1}\right\| \leq \frac{1}{\varepsilon-\sigma}. \tag{6.6.49}$$

Also,

$$\phi^{-1} \circ(\theta-\phi) \circ \theta^{-1}=\phi^{-1}-\theta^{-1} \tag{6.6.50}$$

(see Problem 11).

Hence by Corollary 4 in §2, recalling that $\left\|\phi^{-1}\right\|=1/\varepsilon$, we get

$$\left\|\theta^{-1}-\phi^{-1}\right\| \leq\left\|\phi^{-1}\right\| \cdot\|\theta-\phi\| \cdot\left\|\theta^{-1}\right\| \leq \frac{\sigma}{\varepsilon(\varepsilon-\sigma)} \rightarrow 0 \text{ as } \sigma \rightarrow 0. \quad \square \tag{6.6.51}$$

6.6.E: Problems on Bijective Linear Maps and Jacobians

Exercise $6.6.E.1$

(i) Can a functional determinant $f=\det(v_{ik})$ (see Note 1) be continuous or differentiable even if the functions v_{ik} are not?
(ii) Must a Jacobian map J_f be continuous or differentiable if f is?
Give proofs or counterexamples.

Exercise $6.6.E.2$

$\Rightarrow$ Prove rule (b) on determinants. More generally, show that if $f(\vec{x})=\vec{x}$ on an open set $A \subseteq E^n\left(C^n\right)$, then $J_f=1$ on A.

Exercise $6.6.E.3$

Let $f: E^n \rightarrow E^n$ (or $C^n \rightarrow C^n$), $f=\left(f_1, \ldots, f_n\right)$.
Suppose each f_k depends on x_k only, i.e.,

$$f_k(\vec{x})=f_k(\vec{y}) \text{ if } x_k=y_k, \tag{6.6.E.1}$$

regardless of the other coordinates x_i, y_i. Prove that $J_f=\prod_{k=1}^n D_k f_k$.
[Hint: Show that $D_k f_i=0$ if $i \neq k$.]

Exercise $6.6.E.4$

In Corollary 1, show that

$$J_h(\vec{p})=\prod_{k=1}^n D_k f_k(\vec{p}) \cdot J_g(\vec{q}) \tag{6.6.E.2}$$

if f also has the property specified in Problem 3. Then do all in "variables," with $y_k=y_k\left(x_k\right)$ instead of f_k.

Exercise $6.6.E.5$

Let $E^{\prime}=E^1$ in Note 1. Prove that if all the v_{ik} are differentiable at p, then $f^{\prime}(p)$ is the sum of n determinants, each arising from det $\left(v_{ik}\right)$, by replacing the terms of one column by their derivatives.
[Hint: Use Problem 6 in Chapter 5, §1.]

Exercise $6.6.E.6$

Do Problem 5 for partials of f, with $E^{\prime}=E^n\left(C^n\right)$, and for directionals $D_{\vec{u}} f$, in any normed space $E^{\prime}$. (First, prove formulas analogous to Problem 6 in Chapter 5, §1; use Note 3 in §1.) Finally, do it for the differential, $d f(\vec{p} ; \cdot)$.

Exercise $6.6.E.7$

In Note 1 of §4, express the matrices in terms of partials (see Theorem 4 in §3). Invent a "variable" notation for such matrices, imitating Jacobians (Corollary 3).

Exercise $6.6.E.8$

(i) Show that

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \alpha)}=-r^2 \sin \alpha \tag{6.6.E.3}$$

if

$$\begin{aligned} &x=r \cos \theta, \\ &y=r \sin \theta \sin \alpha, \text{ and} \\ &z=r \cos \alpha \end{aligned} \tag{6.6.E.4}$$

(This transformation is passage to polars in E^3; see Figure 27, where $r=OP, \angle XOA=\theta,$ and $\angle AOP=\alpha.$)
(ii) What if $x=r \cos \theta, y=r \sin \theta,$ and $z=z$ remains unchanged (passage to cylindric coordinates)?
(iii) Same for $x=e^r \cos \theta, y=e^r \sin \theta,$ and $z=z$.

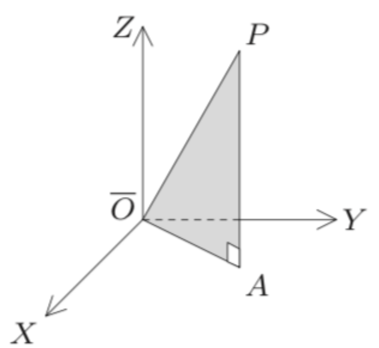

FIGURE 27

Exercise 6.6.E.9

Is $f=(f_1, f_2): E^2 \rightarrow E^2$ one-to-one or bijective, and is $J_f \neq 0,$ if
(i) $f_1(x, y)=e^x \cos y$ and $f_2(x, y)=e^x \sin y$;
(ii) $f_1(x, y)=x^2-y^2$ and $f_2(x, y)=2xy$?

Exercise 6.6.E.10

Define $f: E^3 \rightarrow E^3$ (or $C^3 \rightarrow C^3$)

$$f(\vec{x})=\frac{\vec{x}}{1+\sum_{k=1}^3 x_k} \tag{6.6.E.5}$$

on

$$A=\left\{\vec{x} \mid \sum_{k=1}^3 x_k \neq-1\right\} \tag{6.6.E.6}$$

and $f=\overrightarrow{0}$ on $-A$. Prove the following.
(i) f is one-to-one on A (find f^{-1}!).
(ii) $J_f(\vec{x})=\frac{1}{\left(1+\sum_{k=1}^3 x_k\right)^4}$.
(iii) Describe $-A$ geometrically.

Exercise $6.6.E.11$

Given any sets A, B and maps $f, g: A \rightarrow E', h: E' \rightarrow E,$ and $k: B \rightarrow A,$ prove that
(i) $(f \pm g) \circ k = f \circ k \pm g \circ k,$ and
(ii) $h \circ (f \pm g) = h \circ f \pm h \circ g$ if h is linear.
Use these distributive laws to verify that

$$\phi^{-1} \circ (\theta - \phi) \circ \theta^{-1} = \phi^{-1} - \theta^{-1} \tag{6.6.E.7}$$

In Corollary 3.
[Hint: First verify the associativity of mapping composition.]

Exercise $6.6.E.12$

Prove that if $\phi: E' \rightarrow E$ is linear and one-to-one, so is $\phi^{-1}: E'' \rightarrow E',$ where $E'' = \phi[E']$.

Exercise $6.6.E.13$

Let $\vec{v}_1, \ldots, \vec{v}_n$ be the column vectors in $\det[\phi]$. Prove that $\det[\phi]$ turns into
(i) $c \cdot \det[\phi]$ if one of the $\vec{v}_k$ is multiplied by a scalar c;
(ii) $-\det[\phi]$, if any two of the $\vec{v}_k$ are interchanged (consider λ in formula (1)).
Furthermore, show that
(iii) $\det[\phi]$ does not change if some $\vec{v}_k$ is replaced by $\vec{v}_k + c\vec{v}_i (i \neq k)$;
(iv) $\det[\phi] = 0$ if some $\vec{v}_k$ is $\overrightarrow{0}$, or if two of the $\vec{v}_k$ are the same.

6.7: Inverse and Implicit Functions. Open and Closed Maps

I. "If $f \in CD^1$ at $\vec{p}$, then f resembles a linear map (namely df) at $\vec{p}$." Pursuing this basic idea, we first make precise our notion of "$f \in CD^1$ at $\vec{p}$."

Definition 1

A map $f : E' \rightarrow E$ is continuously differentiable, or of class CD^1 (written $f \in CD^1$), at $\vec{p}$ iff the following statement is true:

$$\begin{array}{l}\text{Given any } \varepsilon>0, \text{ there is } \delta>0 \text{ such that } f \text{ is differentiable on the} \\ \text{globe } \overline{G}=\overline{G_{\vec{p}}(\delta)}, \text{ with} \\ \qquad \|df(\vec{x};\cdot)-df(\vec{p};\cdot)\|<\varepsilon \text{ for all } \vec{x} \in \overline{G}.\end{array} \tag{6.7.1}$$

By Problem 10 in §5, this definition agrees with Definition 1 §5, but is no longer limited to the case $E'=E^n\left(C^n\right)$. See also Problems 1 and 2 below.

We now obtain the following result.

Theorem 6.7.1

Let E' and E be complete. If $f : E' \rightarrow E$ is of class CD^1 at $\vec{p}$ and if $df(\vec{p};\cdot)$ is bijective (§6), then f is one-to-one on some globe $\overline{G}=\overline{G_{\vec{p}}}(\delta)$.

Thus f "locally" resembles df $(\vec{p};\cdot)$ in this respect.

Proof

Set $\phi=df(\vec{p};\cdot)$ and

$$\left\|\phi^{-1}\right\|=\frac{1}{\varepsilon} \tag{6.7.2}$$

(cf. Theorem 2 of §6).

By Definition 1, fix $\delta>0$ so that for $\vec{x} \in \overline{G}=\overline{G_{\vec{p}}(\delta)}$.

$$\|df(\vec{x};\cdot)-\phi\|<\frac{1}{2}\varepsilon. \tag{6.7.3}$$

Then by Note 5 in §2,

$$(\forall \vec{x} \in \overline{G})\left(\forall \vec{u} \in E'\right) \quad |df(\vec{x};\vec{u})-\phi(\vec{u})| \leq \frac{1}{2}\varepsilon|\vec{u}|. \tag{6.7.4}$$

Now fix any $\vec{r}, \vec{s} \in \overline{G}, \vec{r} \neq \vec{s}$, and set $\vec{u}=\vec{r}-\vec{s} \neq 0$. Again, by Note 5 in §2,

$$|\vec{u}|=\left|\phi^{-1}(\phi(\vec{u}))\right| \leq\left\|\phi^{-1}\right\||\phi(\vec{u})|=\frac{1}{\varepsilon}|\phi(\vec{u})|; \tag{6.7.5}$$

so

$$0<\varepsilon|\vec{u}| \leq|\phi(\vec{u})|. \tag{6.7.6}$$

By convexity, $\overline{G} \supseteq I=L[\vec{s}, \vec{r}]$, so (1) holds for $\vec{x} \in I, \vec{x}=\vec{s}+t\vec{u}, 0 \leq t \leq 1$.

Noting this, set

$$h(t)=f(\vec{s}+t\vec{u})-t\phi(\vec{u}), \quad t \in E^1. \tag{6.7.7}$$

Then for $0 \leq t \leq 1$,

$$\begin{aligned} h'(t) &=D_{\vec{u}} f(\vec{s}+t\vec{u})-\phi(\vec{u}) \\ &=df(\vec{s}+t\vec{u};\vec{u})-\phi(\vec{u}). \end{aligned}$$

(Verify!) Thus by (1) and (2),

$$\begin{aligned}\sup_{0 \leq t \leq 1}\left|h^{\prime}(t)\right| &= \sup_{0 \leq t \leq 1}|d f(\vec{s}+t \vec{u} ; \vec{u})-\phi(\vec{u})| \\ &\leq \frac{\varepsilon}{2}|\vec{u}| \leq \frac{1}{2}|\phi(\vec{u})|.\end{aligned}$$

(Explain!) Now, by Corollary 1 in Chapter 5, §4,

$$|h(1)-h(0)| \leq(1-0) \cdot \sup_{0 \leq t \leq 1}\left|h^{\prime}(t)\right| \leq \frac{1}{2}|\phi(\vec{u})|. \tag{6.7.8}$$

As $h(0)=f(\vec{s})$ and

$$h(1)=f(\vec{s}+\vec{u})-\phi(\vec{u})=f(\vec{r})-\phi(\vec{u}), \tag{6.7.9}$$

we obtain (even if $\vec{r}=\vec{s}$)

$$|f(\vec{r})-f(\vec{s})-\phi(\vec{u})| \leq \frac{1}{2}|\phi(\vec{u})| \quad(\vec{r}, \vec{s} \in \overline{G}, \vec{u}=\vec{r}-\vec{s}). \tag{6.7.10}$$

But by the triangle law,

$$|\phi(\vec{u})|-|f(\vec{r})-f(\vec{s})| \leq|f(\vec{r})-f(\vec{s})-\phi(\vec{u})|. \tag{6.7.11}$$

Thus

$$|f(\vec{r})-f(\vec{s})| \geq \frac{1}{2}|\phi(\vec{u})| \geq \frac{1}{2} \varepsilon|\vec{u}|=\frac{1}{2} \varepsilon|\vec{r}-\vec{s}| \tag{6.7.12}$$

by (2).

Hence $f(\vec{r}) \neq f(\vec{s})$ whenever $\vec{r} \neq \vec{s}$ in $\overline{G}$; so f is one-to-one on $\overline{G}$, as claimed. □

Corollary 6.7.1

Under the assumptions of Theorem 1, the maps f and f^{-1} (the inverse of f restricted to $\overline{G}$) are uniformly continuous on $\overline{G}$ and $f[\overline{G}]$, respectively.

Proof

By (3),

$$\begin{aligned}|f(\vec{r})-f(\vec{s})| &\leq|\phi(\vec{u})|+\frac{1}{2}|\phi(\vec{u})| \\ &\leq|2 \phi(\vec{u})| \\ &\leq 2\|\phi\||\vec{u}| \\ &=2\|\phi\||\vec{r}-\vec{s}| \quad(\vec{r}, \vec{s} \in \overline{G}).\end{aligned}$$

This implies uniform continuity for f. (Why?)

Next, let $g=f^{-1}$ on $H=f[\overline{G}]$.

If $\vec{x}, \vec{y} \in H$, let $\vec{r}=g(\vec{x})$ and $\vec{s}=g(\vec{y})$; so $\vec{r}, \vec{s} \in \overline{G}$, with $\vec{x}=f(\vec{r})$ and $\vec{y}=f(\vec{s})$. Hence by (4),

$$|\vec{x}-\vec{y}| \geq \frac{1}{2} \varepsilon|g(\vec{x})-g(\vec{y})|, \tag{6.7.13}$$

proving all for g, too. □

Again, f resembles ϕ which is uniformly continuous, along with ϕ^{-1}.

II. We introduce the following definition.

Definition 2

A map $f:(S,\rho)\to(T,\rho')$ is closed (open) on $D\subseteq S$ iff, for any $X\subseteq D$ the set $f[X]$ is closed (open) in T whenever X is so in S.

Note that continuous maps have such a property for inverse images (Problem 15 in Chapter 4, §2).

Corollary 6.7.2

Under the assumptions of Theorem 1, f is closed on $\overline{G}$, and so the set $f[\overline{G}]$ is closed in E.

Similarly for the map f^{-1} on $f[\overline{G}]$.

Proof for $E'=E=E^n\ (C^n)$ (for the general case, see Problem 6)

Given any closed $X\subseteq\overline{G}$, we must show that $f[X]$ is closed in E.

Now, as $\overline{G}$ is closed and bounded, it is compact (Theorem 4 of Chapter 4, §6).

So also is X (Theorem 1 in Chapter 4, §6), and so is $f[X]$ (Theorem 1 of Chapter 4, §8).

By Theorem 2 in Chapter 4, §6, $f[X]$ is closed, as required. $\square$

For the rest of this section, we shall set $E'=E=E^n\ (C^n)$.

Theorem 6.7.2

If $E'=E=E^n\ (C^n)$ in Theorem 1, with other assumptions unchanged, then f is open on the globe $G=G_{\vec{p}}(\delta)$, with δ sufficiently small.

Proof

We first prove the following lemma.

Lemma

$f[G]$ contains a globe $G_{\vec{q}}(\alpha)$ where $\vec{q}=f(\vec{p})$.

Proof

Indeed, let

$$\alpha=\frac{1}{4}\varepsilon\delta, \tag{6.7.14}$$

where δ and ε are as in the proof of Theorem 1. (We continue the notation and formulas of that proof.)

Fix any $\vec{c}\in G_{\vec{q}}(\alpha)$; so

$$|\vec{c}-\vec{q}|<\alpha=\frac{1}{4}\varepsilon\delta. \tag{6.7.15}$$

Set $h=|f-\vec{c}|$ on E'. As f is uniformly continuous on $\overline{G}$, so is h.

Now, $\overline{G}$ is compact in $E^n\ (C^n)$; so Theorem 2(ii) in Chapter 4, §8, yields a point $\vec{r}\in\overline{G}$ such that

$$h(\vec{r})=\min h[\overline{G}]. \tag{6.7.16}$$

We claim that $\vec{r}$ is in G (the interior of $\overline{G}$).

Otherwise, $|\vec{r}-\vec{p}|=\delta$; for by (4),

$$2\alpha=\frac{1}{2}\varepsilon\delta=\frac{1}{2}\varepsilon|\vec{r}-\vec{p}| \leq|f(\vec{r})-f(\vec{p})| \leq|f(\vec{r})-\vec{c}|+|\vec{c}-f(\vec{p})| =h(\vec{r})+h(\vec{p}).$$

But

$$h(\vec{p})=|\vec{c}-f(\vec{p})|=|\vec{c}-\vec{q}|<\alpha; \tag{6.7.17}$$

and so (7) yields

$$h(\vec{p})<\alpha<h(\vec{r}), \tag{6.7.18}$$

contrary to the minimality of $h(\vec{r})$ (see (6)). Thus $|\vec{r}-\vec{p}|$ cannot equal δ.

We obtain $|\vec{r}-\vec{p}|<\delta,$ so $\vec{r}\in G_{\vec{p}}(\delta)=G$ and $f(\vec{r})\in f[G]$. We shall now show that $\vec{c}=f(\vec{r})$.

To this end, we set $\vec{v}=\vec{c}-f(\vec{r})$ and prove that $\vec{v}=\overrightarrow{0}$. Let

$$\vec{u}=\phi^{-1}(\vec{v}), \tag{6.7.19}$$

where

$$\phi=df(\vec{p};\cdot), \tag{6.7.20}$$

as before. Then

$$\vec{v}=\phi(\vec{u})=df(\vec{p};\vec{u}). \tag{6.7.21}$$

With $\vec{r}$ as above, fix some

$$\vec{s}=\vec{r}+t\vec{u} \quad (0<t<1) \tag{6.7.22}$$

with t so small that $\vec{s}\in G$ also. Then by formula (3),

$$|f(\vec{s})-f(\vec{r})-\phi(t\vec{u})|\leq\frac{1}{2}|t\vec{v}|; \tag{6.7.23}$$

also,

$$|f(\vec{r})-\vec{c}+\phi(t\vec{u})|=(1-t)|\vec{v}|=(1-t)h(\vec{r}) \tag{6.7.24}$$

by our choice of $\vec{v},\vec{u}$ and h. Hence by the triangle law,

$$h(\vec{s})=|f(\vec{s})-\vec{c}|\leq\left(1-\frac{1}{2}t\right)h(\vec{r}). \tag{6.7.25}$$

(Verify!)

As $0<t<1,$ this implies $h(\vec{r})=0$ (otherwise, $h(\vec{s})<h(\vec{r}),$ violating (6)).

Thus, indeed,

$$|\vec{v}|=|f(\vec{r})-\vec{c}|=0, \tag{6.7.26}$$

i.e.,

$$\vec{c}=f(\vec{r})\in f[G] \quad \text{for } \vec{r}\in G. \tag{6.7.27}$$

But $\vec{c}$ was an arbitrary point of $G_{\vec{q}}(\alpha)$. Hence

$$G_{\vec{q}}(\alpha)\subseteq f[G], \tag{6.7.28}$$

proving the lemma. □

Proof of Theorem 2. The lemma shows that $f(\vec{p})$ is in the interior of $f[G]$ if $\vec{p}, f, df(\vec{p};\cdot)$, and δ are as in Theorem 1.

But Definition 1 implies that here $f \in CD^1$ on all of G (see Problem 1).

Also, $df(\vec{x};\cdot)$ is bijective for any $\vec{x} \in G$ by our choice of G and Theorems 1 and 2 in §6.

Thus f maps all $\vec{x} \in G$ onto interior points of $f[G]$; i.e., f maps any open set $X \subseteq G$ onto an open $f[X]$, as required. □

Note 1. A map

$$f:(S,\rho) \underset{\text{onto}}{\longleftrightarrow} (T,\rho') \tag{6.7.29}$$

is both open and closed ("clopen") iff f^{-1} is continuous - see Problem 15(iv)(v) in Chapter 4, §2, interchanging f and f^{-1}.

Thus $\phi = df(\vec{p};\cdot)$ in Theorem 1 is "clopen" on all of E'.

Again, f locally resembles $df(\vec{p};\cdot)$.

III. The Inverse Function Theorem. We now further pursue these ideas.

Theorem 6.7.3 (inverse functions)

Under the assumptions of Theorem 2, let g be the inverse of f_G $\left(f \text{ restricted to } G = G_{\vec{p}}(\delta)\right)$.

Then $g \in CD^1$ on $f[G]$ and $dg(\vec{y};\cdot)$ is the inverse of $df(\vec{x};\cdot)$ whenever $\vec{x} = g(\vec{y}), \vec{x} \in G$.

Briefly: "The differential of the inverse is the inverse of the differential."

Proof

Fix any $\vec{y} \in f[G]$ and $\vec{x} = g(\vec{y});$ so $\vec{y} = f(\vec{x})$ and $\vec{x} \in G.$ Let $U = df(\vec{x};\cdot)$.

As noted above, U is bijective for every $\vec{x} \in G$ by Theorems 1 and 2 in §6; so we may set $V = U^{-1}.$ We must show that $V = dg(\vec{y};\cdot)$.

To do this, give $\vec{y}$ an arbitrary (variable) increment $\Delta\vec{y}$, so small that $\vec{y} + \Delta\vec{y}$ stays in $f[G]$ (an open set by Theorem 2).

As g and f_G are one-to-one, $\Delta\vec{y}$ uniquely determines

$$\Delta\vec{x} = g(\vec{y} + \Delta\vec{y}) - g(\vec{y}) = \vec{t}, \tag{6.7.30}$$

and vice versa:

$$\Delta\vec{y} = f(\vec{x} + \vec{t}) - f(\vec{x}). \tag{6.7.31}$$

Here $\Delta\vec{y}$ and $\vec{t}$ are the mutually corresponding increments of $\vec{y} = f(\vec{x})$ and $\vec{x} = g(\vec{y}).$ By continuity, $\vec{y} \rightarrow \overrightarrow{0}$ iff $\vec{t} \rightarrow \overrightarrow{0}$.

As $U = df(\vec{x};\cdot)$,

$$\lim_{\vec{t} \rightarrow \vec{0}} \frac{1}{|\vec{t}|}|f(\vec{x} + \vec{t}) - f(\vec{t}) - U(\vec{t})| = 0, \tag{6.7.32}$$

or

$$\lim_{\vec{t} \rightarrow \overrightarrow{0}} \frac{1}{|\vec{t}|}|F(\vec{t})| = 0, \tag{6.7.33}$$

where

$$F(\vec{t}) = f(\vec{x} + \vec{t}) - f(\vec{t}) - U(\vec{t}). \tag{6.7.34}$$

As $V = U^{-1},$ we have

$$V(U(\vec{t})) = \vec{t} = g(\vec{y} + \Delta\vec{y}) - g(\vec{y}). \tag{6.7.35}$$

So from (9),

$$\begin{aligned} V(F(\vec{t})) &= V(\Delta\vec{y}) - \vec{t} \\ &= V(\Delta\vec{y}) - [g(\vec{y} + \Delta\vec{y}) - g(\vec{y})]; \end{aligned}$$

that is,

$$\frac{1}{|\Delta\vec{y}|}|g(\vec{y}+\Delta\vec{y})-g(\vec{y})-V(\Delta\vec{y})|=\frac{|V(F(\vec{t}))|}{|\Delta\vec{y}|}, \quad \Delta\vec{y}\neq\overrightarrow{0}. \tag{6.7.36}$$

Now, formula (4), with $\vec{r}=\vec{x}, \vec{s}=\vec{x}+\vec{t},$ and $\vec{u}=\vec{t}$, shows that

$$|f(\vec{x}+\vec{t})-f(\vec{x})|\geq\frac{1}{2}\varepsilon|\vec{t}|; \tag{6.7.37}$$

i.e., $|\Delta\vec{y}|\geq\frac{1}{2}\varepsilon|\vec{t}|$. Hence by (8),

$$\frac{|V(F(\vec{t}))|}{|\Delta\vec{y}|}\leq\frac{|V(F(\vec{t})|}{\frac{1}{2}\varepsilon|\vec{t}|}=\frac{2}{\varepsilon}\left|V\left(\frac{1}{|\vec{t}|}F(\vec{t})\right)\right|\leq\frac{2}{\varepsilon}\|V\|\frac{1}{|\vec{t}|}|F(\vec{t})|\rightarrow 0 \text{ as } \vec{t}\rightarrow\overrightarrow{0}. \tag{6.7.38}$$

Since $\vec{t}\rightarrow\overrightarrow{0}$ as $\Delta\vec{y}\rightarrow\overrightarrow{0}$ (change of variables!), the expression (10) tends to 0 as $\Delta\vec{y}\rightarrow\overrightarrow{0}$.

By definition, then, g is differentiable at $\vec{y}$, with $dg(\vec{y};)=V=U^{-1}$.

Moreover, Corollary 3 in §6, applies here. Thus

$$(\forall\delta'>0)\,(\exists\delta''>0) \quad \|U-W\|<\delta''\Rightarrow\left\|U^{-1}-W^{-1}\right\|<\delta'. \tag{6.7.39}$$

Taking here $U^{-1}=dg(\vec{y})$ and $W^{-1}=dg(\vec{y}+\Delta\vec{y})$, we see that $g\in CD^{1}$ near $\vec{y}$. This completes the proof. □

Note 2. If $E'=E=E^{n}\left(C^{n}\right)$, the bijectivity of $\phi=df(\vec{p};\cdot)$ is equivalent to

$$\operatorname{det}[\phi]=\operatorname{det}[f'(\vec{p})]\neq 0 \tag{6.7.40}$$

(Theorem 1 of §6).

In this case, the fact that f is one-to-one on $G=G_{\vec{p}}(\delta)$ means, componentwise (see Note 3 in §6), that the system of n equations

$$f_{i}(\vec{x})=f\left(x_{1},\ldots,x_{n}\right)=y_{i}, \quad i=1,\ldots,n, \tag{6.7.41}$$

has a unique solution for the n unknowns x_{k} as long as

$$(y_{1},\ldots,y_{n})=\vec{y}\in f[G]. \tag{6.7.42}$$

Theorem 3 shows that this solution has the form

$$x_{k}=g_{k}(\vec{y}), \quad k=1,\ldots,n, \tag{6.7.43}$$

where the g_{k} are of class CD^{1} on $f[G]$ provided the f_{i} are of class CD^{1} near $\vec{p}$ and det $\left[f'(\vec{p})\right]\neq 0$. Here

$$\operatorname{det}[f'(\vec{p})]=J_{f}(\vec{p}), \tag{6.7.44}$$

as in §6.

Thus again f "locally" resembles a linear map, $\phi=df(\vec{p};\cdot)$.

IV. The Implicit Function Theorem. Generalizing, we now ask, what about solving n equations in $n+m$ unknowns $x_{1},\ldots,x_{n},y_{1},\ldots,y_{m}$? Say, we want to solve

$$f_{k}\left(x_{1},\ldots,x_{n},y_{1},\ldots,y_{m}\right)=0, \quad k=1,2,\ldots,n, \tag{6.7.45}$$

for the first n unknowns (or variables) x_{k}, thus expressing them as

$$x_{k}=H_{k}\left(y_{1},\ldots,y_{m}\right), \quad k=1,\ldots,n, \tag{6.7.46}$$

with $H_{k}:E^{m}\rightarrow E^{1}$ or $H_{k}:C^{m}\rightarrow C$.

Let us set $\vec{x}=\left(x_{1},\ldots,x_{n}\right), \vec{y}=\left(y_{1},\ldots,y_{m}\right)$, and

$$(\vec{x},\vec{y})=\left(x_{1},\ldots,x_{n},y_{1},\ldots,y_{m}\right) \tag{6.7.47}$$

so that $(\vec{x}, \vec{y}) \in E^{n+m}\left(C^{n+m}\right)$.

Thus the system of equations (11) simplifies to

$$f_k(\vec{x}, \vec{y}) = 0, \quad k = 1, \ldots, n \tag{6.7.48}$$

or

$$f(\vec{x}, \vec{y}) = \overrightarrow{0}, \tag{6.7.49}$$

where $f = (f_1, \ldots, f_n)$ is a map of $E^{n+m}\left(C^{n+m}\right)$ into $E^n\left(C^n\right)$; f is a function of $n+m$ variables, but it has n components f_k; i.e.,

$$f(\vec{x}, \vec{y}) = f\left(x_1, \ldots, x_n, y_1, \ldots, y_m\right) \tag{6.7.50}$$

is a vector in $E^n\left(C^n\right)$.

Theorem 6.7.4 (implicit functions)

Let $E' = E^{n+m}\left(C^{n+m}\right), E = E^n\left(C^n\right)$, and let $f : E' \rightarrow E$ be of class CD^1 near

$$(\vec{p}, \vec{q}) = \left(p_1, \ldots, p_n, q_1, \ldots, q_m\right), \quad \vec{p} \in E^n\left(C^n\right), \vec{q} \in E^m\left(C^m\right). \tag{6.7.51}$$

Let $[\phi]$ be the $n \times n$ matrix

$$\left(D_j f_k(\vec{p}, \vec{q})\right), \quad j, k = 1, \ldots, n. \tag{6.7.52}$$

If $\operatorname{det}[\phi] \neq 0$ and if $f(\vec{p}, \vec{q}) = \overrightarrow{0}$, then there are open sets

$$P \subseteq E^n\left(C^n\right) \text{ and } Q \subseteq E^m\left(C^m\right), \tag{6.7.53}$$

with $\vec{p} \in P$ and $\vec{q} \in Q$, for which there is a unique map

$$H : Q \rightarrow P \tag{6.7.54}$$

with

$$f(H(\vec{y}), \vec{y}) = \overrightarrow{0} \tag{6.7.55}$$

for all $\vec{y} \in Q$; furthermore, $H \in CD^1$ on Q.

Thus $\vec{x} = H(\vec{y})$ is a solution of (11) in vector form.

Proof

With the above notation, set

$$F(\vec{x}, \vec{y}) = (f(\vec{x}, \vec{y}), \vec{y}), \quad F : E' \rightarrow E'. \tag{6.7.56}$$

Then

$$F(\vec{p}, \vec{q}) = (f(\vec{p}, \vec{q}), \vec{q}) = (\overrightarrow{0}, \vec{q}), \tag{6.7.57}$$

since $f(\vec{p}, \vec{q}) = \overrightarrow{0}$.

As $f \in CD^1$ near $(\vec{p}, \vec{q})$, so is F (verify componentwise via Problem 9(ii) in §3 and Definition 1 of §5).

By Theorem 4, §3, $\operatorname{det}\left[F'(\vec{p}, \vec{q})\right] = \operatorname{det}[\phi] \neq 0$ (explain!).

Thus Theorem 1 above shows that F is one-to-one on some globe G about $(\vec{p}, \vec{q})$.

Clearly G contains an open interval about $(\vec{p}, \vec{q})$. We denote it by $P \times Q$ where $\vec{p} \in P, \vec{q} \in Q; P$ is open in $E^n\left(C^n\right)$ and Q is open in $E^m\left(C^m\right)$.

By Theorem 3, $F_{P \times Q}$ $(F$ restricted to $P \times Q)$ has an inverse

$$g : A \underset{\text{onto}}{\longleftrightarrow} P \times Q, \tag{6.7.58}$$

where $A=F[P\times Q]$ is open in E' (Theorem 2), and $g\in CD^1$ on A. Let the map $u=(g_1,\ldots,g_n)$ comprise the first n components of g (exactly as f comprises the first n components of F).

Then

$$g(\vec{x},\vec{y})=(u(\vec{x},\vec{y}),\vec{y}) \tag{6.7.59}$$

exactly as $F(\vec{x},\vec{y})=(f(\vec{x},\vec{y}),\vec{y})$. Also, $u:A\rightarrow P$ is of class CD^1 on A, as g is (explain!).

Now set

$$H(\vec{y})=u(\overrightarrow{0},\vec{y}); \tag{6.7.60}$$

here $\vec{y}\in Q$, while

$$(\overrightarrow{0},\vec{y})\in A=F[P\times Q], \tag{6.7.61}$$

for F preserves $\vec{y}$ (the last m coordinates). Also set

$$\alpha(\vec{x},\vec{y})=\vec{x}. \tag{6.7.62}$$

Then $f=\alpha\circ F$ (why?), and

$$f(H(\vec{y}),\vec{y})=f(u(\overrightarrow{0},\vec{y}),\vec{y})=f(g(\overrightarrow{0},\vec{y}))=\alpha(F(g(\overrightarrow{0},\vec{y}))=\alpha(\overrightarrow{0},\vec{y})=\overrightarrow{0} \tag{6.7.63}$$

by our choice of α and g (inverse to F). Thus

$$f(H(\vec{y}),\vec{y})=\overrightarrow{0},\quad \vec{y}\in Q, \tag{6.7.64}$$

as desired.

Moreover, as $H(\vec{y})=u(\overrightarrow{0},\vec{y}),$ we have

$$\frac{\partial}{\partial y_i}H(\vec{y})=\frac{\partial}{\partial y_i}u(\overrightarrow{0},\vec{y}),\quad \vec{y}\in Q, i\leq m. \tag{6.7.65}$$

As $u\in CD^1$, all $\partial u/\partial y_i$ are continuous (Definition 1 in §5); hence so are the $\partial H/\partial y_i$. Thus by Theorem 3 in §3, $H\in CD^1$ on Q.

Finally, H is unique for the given P,Q; for

$$\begin{aligned} f(\vec{x},\vec{y})=\overrightarrow{0} &\Longrightarrow (f(\vec{x},\vec{y}),\vec{y})=(\overrightarrow{0},\vec{y}) \\ &\Longrightarrow F(\vec{x},\vec{y})=(\overrightarrow{0},\vec{y}) \\ &\Longrightarrow g(F(\vec{x},\vec{y}))=g(\overrightarrow{0},\vec{y}) \\ &\Longrightarrow (\vec{x},\vec{y})=g(\overrightarrow{0},\vec{y})=(u(\overrightarrow{0},\vec{y}),\vec{y}) \\ &\Longrightarrow \vec{x}=u(\overrightarrow{0},\vec{y})=H(\vec{y}). \end{aligned}$$

Thus $f(\vec{x},\vec{y})=\overrightarrow{0}$ implies $\vec{x}=H(\vec{y});$ so $H(\vec{y})$ is the only solution for $\vec{x}$. $\square$

Note 3. H is said to be implicitly defined by the equation $f(\vec{x},\vec{y})=\overrightarrow{0}$. In this sense we say that $H(\vec{y})$ is an implicit function, given by $f(\vec{x},\vec{y})=\overrightarrow{0}$.

Similarly, under suitable assumptions, $f(\vec{x},\vec{y})=\overrightarrow{0}$ defines $\vec{y}$ as a function of $\vec{x}$.

Note 4. While H is unique for a given neighborhood $P\times Q$ of $(\vec{p},\vec{q})$, another implicit function may result if $P\times Q$ or $(\vec{p},\vec{q})$ is changed.

For example, let

$$f(x, y)=x^{2}+y^{2}-25 \tag{6.7.66}$$

(a polynomial; hence $f \in CD^1$ on all of E^2). Geometrically, $x^2+y^2-25=0$ describes a circle.

Solving for x, we get $x=\pm\sqrt{25-y^2}$. Thus we have two functions:

$$H_{1}(y)=+\sqrt{25-y^{2}} \tag{6.7.67}$$

and

$$H_{2}(y)=-\sqrt{25-y^{2}}. \tag{6.7.68}$$

If $P \times Q$ is in the upper part of the circle, the resulting function is H_1. Otherwise, it is H_2. See Figure 28.

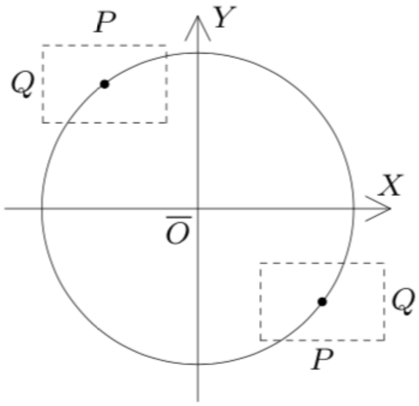

FIGURE 28

V. Implicit Differentiation. Theorem 4 only states the existence (and uniqueness) of a solution, but does not show how to find it, in general.

The knowledge itself that $H \in CD^1$ exists, however, enables us to use its derivative or partials and compute it by implicit differentiation, known from calculus.

Examples

(a) Let $f(x, y)=x^2+y^2-25=0,$ as above.

This time treating y as an implicit function of $x, y=H(x)$, and writing y' for $H'(x)$, we differentiate both sides of (x^{2}+y^{2}-25=0\) with respect to x, using the chain rule for the term $y^2=[H(x)]^2$.

This yields $2x+2yy'=0$, whence $y'=-x/y$.

Actually (see Note 4), two functions are involved: $y=\pm\sqrt{25-x^2}$; but both satisfy $x^2+y^2-25=0$; so the result $y'=-x/y$ applies to both.

Of course, this method is possible only if the derivative y' is known to exist. This is why Theorem 4 is important.

(b) Let

$$f(x, y, z)=x^{2}+y^{2}+z^{2}-1=0, \quad x, y, z \in E^{1}. \tag{6.7.69}$$

Again f satisfies Theorem 4 for suitable $x, y,$ and z.

Setting $z=H(x, y)$, differentiate the equation $f(x, y, z)=0$ partially with respect to x and y. From the resulting two equations, obtain $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

6.7.E: Problems on Inverse and Implicit Functions, Open and Closed Maps

Exercise $6.7.E.1$

Discuss: In Definition 1, $\overline{G}$ can equivalently be replaced by $G=G_{\vec{p}}(\delta)$ (an open globe).

Exercise $6.7.E.2$

Prove that if the set D is open (closed) in (S,ρ), then the map $f:S\rightarrow T$ is open (closed, respectively) on D iff $f_D(f$ restricted to $D)$ has this property as a map of D into $f[D]$.
[Hint: Use Theorem 4 in Chapter 3, §12.]

Exercise $6.7.E.3$

Complete the missing details in the proofs of Theorems 1-4.

Exercise $6.7.E.3'$

Verify footnotes 2 and 3.

Exercise $6.7.E.4$

Show that a map $f:E'\rightarrow E$ may fail to be one-to-one on all of E' even if f satisfies Theorem 1 near every $\vec{p}\in E'$. Nonetheless, show that this cannot occur if $E'=E=E^1$.
[Hints: For the first part, take $E'=C, f(x+iy)=e^x(\cos y+i\sin y)$. For the second, use Theorem 1 in Chapter 5, §2.]

Exercise $6.7.E.4'$

(i) For maps $f:E^1\rightarrow E^1$, prove that the existence of a bijective $df(p;\cdot)$ is equivalent to $f'(p)\neq 0$.
(ii) Let

$$f(x)=x+x^2\sin\frac{1}{x},\quad f(0)=0. \tag{6.7.E.1}$$

Show that $f'(0)\neq 0$, and $f\in CD^1$ near any $p\neq 0$; yet f is not one-to-one near 0. What is wrong?

Exercise $6.7.E.5$

Show that a map $f:E^n\left(C^n\right)\rightarrow E^n\left(C^n\right), f\in CD^1,$ may be bijective even if $\det[f'(\vec{p})]=0$ at some $\vec{p}$, but then f^{-1} cannot be differentiable at $\vec{q}=f(\vec{p})$.
[Hint: For the first clause, take $f(x)=x^3, p=0;$ for the second, note that if f^{-1} is differentiable at $\vec{q}$, then Note 2 in §4 implies that det $[df(\vec{p};\cdot)]\cdot\det\left[df^{-1}(\vec{q};\cdot)\right]=1\neq 0,$ since $f\circ f^{-1}$ is the identity map.]

Exercise $6.7.E.6$

Prove Corollary 2 for the general case of complete E' and E.
[Outline: Given a closed $X\subseteq\overline{G},$ take any convergent sequence $\{\vec{y}_n\}\subseteq f[X]$. By Problem 8 in Chapter 4, §8, $f^{-1}\left(\vec{y}_n\right)=\vec{x}_n$ is a Cauchy sequence in X (why?). By the completeness of E', $(\exists\vec{x}\in X)\vec{x}_n\rightarrow\vec{x}$ (Theorem 4 of Chapter 3, §16). Infer that $\lim\vec{y}_n=f(\vec{x})\in f[X]$, so $f[X]$ is closed.]

Exercise $6.7.E.7$

Prove that "the composite of two open (closed) maps is open (closed)." State the theorem precisely. Prove it also for the uniform Lipschitz property.

Exercise $6.7.E.8$

Prove in detail that $f:(S, \rho) \rightarrow (T, \rho')$ is open on $D \subseteq S$ iff f maps the interior of D into that of $f[D]$; that is, $f\left[D^{0}\right] \subseteq (f[D])^{0}$.

Exercise $6.7.E.9$

Verify by examples that f may be:
(i) closed but not open;
(ii) open but not closed.
[Hints: (i) Consider $f=$ constant. (ii) Define $f: E^{2} \rightarrow E^{1}$ by $f(x, y)=x$ and let

$$D=\left\{(x, y) \in E^{2} | y=\frac{1}{x}, x>0\right\}; \tag{6.7.E.2}$$

use Theorem 4(iii) in Chapter 3, §16 and continuity to show that D is closed in E^{2}, but $f[D]=(0,+\infty)$ is not closed in E^{1}. However, f is open on all of E^{2} by Problem 8. (Verify!)]

Exercise $6.7.E.10$

Continuing Problem 9(ii), define $f: E^{n} \rightarrow E^{1}$ (or $C^{n} \rightarrow C$) by $f(\vec{x})=x_{k}$ for a fixed $k \leq n$ (the "kth projection map"). Show that f is open, but not closed, on $E^{n}\left(C^{n}\right)$.

Exercise $6.7.E.11$

(i) In Example (a), take $(p, q)=(5,0)$ or $(-5,0)$. Are the conditions of Theorem 4 satisfied? Do the conclusions hold?
(ii) Verify Example (b).

Exercise $6.7.E.12$

(i) Treating z as a function of x and y, given implicitly by

$$f(x, y, z)=z^{3}+x z^{2}-y z=0, \quad f: E^{3} \rightarrow E^{1}, \tag{6.7.E.3}$$

discuss the choices of P and Q that satisfy Theorem 4. Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.
(ii) Do the same for $f(x, y, z)=e^{x y z}-1=0$.

Exercise $6.7.E.13$

Given $f: E^{n}\left(C^{n}\right) \rightarrow E^{m}\left(C^{m}\right), n>m,$ prove that if $f \in C D^{1}$ on a globe G, f cannot be one-to-one.
[Hint for $f: E^{2} \rightarrow E^{1}:$ If, say, $D_{1} f \neq 0$ on G, set $F(x, y)=(f(x, y), y)$.]

Exercise $6.7.E.14$

Suppose that f satisfies Theorem 1 for every $\vec{p}$ in an open set $A \subseteq E'$, and is one-to-one on A (cf. Problem 4). Let $g=f_{A}^{-1}$ (restrict f to A and take its inverse). Show that f and g are open and of class $C D^{1}$ on A and $f[A]$, respectively.

Exercise 6.7.E.15

Given $\vec{v} \in E$ and a scalar $c \neq 0$, define $T_{\vec{v}} : E \rightarrow E$ ("translation by $\vec{v}$") and $M_c : E \rightarrow E$ ("dilation by c"), by setting

$$T_{\vec{v}}(\vec{x}) = \vec{x} + \vec{v} \text{ and } M_c(\vec{x}) = c\vec{x}. \tag{6.7.E.4}$$

Prove the following.
(i) $T_{\vec{v}}$ and $T_{\vec{v}}^{-1} \, (= T_{-\vec{v}})$ are bijective, continuous, and "clopen" on E; so also are M_c and $M_c^{-1} \left(= M_{1/c}\right)$.
(ii) Similarly for the Lipschitz property on E.
(iii) If $G = G_{\vec{q}}(\delta) \subset E,$ then $T_{\vec{v}}[G] = G_{\vec{q}+\vec{v}}(\delta)$, and $M_c[G] = G_{c\vec{q}}(|c\delta|)$.
(iv) If $f : E' \rightarrow E$ is linear, and $\vec{v} = f(\vec{p})$ for some $\vec{p} \in E'$, then $T_{\vec{v}} \circ f = f \circ T'_{\vec{p}}$ and $M_c \circ f = f \circ M'_c$, where $T'_{\vec{p}}$ and M'_c are the corresponding maps on E'. If, further, f is continuous at $\vec{p}$, it is continuous on all of E'.
[Hint for (iv): Fix any $\vec{x} \in E'$. Set $\vec{v} = f(\vec{x} - \vec{p}), g = T_{\vec{v}} \circ f \circ T'_{\vec{p}-\vec{x}}$. Verify that $g = f, T'_{\vec{p}-\vec{x}}(\vec{x}) = \vec{p}$, and g is continuous at $\vec{x}$.]

Exercise 6.7.E.16

Show that if $f : E' \rightarrow E$ is linear and if $f\left[G^*\right]$ is open in E for some $G^* = G_{\vec{p}}(\delta) \subseteq E'$, then
(i) f is open on all of E';
(ii) f is onto E.
[Hints: (i) By Problem 8, it suffices to show that the set $f[G]$ is open, for any globe G (why?). First take $G = G_{\vec{0}}(\delta)$. Then use Problems 7 and 15(i)-(iv), with suitable $\vec{v}$ and c.
(ii) To prove $E = f\left[E'\right]$, fix any $\vec{y} \in E$. As \(f=G_{\overrightarrow{0} (\delta\) is open, it contains a globe $G' = G_{\vec{0}}(r)$. For small $c, c\vec{y} \in G' \subseteq f\left[E'\right]$. Hence $\vec{y} \in f\left[E'\right]$ (Problem 10 in §2).]

Exercise 6.7.E.17

Continuing Problem 16, show that if f is also one-to-one on G^*, then

$$f : E' \underset{\text{onto}}{\longleftrightarrow} E, \tag{6.7.E.5}$$

$f \in L\left(E', E\right), f^{-1} \in L\left(E, E'\right), f$ is clopen on E', and f^{-1} is so on E.
[Hints: To prove that f is one-to-one on E', let $f(\vec{x}) = \vec{y}$ for some $\vec{x}, \vec{x}' \in E'$. Show that

$$(\exists c, \varepsilon > 0) \quad c\vec{y} \in G_{\vec{0}}(\varepsilon) \subseteq f\left[G_{\vec{0}}(\delta)\right] \text{ and } f(c\vec{x} + \vec{p}) = f\left(c\vec{x}' + \vec{p}\right) \in f\left[G_{\vec{p}}(\delta)\right] = f\left[G^*\right]. \tag{6.7.E.6}$$

Deduce that $c\vec{x} + \vec{p} = c\vec{x}' + \vec{p}$ and $\vec{x} = \vec{x}'$. Then use Problem 15(v) in Chapter 4, §2, and Note 1.]

Exercise 6.7.E.18

A map

$$f : (S, \rho) \underset{\text{onto}}{\longleftrightarrow} (T, \rho') \tag{6.7.E.7}$$

is said to be bicontinuous, or a homeomorphism, (from S onto T) iff both f and f^{-1} are continuous. Assuming this, prove the following.
(i) $x_n \rightarrow p$ in S iff $f\left(x_n\right) \rightarrow f(p)$ in T;
(ii) A is closed (open, compact, perfect) in S iff $f[A]$ is so in T;
(iii) $B = \overline{A}$ in S iff $f[B] = \overline{f[A]}$ in T;
(iv) $B = A^0$ in S iff $f[B] = (f[A])^0$ in T;

(v) A is dense in B (i.e., $A \subseteq B \subseteq \overline{A} \subseteq S$) in (S, ρ) iff $f[A]$ is dense in $f[B] \subseteq (T, \rho')$.
[Hint: Use Theorem 1 of Chapter 4, §2, and Theorem 4 in Chapter 3, §16, for closed sets; see also Note 1.]

Exercise $6.7.E.19$

Given $A, B \subseteq E, \vec{v} \in E$ and a scalar c, set

$$A+\vec{v}=\{\vec{x}+\vec{v} \mid \vec{x} \in A\} \text{ and } cA=\{c\vec{x} \mid \vec{x} \in A\}. \tag{6.7.E.8}$$

Assuming $c \neq 0$, prove that
(i) A is closed (open, compact, perfect) in E iff $cA+\vec{v}$ is;
(ii) $B=\overline{A}$ iff $cB+\vec{v}=\overline{cA+\vec{v}}$;
(iii) $B=A^0$ iff $cB+\vec{v}=(cA+\vec{v})^0$;
(iv) A is dense in B iff $cA+\vec{v}$ is dense in $cB+\vec{v}$.
[Hint: Apply Problem 18 to the maps $T_{\vec{v}}$ and M_c of Problem 15, noting that $A+\vec{v}=T_{\vec{v}}[A]$ and $cA=M_c[A]$.]

Exercise $6.7.E.20$

Prove Theorem 2, for a reduced δ, assuming that only one of E' and $E$is $E^n\left(C^n\right)$, and the other is just complete.
[Hint: If, say, $E=E^n\left(C^n\right)$, then $f[\overline{G}]$ is compact (being closed and bounded), and so is $\overline{G}=f^{-1}[f[\overline{G}]]$. (Why?) Thus the Lemma works out as before, i.e., $f[G] \supseteq G_{\vec{q}}(\alpha)$.
Now use the continuity of f to obtain a globe $G'=G_{\vec{p}}\left(\delta'\right) \subseteq G$ such that $f\left[G'\right] \subseteq G_{\vec{q}}(\alpha)$. Let $g=f_G^{-1}$, further restricted to $G_{\vec{q}}(\alpha)$. Apply Problem 15(v) in Chapter 4, §2, to g, with $S=G_{\vec{q}}(\alpha), T=E'$.]

6.8: Baire Categories. More on Linear Maps

We pause to outline the theory of so-called sets of Category I or Category II, as introduced by Baire. It is one of the most powerful tools in higher analysis. Below, (S, ρ) is a metric space.

Definition 1

A set $A \subseteq (S, \rho)$ is said to be nowhere dense (in S) iff its closure $\overline{A}$ has no interior points (i.e., contains no globes): $(\overline{A})^0 = \emptyset$.

Equivalently, the set A is nowhere dense iff every open set $G^* \neq \emptyset$ in S contains a globe $\overline{G}$ disjoint from A. (Why?)

Definition 2

A set $A \subseteq (S, \rho)$ is meagre, or of Category I (in S), iff

$$A=\bigcup_{n=1}^{\infty} A_n, \tag{6.8.1}$$

for some sequence of nowhere dense sets A_n.

Otherwise, A is said to be nonmeagre or of Category II.

A is residual iff $-A$ is meagre, but A is not.

Examples

(a) $\emptyset$ is nowhere dense.

(b) Any finite set in a normed space E is nowhere dense.

(c) The set N of all naturals in E^1 is nowhere dense.

(d) So also is Cantor's set P (Problem 17 in Chapter 3, §14); indeed, P is closed $(P=\overline{P})$ and has no interior points (verify!), so $(\overline{P})^0 = P^0 = \emptyset$.

(e) The set R of all rationals in E^1 is meagre; for it is countable (see Chapter 1, §9), hence a countable union of nowhere dense singletons $\{r_n\}$, $r_n \in R$. But R is not nowhere dense; it is even dense in E^1, since $\overline{R} = E^1$ (see Definition 2, in Chapter 3, §14). Thus a meagre set need not be nowhere dense. (But all nowhere dense sets are meagre why?)

Examples (c) and (d) show that a nowhere dense set may be infinite (even uncountable). Yet, sometimes nowhere dense sets are treated as "small" or "negligible," in comparison with other sets. Most important is the following theorem.

Theorem 6.8.1 (Baire)

In a complete metric space (S, ρ), every open set $G^* \neq \emptyset$ is nonmeagre. Hence the entire space S is residual.

Proof

Seeking a contradiction, suppose G^* is meagre, i.e.,

$$G^*=\bigcup_{n=1}^{\infty} A_n \tag{6.8.2}$$

for some nowhere dense sets A_n. Now, as A_1 is nowhere dense, G^* contains a closed globe

$$\overline{G}_1=\overline{G_{x_1}(\delta_1)} \subseteq -A_1. \tag{6.8.3}$$

Again, as A_2 is nowhere dense, G_1 contains a globe

$$\overline{G}_2=\overline{G_{x_2}(\delta_2)} \subseteq -A_2, \quad \text{with } 0<\delta_2 \leq \frac{1}{2}\delta_1. \tag{6.8.4}$$

By induction, we obtain a contracting sequence of closed globes

$$\overline{G}_n=\overline{G_{x_n}\left(\delta_n\right)}, \quad \text{with } 0<\delta_n \leq \frac{1}{2^n} \delta_1 \rightarrow 0. \tag{6.8.5}$$

As S is complete, so are the $\overline{G}_n$ (Theorem 5 in Chapter 3, §17). Thus, by Cantor's theorem (Theorem 5 of Chapter 4, §6), there is

$$p \in \bigcap_{n=1}^{\infty} \overline{G}_n. \tag{6.8.6}$$

As $G^* \supseteq \overline{G}_n,$ we have $p \in G^*$. But, as $\overline{G}_n \subseteq -A_n,$ we also have $(\forall n) p \notin A_n$; hence

$$p \notin \bigcup_{n=1}^{\infty} A_n=G^* \tag{6.8.7}$$

(the desired contradiction!). □

We shall need a lemma based on Problems 15 and 19 in §7. (Review them!)

lemma

Let $f \in L\left(E^{\prime}, E\right), E^{\prime}$ complete. Let $G=G_0(1)$ be the unit globe in $E^{\prime}$. If $\overline{f[G]}$ (closure of $f[G]$ in E) contains a globe $G_0=G_0(r) \subset E,$ then $G_0 \subseteq f[G]$.

Note. Recall that we "arrow" only vectors from $E^{\prime}$ (e.g., $\overrightarrow{0}$), but not those from E (e.g., 0).

Proof

Let $A=f[G] \cap G_0 \subseteq G_0$. We claim that A is dense in G_0; i.e., $G_0 \subseteq \overline{A}$. Indeed, by assumption, any $q \in G_0$ is in $f[G]$. Thus by Theorem 3 in Chapter 3, §16, any G_q meets $f[G] \cap G_0=A$ if $q \in G_0$. Hence

$$\left(\forall q \in G_0\right) \quad q \in \overline{A}, \tag{6.8.8}$$

i.e., $G_0 \subseteq \overline{A},$ as claimed.

Now fix any $q_0 \in G_0=G_0(r)$ and a real $c(0<c<1)$. As A is dense in G_0,

$$A \cap G_{q_0}(c r) \neq \emptyset; \tag{6.8.9}$$

so let $q_1 \in A \cap G_{q_0}(c r) \subseteq f[G]$. Then

$$\left|q_1-q_0\right|<c r, \quad q_0 \in G_{q_1}(c r). \tag{6.8.10}$$

As $q_1 \in f[G]$, we can fix some $\vec{p}_1 \in G=G_0(1),$ with $f\left(\vec{p}_1\right)=q_1$. Also, by Problems 19(iv) and 15(iii) in §7, $c A+q_1$ is dense in $c G_0+q_1=G_{q_1}(c r)$. But $q_0 \in G_{q_1}(c r)$. Thus

$$G_{q_0}\left(c^2 r\right) \cap\left(c A+q_1\right) \neq \emptyset; \tag{6.8.11}$$

so let $q_2 \in G_{q_0}\left(c^2 r\right) \cap\left(c A+q_1\right),$ so $q_0 \in G_{q_2}\left(c^2 r\right),$ etc.

Inductively, we fix for each $n>1$ some $q_n \in G_{q_0}\left(c^n r\right),$ with

$$q_n \in c^{n-1} A+q_{n-1}, \tag{6.8.12}$$

i.e.,

$$q_n-q_{n-1} \in c^{n-1} A. \tag{6.8.13}$$

As $A \subseteq f\left[G_0(1)\right]$, linearity yields

$$q_n-q_{n-1} \in f\left[c^{n-1} G_0(1)\right]=f\left[G_0\left(c^{n-1}\right)\right], \quad n>1. \tag{6.8.14}$$

Thus for each $n>1$, there is $\vec{p}_n \in G_0(c^{n-1})$, (i.e., $|\vec{p}_n|<c^{n-1}$) such that $f(\vec{p}_n)=q_n-q_{n-1}$. Now, as $|\vec{p}_n|<c^{n-1}$ and $0<c<1$,

$$\sum_1^\infty |\vec{p}_n|<+\infty; \tag{6.8.15}$$

so by the completeness of E', $\sum \vec{p}_n$ converges in E' (Theorem 1 in Chapter 4, §13). Let $\vec{p}=\sum_{k=1}^\infty \vec{p}_k$; then

$$\begin{aligned} f(\vec{p}) &= f\left(\lim_{n\to\infty}\sum_{k=1}^n \vec{p}_k\right)=\lim_{n\to\infty} f\left(\sum_{k=1}^n \vec{p}_k\right) \\ &= \lim_{n\to\infty}\sum_{k=1}^n f\left(\vec{p}_k\right) \text{ for } f\in L\left(E',E\right). \end{aligned}$$

But $f(\vec{p}_k)=q_k-q_{k-1}(k>1)$, and $f(\vec{p}_1)=q_1$; so

$$\sum_{k=1}^n f\left(\vec{p}_k\right)=q_1+\sum_{k=2}^n\left(q_k-q_{k-1}\right)=q_n. \tag{6.8.16}$$

Thus

$$f(\vec{p})=\lim_{n\to\infty}\sum_{k=1}^n f\left(\vec{p}_k\right)=\lim_{n\to\infty} q_n=q_0. \tag{6.8.17}$$

Moreover, $|\vec{p}_k|<c^{k-1}(k\geq 1)$. Thus

$$|\vec{p}|\leq\sum_{k=1}^\infty|\vec{p}_k|<\sum_{k=1}^\infty c^{k-1}=\frac{1}{1-c}; \tag{6.8.18}$$

i.e.,

$$\vec{p}\in G_{\vec{0}}\left(\frac{1}{1-c}\right). \tag{6.8.19}$$

But $q_0=f(\vec{p})$; so

$$q_0\in f\left[G_{\vec{0}}\left(\frac{1}{1-c}\right)\right]. \tag{6.8.20}$$

As $q_0\in G_0(r)$ was arbitrary, we have

$$G_0(r)\subseteq f\left[G_0\left(\frac{1}{1-c}\right)\right], \tag{6.8.21}$$

or by linearity,

$$G_0(r(1-c))\subseteq f\left[G_0(1)\right]=f[G]. \tag{6.8.22}$$

This holds for any $c\in(0,1)$. Hence

$$f[G]\supseteq\bigcup_{0<c<1} G_0(r(1-c))=G_0(r). \quad \text{(Verify!)} \tag{6.8.23}$$

Thus all is proved. $\square$

We can now establish an important result due to S. Banach.

Theorem 6.8.2 (Banach)

Let $f \in L(E', E)$, with E' complete. Then $f[E']$ is meagre in E or $f[E'] = E$, according to whether $f[G_0(1)]$ is or is not nowhere dense.

Proof

If $f[G_0(1)]$ is nowhere dense in E, so also is $f[G_0(n)], n > 0$. (Verify by Problems 15 and 19 in §7.) But then

$$f[E'] = f\left[\bigcup_{n=1}^{\infty} G_0(n)\right] = \bigcup_{n=1}^{\infty} f\left[G_{\vec{0}}(n)\right] \tag{6.8.24}$$

is a countable union of nowhere dense sets, hence meagre, by definition.

Now suppose $f[G_0(1)]$ is not nowhere dense; so $\overline{f[G_0(1)]}$ contains some $G_q(r) \subseteq E$. We may assume $q \in f\left[G_{\vec{0}}(1)\right]$ (if not, replace q by a close point from $f[G_0(1)]$). Then $q = f(\vec{p})$ for some $\vec{p} \in G_0(1)$. The latter implies

$$|-\vec{p}| = |\vec{p}| = \rho(\vec{p}, \vec{0}) < 1; \tag{6.8.25}$$

so

$$G_{-\vec{p}}(1) \subseteq G_{\vec{0}}(2). \tag{6.8.26}$$

Also, as $\overline{f\left[G_{\bar{0}}(1)\right]} \supseteq G_q(r)$, translation by $-q = f(-\vec{p})$ yields

$$\overline{f\left[G_{\bar{0}}(1)\right]} + f(-\vec{p}) \supseteq G_q(r) - q = G_0(r), \tag{6.8.27}$$

i.e.,

$$G_0(r) \subseteq \overline{f\left[G_{-\vec{p}}(1)\right]} \subseteq \overline{f\left[G_{\vec{0}}(2)\right]}. \tag{6.8.28}$$

Hence $\overline{f\left[G_{\vec{0}}(1)\right]} \supseteq G_0\left(\frac{1}{2}r\right)$ (why?); so, by the Lemma

$$f\left[G_{\vec{0}}(1)\right] \supseteq G_0\left(\frac{1}{2}r\right) \text{ in } E. \tag{6.8.29}$$

This implies $f\left[G_{\vec{0}}(2n)\right] \supseteq G_0(nr)$, and so

$$f[E'] \supseteq \bigcup_{n=1}^{\infty} G_0(nr) = E, \tag{6.8.30}$$

i.e., $f[E'] = E$, as required. Thus the theorem is proved. $\square$

Theorem 6.8.3 (Open map principle)

Let $f \in L(E', E)$, with E' and E complete. Then the map f is open on E' iff $f[E'] = E$, i.e., iff f is onto E.

Proof

If $f[E'] = E$, then by Theorem 1, $f[E']$ is nonmeagre in E, as is E itself. Thus by Theorem 2, $f\left[G_{\vec{0}}(1)\right]$ is not nowhere dense, and (1) follows as before. Hence by Problems 15(iii) and 19 in §7, $f\left[G_{\vec{p}}\right] \supseteq$ some G_q whenever $q = f(\vec{p})$. (Why?) Therefore, $G_{\vec{p}} \subseteq A \subseteq E'$ implies

$$G_{f(\vec{p})} \subseteq f\left[G_{\vec{p}}\right] \subseteq f[A]; \tag{6.8.31}$$

i.e., f maps any interior point $\vec{p} \in A$ into such a point of $f[A]$. By Problem 8 in §7, f is open on E'.

Conversely, if so, then $f\left[E^{\prime}\right]$ is an open set $\neq \emptyset$ in E, a complete space; so by Theorems 1 and 2, $f\left[E^{\prime}\right]$ is nonmeagre and equals E. (See also Problem 16(ii) in §7.) □

Note 1. Theorem 3 holds even if f is not one-to-one.

Note 2. If in Theorem 3, however, f is bijective, it is open on $E^{\prime}$, and so $f^{-1} \in L\left(E, E^{\prime}\right)$ by Note 1 in §7. (This is the promised general proof of Corollary 2 in §6.)

Theorem 6.8.4 (Banach-Steinhaus uniform boundedness principle)

Let $E^{\prime} be$ complete. Let $\mathcal{N}$ be a family of maps $f \in L\left(E^{\prime}, E\right)$ such that

$$\left(\forall x \in E^{\prime}\right)\left(\exists k \in E^{1}\right)(\forall f \in \mathcal{N}) \quad|f(\vec{x})|<k. \tag{6.8.32}$$

("$\mathcal{N}$ is bounded at each $\vec{x}$.")

Then $\mathcal{N}$ is "norm-bounded," i.e.,

$$\left(\exists K \in E^{1}\right)(\forall f \in \mathcal{N}) \quad\|f\|<K, \tag{6.8.33}$$

with $\|\quad\|$ as in §2.

Proof

It suffices to show that $\mathcal{N}$ is "uniformly" bounded on some globe,

$$\left(\exists c \in E^{1}\right)\left(\exists G=G_{\vec{p}}(r)\right)(\forall f \in \mathcal{N})(\forall \vec{x} \in G) \quad|f(\vec{x})| \leq c. \tag{6.8.34}$$

For then $|\vec{x}-\vec{p}| \leq r$ implies

$$2 c>|f(\vec{x})-f(\vec{p})|=|f(\vec{x}-\vec{p})|, \tag{6.8.35}$$

or (setting $\vec{x}-\vec{p}=r \vec{y})|\vec{y}|<1$ implies

$$(\forall f \in \mathcal{N}) \quad|f(\vec{y})|<\frac{2 c}{r} \quad \text{(why?)}; \tag{6.8.36}$$

so

$$(\forall f \in \mathcal{N}) \quad\|f\|=\sup _{|\vec{y}| \leq 1}|f(\vec{y})|<\frac{2 c}{r}. \tag{6.8.37}$$

Thus, seeking a contradiction, suppose (3) fails and assume its negation:

$$\left(\forall c \in E^{1}\right)\left(\forall G=G_{\vec{p}}(r)\right)(\exists f \in \mathcal{N})\left(\exists \vec{x} \in G=G_{\vec{p}}(r)\right) \quad|f(\vec{x})|>c. \tag{6.8.38}$$

Then for $c=1$, we can fix some $f_{1} \in \mathcal{N}$ and $G_{\vec{x}_{1}}\left(r_{1}\right)$ such that $0<r_{1}<1$ and

$$\left|f_{1}\left(\vec{x}_{1}\right)\right|>1. \tag{6.8.39}$$

By the continuity of the norm $||$, we can choose r_{1} so small that

$$\left(\forall \vec{x} \in \overline{G_{\vec{x}_{1}}\left(r_{1}\right)}\right) \quad|f(\vec{x})|>1. \tag{6.8.40}$$

Again by (4), we fix $f_{2} \in \mathcal{N}$ and $\vec{x}_{2} \in G_{\vec{x}_{1}}\left(r_{1}\right)$ such that $\left|f_{2}\right|>2$ on some globe

$$\overline{G_{\vec{x}_{2}}\left(r_{2}\right)} \subseteq G_{\vec{x}_{1}}\left(r_{1}\right), \tag{6.8.41}$$

with $0<r_{2}<1 / 2.$ Inductively, we thus form a contracting sequence of closed globes

$$\overline{G_{\vec{x}_{n}}\left(r_{n}\right)}, \quad 0<r_{n}<\frac{1}{n}, \tag{6.8.42}$$

and a sequence $\left\{f_{n}\right\} \subseteq \mathcal{N}$, such that

$$(\forall n) \quad |f_n| > n \text{ on } \overline{G_{\vec{x}_n}(r_n)} \subseteq E'. \tag{6.8.43}$$

As E' is complete, so are the closed globes $\overline{G_{\vec{x}_n}(r_n)} \subseteq E'$. Also, $0 < r_n < 1/n \rightarrow 0$. Thus by Cantor's theorem (Theorem 5 of Chapter 4, §6), there is

$$\vec{x}_0 \in \bigcap_{n=1}^{\infty} \overline{G_{\vec{x}_n}(r_n)}. \tag{6.8.44}$$

As $\vec{x}_0$ is in each $\overline{G_{\vec{x}_n}(r_n)}$, we have

$$(\forall n) \quad |f_n(\vec{x}_0)| > n; \tag{6.8.45}$$

so $\mathcal{N}$ is not bounded at $\vec{x}_0$, contrary to (2). This contradiction completes the proof. □

Note 3. Complete normed spaces are also called Banach spaces.

6.8.E: Problems on Baire Categories and Linear Maps

Exercise $6.8.E.1$

Verify the equivalence of the various formulations in Definition 1. Discuss: A is nowhere dense iff it is not dense in any open set $\neq \emptyset$.

Exercise $6.8.E.2$

Verify Examples (a) to (e). Show that Cantor's set P is uncountable.
[Hint: Each $p \in P$ corresponds to a "ternary fraction," $p=\sum_{n=1}^{\infty} x_n / 3^n$, also written $0. x_1, x_2, \ldots, x_n, \ldots$, where $x_n=0$ or $x_n=2$ according to whether p is to the left, or to the right, of the nearest "removed" open interval of length $1 / 3^n$. Imitate the proof of Theorem 3 in Chapter 1, §9, for uncountability. See also Chapter 1, §9, Problem 2(ii).]

Exercise $6.8.E.3$

Complete the missing details in the proof of Theorems 1 to 4.

Exercise $6.8.E.4$

Prove the following.
(i) If $B \subseteq A$ and A is nowhere dense or meagre, so is B.
(ii) If $B \subseteq A$ and B is nonmeagre, so is A.
[Hint: Assume A is meagre and use (i)).]
(iii) Any finite union of nowhere dense sets is nowhere dense. Disprove it for infinite unions.
(iv) Any countable union of meagre sets is meagre.

Exercise $6.8.E.5$

Prove that in a discrete space (S, ρ), only $\emptyset$ is meagre.
[Hint: Use Problem 8 in Chapter 3, §17, Example 7 in Chapter 3, §12, and our present Theorem 1.]

Exercise $6.8.E.6$

Use Theorem 1 to give a new proof for the existence of irrationals in E^1.
[Hint: The rationals R are a meagre set, while E^1 is not.]

Exercise $6.8.E.7$

What is wrong about this "proof" that every closed set $F \neq \emptyset$ in a complete space (S, ρ) is residual: "By Theorem 5 of Chapter 3, §17, F is complete as a subspace. Thus by Theorem 1, F is residual." Give counterexamples!

Exercise $6.8.E.8$

We call K a $\mathcal{G}_\delta$-set and write $K \in \mathcal{G}_\delta$ iff $K=\bigcap_{n=1}^{\infty} G_n$ for some open sets G_n.
(i) Prove that if K is a $\mathcal{G}_\delta$-set, and if K is dense in a complete metric space (S, ρ), i.e., $\overline{K}=S,$ then K is residual in S.
[Hint: Let $F_n=-G_n$. Verify that $(\forall n) G_n$ is dense in S, and F_n is nowhere dense. Deduce that $-K=-\bigcap G_n=\cup F_n$ is meagre. Use Theorem 1.]
(ii) Infer that R (the rationals) is not a $\mathcal{G}_\delta$-set in E^1 (cf. Example(c)).

Exercise 6.8.E.9

Show that, in a complete metric space (S, ρ), a meagre set A cannot have interior points.
[Hint: Otherwise, A would obtain a globe G. Use Theorem 1 and Problem 4(ii).]

Exercise 6.8.E.10

(i) A singleton $\{p\} \subseteq(S, \rho)$ is nowhere dense if S clusters at p; otherwise, it is nonmeagre in S (being a globe, and not a union of nowhere dense sets).
(ii) If $A \subseteq S$ clusters at each $p \in A$, any countable set $B \subseteq A$ is meagre in S.

Exercise 6.8.E.11

(i) Show that if $\emptyset \neq A \in \mathcal{G}_{\delta}$ (see Problem 8) in a complete space (S, ρ), and A clusters at each $p \in A$, then A is uncountable.
(ii) Prove that any nonempty perfect set (Chapter 3, §14) in a complete space is uncountable.
(iii) How about R (the rationals) in E^{1} and in R as a subspace of E^{1}? What is wrong?
[Hints: (i) The subspace $(\overline{A}, \rho)$ is complete (why?); so A is nonmeagre in $\overline{A}$, by Problem 8. Use Problem 10(ii). (ii) Use Footnote 3.]

Exercise 6.8.E.12

If G is open in (S, ρ), then $\overline{G}-G$ is nowhere dense in S.
[Hint: $\overline{G}-G=\overline{G} \cap(-G)$ is closed; so

$$\overline{(\overline{G}-G)^{0}}=(\overline{G}-G)^{0}=(\overline{G} \cap-G)^{0}=\emptyset \tag{6.8.E.1}$$

by Problem 15 in Chapter 3, §12 and Problem 15 in Chapter 3, §16.]

Exercise 6.8.E.13

("Simplified" uniform boundedness theorem.) Let $f_{n}:(S, \rho) \rightarrow\left(T, \rho^{\prime}\right)$ be continuous for $n=1,2, \ldots$, with S complete. If $\left\{f_{n}(x)\right\}$ is a bounded sequence in T for each $x \in S$, then $\left\{f_{n}\right\}$ is uniformly bounded on some open $G \neq \emptyset$:

$$(\forall p \in T)(\exists k)(\forall n)(\forall x \in G) \quad \rho^{\prime}\left(p, f_{n}(x)\right) \leq k. \tag{6.8.E.2}$$

[Outline: Fix $p \in T$ and $(\forall n)$ set

$$F_{n}=\left\{x \in S \mid(\forall m) n \geq \rho^{\prime}\left(p, f_{m}(x)\right)\right\}. \tag{6.8.E.3}$$

Use the continuity of f_{m} and of $\rho^{\prime}$ to show that F_{n} is closed in S, and $S=\bigcup_{n=1}^{\infty} F_{n}$. By Theorem 1, S is nonmeagre; so at least one F_{n} is not nowhere dense-call it F, so $(\overline{F})^{0}=F^{0} \neq \emptyset$. Set $G=F^{0}$ and show that G is as required.]

Exercise 6.8.E.14

Let $f_{n}:(S, \rho) \rightarrow\left(T, \rho^{\prime}\right)$ be continuous for $n=1,2, \ldots$. Show that if $f_{n} \rightarrow f$ (pointwise) on S, then f is continuous on $S-Q$, with Q meagre in S.
[Outline: $(\forall k, m)$ let

$$A_{k m}=\bigcup_{m=n}^{\infty}\left\{x \in S \mid \rho^{\prime}\left(f_{n}(x), f_{m}(x)\right)>\frac{1}{k}\right\}. \tag{6.8.E.4}$$

By the continuity of ρ', f_n and f_m, A_{km} is open in S. (Why?) So by Problem 12, $\bigcup_{m=1}^{\infty}\left(\overline{A_{km}}-A_{km}\right)$ is meagre for $k=1,2,\ldots$.
Also, as $f_n \rightarrow f$ on $S, \bigcap_{m=1}^{\infty} A_{km}=\emptyset$. (Verify!) Thus

$$(\forall k) \quad \bigcap_{m=1}^{\infty} \overline{A_{km}} \subseteq \bigcup_{m=1}^{\infty}\left(\overline{A_{km}}-A_{km}\right). \tag{6.8.E.5}$$

(Why?) Hence the set $Q=\bigcup_{k=1}^{\infty} \bigcap_{m=1}^{\infty} \overline{A_{km}}$ is meagre in S.
Moreover, $S-Q=\bigcap_{k=1}^{\infty} \cup_{m=1}^{\infty}\left(-A_{km}\right)^{0}$ by Problem 16 in Chapter 3, §16. Deduce that if $p \in S-Q$, then

$$(\forall \varepsilon>0)\left(\exists m_{0}\right)\left(\exists G_{p}\right)\left(\forall n, m \geq m_{0}\right)\left(\forall x \in G_{0}\right) \quad \rho'\left(f_{m}(x), f_{n}(x)\right)<\varepsilon. \tag{6.8.E.6}$$

Keeping m fixed, let $n \rightarrow \infty$ to get

$$(\forall \varepsilon>0)\left(\exists m_{0}\right)\left(\exists G_{p}\right)\left(\forall m \geq m_{0}\right)\left(\forall x \in G_{p}\right) \quad \rho'\left(f_{m}(x), f(x)\right) \leq \varepsilon. \tag{6.8.E.7}$$

Now modify the proof of Theorem 2 of Chapter 4, §12, to show that this implies the continuity of f at each $p \in S-Q$.]

6.9: Local Extrema. Maxima and Minima

We say that $f : E' \rightarrow E^1$ has a local maximum (minimum) at $\vec{p} \in E'$ iff $f(\vec{p})$ is the largest (least) value of f on some globe G about $\vec{p}$; more precisely, iff

$$(\forall \vec{x} \in G) \quad \Delta f = f(\vec{x}) - f(\vec{p}) < 0 (> 0). \tag{6.9.1}$$

We speak of an improper extremum if we only have $\Delta f \leq 0 (\geq 0)$ on G. In any case, all depends on the sign of Δf.

From Problem 6 in §1, recall the following necessary condition.

Theorem 6.9.1

If $f : E' \rightarrow E^1$ has a local extremum at $\vec{p}$ then $D_{\vec{u}} f(\vec{p}) = 0$ for all $\vec{u} \neq \overrightarrow{0}$ in E'.

In the case $E' = E^n \left(C^n\right)$, this means that $d^1 f(\vec{p}; \cdot) = 0$ on E'.

(Recall that $d^1 f(\vec{p}; \vec{t}) = \sum_{k=1}^n D_k f(\vec{p}) t_k$. It vanishes if the $D_k f(\vec{p})$ do.

Note 1. This condition is only necessary, not sufficient. For example, if $f(x, y) = xy$, then $d^1 f(\overrightarrow{0}; \cdot) = 0$; yet f has no extremum at $\overrightarrow{0}$. (Verify!)

Sufficient conditions were given in Theorem 2 of §5, for $E' = E^1$. We now take up $E' = E^2$.

Theorem 6.9.2

Let $f : E^2 \rightarrow E^1$ be of class CD^2 on a globe $G = G_{\vec{p}}(\delta)$. Suppose $d^1 f(\vec{p}; \cdot) = 0$ on E^2. Set $A = D_{11} f(\vec{p}), B = D_{12} f(\vec{p})$, and $C = D_{22} f(\vec{p})$.

Then the following statements are true.

(i) If $AC > B^2$, has a maximum or minimum at $\vec{p}$, according to whether

$A < 0$ or $A > 0$.

(ii) If $AC < B^2$, f has no extremum at $\vec{p}$.

The case $AC = B$ is unresolved.

Proof

Let $\vec{x} \in G$ and $\vec{u} = \vec{x} - \vec{p} \neq \overrightarrow{0}$.

As $d^1 f(\vec{p}; \cdot) = 0$, Theorem 2 in §5, yields

$$\Delta f = f(\vec{x}) - f(\vec{p}) = R_1 = \frac{1}{2} d^2 f(\vec{s}; \vec{u}), \tag{6.9.2}$$

with $\vec{s} \in L(\vec{p}, \vec{x}) \subseteq G$ (see Corollary 1 of §5). As $f \in CD^2$, we have $D_{12} f = D_{21} f$ on G (Theorem 1 in §5). Thus by formula (4) in §5,

$$\Delta f = \frac{1}{2} d^2 f(\vec{s}; \vec{u}) = \frac{1}{2}\left[D_{11} f(\vec{s}) u_1^2 + 2 D_{12} f(\vec{s}) u_1 u_2 + D_{22} f(\vec{s}) u_2^2\right]. \tag{6.9.3}$$

Now, as the partials involved are continuous, we can choose $G = G_{\vec{p}}(\delta)$ so small that the sign of expression (1) will not change if $\vec{s}$ is replaced by $\vec{p}$. Then the crucial sign of Δf on G coincides with that of

$$D = A u_1^2 + 2 B u_1 u_2 + C u_2^2 \tag{6.9.4}$$

(with A, B, and C as stated in the theorem).

From (2) we obtain, by elementary algebra,

$$AD = (Au_1 + Bu_2)^2 + (AC - B^2)\,u_2^2,$$
$$CD = (Cu_1 + Bu_2)^2 + (AC - B^2)\,u_2^2.$$

Clearly, if $AC > B^2$, the right-side expression in (3) is >0; so $AD > 0$, i.e., D has the same sign as A.

Hence if $A < 0$, we also have $\Delta f < 0$ on G, and f has a maximum at $\vec{p}$. If $A > 0$, then $\Delta f > 0$, and f has a minimum at $\vec{p}$.

Now let $AC < B^2$. We claim that no matter how small $G = G_{\vec{p}}(\delta)$, Δf changes sign as $\vec{x}$ varies in G, and so f has no extremum at $\vec{p}$.

Indeed, we have $\vec{x} = \vec{p} + \vec{u}, \vec{u} = (u_1, u_2) \neq \overrightarrow{0}$. If $u_2 = 0$, (3) shows that D and Δf have the same sign as $A (A \neq 0)$.

But if $u_2 \neq 0$ and $u_1 = -Bu_2/A$ (assuming $A \neq 0$), then D and Δf have the sign opposite to that of A; and $\vec{x}$ is still in G if u_2 is small enough (how small?).

One proceeds similarly if $C \neq 0$ (interchange A and C, and use (3').

Finally, if $A = C = 0$, then by (2), $D = 2Bu_1u_2$ and $B \neq 0$ (since $AC < B^2$). Again D and Δf change sign as u_1u_2 does; so f has no extremum at $\vec{p}$. Thus all is proved. $\square$

Briefly, the proof utilizes the fact that the trinomial (2) is sign-changing iff its discriminant $B^2 - AC$ is positive, i.e., $\begin{vmatrix} A & B \\ B & C \end{vmatrix} < 0$.

Note 2. Functions $f : C \rightarrow E^1$ (of one complex variable) are likewise covered by Theorem 2 if one treats them as functions on E^2 (of two real variables).

Functions of n variables. Here we must rely on the algebraic theory of so-called symmetric quadratic forms, i.e., polynomials $P : E^n \rightarrow E^1$ of the form

$$P(\vec{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} u_i u_j, \tag{6.9.5}$$

where $\vec{u} = (u_i, \ldots, u_n) \in E^n$ and $a_{ij} = a_{ji} \in E^1$.

We take for granted a theorem due to J. J. Sylvester (see S. Perlis, Theory of Matrices, 1952, p. 197), which may be stated as follows.

Let $P : E^n \rightarrow E^1$ be a symmetric quadratic form,

$$P(\vec{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} u_i u_j. \tag{6.9.6}$$

(i) $P > 0$ on all of $E^n - \{\overrightarrow{0}\}$ iff the following n determinants A_k are positive:

$$A_k = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1k} \\ a_{21} & a_{22} & \ldots & a_{2k} \\ \ldots & \ldots\ldots\ldots & \ldots & a_{2k} \\ a_{k1} & a_{k2} & \ldots & a_{kk} \end{vmatrix}, \quad k = 1, 2, \ldots, n. \tag{6.9.7}$$

(ii) We have $P < 0$ on $E^n - \{\overrightarrow{0}\}$ iff $(-1)^k A_k > 0$ for $k = 1, 2, \ldots, n$.

Now we can extend Theorem 2 to the case $f : E^n \rightarrow E^1$. (This will also cover $f : C^n \rightarrow E^1$, treated as $f : E^{2n} \rightarrow E^1$.) The proof resembles that of Theorem 2.

Theorem 6.9.3

Let $f : E^n \to E^1$ be of class CD^2 on some $G = G_{\vec{p}}(\delta)$. Suppose $df(\vec{p}; \cdot) = 0$ on E^n. Define the A_k as in (4), with $a_{ij} = D_{ij} f(\vec{p}), i, j, k \leq n$ Then the following statements hold.

(i) f has a local minimum at $\vec{p}$ if $A_k > 0$ for $k = 1, 2, \ldots, n$.

(ii) f has a local maximum at $\vec{p}$ if $(-1)^k A_k > 0$ for $k = 1, \ldots, n$.

(iii) f has no extremum at $\vec{p}$ if the expression

$$P(\vec{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} u_i u_j \tag{6.9.8}$$

is > 0 for some $\vec{u} \in E^n$ and < 0 for others (i.e., P changes sign on E^n).

Proof

Let again $\vec{x} \in G, \vec{u} = \vec{x} - \vec{p} \neq \overrightarrow{0}$, and use Taylor's theorem to obtain

$$\Delta f = f(\vec{x}) - f(\vec{p}) = R_1 = \frac{1}{2} d^2 f(\vec{s}; \vec{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} D_{ij} f(\vec{s}) u_i u_j, \tag{6.9.9}$$

with $\vec{s} \in L(\vec{x}, \vec{p})$.

As $f \in CD^2$, the partials $D_{ij} f$ are continuous on G. Thus we can make G so small that the sign of the last double sum does not change if $\vec{s}$ is replaced by $\vec{p}$. Hence the sign of Δf on G is the same as that of $P(\vec{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} u_i u_j$, with the a_{ij} as stated in the theorem.

The quadratic form P is symmetric since $a_{ij} = a_{ji}$ by Theorem 1 in §5. Thus by Sylvester's theorem stated above, one easily obtains our assertions (i) and (ii). Indeed, they are immediate from clauses (i) and (ii) of that theorem.

Now, for (iii), suppose $P(\vec{u}) > 0 > P(\vec{v})$, i.e.,

$$\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} u_i u_j > 0 > \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} v_i v_j \quad \text{for some } \vec{u}, \vec{v} \in E^n - \{\overrightarrow{0}\}. \tag{6.9.10}$$

If here $\vec{u}$ and $\vec{v}$ are replaced by $t\vec{u}$ and $t\vec{v} (t \neq 0)$, then $u_i u_j$ and $v_i v_j$ turn into $t^2 u_i u_j$ and $t^2 v_i v_j$, respectively. Hence

$$P(t\vec{u}) = t^2 P(\vec{u}) > 0 > t^2 P(\vec{v}) = P(t\vec{v}). \tag{6.9.11}$$

Now, for any $t \in (0, \delta / |\vec{u}|)$, the point $\vec{x} = \vec{p} + t\vec{u}$ lies on the $\vec{u}$-directed line through $\vec{p}$, inside $G = G_{\vec{p}}(\delta)$. (Why?) Similarly for the point $\vec{x}' = \vec{p} + t\vec{v}$.

Hence for such $\vec{x}$ and $\vec{x}'$, Taylor's theorem again yields formulas analogous to (5) for some $\vec{s} \in L(\vec{p}, \vec{x})$ and $\vec{s}' \in L\left(\vec{p}, \vec{x}'\right)$ lying on the same two lines. It again follows that for small δ,

$$f(\vec{x}) - f(\vec{p}) > 0 > f\left(\vec{x}'\right) - f(\vec{p}), \tag{6.9.12}$$

just as $P(\vec{u}) > 0 > P(\vec{v})$.

Thus Δf changes sign on $G_{\vec{p}}(\delta)$, and (iii) is proved. □

Note 3. Still unresolved are cases in which $P(\vec{u})$ vanishes for some $\vec{u} \neq \overrightarrow{0}$, without changing its sign; e.g., $P(\vec{u}) = (u_1 + u_2 + u_3)^2 = 0$ for $\vec{u} = (1, 1, -2)$. Then the answer depends on higher-order terms of the Taylor formula. In particular, if $d^1 f(\vec{p}; \cdot) = d^2 f(\vec{p}; \cdot) = 0$ on E^n, then $\Delta f = R_2 = \frac{1}{6} d^3 f(\vec{p}; \vec{s})$, etc.

Note 4. The largest or least value of f on a set A (sometimes called the absolute maximum or minimum) may occur at sominterior (e.g., boundary) point $\vec{p} \in A$, and then fails to be among the local extrema (where, by definition, a globe $G_{\vec{p}} \subseteq A$ is presupposed). Thus to find absolute extrema, one must also explore the behaviour of f at noninterior points of A.

By Theorem 1, local extrema can occur only at so-called critical points $\vec{p}$, i.e., those at which all directional derivatives vanish (or fail to exist, in which case $D_{\vec{u}} f(\vec{p})=0$ by convention).

In practice, to find such points in $E^n\left(C^n\right)$, one equates the partials $D_k f\ (k \leq n)$ to 0. Then one uses Theorems 2 and 3 or other considerations to determine whether an extremum really exists.

Examples

(A) Find the largest value of

$$f(x, y)=\sin x+\sin y-\sin (x+y) \tag{6.9.13}$$

on the set $A \subseteq E^2$ bounded by the lines $x=0, y=0$ and $x+y=2 \pi$.

We have

$$D_1 f(x, y)=\cos x-\cos (x+y) \text{ and } D_2 f(x, y)=\cos y-\cos (x+y). \tag{6.9.14}$$

Inside the triangle A, both partials vanish only at the point $\left(\frac{2 \pi}{3}, \frac{2 \pi}{3}\right)$ at which $f=\frac{3}{2} \sqrt{3}$. On the boundary of A (i.e., on the lines $x=0, y=0$ and $x+y=2 \pi$), $f=0$. Thus even without using Theorem 2, it is evident that f attains its largest value,

$$f\left(\frac{2 \pi}{3}, \frac{2 \pi}{3}\right)=\frac{3}{2} \sqrt{3}, \tag{6.9.15}$$

at this unique critical point.

(B) Find the largest and the least value of

$$f(x, y, z)=a^2 x^2+b^2 y^2+c^2 z^2-\left(a x^2+b y^2+c z^2\right)^2, \tag{6.9.16}$$

on the condition that $x^2+y^2+z^2=1$ and $a>b>c>0$.

As $z^2=1-x^2-y^2$, we can eliminate z from $f(x, y, z)$ and replace f by $F: E^2 \rightarrow E^1$:

$$F(x, y)=\left(a^2-c^2\right) x^2+\left(b^2-c^2\right) y^2+c^2-\left[(a-c) x^2+(b-c) y^2+c\right]^2. \tag{6.9.17}$$

(Explain!) For F, we seek the extrema on the disc $\overline{G}=\overline{G}_0(1) \subset E^2$, where $x^2+y^2 \leq 1$ (so as not to violate the condition $x^2+y^2+z^2=1$) .

Equating to 0 the two partials

$$\begin{aligned} & D_1 F(x, y)=2 x(a-c)\left\{(a+c)-2\left[(a-c) x^2+(b-c) y^2+c\right]^2\right\}=0, \\ & D_2 F(x, y)=2 y(b-c)\left\{(b+c)-2\left[(a-c) x^2+(b-c) y^2+c\right]^2\right\}=0 \end{aligned} \tag{6.9.18}$$

and solving this system of equations, we find these critical points inside G :

(1) $x=y=0\ (F=0)$;

(2) $x=0, y=\pm 2^{-\frac{1}{2}}\left(F=\frac{1}{4}(b-c)^2\right)$; and

(3) $x=\pm 2^{-\frac{1}{2}}, y=0\left(F=\frac{1}{4}(a-c)^2\right)$.

(Verify!)

Now, for the boundary of $\overline{G}$, i.e., the circle $x^2+y^2=1$, repeat this process: substitute $y^2=1-x^2$ in the formula for $F(x, y)$, thus reducing it to

$$h(x)=\left(a^2-b^2\right) x^2+b^2+\left[(a-b) x^2+b\right]^2, \quad h: E^1 \rightarrow E^1, \tag{6.9.19}$$

on the interval $[-1,1] \subset E^1$. In $(-1,1)$ the derivative

$$h^{\prime}(x)=2(a-b) x\left(1-2 x^2\right) \tag{6.9.20}$$

vanishes only when

(4) $x=0\ (h=0)$, and

(5) $x=\pm 2^{-\frac{1}{2}}\ \left(h=\frac{1}{4}(a-b)^{2}\right)$.

Finally, at the endpoints of $[-1,1]$, we have

(6) $x=\pm 1\ (h=0)$.

Comparing the resulting function values in all six cases, we conclude that the least of them is 0, while the largest is $\frac{1}{4}(a-c)^{2}$. These are the desired least and largest values of f, subject to the conditions stated. They are attained, respectively, at the points

$$(0,0,\pm 1),(0,\pm 1,0),(\pm 1,0,0), \text{ and } \left(\pm 2^{-\frac{1}{2}}, 0, \pm 2^{-\frac{1}{2}}\right). \tag{6.9.21}$$

Again, the use of Theorems 2 and 3 was redundant. However, we suggest as an exercise that the reader test the critical points of F by using Theorem 2.

Caution. Theorems 1 to 3 apply to functions of independent variables only. In Example (B), x, y, z were made interdependent by the imposed equation

$$x^{2}+y^{2}+z^{2}=1 \tag{6.9.22}$$

(which geometrically limits all to the surface of $G_{\vec{0}}(1)$ in E^{3}), so that one of them, z, could be eliminated. Only then can Theorems 1 to 3 be used.

6.9.E: Problems on Maxima and Minima

Exercise $6.9.E.1$

Verify Note 1.

Exercise $6.9.E.1'$

Complete the missing details in the proof of Theorems 2 and 3.

Exercise $6.9.E.2$

Verify Examples (A) and (B). Supplement Example (A) by applying Theorem 2.

Exercise $6.9.E.3$

Test f for extrema in E^2 if $f(x, y)$ is
(i) $\frac{x^2}{2p} + \frac{y^2}{2q} (p > 0, q > 0)$;
(ii) $\frac{x^2}{2p} - \frac{y^2}{2q} (p > 0, q > 0)$;
(iii) $y^2 + x^4$;
(iv) $y^2 + x^3$.

Exercise $6.9.E.4$

(i) Find the maximum volume of an interval $A \subset E^3$ (see Chapter 3, §7) whose edge lengths x, y, z have a prescribed sum: $x + y + z = a$.
(ii) Do the same in E^4 and in E^n; show that A is a cube.
(iii) Hence deduce that

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{1}{n} \sum_1^n x_k \quad (x_k \geq 0), \tag{6.9.E.1}$$

i.e., the geometric mean of n nonnegative numbers is $\leq$ their arithmetic mean.

Exercise $6.9.E.5$

Find the minimum value for the sum $f(x, y, z, t) = x + y + z + t$ of four positive numbers on the condition that $xyzt = c^4$ (constant).
[Answer: $x = y = z = t = c$; $f_{\max} = 4c$.]

Exercise $6.9.E.6$

Among all triangles inscribed in a circle of radius R, find the one of maximum area.
[Hint: Connect the vertices with the center. Let x, y, z be the angles at the center. Show that the area of the triangle $= \frac{1}{2} R^2 (\sin x + \sin y + \sin z)$, with $z = 2\pi - (x + y)$.]

Exercise $6.9.E.7$

Among all intervals $A \subset E^3$ inscribed in the ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1 \tag{6.9.E.2}$$

find the one of largest volume.
[Answer: the edge lengths are $\frac{2a}{\sqrt{3}}, \frac{2b}{\sqrt{3}}, \frac{2c}{\sqrt{3}}$.]

Exercise $6.9.E.8$

Let $P_i=(a_i\cdot b_i), i=1,2,3,$ be 3 points in E^2 forming a triangle in which one angle (say, $easureangle P_1$) is $\geq 2\pi/3$. Find a point $P=(x,y)$ for which the sum of the distances,

$$PP_1+PP_2+PP_3=\sum_{i=1}^{3}\sqrt{(x-a_i)^2+(y-b_i)^2}, \tag{6.9.E.3}$$

is the least possible.
[Outline: Let $f(x,y)=\sum_{i=1}^{3}\sqrt{(x-a_i)^2+(y-b_i)^2}$.
Show that f has no partial derivatives at P_1, P_2, or P_3 (and so P_1, P_2, and P_3 are critical points at which an extremum may occur), while at other points P, partials do exist but never vanish simultaneously, so that there are no other critical points.
Indeed, prove that $D_1f(P)=0=D_2f(P)$ would imply that

$$\sum_{i=1}^{3}\cos\theta_i=0=\sum_{1}^{3}\sin\theta_I, \tag{6.9.E.4}$$

where θ_i is the angle between $\overline{PP_i}$ and the x-axis; hence

$$\sin(\theta_1-\theta_2)=\sin(\theta_2-\theta_3)=\sin(\theta_3-\theta_1) \quad (\text{why?}), \tag{6.9.E.5}$$

and so $\theta_1-\theta_2=\theta_2-\theta_3=\theta_3-\theta_1=2\pi/3, \quad$ contrary to $\angle P_1\geq 2\pi/3$. (Why?)
From geometric considerations, conclude that f has an absolute minimum at P_1.
(This shows that one cannot disregard points at which f has no partials.)]

Exercise $6.9.E.9$

Continuing Problem 8, show that if none of $\angle P_1, \angle P_2,$ and $\angle P_3$ is $\geq 2\pi/3,$ then f attains its least value at some P (inside the triangle) such that $\angle P_1PP_2=\angle P_2PP_3=\angle P_3PP_1=2\pi/3$.
[Hint: Verify that $D_1f=0=D_2f$ at P.
Use the law of cosines to show that $P_1P_2>PP_2+\frac{1}{2}PP_1$ and $P_1P_3>PP_3+\frac{1}{2}PP_1$.
Adding, obtain $P_1P_3+P_1P_2>PP_1+PP_2+PP_3, \quad$ i.e., $f(P_1)>f(P)$. Similarly, $f(P_2)>f(P)$ and $f(P_3)>f(P)$.
Combining with Problem 8, obtain the result.]

Exercise $6.9.E.10$

In a circle of radius R inscribe a polygon with $n+1$ sides of maximum area.
[Outline: Let $x_1, x_2, \ldots, x_{n+1}$ be the central angles subtended by the sides of the polygon. Then its area A is

$$\frac{1}{2}R^2\sum_{k=1}^{n+1}\sin x_k, \tag{6.9.E.6}$$

with $x_{n+1}=2\pi-\sum_{k=1}^{n}x_k$. (Why?) Thus all reduces to maximizing

$$f\left(x_{1}, \ldots, x_{n}\right)=\sum_{k=1}^{n} \sin x_{k}+\sin \left(2 \pi-\sum_{k=1}^{r_k} x_{k}\right), \tag{6.9.E.7}$$

on the condition that $0 \leq x_{k}$ and $\sum_{k=1}^{n} x_{k} \leq 2 \pi$. (Why?)
These inequalities define a bounded set $D \subset E^{n}$ (called a simplex). Equating all partials of f to 0, show that the only critical point interior to D is $\vec{x}=\left(x_{1}, \ldots, x_{n}\right)$, with $x_{k}=\frac{2 \pi}{n+1}, k \leq n$ (implying that $x_{n+1}=\frac{2 \pi}{n+1}$, too). For that $\vec{x}$, we get

$$f(\vec{x})=(n+1) \sin [2 \pi /(n+1)]. \tag{6.9.E.8}$$

This value must be compared with the "boundary" values of f, on the "faces" of the simplex D (see Note 4).
Do this by induction. For $n=2$, Problem 6 shows that $f(\vec{x})$ is indeed the largest when all x_{k} equal $\frac{2 \pi}{n+1}$. Now let D_{n} be the "face" of D, where $x_{n}=0$. On that face, treat f as a function of only $n-1$ variables, $x_{1}, \ldots, x_{n-1}$.
By the inductive hypothesis, the largest value of f on D_{n} is $n \sin (2 \pi / n)$. Similarly for the other "faces." As $n \sin (2 \pi / n)<(n+1) \sin 2 \pi /(n+1)$, the induction is complete.
Thus, the area A is the largest when the polygon is regular, for which

$$A=\frac{1}{2} R^{2}(n+1) \sin \frac{2 \pi}{n+1} .] \tag{6.9.E.9}$$

Exercise $6.9.E.11$

Among all triangles of a prescribed perimeter $2p$, find the one of maximum area.
[Hint: Maximize $p(p-x)(p-y)(p-z)$ on the condition that $x+y+z=2 p$.]

Exercise $6.9.E.12$

Among all triangles of area A, find the one of smallest perimeter.

Exercise $6.9.E.13$

Find the shortest distance from a given point $\vec{p} \in E^{n}$ to a given plane $\vec{u} \cdot \vec{x}=c$ (Chapter 3, §§4-6). Answer:

$$\pm \frac{\vec{u} \cdot \vec{p}-c}{|\vec{u}|}. \tag{6.9.E.10}$$

[Hint: First do it in E^{3}, writing (x, y, z) for $\vec{x}$.]

6.10: More on Implicit Differentiation. Conditional Extrema

I. Implicit differentiation was sketched in §7. Under suitable assumptions (Theorem 4 in §7), one can differentiate a given system of equations,

$$g_k\left(x_1, \ldots, x_n, y_1, \ldots, y_m\right)=0, \quad k=1,2, \ldots, n, \tag{6.10.1}$$

treating the x_j as implicit functions of the y_i without seeking an explicit solution of the form

$$x_j=H_j\left(y_1, \ldots, y_m\right). \tag{6.10.2}$$

This yields a new system of equations from which the partials $D_i H_j=\frac{\partial x_j}{\partial y_i}$ can be found directly.

We now supplement Theorem 4 in §7 (review it!) by showing that this new system is linear in the partials involved and that its determinant is $\neq 0$. Thus in general, it is simpler to solve than (1).

As in Part IV of §7, we set

$$(\vec{x}, \vec{y})=\left(x_1, \ldots, x_n, y_1, \ldots, y_m\right) \text{ and } g=\left(g_1, \ldots, g_n\right), \tag{6.10.3}$$

replacing the f of §7 by g. Then equations (1) simplify to

$$g(\vec{x}, \vec{y})=\overrightarrow{0}, \tag{6.10.4}$$

where $g: E^{n+m} \rightarrow E^n$ (or $g: C^{n+m} \rightarrow C^n$).

Theorem 6.10.1 (implicit differentiation)

Adopt all assumptions of Theorem 4 in §7, replacing f by g and setting $H=\left(H_1, \ldots, H_n\right)$,

$$D_j g_k(\vec{p}, \vec{q})=a_{jk}, \quad j \leq n+m, \quad k \leq n. \tag{6.10.5}$$

Then for each $i=1, \ldots, m$, we have n linear equations,

$$\sum_{j=1}^n a_{jk} D_i H_j(\vec{q})=-a_{n+i, k}, \quad k \leq n, \tag{6.10.6}$$

with

$$\det\left(a_{jk}\right) \neq 0, \quad(j, k \leq n), \tag{6.10.7}$$

that uniquely determine the partials $D_i H_j(\vec{q})$ for $j=1,2, \ldots, n$.

Proof

As usual, extend the map $H: Q \rightarrow P$ of Theorem 4 in §7 to $H: E^m \rightarrow E^n$ (or $C^m \rightarrow C^n$) by setting $H=\overrightarrow{0}$ on $-Q$. Also, define $\sigma: E^m \rightarrow E^{n+m}\left(C^m \rightarrow C^{n+m}\right)$ by

$$\sigma(\vec{y})=(H(\vec{y}), \vec{y})=\left(H_1(\vec{y}), \ldots, H_n(\vec{y}), y_1, \ldots, y_m\right), \quad \vec{y} \in E^m\left(C^m\right). \tag{6.10.8}$$

Then σ is differentiable at $\vec{q} \in Q$, as are its $n+m$ components. (Why?) since $\vec{x}=H(\vec{y})$ is a solution of (2), equations (1) and (2) become identities when $\vec{x}$ is replaced by $H(\vec{y})$. Also, $\sigma(\vec{q})=(H(\vec{q}), \vec{q})=(\vec{p}, \vec{q})$ since $H(\vec{q})=\vec{p}$. Moreover,

$$g(\sigma(\vec{y}))=g(H(\vec{y}), \vec{y})=\overrightarrow{0} \text{ for } \vec{y} \in Q; \tag{6.10.9}$$

i.e., $g \circ \sigma=\overrightarrow{0}$ on Q.

Now, by assumption, $g \in CD^1$ at $(\vec{p}, \vec{q})$; so the chain rule (Theorem 2 in §4) applies, with $f, \vec{p}, \vec{q}, n$, and m replaced by $\sigma, \vec{q},(\vec{p}, \vec{q}), m$, and $n+m$, respectively.

As $h=g \circ \sigma=\overrightarrow{0}$ on Q, an open set, the partials of h vanish on Q. So by Theorem 2 of §4, writing σ_j for the jth component of σ,

$$\vec{0}=\sum_{j=1}^{n+m} D_j g(\vec{p}, \vec{q}) \cdot D_i \sigma_j(\vec{q}), \quad i \leq m. \tag{6.10.10}$$

By (4), $\sigma_j=H_j$ if $j \leq n$, and $\sigma_j(\vec{y})=y_i$ if $j=n+i$. Thus $D_i \sigma_j=D_i H_j$ $j \leq n$; but for $j>n$, we have $D_i \sigma_j=1$ if $j=n+i$, and $D_i \sigma_j=0$ otherwise. Hence by (5),

$$\vec{0}=\sum_{j=1}^n D_j g(\vec{p}, \vec{q}) \cdot D_i H_j(\vec{q})+D_{n+i} g(\vec{p}, \vec{q}), \quad i=1,2, \ldots, m. \tag{6.10.11}$$

As $g=\left(g_1, \ldots, g_n\right)$, each of these vector equations splits into n scalar ones:

$$0=\sum_{j=1}^n D_j g_k(\vec{p}, \vec{q}) \cdot D_i H_j(\vec{q})+D_{n+i} g_k(\vec{p}, \vec{q}), \quad i \leq m, k \leq n. \tag{6.10.12}$$

With $D_j g_k(\vec{p}, \vec{q})=a_{jk}$, this yields (3), where $\det\left(a_{jk}\right)=\det\left(D_j g_k(\vec{p}, \vec{q})\right) \neq 0$ by hypothesis (see Theorem 4 in §7).

Thus all is proved. □

Note 1. By continuity (Note 1 in §6), we have det $D_j g_k(\vec{x}, \vec{y})) \neq 0$ for all $(\vec{x}, \vec{y})$ in a sufficiently small neighborhood of $(\vec{p}, \vec{q})$. Thus Theorem 1 holds also with $(\vec{p}, \vec{q})$ replaced by such $(\vec{x}, \vec{y})$. In practice, one does not have to memorize (3), but one obtains it by implicitly differentiating equations (1).

II. We shall now apply Theorem 1 to the theory of conditional extrema.

Definition 1

We say that $f: E^{n+m} \rightarrow E^1$ has a local conditional maximum (minimum) at $\vec{p} \in E^{n+m}$, with constraints

$$g=\left(g_1, \ldots, g_n\right)=\vec{0} \tag{6.10.13}$$

$\left(g: E^{n+m} \rightarrow E^n\right)$ iff in some neighborhood G of $\vec{p}$ we have

$$\Delta f=f(\vec{x})-f(\vec{p}) \leq 0 \quad (\geq 0, \text{ respectively}) \tag{6.10.14}$$

for all $\vec{x} \in G$ for which $g(\vec{x})=\vec{0}$.

In §9 (Example (B) and Problems), we found such conditional extrema by using the constraint equations $g=\vec{0}$ to eliminate some variables and thus reduce all to finding the unconditional extrema of a function of fewer (independent) variables.

Often, however, such elimination is cumbersome since it involves solving a system (1) of possibly nonlinear equations. It is here that implicit differentiation (based on Theorem 1) is useful.

Lagrange invented a method (known as that of multipliers) for finding the critical points at which such extrema may exist; to wit, we have the following:

Given $f: E^{n+m} \rightarrow E^1$, set

$$F=f+\sum_{k=1}^n c_k g_k, \tag{6.10.15}$$

where the constants c_k are to be determined and g_k are as above.

Then find the partials $D_j F(j \leq n+m)$ and solve the system of $2n+m$ equations

$$D_j F(\vec{x})=0, \quad j \leq n+m, \quad \text{and} \quad g_k(\vec{x})=0, \quad k \leq n, \tag{6.10.16}$$

for the $2n+m$ "unknowns" $x_j(j \leq n+m)$ and $c_k(k \leq n)$, the c_k originating from (7).

Any $\vec{x}$ satisfying (8), with the c_k so determined is a critical point (still to be tested). The method is based on Theorem 2 below, where we again write $(\vec{p}, \vec{q})$ for $\vec{p}$ and $(\vec{x}, \vec{y})$ for $\vec{x}$ (we call it "double notation").

Theorem 6.10.2 (Lagrange multipliers)

Suppose $f : E^{n+m} \rightarrow E^1$ is differentiable at

$$(\vec{p}, \vec{q}) = (p_1, \ldots, p_n, q_1, \ldots, q_m) \tag{6.10.17}$$

and has a local extremum at $(\vec{p}, \vec{q})$ subject to the constraints

$$g = (g_1, \ldots, g_n) = \overrightarrow{0}, \tag{6.10.18}$$

with g as in Theorem 1, $g : E^{n+m} \rightarrow E^n$. Then

$$\sum_{k=1}^{n} c_k D_j g_k(\vec{p}, \vec{q}) = -D_j f(\vec{p}, \vec{q}), \quad j = 1, 2, \ldots, n+m, \tag{6.10.19}$$

for certain multipliers c_k (determined by the first n equations in (9)).

Proof

These n equations admit a unique solution for the c_k, as they are linear, and

$$\det(D_j g_k(\vec{p}, \vec{q})) \neq 0 \quad (j, k \leq n) \tag{6.10.20}$$

by hypothesis. With the c_k so determined, (9) holds for $j \leq n$. It remains to prove (9) for $n < j \leq n+m$.

Now, since f has a conditional extremum at $(\vec{p}, \vec{q})$ as stated, we have

$$f(\vec{x}, \vec{y}) - f(\vec{p}, \vec{q}) \leq 0 \quad (\text{ or } \quad \geq 0) \tag{6.10.21}$$

for all $(\vec{x}, \vec{y}) \in P \times Q$ with $g(\vec{x}, \vec{y}) = \overrightarrow{0}$, provided we make the neighborhood $P \times Q$ small enough.

Define H and σ as in the previous proof (see (4)); so $\vec{x} = H(\vec{y})$ is equivalent to $g(\vec{x}, \vec{y}) = \overrightarrow{0}$ for $(\vec{x}, \vec{y}) \in P \times Q$.

Then, for all such $(\vec{x}, \vec{y})$, with $\vec{x} = H(\vec{y})$, we surely have $g(\vec{x}, \vec{y}) = \overrightarrow{0}$ and also

$$f(\vec{x}, \vec{y}) = f(H(\vec{y}), \vec{y}) = f(\sigma(\vec{y})). \tag{6.10.22}$$

Set $h = f \circ \sigma, h : E^m \rightarrow E^1$. Then (10) reduces to

$$h(\vec{y}) - h(\vec{q}) \leq 0 (\text{ or } \geq 0) \quad \text{for all } \vec{y} \in Q. \tag{6.10.23}$$

This means that h has an unconditional extremum at $\vec{q}$, an interior point of Q. Thus, by Theorem 1 in §9,

$$D_i h(\vec{q}) = 0, \quad i = 1, \ldots, m. \tag{6.10.24}$$

Hence, applying the chain rule (Theorem 2 of §4) to $h = f \circ \sigma$, we get, much as in the previous proof,

$$\begin{aligned} 0 &= \sum_{j=1}^{n+m} D_j f(\vec{p}, \vec{q}) D_i \sigma_j(\vec{q}) \\ &= \sum_{j=1}^{n} D_j f(\vec{p}, \vec{q}) D_i H_j(\vec{q}) + D_{n+i} f(\vec{p}, \vec{q}), \quad i \leq m. \end{aligned}$$

(Verify!)

Next, as g by hypothesis satisfies Theorem 1, we get equations (3) or equivalently (6). Multiplying (6) by c_k, adding and combining with (11), we obtain

$$\begin{array}{c} \sum_{j=1}^{n} \left[D_j f(\vec{p}, \vec{q}) + \sum_{k=1}^{n} c_k D_j g_k(\vec{p}, \vec{q}) \right] D_i H_j(\vec{q}) \\ + D_{n+i} f(\vec{p}, \vec{q}) + \sum_{k=1}^{n} c_k D_{n+i} g_k(\vec{p}, \vec{q}) = 0, \quad i \leq m. \end{array} \tag{6.10.25}$$

(Verify!) But the square-bracketed expression is 0; for we chose the c_k so as to satisfy (9) for $j \leq n$. Thus all simplifies to

$$\sum_{k=1}^{n} c_k D_{n+i} g_k(\vec{p}, \vec{q}) = -D_{n+i} f(\vec{p}, \vec{q}), \quad i = 1, 2, \ldots, m. \tag{6.10.26}$$

Hence (9) holds for $n < j \leq n+m$, too, and all is proved. □

Remarks. Lagrange's method has the advantage that all variables (the x_k and y_i) are treated equally, without singling out the dependent ones. Thus in applications, one uses only F, i.e., f and g (not H).

One can also write $\vec{x} = (x_1, \ldots, x_{n+m})$ for $(\vec{x}, \vec{y}) = (x_1, \ldots, x_n, y_1, \ldots, y_m)$ (the "double" notation was good for the proof only).

On the other hand, one still must solve equations (8).

Theorem 2 yields only a necessary condition (9) for extrema with constraints. There also are various sufficient conditions, but mostly one uses geometric and other considerations instead (as we did in §9). Therefore, we limit ourselves to one proposition (using "single" notation this time).

Theorem 6.10.3 (sufficient conditions)

Let

$$F = f + \sum_{k=1}^{n} c_k g_k, \tag{6.10.27}$$

with $f : E^{n+m} \to E^1, g : E^{n+m} \to E^n$, and c_k as in Theorem 2.

Then f has a maximum (minimum) at $\vec{p} = (p_1, \ldots, p_{n+m})$ (with constraints $g = (g_1, \ldots, g_n) = \overrightarrow{0}$ whenever F does. (A fortiori, this is the case if F has an unconditional extremum at $\vec{p}$.)

Proof

Suppose F has a maximum at $\vec{p}$, with constraints $g = \overrightarrow{0}$. Then

$$0 \geq F(\vec{x}) - F(\vec{p}) = f(\vec{x}) - f(\vec{p}) + \sum_{k=1}^{n} c_k \left[g_k(\vec{x}) - g_k(\vec{p})\right] \tag{6.10.28}$$

for those $\vec{x}$ near $\vec{p}$ (including $\vec{x} = \vec{p}$) for which $g(\vec{x}) = \overrightarrow{0}$.

But for such $\vec{x}, g_k(\vec{x}) = g_k(\vec{p}) = 0, c_k \left[g_k(\vec{x}) - g_k(\vec{p})\right] = 0,$ and so

$$0 \geq F(\vec{x}) - F(\vec{p}) = f(\vec{x}) - f(\vec{p}). \tag{6.10.29}$$

Hence f has a maximum at $\vec{p}$, with constraints as stated.

Similarly, $\Delta F = \Delta f$ in case F has a conditional minimum at $\vec{p}$. □

Example 1

Find the local extrema of

$$f(x, y, z, t) = x + y + z + t \tag{6.10.30}$$

on the condition that

$$g(x, y, z, t) = xyzt - a^4 = 0, \tag{6.10.31}$$

with $a > 0$ and $x, y, z, t > 0$. (Note that inequalities do not count as "constraints" in the sense of Theorems 2 and 3.) Here one can simply eliminate $t = a^4/(xyz)$, but it is still easier to use Lagrange's method.

Set $F(x, y, z, t) = x + y + z + t + cxyzt$. (We drop a^4 since it will anyway disappear in differentiation.) Equations (8) then read

$$0=1+cyzt=1+cxzt=1+cxyt=1+cxyz, \quad xyzt-a^{4}=0. \tag{6.10.32}$$

Solving for x, z, t and c, we get $c=-a^{-3}, x=y=z=t=a$.

Thus $F(x, y, z, t)=x+y+z+t-xyzt/a^{3},$ and the only critical point is $\vec{p}=(a, a, a, a)$. (Verify!)

By Theorem 3, one can now explore the sign of $F(\vec{x})-F(\vec{p}),$ where $\vec{x}=(x, y, z, t)$. For $\vec{x}$ near $\vec{p}$, it agrees with the sign of $d^{2} F(\vec{p} ; \cdot)$. (See proof of Theorem 2 in §9.) We shall do it below, using yet another device, to be explained now.

Elimination of dependent differentials. If all partials of F vanish at $\vec{p}$ (e.g., if $\vec{p}$ satisfies (9), then $d^{1} F(\vec{p} ; \cdot)=0$ on E^{n+m} (briefly $dF \equiv 0$).

Conversely, if $d^{1} f(\vec{p} ; \cdot)=0$ on a globe $G_{\vec{p}}$, for some function f on n independent variables, then

$$D_{k} f(\vec{p})=0, \quad k=1,2, \ldots, n, \tag{6.10.33}$$

since $d^{1} f(\vec{p} ; \cdot)$ (a polynomial!) vanishes at infinitely many points if its coefficients $D_{k} f(\vec{p})$ vanish. (The latter fails, however, if the variables are interdependent.)

Thus, instead of working with the partials, one can equate to 0 the differential dF or df. Using the "variable" notation and the invariance of df (Note 4 in §4), one then writes $dx, dy, \ldots$ for the "differentials" of dependent and independent variables alike, and tries to eliminate the differentials of the dependent variables. We now redo Example 1 using this method.

Example 2

With f and g as in Example 1, we treat t as the dependent variable, i.e., an implicit function of x, y, z,

$$t=a^{4}/(xyz)=H(x, y, z), \tag{6.10.34}$$

and differentiate the identity $xyzt-a^{4}=0$ to obtain

$$0=yzt dx+xzt dy+xyt dz+xyz dt; \tag{6.10.35}$$

so

$$dt=-t\left(\frac{dx}{x}+\frac{dy}{y}+\frac{dz}{z}\right). \tag{6.10.36}$$

Substituting this value of dt in $df=dx+dy+dz+dt=0$ (the equation for critical points), we eliminate dt and find:

$$\left(1-\frac{t}{x}\right) dx+\left(1-\frac{t}{y}\right) dy+\left(1-\frac{t}{z}\right) dz \equiv 0. \tag{6.10.37}$$

As x, y, z are independent variables, this identity implies that the coefficients of $dx, dy,$ and dz must vanish, as pointed out above. Thus

$$1-\frac{t}{x}=1-\frac{t}{y}=1-\frac{t}{z}=0. \tag{6.10.38}$$

Hence $x=y=z=t=a$. (Why?) Thus again, the only critical point is $\vec{p}=(a, a, a, a)$.

Now, returning to Lagrange's method, we use formula (5) in §5 to compute

$$d^{2} F=-\frac{2}{a}(dx dy+dx dz+dz dt+dx dt+dy dz+dy dt). \tag{6.10.39}$$

(Verify!)

We shall show that this expression is sign-constant (if $xyzt=a^{4}$), near the critical point $\vec{p}$. Indeed, setting $x=y=z=t=a$ in (12), we get $dt=-(dx+dy+dz),$ and (13) turns into

$$\begin{aligned} -\frac{2}{a}&\left[dx dy+dx dz+dy dz-(dx+dy+dz)^{2}\right] \\ &=\frac{1}{a}\left[dx^{2}+dy^{2}+dz^{2}+(dx+dy+dz)^{2}\right]=d^{2} F. \end{aligned}$$

This expression is >0 (for dx, dy, and dz are not all 0). Thus f has
a local conditional minimum at $\vec{p}=(a, a, a, a)$.

Caution; here we cannot infer that $f(\vec{p})$ is the least value of f under the imposed conditions: $x, y, z>0$ and $xyzt=a^4$.

The simplification due to the Cauchy invariant rule (Note 4 in §4) makes the use of the "variable" notation attractive, though caution ismandatory.

Note 2. When using Theorem 2, it suffices to ascertain that some n equations from (9) admit a solution for the c_k; for then, renumbering the equations, one can achieve that these become the first n equations, as was assumed. This means that the $n \times(n+m)$ matrix $\left(D_j g_k(\vec{p}, \vec{q})\right)$ must be of rank n, i.e., contains an $n \times n$-submatrix (obtained by deleting some columns), with a nonzero determinant.

In the Problems we often use $r, s, t, \ldots$ for Lagrange multipliers.

6.10.E: Further Problems on Maxima and Minima

Exercise $6.10.E.1$

Fill in all details in Examples 1 and 2 and the proofs of all theorems in this section.

Exercise $6.10.E.2$

Redo Example (B) in §9 by Lagrange's method.
[Hint: Set $F(x, y, z)=f(x, y, z)-r\left(x^{2}+y^{2}+z^{2}\right), g(x, y, z)=x^{2}+y^{2}+z^{2}-1$. Compare the values of f at all critical points.]

Exercise $6.10.E.3$

An ellipsoid

$$\frac{x^{2}}{a^{2}}+\frac{y^{2}}{b^{2}}+\frac{z^{2}}{c^{2}}=1 \tag{6.10.E.1}$$

is cut by a plane $u x+v y+w z=0$. Find the semiaxes of the section-ellipse, i.e., the extrema of

$$\rho^{2}=[f(x, y, z)]^{2}=x^{2}+y^{2}+z^{2} \tag{6.10.E.2}$$

under the constraints $g=\left(g_{1}, g_{2}\right)=\overrightarrow{0}$, where

$$g_{1}(x, y, z)=u x+v y+w z \text{ and } g_{2}(x, y, z)=\frac{x^{2}}{a^{2}}+\frac{y^{2}}{b^{2}}+\frac{z^{2}}{c^{2}}-1. \tag{6.10.E.3}$$

Assume that $a>b>c>0$ and that not all $u, v, w=0$.
[Outline: By Note 2, explore the rank of the matrix

$$\left(\begin{array}{ccc}x / a^{2} & y / b^{2} & z / c^{2} \\ u & v & z\end{array}\right). \tag{6.10.E.4}$$

(Why this particular matrix?)
Seeking a contradiction, suppose all its 2×2 determinants vanish at all points of the section-ellipse. Then the upper and lower entries in (14) are proportional (why?); so $x^{2} / a^{2}+y^{2} / b^{2}+z^{2} / c^{2}=0$ (a contradiction!).
Next, set

$$F(x, y, z)=x^{2}+y^{2}+z^{2}+r\left(\frac{x^{2}}{a^{2}}+\frac{y^{2}}{b^{2}}+\frac{z^{2}}{c^{2}}\right)+2 s(u x+v y+w z). \tag{6.10.E.5}$$

Equate $d F$ to 0 :

$$x+\frac{r x}{a^{2}}+s u=0, \quad y+\frac{r y}{b^{2}}+s v=0, \quad z+\frac{r z}{c^{2}}+s w=0. \tag{6.10.E.6}$$

Multiplying by $x, y, z,$ respectively, adding, and combining with $g=\overrightarrow{0}$, obtain $r=-\rho^{2}$; so, by (15), for $a, b, c \neq \rho$,

$$x=\frac{-s u a^{2}}{a^{2}-\rho^{2}}, \quad y=\frac{-s v b^{2}}{b^{2}-\rho^{2}}, \quad z=\frac{-s w c^{2}}{c^{2}-\rho^{2}}. \tag{6.10.E.7}$$

Find s, x, y, z, then compare the ρ-values at critical points.]

Exercise $6.10.E.4$

Find the least and the largest values of the quadratic form

$$f(\vec{x})=\sum_{i,k=1}^{n} a_{ik} x_i x_k \quad \left(a_{ik}=a_{ki}\right) \tag{6.10.E.8}$$

on the condition that $g(\vec{x})=|\vec{x}|^2-1=0\left(f, g: E^n \rightarrow E^1\right)$.
[Outline: Let $F(\vec{x})=f(\vec{x})-t\left(x_1^2+x_2^2+\ldots+x_n^2\right)$. Equating dF to 0, obtain

$$\begin{array}{l}\left(a_{11}-t\right) x_1+a_{12} x_2+\ldots+a_{1n} x_n=0, \\ a_{21} x_1+\left(a_{22}-t\right) x_2+\ldots+a_{2n} x_n=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1} x_1+a_{n2} x_2+\ldots+\left(a_{nn}-t\right) x_n=0.\end{array} \tag{6.10.E.9}$$

Using Theorem 1(iv) in §6, derive the so-called characteristic equation of f,

$$\left|\begin{array}{cccc}a_{11}-t & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22}-t & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{2n} & \ldots & a_{nn}-t\end{array}\right|=0, \tag{6.10.E.10}$$

of degree n in t. If t is one of its n roots (known to be real), then equations (16) admit a nonzero solution for $\vec{x}=\left(x_1, \ldots, x_n\right)$; by replacing $\vec{x}$ by $\vec{x}/|\vec{x}|$ if necessary, $\vec{x}$ satisfies also the constraint equation $g(\vec{x})=|\vec{x}|^2-1=0$. (Explain!) Thus each root t of (17) yields a critical point $\vec{x}_t=\left(x_1, \ldots, x_n\right)$.
Now, to find $f\left(\vec{x}_t\right)$, multiply the kth equation in (16) by $x_k, k=1, \ldots, n$, and add to get

$$0=\sum_{i,k=1}^{n} a_{ik} x_i x_k-t \sum_{k=1}^{n} x_k^2=f\left(\vec{x}_t\right)-t. \tag{6.10.E.11}$$

Hence $f\left(\vec{x}_t\right)=t$.
Thus the values of f at the critical points $\vec{x}_t$ are simply the roots of (17). The largest (smallest) root is also the largest (least) value of f on $S=\left\{\vec{x} \in E^n || \vec{x} \mid=1\right\}$ (Explain!)]

Exercise $6.10.E.5$

Use the method of Problem 4 to find the semiaxes of
(i) the quadric curve in E^2, centered at $\overrightarrow{0}$, given by $\sum_{i,k=1}^2 a_{ik} x_i x_k=1$; and
(ii) the quadric surface $\sum_{i,k=1}^3 a_{ik} x_i x_k=1$ in E^3, centered at (\overrightarrow{0}\).
Assume $a_{ik}=a_{kI}$.
[Hint: Explore the extrema of $f(\vec{x})=|\vec{x}|^2$ on the condition that

$$g(\vec{x})=\sum_{i,k} a_{ik} x_i x_k-1=0.] \tag{6.10.E.12}$$

Exercise $6.10.E.6$

Using Lagrange's method, redo Problems 4, 5, 6, 7, 11, 12, and 13 of §9.

Exercise $6.10.E.7$

In E^2, find the shortest distance from $\overrightarrow{0}$ to the parabola $y^2 = 2(x+a)$.

Exercise $6.10.E.8$

In E^3, find the shortest distance from $\overrightarrow{0}$ to the intersection line of two planes given by the formulas $\vec{u} \cdot \vec{x} = a$ and $\vec{v} \cdot \vec{x} = b$ with $\vec{u}$ and $\vec{v}$ different from $\overrightarrow{0}$. (Rewrite all in coordinate form!)

Exercise $6.10.E.9$

In E^n, find the largest value of $|\vec{a} \cdot \vec{x}|$ if $|\vec{x}| = 1$. Use Lagrange's method.

Exercise $6.10.E.10*$

(Hadamard's theorem.) If $A = \operatorname{det}(x_{ik})(i, k \leq n)$, then

$$|A| \leq \prod_{i=1}^{n} |\vec{x}_i|, \tag{6.10.E.13}$$

where $\vec{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in})$.
[Hints: Set $a_i = |\vec{x}_i|$. Treat A as a function of n^2 variables. Using Lagrange's method, prove that, under the n constraints $|\vec{x}_i|^2 - a_i^2 = 0$, A cannot have an extremum unless $A^2 = \operatorname{det}(y_{ik})$, with $y_{Ik} = 0$ (if $i \neq k$) and \(y_{ii}=a_{i}^{2}.]

CHAPTER OVERVIEW

7: Volume and Measure

I. Our theory of set families leads quite naturally to a generalization of metric spaces. As we know, in any such space (S, ρ), there is a family $\mathcal{G}$ of open sets, and a family $\mathcal{F}$ of all closed sets. In Chapter 3, §12, we derived the following two properties.

(i) $\mathcal{G}$ is closed under any (even uncountable) unions and under finite intersections (Chapter 3, §12, Theorem 2). Moreover,

$$\emptyset \in \mathcal{G} \text{ and } S \in \mathcal{G}. \tag{7.1}$$

(ii) $\mathcal{F}$ has these properties, with "unions" and "intersections" interchanged (Chapter 3, §12, Theorem 3). Moreover, by definition,

$$A \in \mathcal{F} \text{ iff } -A \in \mathcal{G}. \tag{7.2}$$

Now, quite often, it is not so important to have distances (i.e., a metric) defined in S, but rather to single out two set families, $\mathcal{G}$ and $\mathcal{F}$, with properties (i) and (ii), in a suitable manner. For examples, see Problems 1 to 4 below. Once $\mathcal{G}$ and $\mathcal{F}$ are given, one does not need a metric to define such notions as continuity, limits, etc. (See Problems 2 and 3.) This leads us to the following definition.

7.1: More on Intervals in E^n. Semirings of Sets

I. As a prologue, we turn to intervals in E^n (Chapter 3, §7).

Theorem 7.1.1

If A and B are intervals in E^n, then

(i) $A \cap B$ is an interval ($\emptyset$ counts as an interval);

(ii) $A - B$ is the union of finitely many disjoint intervals (but need not be an interval itself).

Proof

The easy proof for E^1 is left to the reader.

An interval in E^2 is the cross-product of two line intervals.

Let

$$A = X \times Y \text{ and } B = X' \times Y', \tag{7.1.1}$$

where X, Y, X', and Y' are intervals in E^1. Then (see Figure 29)

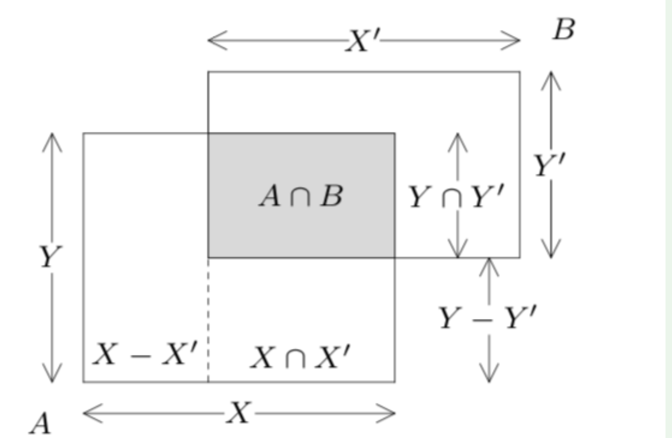

FIGURE 29

and

$$A - B = [(X - X') \times Y] \cup [(X \cap X') \times (Y - Y')]; \tag{7.1.2}$$

see Problem 8 in Chapter 1, §§1-3.

As the theorem holds in E^1,

$$X \cap X' \text{ and } Y \cap Y' \tag{7.1.3}$$

are intervals in E^1, while

$$X - Y' \text{ and } Y - Y' \tag{7.1.4}$$

are finite unions of disjoint line intervals. (In Figure 29 they are just intervals, but in general they are not.)

It easily follows that $A \cap B$ is an interval in E^2, while $A - B$ splits into finitely many such intervals. (Verify!) Thus the theorem holds in E^2.

Finally, for E^n, use induction. An interval in E^n is the cross-product of an interval in E^{n-1} by a line interval. Thus if the theorem holds in E^{n-1}, the same argument shows that it holds in E^n, too. (Verify!)

This completes the inductive proof. $\square$

Actually, Theorem 1 applies to many other families of sets (not necessarily intervals or sets in E^n). We now give such families a name.

Definition 1

A family $\mathcal{C}$ of arbitrary sets is called a semiring iff

(i) $\emptyset \in \mathcal{C}$ ($\emptyset$ is a member), and

(ii) for any sets A and B from $\mathcal{C}$, we have $A \cap B \in \mathcal{C}$, while $A-B$ is the union of finitely many disjoint sets from $\mathcal{C}$.

Briefly: $\mathcal{C}$ is a semiring iff it satisfies Theorem 1.

Note that here $\mathcal{C}$ is not just a set, but a whole family of sets. Recall (Chapter §§1-3) that a set family (family of sets) is a set $\mathcal{M}$ whose members are other sets. If A is a member of $\mathcal{M}$, we call A an $\mathcal{M}$-set and write $A \in \mathcal{M}$ (not $A \subseteq \mathcal{M}$).

Sometimes we use index notation:

$$\mathcal{M}=\{X_i \mid i \in I\}, \tag{7.1.5}$$

briefly

$$\mathcal{M}=\{X_i\}, \tag{7.1.6}$$

where the X_i are $\mathcal{M}$-sets distinguished from each other by the subscripts i varying over some index set I.

A set family $\mathcal{M}=\{X_i\}$ and its union

$$\bigcup_i X_i \tag{7.1.7}$$

are said to be disjoint iff

$$X_i \cap X_j=\emptyset \text{ whenever } i \neq j. \tag{7.1.8}$$

Notation:

$$\bigcup X_i \text{ (disjoint)}. \tag{7.1.9}$$

In our case, $A \in \mathcal{C}$ means that A is a $\mathcal{C}$-set (a member of the semiring $\mathcal{C}$).

The formula

$$(\forall A, B \in \mathcal{C}) \quad A \cap B \in \mathcal{C} \tag{7.1.10}$$

means that the intersection of two $\mathcal{C}$-sets in a $\mathcal{C}$-set itself.

Henceforth, we will often speak of semirings $\mathcal{C}$ in general. In particular, this will apply to the case $\mathcal{C}=\{\text{intervals}\}$. Always keep this case in mind!

Note 1. By Theorem 1, the intervals in E^n form a semiring. So also do the half-open and the half-closed intervals separately (same proof!), but not the open (or closed) ones. (Why?)

Caution. The union and difference of two $\mathcal{C}$-sets need not be a $\mathcal{C}$-set. To remedy this, we now enlarge $\mathcal{C}$.

Definition 2

We say that a set A (from $\mathcal{C}$ or not) is $\mathcal{C}$-simple and write

$$A \in \mathcal{C}_s^{\prime} \tag{7.1.11}$$

iff A is a finite union of disjoint $\mathcal{C}$-sets (such as $A-B$ in Theorem 1).

Thus $\mathcal{C}_s^{\prime}$ is the family of all $\mathcal{C}$-simple sets.

Every $\mathcal{C}$-set is also a $\mathcal{C}_s^{\prime}$-set, i.e., a $\mathcal{C}$-simple one. (Why?) Briefly:

$$\mathcal{C} \subseteq \mathcal{C}_s'. \tag{7.1.12}$$

If $\mathcal{C}$ is the set of all intervals, a $\mathcal{C}$-simple set may look as in Figure 30.

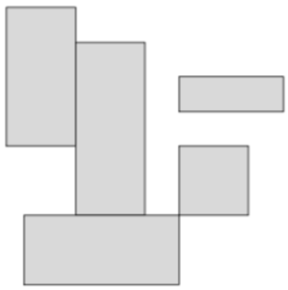

FIGURE 30

Theorem 7.1.2

If $\mathcal{C}$ is a semiring, and if A and B are $\mathcal{C}$-simple, so also are

$$A \cap B, A-B, \text{ and } A \cup B. \tag{7.1.13}$$

In symbols,

$$\left(\forall A, B \in \mathcal{C}_s'\right) \quad A \cap B \in \mathcal{C}_s', A-B \in \mathcal{C}_s', \text{ and } A \cup B \in \mathcal{C}_s'. \tag{7.1.14}$$

Proof

We give a proof outline and suggest the proof as an exercise. Before attempting it, the reader should thoroughly review the laws and problems of Chapter 1, §§1-3.

(1) To prove $A \cap B \in \mathcal{C}_s'$, let

$$A=\bigcup_{i=1}^m A_i (\text{disjoint}) \text{ and } B=\bigcup_{k=1}^n B_k \text{ (disjoint)}, \tag{7.1.15}$$

with $A_i, B_k \in \mathcal{C}$. Verify that

$$A \cap B=\bigcup_{k=1}^n \bigcup_{i=1}^m\left(A_i \cap B_k\right) \text{ (disjoint)}, \tag{7.1.16}$$

and so $A \cap B \in \mathcal{C}_s'$.

(2) Next prove that $A-B \in \mathcal{C}_s'$ if $A \in \mathcal{C}_s'$ and $B \in \mathcal{C}$.

Indeed, if

$$A=\bigcup_{i=1}^m A_i \text{ (disjoint)}, \tag{7.1.17}$$

then

$$A-B=\bigcup_{i=1}^m A_i-B=\bigcup_{i=1}^m\left(A_i-B\right) \text{ (disjoint)}. \tag{7.1.18}$$

Verify and use Definition 2.

(3) Prove that

$$\left(\forall A, B \in \mathcal{C}_s'\right) \quad A-B \in \mathcal{C}_s'; \tag{7.1.19}$$

we suggest the following argument.

Let

$$B=\bigcup_{k=1}^{n} B_k, \quad B_k \in \mathcal{C}. \tag{7.1.20}$$

Then

$$A-B=A-\bigcup_{k=1}^{n} B_k=\bigcap_{k=1}^{n}\left(A-B_k\right) \tag{7.1.21}$$

by duality laws. But $A-B_k$ is $\mathcal{C}$-simple by step (2). Hence so is

$$A-B=\bigcap_{k=1}^{m}\left(A-B_k\right) \tag{7.1.22}$$

by step (1) plus induction.

(4) To prove $A \cup B \in \mathcal{C}_s^{\prime}$, verify that

$$A \cup B=A \cup(B-A), \tag{7.1.23}$$

where $B-A \in \mathcal{C}_s^{\prime}$, by (3).

Note 2. By induction, Theorem 2 extends to any finite number of $\mathcal{C}_s^{\prime}$-sets. It is a kind of "closure law."

We thus briefly say that $\mathcal{C}_s^{\prime}$ is closed under finite unions, intersections, and set differences. Any (nonempty) set family with these properties is called a set ring (see also §3).

Thus Theorem 2 states that if $\mathcal{C}$ is a semiring, then $\mathcal{C}_s^{\prime}$ is a ring.

Caution. An infinite union of $\mathcal{C}$-simple sets need not be $\mathcal{C}$-simple. Yet we may consider such unions, as we do next.

In Corollary 1 below, $\mathcal{C}_s^{\prime}$ may be replaced by any set ring $\mathcal{M}$.

Corollary 7.1.1

If $\{A_n\}$ is a finite or infinite sequence of sets from a semiring $\mathcal{C}$ (or from a ring $\mathcal{M}$ such as $\mathcal{C}_s^{\prime}$), then there is a disjoint sequence of $\mathcal{C}$-simple sets (or $\mathcal{M}$-sets) $B_n \subseteq A_n$ such that

$$\bigcup_n A_n=\bigcup_n B_n. \tag{7.1.24}$$

Proof

Let $B_1=A_1$ and for $n=1,2,\ldots$,

$$B_{n+1}=A_{n+1}-\bigcup_{k=1}^{n} A_k, \quad A_k \in \mathcal{C}. \tag{7.1.25}$$

By Theorem 2, the B_n are $\mathcal{C}$-simple (as are A_{n+1} and $\bigcup_{k=1}^{n} A_k$). Show that they are disjoint (assume the opposite and find a contradiction) and verify that $\bigcup A_n=\bigcup B_n$: If $x \in \bigcup A_n$, take the least n for which $x \in A_n$. Then $n>1$ and

$$x \in A_n-\bigcup_{k=1}^{n-1} A_k=B_n, \tag{7.1.26}$$

or $n=1$ and $x \in A_1=B_1$. □

Note 3. In Corollary 1, $B_n \in \mathcal{C}_s^{\prime}$, i.e., $B_n=\bigcup_{i=1}^{m_n} C_{ni}$ for some disjoint sets $C_{ni} \in \mathcal{C}$. Thus

$$\bigcup_n A_n=\bigcup_n \bigcup_{i=1}^{m_n} C_{ni} \tag{7.1.27}$$

is also a countable disjoint union of $\mathcal{C}$-sets.

II. Recall that the volume of intervals is additive (Problem 9 in Chapter 3, §7). That is, if $A \in \mathcal{C}$ is split into finitely many disjoint subintervals, then vA (the volume of A) equals the sum of the volumes of the parts.

We shall need the following lemma.

lemma 1

Let $X_1, X_2, \ldots, X_m \in \mathcal{C}$ (intervals in E^n). If the X_i are mutually disjoint, then

(i) $\bigcup_{i=1}^m X_i \subseteq Y \in \mathcal{C}$ implies $\sum_{i=1}^m vX_i \leq vY$; and

(ii) $\bigcup_{i=1}^m X_i \subseteq \bigcup_{k=1}^p Y_k$ (with $Y_k \in \mathcal{C}$) implies $\sum_{i=1}^m vX_i \leq \sum_{k=1}^p vY_k$.

Proof

(i) By Theorem 2, the set

$$Y-\bigcup_{i=1}^m X_i \tag{7.1.28}$$

is $\mathcal{C}$-simple; so

$$Y-\bigcup_{i=1}^m X_i=\bigcup_{j=1}^q C_j \tag{7.1.29}$$

for some disjoint intervals C_j. Hence

$$Y=\bigcup X_i \cup \bigcup C_j \text{ (all disjoint).} \tag{7.1.30}$$

Thus by additivity,

$$vY=\sum_{i=1}^m vX_i+\sum_{j=1}^q vC_j \geq \sum_{i=1}^m vX_i, \tag{7.1.31}$$

as claimed.

(ii) By set theory (Problem 9 in Chapter 1, §§1-3),

$$X_i \subseteq \bigcup_{k=1}^p Y_k \tag{7.1.32}$$

implies

$$X_i=X_i \cap \bigcup_{k=1}^p Y_k=\bigcup_{k=1}^p\left(X_i \cap Y_k\right). \tag{7.1.33}$$

If it happens that the Y_k are mutually disjoint also, so certainly are the smaller intervals $X_i \cap Y_k$; so by additivity,

$$vX_i=\sum_{k=1}^p v\left(X_i \cap Y_k\right). \tag{7.1.34}$$

Hence

$$\sum_{i=1}^m vX_i=\sum_{i=1}^m \sum_{k=1}^p v\left(X_i \cap Y_k\right)=\sum_{k=1}^p\left[\sum_{i=1}^m v\left(X_i \cap Y_k\right)\right]. \tag{7.1.35}$$

But by (i),

$$\sum_{i=1}^m v\left(X_i \cap Y_k\right) \leq vY_k \text{ (why?);} \tag{7.1.36}$$

so

$$\sum_{i=1}^{m} vX_i \leq \sum_{k=1}^{p} vY_k, \tag{7.1.37}$$

as required.

If, however, the Y_k are not disjoint, Corollary 1 yields

$$\bigcup Y_k = \bigcup B_k \text{ (disjoint)}, \tag{7.1.38}$$

with

$$Y_k \supseteq B_k = \bigcup_{j=1}^{m_k} C_{kj} \text{(disjoint)}, \quad C_{kj} \in \mathcal{C}. \tag{7.1.39}$$

By (i),

$$\sum_{j=1}^{m_k} vC_{kj} \leq vY_k. \tag{7.1.40}$$

As

$$\bigcup_{i=1}^{m} X_i \subseteq \bigcup_{k=1}^{p} Y_k = \bigcup_{k=1}^{p} B_k = \bigcup_{k=1}^{p} \bigcup_{j=1}^{m_k} C_{kj} \text{ (disjoint)}, \tag{7.1.41}$$

all reduces to the previous disjoint case. □

Corollary 7.1.2

Let $A \in \mathcal{C}_s'$ ($\mathcal{C} =$ intervals in E^n). If

$$A = \bigcup_{i=1}^{m} X_i \text{(disjoint)} = \bigcup_{k=1}^{p} Y_k \text{ (disjoint)} \tag{7.1.42}$$

with $X_i, Y_k \in \mathcal{C}$, then

$$\sum_{i=1}^{m} vX_i = \sum_{k=1}^{p} vY_k. \tag{7.1.43}$$

(Use part (ii) of the lemma twice.)

Thus we can (and do) unambiguously define vA to be either of these sums.

7.1.E: Problems on Intervals and Semirings

Exercise $7.1.E.1$

Complete the proof of Theorem 1 and Note 1.

Exercise $7.1.E.1'$

Prove Theorem 2 in detail.

Exercise $7.1.E.2$

Fill in the details in the proof of Corollary 1.

Exercise $7.1.E.2'$

Prove Corollary 2.

Exercise $7.1.E.3$

Show that, in the definition of a semiring, the condition $\emptyset \in \mathcal{C}$ is equivalent to $\mathcal{C} \neq \emptyset$.
$[\text{Hint: Consider } \emptyset = A - A = \cup_{i=1}^{m} A_i \ (A, A_i \in \mathcal{C}) \text{ to get } \emptyset = A_i \in \mathcal{C}.]$

Exercise $7.1.E.4$

Given a set S, show that the following are semirings or rings.
(a) $\mathcal{C} = \{$ all subsets of $S\}$;
(b) $\mathcal{C} = \{$ all finite subsets of $S\}$;
(c) $\mathcal{C} = \{\emptyset\}$;
(d) $\mathcal{C} = \{\emptyset$ and all singletons in $S\}$.
Disprove it for $\mathcal{C} = \{\emptyset$ and all $two-point$ sets in $S\}$, $S = \{1, 2, 3, \ldots\}$.
In $(a)-(c),$ show that $\mathcal{C}_s' = \mathcal{C}.$ Disprove it for (d).

Exercise $7.1.E.5$

Show that the cubes in $E^n (n > 1)$ do not form a semiring.

Exercise $7.1.E.6$

Using Corollary 2 and the definition thereafter, show that volume is additive for $\mathcal{C}$ -simple sets. That is,

$$\text{if } A = \bigcup_{i=1}^{m} A_i (\text{disjoint}) \text{ then } vA = \sum_{i=1}^{m} vA_i \quad \left(A, A_i \in \mathcal{C}_s'\right). \tag{7.1.E.1}$$

Exercise $7.1.E.7$

Prove the lemma for $\mathcal{C}$-simple sets.
$[\text{Hint: Use Problem 6 and argue as before.}]$

Exercise $7.1.E.8$

Prove that if $\mathcal{C}$ is a semiring, then $\mathcal{C}_s'(\mathcal{C}$ -simple sets $)=\mathcal{C}_s,$ the family of all finite unions of $\mathcal{C}$ -sets (disjoint or not).
$[\text{Hint: Use Theorem 2.}]$

7.2: C σ Cσ -Sets. Countable Additivity. Permutable Series

We now want to further extend the definition of volume by considering countable unions of intervals, called $\mathcal{C}_\sigma$-sets ($\mathcal{C}$ being the semiring of all intervals in E^n).

We also ask, if A is split into countably many such sets, does additivity still hold? This is called countable additivity or σ-additivity (the σ is used whenever countable unions are involved).

We need two lemmas in addition to that of §1.

lemma 1

If B is a nonempty interval in E^n, then given $\varepsilon>0,$ there is an open interval C and a closed one A such that

$$A \subseteq B \subseteq C \tag{7.2.1}$$

and

$$v C-\varepsilon<v B<v A+\varepsilon. \tag{7.2.2}$$

Proof

Let the endpoints of B be

$$\bar{a}=\left(a_1, \ldots, a_n\right) \text{ and } \bar{b}=\left(b_1, \ldots, b_n\right). \tag{7.2.3}$$

For each natural number i, consider the open interval C_i, with endpoints

$$\left(a_1-\frac{1}{i}, a_2-\frac{1}{i}, \ldots, a_n-\frac{1}{i}\right) \text{ and }\left(b_1+\frac{1}{i}, b_2+\frac{1}{i}, \ldots, b_n+\frac{1}{i}\right). \tag{7.2.4}$$

Then $B \subseteq C_i$ and

$$v C_i=\prod_{k=1}^n\left[b_k+\frac{1}{i}-\left(a_k-\frac{1}{i}\right)\right]=\prod_{k=1}^n\left(b_k-a_k+\frac{2}{i}\right). \tag{7.2.5}$$

Making $i \rightarrow \infty$, we get

$$\lim _{i \rightarrow \infty} v C_i=\prod_{k=1}^n\left(b_k-a_k\right)=v B. \tag{7.2.6}$$

(Why?) Hence by the sequential limit definition, given $\varepsilon>0,$ there is a natural i such that

$$v C_i-v B<\varepsilon, \tag{7.2.7}$$

or

$$v C_i-\varepsilon<v B. \tag{7.2.8}$$

As C_i is open and $\supseteq B,$ it is the desired interval C.

Similarly, one finds the closed interval $A \subseteq B$. (Verify!) $\square$

lemma 2

Any open set $G \subseteq E^n$ is a countable union of open cubes A_k and also a disjoint countable union of half-open intervals.

(See also Problem 2 below.)

Proof

If $G=\emptyset,$ take all $A_k=\emptyset$.

If $G \neq \emptyset,$ every point $p \in G$ has a cubic neighborhood

$$C_p \subseteq G, \tag{7.2.9}$$

centered at p (Problem 3 in Chapter 3, §12). By slightly shrinking this C_p, one can make its endpoints rational, with p still in it (but not necessarily its center), and make C_p open, half-open, or closed, as desired. (Explain!)

Choose such a cube C_p for every $p \in G$; so

$$G \subseteq \bigcup_{p \in G} C_p. \tag{7.2.10}$$

But by construction, G contains all C_p, so that

$$G = \bigcup_{p \in G} C_p. \tag{7.2.11}$$

Moreover, because the coordinates of the endpoints of all C_p are rational, the set of ordered pairs of endpoints of the C_p is countable, and thus, while the set of all $p \in G$ is uncountable, the set of distinct C_p is countable. Thus one can put the family of all C_p in a sequence and rename it $\{A_k\}$:

$$G = \bigcup_{k=1}^{\infty} A_k. \tag{7.2.12}$$

If, further, the A_k are half-open, we can use Corollary 1 and Note 3, both from §1, to make the union disjoint (half-open intervals form a semiring!). □

Now let $\mathcal{C}_\sigma$ be the family of all possible countable unions of intervals in E^n, such as G in Lemma 2 (we use $\mathcal{C}_s$ for all finite unions). Thus $A \in \mathcal{C}_\sigma$ means that A is a $\mathcal{C}_\sigma$-set, i.e.,

$$A = \bigcup_{i=1}^{\infty} A_i \tag{7.2.13}$$

for some sequence of intervals $\{A_i\}$. Such are all open sets in E^n, but there also are many other $\mathcal{C}_\sigma$-sets.

We can always make the sequence $\{A_i\}$ infinite (add null sets or repeat a term!).

By Corollary 1 and Note 3 of §1, we can decompose any $\mathcal{C}_\sigma$-set A into countably many disjoint intervals. This can be done in many ways. However, we have the following result.

Theorem 7.2.1

If

$$A = \bigcup_{i=1}^{\infty} A_i \text{ (disjoint)} = \bigcup_{k=1}^{\infty} B_k \text{ (disjoint)} \tag{7.2.14}$$

for some intervals A_i, B_k in E^n, then

$$\sum_{i=1}^{\infty} vA_i = \sum_{k=1}^{\infty} vB_k. \tag{7.2.15}$$

Thus we can (and do) unambiguously define either of these sums to be the volume vA of the $\mathcal{C}_\sigma$-set A.

Proof

We shall use the Heine-Borel theorem (Problem 10 in Chapter 4, §6; review it!).

Seeking a contradiction, let (say)

$$\sum_{i=1}^{\infty} vA_i > \sum_{k=1}^{\infty} vB_k, \tag{7.2.16}$$

so, in particular,

$$\sum_{k=1}^{\infty} vB_k < +\infty. \tag{7.2.17}$$

As

$$\sum_{i=1}^{\infty} vA_i = \lim_{m\to\infty} \sum_{i=1}^{m} vA_i, \tag{7.2.18}$$

there is an integer m for which

$$\sum_{i=1}^{m} vA_i > \sum_{k=1}^{\infty} vB_k. \tag{7.2.19}$$

We fix that m and set

$$2\varepsilon = \sum_{i=1}^{m} vA_i - \sum_{k=1}^{\infty} vB_k > 0. \tag{7.2.20}$$

Dropping "empties" (if any), we assume $A_i \neq \emptyset$ and $B_k \neq \emptyset$.

Then Lemma 1 yields open intervals $Y_k \supseteq B_k,$ with

$$vB_k > vY_k - \frac{\varepsilon}{2^k}, \quad k = 1, 2, \ldots, \tag{7.2.21}$$

and closed ones $X_i \subseteq A_i,$ with

$$vX_i + \frac{\varepsilon}{m} > vA_i; \tag{7.2.22}$$

so

$$\begin{aligned} 2\varepsilon = \sum_{i=1}^{m} vA_i - \sum_{k=1}^{\infty} vB_k \; &< \sum_{i=1}^{m}\left(vX_i + \frac{\varepsilon}{m}\right) - \sum_{k=1}^{\infty}\left(vY_k - \frac{\varepsilon}{2^k}\right) \\ &= \sum_{i=1}^{m} vX_i - \sum_{k=1}^{\infty} vY_k + 2\varepsilon. \end{aligned}$$

Thus

$$\sum_{i=1}^{m} vX_i > \sum_{k=1}^{\infty} vY_k. \tag{7.2.23}$$

(Explain in detail!)

Now, as

$$X_i \subseteq A_i \subseteq A = \bigcup_{k=1}^{\infty} B_k \subseteq \bigcup_{k=1}^{\infty} Y_k, \tag{7.2.24}$$

each of the closed intervals X_i is covered by the open sets Y_k.

By the Heine-Borel theorem, $\bigcup_{i=1}^{m} X_i$ is already covered by a finite number of the Y_k, say,

$$\bigcup_{i=1}^{m} X_i \subseteq \bigcup_{k=1}^{p} Y_k. \tag{7.2.25}$$

The X_i are disjoint, for even the larger sets A_i are. Thus by Lemma 1(ii) in §1,

$$\sum_{i=1}^{m} vX_i \leq \sum_{k=1}^{p} vY_k \leq \sum_{k=1}^{\infty} vY_k, \tag{7.2.26}$$

contrary to (1). This contradiction completes the proof. □

Corollary 7.2.1

If

$$A=\bigcup_{k=1}^{\infty} B_{k} \text{ (disjoint)} \tag{7.2.27}$$

for some intervals B_k, then

$$vA=\sum_{k=1}^{\infty} vB_{k}. \tag{7.2.28}$$

Indeed, this is simply the definition of vA contained in Theorem 1.

Note 1. In particular, Corollary 1 holds if A is an interval itself. We express this by saying that the volume of intervals is σ-additive or countably additive. This also shows that our previous definition of volume (for intervals) agrees with the definition contained in Theorem 1 (for $\mathcal{C}_{\sigma}$-sets).

Note 2. As all open sets are $\mathcal{C}_{\sigma}$-sets (Lemma 2), volume is now defined for any open set $A \subseteq E^{n}$ (in particular, for $A=E^{n}$).

Corollary 7.2.2

If A_i, B_k are intervals in E^n, with

$$\bigcup_{i=1}^{\infty} A_{i} \subseteq \bigcup_{k=1}^{\infty} B_{k}, \tag{7.2.29}$$

then provided the A_i are mutually disjoint,

$$\sum_{i=1}^{\infty} vA_{i} \leq \sum_{k=1}^{\infty} vB_{k}. \tag{7.2.30}$$

Proof

The proof is as in Theorem 1 (but the B_k need not be disjoint here).

Corollary 7.2.3 ("σ-subadditivity" of the volume)

If

$$A \subseteq \bigcup_{k=1}^{\infty} B_{k}, \tag{7.2.31}$$

where $A \in \mathcal{C}_{\sigma}$ and the B_k are intervals in E^n, then

$$vA \leq \sum_{k=1}^{\infty} vB_{k}. \tag{7.2.32}$$

Proof

Set

$$A=\bigcup_{i=1}^{\infty} A_{i}(\text{ disjoint }), A_{i} \in \mathcal{C}, \tag{7.2.33}$$

and use Corollary 2. $\square$

Corollary 7.2.4 ("monotonicity")

If $A, B \in \mathcal{C}_\sigma$, with

$$A \subseteq B, \tag{7.2.34}$$

then

$$vA \leq vB. \tag{7.2.35}$$

("Larger sets have larger volumes.")

This is simply Corollary 3, with $\bigcup_k B_k = B$.

Corollary 7.2.5

The volume of all of E^n is ∞ (we write ∞ for $+\infty$).

Proof

We have $A \subseteq E^n$ for any interval A.

Thus, by Corollary 4, $vA \leq vE^n$.

As vA can be chosen arbitrarily large, vE^n must be infinite. □

Corollary 7.2.6

For any countable set $A \subset E^n, vA = 0$. In particular, $v\emptyset = 0$.

Proof

First let $A = \{\bar{a}\}$ be a singleton. Then we may treat A as a degenerate interval $[\bar{a}, \bar{a}]$. As all its edge lengths are 0, we have $vA = 0$.

Next, if $A = \{\bar{a}_1, \bar{a}_2, \ldots\}$ is a countable set, then

$$A = \bigcup_k \{\bar{a}_k\}; \tag{7.2.36}$$

so

$$vA = \sum_k v\{\bar{a}_k\} = 0 \tag{7.2.37}$$

by Corollary 1.

Finally, $\emptyset$ is the degenerate open interval $(\bar{a}, \bar{a})$; so $v\emptyset = 0$. □

Note 3. Actually, all these propositions hold also if all sets involved are $\mathcal{C}_\sigma$-sets, not just intervals (split each $\mathcal{C}_\sigma$-set into disjoint intervals!).

Permutable Series. Since σ-additivity involves countable sums, it appears useful to generalize the notion of a series.

We say that a series of constants,

$$\sum a_n, \tag{7.2.38}$$

is permutable iff it has a definite (possibly infinite) sum obeying the general commutative law:

Given any one-one map

$$u : N \xleftrightarrow{\text{onto}} N \tag{7.2.39}$$

(N = the naturals), we have

$$\sum_{n} a_{n}=\sum_{n} a_{u_{n}}, \tag{7.2.40}$$

where $u_{n}=u(n)$.

(Such are all positive and all absolutely convergent series in a complete space E; see Chapter 4, §13.) If the series is permutable, the sum does not depend on the choice of the map u.

Thus, given any $u: N \stackrel{\text{onto}}{\longleftrightarrow} J$ (where J is a countable index set) and a set

$$\left\{a_{i} \mid i \in J\right\} \subseteq E \tag{7.2.41}$$

(where E is E^{*} or a normed space), we can define

$$\sum_{i \in J} a_{i}=\sum_{n=1}^{\infty} a_{u_{n}} \tag{7.2.42}$$

if $\sum_{n} a_{u_{n}}$ is permutable.

In particular, if

$$J=N \times N \tag{7.2.43}$$

(a countable set, by Theorem 1 in Chapter 1, §9), we call

$$\sum_{i \in J} a_{i} \tag{7.2.44}$$

a double series, denoted by symbols like

$$\sum_{n, k} a_{k n} \quad(k, n \in N). \tag{7.2.45}$$

Note that

$$\sum_{i \in J}\left|a_{i}\right| \tag{7.2.46}$$

is always defined (being a positive series).

If

$$\sum_{i \in J}\left|a_{i}\right|<\infty, \tag{7.2.47}$$

we say that $\sum_{i \in J} a_{i}$ converges absolutely.

For a positive series, we obtain the following result.

Theorem 7.2.2

(i) All positive series in E^{*} are permutable.

(ii) For positive double series in E^{*}, we have

$$\sum_{n, k=1}^{\infty} a_{n k}=\sum_{n=1}^{\infty}\left(\sum_{k=1}^{\infty} a_{n k}\right)=\sum_{k=1}^{\infty}\left(\sum_{n=1}^{\infty} a_{n k}\right). \tag{7.2.48}$$

Proof

(i) Let

$$s=\sum_{n=1}^{\infty} a_{n} \text{ and } s_{m}=\sum_{n=1}^{m} a_{n} \quad\left(a_{n} \geq 0\right). \tag{7.2.49}$$

Then clearly

$$s_{m+1} = s_m + a_{m+1} \geq s_m; \tag{7.2.50}$$

i.e., $\{s_m\} \uparrow$, and so

$$s = \lim_{m\to\infty} s_m = \sup_m s_m \tag{7.2.51}$$

by Theorem 3 in Chapter 3, §15.

Hence s certainly does not exceed the lub of all possible sums of the form

$$\sum_{i\in I} a_i, \tag{7.2.52}$$

where I is a finite subset of N (the partial sums s_m are among them). Thus

$$s \leq \sup \sum_{i\in I} a_i, \tag{7.2.53}$$

over all finite sets $I \subset N$.

On the other hand, every such $\sum_{i\in I} a_i$ is exceeded by, or equals, some s_m. Hence in (4), the reverse inequality holds, too, and so

$$s = \sup \sum_{i\in I} a_i. \tag{7.2.54}$$

But sup $\sum_{i\in I} a_i$ clearly does not depend on any arrangement of the a_{i}. Therefore, the series $\sum a_n$ is permutable, and assertion (i) is proved.

Assertion (ii) follows similarly by considering sums of the form $\sum_{i\in I} a_i$ where I is a finite subset of $N \times N$, and showing that the lub of such sums equals each of the three expressions in (3). We leave it to the reader. $\square$

A similar formula holds for absolutely convergent series (see Problems).

7.2.E: Problems on C σ Cσ -Sets, σ σ -Additivity, and Permutable Series

Exercise $7.2.E.1$

Fill in the missing details in the proofs of this section.

Exercise $7.2.E.1'$

Prove Note 3.

Exercise $7.2.E.2$

Show that every open set $A \neq \emptyset$ in E^n is a countable union of disjoint half-open cubes.
[Outline: For each natural m, show that E^n is split into such cubes of edge length 2^{-m} by the hyperplanes

$$x_k = \frac{i}{2^m} \quad i = 0, \pm 1, \pm 2, \ldots ; k = 1, 2, \ldots, n, \tag{7.2.E.1}$$

and that the family $\mathcal{C}_m$ of such cubes is countable.
For $m > 1$, let $C_{m1}, C_{m2}, \ldots$ be the sequence of those cubes from $\mathcal{C}_m$ (if any) that lie in A but not in any cube C_{sj} with $s < m$.
As A is open, $x \in A$ iff $x \in$ some C_{mj}.]

Exercise $7.2.E.3$

Prove that any open set $A \subseteq E^1$ is a countable union of disjoint (possibly infinite) open intervals.
[Hint: By Lemma 2, $A = \bigcup_n (a_n, b_n)$. If, say, (a_1, b_1) overlaps with some (a_m, b_m), replace both by their union. Continue inductively.

Exercise $7.2.E.4$

Prove that $\mathcal{C}_\sigma$ is closed under finite intersections and countable unions.

Exercise $7.2.E.5$

(i) Find $A, B \in \mathcal{C}_\sigma$ such that $A - B \notin \mathcal{C}_\sigma$
(ii) Show that $\mathcal{C}_\sigma$ is not a semiring.
[Hint: Try $A = E^1, B = R$ (the rationals).]
Note. In the following problems, J is countably infinite, $a_i \in E(E$ complete$)$.

Exercise $7.2.E.6$

Prove that

$$\sum_{i \in J} |a_i| < \infty \tag{7.2.E.2}$$

iff for every $\varepsilon > 0$, there is a finite set

$$F \subset J \quad (F \neq \emptyset) \tag{7.2.E.3}$$

such that

$$\sum_{i\in I}|a_i|<\varepsilon \tag{7.2.E.4}$$

for every finite $I\subset J-F$.
[Outline: By Theorem 2, fix $u: N onto \Longleftrightarrow J$ with

$$\sum_{i\in J}|a_i|=\sum_{n=1}^{\infty}|a_{u_n}|. \tag{7.2.E.5}$$

By Cauchy's criterion,

$$\sum_{n=1}^{\infty}|a_{u_n}|<\infty \tag{7.2.E.6}$$

iff

$$(\forall\varepsilon>0)(\exists q)(\forall n>m>q)\quad \sum_{k=m}^{n}|a_{u_k}|<\varepsilon. \tag{7.2.E.7}$$

Let $F=\{u_1,\ldots,u_q\}$. If I is as above,

$$(\exists n>m>q)\quad \{u_m,\ldots,u_n\}\supseteq I; \tag{7.2.E.8}$$

so

$$\sum_{i\in I}|a_i|\leq\sum_{k=m}^{n}|a_{u_k}|<\varepsilon.\Bigg] \tag{7.2.E.9}$$

Exercise $7.2.E.7$

Prove that if

$$\sum_{i\in J}|a_i|<\infty, \tag{7.2.E.10}$$

then for every $\varepsilon>0,$ there is a finite $F\subset J(F\neq\emptyset)$ such that

$$\left|\sum_{i\in J}a_i-\sum_{i\in K}a_i\right|<\varepsilon \tag{7.2.E.11}$$

for each finite $K\supset F(K\subset J)$.
[Hint: Proceed as in Problem 6, with $I=K-F$ and q so large that

$$\left|\sum_{i\in J}a_i-\sum_{i\in F}a_i\right|<\frac{1}{2}\varepsilon\quad\text{and}\quad\left|\sum_{i\in F}a_i\right|<\frac{1}{2}\varepsilon.] \tag{7.2.E.12}$$

Exercise $7.2.E.8$

Show that if

$$J=\bigcup_{n=1}^{\infty} I_n (\text{disjoint}), \tag{7.2.E.13}$$

then

$$\sum_{i \in J}\left|a_i\right|=\sum_{n=1}^{\infty} b_n, \text{ where } b_n=\sum_{i \in I_n}\left|a_i\right| . \tag{7.2.E.14}$$

(Use Problem 8' below.)

Exercise $7.2.E.8'$

Show that

$$\sum_{i \in J}\left|a_i\right|=\sup_F \sum_{i \in F}\left|a_i\right| \tag{7.2.E.15}$$

over all finite sets $F \subset J(F \neq \emptyset)$.
$[\text{Hint: Argue as in Theorem 2.}]$

Exercise $7.2.E.9$

Show that if $\emptyset \neq I \subseteq J,$ then

$$\sum_{i \in I}\left|a_i\right| \leq \sum_{i \in J}\left|a_i\right| . \tag{7.2.E.16}$$

$[\text{Hint: Use Problem } 8' \text{ and Corollary 2 of Chapter 2, §§} 8-9.]$

Exercise $7.2.E.10$

Continuing Problem 8, prove that if

$$\sum_{i \in J}\left|a_i\right|=\sum_{n=1}^{\infty} b_n<\infty, \tag{7.2.E.17}$$

then

$$\sum_{i \in J} a_i=\sum_{n=1}^{\infty} c_n \text{ with } c_n=\sum_{i \in I_n} a_i. \tag{7.2.E.18}$$

[Outline: By Problem 9,

$$(\forall n) \sum_{i \in I_n}\left|a_i\right|<\infty; \tag{7.2.E.19}$$

so

$$c_n=\sum_{i \in I_n} a_i \tag{7.2.E.20}$$

and

$$\sum_{n=1}^{\infty} c_n \tag{7.2.E.21}$$

converge absolutely.
Fix ε and F as in Problem 7. Choose the largest $q \in N$ with

$$F \cap I_q \neq \emptyset \tag{7.2.E.22}$$

(why does it exist?), and fix any $n > q$. By Problem 7, $(\forall k \leq n)$

$$(\forall k \leq n)\,(\exists \text{ finite } F_k | J \supseteq F_k \supseteq F \cap I_q)$$
$$(\forall \text{ finite } H_k | I_k \supseteq H_k \supseteq F_k) \left| \sum_{i \in H_k} a_i - \sum_{k=1}^{n} c_k \right| < \frac{1}{2}\varepsilon \,.$$

(Explain!) Let

$$K = \bigcup_{k=1}^{n} H_k; \tag{7.2.E.23}$$

so

$$\left| \sum_{k=1}^{n} c_k - \sum_{i \in J} a_i \right| < \varepsilon \tag{7.2.E.24}$$

and $K \supset F.$ By Problem 7,

$$\left| \sum_{i \in K} a_i - \sum_{i \in J} a_i \right| < \varepsilon. \tag{7.2.E.25}$$

Deduce

$$\left| \sum_{k=1}^{n} c_k - \sum_{i \in J} a_i \right| < 2\varepsilon. \tag{7.2.E.26}$$

$\text{Let } n \to \infty; \text{ then } \varepsilon \to 0.]$

Exercise $7.2.E.11$

(Double series.) Prove that if one of the expressions

$$\sum_{n,k=1}^{\infty} |a_{nk}|, \quad \sum_{n=1}^{\infty}\left(\sum_{k=1}^{\infty} |a_{nk}|\right), \quad \sum_{k=1}^{\infty}\left(\sum_{n=1}^{\infty} |a_{nk}|\right) \tag{7.2.E.27}$$

is finite, so are the other two, and

$$\sum_{n,k} a_{nk} = \sum_{n}\left(\sum_{k} a_{nk}\right) = \sum_{k}\left(\sum_{n} a_{nk}\right), \tag{7.2.E.28}$$

with all series involved absolutely convergent.
[Hint: Use Problems 8 and 10, with $J = N \times N$,

$$I_n = \{(n,k) \in J \mid k = 1,2,\ldots\} \text{ for each } n; \tag{7.2.E.29}$$

so

$$b_n = \sum_{k=1}^{\infty} |a_{nk}| \text{ and } c_n = \sum_{k=1}^{\infty} a_{nk}. \tag{7.2.E.30}$$

Thus obtain

$$\sum_{n,k} a_{nk} = \sum_{n}\sum_{k} a_{nk}. \tag{7.2.E.31}$$

Similarly,

$$\sum_{n,k} a_{nk} = \sum_{k}\sum_{n} a_{n,k}. \Bigg] \tag{7.2.E.32}$$

7.3: More on Set Families

Lebesgue extended his theory far beyond $\mathcal{C}_\sigma$-sets. For a deeper insight, we shall consider set families in more detail, starting with set rings. First, we rephrase and supplement our former definition of that notion, given in §1.

Definition 1

A family $\mathcal{M}$ of subsets of a set S is a ring or set ring (in S) iff

(i) $\emptyset \in \mathcal{M}$, i.e., the empty set is a member; and

(ii) $\mathcal{M}$ is closed under finite unions and differences:

$$(\forall X, Y \in \mathcal{M}) \quad X \cup Y \in \mathcal{M} \text{ and } X - Y \in \mathcal{M}. \tag{7.3.1}$$

(For intersections, see Theorem 1 below.)

If $\mathcal{M}$ is also closed under countable unions, we call it a σ-ring (in S). Then

$$\bigcup_{i=1}^{\infty} X_i \in \mathcal{M} \tag{7.3.2}$$

whenever

$$X_i \in \mathcal{M} \text{ for } i = 1, 2, \ldots. \tag{7.3.3}$$

If S itself is a member of a ring (σ-ring) $\mathcal{M}$, we call $\mathcal{M}$ a set field (σ-field), or a set algebra (σ-algebra), in S.

Note that S is only a member of $\mathcal{M}$, $S \in \mathcal{M}$, not to be confused with $\mathcal{M}$ itself.

The family of all subsets of S (the so-called power set of S) is denoted by 2^S or $\mathcal{P}(S)$.

Examples

(a) In any set S, 2^S is a σ-field. (Why?)

(b) The family $\{\emptyset\}$, consisting of $\emptyset$ alone, is a σ-ring; $\{\emptyset, S\}$ is a σ-field in S. (Why?)

(c) The family of all finite (countable) subsets of S is a ring (σ-ring) in S.

(d) For any semiring $\mathcal{C}$, $\mathcal{C}_s'$ is a ring (Theorem 2 in §1). Not so for $\mathcal{C}_\sigma$ (Problem 5 in §2).

Theorem 7.3.1

Any set ring is closed under finite intersections.

A σ-ring is closed under countable intersections.

Proof

Let $\mathcal{M}$ be a σ-ring (the proof for rings is similar).

Given a sequence $\{A_n\} \subseteq \mathcal{M}$, we must show that $\bigcap_n A_n \in \mathcal{M}$.

Let

$$U = \bigcup_n A_n. \tag{7.3.4}$$

By Definition 1,

$$U \in \mathcal{M} \text{ and } U - A_n \in \mathcal{M}, \tag{7.3.5}$$

as $\mathcal{M}$ is closed under these operations. Hence

$$\bigcup_n (U - A_n) \in \mathcal{M} \tag{7.3.6}$$

and

$$U - \bigcup_n (U - A_n) \in \mathcal{M}, \tag{7.3.7}$$

or, by duality,

$$\bigcap_n [U - (U - A_n)] \in \mathcal{M}, \tag{7.3.8}$$

i.e.,

$$\bigcap_n A_n \in \mathcal{M}. \quad \square \tag{7.3.9}$$

Corollary 7.3.1

Any set ring (field, σ-ring, σ-field) is also a semiring.

Indeed, by Theorem 1 and Definition 1, if $\mathcal{M}$ is a ring, then $\emptyset \in \mathcal{M}$ and

$$(\forall A, B \in \mathcal{M}) \quad A \cap B \in \mathcal{M} \text{ and } A - B \in \mathcal{M}. \tag{7.3.10}$$

Here we may treat $A - B$ as $(A - B) \cup \emptyset,$ a union of two disjoint $\mathcal{M}$-sets. Thus $\mathcal{M}$ has all properties of a semiring.

Similarly for σ-rings, fields, etc.

In §1 we saw that any semiring $\mathcal{C}$ can be enlarged to become a ring, $\mathcal{C}_s'$. More generally, we obtain the following result.

Theorem 7.3.2

For any set family $\mathcal{M}$ in a space S $\left(\mathcal{M} \subseteq 2^S\right)$, there is a unique "smallest" set ring $\mathcal{R}$ such that

$$\mathcal{R} \supseteq \mathcal{M} \tag{7.3.11}$$

("smallest" in the sense that

$$\mathcal{R} \subseteq \mathcal{R}' \tag{7.3.12}$$

for any other ring $\mathcal{R}'$ with $\mathcal{R}' \supseteq \mathcal{M}$).

The $\mathcal{R}$ of Theorem 2 is called the ring generated by $\mathcal{M}$. Similarly for σ-rings, fields, and σ-fields in S.

Proof

We give the proof for σ-fields; it is similar in the other cases.

There surely are σ-fields in S that contain $\mathcal{M}$; e.g., take 2^S. Let $\{\mathcal{R}_i\}$ be the family of all possible σ-fields in S such that $\mathcal{R}_i \supseteq \mathcal{M}.$ Let

$$\mathcal{R} = \bigcap_i \mathcal{R}_i. \tag{7.3.13}$$

We shall show that this $\mathcal{R}$ is the required "smallest" σ-field containing $\mathcal{M}$.

Indeed, by assumption,

$$\mathcal{M} \subseteq \bigcap_i \mathcal{R}_i = \mathcal{R}. \tag{7.3.14}$$

We now verify the σ-field properties for $\mathcal{R}$.

(1) We have that

$$(\forall i) \quad \emptyset \in \mathcal{R}_i \text{ and } S \in \mathcal{R}_i \tag{7.3.15}$$

(for $\mathcal{R}_i$ is a σ-field, by assumption). Hence

$$\emptyset \in \bigcap_i \mathcal{R}_i = \mathcal{R}. \tag{7.3.16}$$

Similarly, $S \in \mathcal{R}$. Thus

$$\emptyset, S \in \mathcal{R}. \tag{7.3.17}$$

(2) Suppose

$$X, Y \in \mathcal{R} = \bigcap_i \mathcal{R}_i. \tag{7.3.18}$$

Then X, Y are in every $\mathcal{R}_i$, and so is $X - Y$. Hence $X - Y$ is in

$$\bigcap_i \mathcal{R}_i = \mathcal{R}. \tag{7.3.19}$$

Thus $\mathcal{R}$ is closed under differences.

(3) Take any sequence

$$\{A_n\} \subseteq \mathcal{R} = \bigcap_i \mathcal{R}_i. \tag{7.3.20}$$

Then all A_n are in each $\mathcal{R}_i$. $\bigcup_n A_n$ is in each $\mathcal{R}_i$; so

$$\bigcup_n A_n \in \mathcal{R}. \tag{7.3.21}$$

Thus $\mathcal{R}$ is closed under countable unions.

We see that $\mathcal{R}$ is indeed a σ-field in S, with $\mathcal{M} \subseteq \mathcal{R}$. As $\mathcal{R}$ is the intersection of all $\mathcal{R}_i$(i.e., all σ-fields $\supseteq \mathcal{M}$), we have

$$(\forall i) \quad \mathcal{R} \subseteq \mathcal{R}_i; \tag{7.3.22}$$

so $\mathcal{R}$ is the smallest of such σ-fields.

It is unique; for if $\mathcal{R}'$ is another such σ-field, then

$$\mathcal{R} \subseteq \mathcal{R}' \subseteq \mathcal{R} \tag{7.3.23}$$

(as both $\mathcal{R}$ and $\mathcal{R}'$ are "smallest"); so

$$\mathcal{R} = \mathcal{R}'. \quad \square \tag{7.3.24}$$

Note 1. This proof also shows that the intersection of any family $\{\mathcal{R}_i\}$ of σ-fields is a σ-field. Similarly for σ-rings, fields, and rings.

Corollary 7.3.2

The ring $\mathcal{R}$ generated by a semiring $\mathcal{C}$ coincides with

$$\mathcal{C}_s = \{\text{all finite unions of } \mathcal{C} - \text{sets}\} \tag{7.3.25}$$

and with

$$\mathcal{C}_s' = \{\text{disjoint finite unions of } \mathcal{C} - \text{sets}\}. \tag{7.3.26}$$

Proof

By Theorem 2 in §1, $\mathcal{C}_s'$ is a ring $\supseteq \mathcal{C}$; and

$$\mathcal{C}_s' \subseteq \mathcal{C}_s \subseteq \mathcal{R} \tag{7.3.27}$$

(for $\mathcal{R}$ is closed under finite unions, being a ring $\supseteq \mathcal{C}$).

Moreover, as $\mathcal{R}$ is the smallest ring $\supseteq \mathcal{C}$, we have

$$\mathcal{R} \subseteq \mathcal{C}_s' \subseteq \mathcal{C}_s \subseteq \mathcal{R}. \tag{7.3.28}$$

Hence

$$\mathcal{R} = \mathcal{C}_s' = \mathcal{C}_s, \tag{7.3.29}$$

as claimed. □

It is much harder to characterize the σ-ring generated by a semiring. The following characterization proves useful in theory and as an exercise.

Theorem 7.3.3

The σ-ring $\mathcal{R}$ generated by a semiring $\mathcal{C}$ coincides with the smallest set family $\mathcal{D}$ such that

(i) $\mathcal{D} \supseteq \mathcal{C}$;

(ii) $\mathcal{D}$ is closed under countable disjoint unions;

(iii) $J - X \in \mathcal{D}$ whenever $X \in \mathcal{D}, J \in \mathcal{C},$ and $X \subseteq J$.

Proof

We give a proof outline, leaving the details to the reader.

(1) The existence of a smallest such $\mathcal{D}$ follows as in Theorem 2. Verify!

(2) Writing briefly AB for $A \cap B$ and A' for $-A$, prove that

$$(A - B)C = A - (AC' \cup BC). \tag{7.3.30}$$

(3) For each $I \in \mathcal{D},$ set

$$\mathcal{D}_I = \{A \in \mathcal{D} | AI \in \mathcal{D}, A - I \in \mathcal{D}\}. \tag{7.3.31}$$

Then prove that if $I \in \mathcal{C},$ the set family $\mathcal{D}_I$ has the properties (i)-(iii) specified in the theorem. (Use the set identity (2) for property (iii).)

Hence by the minimality of $\mathcal{D}, \mathcal{D} \subseteq \mathcal{D}_I$. Therefore,

$$(\forall A \in \mathcal{D})(\forall I \in \mathcal{C}) \quad AI \in \mathcal{D} \text{ and } A - I \in \mathcal{D}. \tag{7.3.32}$$

(4) Using this, show that $\mathcal{D}_I$ satisfies (i)-(iii) for any$I \in \mathcal{D}$.

Deduce

$$\mathcal{D} \subseteq \mathcal{D}_I; \tag{7.3.33}$$

so $\mathcal{D}$ is closed under finite intersections and differences.

Combining with property (ii), show that $\mathcal{D}$ is a σ-ring (see Problem 12 below).

By its minimality, $\mathcal{D}$ is the smallest σ-ring $\supseteq \mathcal{C}$ (for any other such σ-ring clearly satisfies (i)-(iii)).

Thus $\mathcal{D} = \mathcal{R},$ as claimed. □

Definition 2

Given a set family $\mathcal{M}$, we define (following Hausdorff)

(a) $\mathcal{M}_\sigma = \{$all countable unions of $\mathcal{M}$-sets$\}$ (cf. $\mathcal{C}_\sigma$ in §2);

(b) $\mathcal{M}_\delta = \{$all countable intersections of $\mathcal{M}$-sets$\}$.

We use $\mathcal{M}_s$ and $\mathcal{M}_d$ for similar notions, with "countable" replaced by "finite."

Clearly,

$$\mathcal{M}_\sigma \supseteq \mathcal{M}_s \supseteq \mathcal{M} \tag{7.3.34}$$

and

$$\mathcal{M}_\delta \supseteq \mathcal{M}_d \supseteq \mathcal{M}. \tag{7.3.35}$$

Why?

Note 2. Observe that $\mathcal{M}$ is closed under finite (countable) unions iff

$$\mathcal{M} = \mathcal{M}_s \left(\mathcal{M} = \mathcal{M}_\sigma\right). \tag{7.3.36}$$

Verify! Interpret $\mathcal{M} = \mathcal{M}_d \left(\mathcal{M} = \mathcal{M}_\sigma\right)$ similarly.

In conclusion, we generalize Theorem 1 in §1.

Definition 3

The product

$$\mathcal{M} \dot{\times} \mathcal{N} \tag{7.3.37}$$

of two set families $\mathcal{M}$ and $\mathcal{N}$ is the family of all sets of the form

$$A \times B, \tag{7.3.38}$$

with $A \in \mathcal{M}$ and $B \in \mathcal{N}$.

(The dot in $\dot{\times}$ is to stress that $\mathcal{M} \dot{\times} \mathcal{N}$ is not really a Cartesian product.)

Theorem 7.3.4

If $\mathcal{M}$ and $\mathcal{N}$ are semirings, so is $\mathcal{M} \dot{\times} \mathcal{N}$.

The proof runs along the same lines as that of Theorem 1 in §1, via the set identities

$$(X \times Y) \cap (X' \times Y') = (X \cap X') \times (Y \cap Y') \tag{7.3.39}$$

and

$$(X \times Y) - (X' \times Y') = [(X - X') \times Y] \cup [(X \cap X') \times (Y - Y')]. \tag{7.3.40}$$

Proof

The details are left to the reader.

Note 3. As every ring is a semiring (Corollary 1), the product of two rings (fields, σ-rings, σ-fields) is a semiring. However, see Problem 6 below.

7.3.E: Problems on Set Families

Exercise $7.3.E.1$

1. Verify Examples (a),(b), and (c).

Exercise $7.3.E.1'$

Prove Theorem 1 for rings.

Exercise $7.3.E.2$

Show that in Definition 1 "$\emptyset \in \mathcal{M}$" may be replaced by "$\mathcal{M} \neq \emptyset$."

[Hint: $\emptyset=A-A$.]

Exercise $7.3.E.3$

⇒ Prove that $\mathcal{M}$ is a field $(\sigma$-field if $\mathcal{M} \neq \emptyset, \mathcal{M}$ is closed under finite (countable) unions, and

$$(\forall A \in \mathcal{M}) \quad-A \in \mathcal{M}. \tag{7.3.E.1}$$

[Hint: $A-B=-(-A \cup B) ; S=-\emptyset$.]

Exercise $7.3.E.4$

Prove Theorem 2 for set fields.

Exercise $7.3.E.*4'$

Does Note 1 apply to semirings?

Exercise $7.3.E.5$

Prove Note 2.

Exercise $7.3.E.5'$

Prove Theorem 3 in detail.

Exercise $7.3.E.6$

Prove Theorem 4 and show that the product $\mathcal{M} \dot{\times} \mathcal{N}$ of two rings need not be a ring.
[Hint: Let $S=E^1$ and $\mathcal{M}=\mathcal{N}=2^S$. Take A, B as in Theorem 1 of §1. Verify that $A-B \notin \mathcal{M}, \mathcal{M} \dot{\times} \mathcal{N}$.]

Exercise $7.3.E.7$

⇒ Let $\mathcal{R}, \mathcal{R}'$ be the rings (σ-rings, fields, σ-fields) generated by $\mathcal{M}$ and $\mathcal{N}$, respectively. Prove the following.
(i) If $\mathcal{M} \subseteq \mathcal{N}$, then $\mathcal{R} \subseteq \mathcal{R}'$.
(ii) If $\mathcal{M} \subseteq \mathcal{N} \subseteq \mathcal{R}$, then $\mathcal{R}=\mathcal{R}'$.
(iii) If

$$\mathcal{M}=\{\text{open intervals in } E^n\} \tag{7.3.E.2}$$

and

$$\mathcal{N}=\{\text{all open sets in } E^{n}\}, \tag{7.3.E.3}$$

then $\mathcal{R}=\mathcal{R}^{\prime}$.
[Hint: Use Lemma 2 in §2 for (iii). Use the minimality of $\mathcal{R}$ and $\mathcal{R}^{\prime}$.]

Exercise $7.3.E.8$

Is any of the following a semiring, ring, σ-ring, field, or σ-field? Why?
(a) All infinite intervals in E^{1}.
(b) All open sets in a metric space (S, ρ).
(c) All closed sets in (S, ρ).
(d) All "clopen" sets in (S, ρ).
(e) $\left\{X \in 2^{S} \mid-X \text{ finite}\right\}$.
(f) $\left\{X \in 2^{S} \mid-X \text{ countable}\right\}$.

Exercise $7.3.E.9$

$\Rightarrow$ Prove that for any sequence $\{A_{n}\}$ in a ring $\mathcal{R}$, there is
(a) an expanding sequence $\{B_{n}\} \subseteq \mathcal{R}$ such that

$$(\forall n) \quad B_{n} \supseteq A_{n} \tag{7.3.E.4}$$

and

$$\bigcup_{n} B_{n}=\bigcup_{n} A_{n}; \text{ and} \tag{7.3.E.5}$$

(b) a contracting sequence $C_{n} \subseteq A_{n}$, with

$$\bigcap_{n} C_{n}=\bigcap_{n} A_{n}. \tag{7.3.E.6}$$

(The latter holds in semirings, too.)
[Hint: Set $B_{n}=\bigcup_{1}^{n} A_{k}, C_{n}=\bigcap_{1}^{n} A_{k}$.]

Exercise $7.3.E.10$

$\Rightarrow$ The symmetric difference, $A \triangle B$, of two sets is defined

$$A \triangle B=(A-B) \cup(B-A). \tag{7.3.E.7}$$

Inductively, we also set

$$\triangle_{k=1}^{1} A_{k}=A_{1} \tag{7.3.E.8}$$

and

$$\triangle_{k=1}^{n+1} A_{k}=\left(\triangle_{k=1}^{n} A_{k}\right) \triangle A_{n+1}. \tag{7.3.E.9}$$

Show that symmetric differences

(i) are commutative,
(ii) are associative, and
(iii) satisfy the distributive law:

$$(A \triangle B) \cap C=(A \cap C) \triangle(B \cap C). \tag{7.3.E.10}$$

[Hint for (ii): Set $A'=-A, A-B=A \cap B'$. Expand $(A \triangle B) \triangle C$ into an expression symmetric with respect to $A, B,$ and C.]

Exercise $7.3.E.11$

Prove that $\mathcal{M}$ is a ring iff
(i) $\emptyset \in \mathcal{M}$;
(ii) $(\forall A, B \in \mathcal{M}) A \triangle B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$ (see Problem 10); equivalently,
(ii') $A \triangle B \in \mathcal{M}$ and $A \cup B \in \mathcal{M}$.
[Hint: Verify that

$$A \cup B=(A \triangle B) \triangle(A \cap B) \tag{7.3.E.11}$$

and

$$A-B=(A \cup B) \triangle B, \tag{7.3.E.12}$$

while

$$A \cap B=(A \cup B) \triangle(A \triangle B).] \tag{7.3.E.13}$$

Exercise $7.3.E.12$

Show that a set family $\mathcal{M} \neq \emptyset$ is a σ-ring iff one of the following conditions holds.
(a) $\mathcal{M}$ is closed under countable unions and proper differences $(X-Y$ with $X \supseteq Y)$;
(b) $\mathcal{M}$ is closed under countable disjoint unions, proper differences, and finite intersections; or
(c) $\mathcal{M}$ is closed under countable unions and symmetric differences (see Problem 10).
[Hints: (a) $X-Y=(X \cup Y)-Y$, a proper difference.
(b) $X-Y=X-(X \cap Y)$ reduces any difference to a proper one; then

$$X \cup Y=(X-Y) \cup(Y-X) \cup(X \cap Y) \tag{7.3.E.14}$$

shows that $\mathcal{M}$ is closed under all finite unions; so $\mathcal{M}$ is a ring. Now use Corollary 1 in §1 for countable unions.
(c) Use Problem 11.]

Exercise $7.3.E.13$

From Problem 10, treating $\triangle$ as addition and $\cap$ as multiplication, show that any set ring $\mathcal{M}$ is an algebraic ring with unity, i.e., satisfies the six field axioms (Chapter 2, §§1-4), except $V(b)$ (existence of multiplicative inverses).

Exercise $7.3.E.14$

A set family $\mathcal{H}$ is said to be hereditary iff

$$(\forall X \in \mathcal{H})(\forall Y \subseteq X) \quad Y \in H. \tag{7.3.E.15}$$

Prove the following.

(a) For every family $\mathcal{M} \subseteq 2^S$, there is a "smallest" hereditary ring $\mathcal{H} \supseteq \mathcal{M}$ ($\mathcal{H}$ is said to be generated by $\mathcal{M}$). Similarly for σ-rings, fields, and σ-fields.
(b) The hereditary σ-ring generated by $\mathcal{M}$ consists of those sets which can be covered by countably many $\mathcal{M}$-sets.

Exercise $7.3.E.15$

Prove that the field (σ-field in S, generated by a ring (σ-ring $\mathcal{R}$, consists exactly of all $\mathcal{R}$-sets and their complements in S.

Exercise $7.3.E.16$

Show that the ring $\mathcal{R}$ generated by a set family $\mathcal{C} \neq \emptyset$ consists of all sets of the form

$$\triangle_{k=1}^{n} A_k \tag{7.3.E.16}$$

(see Problem 10), where each $A_k \in \mathcal{C}_d$ (finite intersection of $\mathcal{C}$-sets).
[Outline: By Problem 11, $\mathcal{R}$ must contain the family (call it $\mathcal{M}$) of all such $\triangle_{k=1}^{n} A_k$. (Why?) It remains to show that $\mathcal{M}$ is a ring $\supseteq \mathcal{C}$.
Write $A+B$ for $A \triangle B$ and AB for $A \cap B$; so each $\mathcal{M}$-set is a "sum" of finitely many "products"

$$A_1 A_2 \cdots A_n. \tag{7.3.E.17}$$

By algebra, the "sum" and "product" of two such "polynomials" is such a polynomial itself. Thus

$$(\forall X, Y \in \mathcal{M}) \quad X \triangle Y \text{ and } X \cap Y \in \mathcal{M}. \tag{7.3.E.18}$$

Now use Problem 11.]

Exercise $7.3.E.17$

AUse Problem 16 to obtain a new proof of Theorem 2 in §1 and Corollary 2 in the present section.
[Hints: For semirings, $C=\mathcal{C}_d$. (Why?) Thus in Problem 16, $A_k \in \mathcal{C}$.
Also,

$$(\forall A, B \in \mathcal{C}) \quad A \triangle B=(A-B) \cup(B-A) \tag{7.3.E.19}$$

where $A-B$ and $B-A$ are finite disjoint unions of $\mathcal{C}$-sets. (Why?)
Deduce that $A \triangle B \in \mathcal{C}_s^{\prime}$ and, by induction,

$$\triangle_{k=1}^{n} A_k \in \mathcal{C}_s^{\prime}; \tag{7.3.E.20}$$

so $\mathcal{R} \subseteq \mathcal{C}_s^{\prime} \subseteq \mathcal{R}$. (Why?)]

Exercise $7.3.E.18$

Given a set A and a set family $\mathcal{M}$, let
\[A \cap{\dot\} \mathcal{M}\]
be the family of all sets $A \cap X$, with $X \in \mathcal{M}$; similarly,

$$\mathcal{N} \dot{\cup}(\mathcal{M} \dot{-} A)=\{\text{ all sets } Y \cup(X-A), \text{ with } Y \in \mathcal{N}, X \in \mathcal{M}\}, \text{ etc. } \tag{7.3.E.21}$$

Show that if $\mathcal{M}$ generates the ring $\mathcal{R}$, then A \cap{\dot} \mathcal{M} generates the ring

\mathcal{R}^{\prime}=A \cap{\dot} \mathcal{R}.

Similarly for σ-rings, fields, σ-fields.
[Hint for rings: Prove the following.
(i) $A \cap \mathcal{R}$ is a ring.
(ii) $\mathcal{M} \subseteq \mathcal{R}^{\prime} \cup(\mathcal{R} \pm A)$, with $\mathcal{R}^{\prime}$ as above.
(iii) $\mathcal{R} \cup(\mathcal{R} \div A)$ is a ring (call it $\mathcal{N}$) .
(iv) By (ii), $\mathcal{R} \subseteq \mathcal{N}$, so A \cap \mathcal{R} \subseteq A \cap \mathcal{N} \subseteq \mathcal{R}^{\prime}\right.
(v) $A \cap \mathcal{R} \supseteq \mathcal{R}^{\prime}($ for $A \cap \mathcal{R} \supseteq A \cap \mathcal{M})$.
Hence $\mathcal{R}^{\prime}=A \cap \mathcal{R}$.]

7.4: Set Functions. Additivity. Continuity

I. The letter "v" in vA may be treated as a certain function symbol that assigns a numerical value (called "volume") to the set A. So far we have defined such "volumes" for all intervals, then for $\mathcal{C}$-simple sets, and even for $\mathcal{C}_\sigma$-sets in E^n.

Mathematically this means that the volume function v has been defined first on $\mathcal{C}$ (the intervals), then on $\mathcal{C}_s'$ ($\mathcal{C}$-simple sets), and finally on $\mathcal{C}_\sigma$.

Thus we have a function v which assigns values ("volumes") not just to single points, as ordinary "point functions" do, but to whole sets, each set being treated as one thing.

In other words, the domain of the function v is not just a set of points, but a set family ($\mathcal{C}, \mathcal{C}_s'$, or $\mathcal{C}_\sigma$).

The "volumes" assigned to such sets are the function values (for $\mathcal{C}$ and $\mathcal{C}_s'$-sets they are real numbers; for $\mathcal{C}_\sigma$-sets they may reach $+\infty$).This is symbolized by

$$v : \mathcal{C} \rightarrow E^1 \tag{7.4.1}$$

or

$$v : \mathcal{C}_\sigma \rightarrow E^*; \tag{7.4.2}$$

more precisely,

$$v : \mathcal{C}_\sigma \rightarrow [0, \infty], \tag{7.4.3}$$

since volume is nonnegative.

It is natural to call v a set function (as opposed to ordinary point functions). As we shall see, there are many other set functions. The function values need not be real; they may be complex numbers or vectors. This agrees with our general definition of a function as a certain set of ordered pairs (Definition 3 in Chapter 1, §§4-7); e.g.,

$$v=\left(\begin{array}{cccc} A & B & C & \cdots \\ vA & vB & vC & \cdots \end{array}\right). \tag{7.4.4}$$

Here the domain consists of certain sets $A, B, C, \ldots$. This leads us to the following definition.

Definition 1

A set function is a mapping

$$s : \mathcal{M} \rightarrow E \tag{7.4.5}$$

whose domain is a set family $\mathcal{M}$.

The range space E is assumed to be E^1, E^*, C (the complex field), E^n, or another normed space. Thus s may be real, extended real, complex, or vector valued.

To each set $X \in \mathcal{M}$, the function s assigns a unique function value denoted $s(X)$ or sX (which is an element of the range space E).

We say that s is finite on a set family $\mathcal{N} \subseteq \mathcal{M}$ iff

$$(\forall X \in \mathcal{N}) \quad |sX| < \infty; \tag{7.4.6}$$

briefly, $|s| < \infty$ on $\mathcal{N}$. (This is automatic if s is complex or vector valued.)

We call s semifinite if at least one of $\pm\infty$ is excluded as function value, e.g., if $s \geq 0$ on $\mathcal{M}$; i.e.,

$$s : \mathcal{M} \rightarrow [0, \infty]. \tag{7.4.7}$$

(The symbol ∞ stands for $+\infty$ throughout).

Definition 2

A set function

$$s:\mathcal{M}\rightarrow E \tag{7.4.8}$$

is called additive (or finitely additive) on $\mathcal{N}\subseteq\mathcal{M}$ iff for any finite disjoint union $\bigcup_k A_k$, we have

$$\sum_k sA_k = s\left(\bigcup_k A_k\right), \tag{7.4.9}$$

provided $\bigcup_k A_k$ and all the A_k are $\mathcal{N}$-sets.

If this also holds for countable disjoint unions, s is called σ-additive (or countably additive or completely additive) on $\mathcal{N}$.

If $\mathcal{N}=\mathcal{M}$ here, we simply say that s is additive (σ-additive, respectively).

Note 1. As $\bigcup A_k$ is independent of the order of the A_k, σ-additivity pre-supposes and implies that the series

$$\sum sA_k \tag{7.4.10}$$

is permutable (§2) for any disjoint sequence

$$\{A_k\}\subseteq\mathcal{N}. \tag{7.4.11}$$

(The partial sums do exist, by our conventions (2*) in Chapter 4, §4.)

The set functions in the examples below are additive; v is even σ-additive (Corollary 1 in §2).

Examples (b)-(d) show that set functions may arise from ordinary "point functions."

Examples

(a) The volume function $v:\mathcal{C}\rightarrow E^1$ on $\mathcal{C}$ (= intervals in E^n), discussed above, is called the Lebesgue premeasure (in E^n).

(b) Let $\mathcal{M}=\{\text{all finite intervals } I\subset E^1\}$.

Given $f:E^1\rightarrow E,$ set

$$(\forall I\in\mathcal{M})\quad sI=V_f[\bar{I}], \tag{7.4.12}$$

the total variation of f on the closure of I (Chapter 5, §7).

Then $s:\mathcal{M}\rightarrow[0,\infty]$ is additive by Theorem 1 of Chapter 5, §7.

(c) Let $\mathcal{M}$ and f be as in Example (b).

Suppose f has an antiderivative (Chapter 5, §5) on E^1. For each interval X with endpoints $a,b\in E^1(a\leq b),$ set

$$sX=\int_a^b f. \tag{7.4.13}$$

This yields a set function $s:\mathcal{M}\rightarrow E$ (real, complex, or vector valued), additive by Corollary 6 in Chapter 5, §5.

(d) Let $\mathcal{C}=\{\text{all finite intervals in } E^1\}$.

Suppose

$$\alpha:E^1\rightarrow E^1 \tag{7.4.14}$$

has finite one-sided limits

$$\alpha(p+)\text{ and }\alpha(p-) \tag{7.4.15}$$

at each $p\in E^1$. The Lebesgue-Stieltjes (LS) function

$$s_\alpha:\mathcal{C}\rightarrow E^1 \tag{7.4.16}$$

(important for Lebesgue-Stieltjes integration) is defined as follows.

Set $s_\alpha \emptyset = 0$. For nonvoid intervals, including $[a, a] = \{a\}$, set

$$\begin{aligned} s_\alpha[a, b] &= \alpha(b+) - \alpha(a-), \\ s_\alpha(a, b] &= \alpha(b+) - \alpha(a+), \\ s_\alpha[a, b) &= \alpha(b-) - \alpha(a-), \text{ and} \\ s_\alpha(a, b) &= \alpha(b-) - \alpha(a+). \end{aligned}$$

For the properties of s_α see Problem 7ff. , below.

(e) Let mX be the mass concentrated in the part X of the physical space S. Then m is a nonnegative set function defined on

$$2^S = \{\text{ all subsets } X \subseteq S\} \ (\S 3). \tag{7.4.17}$$

If instead mX were the electric load of X, then m would be sign changing.

II. The rest of this section is redundant for a "limited approach."

lemmas

Let $s : \mathcal{M} \rightarrow E$ be additive on $\mathcal{N} \subseteq \mathcal{M}$. Let

$$A, B \in \mathcal{N}, A \subseteq B. \tag{7.4.18}$$

Then we have the following.

(1) If $|sA| < \infty$ and $B - A \in \mathcal{N}$, then

$$s(B-A) = sB - sA \text{ (”subtractivity”)}. \tag{7.4.19}$$

(2) If $\emptyset \in \mathcal{N}$, then $s\emptyset = 0$ provided $|sX| < \infty$ for at least one $X \in \mathcal{N}$.

(3) If $\mathcal{N}$ is a semiring, then $sA = \pm\infty$ implies $|sB| = \infty$. Hence

$$|sB| < \infty \Rightarrow |sA| < \infty. \tag{7.4.20}$$

If further s is semifinite then

$$sA = \pm\infty \Rightarrow sB = \pm\infty \tag{7.4.21}$$

(same sign).

Proof

(1) As $B \supseteq A$, we have

$$B = (B-A) \cup A \text{ (disjoint)}; \tag{7.4.22}$$

so by additivity,

$$sB = s(B-A) + sA. \tag{7.4.23}$$

If $|sA| < \infty$, we may transpose to get

$$sB - sA = s(B-A), \tag{7.4.24}$$

as claimed.

(2) Hence

$$s\emptyset = s(X-X) = sX - sX = 0 \tag{7.4.25}$$

if $X, \emptyset \in \mathcal{N}$, and $|sX| < \infty$.

(3) If $\mathcal{N}$ is a semiring, then

$$B-A=\bigcup_{k=1}^{n} A_{k} \text{ (disjoint)} \tag{7.4.26}$$

for some $\mathcal{N}$-sets A_k; so

$$B=A \cup \bigcup_{k=1}^{n} A_{k} \text{ (disjoint).} \tag{7.4.27}$$

By additivity,

$$s B=s A+\sum_{k=1}^{n} s A_{k}; \tag{7.4.28}$$

so by our conventions,

$$|s A|=\infty \Rightarrow|s B|=\infty. \tag{7.4.29}$$

If, further, s is semifinite, one of $\pm\infty$ is excluded. Thus sA and sB, if infinite, must have the same sign. This completes the proof. □

In §§1 and 2, we showed how to extend the notion of volume from intervals to a larger set family, preserving additivity. We now generalize this idea.

Theorem 7.4.1

If

$$s: \mathcal{C} \rightarrow E \tag{7.4.30}$$

is additive on $\mathcal{C}$, an arbitrary semiring, there is a unique set function

$$\bar{s}: \mathcal{C}_{s} \rightarrow E, \tag{7.4.31}$$

additive on $\mathcal{C}_s$, with $\bar{s}=s$ on $\mathcal{C}$, i.e.,

$$\bar{s} X=s X \text{ for } X \in \mathcal{C}. \tag{7.4.32}$$

We call $\bar{s}$ the additive extension of s to $\mathcal{C}_{s}=\mathcal{C}_{s}^{\prime}$ (Corollary 2 in §3).

Proof

If $s \geq 0(s: \mathcal{C} \rightarrow[0, \infty])$, proceed as in Lemma 1 and Corollary 2, all of §1.

The general proof (which may be omitted or deferred) is as follows.

Each $X \in \mathcal{C}_{s}^{\prime}$ has the form

$$X=\bigcup_{i=1}^{m} X_{i} \text{(disjoint)}, \quad X_{i} \in \mathcal{C}. \tag{7.4.33}$$

Thus if $\bar{s}$ is to be additive, the only way to define it is to set

$$\bar{s} X=\sum_{i=1}^{m} s X_{i}. \tag{7.4.34}$$

This already makes $\bar{s}$ unique, provided we show that

$$\sum_{i=1}^{m} s X_{i} \tag{7.4.35}$$

does not depend on the particular decomposition

$$X=\bigcup_{i=1}^{m} X_{i} \tag{7.4.36}$$

(otherwise, all is ambiguous).

Then take any other decomposition

$$X=\bigcup_{k=1}^{n} Y_{k} \text{ (disjoint)}, \quad Y_{k} \in \mathcal{C}. \tag{7.4.37}$$

Additivity implies

$$(\forall i, k) \quad s X_{i}=\sum_{k=1}^{n} s\left(X_{i} \cap Y_{k}\right) \text{ and } s Y_{k}=\sum_{i=1}^{m} s\left(X_{i} \cap Y_{k}\right). \tag{7.4.38}$$

(Verify!) Hence

$$\sum_{i=1}^{m} s X_{i}=\sum_{i, k} s\left(X_{i} \cap Y_{k}\right)=\sum_{k=1}^{n} s Y_{k}. \tag{7.4.39}$$

Thus, indeed, it does not matter which particular decomposition we choose, and our definition of $\bar{s}$ is unambiguous.

If $X \in \mathcal{C},$ we may choose (say)

$$X=\bigcup_{i=1}^{1} X_{i}, X_{1}=X; \tag{7.4.40}$$

so

$$\bar{s} X=s X_{1}=s X; \tag{7.4.41}$$

i.e., $\bar{s}=s$ on $\mathcal{C}$, as required.

Finally, for the additivity of $\bar{s}$, let

$$A=\bigcup_{k=1}^{m} B_{k} \text{ (disjoint)}, \quad A, B_{k} \in \mathcal{C}_{s}^{\prime}. \tag{7.4.42}$$

Here we may set

$$B_{k}=\bigcup_{i=1}^{n_{k}} C_{k i} \text{ (disjoint)}, \quad C_{k i} \in \mathcal{C}. \tag{7.4.43}$$

Then

$$A=\bigcup_{k, i} C_{k i} \text{ (disjoint)}; \tag{7.4.44}$$

so by our definition of $\bar{s}$,

$$\bar{s} A=\sum_{k, i} s C_{k i}=\sum_{k=1}^{m}\left(\sum_{i=1}^{n_{k}} s C_{k i}\right)=\sum_{k=1}^{m} \bar{s} B_{k}, \tag{7.4.45}$$

as required. $\square$

Continuity. We write $X_{n} \nearrow X$ to mean that

$$X=\bigcup_{n=1}^{\infty} X_{n} \tag{7.4.46}$$

and $\left\{X_{n}\right\} \uparrow$, i.e.,

$$X_n \subseteq X_{n+1}, \quad n=1,2,\ldots. \tag{7.4.47}$$

Similarly, $X_n \searrow X$ iff

$$X=\bigcap_{n=1}^{\infty} X_n \tag{7.4.48}$$

and $\{X_n\}\downarrow$, i.e.,

$$X_n \supseteq X_{n+1}, \quad n=1,2,\ldots. \tag{7.4.49}$$

In both cases, we set

$$X=\lim_{n\to\infty} X_n. \tag{7.4.50}$$

This suggests the following definition.

Definition 3

A set function $s:\mathcal{M}\to E$ is said to be

(i) left continuous (on $\mathcal{M}$) iff

$$sX=\lim_{n\to\infty} sX_n \tag{7.4.51}$$

whenever $X_n \nearrow X$ and $X, X_n \in \mathcal{M}$;

(ii) right continuous iff

$$sX=\lim_{n\to\infty} sX_n \tag{7.4.52}$$

whenever $X_n \searrow X$, with $X, X_n \in \mathcal{M}$ and $|sX_j|<\infty$.

Thus in case (i),

$$\lim_{n\to\infty} sX_n = s\bigcup_{n=1}^{\infty} X_n \tag{7.4.53}$$

if all X_n and $\bigcup_{n=1}^{\infty} X_n$ are $\mathcal{M}$-sets.

In case (ii),

$$\lim_{n\to\infty} sX_n = s\bigcap_{n=1}^{\infty} X_n \tag{7.4.54}$$

if all X_n and $\bigcap_{n=1}^{\infty} X_n$ are in $\mathcal{M}$, and $|sX_1|<\infty$.

Note 2. The last restriction applies to right continuity only. (We choose simply to exclude from consideration sequences $\{X_n\}\downarrow$, with $|sX_1|=\infty$; see Problem 4.)

Theorem 7.4.2

If $s:\mathcal{C}\to E$ is σ-additive and semifinite on $\mathcal{C}$, a semiring, then s is both left and right continuous (briefly, continuous).

Proof

We sketch the proof for rings; for semirings, see Problem 1.

Left continuity. Let $X_n \nearrow X$ with $X_n, X\in\mathcal{C}$ and

$$X=\bigcup_{n=1}^{\infty} X_n. \tag{7.4.55}$$

If $sX_n=\pm\infty$ for some n, then (Lemma 3)

$$sX = sX_m = \pm\infty \text{ for } m \geq n, \tag{7.4.56}$$

since $X \supseteq X_m \supseteq X_n$; so

$$\lim sX_m = \pm\infty = sX, \tag{7.4.57}$$

as claimed.

Thus assume all sX_n finite; so $s\emptyset = 0$, by Lemma 2.

Set $X_0 = \emptyset$. As is easily seen,

$$X = \bigcup_{n=1}^{\infty} X_n = \bigcup_{n=1}^{\infty} \left(X_n - X_{n-1}\right) \text{ (disjoint)}, \tag{7.4.58}$$

and

$$(\forall n) \quad X_n - X_{n-1} \in \mathcal{C} \text{ (a ring)}. \tag{7.4.59}$$

Also,

$$(\forall m \geq n) \quad X_m = \bigcup_{n=1}^{m} \left(X_n - X_{n-1}\right) \text{ (disjoint)}. \tag{7.4.60}$$

(Verify!) Thus by additivity,

$$sX_m = \sum_{n=1}^{m} s\left(X_n - X_{n-1}\right), \tag{7.4.61}$$

and by the assumed σ-additivity,

$$\begin{aligned} sX = s\bigcup_{n=1}^{\infty}\left(X_n - X_{n-1}\right) &= \sum_{n=1}^{\infty} s\left(X_n - X_{n-1}\right) \\ &= \lim_{m\to\infty} \sum_{n=1}^{m} s\left(X_n - X_{n-1}\right) = \lim_{m\to\infty} sX_m, \end{aligned}$$

as claimed.

Right continuity. Let $X_n \searrow X$ with $X, X_n \in \mathcal{C}$,

$$X = \bigcap_{n=1}^{\infty} X_n, \tag{7.4.62}$$

and

$$\left|sX_1\right| < \infty. \tag{7.4.63}$$

As $X \subseteq X_n \subseteq X_1$, Lemma 3 yields that

$$(\forall n) \quad \left|sX_n\right| < \infty \tag{7.4.64}$$

and $|sX| < \infty$.

As

$$X = \bigcap_{k=1}^{\infty} X_k, \tag{7.4.65}$$

we have

$$(\forall n) \quad X_n = X \cup \bigcup_{k=n+1}^{\infty} \left(X_{k-1} - X_k\right) \text{ (disjoint)}. \tag{7.4.66}$$

(Verify!) Thus by σ-additivity,

$$(\forall n) \quad sX_n = sX + \sum_{k=n+1}^{\infty} s\left(X_{k-1} - X_k\right), \tag{7.4.67}$$

with $|sX| < \infty, |sX_n| < \infty$ (see above).

Hence the sum

$$\sum_{k=n+1}^{\infty} s\left(X_{k-1} - X_k\right) = sX_n - sX \tag{7.4.68}$$

is finite. Therefore, it tends to 0 as $n \rightarrow \infty$ (being the "remainder term" of a convergent series). Thus $n \rightarrow \infty$ yields

$$\lim_{n \rightarrow \infty} sX_n = sX + \lim \sum_{k=n+1}^{\infty} s\left(X_{k-1} - X_k\right) = sX, \tag{7.4.69}$$

as claimed. □

7.4.E: Problems on Set Functions

Exercise $7.4.E.1$

Prove Theorem 2 in detail for semirings.
[Hint: We know that

$$X_n - X_{n-1} = \bigcup_{i=1}^{m_n} Y_{ni} \text{ (disjoint)} \tag{7.4.E.1}$$

for some $Y_{ni} \in \mathcal{C}$, so

$$\bar{s}\left(X_n - X_{n-1}\right) = \sum_{i=1}^{m_n} s Y_{ni}, \tag{7.4.E.2}$$

with $\bar{s}$ as in Theorem 1.]

Exercise $7.4.E.2$

Let s be additive on $\mathcal{M}$, a ring. Prove that s is also σ-additive provided s is either
(i) left continuous, or
(ii) finite on $\mathcal{M}$ and right-continuous at $\emptyset$; i.e.,

$$\lim_{n \rightarrow \infty} s X_n = 0 \tag{7.4.E.3}$$

when $X_n \searrow \emptyset \left(X_n \in \mathcal{M}\right)$.
[Hint: Let

$$A = \bigcup_n A_n \text{ (disjoint)}, \quad A, A_n \in \mathcal{M}. \tag{7.4.E.4}$$

Set

$$X_n = \bigcup_{k=1}^n A_k, Y_n = A - X_n. \tag{7.4.E.5}$$

Verify that $X_n, Y_n \in \mathcal{M}, X_n \nearrow A, Y_n \searrow \emptyset$.
In case (i),

$$s A = \lim s X_n = \sum_{k=1}^{\infty} s A_k. \tag{7.4.E.6}$$

(Why?)
For (ii), use the Y_n.]

Exercise $7.4.E.3$

Let

$$\mathcal{M} = \left\{\text{all intervals in the rational field } R \subset E^1\right\}. \tag{7.4.E.7}$$

Let

$$sX = b - a \tag{7.4.E.8}$$

if a, b are the endpoints of $X \in \mathcal{M}$ $(a, b \in R, a \leq b)$. Prove that
(i) $\mathcal{M}$ is a semiring;
(ii) s is continuous;
(iii) s is additive but not σ-additive; thus Problem 2 fails for semirings.
[Hint: R is countable. Thus each $X \in \mathcal{M}$ is a countable union of singletons $\{x\} = [x, x]$; hence $sX = 0$ if s were σ-additive.]

Exercise $7.4.E.3'$

Let $N = \{\text{naturals}\}$. Let

$$\mathcal{M} = \{\text{all finite subsets of } N \text{ and their complements in } N\}. \tag{7.4.E.9}$$

If $X \in \mathcal{M}$, let $sX = 0$ if X is finite, and $sX = \infty$ otherwise. Show that
(i) $\mathcal{M}$ is a set field;
(ii) s is right continuous and additive, but not σ-additive.
Thus Problem 2 (ii) fails if s is not finite.

Exercise $7.4.E.4$

Let

$$\mathcal{C} = \left\{\text{finite and infinite intervals in } E^1\right\}. \tag{7.4.E.10}$$

If a, b are the endpoints of an interval X $(a, b \in E^*, a < b)$, set

$$sX = \begin{cases} b - a, & a < b, \\ 0, & a = b. \end{cases} \tag{7.4.E.11}$$

Show that s is σ-additive on $\mathcal{C}$, a semiring.
Let

$$X_n = (n, \infty); \tag{7.4.E.12}$$

so $sX_n = \infty - n = \infty$ and $X_n \searrow \emptyset$. (Verify!) Yet

$$\lim sX_n = \infty \neq s\emptyset. \tag{7.4.E.13}$$

Does this contradict Theorem 2?

Exercise $7.4.E.5$

Fill in the missing proof details in Theorem 1.

Exercise 7.4.E.6

Let s be additive on $\mathcal{M}$. Prove the following.
(i) If $\mathcal{M}$ is a ring or semiring, so is

$$\mathcal{N}=\{X\in\mathcal{M}||sX|<\infty\} \tag{7.4.E.14}$$

if $\mathcal{N}\neq\emptyset$.
(ii) If $\mathcal{M}$ is generated by a set family $\mathcal{C}$, with $|s|<\infty$ on $\mathcal{C}$, then $|s|<\infty$ on $\mathcal{M}$.
[Hint: Use Problem 16 in §3.]

Exercise 7.4.E.7

$\Rightarrow$ (Lebesgue-Stieltjes set functions.) Let α and s_α be as in Example (d). Prove the following.
(i) $s_\alpha\geq 0$ on $\mathcal{C}$ iff $\alpha\uparrow$ on E^1 (see Theorem 2 in Chapter 4, §5).
(ii) $s_\alpha\{p\}=s_\alpha[p,p]=0$ iff α is continuous at p.
(iii) s_α is additive.
[Hint: If

$$A=\bigcup_{i=1}^{n}A_i \text{ (disjoint)}, \tag{7.4.E.15}$$

the intervals A_{i-1}, A_i must be adjacent. For two such intervals, consider all cases like

$$(a,b]\cup(b,c),[a,b)\cup[b,c], \text{ etc.} \tag{7.4.E.16}$$

Then use induction on n.]
(iv) If α is right continuous at a and b, then

$$s_\alpha(a,b]=\alpha(b)-\alpha(b). \tag{7.4.E.17}$$

If α is continuous at a and b, then

$$s_\alpha[a,b]=s_\alpha(a,b]=s_\alpha[a,b)=s_\alpha(a,b). \tag{7.4.E.18}$$

(v) If $\alpha\uparrow$ on E^1, then s_α satisfies Lemma 1 and Corollary 2 in §1 (same proof), as well as Lemma 1, Theorem 1, Corollaries 1-4, and Note 3 in §2 (everything except Corollaries 5 and 6).
[Hint: Use (i) and (iii). For Lemma 1 in §2, take first a half-open $B=(a,b];$ use the definition of a right-side limit along with Theorems 1 and 2 in Chapter 4, §5, to prove

$$(\forall\varepsilon>0)(\exists c>b)\quad 0\leq\alpha(c-)-\alpha(b+)<\varepsilon; \tag{7.4.E.19}$$

then set $C=(a,c).$ Similarly for $B=[a,b),$ etc. and for the closed interval $A\subseteq B$.]
(vi) If $\alpha(x)=x$ then $s_\alpha=v,$ the volume (or length) function in E^1.

Exercise 7.4.E.8

Construct LS set functions (Example (d)), with $\alpha\uparrow$ (see Problem 7(v)), so that
(i) $s_\alpha[0,1]\neq s_\alpha[1,2]$;
(ii) $s_\alpha E^1=1$ (after extending s_α to $\mathcal{C}_\sigma-setsin\backslash(E^1$);
(ii') $s_\alpha E^1=c$ for a fixed $c\in(0,\infty)$;
(iii) $s_\alpha\{0\}=1$ and $s_\alpha[0,1]>s_\alpha(0,1]$.

Describe s_α if $\alpha(x)=[x]$ (the integral part of x).
[Hint: See Figure 16 in Chapter 4, §1.]

Exercise $7.4.E.9$

For an arbitrary $\alpha: E^1 \rightarrow E^1$, define $\sigma_\alpha: \mathcal{C} \rightarrow E^1$ by

$$\sigma_\alpha[a, b]=\sigma_\alpha(a, b]=\sigma_\sigma[a, b)=\sigma_\alpha(a, b)=\alpha(b)-\alpha(a) \tag{7.4.E.20}$$

(the original Stieltjes method). Prove that σ_α is additive but not σ-additive unless α is continuous (for Theorem 2 fails).

7.5: Nonnegative Set Functions. Premeasures. Outer Measures

We now concentrate on nonnegative set functions

$$m : \mathcal{M} \rightarrow [0, \infty] \tag{7.5.1}$$

(we mostly denote them by m or μ). Such functions have the advantage that

$$\sum_{n=1}^{\infty} m X_n \tag{7.5.2}$$

exists and is permutable (Theorem 2 in §2) for any sets $X_n \in \mathcal{M}$, since $m X_n \geq 0$. Several important notions apply to such functions (only). They "mimic" §§1 and 2.

Definition 1

A set function

$$m : \mathcal{M} \rightarrow [0, \infty] \tag{7.5.3}$$

is said to be

(i) monotone (on $\mathcal{M}$) iff

$$m X \leq m Y \tag{7.5.4}$$

whenever

$$X \subseteq Y \text{ and } X, Y \in \mathcal{M}; \tag{7.5.5}$$

(ii) (finitely) subadditive (on $\mathcal{M}$) iff for any finite union

$$\bigcup_{k=1}^{n} Y_k, \tag{7.5.6}$$

we have

$$m X \leq \sum_{k=1}^{m} m Y_k \tag{7.5.7}$$

whenever $X, Y_k \in \mathcal{M}$ and

$$X \subseteq \bigcup_{k=1}^{n} Y_k \text{ (disjoint or not)}; \tag{7.5.8}$$

(iii) σ-subadditive (on $\mathcal{M}$) iff (1) holds for countable unions, too.

Recall that $\{Y_k\}$ is called a covering of X iff

$$X \subseteq \bigcup_{k} Y_k. \tag{7.5.9}$$

We call it an $\mathcal{M}$-covering of X if all Y_k are $\mathcal{M}$-sets. We now obtain the following corollary.

Corollary 7.5.1

Subadditivity implies monotonicity.

Take $n = 1$ in formula (1).

Corollary 7.5.2

If $m : \mathcal{C} \rightarrow [0, \infty]$ is additive (σ-additive) on $\mathcal{C}$, a semiring, then m is also subadditive (σ-subadditive, respectively), hence monotone, on $\mathcal{C}$.

Proof

The proof is a mere repetition of the argument used in Lemma 1 in §1.

Taking $n = 1$ in formula (ii) there, we obtain finite subadditivity.

For σ-subadditivity, one only has to use countable unions instead of finite ones.

Note 1. The converse fails: subadditivity does not imply additivity.

Note 2. Of course, Corollary 2 applies to rings, too (see Corollary 1 in §3).

Definition 2

A premeasures is a set function

$$\mu : \mathcal{C} \rightarrow [0, \infty] \tag{7.5.10}$$

such that

$$\emptyset \in \mathcal{C} \text{ and } \mu\emptyset = 0. \tag{7.5.11}$$

($\mathcal{C}$ may, but need not, be a semiring.)

A premeasure space is a triple

$$(S, \mathcal{C}, \mu), \tag{7.5.12}$$

where $\mathcal{C}$ is a family of subsets of S (briefly, $\mathcal{C} \subseteq 2^S$) and

$$\mu : \mathcal{C} \rightarrow [0, \infty] \tag{7.5.13}$$

is a premeasure. In this case, $\mathcal{C}$-sets are also called basic sets.

If

$$A \subseteq \bigcup_n B_n, \tag{7.5.14}$$

with $B_n \in \mathcal{C}$, the sequence $\{B_n\}$ is called a basic covering of A, and

$$\sum_n \mu B_n \tag{7.5.15}$$

is a basic covering value of A; $\{B_n\}$ may be finite or infinite.

Examples

(a) The volume function v on $\mathcal{C}$ (= intervals in E^n) is a premeasure, as $v \geq 0$ and $v\emptyset = 0$. $(E^n, \mathcal{C}, v)$ is the Lebesgue premeasure space.

(b) The LS set function s_α is a premeasure if $\alpha \uparrow$ (see Problem 7 in §4). We call it the α-induced Lebesgue-Stieltjes (LS) premeasure in E^1.

We now develop a method for constructing σ-subadditive premeasures. (This is a first step toward achieving σ-additivity; see §4.)

Definition 3

For any premeasure space $(S, \mathcal{C}, \mu)$, we define the μ-induced outer measure m^* on 2^S (= all subsets of S) by setting, for each $A \subseteq S$,

$$m^* A = \inf\left\{\sum_n \mu B_n | A \subseteq \bigcup_n B_n, B_n \in \mathcal{C}\right\}, \tag{7.5.16}$$

i.e., $m^* A$ (called the outer measure of A) is the glb of all basic covering values of A.

If $\mu = v, m^*$ is called the Lebesgue outer measure in E^n.

Note 3. If A has no basic coverings, we set $m^* A = \infty$. More generally, we make the convention that $\inf \emptyset = +\infty$.

Note 4. By the properties of the glb, we have

$$(\forall A \subseteq S) \quad 0 \leq m^* A. \tag{7.5.17}$$

If $A \in \mathcal{C}$, then $\{A\}$ is a basic covering; so

$$m^* A \leq \mu A. \tag{7.5.18}$$

In particular, $m^* \emptyset = \mu \emptyset = 0$.

Theorem 7.5.1

The set function m^* so defined is σ-subadditive on 2^S.

Proof

Given

$$A \subseteq \bigcup_n A_n \subset S, \tag{7.5.19}$$

we must show that

$$m^* A \leq \sum_n m^* A_n. \tag{7.5.20}$$

This is trivial if $m^* A_n = \infty$ for some n. Thus assume

$$(\forall n) \quad m^* A_n < \infty \tag{7.5.21}$$

and fix $\varepsilon > 0$.

By Note 3, each A_n has a basic covering

$$\{B_{nk}\}, \quad k = 1, 2, \ldots \tag{7.5.22}$$

(otherwise, $m^* A_n = \infty$.) By properties of the glb, we can choose the B_{nk} so that

$$(\forall n) \quad \sum_k \mu B_{nk} < m^* A_n + \frac{\varepsilon}{2^n}. \tag{7.5.23}$$

(Explain from (2)). The sets B_{nk} (for all n and all k) form a countable basic covering of all A_n, hence of A. Thus by Definition 3,

$$m^* A \leq \sum_n \left(\sum_k \mu B_{nk}\right) \leq \sum_n \left(m^* A_n + \frac{\varepsilon}{2^n}\right) \leq \sum^n m^* A_n + \varepsilon. \tag{7.5.24}$$

As ε is arbitrary, we can let $\varepsilon \rightarrow 0$ to obtain the desired result. □

Note 5. In view of Theorem 1, we now generalize the notion of an outer measure in S to mean any σ-subadditive premeasure defined on all of 2^S.

By Note 4, $m^* \leq \mu$ on $\mathcal{C}$, not $m^* = \mu$ in general. However, we obtain the following result.

Theorem 7.5.2

With m^* as in Definition 3, we have $m^* = \mu$ on $\mathcal{C}$ iff μ is σ-subadditive on $\mathcal{C}$. Hence, in this case, m^* is an extension of μ.

Proof

Suppose μ is σ-subadditive and fix any $A \in \mathcal{C}$. By Note 4,

$$m^* A \leq \mu A. \tag{7.5.25}$$

We shall show that

$$\mu A \leq m^* A, \tag{7.5.26}$$

too, and hence $\mu A = m^* A$.

Now, as $A \in \mathcal{C}$, A surely has basic coverings, e.g., $\{A\}$. Take any basic covering:

$$A \subseteq \bigcup_n B_n, \quad B_n \in \mathcal{C}. \tag{7.5.27}$$

As μ is σ-subadditive,

$$\mu A \leq \sum_n \mu B_n. \tag{7.5.28}$$

Thus μA does not exceed any basic covering values of A; so it cannot exceed their glb, $m^* A$. Hence $\mu = m^*$, indeed.

Conversely, if $\mu = m^*$ on $\mathcal{C}$, then the σ-subadditivity of m^* (Theorem 1) implies that of μ (on $\mathcal{C}$). Thus all is proved. $\square$

Note 6. If, in (2), we allow only finite basic coverings, then the μ-induced set function is called the μ-induced outer content, c^*. It is only finitely subadditive, in general.

In particular, if $\mu = v$ (Lebesgue premeasure), we speak of the Jordan outer content in E^n. (It is superseded by Lebesgue theory but still occurs in courses on Riemann integration.)

We add two more definitions related to the notion of coverings.

Definition 4

A set function $s : \mathcal{M} \rightarrow E\ (\mathcal{M} \subseteq 2^S)$ is called σ-finite iff every $X \in \mathcal{M}$ can be covered by a sequence of $\mathcal{M}$-sets X_n, with

$$|s X_n| < \infty \quad (\forall n). \tag{7.5.29}$$

Any set $A \subseteq S$ which can be so covered is said to be σ-finite with respect to s (briefly, (s) σ-finite).

If the whole space S can be so covered, we say that s is totally σ-finite.

For example, the Lebesgue premeasure v on E^n is totally σ-finite.

Definition 5

A set function $s : \mathcal{M} \rightarrow E^*$ is said to be regular with respect to a set family $\mathcal{A}$ (briefly, $\mathcal{A}$-regular) iff for each $A \in \mathcal{M}$,

$$sA = \inf\{sX \mid A \subseteq X, X \in \mathcal{A}\}; \tag{7.5.30}$$

that is, sA is the glb of all sX, with $A \subseteq X$ and $X \in \mathcal{A}$.

These notions are important for our later work. At present, we prove only one theorem involving Definitions 3 and 5.

Theorem 7.5.3

For any premeasure space $(S, \mathcal{C}, \mu)$, the μ-induced outer measure m^* is $\mathcal{A}$-regular whenever

$$\mathcal{C}_\sigma \subseteq \mathcal{A} \subseteq 2^S. \tag{7.5.31}$$

Thus in this case,

$$(\forall A \subseteq S) \quad m^* A = \inf \{m^* X | A \subseteq X, X \in \mathcal{A}\}. \tag{7.5.32}$$

Proof

As m^* is monotone, $m^* A$ is surely a lower bound of

$$\{m^* X | A \subseteq X, X \in \mathcal{A}\}. \tag{7.5.33}$$

We must show that there is no greater lower bound.

This is trivial if $m^* A = \infty$.

Thus let $m^* A < \infty$; so A has basic coverings (Note 3). Now fix any $\varepsilon > 0$.

By formula (2), there is a basic covering $\{B_n\} \subseteq \mathcal{C}$ such that

$$A \subseteq \bigcup_n B_n \tag{7.5.34}$$

and

$$m^* A + \varepsilon > \sum_n \mu B_n \geq \sum_n m^* B_n \geq m^* \bigcup_n B_n. \tag{7.5.35}$$

(m^* is σ-subadditive!)

Let

$$X = \bigcup_n B_n. \tag{7.5.36}$$

Then X is in $\mathcal{C}_\sigma$, hence in $\mathcal{A}$, and $A \subseteq X$. Also,

$$m^* A + \varepsilon > m^* X. \tag{7.5.37}$$

Thus $m^* A + \varepsilon$ is not a lower bound of

$$\{m^* X | A \subseteq X, X \in \mathcal{A}\}. \tag{7.5.38}$$

This proves (4). □

7.5.E: Problems on Premeasures and Related Topics

Exercise $7.5.E.1$

Fill in the missing details in the proofs, notes, and examples of this section.

Exercise $7.5.E.2$

Describe m^* on 2^S induced by a premeasure $\mu: \mathcal{C} \rightarrow E^*$ such that each of the following hold.
(a) $\mathcal{C}=\{S, \emptyset\}, \mu S=1$.
(b) $\mathcal{C}=\{S, \emptyset, \text{and all singletons}\} ; \mu S=\infty, \mu\{x\}=1$.
(c) $\mathcal{C}$ as in (b), with S uncountable; $\mu S=1$, and $\mu X=0$ otherwise.
(d) $\mathcal{C}=\{\text{all proper subsets of } S\} ; \mu X=1$ when $\emptyset \subset X \subset S ; \mu \emptyset=0$.

Exercise $7.5.E.3$

Show that the premeasures

$$v^{\prime}: \mathcal{C}^{\prime} \rightarrow[0, \infty] \tag{7.5.E.1}$$

induce one and the same (Lebesgue) outer measure m^* in E^n, with $v^{\prime}=v$ (volume, as in §2):
(a) $\mathcal{C}^{\prime}=\{\text{open intervals}\}$;
(b) $\mathcal{C}^{\prime}=\{\text{half-open intervals}\}$;
(c) $\mathcal{C}^{\prime}=\{\text{closed intervals}\}$;
(d) $\mathcal{C}^{\prime}=\mathcal{C}_\sigma$;
(e) $\mathcal{C}^{\prime}=\{\text{open sets}\}$;
(f) $\mathcal{C}^{\prime}=\{\text{half-open cubes}\}$;
[Hints: (a) Let $m^{\prime}$ be the $v^{\prime}$-induced outer measure; let $\mathcal{C}=\{\text{all intervals}\}$. As $\mathcal{C}^{\prime} \subseteq \mathcal{C}, m^{\prime} A \geq m^* A$. (Why?) Also,

$$(\forall \varepsilon>0)\left(\exists\left\{B_k\right\} \subseteq \mathcal{C}\right) \quad A \subseteq \bigcup_k B_k \text{ and } \sum v B_k \leq m^* A+\varepsilon. \tag{7.5.E.2}$$

(Why?) By Lemma 1 in §2,

$$\left(\exists\left\{C_k\right\} \subseteq \mathcal{C}^{\prime}\right) \quad B_k \subseteq C_k \text{ and } v B_k+\frac{\varepsilon}{2^k}>v^{\prime} C_k. \tag{7.5.E.3}$$

Deduce that $m^* A \geq m^{\prime} A, m^*=m^{\prime}$. Similarly for (b) and (c). For (d), use Corollary 1 and Note 3 in §1. For (e), use Lemma 2 in §2. For (f), use Problem 2 in §2.]

Exercise $7.5.E.3^{\prime}$

Do Problem 3(a)-(c), with m^* replaced by the Jordan outer content c^* (Note 6).

Exercise $7.5.E.4$

Do Problem 3, with v and m^* replaced by the LS premeasure and outer measure. (Use Problem 7 in §4.)

Exercise $7.5.E.5$

Show that a set $A \subseteq E^n$ is bounded iff its outer Jordan content is finite.

Exercise $7.5.E.6$

Find a set $A \subseteq E^1$ such that
(i) its Lebesgue outer measure is 0 $(m^* A = 0)$, while its Jordan outer content $c^* A = \infty$;
(ii) $m^* A = 0, c^* A = 1$ (see Corollary 6 in §2).

Exercise $7.5.E.7$

Let

$$\mu_1, \mu_2 : \mathcal{C} \rightarrow [0, \infty] \tag{7.5.E.4}$$

be two premeasures in S and let m_1^* and m_2^* be the outer measures induced by them.
Prove that if $m_1^* = m_2^*$ on $\mathcal{C}$, then $m_1^* = m_2^*$ on all of 2^S.

Exercise $7.5.E.8$

With the notation of Definition 3 and Note 6, prove the following.
(i) If $A \subseteq B \subseteq S$ and $m^* B = 0$, then $m^* A = 0$; similarly for c^*.
[Hint: Use monotonicity.]
(ii) The set family

$$\{X \subseteq S | c^* A = 0\} \tag{7.5.E.5}$$

is a hereditary set ring, i.e., a ring $\mathcal{R}$ such that

$$(\forall B \in \mathcal{R})(\forall A \subseteq B) \quad A \in \mathcal{R}. \tag{7.5.E.6}$$

(iii) The set family

$$\{X \subseteq S | m^* X = 0\} \tag{7.5.E.7}$$

is a hereditary σ-ring.
(iv) So also is

$$\mathcal{H} = \{\text{those } X \subseteq S \text{ that have basic coverings}\}; \tag{7.5.E.8}$$

thus $\mathcal{H}$ is the hereditary σ-ring generated by $\mathcal{C}$ (see Problem 14 in §3).

Exercise $7.5.E.9$

Continuing Problem 8(iv), prove that if μ is σ-finite (Definition 4), so is m^* when restricted to $\mathcal{H}$.
Show, moreover, that if $\mathcal{C}$ is a semiring, then each $X \in \mathcal{H}$ has a basic covering $\{Y_n\}$, with $m^* Y_n < \infty$ and with all Y_n disjoint.
[Hint: Show that

$$X \subseteq \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} B_{nk} \tag{7.5.E.9}$$

for some sets $B_{nk} \in \mathcal{C}$, with $\mu B_{nk} < \infty$. Then use Note 4 in §5 and Corollary 1 of §1.]

Exercise 7.5.E.10

Show that if

$$s: \mathcal{C} \rightarrow E^{*} \tag{7.5.E.10}$$

is σ-finite and additive on $\mathcal{C}$, a semiring, then the σ-ring $\mathcal{R}$ generated by $\mathcal{C}$ equals the σ-ring $\mathcal{R}^{\prime}$ generated by

$$\mathcal{C}^{\prime}=\{X \in \mathcal{C}|| s X \mid<\infty\} \tag{7.5.E.11}$$

(cf. Problem 6 in §4).
[Hint: By σ-finiteness,

$$(\forall X \in \mathcal{C})\left(\exists\left\{A_{n}\right\} \subseteq \mathcal{C}|| s A_{n} \mid<\infty\right) \quad X \subseteq \bigcup_{n} A_{n} ; \tag{7.5.E.12}$$

so

$$X=\bigcup_{n}\left(X \cap A_{n}\right), \quad X \cap A_{n} \in \mathcal{C}^{\prime}. \tag{7.5.E.13}$$

(Use Lemma 3 in §4.)
Thus $(\forall X \in \mathcal{C}) X$ is a countable union of $\mathcal{C}^{\prime}$-sets; so $\mathcal{C} \subseteq \mathcal{R}^{\prime}$. Deduce $\mathcal{R} \subseteq \mathcal{R}^{\prime}$. Proceed.]

Exercise 7.5.E.11

With all as in Theorem 3, prove that if A has basic coverings, then

$$\left(\exists B \in \mathcal{A}_{\delta}\right) \quad A \subseteq B \text { and } m^{*} A=m^{*} B. \tag{7.5.E.14}$$

[Hint: By formula (4),

$$(\forall n \in N)\left(\exists X_{n} \in \mathcal{A} \mid A \subseteq X_{n}\right) \quad m^{*} A \leq m X_{n} \leq m^{*} A+\frac{1}{n}. \tag{7.5.E.15}$$

(Explain!) Set

$$B=\bigcap_{n=1}^{\infty} X_{n} \in \mathcal{A}_{\delta}. \tag{7.5.E.16}$$

Proceed. For $\mathcal{A}_{\delta}$, see Definition 2(b) in §3.]

Exercise 7.5.E.12

Let $(S, \mathcal{C}, \mu)$ and m^{*} be as in Definition 3. Show that if $\mathcal{C}$ is a σ-field in S, then

$$(\forall A \subseteq S)(\exists B \in \mathcal{C}) \quad A \subseteq B \text { and } m^{*} A=\mu B. \tag{7.5.E.17}$$

[Hint: Use Problem 11 and Note 3.]

Exercise $7.5.E.13$

$\Rightarrow^*$ Show that if

$$s : \mathcal{C} \rightarrow E \tag{7.5.E.18}$$

is σ-finite and σ-additive on $\mathcal{C}$, a semiring, then s has at most one σ-additive extension to the σ-ring $\mathcal{R}$ generated by $\mathcal{C}$.
(Note that s is automatically σ-finite if it is finite, e.g., complex or vector valued.)
[Outline: Let

$$s^{\prime}, s^{\prime \prime} : \mathcal{R} \rightarrow E \tag{7.5.E.19}$$

be two σ-additive extensions of s. By Problem 10, $\mathcal{R}$ is also generated by

$$\mathcal{C}^{\prime}=\{X \in \mathcal{C}|| s X \mid<\infty\}. \tag{7.5.E.20}$$

Now set

$$\mathcal{R}^{*}=\left\{X \in \mathcal{R} | s^{\prime} X=s^{\prime \prime} X\right\}. \tag{7.5.E.21}$$

Show that $\mathcal{R}^*$ satisfies properties (i)-(iii) of Theorem 3 in §3, with $\mathcal{C}$ replaced by $\mathcal{C}'$; so $\mathcal{R}=\mathcal{R}^*$.]

Exercise $7.5.E.14$

Let $m_n^*(n=1,2,\ldots)$ be outer measures in S such that

$$(\forall X \subseteq S)(\forall n) \quad m_{n}^{*} X \leq m_{n+1}^{*} X. \tag{7.5.E.22}$$

Set

$$\mu^{*}=\lim _{n \rightarrow \infty} m_{n}^{*}. \tag{7.5.E.23}$$

Show that μ^* is an outer measure in S (see Note 5).

Exercise $7.5.E.15$

An outer measure m^* in a metric space (S,ρ) is said to have the Carathéodory property (CP) iff

$$m^{*}(X \cup Y) \geq m^{*} X+m^{*} Y \tag{7.5.E.24}$$

whenever $\rho(X,Y)>0$, where

$$\rho(X, Y)=\inf \{\rho(x, y) | x \in X, y \in Y\}. \tag{7.5.E.25}$$

For such m^*, prove that

$$m^{*}\left(\bigcup_{k} X_{k}\right)=\sum_{k} m^{*} X_{k} \tag{7.5.E.26}$$

if $\{X_k\} \subseteq 2^S$ and

$$\rho\left(X_{i}, X_{k}\right)>0 \quad(i \neq k). \tag{7.5.E.27}$$

[Hint: For finite unions, use the CP, subadditivity, and induction. Deduce that

$$(\forall n) \sum_{k=1}^{n} m^{*} X_{k} \leq m^{*} \bigcup_{k=1}^{\infty} X_{k}. \tag{7.5.E.28}$$

Let $n \rightarrow \infty$. Proceed.]

Exercise 7.5.E.16

Let $(S, \mathcal{C}, \mu)$ and m^{*} be as in Definition 3, with ρ a metric for S. Let μ_{n} be the restriction of μ to the family $\mathcal{C}_{n}$ of all $X \in \mathcal{C}$ of diameter

$$d X \leq \frac{1}{n}. \tag{7.5.E.29}$$

Let m_{n}^{*} be the μ_{n}-induced outer measure in S.
Prove that
(i) $\{m_{n}^{*}\} \uparrow$ as in Problem 14;
(ii) the outer measure

$$\mu^{*}=\lim_{n \rightarrow \infty} m_{n}^{*} \tag{7.5.E.30}$$

has the CP (see Problem 15), and

$$\mu^{*} \geq m^{*} \text{ on } 2^{S}. \tag{7.5.E.31}$$

[Outline: Let $\rho(X, Y)>\varepsilon>0(X, Y \subseteq S)$.
If for some $n, X \cup Y$ has no basic covering from $\mathcal{C}_{n}$, then

$$\mu^{*}(X \cup Y) \geq m_{n}^{*}(X \cup Y)=\infty \geq \mu^{*} X+\mu^{*} Y, \tag{7.5.E.32}$$

and the CP follows. (Explain!)
Thus assume

$$\left(\forall n>\frac{1}{\varepsilon}\right)(\forall k)\left(\exists B_{n k} \in \mathcal{C}_{n}\right) \quad X \cup Y \subseteq \bigcup_{k=1}^{\infty} B_{n k}. \tag{7.5.E.33}$$

One can choose the $B_{n k}$ so that

$$\sum_{k=1}^{\infty} \mu B_{n k} \leq m_{n}^{*}(X \cup Y)+\varepsilon. \tag{7.5.E.34}$$

(Why?) As

$$d B_{n k} \leq \frac{1}{n}<\varepsilon, \tag{7.5.E.35}$$

some $B_{n k}$ cover X only, others Y only. (Why?) Deduce that

$$\left(\forall n>\frac{1}{\varepsilon}\right) \quad m_{n}^{*} X+m_{n}^{*} Y \leq \sum_{k=1}^{\infty} \mu_{n} B_{n k} \leq m_{n}^{*}(X \cup Y)+\varepsilon. \tag{7.5.E.36}$$

Let $\varepsilon \rightarrow 0$ and then $n \rightarrow \infty$.
Also, $m^{*} \leq m_{n}^{*} \leq \mu^{*}$. (Why?)]

Exercise $7.5.E.17$

Continuing Problem 16, suppose that
$(\forall \varepsilon>0)(\forall n, k)(\forall B \in \mathcal{C})\left(\exists B_{n k} \in \mathcal{C}_{n}\right)$

$$B \subseteq \bigcup_{k=1}^{\infty} B_{n k} \text { and } \mu B+\varepsilon \geq \sum_{k=1}^{\infty} \mu B_{n k}. \tag{7.5.E.37}$$

Show that

$$m^{*}=\lim_{n \rightarrow \infty} \mu_{n}^{*}=\mu^{*}, \tag{7.5.E.38}$$

so m^{*} itself has the CP.
[Hints: It suffices to prove that $m^{*} A \geq \mu^{*} A$ if $m^{*} A<\infty$. (Why?)
Now, given $\varepsilon>0, A$ has a covering

$$\left\{B_{i}\right\} \subseteq c \tag{7.5.E.39}$$

such that

$$m^{*} A+\varepsilon \geq \sum \mu B_{i}. \tag{7.5.E.40}$$

(Why?) By assumption,

$$(\forall n) \quad B_{i} \subseteq \bigcup_{k=1}^{\infty} B_{n k}^{i} \in \mathcal{C}_{n} \text { and } \mu B_{i}+\frac{\varepsilon}{2^{i}} \geq \sum_{k=1}^{\infty} \mu B_{n k}^{i}. \tag{7.5.E.41}$$

Deduce that

$$m^{*} A+\varepsilon>\sum \mu B_{i} \geq \sum_{i=1}^{\infty}\left(\sum_{k=1}^{\infty} \mu B_{n k}^{i}-\frac{\varepsilon}{2^{i}}\right)=\sum_{i, k} \mu B_{n k}^{i}-\varepsilon \geq m_{n}^{*} A-\varepsilon. \tag{7.5.E.42}$$

Let $\varepsilon \rightarrow 0$; then $n \rightarrow \infty$.]

Exercise $7.5.E.18$

Using Problem 17, show that the Lebesgue and Lebesgue-Stieltjes outer measures have the CP.

7.6: Measure Spaces. More on Outer Measures

I. In §5, we considered premeasure spaces, stressing mainly the idea of σ-subadditivity (Note 5 in §5). Now we shall emphasize σ-additivity.

Definition 1

A premeasure

$$m : \mathcal{M} \rightarrow [0, \infty] \tag{7.6.1}$$

is called a measure (in S) iff $\mathcal{M}$ is a σ-ring (in S), and m is σ-additive on $\mathcal{M}$.

If so, the system

$$(S, \mathcal{M}, m) \tag{7.6.2}$$

is called a measure space; mX is called the measure of $X \in \mathcal{M}$; $\mathcal{M}$-sets are called m-measurable sets.

Note that m is nonnegative and $m\emptyset = 0$, as m is a premeasure (Definition 2 in §5).

Corollary 7.6.1

Measures are σ-additive, σ-subadditive, monotone, and continuous.

Proof

Use Corollary 2 in §5 and Theorem 2 in §4, noting that $\mathcal{M}$ is a σ-ring. $\square$

Corollary 7.6.2

In any measure space $(S, \mathcal{M}, m)$, the union and intersection of any sequence of m-measurable sets is m-measurable itself. So also is $X - Y$ if $X, Y \in \mathcal{M}$.

This is obvious since $\mathcal{M}$ is a σ-ring.

As measures and other premeasures are understood to be ≥ 0, we often write

$$m : \mathcal{M} \rightarrow E^{*} \tag{7.6.3}$$

for

$$m : \mathcal{M} \rightarrow [0, \infty]. \tag{7.6.4}$$

We also briefly say "measurable" for "m-measurable."

Note that $\emptyset \in \mathcal{M}$, but not always $S \in \mathcal{M}$.

Examples

(a) The volume of intervals in E^n is a σ-additive premeasure, but not a measure since its domain (the intervals) is not a σ-ring.

(b) Let $\mathcal{M} = 2^S$. Define

$$(\forall X \subseteq S) \quad mX = 0. \tag{7.6.5}$$

Then m is trivially a measure (the zero-measure). Here each set $X \subseteq S$ is measurable, with $mX = 0$.

(c) Let again $\mathcal{M} = 2^S$. Let mX be the number of elements in X, if finite, and $mX = \infty$ otherwise.

Then m is a measure ("counting measure"). Verify!

(d) Let $\mathcal{M} = 2^S$. Fix some $p \in S$. Let

$$mX = \begin{cases} 1 & \text{if } p \in X, \\ 0 & \text{otherwise} . \end{cases} \tag{7.6.6}$$

Then m is a measure (it describes a "unit mass" concentrated at p).

(e) A probability space is a measure space $(S, \mathcal{M}, m)$, with

$$S \in \mathcal{M} \text{ and } mS = 1. \tag{7.6.7}$$

In probability theory, measurable sets are called events; mX is called the probability of X, often denoted by pX or similar symbols.

In Examples (b), (c), and (d),

$$\mathcal{M} = 2^S \text{ (all subsets of } S). \tag{7.6.8}$$

More often, however,

$$\mathcal{M} \neq 2^S, \tag{7.6.9}$$

i.e., there are nonmeasurable sets $X \subseteq S$ for which mX is not defined.

Of special interest are sets $X \in \mathcal{M}$, with $mX = 0$, and their subsets. We call them m-null or null sets. One would like them to be measurable, but this is not always the case for subsets of X.

This leads us to the following definition.

Definition 2

A measure $m : \mathcal{M} \rightarrow E^*$ is called complete iff all null sets (subsets of sets of measure zero) are measurable.

We now develop a general method for constructing complete measures.

II. From §5 (Note 5) recall that an outer measure in S is a σ-subadditive premeasure defined on all of 2^S (even if it is not derived via Definition 3 in §5). In Examples (b), (c), and (d), m is both a measure and an outer measure. (Why?)

An outer measure

$$m^* : 2^S \rightarrow E^* \tag{7.6.10}$$

need not be additive; but consider this fact:

$$\text{Any set } A \subseteq S \text{ splits } S \text{ into two parts: } A \text{ itself and } -A. \tag{7.6.11}$$

It also splits any other set X into $X \cap A$ and $X - A$; indeed,

$$X = (X \cap A) \cup (X - A) \text{ (disjoint)}. \tag{7.6.12}$$

We want to single out those sets A for which m^* behaves "additively," i.e., so that

$$m^* X = m^*(X \cap A) + m^*(X - A). \tag{7.6.13}$$

This motivates our next definition.

Definition 3

Given an outer measure $m^* : 2^S \rightarrow E^*$ and a set $A \subseteq S$, we say that A is m^*-measurable iff all sets $X \subseteq S$ are split "additively" by A; that is,

$$(\forall X \subseteq S) \quad m^* X = m^*(X \cap A) + m^*(X - A). \tag{7.6.14}$$

As is easily seen (see Problem 1), this is equivalent to

$$(\forall X \subseteq A)(\forall Y \subseteq -A) \quad m^*(X \cup Y) = m^* X + m^* Y. \tag{7.6.15}$$

The family of all m^*-measurable sets is usually denoted by $\mathcal{M}^*$. The system $(S, \mathcal{M}^*, m^*)$ is called an outer measure space.

Note 1. Definition 3 applies to outer measures only. For measures, "m-measurable" means simply "member of the domain of m" (Definition 1).

Note 2. In (1) and (2), we may equivalently replace the equality sign $(=)$ by $(\geq)$. Indeed, X is covered by

$$\{X \cap A, X - A\}, \tag{7.6.16}$$

and $X \cup Y$ is covered by $\{X, Y\}$; so the reverse inequality $(\leq)$ anyway holds, by subadditivity.

Our main objective is to prove the following fundamental theorem.

Theorem 7.6.1

In any outer measure space

$$\left(S, \mathcal{M}^{*}, m^{*}\right), \tag{7.6.17}$$

the family $\mathcal{M}^{*}$ of all m^{*}-measurable sets is a σ-field in S, and m^{*}, when restricted to $\mathcal{M}^{*}$, is a complete measure (denoted by m and called the m^{*}-induced measure; so $m^{*}=m$ on $\mathcal{M}^{*}$).

Proof

We split the proof into several steps (lemmas).

lemma 1

$\mathcal{M}^{*}$ is closed under complementation:

$$\left(\forall A \in \mathcal{M}^{*}\right) \quad -A \in \mathcal{M}^{*}. \tag{7.6.18}$$

Indeed, the measurability criterion (2) is same for A and $-A$ alike.

lemma 2

$\emptyset$ and S are $\mathcal{M}^{*}$ sets. So are all sets of outer measure 0.

Proof

Let $m^{*} A=0$. To prove $A \in \mathcal{M}^{*}$, use (2) and Note 2.

Thus take any $X \subseteq A$ and $Y \subseteq -A$. Then by monotonicity,

$$m^{*} X \leq m^{*} A=0 \tag{7.6.19}$$

and

$$m^{*} Y \leq m^{*}(X \cup Y). \tag{7.6.20}$$

Thus

$$m^{*} X+m^{*} Y=0+m^{*} Y \leq m^{*}(X \cup Y), \tag{7.6.21}$$

as required.

In particular, as $m^{*} \emptyset=0, \emptyset$ is m^{*}-measurable $\left(\emptyset \in \mathcal{M}^{*}\right)$.

So is S (the complement of $\emptyset$) by Lemma 1. $\square$

lemma 3

$\mathcal{M}^{*}$ is closed under finite unions:

$$\left(\forall A, B \in \mathcal{M}^{*}\right) \quad A \cup B \in \mathcal{M}^{*}. \tag{7.6.22}$$

Proof

This time we shall use formula (1). By Note 2, it suffices to show that

$$(\forall X \subseteq S) \quad m^* X \geq m^*(X \cap (A \cup B)) + m^*(X - (A \cup B)). \tag{7.6.23}$$

Fix any $X \subseteq S$; as $A \in \mathcal{M}^*$, we have

$$m^* X = m^*(X \cap A) + m^*(X - A). \tag{7.6.24}$$

Similarly, as $B \in \mathcal{M}^*$, we have (replacing X by $X - A$ in (1))

$$\begin{aligned} m^*(X - A) &= m^*((X - A) \cap B) + m^*(X - A - B) \\ &= m^*(X \cap -A \cap B) + m^*(X - (A \cup B)), \end{aligned}$$

since

$$X - A = X \cap -A \tag{7.6.25}$$

and

$$X - A - B = X - (A \cup B). \tag{7.6.26}$$

Combining (4) with (3), we get

$$m^* X = m^*(X \cap A) + m^*(X \cap -A \cap B) + m^*(X - (A \cup B)). \tag{7.6.27}$$

Now verify that

$$(X \cap A) \cup (X \cap -A \cap B) \supseteq X \cap (A \cup B). \tag{7.6.28}$$

As m is subadditive, this yields

$$m^*(X \cap A) + m^*(X \cap -A \cap B) \geq m^*(X \cap (A \cup B)). \tag{7.6.29}$$

Combining with (5), we get

$$m^* X \geq m^*(X \cap (A \cup B)) + m^*(X - (A \cup B)), \tag{7.6.30}$$

so that $A \cup B \in \mathcal{M}^*$, indeed. $\square$

Induction extends Lemma 3 to all finite unions of $\mathcal{M}^*$-sets.

Note that by Problem 3 in §3, $\mathcal{M}^*$ is a set field, hence surely a ring. Thus Corollary 1 in §1 applies to it. (We use it below.)

lemma 4

Let

$$X_k \subseteq A_k \subseteq S, \quad k = 0, 1, 2, \ldots, \tag{7.6.31}$$

with all A_k pairwise disjoint.

Let $A_k \in \mathcal{M}^*$ for $k \geq 1$. (A_0 and the X_k need not be $\mathcal{M}^*$-sets.) Then

$$m^*\left(\bigcup_{k=0}^{\infty} X_k\right) = \sum_{k=0}^{\infty} m^* X_k. \tag{7.6.32}$$

Proof

We start with two sets, A_0 and A_1; so

$$A_1 \in \mathcal{M}^*, A_0 \cap A_1 = \emptyset, X_0 \subseteq A_0, \text{ and } X_1 \subseteq A_1. \tag{7.6.33}$$

As $A_0 \cap A_1 = \emptyset$, we have $A_0 \subseteq -A_1$; hence also $X_0 \subseteq -A_1$.

since $A_1 \in \mathcal{M}^*$, we use formula (2), with

$$X=X_1\subseteq A_1 \text{ and } Y=X_0\subseteq -A, \tag{7.6.34}$$

to obtain

$$m^*\left(X_0\cup X_1\right)=m^*X_0+m^*X_1. \tag{7.6.35}$$

Thus (6) holds for two sets.

Induction now easily yields

$$(\forall n)\sum_{k=0}^{n} m^*X_k=m^*\left(\bigcup_{k=0}^{n} X_k\right)\leq m^*\left(\bigcup_{k=0}^{\infty} X_k\right) \tag{7.6.36}$$

by monotonicity of m^*. Now let $n\rightarrow\infty$ and pass to the limit to get

$$\sum_{k=0}^{\infty} m^*X_k\leq m^*\left(\bigcup_{k=0}^{\infty} X_k\right). \tag{7.6.37}$$

As $\bigcup X_k$ is covered by the X_k, the σ-subadditivity of m^* yields the reverse inequality as well. Thus (6) is proved. $\square$

Proof of Theorem 1. As we noted, $\mathcal{M}^*$ is a field. To show that it is also closed under countable unions (a σ-field), let

$$U=\bigcup_{k=1}^{\infty} A_k,\quad A_k\in\mathcal{M}^*. \tag{7.6.38}$$

We have to prove that $U\in\mathcal{M}^*;$ or by (2) and Note 2,

$$(\forall X\subseteq U)(\forall Y\subseteq -U)\quad m^*(X\cup Y)\geq m^*X+m^*Y. \tag{7.6.39}$$

We may safely assume that the A_k are disjoint. (If not, replace them by disjoint sets $B_k\in\mathcal{M}^*$, as in Corollary 1 §1.)

To prove (7), fix any $X\subseteq U$ and $Y\subseteq -U,$ and let

$$X_k=X\cap A_k\subseteq A_k, \tag{7.6.40}$$

$A_0=-U,$ and $X_0=Y,$ satisfying all assumptions of Lemma 4. Thus by (6), writing the first term separately, we have

$$m^*\left(Y\cup\bigcup_{k=1}^{\infty} X_k\right)=m^*Y+\sum_{k=1}^{\infty} m^*X_k. \tag{7.6.41}$$

But

$$\bigcup_{k=1}^{\infty} X_k=\bigcup_{k=1}^{\infty}\left(X\cap A_k\right)=X\cap\bigcup_{k=1}^{\infty} A_k=X\cap U=X \tag{7.6.42}$$

(as $X\subseteq U$). Also, by σ-subadditivity,

$$\sum m^*X_k\geq m^*\bigcup X_k=m^*X. \tag{7.6.43}$$

Therefore, (8) implies (7); so $\mathcal{M}^*$ is a σ-field.

Moreover, m^* is σ-additive on $\mathcal{M}^*$, as follows from Lemma 4 by taking

$$X_k=A_k\in\mathcal{M}^*, A_0=\emptyset. \tag{7.6.44}$$

Thus m^* acts as a measure on $\mathcal{M}^*$.

By Lemma 2, m^* is complete; for if X is "null" ($X\subseteq A$ and $m^*A=0$), then $m^*X=0;$ so $X\in\mathcal{M}^*$, as required.

Thus all is proved. $\square$

We thus have a standard method for constructing measures: From a premeasure

$$\mu:\mathcal{C}\rightarrow E^* \tag{7.6.45}$$

in S, we obtain the μ-induced outer measure

$$m^{*}: 2^{S} \rightarrow E^{*}\ (\S 5); \tag{7.6.46}$$

this, in turn, induces a complete measure

$$m: \mathcal{M}^{*} \rightarrow E^{*}. \tag{7.6.47}$$

But we need more: We want m to be an extension of μ, i.e.,

$$m=\mu \text{ on } \mathcal{C}, \tag{7.6.48}$$

with $\mathcal{C} \subseteq \mathcal{M}^{*}$ (meaning that all $\mathcal{C}$-sets are m^{*}-measurable). We now explore this question.

lemma 5

Let $(S, \mathcal{C}, \mu)$ and m^{*} be as in Definition 3 of §5. Then for a set $A \subseteq S$ to be m^{*}-measurable, it suffices that

$$m^{*} X \geq m^{*}(X \cap A)+m^{*}(x-A) \quad \text{for all } X \in \mathcal{C}. \tag{7.6.49}$$

Proof

We must show that (9) holds for any $X \subseteq S,$ even not a $\mathcal{C}$-set.

This is trivial if $m^{*} X=\infty$. Thus assume $m^{*} X<\infty$ and fix any $\varepsilon>0$.

By Note 3 in §5, X must have a basic covering $\left\{B_{n}\right\} \subseteq \mathcal{C}$ so that

$$X \subseteq \bigcup_{n} B_{n} \tag{7.6.50}$$

and

$$m^{*} X+\varepsilon>\sum \mu B_{n} \geq \sum m^{*} B_{n}. \tag{7.6.51}$$

(Explain!)

Now, as $X \subseteq \cup B_{n},$ we have

$$X \cap A \subseteq \bigcup B_{n} \cap A=\bigcup\left(B_{n} \cap A\right). \tag{7.6.52}$$

Similarly,

$$X-A=X \cap-A \subseteq \bigcup\left(B_{n}-A\right). \tag{7.6.53}$$

Hence, as m^{*} is σ-subadditive and monotone, we get

$$\begin{aligned} m^{*}(X \cap A)+m^{*}(X-A) & \leq m^{*}\left(\bigcup\left(B_{n} \cap A\right)\right)+m^{*}\left(\bigcup\left(B_{n}-A\right)\right) \\ & \leq \sum\left[m^{*}\left(B_{n} \cap A\right)+m^{*}\left(B_{n}-A\right)\right]. \end{aligned}$$

But by assumption, (9) holds for any $\mathcal{C}$-set, hence for each B_{n}. Thus

$$m^{*}\left(B_{n} \cap A\right)+m^{*}\left(B_{n}-A\right) \leq m^{*} B_{n}, \tag{7.6.54}$$

and (11) yields

$$m^{*}(X \cap A)+m^{*}(X-A) \leq \sum\left[m^{*}\left(B_{n} \cap A\right)+m^{*}\left(B_{n}-A\right)\right] \leq \sum m^{*} B_{n}. \tag{7.6.55}$$

Therefore, by (10),

$$m^{*}(X \cap A)+m^{*}(X-A) \leq m^{*} X+\varepsilon. \tag{7.6.56}$$

Making $\varepsilon \rightarrow 0,$ we prove (10) for any $X \subseteq S,$ so that $A \in \mathcal{M}^{*},$ as required. □

Theorem 7.6.2

Let the premeasure

$$\mu : \mathcal{C} \rightarrow E^* \tag{7.6.57}$$

be σ-additive on $\mathcal{C}$, a semiring in S. Let m^* be the μ-induced outer measure, and

$$m : \mathcal{M}^* \rightarrow E^* \tag{7.6.58}$$

be the m^*-induced measure. Then

(i) $\mathcal{C} \subseteq \mathcal{M}^*$ and

(ii) $\mu = m^* = m$ on $\mathcal{C}$.

Thus m is a σ-additive extension of μ (called its Lebesgue extension) to $\mathcal{M}^*$.

Proof

By Corollary 2 in §5, μ is also σ-subadditive on the semiring $\mathcal{C}$. Thus by Theorem 2 in §5, $\mu = m^*$ on $\mathcal{C}$.

To prove that $\mathcal{C} \subseteq \mathcal{M}^*$, we fix $A \in \mathcal{C}$ and show that A satisfies (9), so that $A \in \mathcal{M}^*$.

Thus take any $X \in \mathcal{C}$. As $\mathcal{C}$ is a semiring, $X \cap A \in \mathcal{C}$ and

$$X - A = \bigcup_{k=1}^{n} A_k \text{ (disjoint)} \tag{7.6.59}$$

for some sets $A_k \in \mathcal{C}$. Hence

$$\begin{aligned} m^*(X \cap A) + m^*(X - A) &= m^*(X \cap A) + m^* \bigcup_{k=1}^{n} A_k \\ &\leq m^*(X \cap A) + \sum_{k=1}^{n} m^* A_k. \end{aligned}$$

As

$$X = (X \cap A) \cup (X - A) = (X \cap A) \cup \bigcup A_k \text{ (disjoint)}, \tag{7.6.60}$$

the additivity of μ and the equality $\mu = m^*$ on $\mathcal{C}$ yield

$$m^* X = m^*(X \cap A) + \sum_{k=1}^{n} m^* A_k. \tag{7.6.61}$$

Hence by (12),

$$m^* X \geq m^*(X \cap A) + m^*(X - A); \tag{7.6.62}$$

so by Lemma 5, $A \in \mathcal{M}^*$, as required.

Also, by definition, $m = m^*$ on $\mathcal{M}^*$, hence on $\mathcal{C}$. Thus

$$\mu = m^* = m \text{ on } \mathcal{C}, \tag{7.6.63}$$

as claimed. □

Note 3. In particular, Theorem 2 applies if

$$\mu : \mathcal{M} \rightarrow E^* \tag{7.6.64}$$

is a measure (so that $\mathcal{C} = \mathcal{M}$ is even a σ-ring).

Thus any such μ can be extended to a complete measure m (its Lebesgue extension) on a σ-field

$$\mathcal{M}^* \supseteq \mathcal{M} \tag{7.6.65}$$

via the μ-induced outer measure (call it μ^* this time), with

$$\mu^* = m = \mu \text{ on } \mathcal{M}. \tag{7.6.66}$$

Moreover,

$$\mathcal{M}^* \supseteq \mathcal{M} \supseteq \mathcal{M}_\sigma \tag{7.6.67}$$

(see Note 2 in §3); so μ^* is $\mathcal{M}$-regular and $\mathcal{M}^*$-regular (Theorem 3 of §5).

Note 4. A reapplication of this process to m does not change m (Problem 16).

7.6.E: Problems on Measures and Outer Measures

Exercise $7.6.E.1$

Show that formulas (1) and (2) are equivalent.
[Hints: (i) Assume (1) and let $X \subseteq A, Y \subseteq -A$.
As X in (1) is arbitrary, we may replace it by $X \cup Y$. Simplifying, obtain (2) on noting that $X \cap A=X, X \cap-A=\emptyset, Y \cap A=\emptyset,$ and $Y \cap-A=Y$.
(ii) Assume (2). Take any X and substitute $X \cap A$ and $X-A$ for X and Y in (2).]

Exercise $7.6.E.2$

Given an outer measure space $\left(S, \mathcal{M}^{*}, m^{*}\right)$ and $A \subseteq S$, set

$$A \cap \mathcal{M}^{*}=\left\{A \cap X | X \in \mathcal{M}^{*}\right\} \tag{7.6.E.1}$$

(SYMBOL!)
(all sets of the form $A \cap X$ with $X \in \mathcal{M}^{*}$).
Prove that $A \cap \mathcal{M}^{*}$ is a σ-field in A, and m^{*} is σ-additive on it. (SYMBOL!)
[Hint: Use Lemma 4, with $X_{k}=A \cap A_{k} \in A \cap \mathcal{M}^{*}$.] (SYMBOL!)

Exercise $7.6.E.3$

Prove Lemmas 1 and 2, using formula (1).

Exercise $7.6.E.3'$

Prove Corollary 1.

Exercise $7.6.E.4$

Verify Examples (b),(c), and (d). Why is m an outer measure as well?
[Hint: Use Corollary 2 in §5.]

Exercise $7.6.E.5$

Fill in all details (induction, etc.) in the proofs of this section.

Exercise $7.6.E.6$

Verify that m^{*} is an outer measure and describe $\mathcal{M}^{*}$ under each of the following conditions.
(a) $m^{*} A=1$ if $\emptyset \subset A \subseteq S ; m^{*} \emptyset=0$.
(b) $m^{*} A=1$ if $\emptyset \subset A \subset S ; m^{*} S=2 ; m^{*} \emptyset=0$.
(c) $m^{*} A=0$ if $A \subseteq S$ is countable; $m^{*} A=1$ otherwise (S is uncountable).
(d) $S=N$ (naturals); $m^{*} A=1$ if A is infinite; $m^{*} A=\frac{n}{n+1}$ if A has n elements.

Exercise $7.6.E.7$

Prove the following.
(i) An outer measure m^{*} is $\mathcal{M}^{*}$-regular (Definition 5 in §5) iff

$$\left(\forall A \subseteq S\right)\left(\exists B \in \mathcal{M}^{*}\right) \quad A \subseteq B \text{ and } m^{*} A=m B. \tag{7.6.E.2}$$

B is called a measurable cover of A.
[Hint: If

$$m^{*} A=\inf \left\{m X \mid A \subseteq X \in \mathcal{M}^{*}\right\}, \tag{7.6.E.3}$$

then

$$(\forall n)\left(\exists X_{n} \in \mathcal{M}^{*}\right) \quad A \subseteq X_{n} \text { and } m X_{n} \leq m^{*} A+\frac{1}{n}. \tag{7.6.E.4}$$

Set $B=\bigcap_{n=1}^{\infty} X_{n}$.]
(ii) If m^{*} is as in Definition 3 of §5, with $\mathcal{C} \subseteq \mathcal{M}^{*}$, then m^{*} is $\mathcal{M}^{*}$-regular.

Exercise $7.6.E.8$

Show that if m^{*} is $\mathcal{M}^{*}$-regular (Problem 7), it is left continuous.
[Hints: Let $\left\{A_{n}\right\} \uparrow$; let B_{n} be a measurable cover of A_{n}; set

$$C_{n}=\bigcap_{k=n}^{\infty} B_{k}. \tag{7.6.E.5}$$

Verify that $\left\{C_{n}\right\} \uparrow, B_{n} \supseteq C_{n} \supseteq A_{n},$ and $m C_{n}=m^{*} A_{n}$.
By the left continuity of m (Theorem 2 in §4),

$$\lim m^{*} A_{n}=\lim m C_{n}=m \bigcup_{n=1}^{\infty} C_{n} \geq m^{*} \bigcup_{n=1}^{\infty} A_{n}. \tag{7.6.E.6}$$

Prove the reverse inequality as well.]

Exercise $7.6.E.9$

Continuing Problems 6-8, verify the following.
(i) In 6(a), with $S=N, m^{*}$ is $\mathcal{M}^{*}$-regular, but not right continuous.
Hint: Take $A_{n}=\{x \in N \mid x \geq n\}$.
(ii) In 6(b), with $S=N, m^{*}$ is neither $\mathcal{M}^{*}$-regular nor left continuous.
(iii) In 6(d), m^{*} is not $\mathcal{M}^{*}$-regular; yet it is left continuous. (Thus Problem 8 is not a necessary condition.)

Exercise $7.6.E.10$

In Problem 2, let n^{*} be the restriction of m^{*} to 2^{A}. Prove the following.
(a) n^{*} is an outer measure in A.
(b) $A \cap \mathcal{M}^{*} \subseteq \mathcal{N}^{*}=\left\{n^{*}\text{-measurable sets}\right\}$. (SYMBOL!)
(c) $A \cap \mathcal{M}^{*}=\mathcal{N}^{*}$ if $A \in \mathcal{M}^{*}$, or if m^{*} is $\mathcal{M}^{*}$-regular (see Problem 7) and finite. (SYMBOL!)
(d) n^{*} is $\mathcal{N}^{*}$-regular if m^{*} is $\mathcal{M}^{*}$-regular.

Exercise $7.6.E.11$

Show that if m^{*} is $\mathcal{M}^{*}$-regular and finite, then $A \subseteq S$ is m^{*}-measurable iff

$$m S=m^{*} A+m^{*}(-A). \tag{7.6.E.7}$$

[Hint: Assume the latter. By Problem 7,

$$(\forall X \subseteq S)\,(\exists B \in \mathcal{M}^*, B \supseteq X) \quad m^* X = mB; \tag{7.6.E.8}$$

so

$$m^* A = m^*(A \cap B) + m^*(A - B). \tag{7.6.E.9}$$

Similarly for $-A$. Deduce that

$$m^*(A \cap B) + m^*(A - B) + m^*(B - A) + m^*(-A - B) = mS = mB + m(-B); \tag{7.6.E.10}$$

hence

$$m^* X = mB \geq m^*(B \cap A) + m^*(B - A) \geq m^*(X \cap A) + m^*(X - A), \tag{7.6.E.11}$$

so $A \in \mathcal{M}^*$.]

Exercise $7.6.E.12$

Using Problem 15 in §5, prove that if m^* has the CP then each open set $G \subseteq S$ is in $\mathcal{M}^*$.
[Outline: Show that

$$(\forall X \subseteq G)(\forall Y \subseteq -G) \quad m^*(X \cup Y) \geq m^* X + m^* Y, \tag{7.6.E.12}$$

assuming $m^* X < \infty$. (Why?) Set

$$D_0 = \{x \in X | \rho(x, -G) \geq 1\} \tag{7.6.E.13}$$

and

$$D_k = \left\{x \in X | \frac{1}{k+1} \leq \rho(x, -G) < \frac{1}{k}\right\}, \quad k \geq 1. \tag{7.6.E.14}$$

Prove that

$$X = \bigcup_{k=0}^{\infty} D_k \tag{7.6.E.15}$$

and

$$\rho\left(D_k, D_{k+2}\right) > 0; \tag{7.6.E.16}$$

so by Problem 15 in §5,

$$\sum_{n=0}^{\infty} m^* D_{2n} = m^* \bigcup_{n=0}^{\infty} D_{2n} \leq m^* \bigcup_{n=0}^{\infty} D_n = m^* X < \infty. \tag{7.6.E.17}$$

Similarly,

$$\sum_{n=0}^{\infty} m^* D_{2n+1} \leq m^* X < \infty. \tag{7.6.E.18}$$

Hence

$$\sum_{n=0}^{\infty} m^* D_n < \infty; \tag{7.6.E.19}$$

so

$$\lim_{n \to \infty} \sum_{k=n}^{\infty} m^* D_k = 0. \tag{7.6.E.20}$$

(Why?) Thus

$$(\forall \varepsilon > 0)(\exists n) \sum_{k=n}^{\infty} m^* D_k < \varepsilon. \tag{7.6.E.21}$$

Also,

$$X = \bigcup_{k=0}^{\infty} D_k = \bigcup_{k=0}^{n-1} D_k \cup \bigcup_{k=n}^{\infty} D_k; \tag{7.6.E.22}$$

so

$$m^* X \leq m^* \bigcup_{k=0}^{n-1} D_k + \sum_{k=n}^{\infty} m^* D_k < m^* \bigcup_{k=0}^{n-1} D_k + \varepsilon. \tag{7.6.E.23}$$

Adding $m^* Y$ on both sides, get

$$m^* X + m^* Y \leq m^* \bigcup_{k=0}^{n-1} D_k + m^* Y + \varepsilon. \tag{7.6.E.24}$$

Moreover,

$$\rho\left(\bigcup_{k=0}^{n-1} D_k, Y\right) > 0, \tag{7.6.E.25}$$

for $Y \subseteq -G$ and

$$\rho\left(D_k, -G\right) \geq \frac{1}{k+1}. \tag{7.6.E.26}$$

Hence by the CP,

$$m^* Y + \sum_{k=0}^{n-1} m^* D_k = m^*\left(Y \cup \bigcup_{k=0}^{n-1} D_k\right) < m^*(Y \cup X). \tag{7.6.E.27}$$

(Why?) Combining with (iii), obtain

$$m^*X + m^*Y \leq m^*(X \cup Y) + \varepsilon. \tag{7.6.E.28}$$

Now let $\varepsilon \to 0$.]

Exercise $7.6.E.13$

$\Rightarrow$ Show that if $m : \mathcal{M} \to E^*$ is a measure, there is $P \in \mathcal{M}$, with

$$mP = \max\{mX \mid X \in \mathcal{M}\}. \tag{7.6.E.29}$$

[Hint: Let

$$k = \sup\{mX \mid X \in \mathcal{M}\} \tag{7.6.E.30}$$

in E^*. As $k \geq 0$, there is a sequence $r_n \nearrow k, r_n < k$. (If $k = \infty$, set $r_n = n$; if $k < \infty, r_n = k - \frac{1}{n}$.) By lub properties,

$$(\forall n)\,(\exists X_n \in \mathcal{M}) \quad r_n < mX_n \leq k, \tag{7.6.E.31}$$

with $\{X_n\} \uparrow$ (Problem 9 in §3). Set

$$P = \bigcup_{n=1}^{\infty} X_n. \tag{7.6.E.32}$$

Show that

$$mP = \lim_{n \to \infty} mX_n = k.] \tag{7.6.E.33}$$

Exercise $7.6.E.14$

$\Rightarrow^*$ Given a measure $m : \mathcal{M} \to E^*$, let

$$\overline{\mathcal{M}} = \{\text{all sets of the form } X \cup Z \text{ where } X \in \mathcal{M} \text{ and } Z \text{ is } m\text{-null}\}. \tag{7.6.E.34}$$

Prove that $\overline{\mathcal{M}}$ is a σ-ring $\supseteq \mathcal{M}$.
[Hint: To prove that

$$(\forall A, B \in \overline{\mathcal{M}}) \quad A - B \in \overline{\mathcal{M}}, \tag{7.6.E.35}$$

suppose first $A \in \mathcal{M}$ and B is "null," i.e., $B \subseteq U \in \mathcal{M}, mU = 0$.
Show that

$$A - B = X \cup Z, \tag{7.6.E.36}$$

with $X = A - U \in \mathcal{M}$ and $Z = A \cap U - B\, m$-null (Z is shaded in Figure 31).

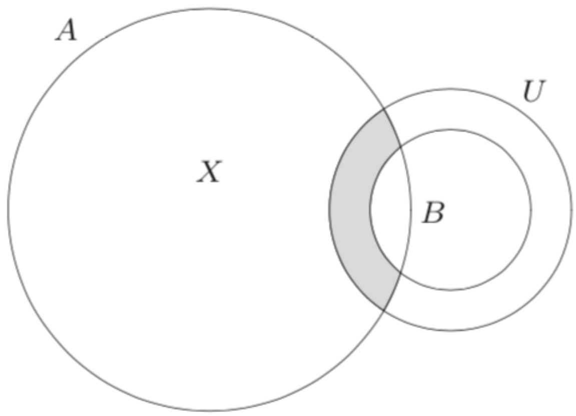

FIGURE 31

Next, if $A, B \in \overline{\mathcal{M}}$, let $A = X \cup Z \;\; B = X' \cup Z'$, where $X, X' \in \mathcal{M}$ and Z, Z' are m-null. Hence

$$\begin{aligned} A - B &= (X \cup Z) - B \\ &= (X - B) \cup (Z - B) \\ &= (X - B) \cup Z'', \end{aligned}$$

where

$$Z'' = Z - B \tag{7.6.E.37}$$

is m-null. Also, $B = X' \cup Z'$ implies

$$X - B = (X - X') - Z' \in \overline{\mathcal{M}}, \tag{7.6.E.38}$$

by the first part of the proof.
Deduce that

$$A - B = (X - B) \cup Z'' \in \overline{\mathcal{M}} \tag{7.6.E.39}$$

(after checking closure under unions).]

Exercise $7.6.E.15$

$\Rightarrow^*$ Continuing Problem 14, define $\overline{m} : \overline{\mathcal{M}} \rightarrow E^*$ by setting $\overline{m}A = mX$ whenever $A = X \cup Z$, with $X \in \mathcal{M}$ and Z m-null. (Show that $\overline{m}A$ does not depend on the particular representation of A as $X \cup Z$.)
Prove the following.
(i) $\overline{m}$ is a complete measure (called the completion of m), with $\overline{m} = m$ on $\mathcal{M}$.
(ii) $\overline{m}$ is the least complete extension of m; that is, if $n : \mathcal{N} \rightarrow E^*$ is another complete measure, with $\mathcal{M} \subseteq \mathcal{N}$ and $n = m$ on $\mathcal{M}$, then $\overline{\mathcal{M}} \subseteq \mathcal{N}$ and $n = \overline{m}$ on $\overline{\mathcal{M}}$.
(iii) $m = \overline{m}$ iff m is complete.

Exercise $7.6.E.16*$

Show that if $m : \mathcal{M}^* \rightarrow E^*$ is induced by an $\mathcal{M}^*$-regular outer measure μ^*, then m equals its Lebesgue extension m' and completion $\overline{m}$ (see Problem 15).
[Hint: By Definition 3 in §5, m induces an outer measure m^*. By Theorem 3 in §5,

$$m^*A = \inf\{mX | A \subseteq X \in \mathcal{M}^*\} = \mu^*A \tag{7.6.E.40}$$

(for μ^* is $\mathcal{M}^*$-regular).
As $m^* = \mu^*$, we get $m' = m$. Also, $m = \overline{m}$, by Problem 15(iii).]

Exercise $7.6.E.17*$

Prove that if a measure $\mu : \mathcal{M} \rightarrow E^*$ is σ-finite (Definition 4 in §5), with $S \in \mathcal{M}$, then its Lebesgue extension $m : \mathcal{M}^* \rightarrow E^*$ equals its completion $\overline{\mu}$ (see Problem 15).
[Outline: It suffices to prove $\mathcal{M}^* \subseteq \overline{\mathcal{M}}$. (Why?)
To start with, let $A \in \mathcal{M}^*, mA < \infty$. By Problem 12 in §5,

$$(\exists B \in \mathcal{M}) \quad A \subseteq B \text{ and } m^*A = mA = mB < \infty; \tag{7.6.E.41}$$

so

$$m(B - A) = mB - mA = 0. \tag{7.6.E.42}$$

Also,

$$(\exists H \in \mathcal{M}) \quad B - A \subseteq H \text{ and } \mu H = m(B - A) = 0. \tag{7.6.E.43}$$

Thus $B - A$ is μ-null; so $B - A \in \overline{\mathcal{M}}$. (Why?) Deduce that

$$A = B - (B - A) \in \overline{\mathcal{M}}. \tag{7.6.E.44}$$

Thus $\overline{\mathcal{M}}$ contains any $A \in \mathcal{M}^*$ with $mA < \infty$. Use the σ-finiteness of μ to show

$$(\forall x \in \mathcal{M}^*) (\exists \{A_n\} \subseteq \mathcal{M}^*) \quad mA_n < \infty \text{ and } X = \bigcup_n A_n \in \overline{\mathcal{M}}.\Bigg] \tag{7.6.E.45}$$

7.7: Topologies. Borel Sets. Borel Measures

I. Our theory of set families leads quite naturally to a generalization of metric spaces. As we know, in any such space (S, ρ), there is a family $\mathcal{G}$ of open sets, and a family $\mathcal{F}$ of all closed sets. In Chapter 3, §12, we derived the following two properties.

(i) $\mathcal{G}$ is closed under any (even uncountable) unions and under finite intersections (Chapter 3, §12, Theorem 2). Moreover,

$$\emptyset \in \mathcal{G} \text{ and } S \in \mathcal{G}, \tag{7.7.1}$$

(ii) $\mathcal{F}$ has these properties, with "unions" and "intersections" interchanged (Chapter 3, §12, Theorem 3). Moreover, by definition,

$$A \in \mathcal{F} \text{ iff } -A \in \mathcal{G}. \tag{7.7.2}$$

Now, quite often, it is not so important to have distances (i.e., a metric) defined in S, but rather to single out two set families, $\mathcal{G}$ and $\mathcal{F}$, with properties (i) and (ii), in a suitable manner. For examples, see Problems 1 to 4 below. Once $\mathcal{G}$ and $\mathcal{F}$ are given, one does not need a metric to define such notions as continuity, limits, etc. (See Problems 2 and 3.) This leads us to the following definition.

Definition 1

A topology for a set S is any set family $\mathcal{G} \subseteq 2^S$, with properties (i).

The pair $(S, \mathcal{G})$ then is called a topological space. If confusion is unlikely, we simply write S for $(S, \mathcal{G})$.

$\mathcal{G}$-sets are called open sets; their complements form the family $\mathcal{F}$ (called cotopology) of all closed sets in S; $\mathcal{F}$ satisfies (ii) (the proof is as in Theorem 3 of Chapter 3, §12).

Any metric space may be treated as a topological one (with $\mathcal{G}$ defined as in Chapter 3, §12), but the converse is not true. Thus $(S, \mathcal{G})$ is more general.

Note 1. By Problem 15 in Chapter 4, §2, a map

$$f:(S, \rho) \rightarrow (T, \rho') \tag{7.7.3}$$

is continuous iff $f^{-1}[B]$ is open in S whenever B is open in T.

We adopt this as a definition, for topological spaces S, T.

Many other notions (neighborhoods, limits, etc.) carry over from metric spaces by simply treating G_p as "an open set containing p." (See Problem 3.)

Note 2. By (i), $\mathcal{G}$ is surely closed under countable unions. Thus by Note 2 in §3,

$$\mathcal{G} = \mathcal{G}_\sigma. \tag{7.7.4}$$

Also, $\mathcal{G} = \mathcal{G}_d$ and

$$\mathcal{F}_\delta = \mathcal{F} = \mathcal{F}_s, \tag{7.7.5}$$

but not

$$\mathcal{G} = \mathcal{G}_\delta \text{ or } \mathcal{F} = \mathcal{F}_\sigma \tag{7.7.6}$$

in general.

$\mathcal{G}$ and $\mathcal{F}$ need not be rings or σ-rings (closure fails for differences). But by Theorem 2 in §3, $\mathcal{G}$ and $\mathcal{F}$ can be "embedded" in a smallest σ-ring. We name it in the following definition.

Definition 2

The σ-ring $\mathcal{B}$ generated by a topology $\mathcal{G}$ in S is called the Borel field in S. (It is a σ-field, as $S \in \mathcal{G} \subseteq \mathcal{B}$.)

Equivalently, $\mathcal{B}$ is the least σ-ring $\supseteq \mathcal{F}$. (Why?)

$\mathcal{B}$-sets are called Borel sets in $(S, \mathcal{G})$.

As $\mathcal{B}$ is closed under countable unions and intersections, we have not only

$$\mathcal{B} \supseteq \mathcal{G} \text{ and } \mathcal{B} \supseteq \mathcal{F}, \tag{7.7.7}$$

but also

$$\mathcal{B} \supseteq \mathcal{G}_{\delta}, \mathcal{B} \supseteq \mathcal{F}_{\sigma}, \mathcal{B} \supseteq \mathcal{G}_{\delta\sigma} \left[\text{ i.e. }, \left(\mathcal{G}_{\delta}\right)_{\sigma}\right], \mathcal{B} \supseteq \mathcal{F}_{\sigma\delta}, \text{ etc. } \tag{7.7.8}$$

Note that

$$\mathcal{G}_{\delta\delta} = \mathcal{G}_{\delta}, \mathcal{F}_{\sigma\sigma} = \mathcal{F}_{\sigma}, \text{ etc. (Why?)} \tag{7.7.9}$$

II. Special notions apply to measures in metric and topological spaces.

Definition 3

A measure $m : \mathcal{M} \rightarrow E^{*}$ in $(S, \mathcal{G})$ is called topological iff $\mathcal{G} \subseteq \mathcal{M},$ i.e., all open sets are measurable; m is a Borel measure iff $\mathcal{M} = \mathcal{B}$.

Note 3. If $\mathcal{G} \subseteq \mathcal{M}$ (a σ-ring), then also $\mathcal{B} \subseteq \mathcal{M}$ since $\mathcal{B}$ is, by definition, the least σ-ring $\supseteq \mathcal{G}$.

Thus m is topological iff $\mathcal{B} \subseteq \mathcal{M}$ (hence surely $\mathcal{F} \subseteq \mathcal{M}, \mathcal{G}_{\delta} \subseteq \mathcal{M}, \mathcal{F}_{\sigma} \subseteq \mathcal{M}$, etc.).

It also follows that any topological measure can be restricted to $\mathcal{B}$ to obtain a Borel measure, called its Borel restriction.

Definition 4

A measure $m : \mathcal{M} \rightarrow E^{*}$ in $(S, \mathcal{G})$ is called regular iff it is regular with respect to $\mathcal{M} \cap \mathcal{G},$ the measurable open sets; i.e.,

$$(\forall A \in \mathcal{M}) \quad mA = \inf\{mX \mid A \subseteq X \in \mathcal{M} \cap \mathcal{G}\}. \tag{7.7.10}$$

If m is topological $(\mathcal{G} \subseteq \mathcal{M}),$ this simplifies to

$$mA = \inf\{mX \mid A \subseteq X \in \mathcal{G}\}, \tag{7.7.11}$$

i.e., m is $\mathcal{G}$-regular (Definition 5 in §5).

Definition 5

A measure m is strongly regular iff for any $A \in \mathcal{M}$ and $\varepsilon > 0,$ there is an open set $G \in \mathcal{M}$ and a closed set $F \in \mathcal{M}$ such that

$$F \subseteq A \subseteq G, \text{ with } m(A-F) < \varepsilon \text{ and } m(G-A) < \varepsilon; \tag{7.7.12}$$

thus A can be "approximated" by open supersets and closed subsets, both measurable. As is easily seen, this implies regularity.

A kind of converse is given by the following theorem.

Theorem 7.7.1

If a measure $m : \mathcal{M} \rightarrow E^{*}$ in $(S, \mathcal{G})$ is regular and σ-finite (see Definition 4 in §5), with $S \in \mathcal{M},$ then m is also strongly regular.

Proof

Fix $\varepsilon > 0$ and let $mA < \infty$.

By regularity,

$$mA = \inf\{mX \mid A \subseteq X \in \mathcal{M} \cap \mathcal{G}\}; \tag{7.7.13}$$

so there is a set $X \in \mathcal{M} \cap \mathcal{G}$ (measurable and open), with

$$A \subseteq X \text{ and } mX < mA + \varepsilon. \tag{7.7.14}$$

Then

$$m(X-A)=mX-mA<\varepsilon, \tag{7.7.15}$$

and X is the open set G required in (2).

If, however, $mA=\infty,$ use σ-finiteness to obtain

$$A \subseteq \bigcup_{k=1}^{\infty} X_k \tag{7.7.16}$$

for some sets $X_k \in \mathcal{M}, mX_k<\infty;$ so

$$A=\bigcup_k \left(A \cap X_k\right). \tag{7.7.17}$$

Put

$$A_k=A \cap X_k \in \mathcal{M}. \tag{7.7.18}$$

(Why?) Then

$$A=\bigcup_k A_k, \tag{7.7.19}$$

and

$$mA_k \leq mX_k<\infty. \tag{7.7.20}$$

Now, by what was proved above, for each A_k there is an open measurable $G_k \supseteq A_k,$ with

$$m\left(G_k-A_k\right)<\frac{\varepsilon}{2^k}, \tag{7.7.21}$$

Set

$$G=\bigcup_{k=1}^{\infty} G_k. \tag{7.7.22}$$

Then $G \in \mathcal{M} \cap \mathcal{G}$ and $G \supseteq A.$ Moreover,

$$G-A=\bigcup_k G_k-\bigcup_k A_k \subseteq \bigcup_k \left(G_k-A_k\right). \tag{7.7.23}$$

(Verify!) Thus by σ-subadditivity,

$$m(G-A) \leq \sum_k m\left(G_k-A_k\right)<\sum_{k=1}^{\infty} \frac{\varepsilon}{2^k}=\varepsilon, \tag{7.7.24}$$

as required.

To find also the closed set F, consider

$$-A=S-A \in \mathcal{M}. \tag{7.7.25}$$

As shown above, there is an open measurable set $G' \supseteq -A,$ with

$$\varepsilon>m\left(G'-(-A)\right)=m\left(G' \cap A\right)=m\left(A-\left(-G'\right)\right). \tag{7.7.26}$$

Then

$$F=-G' \subseteq A \tag{7.7.27}$$

is the desired closed set, with $m(A-F)<\varepsilon.$ □

Theorem 7.7.2

If $m : \mathcal{M} \rightarrow E^*$ is a strongly regular measure in $(S, \mathcal{G})$, then for any $A \in \mathcal{M}$, there are measurable sets $H \in \mathcal{F}_\sigma$ and $K \in \mathcal{G}_\delta$ such that

$$H \subseteq A \subseteq K \text{ and } m(A-H)=0=m(K-A); \tag{7.7.28}$$

hence

$$mA=mH=mK. \tag{7.7.29}$$

Proof

Let $A \in \mathcal{M}$. By strong regularity, given $\varepsilon_n=1/n$, one finds measurable sets

$$G_n \in \mathcal{G} \text{ and } F_n \in \mathcal{F}, \quad n=1,2,\ldots, \tag{7.7.30}$$

such that

$$F_n \subseteq A \subseteq G_n \tag{7.7.31}$$

and

$$m\left(A-F_n\right)<\frac{1}{n} \text{ and } m\left(G_n-A\right)<\frac{1}{n}, \quad n=1,2,\ldots. \tag{7.7.32}$$

Let

$$H=\bigcup_{n=1}^{\infty} F_n \text{ and } K=\bigcap_{n=1}^{\infty} G_n. \tag{7.7.33}$$

Then $H, K \in \mathcal{M}, H \in \mathcal{F}_\sigma, K \in \mathcal{G}_\delta,$ and

$$H \subseteq A \subseteq K. \tag{7.7.34}$$

Also, $F_n \subseteq H$ and $G_n \supseteq K$.

Hence

$$A-H \subseteq A-F_n \text{ and } K-A \subseteq G_n-A; \tag{7.7.35}$$

so by (4),

$$m(A-H)<\frac{1}{n} \rightarrow 0 \text{ and } m(K-A)<\frac{1}{n} \rightarrow 0. \tag{7.7.36}$$

Finally,

$$mA=m(A-H)+mH=mH, \tag{7.7.37}$$

and similarly $mA=mK$.

Thus all is proved. □

7.7.E: Problems on Topologies, Borel Sets, and Regular Measures

Exercise $7.7.E.1$

Show that $\mathcal{G}$ is a topology in S (in $(a)-(c),$ describe $\mathcal{B}$ also), given
(a) $\mathcal{G}=2^{S}$;
(b) $\mathcal{G}=\{\emptyset, S\}$;
(c) $\mathcal{G}=\{\emptyset$ and all sets in $S,$ containing a fixed point $p\};$ or
(d) $S=E^{*}; \mathcal{G}$ consists of all possible unions of sets of the form $(a, b),(a, \infty]$, and $[-\infty, b)$, with $a, b \in E^{1}$.

Exercise $7.7.E.2$

(S, ρ) is called a pseudometric space (and ρ is a pseudometric) iff the metric laws (i)-(iii) of Chapter 3, s 11 hold, but (i) is weakened to

$$\rho(x, x)=0 \tag{7.7.E.1}$$

(so that $\rho(x, y)$ may be 0 even if $x \neq y$).
(a) Define "globes," "interiors," and "open sets" (i.e., $\mathcal{G}$) as in Chapter 3, §12; then show that $\mathcal{G}$ is a topology for S.
(b) Let $S=E^{2}$ and

$$\rho(\overline{x}, \overline{y})=\left|x_{1}-y_{1}\right|, \tag{7.7.E.2}$$

where $\overline{x}=\left(x_{1}, x_{2}\right)$ and $\overline{y}=\left(y_{1}, y_{2}\right)$. Show that ρ is a pseudometric but not a metric (the Hausdorff properly fails!).

Exercise $7.7.E.3$

Define "neighborhood," "interior, "cluster point," "closure," and "function limit" for topological spaces. Specify some notions (e.g., "diameter," "uniform continuity") that do not carry over (they involve distances).

Exercise $7.7.E.4$

In a topological space $(S, \mathcal{G})$, detine

$$\mathcal{G}^{0}=\mathcal{G}, \mathcal{G}^{1}=\mathcal{G}_{\delta}, \mathcal{G}^{2}=\mathcal{G}_{\delta \sigma}, \ldots \tag{7.7.E.3}$$

and

$$\mathcal{F}^{0}=\mathcal{F}, \mathcal{F}^{1}=\mathcal{F}_{\sigma}, \mathcal{F}^{2}=\mathcal{F}_{\sigma \delta}, \mathcal{F}^{3}=\mathcal{F}_{\sigma \delta \sigma}, \text{ etc.} \tag{7.7.E.4}$$

(Give an inductive definition.) Then prove by induction that
(a) $\mathcal{G}^{n} \subseteq \mathcal{B}, \mathcal{F}^{n} \subseteq \mathcal{B}$;
(b) $\mathcal{G}^{n-1} \subseteq \mathcal{G}^{n}, \mathcal{F}^{n-1} \subseteq \mathcal{F}^{n}$;
(c) $(\forall X \subseteq S) X \in \mathcal{F}^{n}$ iff $-X \in \mathcal{G}^{n}$;
(d) $\left(\forall X, Y \in \mathcal{F}^{n}\right) X \cap Y \in \mathcal{F}^{n}, X \cup Y \in \mathcal{F}^{n};$ same for $\mathcal{G}^{n}$;
(e) $\left(\forall X \in \mathcal{G}^{n}\right)\left(\forall Y \in \mathcal{F}^{n}\right) X-Y \in \mathcal{G}^{n}$ and $Y-X \in \mathcal{F}^{n}$.
[Hint: $X-Y=X \cap-Y$.]

Exercise $7.7.E.5$

For metric and pseudometric spaces (see Problem 2) prove that

$$\mathcal{F}^{n} \subseteq \mathcal{G}^{n+1} \text{ and } \mathcal{G}^{n} \subseteq \mathcal{F}^{n+1} \tag{7.7.E.5}$$

(cf. Problem 4).
[Hint for $\mathcal{F} \subseteq \mathcal{G}_{\delta}$: Let $F \in \mathcal{F}$. Set

$$G_{n}=\bigcup_{p \in F} G_{p}\left(\frac{1}{n}\right); \tag{7.7.E.6}$$

so

$$(\forall n) \quad F \subseteq G_{n} \in \mathcal{G}. \tag{7.7.E.7}$$

Hence

$$F \subseteq \bigcap_{n} G_{n} \in \mathcal{G}_{\delta}. \tag{7.7.E.8}$$

Also,

$$\bigcap_{n} G_{n}=\overline{F}=F \tag{7.7.E.9}$$

by Theorem 3 in Chapter 3, §16. Hence deduce that

$$(\forall F \in \mathcal{F}) \quad F \in \mathcal{G}_{\delta}, \tag{7.7.E.10}$$

so $\mathcal{F} \subseteq \mathcal{G}_{\delta}$; hence $\mathcal{G} \subseteq \mathcal{F}_{\sigma}$ by Problem 4(c). Now use induction.]

Exercise $7.7.E.6$

If m is as in Definition 5, then prove the following.
(i) m is regular.
(ii) $(\forall A \in \mathcal{M}) m A=\sup \{m X | A \supseteq X \in \mathcal{M} \cap \mathcal{F}\}$.
(iii) The latter implies strong regularity if $m<\infty$ and $S \in \mathcal{M}$.

Exercise $7.7.E.7$

Let $\mu: \mathcal{B} \rightarrow E^{*}$ be a Borel measure in a metric space (S, ρ). Set

$$(\forall A \subseteq S) \quad n^{*} A=\inf \{\mu X | A \subseteq X \in \mathcal{G}\}. \tag{7.7.E.11}$$

Prove that
(i) n^{*} is an outer measure in S;
(ii) $n^{*}=\mu$ on $\mathcal{G}$;
(iii) the n^{*}-induced measure, $n: \mathcal{N}^{*} \rightarrow E^{*},$ is topological (so $\mathcal{B} \subseteq \mathcal{N}^{*}$);
(iv) $n \geq \mu$ on $\mathcal{B}$;
(v) $(\forall A \subseteq S)\left(\exists H \in \mathcal{G}_{\delta}\right) A \subseteq H$ and $\mu H=n^{*} A$.
[Hints: (iii) Using Problem 15 in §5 and Problem 12 in §6, let

$$\rho(X,Y) > \varepsilon > 0, \quad U = \bigcup_{x\in X} G_x\left(\frac{1}{2}\varepsilon\right), \quad V = \bigcup_{y\in Y} G_y\left(\frac{1}{2}\varepsilon\right). \tag{7.7.E.12}$$

Verify that $U, V \in \mathcal{G}, U \supseteq X, V \supseteq Y, U \cap V = \emptyset$.
By the definition of n^*,

$$(\exists G \in \mathcal{G}) \quad G \supseteq X \cup Y \text{ and } n^* G \leq n^*(X \cup Y) + \varepsilon; \tag{7.7.E.13}$$

also, $X \subseteq G \cap U$ and $Y \subseteq G \cap V$. Thus by (ii),

$$n^* X \leq \mu(G \cap U) \text{ and } n^* Y \leq \mu(G \cap V). \tag{7.7.E.14}$$

Hence

$$n^* X + n^* Y \leq \mu(G \cap U) + \mu(G \cap V) = \mu((G \cap U) \cup (G \cap V)) \leq \mu G = n^* G \leq n^*(X \cup Y) + \varepsilon. \tag{7.7.E.15}$$

Let $\varepsilon \to 0$ to get the CP : $n^* X + n^* Y \leq n^*(X \cup Y)$.
(iv) We have $(\forall A \in \mathcal{B})$

$$nA = n^* A = \inf\{\mu X | A \subseteq X \in \mathcal{G}\} \geq \inf\{\mu X | A \subseteq X \in \mathcal{B}\} = \mu A. \tag{7.7.E.16}$$

(Why?)
(v) Use the hint to Problem 11 in §5.]

Exercise $7.7.E.8$

From Problem 7 with $m = \mu$, prove that if

$$A \subseteq G \in \mathcal{G}, \tag{7.7.E.17}$$

with $mG < \infty$ and $A \in \mathcal{B}$, then $mA = nA$.
[Hint: A, G, and $(G - A) \in \mathcal{B}$. By Problem 7(iii), $\mathcal{B} \subseteq N^*$ and n is additive on $\mathcal{B}$; so by Problem 7(ii)(iv),

$$nA = nG - n(G - A) \leq mG - m(G - A) = mA \leq nA. \tag{7.7.E.18}$$

Thus $mA = nA$. Explain all!]

Exercise $7.7.E.9$

Let m, n, and n^* be as in Problems 7 and 8. Suppose

$$S = \bigcup_{n=1}^{\infty} G_n, \tag{7.7.E.19}$$

with $G_n \in \mathcal{G}$ and $mG_n < \infty$ (this is called σ^0-finiteness).
Prove that
(i) $m = n$ on $\mathcal{B}$, and
(ii) m and n are strongly regular.
[Hints: Fix $A \in \mathcal{B}$. Show that

$$A = \bigcup A_n \text{ (disjoint)} \tag{7.7.E.20}$$

for some Borel sets $A_n \subseteq G_n$ (use Corollary 1 in §1). By Problem 8, $mA_n = nA_n$ since

$$A_n \subseteq G_n \in \mathcal{G} \tag{7.7.E.21}$$

and $mG_n < \infty$. Now use σ-additivity to find $mA = nA$.
(ii) Use $\mathcal{G}$-regularity, part (i), and Theorem 1.]

Exercise $7.7.E.10$

Continuing Problems 8 and 9, show that n is the Lebesgue extension of m (see Theorem 2 in §6 and Note 3 in §6).
Thus every σ^0-finite Borel measure m in (S, ρ) and its Lebesgue extension are strongly regular.
[Hint: m induces an outer measure m^*, with $m^* = m$ on $\mathcal{B}$. It suffices to show that $m^* = n^*$ on 2^S. (Why?)
So let $A \subseteq S.$ By Problem 7(v),

$$(\exists H \in \mathcal{B}) A \subseteq H \text{ and } n^* A = mH = m^* H. \tag{7.7.E.22}$$

Also,

$$(\exists K \in \mathcal{B}) A \subseteq K \text{ and } m^* A = mK \tag{7.7.E.23}$$

(Problem 12 in §5). Deduce that

$$n^* A \leq n(H \cap K) = m(H \cap K) \leq mH = n^* A \tag{7.7.E.24}$$

and
$n^* A = m(H \cap K) = m^* A$.]

7.8: Lebesgue Measure

We shall now consider the most important example of a measure in E^n, due to Lebesgue. This measure generalizes the notion of volume and assigns "volumes" to a large set family, the "Lebesgue measurable" sets, so that "volume" becomes a complete topological measure. For "bodies" in E^3, this measure agrees with our intuitive idea of "volume."

We start with the volume function $v : \mathcal{C} \rightarrow E^1$ ("Lebesgue premeasure") on the semiring $\mathcal{C}$ of all intervals in E^n (§1). As we saw in §§5 and 6, this premeasure induces an outer measure m^* on all subsets of E^n; and m^*, in turn, induces a measure m on the σ-field $\mathcal{M}^*$ of m^*-measurable sets. These sets are, by definition, the Lebesgue-measurable (briefly L-measurable) sets; m^* and m so defined are the (n-dimensional) Lebesgue outer measure and Lebesgue measure.

Theorem 7.8.1

Lebesgue premeasure v is σ-additive on $\mathcal{C}$, the intervals in E^n. Hence the latter are Lebesgue measurable $(\mathcal{C} \subseteq \mathcal{M}^*)$, and the volume of each interval equals its Lebesgue measure:

$$v = m^* = m \text{ on } \mathcal{C}. \tag{7.8.1}$$

This follows by Corollary 1 in §2 and Theorem 2 of §6

Note 1. As $\mathcal{M}^*$ is a (σ-field §6), it is closed under countable unions, countable intersections, and differences. Thus

$$\mathcal{C} \subseteq \mathcal{M}^* \text{ implies } \mathcal{C}_\sigma \subseteq \mathcal{M}^*; \tag{7.8.2}$$

i.e., any countable union of intervals is L-measurable. Also, $E^n \in \mathcal{M}^*$.

Corollary 7.8.1

Any countable set $A \subset E^n$ is L-measurable, with $mA = 0$.

Proof

The proof is as in Corollary 6 of §2.

Corollary 7.8.2

The Lebesgue measure of E^n is ∞.

Proof

Prove as in Corollary 5 of §2.

Examples

(a) Let

$$R = \left\{\text{rationals in } E^1\right\}. \tag{7.8.3}$$

Then R is countable (Corollary 3 of Chapter 1, §9); so $mR = 0$ by Corollary 1. Similarly for R^n (rational points in E^n).

(b) The measure of an interval with endpoints a, b in E^1 is its length, $b - a$.

Let

$$R_o = \{\text{ all rationals in } [a, b]\}; \tag{7.8.4}$$

so $mR_o = 0$. As $[a, b]$ and R_o are in $\mathcal{M}^*$ (a σ-field), so is

$$[a, b] - R_o, \tag{7.8.5}$$

the irrationals in $[a, b]$. By Lemma 1 in §4, if $b > a$, then

$$m([a, b]-R_o)=m([a, b])-mR_o=m([a, b])=b-a>0=mR_o. \tag{7.8.6}$$

This shows again that the irrationals form a "larger" set than the rationals (cf. Theorem 3 of Chapter 1, §9).

(c) There are uncountable sets of measure zero (see Problems 8 and 10 below).

Theorem 7.8.2

Lebesgue measure in E^n is complete, topological, and totally σ-finite. That is,

(i) all null sets (subsets of sets of measure zero) are L-measurable;

(ii) so are all open sets $(\mathcal{M}^*\supseteq\mathcal{G})$, hence all Borel sets $(\mathcal{M}^*\supseteq\mathcal{B})$; in particular, $\mathcal{M}^*\supseteq\mathcal{F}, \mathcal{M}^*\supseteq\mathcal{G}_\delta, \mathcal{M}^*\supseteq\mathcal{F}_\sigma, \mathcal{M}^*\supseteq\mathcal{F}_{\sigma\delta},$ etc.;

(iii) each $A\in\mathcal{M}^*$ is a countable union of disjoint sets of finite measure.

Proof

(i) This follows by Theorem 1 in §6.

(ii) By Lemma 2 in §2, each open set is in $\mathcal{C}_\sigma$, hence in $\mathcal{M}^*$ (Note 1). Thus $\mathcal{M}^*\supseteq\mathcal{G}.$ But by definition, the Borel field $\mathcal{B}$ is the least σ-ring $\supseteq\mathcal{G}$. Hence $\mathcal{M}^*\supseteq\mathcal{B}^*$.

(iii) As E^n is open, it is a countable union of disjoint half-open intervals,

$$E^n=\bigcup_{k=1}^{\infty} A_k \text{ (disjoint)}, \tag{7.8.7}$$

with $mA_k<\infty$ (Lemma 2 §2). Hence

$$(\forall A\subseteq E^n) \quad A\subseteq\bigcup A_k; \tag{7.8.8}$$

so

$$A=\bigcup_k(A\cap A_k) \text{ (disjoint)}. \tag{7.8.9}$$

If, further, $A\in\mathcal{M}^*,$ then $A\cap A_k\in\mathcal{M}^*,$ and

$$m(A\cap A_k)\leq mA_k<\infty. \text{ (Why?)} \quad \square \tag{7.8.10}$$

Note 2. More generally, a σ-finite set $A\in\mathcal{M}$ in a measure space $(S, \mathcal{M}, \mu)$ is a countable union of disjoint sets of finite measure (Corollary 1 of §1).

Note 3. Not all L-measurable sets are Borel sets. On the other hand, not all sets in E^n are L-measurable (see Problems 6 and 9 below.)

Theorem 7.8.3

(a) Lebesgue outer measure m^* in E^n is $\mathcal{G}$-regular; that is,

$$(\forall A\subseteq E^n) \quad m^*A=\inf\{mX|A\subseteq X\in\mathcal{G}\} \tag{7.8.11}$$

($\mathcal{G}=$ open sets in E^n).

(b) Lebesgue measure m is strongly regular (Definition 5 and Theorems 1 and 2, all in §7).

Proof

By definition, m^*A is the glb of all basing covering values of A. Thus given $\varepsilon>0,$ there is a basic covering $\{B_k\}\subseteq\mathcal{C}$ of nonempty sets B_k such that

$$A \subseteq \bigcup B_k \text{ and } m^* A + \frac{1}{2}\varepsilon \geq \sum_k vB_k. \tag{7.8.12}$$

(Why? What if $m^* A = \infty$?)

Now, by Lemma 1 in §2, fix for each B_k an open interval $C_k \supseteq B_k$ such that

$$vC_k - \frac{\varepsilon}{2^{k+1}} < vB_k. \tag{7.8.13}$$

Then (2) yields

$$m^* A + \frac{1}{2}\varepsilon \geq \sum_k \left(vC_k - \frac{\varepsilon}{2^{k+1}} \right) = \sum_k vC_k - \frac{1}{2}\varepsilon; \tag{7.8.14}$$

so by σ-subadditivity,

$$m \bigcup_k C_k \leq \sum_k mC_k = \sum_k vC_k \leq m^* A + \varepsilon. \tag{7.8.15}$$

Let

$$X = \bigcup_k C_k. \tag{7.8.16}$$

Then X is open (as the C_k are). Also, $A \subseteq X$, and by (3),

$$mX \leq m^* A + \varepsilon. \tag{7.8.17}$$

Thus, indeed, $m^* A$ is the *glb* of all mX, $A \subseteq X \in \mathcal{G}$, proving (a).

In particular, if $A \in \mathcal{M}^*$, (1) shows that m is regular (for $m^* A = mA). Also, by Theorem 2, \(m$ is σ-finite, and $E^n \in \mathcal{M}^*$; so (b) follows by Theorem 1 in §7. □

7.8.E: Problems on Lebesgue Measure

Exercise $7.8.E.1$

Fill in all details in the proof of Theorems 3 and 4.

Exercise $7.8.E.1'$

Prove Note 2.

Exercise $7.8.E.2$

From Theorem 3 deduce that

$$(\forall A \subseteq E^{n})(\exists B \in \mathcal{G}_{\delta}) \quad A \subseteq B \text{ and } m^{*} A = m B. \tag{7.8.E.1}$$

[Hint: See the hint to Problem 7 in §5.]

Exercise $7.8.E.3$

Review Problem 3 in §5.

Exercise $7.8.E.4$

Consider all translates

$$R+p \quad \left(p \in E^{1}\right) \tag{7.8.E.2}$$

of

$$R=\left\{\text{rationals in } E^{1}\right\}. \tag{7.8.E.3}$$

Prove the following.
(i) Any two such translates are either disjoint or identical.
(ii) Each $R+p$ contains at least one element of $[0,1]$.
[Hint for (ii): Fix a rational $y \in(-p, 1-p)$, so $0<y+p<1$. Then $y+p \in R+p$, and $y+p \in[0,1]$.]

Exercise $7.8.E.5$

Continuing Problem 4, choose one element $q \in[0,1]$ from each $R+p$. Let Q be the set of all q so chosen.
Call a translate of $Q, Q+r$, "good" iff $r \in R$ and $|r|<1$. Let U be the union of all "good" translates of Q.
Prove the following.
(a) There are only countably many "good" $Q+r$.
(b) All of them lie in $[-1,2]$.
(c) Any two of them are either disjoint or identical.
(d) $[0,1] \subseteq U \subseteq[-1,2]$; hence $1 \leq m^{*} U \leq 3$.
[Hint for (c): Suppose

$$y \in(Q+r) \cap\left(Q+r'\right). \tag{7.8.E.4}$$

Then

$$y=q+r=q'+r' \quad (q,q'\in Q, r,r'\in R); \tag{7.8.E.5}$$

so $q=q'+(r'-r)$, with $(r'-r)\in R$.
Thus $q\in R+q'$ and $q'=0+q'\in R+q'$. Deduce that $q=q'$ and $r=r'=$; hence $Q+r=Q+r'$.]

Exercise 7.8.E.6

Show that Q in Problem 5 is not L-measurable.
[Hint: Otherwise, by Theorem 4, each $Q+r$ is L-measurable, with $m(Q+r)=mQ$. By 5(a)(c), U is a countable disjoint union of "good" translates.
Deduce that $mU=0$ if $mQ=0$, or $mU=\infty$, contrary to 5(d).]

Exercise 7.8.E.7

Show that if $f:S\rightarrow T$ is continuous, then $f^{-1}[X]$ is a Borel set in S whenever $X\in\mathcal{B}$ in T.
[Hint: Using Note 1 in §7, show that

$$\mathcal{R}=\left\{X\subseteq T | f^{-1}[X]\in\mathcal{B}\text{ in } S\right\} \tag{7.8.E.6}$$

is a σ-ring in T. As $\mathcal{B}$ is the least σ-ring $\supseteq\mathcal{G}, \mathcal{R}\supseteq\mathcal{B}$ (the Borel field in T.]

Exercise 7.8.E.8

Prove that every degenerate interval in E^n has Lebesgue measure 0, even if it is uncountable. Give an example in E^2. Prove uncountability.
[Hint: Take $\bar{a}=(0,0), \bar{b}=(0,1)$. Define $f:E^1\rightarrow E^2$ by $f(x)=(0,x)$. Show that f is one-to-one and that $[\bar{a},\bar{b}]$ is the f-image of $[0,1]$. Use Problem 2 of Chapter 1, §9.]

Exercise 7.8.E.9

Show that not all L-measurable sets are Borel sets in E^n.
[Hint for E^2: With $[\bar{a},\bar{b}]$ and f as in Problem 8, show that f is continuous (use the sequential criterion). As $m[\bar{a},\bar{b}]=0$, all subsets of $[\bar{a},\bar{b}]$ are in $\mathcal{M}^*$ (Theorem 2(i)), hence in $\mathcal{B}$ if we assume $\mathcal{M}^*=\mathcal{B}$. But then by Problem 7 , the same would apply to subsets of $[0,1]$, contrary to Problem 6.
Give a similar proof for $E^n (n>1)$.
Note: In E^1, too, $\mathcal{B}\neq\mathcal{M}^*$, but a different proof is necessary. We omit it.]

Exercise 7.8.E.10

Show that Cantor's set P (Problem 17 in Chapter 3, 14) has Lebesgue measure zero, even though it is uncountable.
[Outline: Let

$$U=[0,1]-P; \tag{7.8.E.7}$$

so U is the union of open intervals removed from $[0,1]$. Show that

$$mU=\frac{1}{2}\sum_{n=1}^{\infty}\left(\frac{2}{3}\right)^n=1 \tag{7.8.E.8}$$

and use Lemma 1 in §4.]

Exercise $7.8.E.11$

Let $\mu: \mathcal{B} \rightarrow E^{*}$ be the Borel restriction of Lebesgue measure m in E^{n} (§7). Prove that
(i) μ in incomplete;
(ii) m is the Lebesgue extension (* and completion, as in Problem 15 of §6) of μ.
[Hints: (i) By Problem 9, some μ-null sets are not in $\mathcal{B}$. (ii) See the proof (end) of Theorem 2 in §9 (the next section).]

Exercise $7.8.E.12$

Prove the following.
(i) All intervals in E^{n} are Borel sets.
(ii) The σ-ring generated by any one of the families $\mathcal{C}$ or $\mathcal{C}'$ in Problem 3 of §5 coincides with the Borel field in E^{n}.
[Hints: (i) Any interval arises from a closed one by dropping some "faces" (degenerate closed intervals). (ii) Use Lemma 2 from §2 and Problem 7 of §3.]

Exercise $7.8.E.13*$

Show that if a measure $m^{\prime}: \mathcal{M}^{\prime} \rightarrow E^{*}$ in E^{n} agrees on intervals with Lebesgue measure $m: \mathcal{M}^{*} \rightarrow E^{*},$ then the following are true.
(i) $m^{\prime}=m$ on $\mathcal{B}$, the Borel field in E^{n}.
(ii) If $m^{\prime}$ is also complete, then $m^{\prime}=m$ on $\mathcal{M}^{*}$.
[Hint: (i) Use Problem 13 of §5 and Problem 12 above.]

Exercise $7.8.E.14$

Show that globes of equal radius have the same Lebesgue measure.
[Hint: Use Theorem 4.]

Exercise $7.8.E.15$

Let $f: E^{n} \rightarrow E^{n},$ with

$$f(\overline{x})=c \overline{x} \quad(0<c<\infty). \tag{7.8.E.9}$$

Prove the following.
(i) $\left(\forall A \subseteq E^{n}\right) m^{*} f[A]=c^{n} m^{*} A$ ($m^{*}=$Lebesgue outer measure).
(ii) $A \in \mathcal{M}^{*}$ iff $f[A] \in \mathcal{M}^{*}$.
[Hint: If, say, $A=(\overline{a}, \overline{b}],$ then $f[A]=(c \overline{a}, c \overline{b}]$. (Why?) Proceed as in Theorem 4, using f^{-1} also.]

Exercise $7.8.E.16$

From Problems 14 and 15 show that
(i) $m G_{\overline{p}}(c r)=c^{n} \cdot m G_{\overline{p}}(r)$;
(ii) $m G_{\overline{p}}(r)=m \overline{G}_{\overline{p}}(r)$;
(iii) $m G_{\overline{p}}(r)=a \cdot m I,$ where I is the cube inscribed in $G_{\overline{p}}(r)$ and

$$a=\left(\frac{1}{2} \sqrt{n}\right)^{n} \cdot m G_{\overline{0}}(1). \tag{7.8.E.10}$$

[Hints: (i) $f\left[G_{\overline{0}}(r)\right]=G_{\overline{0}}(c r)$. (ii) Prove that

$$m G_{\overline{p}} \leq m \overline{G}_{\overline{p}} \leq c^{n} m G_{\overline{p}} \tag{7.8.E.11}$$

if $c>1$. Let $c \rightarrow 1$.]

Exercise 7.8.E.17

Given $a<b$ in E^{1}, let $\left\{r_{n}\right\}$ be the sequence of all rationals in $A=[a, b]$.
Set $(\forall n)$

$$\delta_{n}=\frac{b-a}{2^{n+1}} \tag{7.8.E.12}$$

and

$$G_{n}=\left(a_{n}, b_{n}\right)=(a, b) \cap\left(r_{n}-\frac{1}{2} \delta_{n}, r_{n}+\frac{1}{2} \delta_{n}\right). \tag{7.8.E.13}$$

Let

$$P=A-\bigcup_{n=1}^{\infty} G_{n}. \tag{7.8.E.14}$$

Prove the following.
(i) $\sum_{n=1}^{\infty} \delta_{n}=\frac{1}{2}(b-a)=\frac{1}{2} m A$.
(ii) P is closed; $P^{o}=\emptyset$, yet $m P>0$.
(iii) The G_{n} can be made disjoint (see Problem 3 in §2), with $m P$ still >0.
(iv) Construct such a $P \subseteq A\left(P=\overline{P}, P^{o}=\emptyset\right)$ of prescribed measure $m P=\varepsilon>0$.

Exercise 7.8.E.18

Find an open set $G \subset E^{1}$, with $m G<m \overline{G}<\infty$.
[Hint: $G=\cup_{n=1}^{\infty} G_{n}$ with G_{n} as in Problem 17.]

Exercise 7.8.E.19*

If $A \subseteq E^{n}$ is open and convex, then $m A=m \overline{A}$.
[Hint: Let first $\overline{0} \in A$. Argue as in Problem 16.]

7.9: Lebesgue–Stieltjes Measures

7.9.E: Problems on Lebesgue-Stieltjes Measures

Exercise $7.9.E.1$

Do Problems 7 and 8 in §4 and Problem 3' in §5, if not done before.

Exercise $7.9.E.2$

Prove in detail Theorems 1 to 3 in §8 for LS measures and outer measures.

Exercise $7.9.E.3$

Do Problem 2 in §8 for LS-outer measures in E^1.

Exercise $7.9.E.4$

Prove that $f : E^1 \rightarrow (S, \rho)$ is right (left) continuous at p iff

$$\lim_{n \rightarrow \infty} f(x_n) = f(p) \text{ as } x_n \searrow p \,(x_n \nearrow p)\,. \tag{7.9.E.1}$$

[Hint: Modify the proof of Theorem 1 in Chapter 4, §2.]

Exercise $7.9.E.5$

Fill in all proof details in Theorem 2.
[Hint: Use Problem 4.]

Exercise $7.9.E.6$

In Problem 8(iv) of §4, describe m_α^* and M_α^*.

Exercise $7.9.E.7$

Show that if $\alpha = c$ constant on an open interval $I \subseteq E^1$ then

$$(\forall A \subseteq I) \quad m_\alpha^*(A) = 0. \tag{7.9.E.2}$$

Disprove it for nonopen intervals I (give a counterexample).

Exercise $7.9.E.8$

Let $m' : \mathcal{M} \rightarrow E^*$ be a topological, translation-invariant measure in E^1, with $m'(0,1] = c < \infty$. Prove the following.
(i) $m' = cm$ on the Borel field $\mathcal{B}$. (Here $m : \mathcal{M}^* \rightarrow E^*$ is Lebesgue measure in E^1.)
*(ii) If m' is also complete, then $m' = cm$ on $\mathcal{M}^*$.
(iii) If $0 < c < \infty$, some set $Q \subset [0,1]$ is not m'-measurable.
*(iv) If $\mathcal{M}' = \mathcal{B}$, then cm is the completion of m' (Problem 15 in §6).
[Outline: (i) By additivity and translation invariance,

$$m'(0,r] = cm(0,r] \tag{7.9.E.3}$$

for rational

$$r=\frac{n}{k}, \quad n, k \in N \tag{7.9.E.4}$$

(first take $r=n$, then $r=\frac{1}{k}$, then $r=\frac{n}{k}$).
By right continuity (Theorem 2 in §4), prove it for real $r>0$ (take rationals $r_i \searrow r$).
By translation, $m'=cm$ on half-open intervals. Proceed as in Problem 13 of §8.
(iii) See Problems 4 to 6 in §8. Note that, by Theorem 2, one may assume $m'=m_\alpha$ (a translation-invariant LS measure). As $m_\alpha=cm$ on half-open intervals, Lemma 2 in §2 yields $m_\alpha=cm$ on $\mathcal{G}$ (open sets). Use $\mathcal{G}$-regularity to prove $m_\alpha^*=cm^*$ and $\mathcal{M}_\alpha^*=\mathcal{M}^*$.]

Exercise $7.9.E.9*$

(LS measures in E^n.) Let

$$\mathcal{C}^*=\{\text{alf-open intervals in } E^n\} . \tag{7.9.E.5}$$

For any map $G: E^n \rightarrow E^1$ and any $(\bar{a}, \bar{b}] \in \mathcal{C}^*$, set

$$\begin{aligned} \Delta_k G(\bar{a}, \bar{b}] &= G\left(x_1, \ldots, x_{k-1}, b_k, x_{k+1}, \ldots, x_n\right) \\ &\quad -G\left(x_1, \ldots, x_{k-1}, a_k, x_{k+1}, \ldots, x_n\right), \quad 1 \leq k \leq n . \end{aligned}$$

Given $\alpha: E^n \rightarrow E^1$, set

$$s_\alpha(\bar{a}, \bar{b}]=\Delta_1\left(\Delta_2\left(\cdots\left(\Delta_n \alpha(\bar{a}, \bar{b}]\right) \cdots\right)\right) . \tag{7.9.E.6}$$

For example, in E^2,

$$s_\alpha(a, b]=\alpha\left(b_1, b_2\right)-\alpha\left(b_1, a_2\right)-\left[\alpha\left(a_1, b_2\right)-\alpha\left(a_1, a_2\right)\right] . \tag{7.9.E.7}$$

Show that s_α is additive on $\mathcal{C}^*$. Check that the order in which the Δ_k are applied is immaterial. Set up a formula for s_α in E^3.
[Hint: First take two disjoint intervals

$$(\bar{a}, \bar{q}] \cup(\bar{p}, \bar{b}]=(\bar{a}, \bar{b}], \tag{7.9.E.8}$$

as in Figure 2 in Chapter 3, §7. Then use induction, as in Problem 9 of Chapter 3, §7.]

Exercise $7.9.E.10*$

If s_α in Problem 9 is nonnegative, and α is right continuous in each variable x_k separately, we call α a distribution function, and s_α is called the α-induced LS premeasure in E^n; the LS outer measure m_α^* and measure

$$m_\alpha: \mathcal{M}_\alpha^* \rightarrow E^* \tag{7.9.E.9}$$

in E^n (obtained from s_α as shown in } §§5 and 6) are said to be induced by α.
For s_α, m_α^*, and m_α so defined, redo Problems 1-3 above.

7.10: Generalized Measures. Absolute Continuity

7.10.E: Problems on Generalized Measures

Exercise $7.10.E.1$

Complete the proofs of Theorems 1,4, and 5.

Exercise $7.10.E.1'$

Do it also for the lemmas and Corollary 3.

Exercise $7.10.E.2$

Verify the following.
(i) In Definition 2, one can equivalently replace "countable $\{X_i\}$" by "finite $\{X_i\}$."
(ii) If $\mathcal{M}$ is a ring, Note 1 holds for finite sequences $\{X_i\}$.
(iii) If $s : \mathcal{M} \rightarrow E$ is additive on $\mathcal{M}$, a semiring, so is v_s.
[Hint: Use Theorem 1 from §4.]

Exercise $7.10.E.3$

For any set functions s, t on $\mathcal{M}$, prove that
(i) $\quad v_{|s|} = v_s,$ and
(ii) $v_{st} \leq a v_t,$ provided st is defined and

$$a = \sup\{|sX| | X \in \mathcal{M}\}. \tag{7.10.E.1}$$

Exercise $7.10.E.4$

Given $s, t : \mathcal{M} \rightarrow E,$ show that
(i) $v_{s+t} \leq v_s + v_t$;
(ii) $v_{ks} = |k| v_s$ ((\k\) as in Corollary 2); and
(iii) if $E = E^n\left(C^n\right)$ and

$$s = \sum_{k=1}^{n} s_k \bar{e}_k, \tag{7.10.E.2}$$

then

$$v_{s_k} \leq v_s \leq \sum_{k=1}^{n} v_{sk}. \tag{7.10.E.3}$$

[Hints: (i) If

$$A \supseteq \bigcup X_i \text{ (disjoint)}, \tag{7.10.E.4}$$

with $A_i, X_i \in \mathcal{M},$ verify that

$$\begin{aligned} |(s+t) X_i| &\leq |s X_i| + |t X_i|, \\ \sum |(s+t) X_i| &\leq v_s A + v_t A, \text{ etc.}; \end{aligned} \tag{7.10.E.5}$$

(ii) is analogous.

(iii) Use (ii) and (i), with $|\bar{e}_k| = 1$.]

? Exercise 7.10.E.5

If $g\uparrow, h\uparrow$, and $\alpha = g - h$ on E^1, can one define the signed LS measure s_α by simply setting $s_\alpha = m_g - m_h$ (assuming $m_h < \infty$)?
[Hint: the domains of m_g and m_h may be different. Give an example. How about taking their intersection?]

? Exercise 7.10.E.6

Find an LS measure m_α such that α is continuous and one-to-one, but m_α is not m-finite ($m =$Lebesgue measure).
[Hint: Take

$$\alpha(x) = \begin{cases} \frac{x^3}{|x|}, & x \neq 0, \\ 0, & x = 0, \end{cases} \tag{7.10.E.6}$$

and

$$A = \bigcup_{n=1}^{\infty}\left(n, n + \frac{1}{n^2}\right].\Bigg] \tag{7.10.E.7}$$

? Exercise 7.10.E.7

Construct complex and vector-valued LS measures $s_\alpha : \mathcal{M}_\alpha^* \to E^n\left(C^n\right)$ in E^1.

? Exercise 7.10.E.8

Show that if $s : \mathcal{M} \to E^n\left(C^n\right)$ is additive and bounded on $\mathcal{M}$, a ring, so is v_s.
[Hint: By Problem 4(iii), reduce all to the real case.
Use Problem 2. Given a finite disjoint sequence $\{X_i\} \subseteq \mathcal{M}$, let $U^+\left(U^-\right)$ be the union of those X_i for which $sX_i \geq 0 (sX_i < 0,$ respectively). Show that

$$\sum sX_i = sU^+ - sU^- \leq 2 \sup|s| < \infty.\Big] \tag{7.10.E.8}$$

? Exercise 7.10.E.9

For any $s : \mathcal{M} \to E^*$ and $A \in \mathcal{M}$, set

$$s^+ A = \sup\{sX | A \supseteq X \in \mathcal{M}\} \tag{7.10.E.9}$$

and

$$s^- A = \sup\{-sX | A \supseteq X \in \mathcal{M}\}. \tag{7.10.E.10}$$

Prove that if s is additive and bounded on $\mathcal{M}$, a ring, so are s^+ and s^-; furthermore,

$$\begin{aligned} s^+ &= \frac{1}{2}(v_s + s) \geq 0, \\ s^- &= \frac{1}{2}(v_s - s) \geq 0, \\ s &= s^+ - s^-, \text{ and} \\ v_s &= s^+ + s^-. \end{aligned}$$

[Hints: Use Problem 8. Set

$$s'=\frac{1}{2}(v_s+s)\,. \tag{7.10.E.11}$$

Then $(\forall X\in\mathcal{M}\,|\,X\subseteq A)$

$$\begin{aligned}2sX=sA+sX-s(A-X)&\leq sA+(|sX|+|s(A-X)|)\\&\leq sA+v_sA=2s'A.\end{aligned}$$

Deduce that $s^+A\leq s'A$.
To prove also that $s'A\leq s^+A,$ let $\varepsilon>0.$ By Problems 2 and 8, fix $\{X_i\}\subseteq\mathcal{M}$, with

$$A=\bigcup_{i=1}^{n}X_i\ \text{(disjoint)} \tag{7.10.E.12}$$

and

$$v_sA-\varepsilon<\sum_{i=1}^{n}|sX_i|\,. \tag{7.10.E.13}$$

Show that

$$2s'A-\varepsilon=v_sA+sA-\varepsilon\leq sU^+-sU^-+s\bigcup_{i=1}^{n}X_i=2sU^+ \tag{7.10.E.14}$$

and

$$2s^+A\geq 2sU^+\geq 2s'A-\varepsilon.\,] \tag{7.10.E.15}$$

Exercise $7.10.E.10$

Let

$$\mathcal{K}=\{\text{compact sets in a topological space }(S,\mathcal{G})\} \tag{7.10.E.16}$$

(adopt Theorem 2 in Chapter 4, §7, as a definition). Given

$$s:\mathcal{M}\rightarrow E,\quad \mathcal{M}\subseteq 2^S, \tag{7.10.E.17}$$

we call s compact regular (CR) iff

$$\begin{aligned}&(\forall\varepsilon>0)\,(\forall A\in\mathcal{M})(\exists F\in\mathcal{K})(\exists G\in\mathcal{G})\\&\quad F,G\ \in\mathcal{M},F\subseteq A\subseteq G,\text{ and }v_sG-\varepsilon\leq v_sA\leq v_sF+\varepsilon.\end{aligned}$$

Prove the following.
(i) If $s,t:\mathcal{M}\rightarrow E$ are CR, so are $s\pm t$ and ks (k as in Corollary 2).
(ii) If s is additive and CR on $\mathcal{M}$, a semiring, so is its extension to the ring $\mathcal{M}_s$ (Theorem 1 in §4 and Theorem 4 of §3).
(iii) If $E=E^n\ (C^n)$ and $v_s<\infty$ on $\mathcal{M}$, a ring, then s is CR iff its components s_k are, or in the case $E=E^1$, iff s^+ and s^- are (see Problem 9).
[Hint for (iii): Use (i) and Problem 4(iii). Consider $v_s(G-F)$.]

Exercise $7.10.E.11$

(Aleksandrov.) Show that if $s : \mathcal{M} \rightarrow E$ is CR (see Problem 10) and additive on $\mathcal{M}$, a ring in a topological space S, and if $v_s < \infty$ on $\mathcal{M}$, then v_s and s are σ-additive, and v_s has a unique σ-additive extension $\overline{v}_s$ to the σ-ring $\mathcal{N}$ generated by $\mathcal{M}$. The latter holds for s, too, if $S \in \mathcal{M}$ and $E = E^n\left(C^n\right)$.
[Proof outline: The σ-additivity of v_s results as in Theorem 1 of §2 (first check Lemma 1 in §1 for v_s).
For the σ-additivity of s, let

$$A=\bigcup_{i=1}^{\infty} A_i \text{ (disjoint)}, \quad A, A_i \in \mathcal{M}; \tag{7.10.E.18}$$

then

$$\left|sA-\sum_{i=1}^{r-1} sA_i\right| \leq \sum_{i=r}^{\infty} v_s A_i \rightarrow 0 \tag{7.10.E.19}$$

as $r \rightarrow \infty$, for

$$\sum_{i=1}^{\infty} v_s A_i = v_s \bigcup_{i=1}^{\infty} A_i < \infty. \tag{7.10.E.20}$$

(Explain!) Now, Theorem 2 of §6 extends v_s to a measure on a σ-field

$$\mathcal{M}^* \supseteq \mathcal{N} \supseteq \mathcal{M} \tag{7.10.E.21}$$

(use the minimality of $\mathcal{N}$). Its restriction to $\mathcal{N}$ is the desired $\overline{v}_s$ (unique by Problem 15 in §6).
A similar proof holds for s, too, if $s : \mathcal{M} \rightarrow [0, \infty)$. The case $s : \mathcal{M} \rightarrow E^n\left(C^n\right)$ results via Theorem 5 and Problem 10(iii) provided $S \in \mathcal{M}$; for then by Corollary 1, $v_s S < \infty$ ensures the finiteness of v_s, s^+, and s^- even on $\mathcal{N}$.]

Exercise $7.10.E.12$

Do Problem 11 for semirings $\mathcal{M}$.
[Hint: Use Problem 10(ii).]

7.11: Differentiation of Set Functions

7.11.E: Problems on Vitali Coverings

Exercise $7.11.E.1$

Prove Theorem 1 for globes, filling in all details.
[Hint: Use Problem 16 in §8.]

Exercise $7.11.E.2$

⇒ Show that any (even uncountable) union of globes or nondegenerate cubes $J_i \subset E^n$ is L-measurable.
[Hint: Include in $\mathcal{K}$ each globe (cube) that lies in some J_i. Then Theorem 1 represents $\cup J_I$ as a countable union plus a null set.]

Exercise $7.11.E.3$

Supplement Theorem 1 by proving that

$$m^*\left(A-\bigcup I_k^o\right)=0 \tag{7.11.E.1}$$

and

$$m^* A=m^*\left(A \cap \bigcup I_k^o\right); \tag{7.11.E.2}$$

here $I^o=$ interior of I.

Exercise $7.11.E.4$

Fill in all proof details in Lemmas 1 and 2. Do it also for $\overline{\mathcal{K}}=\{\text{globes}\}$.

Exercise $7.11.E.5$

Given $m Z=0$ and $\varepsilon>0,$ prove that there are open globes

$$G_k^* \subseteq E^n, \tag{7.11.E.3}$$

with

$$Z \subset \bigcup_{k=1}^{\infty} G_k^* \tag{7.11.E.4}$$

and

$$\sum_{k=1}^{\infty} m G_k^*<\varepsilon. \tag{7.11.E.5}$$

[Hint: Use Problem 3(f) in §5 and Problem 16(iii) from §8.]

Exercise $7.11.E.6$

Do Problem 3 in §5 for
(i) $\mathcal{C}'=\{\text{open globes}\}$, and
(ii) $\mathcal{C}'=\{\text{all globes in } E^n\}$.
[Hints for (i): Let $m'=$ outer measure induced by $v':\mathcal{C}'\rightarrow E^1$. From Problem 3(e) in §5, show that

$$(\forall A\subseteq E^n)\quad m'A\geq m^*A. \tag{7.11.E.6}$$

To prove $m'A\leq m^*A$ also, fix $\varepsilon>0$ and an open set $G\supseteq A$ with

$$m^*A+\varepsilon\geq mG \text{ (Theorem 3 of §8)}. \tag{7.11.E.7}$$

Globes inside G cover A in the V-sense (why?); so

$$A\subseteq Z\cup\bigcup G_k \text{ (disjoint)} \tag{7.11.E.8}$$

for some globes G_k and null set Z. With G_k^* as in Problem 5,

$$m'A\leq\sum\left(mG_k+mG_k^*\right)\leq mG+\varepsilon\leq m^*A+2\varepsilon.] \tag{7.11.E.9}$$

Exercise $7.11.E.7$

Suppose $f:E^n\overset{\text{onto}}{\longleftrightarrow}E^n$ is an isometry, i.e., satisfies

$$|f(\bar{x})-f(\bar{y})|=|\bar{x}-\bar{y}|\quad \text{for } \bar{x},\bar{y}\in E^n. \tag{7.11.E.10}$$

Prove that
(i) $(\forall A\subseteq E^n)\, m^*A=m^*f[A]$, and
(ii) $A\in\mathcal{M}^*$ iff $f[A]\in\mathcal{M}^*$.
[Hints: If A is a globe of radius r, so is $f[A]$ (verify!); thus Problems 14 and 16 in §8 apply. In the general case, argue as in Theorem 4 of §8 , replacing intervals by globes (see Problem 6). Note that f^{-1} is an isometry, too.]

Exercise $7.11.E.7'$

From Problem 7 infer that Lebesgue measure in E^n is rotation invariant. (A rotation about $\bar{p}$ is an isometry f such that $f(\bar{p})=\bar{p}$.)

Exercise $7.11.E.8$

A V-covering $\mathcal{K}$ of $A\subseteq E^n$ is called normal iff
(i) $(\forall I\in K) 0<m\bar{I}=mI^o$, and
(ii) for every $\bar{p}\in A$, there is some $c\in(0,\infty)$ and a sequence

$$I_k\rightarrow\bar{p}\quad(\{I_k\}\subseteq\mathcal{K}) \tag{7.11.E.11}$$

such that

$$(\forall k)\,(\exists \text{ cube } J_k\supseteq I_k)\quad c\cdot m^*I_k\geq mJ_k. \tag{7.11.E.12}$$

(We then say that $\bar{p}$ and $\{I_k\}$ are normal; specifically, c-normal.)
Prove Theorems 1 and 2 for any normal $\mathcal{K}$.

[Hints: By Problem 21 of Chapter 3, §16, $dI = d\overline{I}$.
First, suppose $\mathcal{K}$ is uniformly normal, i.e., all $\bar{p} \in A$ are c-normal for the same c.
In the general case, let

$$A_i = \{\bar{x} \in A \mid \bar{x} \text{ is } i\text{-normal}\}, \quad i = 1, 2, \ldots; \tag{7.11.E.13}$$

so $\mathcal{K}$ is uniform for A_i. Verify that $A_i \nearrow A$.
Then select, step by step, as in Theorem 1, a disjoint sequence $\{I_k\} \subseteq \mathcal{K}$ and naturals $n_1 < n_2 < \cdots < n_i < \cdots$ such that

$$(\forall i) \quad m^*\left(A_i - \bigcup_{k=1}^{n_i} I_k\right) < \frac{1}{i}. \tag{7.11.E.14}$$

Let

$$U = \bigcup_{k=1}^{\infty} I_k. \tag{7.11.E.15}$$

Then

$$(\forall i) \quad m^*\left(A_i - U\right) < \frac{1}{i} \tag{7.11.E.16}$$

and

$$A_i - U \nearrow A - U. \tag{7.11.E.17}$$

(Why?) Thus by Problems 7 and 8 in §6,

$$m^*(A - U) \leq \lim_{i \rightarrow \infty} \frac{1}{i} = 0.] \tag{7.11.E.18}$$

Exercise $7.11.E.9$

A V-covering $\overline{K}^*$ of E^n is called universal iff
(i) (\forall I \in \overline{\mathcal{K}}^{*}\right) 0<m \overline{I}=m I^{o}<\infty, and
(ii) whenever a subfamily $\mathcal{K} \subseteq \overline{\mathcal{K}}^*$ covers a set $A \subseteq E^n$ in the V-sense, we have

$$m^*\left(A - \bigcup I_k\right) = 0 \tag{7.11.E.19}$$

for a disjoint sequence

$$\{I_k\} \subseteq \mathcal{K}. \tag{7.11.E.20}$$

Show the following.
(a) $\overline{\mathcal{K}}^* \subseteq \mathcal{M}^*$.
(b) Lemmas 1 and 2 are true with $\overline{\mathcal{K}}$ replaced by any universal $\overline{\mathcal{K}}^*$. (In this case, write $\underline{D}^* s$ and $\overline{D}^* s$ for the analogues of $\underline{D} s$ and $\overline{D}_s$.)
(c) $\underline{D} s = \underline{D}^* s = \overline{D}^* s = \overline{D} s \quad$ a.e.
[Hints: (a) By (i), $I = \overline{I}$ minus a null set $Z \subseteq \overline{I} - I^o$.
(c) Argue as in Lemma 2, but set

$$Q=J\left(\underline{D}^{*} s>u>v>\underline{D} s\right) \tag{7.11.E.21}$$

and

$$\mathcal{K}^{\prime}=\left\{I \in \overline{\mathcal{K}}^{*} \mid I \subseteq G^{\prime}, \frac{s I}{m I}>v\right\} \tag{7.11.E.22}$$

to prove a.e. that $\underline{D}^{*} s \leq \underline{D} s$; similarly for $\underline{D} s \leq D^{*} s$.
Throughout assume that $s: \mathcal{M}^{\prime} \rightarrow E^{*}\left(\mathcal{M}^{\prime} \supseteq \overline{\mathcal{K}} \cup \overline{\mathcal{K}}^{*}\right)$ is a measure in E^{n}, finite on $\overline{\mathcal{K}} \cup \overline{\mathcal{K}}^{*}$.]

Exercise 7.11.E.10

Continuing Problems 8 and 9, verify that
(a) \(\overline{\mathcal{K}}=\{\text {nondegenerate cubes \}\}\) is a normal and universal V-covering of E^{n};
(b) so also is $\overline{\mathcal{K}}^{o}=\{\text{all globes in} E^{n}\}$;
(c) $\overline{\mathcal{C}}=\{\text{nondegenerate intervals}\}$ is normal.
Note that $\overline{\mathcal{C}}$ is not universal.

Exercise 7.11.E.11

Continuing Definition 3, we call q a derivate of s, and write $q \sim D s(\overline{p})$, iff

$$q=\lim _{k \rightarrow \infty} \frac{s I_{k}}{m I_{k}} \tag{7.11.E.23}$$

for some sequence $I_{k} \rightarrow \overline{p}$, with $I_{k} \in \overline{\mathcal{K}}$.
Set

$$D_{\overline{p}}=\left\{q \in E^{*} \mid q \sim D s(\overline{p})\right\} \tag{7.11.E.24}$$

and prove that

$$\underline{D} s(\overline{p})=\min D_{\overline{p}} \text { and } \overline{D} s(\overline{p})=\max D_{\overline{p}}. \tag{7.11.E.25}$$

Exercise 7.11.E.12

Let $\mathcal{K}^{*}$ be a normal V-covering of E^{n} (see Problem 8). Given a measure s in E^{n}, finite on $\mathcal{K}^{*} \cup \overline{\mathcal{K}}$, write

$$q \sim D^{*} s(\overline{p}) \tag{7.11.E.26}$$

iff

$$q=\lim _{k \rightarrow \infty} \frac{s I_{k}}{m I_{k}} \tag{7.11.E.27}$$

for some normal sequence $I_{k} \rightarrow \overline{p}$, with $I_{k} \in \mathcal{K}^{*}$.
Set

$$D_{\overline{p}}^{*}=\left\{q \in E^{*} \mid q \sim D^{*} s(\overline{p})\right\}, \tag{7.11.E.28}$$

and then

$$\underline{D}^{*} s(\bar{p})=\inf D_{\bar{p}}^{*} \text{ and } \overline{D}^{*} s(\bar{p})=\sup D_{\bar{p}}^{*}. \tag{7.11.E.29}$$

Prove that

$$\underline{D} s=\underline{D}^{*} s=\overline{D}^{*} s=\overline{D} s \text{ a.e. on } E^{n}. \tag{7.11.E.30}$$

[Hint: $E^{n}=\bigcup_{i=1}^{\infty} E_{i},$ where

$$E_{i}=\left\{\bar{x} \in E^{n} \mid \bar{x} \text{ is } i\text{-normal}\right\}. \tag{7.11.E.31}$$

On each $E_{i}, \mathcal{K}^{*}$ is uniformly normal. To prove $\underline{D} s=\underline{D}^{*} s$ a.e. on E_{i}, "imitate" Problem 9(c). Proceed.]

CHAPTER OVERVIEW

8: Measurable Functions and Integration

8.1: Elementary and Measurable Functions

From set functions, we now return to point functions

$$f : S \rightarrow (T, \rho') \tag{8.1.1}$$

whose domain D_f consists of points of a set S. The range space T will mostly be E, i.e., E^1, E^*, C, E^n, or another normed space. We assume $f(x) = 0$ unless defined otherwise. (In a general metric space T, we may take some fixed element q for 0.) Thus D_f is all of S, always.

We also adopt a convenient notation for sets:

$$\text{" } A(P) \text{" for " } \{x \in A | P(x)\}. \text{"} \tag{8.1.2}$$

Thus

$$\begin{aligned} A(f \neq a) &= \{x \in A | f(x) \neq a\}, \\ A(f = g) &= \{x \in A | f(x) = g(x)\}, \\ A(f > g) &= \{x \in A | f(x) > g(x)\}, \text{ etc.} \end{aligned}$$

Definition

A measurable space is a set $S \neq \emptyset$ together with a set ring $\mathcal{M}$ of subsets of S, denoted $(S, \mathcal{M})$.

Henceforth, $(S, \mathcal{M})$ is fixed.

Definition

An M-partition of a set A is a countable set family $\mathcal{P} = \{A_i\}$ such that

$$A = \bigcup_i A_i (disjoint), \tag{8.1.3}$$

with $A, A_i \in \mathcal{M}$.

We briefly say "the partition $A = \bigcup A_i$."

An $\mathcal{M}$-partition $\mathcal{P}' = \{B_{ik}\}$ is a refinement of $\mathcal{P} = \{A_i\}$ (or $\mathcal{P}'$ refines $\mathcal{P}$, or $\mathcal{P}'$ is finer than $\mathcal{P}$) iff

$$(\forall i) \quad A_i = \bigcup_k B_{ik} \tag{8.1.4}$$

i.e., each B_{ik} is contained in some A_i.

The intersection $\mathcal{P}' \cap \mathcal{P}''$ of $\mathcal{P}' = \{A_i\}$ and $\mathcal{P}'' = \{B_k\}$ is understood to be the family of all sets of the form

$$A_i \cap B_k, \quad i, k = 1, 2, \ldots \tag{8.1.5}$$

It is an $\mathcal{M}$ -partition that refines both $\mathcal{P}'$ and $\mathcal{P}''$.

Definition

A map (function) $f : S \rightarrow T$ is elementary, or $\mathcal{M}$-elementary, on a set $A \in \mathcal{M}$ iff there is an M-partition $\mathcal{P} = \{A_i\}$ of A such that f is constant $(f = a_i)$ on each A_i.

If $\mathcal{P} = \{A_1, \ldots, A_q\}$ is finite, we say that f is simple, or $\mathcal{M}$-simple, on A.

If the A_i are intervals in E^n, we call f a step function; it is a simple step function if $\mathcal{P}$ is finite.

The function values a_i are elements of T (possibly vectors). They may be infinite if $T = E^*$. Any simple map is also elementary, of course.

Definition

A map $f: S \rightarrow (T, \rho')$ is said to be measurable (or $\mathcal{M}$ -measurable) on a set A in $(S, \mathcal{M})$ iff

$$f=\lim_{m \rightarrow \infty} f_m \quad (\text{ pointwise }) \text{ on } A \tag{8.1.6}$$

for some sequence of functions $f_m: S \rightarrow T,$ all elementary on A. (See Chapter 4, §12 for "pointwise.")

Note 1. This implies $A \in \mathcal{M},$ as follows from Definitions 2 and 3. (Why ?)

Corollary 8.1.1

If $f: S \rightarrow (T, \rho')$ is elementary on A, it is measurable on A.

Proof

Set $f_m=f, m=1,2, \ldots,$ in Definition 4. Then clearly $f_m \rightarrow f$ on A. $square$

Corollary 8.1.2

If f is simple, elementary, or measurable on A in $(S, \mathcal{M})$, it has the same property on any subset $B \subseteq A$ with $B \in \mathcal{M}$.

Proof

Let f be simple on A; so $f=a_i$ on $A_i, i=1,2, \ldots, n$, for some finite $\mathcal{M}$ -partition, $A=\bigcup_{i=1}^n A_i$.

If $A \supseteq B \in \mathcal{M},$ then

$$\{B \cap A_i\}, \quad i=1,2, \ldots, n, \tag{8.1.7}$$

is a finite $\mathcal{M}$ -partition of $B($ why? $),$ and $f=a_i$ on $B \cap A_i$; so f is simple on B.

For elementary maps, use countable partitions.

Now let f be measurable on A, i.e.,

$$f=\lim_{m \rightarrow \infty} f_m \tag{8.1.8}$$

for some elementary maps f_m on A. As shown above, the f_m are elementary on B, too, and $f_m \rightarrow f$ on B; so f is measurable on B. $\square$

Corollary 8.1.3

If f is elementary or measurable on each of the (countably many) sets A_n in $(S, \mathcal{M})$, it has the same property on their union $A=\bigcup_n A_n$.

Proof

Let f be elementary on each A_n (so $A_n \in \mathcal{M}$ by Note 1$)$.

By Corollary 1 of Chapter 7, §1,

$$A=\bigcup A_n=\bigcup B_n \tag{8.1.9}$$

for some disjoint sets $B_n \subseteq A_n\left(B_n \in \mathcal{M}\right)$.

By Corollary 2, f is elementary on each B_n; i.e., constant on sets of some $\mathcal{M}$ -partition $\{B_{ni}\}$ of B_i.

All B_{ni} combined (for all n and all i) form an $\mathcal{M}$-partition of A,

$$A=\bigcup_n B_n=\bigcup_{n, i} B_{ni}. \tag{8.1.10}$$

As f is constant on each B_{ni}, it is elementary on A.

For measurable functions f, slightly modify the method used in Corollary 2. $\square$

Corollary 8.1.4

If $f: S \rightarrow (T, \rho')$ is measurable on A in $(S, \mathcal{M})$, so is the composite map $g \circ f$, provided $g: T \rightarrow (U, \rho'')$ is relatively continuous on $f[A]$.

Proof

By assumption,

$$f = \lim_{m \rightarrow \infty} f_m \text{ (pointwise)} \tag{8.1.11}$$

for some elementary maps f_m on A.

Hence by the continuity of g,

$$g(f_m(x)) \rightarrow g(f(x)), \tag{8.1.12}$$

i.e., $g \circ f_m \rightarrow g \circ f$ (pointwise) on A.

Moreover, all $g \circ f_m$ are elementary on A (for $g \circ f_m$ is constant on any partition set, if f_m is).

Thus $g \circ f$ is measurable on A, as claimed. $\square$

Theorem 8.1.1

If the maps $f, g, h: S \rightarrow E^1(C)$ are simple, elementary, or measurable on A in $(S, \mathcal{M})$, so are $f \pm g, fh, |f|^a$ (for real $a \neq 0$) and f/h (if $h \neq 0$ on A).

Similarly for vector-valued f and g and scalar-valued h.

Proof

First, let f and g be elementary on A. Then there are two $\mathcal{M}$-partitions,

$$A = \bigcup A_i = \bigcup B_k, \tag{8.1.13}$$

such that $f = a_i$ on A_i and $g = b_k$ on B_k, say.

The sets $A_i \cap B_k$ (for all i and k) then form a new $\mathcal{M}$ -partition of A(why?), such that both f and g are constant on each $A_i \cap B_k$(why?); hence so is $f \pm g$.

Thus $f \pm g$ is elementary on A. Similarly for simple functions.

Next, let f and g be measurable on A; so

$$f = \lim f_m \text{ and } g = \lim g_m \text{ (pointwise) on } A \tag{8.1.14}$$

for some elementary maps f_m, g_m.

By what was shown above, $f_m \pm g_m$ is elementary for each m. Also,

$$f_m \pm g_m \rightarrow f \pm g(\text{ pointwise }) \text{ on } A, \tag{8.1.15}$$

Thus $f \pm g$ is measurable on A.

The rest of the theorem follows quite similarly. $\square$

If the range space is E^n (or C^n) , then f has n real (complex) components $f_1, \ldots, f_n$, as in Chapter 4,§3 (Part II). This yields the following theorem.

Theorem 8.1.2

A function $f: S \rightarrow E^n\left(C^n\right)$ is simple, elementary, or measurable on a set A in $(S, \mathcal{M})$ iff all its n component functions $f_1, f_2, \ldots, f_n$ are.

Proof

For simplicity, consider $f: S \rightarrow E^2, f=\left(f_1, f_2\right)$.

If f_1 and f_2 are simple or elementary on A then (exactly as in Theorem 1), one can achieve that both are constant on sets $A_i \cap B_k$ of one and the same $\mathcal{M}$-partition of A. Hence $f=\left(f_1, f_2\right)$, too, is constant on each $A_i \cap B_k$, as required.

Conversely, let

$$f=\bar{c}_i=\left(a_i, b_i\right) \text{ on } C_i \tag{8.1.16}$$

for some $\mathcal{M}$-partition

$$A=\bigcup C_i. \tag{8.1.17}$$

Then by definition, $f_1=a_i$ and $f_2=b_i$ on C_i; so both are elementary (or simple) on A.

In the general case $\left(E^n \text{ or } C^n\right)$, the proof is analogous.

For measurable functions, the proof reduces to limits of elementary maps (using Theorem 2 of Chapter 3, §15). The details are left to the reader. □

Note 2. As $C=E^2$, a complex function $f: S \rightarrow C$ is simple, elementary, or measurable on A iff its real and imaginary parts are.

By Definition 4, a measurable function is a pointwise limit of elementary maps. However, if $\mathcal{M}$ is a σ-ring, one can make the limit uniform. Indeed, we have the following theorem.

Theorem 8.1.3

If $\mathcal{M}$ is a σ-ring, and $f: S \rightarrow\left(T, \rho^{\prime}\right)$ is $\mathcal{M}$-measurable on A, then

$$f=\lim_{m \rightarrow \infty} g_m \text{ (uniformly) on } A \tag{8.1.18}$$

for some finite elementary maps g_m.

Proof

Thus given $\varepsilon>0$, there is a finite elementary map g such that $\rho^{\prime}(f, g)<\varepsilon$ on A.

Theorem 8.1.4

If $\mathcal{M}$ is a σ-ring in S, if

$$f_m \rightarrow f(\text{pointwise}) \text{ on } A \tag{8.1.19}$$

$\left(f_m: S \rightarrow\left(T, \rho^{\prime}\right)\right)$, and if all f_m are $\mathcal{M}$ -measurable on A, so also is f.

Briefly: A pointwise limit of measurable maps is measurable (unlike continuous maps; cf. Chapter 4, §12).

Proof

By the second clause of Theorem 3, each f_m is uniformly approximated by some elementary map g_m on A, so that, taking $\varepsilon=1 / m, m=1,2, \ldots$,

$$\rho^{\prime}\left(f_m(x), g_m(x)\right)<\frac{1}{m} \quad \text{for all } x \in A \text{ and all } m. \tag{8.1.20}$$

Fixing such a g_m for each m, we show that $g_m \rightarrow f(\text{ pointwise })$ on A, as required in Definition 4.

Indeed, fix any $x \in A$. By assumption, $f_m(x) \rightarrow f(x)$. Hence, given $\delta > 0$,

$$(\exists k)(\forall m > k) \quad \rho'(f(x), f_m(x)) < \delta. \tag{8.1.21}$$

Take k so large that, in addition,

$$(\forall m > k) \quad \frac{1}{m} < \delta. \tag{8.1.22}$$

Then by the triangle law and by (1), we obtain for $m > k$ that

$$\begin{aligned} \rho'(f(x), g_m(x)) &\leq \rho'(f(x), f_m(x)) + \rho'(f_m(x), g_m(x)) \\ &< \delta + \frac{1}{m} < 2\delta \end{aligned}.$$

As δ is arbitrary, this implies $\rho'(f(x), g_m(x)) \rightarrow 0$, i.e., $g_m(x) \rightarrow f(x)$ for any (fixed) $x \in A$, thus proving the measurability of f. $\square$

Note 3. If

$$\mathcal{M} = \mathcal{B}(= \text{ Borel field in } S), \tag{8.1.23}$$

we often say "Borel measurable" for $\mathcal{M}$-measurable. If

$$\mathcal{M} = \{\text{ Lebesgue measurable sets in } E^n\}, \tag{8.1.24}$$

we say "Lebesgue (L) measurable" instead. Similarly for "Lebesgue-Stieltjes (LS) measurable."

8.1.E: Problems on Measurable and Elementary Functions in (S,M) (S,M)

Exercise $8.1.E.1$

Fill in all proof details in Corollaries 2 and 3 and Theorems 1 and 2.

Exercise $8.1.E.2$

Show that $\mathcal{P}^{\prime} \cap P^{\prime \prime}$ is as stated at the end of Definition 2.

Exercise $8.1.E.3$

Given $A \subseteq S$ and $f, f_{m} : S \rightarrow\left(T, \rho^{\prime}\right), m=1,2, \ldots,$ let

$$H=A\left(f_{m} \rightarrow f\right) \tag{8.1.E.1}$$

and

$$A_{mn}=A\left(\rho^{\prime}\left(f_{m}, f\right)<\frac{1}{n}\right). \tag{8.1.E.2}$$

Prove that
(i) $H=\bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{m=k}^{\infty} A_{mn};$
(ii) $H \in \mathcal{M}$ if all A_{mn} are in $\mathcal{M}$ and $\mathcal{M}$ is a σ-ring.
[Hint: $x \in H$ iff

$$(\forall n)(\exists k)(\forall m \geq k) \quad x \in A_{mn}. \tag{8.1.E.3}$$

Why?]

Exercise $8.1.E.3'$

Do Problem 3 for $T=E^{*}$ and $f=\pm \infty$ on H.
[Hint: If $f=+\infty$, $A_{mn}=A\left(f_{m}>n\right) \cdot$]

Exercise $8.1.E.4$

$\Rightarrow 4$. Let $f : S \rightarrow T$ be $\mathcal{M}$-elementary on A, with $\mathcal{M}$ a σ -ring in S. Show the following.
(i) $A(f=a) \in \mathcal{M}, A(f \neq a) \in \mathcal{M}$.
(ii) If $T=E^{*}$, then
$A(f<a), A(f \geq a), A(f>a),$ and $A(f \geq a)$
are in $\mathcal{M}$, too.
(iii) $(\forall B \subseteq T) A \cap f^{-1}[B] \in \mathcal{M}$.
[Hint: If

$$A=\bigcup_{i-1}^{\infty} A_{i} \tag{8.1.E.4}$$

and $f=a_{i}$ on $A_{i},$ then $A(f=a)$ is the countable union of those A_{i} for which $a_{i}=a$.]

Exercise 8.1.E.5

Do Problem 4(i) for measurable f.
[Hint: If $f=\lim f_m$ for elementary maps f_m, then

$$H=A(f=a)=A\left(f_m \rightarrow a\right). \tag{8.1.E.5}$$

Express H as in Problem 3, with

$$A_{mn}=A\left(h_m<\frac{1}{n}\right), \tag{8.1.E.6}$$

where $h_m=\rho^{\prime}\left(f_m, a\right)$ is elementary. (Why?) Then use Problems 4(ii) and 3(ii).]

Exercise 8.1.E.6

$\Rightarrow 6$. Given $f, g: S \rightarrow\left(T, \rho^{\prime}\right)$, let $h=\rho^{\prime}(f, g)$, i.e.,

$$h(x)=\rho^{\prime}(f(x), g(x)). \tag{8.1.E.7}$$

Prove that if f and g are elementary, simple, or measurable on A, so is h.
[Hint: Argue as in Theorem 1. Use Theorem 4 in Chapter 3, §15.]

Exercise 8.1.E.7

$\Rightarrow 7$. A set $B \subseteq\left(T, \rho^{\prime}\right)$ is called separable (in T) iff $B \subseteq \overline{D}$ (closure of D) for a countable set $D \subseteq T$.
Prove that if $f: S \rightarrow T$ is $\mathcal{M}$-measurable on A, then $f[A]$ is separable in T.
[Hint: $f=\lim f_m$ for elementary maps f_m; say,

$$f_m=a_{mi} \text{ on } A_{mi} \in \mathcal{M}, \quad i=1,2, \ldots \tag{8.1.E.8}$$

Let D consist of all $a_{mi}(m, i=1,2, \ldots)$; so D is countable (why?) and $D \subseteq T$.
Verify that

$$(\forall y \in f[A])(\exists x \in A) \quad y=f(x)=\lim f_m(x), \tag{8.1.E.9}$$

with $f_m(x) \in D$. Hence

$$(\forall y \in f[A]) \quad y \in \overline{D}, \tag{8.1.E.10}$$

by Theorem 3 of Chapter 3, §16.]

Exercise 8.1.E.8

$\Rightarrow 8$. Continuing Problem 7, prove that if $B \subseteq \overline{D}$ and $D=\left\{q_1, q_2, \ldots\right\}$, then

$$(\forall n) \quad B \subseteq \bigcup_{i=1}^{\infty} G_{q_i}\left(\frac{1}{n}\right), \tag{8.1.E.11}$$

[Hint: If $p \in B \subseteq \overline{D}$, any $G_p\left(\frac{1}{n}\right)$ contains some $q_1 \in D$; so

$$\rho'(p, q_i) < \frac{1}{n}, \text{ or } p \in G_{q_i}\left(\frac{1}{n}\right). \tag{8.1.E.12}$$

Thus

$$(\forall p \in B) \quad p \in \bigcup_{i-1}^{\infty} G_{q_i}\left(\frac{1}{n}\right) \cdot \Bigg] \tag{8.1.E.13}$$

Exercise $8.1.E.9$

Prove Corollaries 2 and 3 and Theorems 1 and 2, assuming that $\mathcal{M}$ is a semiring only.

Exercise $8.1.E.10$

Do Problem 4 for $\mathcal{M}$-simple maps, assuming that $\mathcal{M}$ is a ring only.

8.2: Measurability of Extended-Real Functions

Henceforth we presuppose a measurable space $(S, \mathcal{M})$, where $\mathcal{M}$ is a σ-ring in S. Our aim is to prove the following basic theorem, which is often used as a definition, for extended-real functions $f : S \rightarrow E^*$.

Theorem 8.2.1

A function $f : S \rightarrow E^*$ is measurable on a set $A \in \mathcal{M}$ iff it satisfies one of the following equivalent conditions (hence all of them):

$$\begin{array}{ll}\left(\mathrm{i}^*\right)\left(\forall a \in E^*\right) A(f>a) \in \mathcal{M}; & \left(\mathrm{ii}^*\right)\left(\forall a \in E^*\right) A(f \geq a) \in \mathcal{M}; \\ \left(\mathrm{iii}^*\right)\left(\forall a \in E^*\right) A(f<a) \in \mathcal{M}; & \left(\mathrm{iv}^*\right)\left(\forall a \in E^*\right) A(f \leq a) \in \mathcal{M}.\end{array} \tag{8.2.1}$$

We first prove the equivalence of these conditions by showing that $\left(\mathrm{i}^*\right) \Rightarrow\left(\mathrm{ii}^*\right) \Rightarrow\left(\mathrm{iii}^*\right) \Rightarrow\left(\mathrm{i} v^*\right) \Rightarrow\left(\mathrm{i}^*\right)$, closing the "circle."

$\left(\mathrm{i}^*\right) \Rightarrow\left(\mathrm{ii}^*\right)$. Assume $\left(\mathrm{i}^*\right)$. If $a=-\infty$,

$$A(f \geq a)=A \in \mathcal{M} \tag{8.2.2}$$

by assumption. If $a=+\infty$,

$$A(f \geq a)=A(f=\infty)=\bigcap_{n=1}^{\infty} A(f>n) \in \mathcal{M} \tag{8.2.3}$$

by $\left(\mathrm{i}^*\right)$. And if $a \in E^1$,

$$A(f \geq a)=\bigcap_{n=1}^{\infty} A\left(f>a-\frac{1}{n}\right). \tag{8.2.4}$$

(Verify!) By $\left(\mathrm{i}^*\right)$,

$$A\left(f>a-\frac{1}{n}\right) \in \mathcal{M}; \tag{8.2.5}$$

so $A(f \geq a) \in \mathcal{M}(\text{ a } \sigma\text{-ring! })$.

$\left(\mathrm{ii}^*\right) \Rightarrow\left(\mathrm{iii}^*\right)$. For $\left(\mathrm{i}^*\right)$ and $A \in \mathcal{M}$ imply

$$A(f<a)=A-A(f \geq a) \in \mathcal{M}. \tag{8.2.6}$$

$\left(\mathrm{iii}^*\right) \Rightarrow\left(\mathrm{iv}^*\right)$. If $a \in E^1$,

$$A(f \leq a)=\bigcap_{n=1}^{\infty} A\left(f<a+\frac{1}{n}\right) \in \mathcal{M}. \tag{8.2.7}$$

What if $a=\pm \infty$?

$\left(\mathrm{i} v^*\right) \Rightarrow\left(\mathrm{i}^*\right)$. Indeed, $\left(\mathrm{iv}^*\right)$ and $A \in \mathcal{M}$ imply

$$A(f>a)=A-A(f \leq a) \in \mathcal{M}. \tag{8.2.8}$$

Thus, indeed, each of $\left(\mathrm{i}^*\right)$ to $\left(\mathrm{iv}^*\right)$ implies the others. To finish, we need two lemmas that are of interest in their own right.

Lemma 8.2.1

If the maps $f_m : S \rightarrow E^*(m=1,2, \ldots)$ satisfy conditions $\left(\mathrm{i}^*\right)-\left(\mathrm{iv}^*\right)$, so also do the functions

$$\sup f_m, \inf f_m, \overline{\lim} f_m, \text{ and } \underline{\lim} f_m, \tag{8.2.9}$$

defined pointwise, i.e.,

$$\left(\sup f_m\right)(x)=\sup f_m(x), \tag{8.2.10}$$

and similarly for the others.

Proof

Let $f=\sup f_m$. Then

$$A(f \leq a)=\bigcap_{m=1}^{\infty} A\left(f_m \leq a\right) \quad \text{(Why?)} \tag{8.2.11}$$

But by assumption,

$$A\left(f_m \leq a\right) \in \mathcal{M} \tag{8.2.12}$$

$\left(f_m \text{ satisfies } \left(\mathrm{iv}^*\right)\right)$. Hence $A(f \leq a) \in \mathcal{M}(\text{ for } \mathcal{M} \text{ is a } \sigma\text{-ring })$.

Thus sup f_m satisfies $\left(i^*\right)-\left(iv^*\right)$.

So does inf f_m; for

$$A\left(\inf f_m \geq a\right)=\bigcap_{m=1}^{\infty} A\left(f_m \geq a\right) \in \mathcal{M}. \tag{8.2.13}$$

(Explain!)

So also do $\underline{\lim} f_m$ and $\overline{\lim} f_m$; for by definition,

$$\underline{\lim}\, t f_m=\sup_k g_k, \tag{8.2.14}$$

where

$$g_k=\inf_{m \geq k} f_m \tag{8.2.15}$$

satisfies $\left(\mathrm{i}^*\right)-\left(\mathrm{iv}^*\right)$, as was shown above; hence so does sup $g_k=\underline{\lim} f_m$.

Similarly for $\overline{\lim} f_m$. $\square$

Lemma 8.2.2

If f satisfies $\left(\mathrm{i}^*\right)-\left(\mathrm{iv}^*\right)$, then

$$f=\lim_{m \rightarrow \infty} f_m \text{ (uniformly) on } A \tag{8.2.16}$$

for some sequence of finite functions f_m, all $\mathcal{M}$ -elementary on A.

Moreover, if $f \geq 0$ on A, the f_m can be made nonnegative, with $\left\{f_m\right\} \uparrow$ on A.

Proof

Let $H=A(f=+\infty), K=A(f=-\infty), \quad$ and

$$A_{mk}=A\left(\frac{k-1}{2^m} \leq f<\frac{k}{2^m}\right) \tag{8.2.17}$$

for $m=1,2, \ldots$ and $k=0, \pm 1, \pm 2, \ldots, \pm n, \ldots$

By $\left(\mathrm{i}^*\right)-\left(\mathrm{iv}^*\right)$,

$$H=A(f=+\infty)=A(f \geq+\infty) \in \mathcal{M}, \tag{8.2.18}$$

$K \in \mathcal{M}$, and

$$A_{mk}=A\left(f \leq \frac{k-1}{2^m}\right) \cap A\left(f<\frac{k}{2^m}\right) \in \mathcal{M}. \tag{8.2.19}$$

Now define

$$(\forall m) \quad f_m = \frac{k-1}{2^m} \text{ on } A_{mk}, \tag{8.2.20}$$

$f_m = m$ on H, and $f_m = -m$ on K. Then each f_m is finite and elementary on A since

$$(\forall m) \quad A = H \cup K \cup \bigcup_{k=-\infty}^{\infty} A_{mk} (disjoint) \tag{8.2.21}$$

and f_m is constant on H, K, and A_{mk}.

We now show that $f_m \rightarrow f$ (uniformly) on H, K, and

$$J = \bigcup_{k=-\infty}^{\infty} A_{mk}, \tag{8.2.22}$$

hence on A.

Indeed, on H we have

$$\lim f_m = \lim m = +\infty = f, \tag{8.2.23}$$

and the limit is uniform since the f_m are constant on H.

Similarly,

$$f_m = -m \rightarrow -\infty = f \text{ on } K. \tag{8.2.24}$$

Finally, on A_{mk} we have

$$(k-1)2^{-m} \leq f < k2^{-m} \tag{8.2.25}$$

and $f_m = (k-1)2^{-m};$ hence

$$|f_m - f| < k2^{-m} - (k-1)2^{-m} = 2^{-m}. \tag{8.2.26}$$

Thus

$$|f_m - f| < 2^{-m} \rightarrow 0 \tag{8.2.27}$$

on each A_{mk}, hence on

$$J = \bigcup_{k=-\infty}^{\infty} A_{mk}. \tag{8.2.28}$$

By Theorem 1 of Chapter 4, §12, it follows that $f_m \rightarrow f(\text{ uniformly })$ on J. Thus, indeed, $f_m \rightarrow f(\text{ uniformly })$ on A.

If, further, $f \geq 0$ on A, then $K = \emptyset$ and $A_{mk} = \emptyset$ for $k \leq 0$. Moreover, on passage from m to $m+1$, each $A_{mk} (k > 0)$ splits into two sets. On one, $f_{m+1} = f_m$; on the other, $f_{m+1} > f_m$. (Why?)

Thus $0 \leq f_m \nearrow f(\text{ uniformly })$ on A, and all is proved. □

Theorem 8.2.1 (Restated)

A function $f : S \rightarrow E^*$ is measurable on a set $A \in \mathcal{M}$ iff it satisfies one of the following equivalent conditions (hence all of them):

$$\begin{array}{ll} (\mathrm{i}^*)\ (\forall a \in E^*)\ A(f > a) \in \mathcal{M}; & (\mathrm{ii}^*)\ (\forall a \in E^*)\ A(f \geq a) \in \mathcal{M}; \\ (\mathrm{iii}^*)\ (\forall a \in E^*)\ A(f < a) \in \mathcal{M}; & (\mathrm{iv}^*)\ (\forall a \in E^*)\ A(f \leq a) \in \mathcal{M}. \end{array} \tag{8.2.29}$$

Proof

If f is measurable on A, then by definition, $f = \lim f_m$ (pointwise) for some elementary maps f_m on A.

By Problem 4 (ii) in §1, all f_m satisfy (i*)-(iv*). Thus so does $f$by Lemma 1, for here $f=\lim f_m=\overline{\lim} f_m$.

The converse follows by Lemma 2. This completes the proof. □

Note 1. Lemmas 1 and 2 prove Theorems 3 and 4 of $1, for $f: S \rightarrow E^*$. By using also Theorem 2 in §1, one easily extends this to $f: S \rightarrow E^n\left(C^n\right)$. Verify!

Corollary 8.2.1

If $f: S \rightarrow E^*$ is measurable on A, then

$$\left(\forall a \in E^*\right) \quad A(f=a) \in \mathcal{M} \text{ and } A(f \neq a) \in \mathcal{M}. \tag{8.2.30}$$

Indeed,

$$A(f=a)=A(f \geq a) \cap A(f \leq a) \in \mathcal{M} \tag{8.2.31}$$

and

$$A(f \neq a)=A-A(f=a) \in \mathcal{M}. \tag{8.2.32}$$

Corollary 8.2.2

If $f: S \rightarrow\left(T, \rho^{\prime}\right)$ is measurable on A in $(S, \mathcal{M})$, then

$$A \cap f^{-1}[G] \in \mathcal{M} \tag{8.2.33}$$

for every globe $G=G_q(\delta)$ in $\left(T, \rho^{\prime}\right)$.

Proof

Define $h: S \rightarrow E^1$ by

$$h(x)=\rho^{\prime}(f(x), q). \tag{8.2.34}$$

Then h is measurable on A by Problem 6 in §1. Thus by Theorem 1,

$$A(h<\delta) \in \mathcal{M}. \tag{8.2.35}$$

But as is easily seen,

$$A(h<\delta)=\left\{x \in A \mid \rho^{\prime}(f(x), q)<\delta\right\}=A \cap f^{-1}\left[G_q(\delta)\right]. \tag{8.2.36}$$

Hence the result. □

Definition

Given $f, g: S \rightarrow E^*,$ we define the maps $f \vee g$ and $f \wedge g$ on S by

$$(f \vee g)(x)=\max \{f(x), g(x)\} \tag{8.2.37}$$

and

$$(f \wedge g)(x)=\min \{f(x), g(x)\}; \tag{8.2.38}$$

similarly for $f \vee g \vee h, f \wedge g \wedge h,$ etc.

We also set

$$f^{+}=f \vee 0 \text{ and } f^{-}=-f \vee 0. \tag{8.2.39}$$

Clearly, $f^{+} \geq 0$ and $f^{-} \geq 0$ on S. Also, $f=f^{+}-f^{-}$ and $|f|=f^{+}+f^{-}$.

(Why?) We now obtain the following theorem.

Theorem 8.2.2

If the functions $f, g: S \rightarrow E^{*}$ are simple, elementary, or measurable on A, so also are $f \pm g, f g, f \vee g, f \wedge g, f^{+}, f^{-},$ and $|f|^{a}(a \neq 0)$.

Proof

If f and g are finite, this follows by Theorem 1 of §1 on verifying that

$$f \vee g=\frac{1}{2}(f+g+|f-g|) \tag{8.2.40}$$

and

$$f \wedge g=\frac{1}{2}(f+g-|f-g|) \tag{8.2.41}$$

on S. (Check it!)

Otherwise, consider

$$A(f=+\infty), A(f=-\infty), A(g=+\infty), \text{ and } A(g=-\infty). \tag{8.2.42}$$

By Theorem 1, these are $\mathcal{M}$-sets; hence so is their union U.

On each of them $f \vee g$ and $f \wedge g$ equal f or g; so by Corollary 3 in §1, $f \vee g$ and $f \wedge g$ have the desired properties on U. So also have $f^{+}=f \vee 0$ and $f^{-}=-f \vee 0(\text{ take } g=0)$.

We claim that the maps $f \pm g$ and $f g$ are simple (hence elementary and measurable) on each of the four sets mentioned above, hence on U.

For example, on $A(f=+\infty)$,

$$f \pm g=+\infty(\text{ constant }) \tag{8.2.43}$$

by our conventions $\left(2^{*}\right)$ in Chapter 4, §4. For $f g$, split $A(f=+\infty)$ into three sets $A_{1}, A_{2}, A_{3} \in \mathcal{M}$, with $g>0$ on $A_{1}, g<0$ on A_{2}, and $g=0$ on A_{3}; so $f g=+\infty$ on $A_{1}, f g=-\infty$ on A_{2}, and $f g=0$ on A_{3}. Hence $f g$ is simple on $A(f=+\infty)$.

For $|f|^{a}$, use $U=A(|f|=\infty)$. Again, the theorem holds on U, and also on $A-U$, since f and g are finite on $A-U \in \mathcal{M}$. Thus it holds on $A=(A-U) \cup U$ by Corollary 3 in §1. □

Note 2. Induction extends Theorem 2 to any finite number of functions.

Note 3. Combining Theorem 2 with $f=f^{+}-f^{-},$ we see that $f: S \rightarrow E^{*}$ is simple (elementary, measurable) iff f^{+} and f^{-} are. We also obtain the following result.

Theorem 8.2.3

If the functions $f, g: S \rightarrow E *$ are measurable on $A \in \mathcal{M}$, then $A(f \geq g) \in \mathcal{M}, A(f<g) \in \mathcal{M}, A(f=g) \in \mathcal{M},$ and $A(f \neq g) \in \mathcal{M}$.

8.2.E: Problems on Measurable Functions in (S,M,m) (S,M,m)

Exercise $8.2.E.1$

In Theorem 1, give the details in proving the equivalence of $\left(i^{*}\right)-\left(i v^{*}\right)$.

Exercise $8.2.E.2$

Prove Note 1.

Exercise $8.2.E.2'$

Prove that $f=f^{+}-f^{-}$ and $|f|=f^{+}+f^{-}$.

Exercise $8.2.E.3$

Complete the proof of Theorem 2, in detail.

Exercise $8.2.E.4$

$\Rightarrow 4$. Prove Theorem 3.
[Hint: By our conventions, $A(f \geq g)=A(f-g \geq 0)$ even if g or f is $\pm \infty$ for some $x \in A$. (Verify all cases!) By Theorems 1 and 2, $A(f-g \geq 0) \in \mathcal{M};$ so $A(f \geq g) \in \mathcal{M},$ and $A(f<g)=A-A(f \geq g) \in \mathcal{M}.$ Proceed.]

Exercise $8.2.E.5$

Show that the measurability of $|f|$ does not imply that of f.
[Hint: Let $f=1$ on Q and $f=-1$ on $A-Q$ for some $Q \notin \mathcal{M}(Q \subset A)$; e.g., use Q of Problem 6 in Chapter 7, §8.]

Exercise $8.2.E.6$

$\Rightarrow 6$. Show that a function $f \geq 0$ is measurable on A iff $f_{m} \nearrow f$ (pointwise) on A for some finite simple maps $f_{m} \geq 0,\left\{f_{m}\right\} \uparrow$.
[Hint: Modify the proof of Lemma 2, setting $H_{m}=A(f \geq m)$ and $f_{m}=m$ on H_{m}, and defining the $A_{m k}$ for $1 \leq k \leq m 2^{m}$ only.]

Exercise $8.2.E.7$

$\Rightarrow 7$. Prove Theorem 3 in §1.
[Outline: By Problems 7 and 8 in $\xi 1$, there are $q_{i} \in T$ such that

$$(\forall n) \quad f[A] \subseteq \bigcup_{i=1}^{\infty} G q_{i}\left(\frac{1}{n}\right). \tag{8.2.E.1}$$

Set

$$A_{n i}=A \cap f^{-1}\left[G_{q_{i}}\left(\frac{1}{n}\right)\right] \in \mathcal{M} \tag{8.2.E.2}$$

by Corollary 2; so $\rho^{\prime}\left(f(x), q_{i}\right)<\frac{1}{n}$ on $A_{n i}$.
By Corollary 1 in Chapter 7, §1

$$A=\bigcup_{i=1}^{\infty} A_{ni}=\bigcup_{i=1}^{\infty} B_{ni}(\text{disjoint}) \tag{8.2.E.3}$$

for some sets $B_{ni}\in\mathcal{M}, B_{ni}\subseteq A_{ni}$. Now define $g_n=q_i$ on B_{ni}; so $\rho'(f,g_n)<\frac{1}{n}$ on each B_{ni}, hence on A. By Theorem 1 in Chapter 4, §12, $g_n\to f$ (uniformly) on A.]

Exercise 8.2.E.8

$\Rightarrow 8$. Prove that $f:S\rightarrow E^1$ is $\mathcal{M}$ -measurable on A iff $A\cap f^{-1}[B]\in\mathcal{M}$ for every Borel set B (equivalently, for every open set B) in E^1. (In the case $f:S\rightarrow E^*$, add: "and for $B=\{\pm\infty\}$. ")
[Outline: Let

$$\mathcal{R}=\left\{X\subseteq E^1 | A\cap f^{-1}[X]\in\mathcal{M}\right\}. \tag{8.2.E.4}$$

Show that $\mathcal{R}$ is a σ -ring in E^1.
Now, by Theorem 1, if f is measurable on $A, \mathcal{R}$ contains all open intervals; for

$$A\cap f^{-1}[(a,b)]=A(f>a)\cap A(f<b). \tag{8.2.E.5}$$

Then by Lemma 2 of Chapter $7, 2, \mathcal{R}\supseteq\mathcal{G}$, hence $\mathcal{R}\supseteq\mathcal{B}$. (Why?)
Conversely, if so,

$$(a,\infty)\in\mathcal{R}\Rightarrow A\cap f^{-1}[(a,\infty)]\in\mathcal{M}\Rightarrow A(f>a)\in\mathcal{M}.] \tag{8.2.E.6}$$

Exercise 8.2.E.9

$\Rightarrow 9$. Do Problem 8 for $f:S\rightarrow E^n$.
[Hint: If $f=(f_1,\ldots,f_n)$ and $B=(\bar{a},\bar{b})\subset E^n$, with $\bar{a}=(a_1,\ldots,a_n)$ and $\bar{b}=(b_1,\ldots,b_n)$, show that

$$f^{-1}[B]=\bigcap_{k=1}^{n} f_k^{-1}\left[(a_k,b_k)\right]. \tag{8.2.E.7}$$

Apply Problem 8 to each $f_k:S\rightarrow E^1$ and use Theorem 2 in §1. Proceed as in Problem 8.]

Exercise 8.2.E.10

Do Problem 8 for $f:S\rightarrow C^n$, treating C^n as E^{2n}.

Exercise 8.2.E.11

Prove that $f:S\rightarrow(T,\rho')$ is measurable on A in $(S,\mathcal{M})$ iff
(i) $A\cap f^{-1}[G]\in\mathcal{M}$ for every open globe $G\subseteq T$, and
(ii) $f[A]$ is separable in $T($ Problem 7 in §1).
[Hint: If so, proceed as in Problem 7 (without assuming measurability of f) to show that $f=\lim g_n$ for some elementary maps g_n on A. For the converse, use Problem 7 in §1 and Corollary 2 in §2.]

Exercise 8.2.E.12

(i) Show that if all of T is separable (Problem 7 in §1), there is a sequence of globes $G_k\subseteq T$ such that each nonempty open set
$B\subseteq T$ is the union of some of these G_k.

(ii) Show that E^n and C^n are separable.
[Hints: (i) Use the $G_{q_i}\left(\frac{1}{n}\right)$ of Problem 8 in §1, putting them in one sequence.
(ii) Take $D=R^n \subset E^n$ in Problem 7 of §1.]

Exercise 8.2.E.13

Do Problem 11 with "globe $G \subseteq T''$ replaced by "Borel set $B \subseteq T$."
[Hints: Treat f as $f: A \rightarrow T', T'=f[A]$, noting that

$$A \cap f^{-1}[B]=A \cap f^{-1}\left[B \cap T'\right]. \tag{8.2.E.8}$$

By Problem 12, if $B \neq \emptyset$ is open in T, then $B \cap T'$ is a countable union of "globes" $G_q \cap T'$ in $\left(T', \rho'\right)$; see Theorem 4 in Chapter 3, §12. Proceed as in Problem 8, replacing E^1 by T.]

Exercise 8.2.E.14

A map $g:\left(T, \rho'\right) \rightarrow\left(U, \rho''\right)$ is said to be of Baire class 0 $\left(g \in \mathrm{B}_0\right)$ on $D \subseteq T$ iff g is relatively continuous on D. Inductively, g is of Baire class $n\left(g \in \mathbf{B}_n, n \geq 1\right)$ iff $g=\lim g_m$ (pointwise) on D for some maps $g_m \in \mathbf{B}_{n-1}$. Show by induction that Corollary 4 in §1 holds also if $g \in \mathbf{B}_n$ on $f[A]$ for some n.

8.3: Measurable Functions in (S,M,m) (S,M,m)

I. Henceforth we shall presuppose not just a measurable space (§1) but a measure space $(S, \mathcal{M}, m)$, where $m : \mathcal{M} \rightarrow E^*$ is a measure on a σ -ring $\mathcal{M} \subseteq 2^S$.

We saw in Chapter 7 that one could often neglect sets of Lebesgue measure zero on $E^n -$ if a property held everywhere except on a set of Lebesgue measure zero, we said it held "almost everywhere." The following definition generalizes this usage.

Definition

We say that a property $P(x)$ holds for almost all $x \in A$ (with respect to the measure m) or almost everywhere (a.e. (m)) on A iff it holds on $A-Q$ for some $Q \in \mathcal{M}$ with $mQ=0$.

Thus we write

$$f_n \rightarrow f(a. e.) \text{ or } f=\lim f_n(a. e. (m)) \text{ on } A \tag{8.3.1}$$

iff $f_n \rightarrow f(\text{ pointwise }) \text{ on } A-Q, mQ=0.$ Of course, "pointwise" implies " $a. e.$ " $(\text{ take } Q=\emptyset),$ but the converse fails.

Definition

We say that $f : S \rightarrow (T, \rho')$ is almost measurable on A iff $A \in \mathcal{M}$ and f is $\mathcal{M}$ -measurable on $A-Q, mQ=0$.

We then also say that f is m -measurable (m being the measure involved) as opposed to $\mathcal{M}$-measurable.

Observe that we may assume $Q \subseteq A$ here (replace Q by $A \cap Q$).

***Note 1.** If m is a generalized measure (Chapter 7, §11), replace $mQ=0$ by $v_m Q=0 \, (v_m = \text{total variation of } m)$ in Definitions 1 and 2 and in the following proofs.

Corollary 8.3.1

If the functions

$$f_n : S \rightarrow (T, \rho'), \quad n=1,2, \ldots \tag{8.3.2}$$

are m-measurable on A, and if

$$f_n \rightarrow f(a. e. (m)) \tag{8.3.3}$$

on A, then f is m-measurable on A.

Proof

By assumption, $f_n \rightarrow f(\text{ pointwise }) \text{ on } A-Q_0, mQ_0=0.$ Also, f_n is $\mathcal{M}$-measurable on

$$A-Q_n, mQ_n=0, \quad n=1,2, \ldots \tag{8.3.4}$$

(The Q_n need not be the same.)

Let

$$Q=\bigcup_{n=0}^{\infty} Q_n; \tag{8.3.5}$$

so

$$mQ \leq \sum_{n=0}^{\infty} mQ_n=0. \tag{8.3.6}$$

By Corollary 2 in §1, all f_n are $\mathcal{M}$-measurable on $A-Q$ (why?), and $f_n \rightarrow f$ (pointwise) on $A-Q,$ as $A-Q \subseteq A-Q_0$.

Thus by Theorem 4 in §1, f is $\mathcal{M}$ -measurable on $A-Q$. As $mQ=0$, this is the desired result. □

Corollary 8.3.2

If $f=g(a.e.(m))$ on A and f is m-measurable on A, so is g.

Proof

By assumption, $f=g$ on $A-Q_1$ and f is $\mathcal{M}$-measurable on $A-Q_2$, with $mQ_1=mQ_2=0$.

Let $Q=Q_1 \cup Q_2$. Then $mQ=0$ and $g=f$ on $A-Q$. (Why?)

By Corollary 2 of §1, f is $\mathcal{M}$-measurable on $A-Q$. Hence so is g, as claimed. □

Corollary 8.3.3

If $f: S \rightarrow (T, \rho')$ is m -measurable on A, then

$$f=\lim_{n \rightarrow \infty} f_n(\text{uniformly}) \text{ on } A-Q(mQ=0), \tag{8.3.7}$$

for some maps f_n, all elementary on $A-Q$.

Proof

Add proof here and it will automatically be hidden

(Compare Corollary 3 with Theorem 3 in §1).

Quite similarly all other propositions of §1 carry over to almost measurable (i.e., m -measurable) functions. Note, however, that the term "measurable" in §§1 and 2 always meant " $\mathcal{M}$ -measurable." This implies m-measurability (take $Q=\emptyset$), but the converse fails. (See Note 2, however.)

We still obtain the following result.

Corollary 8.3.4

If the functions

$$f_n: S \rightarrow E^* \quad (n=1,2, \ldots) \tag{8.3.8}$$

are m-measurable on a set A, so also are

$$\sup f_n, \inf f_n, \overline{\lim} f_n, \text{ and } \underline{\lim} f_n \tag{8.3.9}$$

(Use Lemma 1 of §2).

Similarly, Theorem 2 in §2 carries over to m-measurable functions.

Note 2. If m is complete (such as Lebesgue measure and LS measures) then, for $f: S \rightarrow E^*\left(E^n, C^n\right), m-$ and $\mathcal{M}$-measurability coincide (see Problem 3 below).

II. Measurability and Continuity. To study the connection between these notions, we first state two lemmas, often treated as definitions.

Lemma 8.3.1

A map $f: S \rightarrow E^n\left(C^n\right)$ is $\mathcal{M}$-measurable on A iff

$$A \cap f^{-1}[B] \in \mathcal{M} \tag{8.3.10}$$

for each Borel set (equivalently, open set) B in $E^n\ (C^n)$.

Proof

See Problems $8-10$ in §2 for a sketch of the proof.

Lemma 8.3.2

A map $f:(S,\rho)\rightarrow(T,\rho')$ is relatively continuous on $A\subseteq S$ iff for any open set $B\subseteq(T,\rho')$, the set $A\cap f^{-1}[B]$ is open in (A,ρ) as a subspace of (S,ρ).
(This holds also with "open" replaced by "closed.")

Proof

By Chapter 4, §1, footnote 4, f is relatively continuous on A iff its restriction to A (call it $g:A\rightarrow T$) is continuous in the ordinary sense.
Now, by Problem $15(\text{iv})(\text{v})$ in Chapter 4, §2, with S replaced by A, this means that $g^{-1}[B]$ is open (closed) $in(A,\rho)$ when B is so in (T,ρ'). But

$$g^{-1}[B]=\{x\in A|f(x)\in B\}=A\cap f^{-1}[B]. \tag{8.3.11}$$

(Why?) Hence the result follows. □

Theorem 8.3.1

Let $m:\mathcal{M}\rightarrow E^*$ be a topological measure in (S,ρ). If $f:S\rightarrow\ E^n\ (C^n)$ is relatively continuous on a set $A\in\mathcal{M}$, it is $\mathcal{M}$-measurable on A.

Proof

Let B be open in $E^n\ (C^n)$. By Lemma 2,

$$A\cap f^{-1}[B] \tag{8.3.12}$$

is open $in(A,\rho)$. Hence by Theorem 4 of Chapter 3, §12,

$$A\cap f^{-1}[B]=A\cap U \tag{8.3.13}$$

for some open set U in (S,ρ).
Now, by assumption, A is in $\mathcal{M}$. So is U, as $\mathcal{M}$ is topological $(\mathcal{M}\supseteq\mathcal{G})$.
Hence

$$A\cap f^{-1}[B]=A\cap U\in\mathcal{M} \tag{8.3.14}$$

for any open $B\subseteq E^n\ (C^n)$. The result follows by Lemma 1. □

Note 3. The converse fails. For example, the Dirichlet function (Example (c) in Chapter 4, §1) is L-measurable (even simple) but discontinuous everywhere.
Note 4. Lemma 1 and Theorem 1 hold for a map $f:S\rightarrow(T,\rho')$, too, provided $f[A]$ is separable, i.e.,

$$f[A]\subseteq\overline{D} \tag{8.3.15}$$

for a countable set $D\subseteq T$ (cf. Problem 11 in §2).

***III.** For strongly regular measures (Definition 5 in Chapter 7, §7), we obtain the following theorem.

*Theorem 8.3.2

(Luzin). Let $m : \mathcal{M} \rightarrow E^{*}$ be a strongly regular measure in (S, ρ). Let $f : S \rightarrow (T, \rho')$ be m-measurable on A. Then given $\varepsilon > 0,$ there is a closed set $F \subseteq A (F \in \mathcal{M})$ such that

$$m(A - F) < \varepsilon \tag{8.3.16}$$

and f is relatively continuous on F.
(Note that if $T = E^{*}, \rho'$ is as in Problem 5 of Chapter 3, §11.)

Proof

By assumption, f is $\mathcal{M}$-measurable on a set

$$H = A - Q, mQ = 0; \tag{8.3.17}$$

so by Problem 7 in §1, $f[H]$ is separable in T. We may safely assume that f is $\mathcal{M}$-measurable on S and that all of T is separable. (If not, replace S and T by H and $f[H]$, restricting f to H, and m to $\mathcal{M}$-sets inside H; see also Problems 7 and 8 below.)
Then by Problem 12 of §2, we can fix globes $G_1, G_2, \ldots$ in T such that

$$\text{each open set } B \neq \emptyset \text{ in } T \text{ is the union of a subsequence of } \{G_k\}. \tag{8.3.18}$$

Now let $\varepsilon > 0,$ and set

$$S_k = S \cap f^{-1}[G_k] = f^{-1}[G_k], \quad k = 1, 2, \ldots \tag{8.3.19}$$

By Corollary 2 in §2, $S_k \in \mathcal{M}.$ As m is strongly regular, we find for each S_k an open set

$$U_k \supseteq S_k, \tag{8.3.20}$$

with $U_k \in \mathcal{M}$ and

$$m(U_k - S_k) < \frac{\varepsilon}{2^{k+1}}. \tag{8.3.21}$$

Let $B_k = U_k - S_k, D = \bigcup_k B_k;$ so $D \in \mathcal{M}$ and

$$mD \leq \sum_k mB_k \leq \sum_k \frac{\varepsilon}{2^{k+1}} \leq \frac{1}{2}\varepsilon \tag{8.3.22}$$

and

$$U_k - B_k = S_k = f^{-1}[G_k]. \tag{8.3.23}$$

As $D = \bigcup B_k,$ we have

$$(\forall k) \quad B_k - D = B_k \cap (-D) = \emptyset. \tag{8.3.24}$$

Hence by $(2'),$

$$(\forall k) \quad f^{-1}\left[G_k\right] \cap(-D) \begin{aligned}[t] & =\left(U_k-B_k\right) \cap(-D) \\ & =\left(U_k \cap(-D)\right)-\left(B_k \cap(-D)\right)=U_k \cap(-D)\end{aligned}.$$

Combining this with (1), we have, for each open set $B=\bigcup_i G_{k_i}$ in T,

$$f^{-1}[B] \cap(-D)=\bigcup_i f^{-1}\left[G_{k_i}\right] \cap(-D)=\bigcup_i U_{k_i} \cap(-D). \tag{8.3.25}$$

since the U_{k_i} are open in S (by construction), the set (3) is open in $S-D$ as a subspace of S. By Lemma 2, then, f is relatively continuous on $S-D$, or rather on

$$H-D=A-Q-D \tag{8.3.26}$$

(since we actually substituted S for H in the course of the proof). As $mQ=0$ and $mD<\frac{1}{2}\varepsilon$ by (2),

$$m(H-D)<mA-\frac{1}{2}\varepsilon. \tag{8.3.27}$$

Finally, as m is strongly regular and $H-D \in \mathcal{M}$, there is a closed $\mathcal{M}$-set

$$F \subseteq H-D \subseteq A \tag{8.3.28}$$

such that

$$m(H-D-F)<\frac{1}{2}\varepsilon. \tag{8.3.29}$$

since f is relatively continuous on $H-D$, it is surely so on F. Moreover,

$$A-F=(A-(H-D)) \cup(H-D-F); \tag{8.3.30}$$

so

$$m(A-F) \leq m(A-(H-D))+m(H-D-F)<\frac{1}{2}\varepsilon+\frac{1}{2}\varepsilon=\varepsilon. \tag{8.3.31}$$

This completes the proof. □

*Lemma 8.3.3

Given $[a, b] \subset E^1$ and disjoint closed sets $A, B \subseteq(S, \rho)$, there always is a continuous map $g: S \rightarrow[a, b]$ such that $g=a$ on A and $g=b$.

Proof

If $A=\emptyset$ or $B=\emptyset$, set $g=b$ or $g=a$ on all of S.
If, however, A and B are both nonempty, set

$$g(x)=a+\frac{(b-a)\rho(x, A)}{\rho(x, A)+\rho(x, B)}. \tag{8.3.32}$$

As A is closed, $\rho(x, A)=0$ iff $x \in A$ (Problem 15 in Chapter 3, §14); similarly for B. Thus $\rho(x, A)+\rho(x, B) \neq 0$.

Also, $g=a$ on A, $g=b$ on B, and $a \leq g \leq b$ on S.
For continuity, see Chapter 4, §8, Example (e). $\square$

*Lemma 8.3.4

(Tietze). If $f:(S, \rho) \rightarrow E^{*}$ is relatively continuous on a closed set $F \subseteq S$, there is a function $g: S \rightarrow E^{*}$ such that $g=f$ on F,

$$\inf g[S]=\inf f[F], \sup g[S]=\sup f[F], \tag{8.3.33}$$

and g is continuous on all of S.
(We assume E^{*} metrized as in Problem 5 of Chapter 3, §11. If $|f|<\infty$, the standard metric in E^{1} may be used.)

Proof Outline

First, assume $\inf f[F]=0$ and $\sup f[F]=1$. Set

$$A=F\left(f \leq \frac{1}{3}\right)=F \cap f^{-1}\left[\left[0, \frac{1}{3}\right]\right] \tag{8.3.34}$$

and

$$B=F\left(f \geq \frac{2}{3}\right)=F \cap f^{-1}\left[\left[\frac{2}{3}, 1\right]\right]. \tag{8.3.35}$$

As F is closed $in S,$ so are A and B by Lemma 2. (Why?)
As $B \cap A=\emptyset$, Lemma 3 yields a continuous map $g_{1}: S \rightarrow\left[0, \frac{1}{3}\right]$, with $g_{1}=0$ on A, and $g_{1}=\frac{1}{3}$ on B. Set $f_{1}=f-g_{1}$ on F; so $\left|f_{1}\right| \leq \frac{2}{3}$, and f_{1} is continuous on F.
Applying the same steps to f_{1} (with suitable sets $A_{1}, B_{1} \subseteq F$), find a continuous map g_{2}, with $0 \leq g_{2} \leq \frac{2}{3} \cdot \frac{1}{3}$ on S. Then $f_{2}=f_{1}-g_{2}$ is continuous, and $0 \leq f_{2} \leq\left(\frac{2}{3}\right)^{2}$ on F.
Continuing, obtain two sequences $\left\{g_{n}\right\}$ and $\left\{f_{n}\right\}$ of real functions such that each g_{n} is continuous on S,

$$0 \leq g_{n} \leq \frac{1}{3}\left(\frac{2}{3}\right)^{n-1}, \tag{8.3.36}$$

and $f_{n}=f_{n-1}-g_{n}$ is defined and continuous on F, with

$$0 \leq f_{n} \leq\left(\frac{2}{3}\right)^{n} \tag{8.3.37}$$

there $\left(f_{0}=f\right)$.
We claim that

$$g=\sum_{n=1}^{\infty} g_{n} \tag{8.3.38}$$

is the desired map.
Indeed, the series converges uniformly on S (Theorem 3 of Chapter 4, §12).
As all g_{n} are continuous, so is g (Theorem 2 in Chapter 4, §12). Also,

$$\left|f-\sum_{k=1}^{n} g_{k}\right| \leq\left(\frac{2}{3}\right)^{n} \rightarrow 0 \tag{8.3.39}$$

on $F($ why? $);$ so $f=g$ on F. Moreover,

$$0 \leq g_1 \leq g \leq \sum_{n=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^n=1 \text{ on } S. \tag{8.3.40}$$

Hence $\inf g[S]=0$ and $\sup g[S]=1$, as required.
Now assume

$$\inf f[F]=a<\sup f[F]=b \quad \left(a, b \in E^1\right) \tag{8.3.41}$$

Set

$$h(x)=\frac{f(x)-a}{b-a} \tag{8.3.42}$$

so that $\inf h[F]=0$ and $\sup h[F]=1.$ (Why?)
As shown above, there is a continuous map g_0 on S, with

$$g_0=h=\frac{f-a}{b-a} \tag{8.3.43}$$

on F, $\inf g_0[S]=0$, and $\sup g_0[S]=1.$ Set

$$a+(b-a) g_0=g. \tag{8.3.44}$$

Then g is the required function. (Verify!)
Finally, if $a, b \in E^*(a<b),$ all reduces to the bounded case by considering $H(x)=\arctan f(x)$. □

*Theorem 8.3.3

(Fréchet). Let $m: \mathcal{M} \rightarrow E^*$ be a strongly regular measure in (S, ρ). If $f: S \rightarrow E^*\left(E^n, C^n\right)$ is m -measurable on A, then

$$f=\lim_{i \rightarrow \infty} f_i(a \cdot e.(m)) \text{ on } A \tag{8.3.45}$$

for some sequence of maps f_i continuous on S. (We assume E^* to be metrized as in Lemma 4.)

Proof

We consider $f: S \rightarrow E^*$ (the other cases reduce to E^1 via components).
Taking $\varepsilon=\frac{1}{i}(i=1,2, \ldots)$ in Theorem 2, we obtain for each i a closed $\mathcal{M}$-set $F_i \subseteq A$ such that

$$m\left(A-F_i\right)<\frac{1}{i} \tag{8.3.46}$$

and f is relatively continuous on each F_i. We may assume that $F_i \subseteq F_{i+1}$ (if not, replace F_i by $\bigcup_{k=1}^i F_k$).
Now, Lemma 4 yields for each i a continuous map $f_i: S \rightarrow E^*$ such that $f_i=f$ on F_i. We complete the proof by showing that $f_i \rightarrow f$ (pointwise) on the set

$$B=\bigcup_{i=1}^{\infty} F_i \tag{8.3.47}$$

and that $m(A-B)=0$.
Indeed, fix any $x \in B$. Then $x \in F_i$ for some $i=i_0$, hence also for $i>i_0$ (since $\{F_i\} \uparrow$). As $f_i=f$ on F_i, we have

$$(\forall i>i_0) \quad f_i(x)=f(x), \tag{8.3.48}$$

and so $f_i(x) \rightarrow f(x)$ for $x \in B$. As $F_i \subseteq B$, we get

$$m(A-B) \leq m\left(A-F_i\right)<\frac{1}{i} \tag{8.3.49}$$

for all i. Hence $m(A-B)=0$, and all is proved. □

8.3.E: Problems on Measurable Functions in (S,M,m) (S,M,m)

Exercise $8.3.E.1$

Fill in all proof details in Corollaries 1 to 4.

Exercise $8.3.E.1'$

Verify Notes 3 and 4.

Exercise $8.3.E.2$

Prove Theorems 1 and 2 in §1 and Theorem 2 in §2, for almost measurable functions.

Exercise $8.3.E.3$

Prove Note 2.
[Hint: If $f: S \rightarrow E^{*}$ is $\mathcal{M}$-measurable on $B=A-Q(m Q=0, Q \subseteq A)$, then $A=B \cup Q$ and

$$(\forall a \in E^{*}) \quad A(f>a)=B(f>a) \cup Q(f>a). \tag{8.3.E.1}$$

Here $B(f>a) \in \mathcal{M}$ by Theorem 1 in §2, and $Q(f>a) \in \mathcal{M}$ if m is complete. For $f: S \rightarrow E^{n}\left(C^{n}\right)$, use Theorem 2 of $§1$.]

Exercise $8.3.E.4$

*4. Show that if m is complete and $f: S \rightarrow\left(T, \rho^{\prime}\right)$ is m-measurable on A with $f[A]$ separable in T, then f is $\mathcal{M}$-measurable on A.
[Hint: Use Problem 13 in $§2$.]

Exercise $8.3.E.5$

*5. Prove Theorem 1 for $f: S \rightarrow\left(T, \rho^{\prime}\right)$, assuming that $f[A]$ is separable in T.

Exercise $8.3.E.6$

Given $f_{n} \rightarrow f($ a.e. $)$ on A, prove that $f_{n} \rightarrow g($ a.e. $)$ on A iff $f=g($ a.e. $)$ on A.

Exercise $8.3.E.7$

Given $A \in \mathcal{M}$ in $(S, \mathcal{M}, m)$, let m_{A} be the restriction of m to

$$\mathcal{M}_{A}=\{X \in \mathcal{M} | X \subseteq A\}. \tag{8.3.E.2}$$

Prove that
(i) $\left(A, \mathcal{M}_{A}, m_{A}\right)$ is a measure space (called a subspace of $(S, \mathcal{M}, m)$);
(ii) if m is complete, topological, σ-finite or (strongly) regular, so is m_{A}.

Exercise $8.3.E.8$

(i) Show that if $D \subseteq K \subseteq (T, \rho')$, then the closure of D in the subspace (K, ρ') is $K \cap \bar{D}$, where $\bar{D}$ is the closure of D in (T, ρ').
[Hint: Use Problem 11 in Chapter $3, \S 16.$]
(ii) Prove that if $B \subseteq K$ and if B is separable in (T, ρ'), it is so in (K, ρ').
[Hint: Use Problem 7 from $\xi 1$.]

Exercise $8.3.E.9$

*9. Fill in all proof details in Lemma 4.

Exercise $8.3.E.10$

Simplify the proof of Theorem 2 for the case $mA < \infty$.
[Outline: (i) First, let f be elementary, with $f = a_i$ on $A_i \in \mathcal{M}, A = \cup_i A_i$ (disjoint), $\sum mA_i = mA < \infty$.
Given $\varepsilon > 0$

$$(\exists n) \quad mA - \sum_{i=1}^{n} mA_i < \frac{1}{2}\varepsilon. \tag{8.3.E.3}$$

Each A_i has a closed subset $F_i \in \mathcal{M}$ with $m(A_i - F_i) < \varepsilon / 2n$. (Why?) Now use Problem 17 in Chapter 4, §8, and set $F = \bigcup_{i=1}^{n} F_i$.
(ii) If f is $\mathcal{M}$-measurable on $H = A - Q, mQ = 0$, then by Theorem 3 in $\xi 1$,
$f_n \rightarrow f$ (uniformly) on H for some elementary maps f_n. By (i), each f_n is relatively continuous on a closed $\mathcal{M}$-set $F_n \subseteq H$, with $mH - mF_n < \varepsilon / 2^n$; so all f_n are relatively continuous on $F = \bigcap_{n=1}^{\infty} F_n$. Show that F is the required set.

Exercise $8.3.E.11$

Given $f_n : S \rightarrow (T, \rho'), n = 1, 2, \ldots,$ we say that
(i) $f_n \rightarrow f$ almost uniformly on $A \subseteq S$ iff

$$(\forall \delta > 0)(\exists D \in \mathcal{M} | mD < \delta) \quad f_n \rightarrow f (\text{uniformly}) \text{ on } A - D; \tag{8.3.E.4}$$

(ii) $f_n \rightarrow f$ in measure on A iff

$$\begin{aligned} (\forall \delta, \sigma > 0)(\exists k)(\forall n > k) \, (\exists D_n \in \mathcal{M} | mD_n < \delta) \\ \rho'(f, f_n) < \sigma \text{ on } A - D_n. \end{aligned}$$

Prove the following.
(a) $f_n \rightarrow f$ (uniformly) implies $f_n \rightarrow f$ (almost uniformly), and the latter implies both $f_n \rightarrow f(\text{ in measure})$ and $f_n \rightarrow f(a.e.)$.
(b) Given $f_n \rightarrow f$ (almost uniformly), we have $f_n \rightarrow g$ (almost uniformly) iff $f = g(\text{ a.e. })$; similarly for convergence in measure.
(c) If f and f_n are $\mathcal{M}$-measurable on A, then $f_n \rightarrow f$ in measure on A iff

$$(\forall \sigma > 0) \quad \lim_{n \rightarrow \infty} mA(\rho'(f, f_n) \geq \sigma) = 0. \tag{8.3.E.5}$$

Exercise $8.3.E.12$

Assuming that $f_n : S \rightarrow (T, \rho')$ is m -measurable on A for $n = 1, 2, \ldots$, that $mA < \infty$, and that $f_n \rightarrow f(a.e.)$ on A, prove the following.

(i) Lebesgue's theorem: $f_n \rightarrow f$ (in measure) on A (see Problem 11).

(ii) Egorov's theorem: $f_n \rightarrow f$ (almost uniformly) on A.

[Outline: (i) f_n and f are $\mathcal{M}$ -measurable on $H = A - Q, mQ = 0$ (Corollary 1), with $f_n \rightarrow f$ (pointwise) on H. For all i, k, set

$$H_i(k) = \bigcap_{n=i}^{\infty} H\left(\rho'(f_n, f) < \frac{1}{k}\right) \in \mathcal{M} \tag{8.3.E.6}$$

by Problem 6 in §1. Show that $(\forall k) H_i(k) \nearrow H$; hence

$$\lim_{i \rightarrow \infty} mH_i(k) = mH = mA < \infty; \tag{8.3.E.7}$$

so

$$(\forall \delta > 0)(\forall k)(\exists i_k) \quad m(A - H_{i_k}(k)) < \frac{\delta}{2^k}, \tag{8.3.E.8}$$

proving (i), since

$$(\forall n > i_k) \quad \rho'(f_n, f) < \frac{1}{k} \text{ on } H_{i_k}(k) = A - (A - H_{i_k}(k)). \tag{8.3.E.9}$$

(ii) Continuing, set $(\forall k) D_k = H_{i_k}(k)$ and

$$D = A - \bigcap_{k=1}^{\infty} D_k = \bigcup_{k=1}^{\infty}(A - D_k). \tag{8.3.E.10}$$

Deduce that $D \in \mathcal{M}$ and

$$mD \leq \sum_{k=1}^{\infty} m(A - H_{i_k}(k)) < \sum_{k=1}^{\infty} \frac{\delta}{2^k} = \delta. \tag{8.3.E.11}$$

Now, from the definition of the $H_i(k)$, show that $f_n \rightarrow f$ (uniformly) on $A - D$, proving (ii).]

Exercise $8.3.E.13$

Disprove the converse to Problem 12(i).

[Outline: Assume that $A = [0, 1)$; for all $0 \leq k$ and all $0 \leq i < 2^k$, set

$$g_{ik}(x) = \begin{cases} 1 & \text{if } \frac{i-1}{2^k} \leq x < \frac{i}{2^k} \\ 0 & \text{otherwise} \end{cases} \tag{8.3.E.12}$$

Put the g_{ik} in a single sequence by

$$f_{2^k + i} = g_{ik}. \tag{8.3.E.13}$$

Show that $f_n \rightarrow 0$ in L measure on A, yet for no $x \in A$ does $f_n(x)$ converge as $n \rightarrow \infty$.]

Exercise $8.3.E.14$

Prove that if $f : S \rightarrow (T, \rho')$ is m -measurable on A and $g : T \rightarrow (U, \rho'')$ is relatively continuous on $f[A]$, then $g \circ f : S \rightarrow (U, \rho'')$ is m-measurable on A.
[Hint: Use Corollary 4 in §1.]

8.4: Integration of Elementary Functions

In Chapter 5, integration was treated as antidifferentiation. Now we adopt another, measure-theoretical approach.

Lebesgue's original theory was based on Lebesgue measure (Chapter 7, §8). The more general modern treatment develops the integral for functions $f : S \rightarrow E$ in an arbitrary measure space. Henceforth, $(S, \mathcal{M}, m)$ is fixed, and the range space E is E^1, E^*, C, E^n, or another complete normed space. Recall that in such a space, $\sum_i |a_i| < \infty$ implies that $\sum a_i$ converges and is permutable (Chapter 7, §2).

We start with elementary maps, including simple maps as a special case.

Definition

Let $f : S \rightarrow E$ be elementary on $A \in \mathcal{M}$; so $f = a_i$ on A_i for some $\mathcal{M}$-partition

$$A = \bigcup_i A_i \text{ (disjoint).} \tag{8.4.1}$$

(Note that there may be many such partitions.)

We say that f is integrable (with respect to m), or m-integrable, on A iff

$$\sum |a_i| \, mA_i < \infty. \tag{8.4.2}$$

(The notation "$|a_i| mA_i$" always makes sense by our conventions (2*) in Chapter 4, §4.) If m is Lebesgue measure, then we say that f is Lebesgue integrable, or L-integrable.

We then define $\int_A f$, the m-integral of f on A, by

$$\int_A f = \int_A f dm = \sum_i a_i m A_i. \tag{8.4.3}$$

(The notation "dm" is used to specify the measure m.)

The "classical" notation for $\int_A f dm$ is $\int_A f(x) dm(x)$.

Note 1. The assumption

$$\sum |a_i| \, mA_i < \infty \tag{8.4.4}$$

implies

$$(\forall i) \quad |a_i| \, mA_i < \infty; \tag{8.4.5}$$

so $a_i = 0$ if $mA_i = \infty$, and $mA_i = 0$ if $|a_i| = \infty$. Thus by our conventions, all "bad" terms $a_i m A_i$ vanish. Hence the sum in (1) makes sense and is finite.

Note 2. This sum is also independent of the particular choice of $\{A_i\}$. For if $\{B_k\}$ is another $\mathcal{M}$-partition of A, with $f = b_k$ on B_k, say, then $f = a_i = b_k$ on $A_i \cap B_k$ whenever $A_i \cap B_k \neq \emptyset$. Also,

$$(\forall i) \quad A_i = \bigcup_k (A_i \cap B_k) \text{ (disjoint);} \tag{8.4.6}$$

so

$$(\forall i) \quad a_i m A_i = \sum_k a_i m(A_i \cap B_k), \tag{8.4.7}$$

and hence (see Theorem 2 of Chapter 7, §2, and Problem 11 there)

$$\sum_i a_i m A_i = \sum_i \sum_k a_i m (A_i \cap B_k) = \sum_k \sum_i b_k m (A_i \cap B_k) = \sum_k b_k m B_k. \tag{8.4.8}$$

(Explain!)

This makes our definition (1) unambiguous and allows us to choose any $\mathcal{M}$-partition $\{A_i\}$, with f constant on each A_i, when forming integrals (1).

Corollary 8.4.1

Let $f: S \rightarrow E$ be elementary and integrable on $A \in \mathcal{M}$. Then the following statements are true.

(i) $|f|<\infty$ a.e. on A.

(ii) f and $|f|$ are elementary and integrable on any $\mathcal{M}$-set $B \subseteq A$, and

$$\left|\int_B f\right| \leq \int_B|f| \leq \int_A|f|. \tag{8.4.9}$$

(iii) The set $B=A(f \neq 0)$ is σ-finite (Definition 4 in Chapter 7, §5), and

$$\int_A f=\int_B f. \tag{8.4.10}$$

(iv) If $f=a$ (constant) on A,

$$\int_A f=a \cdot m A. \tag{8.4.11}$$

(v) $\int_A|f|=0$ iff $f=0$ a.e. on A.

(vi) If $m Q=0$, then

$$\int_A f=\int_{A-Q} f \tag{8.4.12}$$

(so we may neglect sets of measure 0 in integrals).

(vii) For any k in the scalar field of E, $k f$ is elementary and integrable, and

$$\int_A k f=k \int_A f. \tag{8.4.13}$$

Note that if f is scalar valued, k may be a vector. If $E=E^*$, we assume $k \in E^1$.

Proof

(i) By Note 1, $|f|=\left|a_i\right|=\infty$ only on those A_i with $m A_i=0$. Let Q be the union of all such A_i. Then $m Q=0$ and $|f|<\infty$ on $A-Q$, proving (i).

(ii) If $\{A_i\}$ is an $\mathcal{M}$-partition of A, $\{B \cap A_i\}$ is one for B. (Verify!) We have $f=a_i$ and $|f|=\left|a_i\right|$ on $B \cap A_i \subseteq A_i$.

Also,

$$\sum\left|a_i\right| m\left(B \cap A_i\right) \leq \sum\left|a_i\right| m A_i<\infty. \tag{8.4.14}$$

(Why?) Thus f and $|f|$ are elementary and integrable on B, and (ii) easily follows by formula (1).

(iii) By Note 1, $f=0$ on A_i if $m A_i=\infty$. Thus $f \neq 0$ on A_i only if $m A_i<\infty$. Let $\{A_{i_k}\}$ be the subsequence of those A_i on which $f \neq 0$; so

$$(\forall k) \quad m A_{i_k}<\infty. \tag{8.4.15}$$

Also,

$$B=A(f \neq 0)=\bigcup_k A_{i_k} \in \mathcal{M} \text{ (σ-finite!)}. \tag{8.4.16}$$

By (ii), f is elementary and integrable on B. Also,

$$\int_B f=\sum_k a_{i_k} m A_{i_k}, \tag{8.4.17}$$

while

$$\int_A f=\sum_i a_i m A_i. \tag{8.4.18}$$

These sums differ only by terms with $a_i=0$. Thus (iii) follows.

The proof of (iv)-(vii) is left to the reader. □

Note 3. If $f: S \rightarrow E^*$ is elementary and sign-constant on A, we also allow that

$$\int_A f=\sum_i a_i m A_i= \pm \infty. \tag{8.4.19}$$

Thus here $\int_A f$ exists even if f is not integrable. Apart from claims of integrability and σ-finiteness, Corollary 1(ii)-(vii) hold for such f, with the same proofs.

✔ Example

Let m be Lebesgue measure in E^1. Define $f=1$ on R (rationals) and $f=0$ on E^1-R; see Chapter 4, §1, Example (c). Let $A=[0,1]$.

By Corollary 1 in Chapter 7, §8, $A \cap R \in \mathcal{M}^*$ and $m(A \cap R)=0$. Also, $A-R \in \mathcal{M}^*$.

Thus $\{A \cap R, A-R\}$ is an $\mathcal{M}^*$-partition of A, with $f=1$ on $A \cap R$ and $f=0$ on $A-R$.

Hence f is elementary and integrable on A, and

$$\int_A f=1 \cdot m(A \cap R)+0 \cdot m(A-R)=0. \tag{8.4.20}$$

Thus f is L-integrable (even though it is nowhere continuous).

Theorem 8.4.1 (additivity)

(i) If $f: S \rightarrow E$ is elementary and integrable or elementary and nonnegative on $A \in \mathcal{M}$, then

$$\int_A f=\sum_k \int_{B_k} f \tag{8.4.21}$$

for any $\mathcal{M}$-partition $\{B_k\}$ of A.

(ii) If f is elementary and integrable on each set B_k of a finite $\mathcal{M}$-partition

$$A=\bigcup_k B_k, \tag{8.4.22}$$

it is elementary and integrable on all of A, and (2) holds again.

Proof

(i) If f is elementary and integrable or elementary and nonnegative on $A=\bigcup_k B_k$, it is surely so on each B_k by Corollary 2 of §1 and Corollary 1(ii) above.

Thus for each k, we can fix an $\mathcal{M}$-partition $B_k=\bigcup_i A_{ki}$, with f constant $(f=a_{ki})$ on $A_{ki}, i=1,2, \ldots$. Then

$$A=\bigcup_k B_k=\bigcup_k \bigcup_i A_{ki} \tag{8.4.23}$$

is an $\mathcal{M}$-partition of A into the disjoint sets $A_{ki} \in \mathcal{M}$.

Now, by definition,

$$\int_{B_k} f=\sum_i a_{ki} m A_{ki} \tag{8.4.24}$$

and

$$\int_A f=\sum_{k,i} a_{ki} m A_{ki}=\sum_k\left(\sum_i a_{ki} m A_{ki}\right)=\sum_k \int_{B_k} f \tag{8.4.25}$$

by rules for double series. This proves formula (2).

(ii) If f is elementary and integrable on $B_k (k=1, \ldots, n),$ then with the same notation, we have

$$\sum_i\left|a_{ki}\right| m A_{ki}<\infty \tag{8.4.26}$$

(by integrability); hence

$$\sum_{k=1}^n \sum_i\left|a_{ki}\right| m A_{ki}<\infty. \tag{8.4.27}$$

This means, however, that f is elementary and integrable on A, and so clause (ii) follows. □

Caution. Clause (ii) fails if the partition $\{B_k\}$ is infinite.

Theorem 8.4.2

(i) If $f, g: S \rightarrow E^*$ are elementary and nonnegative on A, then

$$\int_A(f+g)=\int_A f+\int_A g. \tag{8.4.28}$$

(ii) If $f, g: S \rightarrow E$ are elementary and integrable on A, so is $f \pm g,$ and

$$\int_A(f \pm g)=\int_A f \pm \int_A g. \tag{8.4.29}$$

Proof

Arguing as in the proof of Theorem 1 of §1, we can make f and g constant on sets of one and the same $\mathcal{M}$-partition of A, say, $f=a_i$ and $g=b_i$ on $A_i \in \mathcal{M};$ so

$$f \pm g=a_i \pm b_i \text{ on } A_i, \quad i=1,2, \ldots . \tag{8.4.30}$$

In case (i), $f, g \geq 0$; so integrability is irrelevant by Note 3, and formula (1) yields

$$\int_A(f+g)=\sum_i\left(a_i+b_i\right) m A_i=\sum_i a_i m A_i+\sum b_i m A_i=\int_A f+\int_A g. \tag{8.4.31}$$

In (ii), we similarly obtain

$$\sum_i\left|a_i \pm b_i\right| m A_i \leq \sum\left|a_i\right| m A_i+\sum_i\left|b_i\right| m A_i<\infty. \tag{8.4.32}$$

(Why?) Thus $f \pm g$ is elementary and integrable on A. As before, we also get

$$\int_A(f \pm g)=\int_A f \pm \int_A g, \tag{8.4.33}$$

simply by rules for addition of convergent series. (Verify!) □

Note 4. As we know, the characteristic function C_B of a set $B \subseteq S$ is defined

$$C_B(x)=\begin{cases}1, & x\in B,\\ 0, & x\in S-B.\end{cases} \tag{8.4.34}$$

If $g:S\to E$ is elementary on A, so that

$$g=a_i \text{ on } A_i, 1, 2, \ldots, \tag{8.4.35}$$

for some $\mathcal{M}$-partition

$$A=\bigcup A_i, \tag{8.4.36}$$

then

$$g=\sum_i a_i C_{A_i} \text{ on } A. \tag{8.4.37}$$

(This sum always exists for disjoint sets A_i. Why?) We shall often use this notation.

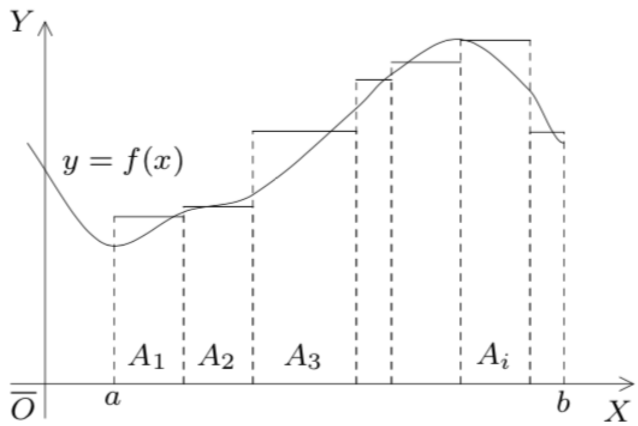

FIGURE 33

If m is Lebesgue measure in E^1, the integral

$$\int_A g=\sum_i a_i m A_i \tag{8.4.38}$$

has a simple geometric interpretation; see Figure 33. Let $A=[a,b]\subset E^1$; let g be bounded and nonnegative on E^1. Each product $a_i m A_i$ is the area of a rectangle with base A_i and altitude a_i. (We assume the A_i to be intervals here.) The total area,

$$\int_A g=\sum_i a_i m A_i, \tag{8.4.39}$$

can be treated as an approximation to the area under some curve $y=f(x)$, where f is approximated by g (Theorem 3 in §1). Integration historically arose from such approximations.

Integration of elementary extended-real functions. Note 3 can be extended to sign-changing functions as follows.

Definition

If

$$f=\sum_i a_i C_{A_i} \quad (a_i\in E^*) \tag{8.4.40}$$

on

$$A=\bigcup_i A_i \quad (A_i\in\mathcal{M}), \tag{8.4.41}$$

we set

$$\int_A f=\int_A f^+-\int_A f^-, \tag{8.4.42}$$

with

$$f^+=f\vee 0\geq 0 \text{ and } f^-=(-f)\vee 0\geq 0; \tag{8.4.43}$$

see §2.

By Theorem 2 in §2, f^+ and f^- are elementary and nonnegative on A; so

$$\int_A f^+ \text{ and } \int_A f^- \tag{8.4.44}$$

are defined by Note 3, and so is

$$\int_A f=\int_A f^+-\int_A f^- \tag{8.4.45}$$

by our conventions (2*) in Chapter 4, §4.

We shall have use for formula (3), even if

$$\int_A f^+=\int_A f^-=\infty; \tag{8.4.46}$$

then we say that $\int_A f$ is unorthodox and equate it to $+\infty$, by convention; cf. Chapter 4, §4. (Other integrals are called orthodox.) Thus for elementary and (extended) real functions, $\int_A f$ is always defined. (We further develop this idea in §5.)

Note 5. With f as above, we clearly have

$$f^+=a_i^+ \text{ and } f^-=a_i^- \text{ on } A_i, \tag{8.4.47}$$

where

$$a_i^+=\max\left(a_i, 0\right) \text{ and } a_i^-=\max\left(-a_i, 0\right). \tag{8.4.48}$$

Thus

$$\int_A f^+=\sum a_i^+\cdot mA_i \text{ and } \int_A f^-=\sum a_i^-\cdot mA_i, \tag{8.4.49}$$

so that

$$\int_A f=\int_A f^+-\int_A f^-=\sum_i a_i^+\cdot mA_i-\sum_i a_i^-\cdot mA_i. \tag{8.4.50}$$

If $\int_A f^+<\infty$ or $\int_A f^-<\infty,$ we can subtract the two series termwise (Problem 14 of Chapter 4, §13) to obtain

$$\int_A f=\sum_i\left(a_i^+-a_i^-\right)mA_i=\sum_i a_i mA_i \tag{8.4.51}$$

for $a_i^+-a_i^-=a_i.$ Thus formulas (3) and (4) agree with our previous definitions.

8.4.E: Problems on Integration of Elementary Functions

8.4.E: Problems on Integration of Elementary Functions

Exercise $8.4.E.1$

Verify Note 2.

Exercise $8.4.E.1'$

Prove Corollary 1(iv) $-$ (vii) .

Exercise $8.4.E.2$

Prove that $\int_A f=0$ if $mA=0$ or $f=0$ on A. Disprove the converse by examples.

Exercise $8.4.E.3$

Find a primitive F for $f=C_R$ in our example. Show that

$$\int_{[0,1]} f dm=F(1)-F(0). \tag{8.4.E.1}$$

Exercise $8.4.E.4$

Fill in the proof details in Theorem 2.
[Hint: Use comparison test for series.]

Exercise $8.4.E.5$

$\Rightarrow 5$. Show that if f and g are elementary and nonnegative with $f\geq g$ on A, then

$$\int_A f\geq \int_A g\geq 0. \tag{8.4.E.2}$$

[Hint: As in Theorem 2, let

$$f=\sum_i a_i C_{A_i} \text{ and } g=\sum_i b_i C_{A_i}. \tag{8.4.E.3}$$

Then $f\geq g\geq 0$ implies $a_i\geq b_i\geq 0.$]

Exercise $8.4.E.6$

$\Rightarrow 6$. Prove that if f and g are elementary and (extended) real on A, then

$$\int_A (f\pm g)=\int_A f\pm \int_A g, \tag{8.4.E.4}$$

provided
(i) $\int_A f$ or $\int_A g$ is finite, or
(ii) $\int_A f$, $f_A g$, and $\int_A f\pm \int_A g$ are all orthodox.
[Outline: As in Theorem 2, let

$$f=\sum_i a_i C_{A_i} \text{ and } g=\sum_i b_i C_{A_i}, \tag{8.4.E.5}$$

so

$$f \pm g=a_i \pm b_i \text{ on } A_i. \tag{8.4.E.6}$$

Now, if

$$\left|\int_A f\right|<\infty, \tag{8.4.E.7}$$

then by Problem 14 in Chapter 4, §13, and formula $(4), \sum a_i m A_i$ converges absolutely; so its termwise addition to any other series does not affect the absolute convergence or divergence of the latter, i.e., the finiteness or infiniteness of its positive and negative parts. For example,

$$\sum_i\left(a_i \pm b_i\right)^{+} m A_i=\infty \tag{8.4.E.8}$$

iff

$$\sum b_i^{+} m A_i=\infty. \tag{8.4.E.9}$$

Thus if

$$\int_A g=\pm \infty, \tag{8.4.E.10}$$

then

$$\int_A(f \pm g)=\int_A g=\pm \infty=\int_A f \pm \int_A g. \tag{8.4.E.11}$$

If both

$$\int_A f, \int_A g \neq \pm \infty, \tag{8.4.E.12}$$

Theorem $2($ ii) applies. In the orthodox infinite case, a similar proof works on noting that either the positive or the negative parts of both series are finite if

$$\int_A f \pm \int_A g \tag{8.4.E.13}$$

is orthodox, too. (Verify!)]

Exercise $8.4.E.7$

Show that if f is elementary and nonnegative on A and

$$\int_A f>p \in E^*, \tag{8.4.E.14}$$

then there is an elementary and nonnegative map g on A such that

$$\int_A f \geq \int_A g > p, \tag{8.4.E.15}$$

$g = 0$ on $A(f = 0)$, and

$$f > g \text{ on } A - A(f = 0). \tag{8.4.E.16}$$

[Hints: Let

$$B = A(f = \infty) \tag{8.4.E.17}$$

and

$$C = A - B; \tag{8.4.E.18}$$

so g_n is elementary and nonnegative on A and

$$g_n = n \text{ on } B \tag{8.4.E.19}$$

and

$$g_n = \left(1 - \frac{1}{n}\right) f \text{ on } C; \tag{8.4.E.20}$$

so g_n is elementary and nonnegative on A and

$$f > g_n \text{ on } A - A(f = 0). \text{ (Why?)} \tag{8.4.E.21}$$

By Theorem 1 and Corollary 1(iv)(vii),

$$\int_A g_n = \int_B g_n + \int_C g_n = \int_B (n) + \int_C \left(1 - \frac{1}{n}\right) f = n \cdot mB + \left(1 - \frac{1}{n}\right) \int_C f. \tag{8.4.E.22}$$

Deduce that

$$\lim_{n \to \infty} \int_A g_n = \int_B f + \int_C f = \int_A f > p; \tag{8.4.E.23}$$

so

$$(\exists n) \quad \int_A g_n > p. \tag{8.4.E.24}$$

Take $g = g_n$ for that n.]

Exercise $8.4.E.8$

Show that if $E = E^*$, Theorem $1(\text{ i })$ holds also if $\int_A f$ is infinite but orthodox.

Exercise $8.4.E.9$

(i) Prove that if f is elementary and integrable on A, so is $-f$, and

$$\int_A(-f)=-\int_A f. \tag{8.4.E.25}$$

(ii) Show that this holds also if f is elementary and (extended) real and $\int_A f$ is orthodox.

8.5: Integration of Extended-Real Functions

We shall now define integrals for arbitrary functions $f: S \rightarrow E^{*}$ in a measure space $(S, \mathcal{M}, m)$. We start with the case $f \geq 0$.

Definition

Given $f \geq 0$ on $A \in \mathcal{M},$ we define the upper and lower integrals,

$$\overline{\int} \text{ and } \underline{\int}, \tag{8.5.1}$$

of f on A (with respect to m) by

$$\overline{\int_{A}} f=\overline{\int_{A}} f d m=\inf _{h} \int_{A} h \tag{8.5.2}$$

over all elementary maps $h \geq f$ on A, and

$$\int_{-A} f=\int_{-A} f d m=\sup _{g} \int_{A} g \tag{8.5.3}$$

over all elementary and nonnegative maps $g \leq f$ on A.
If f is not nonnegative, we use $f^{+}=f \vee 0$ and $f^{-}=(-f) \vee 0$ (§2), and set
\[

$$\begin{aligned} \int_{A} f &= \overline{\int_{A}} f d m=\overline{\int_{A}} f^{+}-\underline{\int_{A}} f^{-} \text{ and} \\ \underline{\int_{A}} f &= \underline{\int_{A}} f d m=\underline{\int_{A}} f^{+}-\overline{\int_{A}} f^{-}. \end{aligned} \tag{8.5.4}$$

\]
By our conventions, these expressions are always defined. The integral $\overline{\int}_{A} f\left(\text{ or } \int_{-A} f\right)$ is called orthodox iff it does not have the form $\infty-\infty$ in (1), e.g., $\overline{\text{ if }} f \geq 0$ (i.e., $f^{-}=0$), or if $\int_{A} f<\infty$. An unorthodox integral equals $+\infty$.
We often write $\int$ for $\overline{\int}$ and call it simply the integral (of f), even if

$$\overline{\int_{A}} f \neq \int_{A} f. \tag{8.5.5}$$

"Classical" notation is $\int_{A} f(x) d m(x)$.

Definition

The function f is called integrable (or m-integrable, or Lebesgue integrable, with respect to m) on A, iff

$$\overline{\int_{A}} f d m=\int_{A} f d m \neq \pm \infty \tag{8.5.6}$$

The process described above is called (abstract) Lebesgue integration as opposed to Riemann integration (B. Riemann, 1826-1866). The latter deals with bounded functions only and allows h and g in $\left(1^{\prime}\right)$ and $\left(1^{\prime \prime}\right)$ to be simple step functions only (see §9). It is inferior to Lebesgue theory.
The values of

$$\overline{\int_A} f dm \text{ and } \underline{\int_A} f dm \tag{8.5.7}$$

depend on m. If m is Lebesgue measure, we speak of Lebesgue integrals, in the stricter sense. If m is Lebesgue-Stieltjes measure, we speak of LS-integrals, and so on.
Note 1. If f is elementary and (extended) real, our present definition of

$$\overline{\int_A} f \tag{8.5.8}$$

agrees with that of §4. For if $f \geq 0$, f itself is the least of all elementary and nonnegative functions

$$h \geq f \tag{8.5.9}$$

and the greatest of all elementary and nonnegative functions

$$g \leq f. \tag{8.5.10}$$

Thus by Problem 5 in §4,

$$\int_A f = \min_{h \geq f} \int_A h = \max_{g \leq f} \int_A g, \tag{8.5.11}$$

i.e.,

$$\int_A f = \overline{\int}_A f = \underline{\int_A} f. \tag{8.5.12}$$

If, however, $f \ngeq 0$, this follows by Definition 2 in §4. This also shows that for elementary and (extended) real maps,

$$\overline{\int_A} f = \underline{\int_A} f \text{ always.} \tag{8.5.13}$$

(See also Theorem 3.)
Note 2. By Definition 1,

$$\underline{\int_A} f \leq \overline{\int_A} f \text{ always.} \tag{8.5.14}$$

For if $f \geq 0$, then for any elementary and nonnegative maps g, h with

$$g \leq f \leq h, \tag{8.5.15}$$

we have

$$\int_A g \leq \int_A h \tag{8.5.16}$$

by Problem 5 in §4. Thus

$$\underline{\int}_A f=\sup_g \int_A g \tag{8.5.17}$$

is a lower bound of all such $\int_A h$, and so

$$\underline{\int}_A f \leq \operatorname{glb} \int_A h=\overline{\int}_A f. \tag{8.5.18}$$

In the general formula (1), too

$$\underline{\int}_A f \leq \overline{\int}_A f, \tag{8.5.19}$$

since

$$\int_{-A} f^{+} \leq \overline{\int}_A f^{+} \text{ and } \int_{-A} f^{-} \leq \bar{J}_A f^{-}. \tag{8.5.20}$$

Theorem 8.5.1

For any functions $f, g: S \rightarrow E^*$ and any set $A \in \mathcal{M},$ we have the following results.
(a) If $f=a(\text{constant})$ on A, then
$\overline{\int}_A f=\underline{\int}_A f=a \cdot m A$.
(b) If $f=0$ on A or $m A=0$, then

$$\overline{\int}_A f=\underline{\int}_A f=0. \tag{8.5.21}$$

(c) If $f \geq g$ on A, then

$$\overline{\int}_A f \geq \overline{\int}_A g \text{ and } \underline{\int}_A f \geq \underline{\int}_A g. \tag{8.5.22}$$

(d) If $f \geq 0$ on A, then

$$\overline{\int}_A f \geq 0 \text{ and } \underline{\int}_A f \geq 0. \tag{8.5.23}$$

Similarly if $f \leq 0$ on A.
(e) If $0 \leq p<\infty,$ then

$$\overline{\int}_A p f=p \overline{\int}_A f \text{ and } \underline{\int}_A p f=p \underline{\int}_A f. \tag{8.5.24}$$

(e') We have

$$\overline{\int}_A(-f)=-\underline{\int}_A f \text{ and } \underline{\int}_A(-f)=-\overline{\int}_A f \tag{8.5.25}$$

if one of the integrals involved in each case is orthodox. Otherwise,

$$\overline{\int}_A(-f)=\infty=\underline{\int}_A f \text{ and } \underline{\int}_A(-f)=\infty=\overline{\int}_A f. \tag{8.5.26}$$

(f) If $f \geq 0$ on A and

$$A \supseteq B, B \in \mathcal{M}, \tag{8.5.27}$$

then

$$\overline{\int}_A f \geq \overline{\int}_B f \text{ and } \underline{\int}_A f \geq \underline{\int}_B f. \tag{8.5.28}$$

(g) We have

$$\left|\overline{\int}_A f\right| \leq \overline{\int}_A |f| \text{ and } \left|\underline{\int}_A f\right| \leq \overline{\int}_A |f| \tag{8.5.29}$$

(but not

$$\left|\underline{\int}_A f\right| \leq \underline{\int}_A |f| \tag{8.5.30}$$

in general).

(h) If $f \geq 0$ on A and $\overline{\int}_A f=0$ (or $f \leq 0$ and $\underline{\int}_A f=0$), then $f=0$ a.e. on A.

Proof

We prove only some of the above, leaving the rest to the reader.

(a) This following by Corollary 1 (iv) in §4.

(b) Use (a) and Corollary 1 (v) in §4.

(c) First, let

$$f \geq g \geq 0 \text{ on } A. \tag{8.5.31}$$

Take any elementary and nonnegative map $H \geq f$ on A. Then $H \geq g$ as well; so by definition,

$$\overline{\int}_A g=\inf_{h \geq g} \int_A h \leq \int_A H. \tag{8.5.32}$$

Thus

$$\overline{\int}_A f \leq \int_A H \tag{8.5.33}$$

for any such H. Hence also

$$\overline{\int}_A g \leq \inf_{H \geq f} \int_A H=\overline{\int}_A f. \tag{8.5.34}$$

Similarly,

$$\underline{\int}_A f \geq \underline{\int}_A g \tag{8.5.35}$$

if $f \geq g \geq 0$.
In the general case, $f \geq g$ implies

$$f^+ \geq g^+ \text{ and } f^- \leq g^-. \text{ (Why?)} \tag{8.5.36}$$

Thus by what was proved above,

$$\overline{\int}_A f^+ \geq \overline{\int}_A g^+ \text{ and } \underline{\int}_A f^- \leq \underline{\int}_A g^-. \tag{8.5.37}$$

Hence

$$\overline{\int}_A f^+ - \underline{\int}_A f^- \geq \overline{\int}_A g^+ - \underline{\int}_{-A} g^-; \tag{8.5.38}$$

i.e.,

$$\overline{\int}_A \geq \overline{\int}_A g. \tag{8.5.39}$$

Similarly, one obtains

$$]underline{\int} Af \geq \underline{\int}_A g. \tag{8.5.40}$$

(d) It is clear that (c) implies (d).
(e) Let $0 \leq p < \infty$ and suppose $f \geq 0$ on A. Take any elementary and nonnegative map

$$h \geq f \text{ on } A. \tag{8.5.41}$$

By Corollary 1 (vii) and Note 3 of §4,

$$\int_A ph = p \int_A h \tag{8.5.42}$$

for any such h. Hence

$$\overline{\int}_A pf = \inf_h \int_A ph = \inf_h p \int_A h = p \overline{\int}_A f. \tag{8.5.43}$$

Similarly,

$$\underline{\int}_A pf = p \underline{\int}_A f. \tag{8.5.44}$$

The general case reduces to the case $f \geq 0$ by formula (1).
(e') Assertion (e') follows from (1) since

$$(-f)^{+}=f^{-}, \quad (-f)^{-}=f^{+}, \tag{8.5.45}$$

and $-(x-y)=y-x \quad$ if $x-y$ is orthodox. (Why?)
(f) Take any elementary and nonnegative map

$$h \geq f \geq 0 \text{ on } A. \tag{8.5.46}$$

By Corollary 1 (ii) and Note 3 of §4,

$$\int_{B} h \geq \int_{A} h \tag{8.5.47}$$

for any such h. Hence

$$\overline{\int}_{B} f=\inf_{h} \int_{B} h \leq \inf_{h} \int_{A} h=\overline{\int}_{A} f. \tag{8.5.48}$$

Similarly for $\underline{\int}$.
(g) This follows from (c) and (e') since $\pm f \leq|f|$ implies

$$\overline{\int}_{A}|f| \geq \overline{\int}_{A} f \geq \underline{\int}_{A} f \tag{8.5.49}$$

and

$$\overline{\int}_{A}|f| \geq \overline{\int}_{A}(-f) \geq-\underline{\int}_{A} f \geq-\overline{\int}_{A} f . \square \tag{8.5.50}$$

For (h) and later work, we need the following lemmas.

Lemma 8.5.1

Let $f: S \rightarrow E^{*}$ and $A \in \mathcal{M}$. Then the following are true.
(i) If

$$\int_{A} f<q \in E^{*}, \tag{8.5.51}$$

there is an elementary and (extended) real map

$$h \geq f \text{ on } A, \tag{8.5.52}$$

with

$$\int_{A} h<q. \tag{8.5.53}$$

(ii) If

$$\int_{A} f>p \in E^{*}, \tag{8.5.54}$$

there is an elementary and (extended) real map

$$g \leq f \text{ on } A, \tag{8.5.55}$$

with

$$\int_A g > p; \tag{8.5.56}$$

moreover, g can be made elementary and nonnegative if $f \geq 0$ on A.

Proof

If $f \geq 0$, this is immediate by Definition 1 and the properties of glb and lub.
If, however, $f \ngeq 0$, and if

$$q > \int_A f = \overline{\int}_A f^+ - \underline{\int}_A f^-, \tag{8.5.57}$$

our conventions yield

$$\infty > \int_A f^+. \text{ (Why?)} \tag{8.5.58}$$

Thus there are $u, v \in E^*$ such that $q = u + v$ and

$$0 \leq \int_A f^+ < u < \infty \tag{8.5.59}$$

and

$$-\int_A f^- < v. \tag{8.5.60}$$

To see why this is so, choose u so close to $\overline{\int}_A f^+$ that

$$q - u > -\underline{\int}_A f^- \tag{8.5.61}$$

and set $v = q - u$.
As the lemma holds for positive functions, we find elementary and nonnegative maps h' and h'', with

$$h' \geq f^+, h'' \leq f^-, \tag{8.5.62}$$

$$\int_A h' < u < \infty \text{ and } \int_A h'' > -v. \tag{8.5.63}$$

Let $h = h' - h''$. Then

$$h \geq f^+ - f^- = f, \tag{8.5.64}$$

and by Problem 6 in §4,

$$\int_A h=\int_A h'-\int_A h'' \quad \left(\text{for } \int_A h' \text{ is finite!}\right). \tag{8.5.65}$$

Hence

$$\int_A h>u+v=q, \tag{8.5.66}$$

and clause (i) is proved in full.
Clause (ii) follows from (i) by Theorem 1(e$'$) if

$$\underline{\int_A} f<\infty. \tag{8.5.67}$$

(Verify!) For the case $\underline{\int_A} f=\infty$, see Problem 3. □

Note 3. The preceding lemma shows that formulas $(1')$ and $(1'')$ hold (and might be used as definitions) even for sign-changing f, g, and h.

Lemma 8.5.2

If $f: S \rightarrow E^*$ and $A \in \mathcal{M}$, there are $\mathcal{M}$ -measurable maps g and h, with

$$g \leq f \leq h \text{ on } A, \tag{8.5.68}$$

such that

$$\overline{\int_A} f=\overline{int_A} h \text{ and } \underline{\int_A} f=\underline{\int_A} g. \tag{8.5.69}$$

We can take $g, h \geq 0$ if $f \geq 0$ on A.

Proof

If

$$\overline{\int_A} f=\infty, \tag{8.5.70}$$

the constant map $h=\infty$ satisfies the statement of the theorem.
If

$$-\infty<\overline{\int_A} f<\infty, \tag{8.5.71}$$

let

$$q_n=\overline{\int_A} f+\frac{1}{n}, \quad n=1,2, \ldots; \tag{8.5.72}$$

so

$$q_n \rightarrow \overline{\int}_A f < q_n. \tag{8.5.73}$$

By Lemma 1, for each n there is an elementary and (extended) real (hence measurable) map $h_n \geq f$ on A, with

$$q_n \geq \int_A h_n \geq \overline{\int}_A f. \tag{8.5.74}$$

Let

$$h = \inf_n h_n \geq f. \tag{8.5.75}$$

By Lemma 1 in §2, h is $\mathcal{M}$-measurable on A. Also,

$$(\forall n) \quad q_n > \int_A h_n \geq \overline{\int}_A h \geq \overline{\int}_A f \tag{8.5.76}$$

by Theorem $1(\mathrm{c})$. Hence

$$\overline{\int}_A f = \lim_{n \rightarrow \infty} q_n \geq \overline{\int}_A h \geq \overline{\int}_A f, \tag{8.5.77}$$

so

$$\overline{\int}_A f = \overline{\int}_A h, \tag{8.5.78}$$

as required.
Finally, if

$$\overline{\int}_A f = -\infty, \tag{8.5.79}$$

the same proof works with $q_n = -n$. (Verify!)
Similarly, one finds a measurable map $g \leq f$, with

$$\underline{\int}_A f = \underline{\int}_A g. \square \tag{8.5.80}$$

Proof of Theorem 1(h). If $f \geq 0$, choose $h \geq f$ as in Lemma 2. Let

$$D = A(h > 0) \text{ and } A_n = A\left(h > \frac{1}{n}\right); \tag{8.5.81}$$

so

$$D = \bigcup_{n=1}^{\infty} A_n (\text{ why ?}) \tag{8.5.82}$$

and $D, A_n \in \mathcal{M}$ by Theorem 1 of §2. Also,

$$0=\overline{\int}_A f=\overline{\int}_A h \geq \int_{A_n}\left(\frac{1}{n}\right)=\frac{1}{n} m A_n \geq 0. \tag{8.5.83}$$

Thus $(\forall n) m A_n=0$. Hence

$$m D=m \bigcup_{n=1}^{\infty} A_n=m A(h>0)=0; \tag{8.5.84}$$

so $0 \leq f \leq h \leq 0(\text { i.e., } f=0)$ a.e. on A.
The case $f \leq 0$ reduces to $(-f) \geq 0$. $\square$

Corollary 8.5.1

If

$$\overline{\int}_A|f|<\infty, \tag{8.5.85}$$

then $|f|<\infty$ a.e. on A, and $A(f \neq 0)$ is σ-finite.

Proof

By Lemma 1, fix an elementary and nonnegative $h \geq|f|$ with

$$\int_A h<\infty \tag{8.5.86}$$

(so h is elementary and integrable).
Now, by Corollary $1(\mathrm{i})-(\text { iii })$ in §4, our assertions apply to h, hence certainly to f. $\square$

Theorem 8.5.2

(additivity). Given $f: S \rightarrow E^*$ and an $\mathcal{M}$-partition $\mathcal{P}=\left\{B_n\right\}$ of $A \in \mathcal{M}$, we have

$$\text{(a)} \overline{\int}_A f=\sum_n \overline{\int}_{B_n} f \quad \text{and} \quad \text{(b)} \underline{\int}_A f=\sum_n \underline{\int}_{B_n} f, \tag{8.5.87}$$

provided

$$\overline{\int}_A f\left(\underline{\int}_A f, \text{ respectively}\right) \tag{8.5.88}$$

is orthodox, or $\mathcal{P}$ is finite.
Hence if f is integrable on each of finitely many disjoint M-sets B_n, it is so on

$$A=\bigcup_n B_n, \tag{8.5.89}$$

and formulas $(2)(\mathrm{a})(\mathrm{b})$ apply.

Proof

Assume first $f \geq 0$ on A. Then by Theorem $1(\mathrm{f})$, if one of

$$\overline{\int}_{B_n} f=\infty, \tag{8.5.90}$$

so is $\overline{\int}_A f$, and all is trivial. Thus assume all $\int_{B_n} f$ are finite.
Then for any $\varepsilon>0$ and $n \in N$, there is an elementary and nonnegative map $h_n \geq f$ on B_n, with

$$\int_{B_n} h_n<\overline{\int}_{B_n} f+\frac{\varepsilon}{2^n}. \tag{8.5.91}$$

(Why?) Now define $h: A \rightarrow E^*$ by $h=h_n$ on $B_n, n=1,2, \ldots$
Clearly, h is elementary and nonnegative on each B_n, hence on A (Corollary 3 in §1), and $h \geq f$ on A. Thus by Theorem 1 of §4,

$$\overline{\int}_A f \leq \int_A h=\sum_n \int_{B_n} h_n \leq \sum_n\left(\overline{\int}_{B_n} f+\frac{\varepsilon}{2^n}\right) \leq \sum_n \overline{\int}_{B_n} f+\varepsilon. \tag{8.5.92}$$

Making $\varepsilon \rightarrow 0$, we get

$$\overline{\int}_A f \leq \sum_n \overline{\int}_{B_n} f. \tag{8.5.93}$$

To prove also

$$\overline{\int}_A f \geq \sum_n \overline{\int}_{B_n} f, \tag{8.5.94}$$

take any elementary and nonnegative map $H \geq f$ on A. Then again,

$$\int_A H=\sum_n \int_{B_n} H \geq \sum_n \overline{\int}_{B_n} f. \tag{8.5.95}$$

As this holds for any such H, we also have

$$\overline{\int}_A f=\inf_H \int_A H \geq \sum_n \overline{\int}_{B_n} f. \tag{8.5.96}$$

This proves formula (a) for $f \geq 0$. The proof of (b) is quite similar.
If $f \ngeq 0$, we have

$$\overline{\int}_A f=\overline{\int}_A f^{+}-\underline{\int}_A f^{-}, \tag{8.5.97}$$

where by the first part of the proof,

$$\overline{\int}_A f^{+}=\sum_n \overline{\int}_{B_n} f^{+} \text { and } \underline{\int}_A f^{-}=\sum_n \underline{\int}_{B_n} f^{-}. \tag{8.5.98}$$

If

$$\overline{\int}_A f \tag{8.5.99}$$

is orthodox, one of these sums must be finite, and so their difference may be rearranged to yield

$$\overline{\int}_A f=\sum_n\left(\overline{\int}_{B_n} f^{+}-\underline{\int}_{B_n} f^{-}\right)=\sum_n \overline{\int}_{B_n} f, \tag{8.5.100}$$

proving (a). Similarly for (b).
This rearrangement works also if $\mathcal{P}$ is finite (i.e., the sums have a finite number of terms). For, then, all reduces to commutativity and associativity of addition, and our conventions $\left(2^{*}\right)$ of Chapter 4, §4. Thus all is proved. □

Corollary 8.5.2

If $mQ=0(Q\in\mathcal{M})$, then for $A\in\mathcal{M}$

$$\overline{\int}_{A-Q} f=\overline{\int}_A f \text{ and } \underline{\int}_{A-Q} f=\underline{\int}_A f. \tag{8.5.101}$$

For by Theorem 2,

$$\overline{\int}_A f=\overline{\int}_{A-Q} f+\overline{\int}_{A\cap Q} f, \tag{8.5.102}$$

where

$$\overline{\int}_{A\cap Q} f=0 \tag{8.5.103}$$

by Theorem 1(b).

Corollary 8.5.3

If

$$\overline{\int}_A f\left(or \underline{\int}_A f\right) \tag{8.5.104}$$

is orthodox, so is

$$\overline{\int}_X f\left(\underline{\int}_X f\right) \tag{8.5.105}$$

whenever $A\supseteq X, X\in\mathcal{M}$.
For if

$$\overline{\int}_{A} f^{+}, \overline{\int}_{A} f^{-}, \underline{\int}_{A} f^{+}, \text{ or } \underline{\int}_{A} f^{-} \text{ is finite,} \tag{8.5.106}$$

it remains so also when A is reduced to X (see Theorem $1(\mathrm{f})$. Hence orthodoxy follows by formula (1).
Note 4. Given $f : S \rightarrow E^{*},$ we can define two additive (by Theorem 2) set functions $\bar{s}$ and $\underline{s}$ by setting for $X \in \mathcal{M}$

$$\bar{s} X=\overline{\int}_{X} f \text{ and } \underline{s} X=\underline{\int}_{X} f. \tag{8.5.107}$$

They are called, respectively, the upper and lower indefinite integrals of f, also denoted by

$$\overline{\int} f \text{ and } \underline{\int} f \tag{8.5.108}$$

$\left(\text{ or } \bar{s}_{f} \text{ and } \underline{s}_{f}\right)$.
By Theorem 2 and Corollary 3, if

$$\overline{\int}_{A} f \tag{8.5.109}$$

is orthodox, then $\bar{s}$ is σ -additive (and semifinite) when restricted to $\mathcal{M}$-sets $X \subseteq A.$ Also

$$\bar{s} \emptyset=\underline{s} \emptyset=0 \tag{8.5.110}$$

by Theorem 1(b).
Such set functions are called signed measures (see Chapter 7, §11). In particular, if $f \geq 0$ on $S, \bar{s}$ and $\underline{s}$ are σ-additive and nonnegative on all of $\mathcal{M}$, hence measures on $\mathcal{M}$.

Theorem 8.5.3

If $f : S \rightarrow E^{*}$ is m-measurable (Definition 2 in §3) on A, then

$$\overline{\int}_{A} f=\underline{\int}_{A} f. \tag{8.5.111}$$

Proof

First, let $f \geq 0$ on A. By Corollary 2, we may assume that f is $\mathcal{M}$-measurable on A (drop a set of measure zero). Now fix $\varepsilon>0$.
Let $A_{0}=A(f=0), A_{\infty}=A(f=\infty),$ and

$$A_{n}=A\left((1+\varepsilon)^{n} \leq f<(1+\varepsilon)^{n+1}\right), \quad n=0, \pm 1, \pm 2, \ldots \tag{8.5.112}$$

Clearly, these are disjoint $\mathcal{M}$ -sets (Theorem 1 of §2), and

$$A=A_{0} \cup A_{\infty} \cup \bigcup_{n=-\infty}^{\infty} A_{n}. \tag{8.5.113}$$

Thus, setting

$$g=\begin{cases}0 & \text{on } A_0\\ \infty & \text{on } A_\infty, \text{ and}\\ (1+\varepsilon)^n & \text{on } A_n (n=0,\pm 1,\pm 2,\ldots)\end{cases} \tag{8.5.114}$$

and
$h=(1+\varepsilon)g$ on A,
we obtain two elementary and nonnegative maps, with

$$g\leq f\leq h \text{ on } A. \text{ (Why?)} \tag{8.5.115}$$

By Note 1,

$$\underline{\int_A} g=\overline{\int_A} g. \tag{8.5.116}$$

Now, if $\int_A g=\infty$, then

$$\overline{\int_A} f\geq \underline{\int_A} f\geq \int_A g \tag{8.5.117}$$

yields

$$\overline{\int_A} f\geq \underline{\int_A} f=\infty. \tag{8.5.118}$$

If, however, $\int_A g<\infty$, then

$$\int_A h=\int_A (1+\varepsilon)g=(1+\varepsilon)\int_A g<\infty; \tag{8.5.119}$$

so g and h are elementary and integrable on A. Thus by Theorem 2(ii) in §4,

$$\int_A h-\int_A g=\int_A (h-g)=\int_A ((1+\varepsilon)g-g)=\varepsilon\int_A g. \tag{8.5.120}$$

Moreover, $g\leq f\leq h$ implies

$$\int_A g\leq \underline{\int_A} f\leq \overline{\int_A} f\leq \int_A h; \tag{8.5.121}$$

so

$$\left|\overline{\int}_A f-\underline{\int_A} f\right|\leq \int_A h-\int_A g\leq \varepsilon\int_A g. \tag{8.5.122}$$

As ε is arbitrary, all is proved for $f\geq 0$.
The case $f\ngeq 0$ now follows by formula (1), since f^+ and f^- are $\mathcal{M}-$measurable (Theorem 2 in §2). □

8.5.E: Problems on Integration of Extended-Real Functions

Exercise 8.5.E.1

Using the formulas in (1) and our conventions, verify that
(i) $\overline{\int}_A f=+\infty$ iff $\overline{\int}_A f^{+}=\infty$;
(ii) $\underline{\int}_A f=\infty$ iff $\underline{\int}_A f^{+}=\infty$; and
(iii) $\overline{\int}_A f=-\infty$ iff $\underline{\int}_A f^{-}=\infty$ and $\overline{\int}_A f^{+}<\infty$.
(iv) Derive a condition similar to (iii) for $\underline{\int}_A f=-\infty$.
(v) Review Problem 6 of Chapter 4, §4.

Exercise 8.5.E.2

Fill in the missing proof details in Theorems 1 to 3 and Lemmas 1 and 2.

Exercise 8.5.E.3

Prove that if $\underline{\int}_A f=\infty$, there is an elementary and (extended) real map $g \leq f$ on A, with $\int_A g=\infty$.
[Outline: By Problem 1, we have

$$\underline{\int}_A f^{+}=\infty. \tag{8.5.E.1}$$

As Lemmas 1 and 2 surely hold for nonnegative functions, fix a measurable $F \leq f^{+}$ $(F \geq 0)$, with

$$\int_A F=\underline{\int}_A f^{+}=\infty. \tag{8.5.E.2}$$

Arguing as in Theorem 3, find an elementary and nonnegative map $g \leq F$, with

$$(1+\varepsilon) \int_A g=\int_A F=\infty; \tag{8.5.E.3}$$

so $\int_A g=\infty$ and $0 \leq g \leq F \leq f^{+}$ on A.
Let

$$A_{+}=A(F>0) \in \mathcal{M} \tag{8.5.E.4}$$

and

$$A_0=A(F=0) \in \mathcal{M} \tag{8.5.E.5}$$

(Theorem 1 in §2). On A_{+},

$$g \leq F \leq f^{+}=f(\text{ why? }), \tag{8.5.E.6}$$

while on $A_0, g=F=0;$ so

$$\int_{A_{+}} g=\int_A g=\infty(\text{why?}). \tag{8.5.E.7}$$

Now redefine $g=-\infty$ on A_0 (only). Show that g is then the required function.]

Exercise $8.5.E.4$

For any $f: S \rightarrow E^*$, prove the following.
(a) If $\overline{\int}_A f<\infty$, then $f<\infty$ a.e. on A.
(b) If $\underline{\int}_A f$ is orthodox and $>-\infty$, then $f>-\infty$ a.e. on A.
[Hint: Use Problem 1 and apply Corollary 1 to f^+; thus prove (a). Then for (b), use Theorem 1(e').]

Exercise $8.5.E.5$

$\Rightarrow 5$. For any $f, g: S \rightarrow E^*$, prove that
(i) $\overline{\int}_A f+\overline{\int}_A g \geq \overline{\int}_A(f+g)$, and
(ii) $\underline{\int}_A(f+g) \geq \underline{\int}_A f+\underline{\int}_A g$ if $\left|\underline{\int}_A g\right|<\infty$.
[Hint: Suppose that

$$\overline{\int}_A f+\overline{\int}_A g<\overline{\int}_A(f+g). \tag{8.5.E.8}$$

Then there are numbers

$$u>\overline{\int}_A f \text{ and } v>\overline{\int}_A g, \tag{8.5.E.9}$$

with

$$u+v \leq overline \int_A(f+g). \tag{8.5.E.10}$$

(Why?) Thus Lemma 1 yields elementary and (extended) real maps $F \geq f$ and $G \geq g$ such that

$$u>\overline{\int}_A F \text{ and } v>\overline{\int}_A G \tag{8.5.E.11}$$

As $f+g \leq F+G$ on A, Theorem $1(\mathrm{c})$ of §5 and Problem 6 of §4 show that

$$\overline{\int}_A(f+g) \leq \int_A(F+G)=\int_A F+\int_A G<u+v, \tag{8.5.E.12}$$

contrary to

$$u+v \leq \overline{\int}_A(f+g). \tag{8.5.E.13}$$

Similarly prove clause (ii).]

Exercise 8.5.E.6

Continuing Problem 5, prove that

$$\overline{\int}_A (f+g) \geq \overline{\int}_A f + \underline{\int}_A g \geq \underline{\int}_A (f+g) \geq \underline{\int}_A f + \underline{\int}_A g \tag{8.5.E.14}$$

provided $\left|\underline{\int}_A g\right| < \infty$.
[Hint for the second inequality: We may assume that

$$\overline{\int}_A (f+g) < \infty \text{ and } \overline{\int}_A f > -\infty. \tag{8.5.E.15}$$

(Why?) Apply Problems 5 and 4(a) to

$$\overline{\int}_A ((f+g)+(-g)). \tag{8.5.E.16}$$

Use Theorem 1 (e$'$).]

Exercise 8.5.E.7

Prove the following.
(i)

$$\overline{\int}_A |f| < \infty \text{ iff } -\infty < \underline{\int}_A f \leq \overline{\int}_A f < \infty. \tag{8.5.E.17}$$

(ii) If $\overline{\int}_A |f| < \infty$ and $\overline{\int}_A |g| < \infty$, then

$$\left|\overline{\int}_A f - \overline{\int}_A g\right| \leq \overline{\int}_A |f-g| \tag{8.5.E.18}$$

and

$$\left|\underline{\int}_A f - \underline{\int}_A g\right| \leq \overline{\int}_A |f-g|. \tag{8.5.E.19}$$

[Hint: Use Problems 5 and 6.]

Exercise 8.5.E.8

Show that any signed measure $\bar{s}_f$ (Note 4) is the difference of two measures: $\bar{s}_f = \bar{s}_{f+} - \bar{s}_{f-}$.

8.6: Integrable Functions. Convergence Theorems

I. Some important theorems apply to integrable functions.

Theorem 8.6.1 (linearity of the integral)

If $f, g: S \rightarrow E^{*}$ are integrable on a set $A \in \mathcal{M}$ in $(S, \mathcal{M}, m)$, so is

$$pf+qg \tag{8.6.1}$$

for any $p, q \in E^{1}$, and

$$\int_{A}(pf+qg)=p\int_{A} f+q\int_{A} g; \tag{8.6.2}$$

in particular,

$$\int_{A}(f \pm g)=\int_{A} f \pm \int_{A} g. \tag{8.6.3}$$

Proof

By Problem 5 in §5,

$$\overline{\int}_{A} f+\overline{\int}_{A} g \geq \overline{\int}_{A}(f+g) \geq \underline{\int}_{A}(f+g) \geq \underline{\int}_{A} f+\underline{\int}_{A} g. \tag{8.6.4}$$

(Here

$$\overline{\int}_{A} f, \underline{\int}_{A} f, \overline{\int}_{A} g, \text{ and } \underline{\int}_{A} g \tag{8.6.5}$$

are finite by integrability; so all is orthodox.)

As

$$\overline{\int}_{A} f=\underline{\int}_{A} f \text{ and } \overline{\int}_{A} g=\underline{\int}_{A} g, \tag{8.6.6}$$

the inequalities turn into equalities, so that

$$\int_{A} f+\int_{A} g=\overline{\int}_{A}(f+g)=\underline{\int}_{A}(f+g). \tag{8.6.7}$$

Using also Theorem 1(e)(e') from §5, we obtain the desired result for any $p, q \in E^{1}$. $\square$

Theorem 8.6.2

A function $f: S \rightarrow E^{*}$ is integrable on A in $(S, \mathcal{M}, m)$ iff

(i) it is m-measurable on A, and

(ii) $\overline{\int}_{A} f$ (equivalently $\overline{\int}_{A}|f|$) is finite.

Proof

If these conditions hold, f is integrable on A by Theorem 3 of §5.

Conversely, let

$$\overline{\int}_{A} f=\underline{\int}_{A} f \neq \pm \infty. \tag{8.6.8}$$

Using Lemma 2 in §5, fix measurable maps g and h $(g \leq f \leq h)$ on A, with

$$\int_{A} g=\int_{A} f=\int_{A} h \neq \pm \infty. \tag{8.6.9}$$

By Theorem 3 in §5, g and h are integrable on A; so by Theorem 1,

$$\int_{A}(h-g)=\int_{A} h-\int_{A} g=0. \tag{8.6.10}$$

As

$$h-g \geq h-f \geq 0, \tag{8.6.11}$$

we get

$$\int_{A}(h-f)=0, \tag{8.6.12}$$

and so by Theorem 1(h) of §5, $h-f=0$ a.e. on A.

Hence f is almost measurable on A, and

$$\int_{A} f \neq \pm \infty \tag{8.6.13}$$

by assumption. From formula (1), we then get

$$\int_{A} f^{+} \text { and } \int_{A} f^{-}<\infty, \tag{8.6.14}$$

and hence

$$\int_{A}|f|=\int_{A}\left(f^{+}+f^{-}\right)=\int_{A} f^{+}+\int_{A} f^{-}<\infty \tag{8.6.15}$$

by Theorem 1 and by Theorem 2 of §2. Thus all is proved. □

Simultaneously, we also obtain the following corollary.

Corollary 8.6.1

A function $f: S \rightarrow E^{*}$ is integrable on A iff f^{+} and f^{-} are.

Corollary 8.6.2

If $f, g: S \rightarrow E^{*}$ are integrable on A, so also are

$$f \vee g, f \wedge g,|f|, \text { and } k f \text { for } k \in E^{1}, \tag{8.6.16}$$

with

$$\int_{A} k f=k \int_{A} f. \tag{8.6.17}$$

Exercise!

For products fg, this holds if f or g is bounded. In fact, we have the following theorem.

Theorem 8.6.3 (weighted law of the mean)

Let f be m-measurable and bounded on A. Set

$$p=\inf f[A] \text{ and } q=\sup f[A]. \tag{8.6.18}$$

Then if g is m-integrable on A, so is fg, and

$$\int_A f|g|=c\int_A|g| \tag{8.6.19}$$

for some $c\in[p,q]$.

If, further, f also has the Darboux property on A (Chapter 4, §9), then $c=f(x_0)$ for some $x_0\in A$.

Proof

By assumption,

$$(\exists k\in E^1)\quad |f|\leq k \tag{8.6.20}$$

on A. Hence if $\int_A|g|=0$,

$$\left|\int_A f|g|\right|\leq\int_A|fg|\leq k\int_A|g|=0; \tag{8.6.21}$$

so any $c\in[p,q]$ yields

$$\int_A f|g|=c\int_A|g|=0. \tag{8.6.22}$$

If, however, $\int_A|g|\neq 0$, the number

$$c=\left(\int_A f|g|\right)/\int_A|g| \tag{8.6.23}$$

is the required constant.

Moreover, as f and g are m-measurable on A, so is fg; and as

$$\left|\int_A fg\right|\leq|c|\int_A|g|<\infty, \tag{8.6.24}$$

fg is integrable on A by Theorem 2.

Finally, if f has the Darboux property and if $p<c<q$ (with p,q as above), then

$$f(x)<c<f(y) \tag{8.6.25}$$

for some $x,y\in A$ (why?); hence by the Darboux property, $f(x_0)=c$ for some $x_0\in A$.

If, however,

$$c\leq\inf f[A]=p, \tag{8.6.26}$$

then

$$(f-c)|g|\geq 0 \tag{8.6.27}$$

and

$$\int_A(f-c)|g|=\int_A f|g|-c\int_A|g|=0 \text{ (why?)}; \tag{8.6.28}$$

so by Theorem 1(h) in §5, $f-c=0$ a.e. on A. Then surely $f(x_0)=c$ for some $x_0\in A$ (except the trivial case $mA=0$). This also implies $c\in f[A]\in[p,q]$.

Proceed similarly in the case $c\geq q$. □

Corollary 8.6.3

If f is integrable on $A \in \mathcal{M}$, it is so on any $B \subseteq A(B \in \mathcal{M})$.

Proof

Apply Theorem 1(f) in §5, and Theorem 3 of §5, to f^{+} and f^{-}. $\square$

II. Convergence Theorems. If $f_n \rightarrow f$ on A (pointwise, a.e., or uniformly), does it follow that

$$\int_A f_n \rightarrow \int_A f? \tag{8.6.29}$$

To give some answers, we need a lemma.

Lemma 8.6.1

If $f \geq 0$ on $A \in \mathcal{M}$ and if

$$\underline{\int}_A f > p \in E^*, \tag{8.6.30}$$

there is an elementary and nonnegative map g on A such that

$$\int_A g > p, \tag{8.6.31}$$

and $g < f$ on A except only at those $x \in A$ (if any) at which

$$f(x) = g(x) = 0. \tag{8.6.32}$$

(We then briefly write $g \subset f$ on A.)

Proof

By Lemma 1 in §5, there is an elementary and nonnegative map $G \leq f$ on A, with

$$\underline{\int}_A f \geq \int_A G > p. \tag{8.6.33}$$

For the rest, proceed as in Problem 7 of §4, replacing f by G there. $\square$

Theorem 8.6.4 (monotone convergence)

If $0 \leq f_n \nearrow f$ (a.e.) on $A \in \mathcal{M}$, i.e.,

$$0 \leq f_n \leq f_{n+1} \quad (\forall n), \tag{8.6.34}$$

and $f_n \rightarrow f$ (a.e.) on A, then

$$\overline{\int}_A f_n \nearrow \overline{\int}_A f. \tag{8.6.35}$$

Proof for $\mathcal{M}$-measurable f_n and f on A.

By Corollary 2 in §5, we may assume that $f_n \nearrow f$ (pointwise) on A (otherwise, drop a null set).

By Theorem 1(c) of §5, $0 \leq f_n \nearrow f$ implies

$$0 \leq \int_A f_n \leq \int_A f, \tag{8.6.36}$$

and so

$$\lim_{n \rightarrow \infty} \int_{A} f_{n} \leq \int_{A} f. \tag{8.6.37}$$

The limit, call it p, exists in E^{*}, as $\left\{\int_{A} f_{n}\right\} \uparrow$. It remains to show that

$$p \geq \overline{\int}_{A} f=\underline{\int}_{A} f. \tag{8.6.38}$$

(We know that

$$\overline{\int}_{A} f=\underline{\int}_{A} f, \tag{8.6.39}$$

by the assumed measurability of f; see Theorem 3 in §5.)

Suppose

$$\underline{\int}_{A} f>p. \tag{8.6.40}$$

Then Lemma 1 yields an elementary and nonnegative map $g \subset f$ on A, with

$$p<\int_{A} g. \tag{8.6.41}$$

Let

$$A_{n}=A\left(f_{n} \geq g\right), \quad n=1,2, \ldots. \tag{8.6.42}$$

Then $A_{n} \in \mathcal{M}$ and

$$A_{n} \nearrow A=\bigcup_{n=1}^{\infty} A_{n}. \tag{8.6.43}$$

For if $f(x)=0,$ then $x \in A_{1}$, and if $f(x)>0,$ then $f(x)>g(x),$ so that $f_{n}(x)>g(x)$ for large n; hence $x \in A_{n}$.

By Note 4 in §5, the set function $s=\int g$ is a measure, hence continuous by Theorem 2 in Chapter 7, §4. Thus

$$\int_{A} g=s A=\lim_{n \rightarrow \infty} s A_{n}=\lim_{n \rightarrow \infty} \int_{A_{n}} g. \tag{8.6.44}$$

But as $g \leq f_{n}$ on A_{n}, we have

$$\int_{A_{n}} g \leq \int_{A_{n}} f_{n} \leq \int_{A} f_{n}. \tag{8.6.45}$$

Hence

$$\int_{A} g=\lim \int_{A_{n}} g \leq \lim \int_{A} f_{n}=p, \tag{8.6.46}$$

contrary to $p<\int_{A} g$. This contradiction completes the proof. □

Lemma 8.6.2 (Fatou)

If $f_{n} \geq 0$ on $A \in \mathcal{M}(n=1,2, \ldots)$, then

$$\overline{\int}_{A} \underline{\lim} f_{n} \leq \underline{\lim} \overline{\int}_{A} f_{n}. \tag{8.6.47}$$

Proof

Let

$$g_n = \inf_{k \geq n} f_k, \quad n = 1, 2, \ldots; \tag{8.6.48}$$

so $f_n \geq g_n \geq 0$ and $\{g_n\} \uparrow$ on A. Thus by Theorem 4,

$$\overline{\int}_A \lim g_n = \lim \overline{\int}_A g_n = \underline{\lim} \overline{\int}_A g_n \leq \underline{\lim} \overline{\int}_A f_n. \tag{8.6.49}$$

But

$$\lim_{n \rightarrow \infty} g_n = \sup_n g_n = \sup_n \inf_{k \geq n} f_k = \underline{\lim}_n f_n. \tag{8.6.50}$$

Hence

$$\overline{\int}_A \underline{\lim} f_n = \overline{\int}_A \lim g_n \leq \underline{\lim} \overline{\int}_A f_n, \tag{8.6.51}$$

as claimed. $\square$

Theorem 8.6.5 (dominated convergence)

Let $f_n : S \rightarrow E$ be m-measurable on $A \in \mathcal{M} (n = 1, 2, \ldots)$. Let

$$f_n \rightarrow f \text{ (a.e.) on } A. \tag{8.6.52}$$

Then

$$\lim_{n \rightarrow \infty} \int_A |f_n - f| = 0, \tag{8.6.53}$$

provided that there is a map $g : S \rightarrow E^1$ such that

$$\int_A g < \infty \tag{8.6.54}$$

and

$$(\forall n) \quad |f_n| \leq g \text{ a.e. on } A. \tag{8.6.55}$$

Proof

Neglecting null sets, we may assume that

$$|f_n| \leq g < \infty \tag{8.6.56}$$

on A and $f_n \rightarrow f$ (pointwise) on A; so $|f| \leq g$ and

$$|f_n - f| \leq |f_n| + |f| \leq 2g \tag{8.6.57}$$

on A. As $|f| < \infty$, we have

$$|f_n - f| \rightarrow 0 \tag{8.6.58}$$

on A. Hence, setting

$$h_n = 2g - |f_n - f| \geq 0, \tag{8.6.59}$$

we get

$$2g = \lim_{n \rightarrow \infty} h_n = \underline{\lim} h_n. \tag{8.6.60}$$

We may also assume that g is measurable on A. (If not, replace it by a measurable $G \geq g$, with

$$\int_A G = \int_A g < \infty, \tag{8.6.61}$$

by Lemma 2 in §5.) Then all

$$h_n = 2g - |f_n - f| \tag{8.6.62}$$

are measurable (even integrable) on A.

Thus by Lemma 2,

$$\begin{aligned} \int_A 2g = \int_A \underline{\lim}\, h_n &\leq \underline{\lim} \int_A (2g - |f_n - f|) \\ &= \underline{\lim} \left(\int_A 2g + \int_A (-|f_n - f|) \right) \\ &= \int_A 2g + \underline{\lim} \left(-\int_A |f_n - f| \right) \\ &= \int_A 2g - \overline{\lim} \int_A |f_n - f| . \end{aligned}$$

(See Problems 5 and 8 in Chapter 2, §13.)

Canceling $\int_A 2g$ (finite!), we have

$$0 \leq -\overline{\lim} \int_A |f_n - f| . \tag{8.6.63}$$

Hence

$$0 \geq \overline{\lim} \int_A |f_n - f| \geq \underline{\lim} \int_A |f_n - f| \geq 0, \tag{8.6.64}$$

as $|f_n - f| \geq 0$. This yields

$$0 = \overline{\lim} \int_A |f_n - f| = \underline{\lim} \int_A |f_n - f| = \lim \int_A |f_n - f| , \tag{8.6.65}$$

as required. □

Note 1. Theorem 5 holds also for complex and vector-valued functions (for $|f_n - f|$ is real).

In the extended-real case, Theorems 1(g) in §5 and Theorems 1 and 2 in §6 yield

$$\left| \int_A f_n - \int_A f \right| = \left| \int_A (f_n - f) \right| \leq \int_A |f_n - f| \rightarrow 0, \tag{8.6.66}$$

i.e.,

$$\int_A f_n \rightarrow \int_A f. \tag{8.6.67}$$

Moreover, f is integrable on A, being measurable (why?), with

$$\int_A |f| \leq \int_A g < \infty. \tag{8.6.68}$$

For complex and vector-valued functions, this will follow from §7. Observe that Theorem 5, unlike Theorem 4, requires the m-measurability of the f_n.

Note 2. Theorem 5 fails if there is no "dominating"

$$g \geq |f_n| \text{ with } \int_A g < \infty, \tag{8.6.69}$$

even if f and the f_n are integrable.

Example

Let m be Lebesgue measure in $A=E^1, f=0,$ and

$$f_n=\begin{cases}1 & \text{ on } [n, n+1], \\ 0 & \text{ elsewhere. }\end{cases} \tag{8.6.70}$$

Then $f_n \rightarrow f$ and $\int_A f_n=1$; so

$$\lim_{n \rightarrow \infty} \int_A f_n=1 \neq 0=\int_A f. \tag{8.6.71}$$

The trouble is that any

$$g \geq f_n \quad(n=1,2, \ldots) \tag{8.6.72}$$

would have to be ≥ 1 on $B=[1, \infty)$; so

$$\int_A g \geq \int_B g=1 \cdot m B=\infty, \tag{8.6.73}$$

instead of $\int_A g<\infty$.

This example also shows that $f_n \rightarrow f$ alone does not imply

$$\int_A f_n \rightarrow \int_A f. \tag{8.6.74}$$

Theorem 8.6.6 (absolute continuity of the integral)

Given $f: S \rightarrow E$ with

$$\overline{\int}_A|f|<\infty \tag{8.6.75}$$

and $\varepsilon>0,$ there is $\delta>0$ such that

$$\overline{\int}_X|f|<\varepsilon \tag{8.6.76}$$

whenever

$$m X<\delta \quad(A \supseteq X, X \in \mathcal{M}). \tag{8.6.77}$$

Proof

By Lemma 2 in §5, fix $h \geq|f|$, measurable on A, with

$$\int_A h=\overline{\int}_A|f|<\infty. \tag{8.6.78}$$

Neglecting a null set, we assume that $|h|<\infty$ on A (Corollary 1 of §5). Now, $(\forall n)$ set

$$g_n(x)=\begin{cases}h(x), & x \in A_n=A(h<n), \\ 0, & x \in-A_n.\end{cases} \tag{8.6.79}$$

Then $g_n \leq n$ and g_n is measurable on A. (Why?)

Also, $g_n \geq 0$ and $g_n \rightarrow h$ (pointwise) on A.

For let $\varepsilon>0,$ fix $x \in A$, and find $k>h(x).$ Then

$$(\forall n \geq k) \quad h(x) \leq n \text{ and } g_n(x)=h(x). \tag{8.6.80}$$

So

$$(\forall n \geq k) \quad |g_n(x) - h(x)| = 0 < \varepsilon. \tag{8.6.81}$$

Clearly, $g_n \leq h$. Hence by Theorem 5

$$\lim_{n \to \infty} \int_A |h - g_n| = 0. \tag{8.6.82}$$

Thus we can fix n so large that

$$\int_A (h - g_n) < \frac{1}{2}\varepsilon. \tag{8.6.83}$$

For that n, let

$$\delta = \frac{\varepsilon}{2n} \tag{8.6.84}$$

and take any $X \subseteq A(X \in \mathcal{M})$, with $mX < \delta$.

As $g_n \leq n$ (see above), Theorem 1(c) in §5 yields

$$\int_X g_n \leq \int_X (n) = n \cdot mX < n\delta = \frac{1}{2}\varepsilon. \tag{8.6.85}$$

Hence as $|f| \leq h$ and

$$\int_X (h - g_n) \leq \int_A (h - g_n) < \frac{1}{2}\varepsilon \tag{8.6.86}$$

(Theorem 1(f) of §5), we obtain

$$\overline{\int}_X |f| \leq \int_X h = \int_X (h - g_n) + \int_X g_n < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon, \tag{8.6.87}$$

as required. □

8.6.E: Problems on Integrability and Convergence Theorems

Exercise $8.6.E.1$

Fill in the missing details in the proofs of this section.

Exercise $8.6.E.2$

(i) Show that if $f: S \rightarrow E^{*}$ is bounded and m-measurable on A, with $m A<\infty$, then f is m -integrable on $A(\text{ Theorem } 2)$ and

$$\int_{A} f=c \cdot m A, \tag{8.6.E.1}$$

where $\inf f[A] \leq c \leq \sup f[A]$.
(ii) Prove that if f also has the Darboux property on A, then

$$\left(\exists x_{0} \in A\right) \quad c=f\left(x_{0}\right). \tag{8.6.E.2}$$

[Hint: Take $g=1$ in Theorem $3.]$
(iii) What results if $A=[a, b]$ and $m=$ Lebesgue measure?

Exercise $8.6.E.3$

Prove Theorem 4 assuming that the f_{n} are measurable on A and that

$$(\exists k) \quad \int_{A} f_{k}>-\infty \tag{8.6.E.3}$$

instead of $f_{n} \geq 0$.
[Hint: As $\left\{f_{n}\right\} \uparrow$, show that

$$(\forall n \geq k) \quad \int_{A} f_{n}>-\infty. \tag{8.6.E.4}$$

If

$$(\exists n) \quad \int_{A} f_{n}=\infty, \tag{8.6.E.5}$$

then

$$\int_{A} f=\lim \int_{A} f_{n}=\infty. \tag{8.6.E.6}$$

Otherwise,

$$(\forall n \geq k) \quad\left|\int_{A} f_{n}\right|<\infty; \tag{8.6.E.7}$$

so f_{n} is integrable. (Why?) By Corollary 1 in §5, assume $\left|f_{n}\right|<\infty$. (Why?) Apply Theorem 4 to $h_{n}=f_{n}-f_{k}(n \geq k)$, considering two cases:

$$\int_A h < \infty \text{ and } \int_A h = \infty.\Big] \tag{8.6.E.8}$$

Exercise $8.6.E.4$

Show that if $f_n \nearrow f$ (pointwise) on $A \in \mathcal{M}$, there are $\mathcal{M}$-measurable maps $F_n \geq f_n$ and $F \geq f$ on A, with $F_n \nearrow F$ (pointwise) on A, such that

$$\int_A F = \overline{\int}_A f \text{ and } \int_A F_n = \overline{\int}_A f_n. \tag{8.6.E.9}$$

[Hint: By Lemma 2 of §5, fix measurable maps $h \geq f$ and $h_n \geq f_n$ with the same integrals. Let

$$F_n = \inf_{k \geq n} (h \wedge h_k), \quad n = 1, 2, \ldots, \tag{8.6.E.10}$$

and $F = \sup_n F_n \leq h$. (Why?) Proceed.]

Exercise $8.6.E.5$

For $A \in \mathcal{M}$ and any (even nonmeasurable) functions $f, f_n : S \rightarrow E^*$, prove the following.
(i) If $f_n \nearrow f(\text{ a.e. })$ on A, then

$$\overline{\int}_A f_n \nearrow \overline{\int}_A f, \tag{8.6.E.11}$$

provided

$$(\exists n) \quad \overline{\int}_A f_n > -\infty. \tag{8.6.E.12}$$

(ii) If $f_n \searrow f(\text{ a.e. })$ on A, then

$$\underline{\int}_A f_n \searrow \underline{\int}_A f, \tag{8.6.E.13}$$

provided

$$(\exists n) \quad \underline{\int}_A f_n < \infty. \tag{8.6.E.14}$$

[Hint: Replace f, f_n by F, F_n as in Problem 4. Then apply Problem 3 to F_n; thus obtain (i). For (ii), use (i) and Theorem 1 (e') in §5. (All is orthodox; why?)]

Exercise $8.6.E.6$

Show by examples that
(i) the conditions

$$\overline{\int}_A f_n > -\infty \text{ and } \underline{\int}_A f_n < \infty \tag{8.6.E.15}$$

in Problem 5 are essential; and
(ii) Problem $5(\mathrm{i})$ fails for lower integrals. What about $5(\mathrm{ii})$?
[Hints: (i) Let $A=(0,1) \subset E^1, m=$ Lebesgue measure, $f_n=-\infty$ on $\left(0, \frac{1}{n}\right), f_n=1$ elsewhere.
(ii) Let $\mathcal{M}=\left\{E^1, \emptyset\right\}, m E^1=1, m \emptyset=0, f_n=1$ on $(-n, n), f_n=0$ elsewhere. If $f=1$ on $A=E^1$, then $f_n \rightarrow f$, but not

$$\underline{\int}_A f_n \rightarrow \underline{\int}_A f. \tag{8.6.E.16}$$

Explain!]

Exercise $8.6.E.7$

Given $f_n: S \rightarrow E^*$ and $A \in \mathcal{M}$, let

$$g_n=\inf_{k \geq n} f_k \text{ and } h_n=\sup_{k \geq n} f_k \quad(n=1,2, \ldots). \tag{8.6.E.17}$$

Prove that
(i) $\overline{\int}_A \underline{\lim} f_n \leq \underline{\lim} \underline{\int}_A f_n$ provided $(\exists n) \overline{\int}_A g_n>-\infty$; and
(ii) $\underline{\int}_A \overline{\lim} f_n \leq \overline{\lim} \underline{\int}_A f_n$ provided $(\exists n) \underline{\int}_A h_n<\infty$.
[Hint: Apply Problem 5 to g_n and h_n.]
(iii) Give examples for which

$$\overline{\int}_A \underline{\lim} f_n \neq \overline{\lim}_A \overline{\int}_A f_n \text{ and } \underline{\int}_A \overline{\lim} f_n \neq \underline{\lim} \underline{\int}_A f_n. \tag{8.6.E.18}$$

(See Note 2).

Exercise $8.6.E.8$

Let $f_n \geq 0$ on $A \in \mathcal{M}$ and $f_n \rightarrow f$ (a.e.) on A. Let $A \supseteq X, X \in \mathcal{M}$.
Prove the following.
(i) If

$$\overline{\int}_A f_n \rightarrow \overline{\int}_A f<\infty, \tag{8.6.E.19}$$

then

$$\overline{\int}_X f_n \rightarrow \overline{\int}_X f. \tag{8.6.E.20}$$

(ii) This fails for sign-changing f_n.
[Hints: If (i) fails, then

$$\underline{\lim}_X \overline{\int}_X f_n<\overline{\int}_X f \text{ or } \underline{\lim}_X \overline{\int}_X f_n>\overline{\int}_X f. \tag{8.6.E.21}$$

Find a subsequence of

$$\left\{\overline{\int}_X f_n\right\} \text{ or } \left\{\overline{\int}_{A-X} f_n\right\} \tag{8.6.E.22}$$

contradicting Lemma 2.
(ii) Let $m =$ Lebesgue measure; $A = (0, 1), X = \left(0, \frac{1}{2}\right)$,

$$f_n = \begin{cases} n & \text{on } \left(0, \frac{1}{2n}\right], \\ -n & \text{on } \left(1 - \frac{1}{2n}, 1\right[. \end{cases} \tag{8.6.E.23}$$

? Exercise 8.6.E.9

$\Rightarrow 9$. (i) Show that if f and g are m-measurable and nonnegative on A, then

$$(\forall a, b \geq 0) \quad \int_A (af + bg) = a \int_A f + b \int_A g. \tag{8.6.E.24}$$

(ii) If, in addition, $\int_A f < \infty$ or $\int_A g < \infty$, this formula holds for any $a, b \in E^1$.
[Hint: Proceed as in Theorem 1.]

? Exercise 8.6.E.10

$\Rightarrow 10$. If

$$f = \sum_{n=1}^{\infty} f_n, \tag{8.6.E.25}$$

with all f_n measurable and nonnegative on A, then

$$\int_A f = \sum_{n=1}^{\infty} \int_A f_n. \tag{8.6.E.26}$$

[Hint: Apply Theorem 4 to the maps

$$g_n = \sum_{k=1}^{n} f_k \nearrow f. \tag{8.6.E.27}$$

Use Problem 9.]

? Exercise 8.6.E.11

If

$$q = \sum_{n=1}^{\infty} \int_A |f_n| < \infty \tag{8.6.E.28}$$

and the f_n are m-measurable on A, then

$$\sum_{n=1}^{\infty} |f_n| < \infty (a.e.) \text{ on } A \tag{8.6.E.29}$$

and $f=\sum_{n=1}^{\infty} f_n$ is m-integrable on A, with

$$\int_A f=\sum_{n=1}^{\infty} \int_A f_n. \tag{8.6.E.30}$$

[Hint: Let $g=\sum_{n=1}^{\infty}\left|f_n\right|$. By Problem 10,

$$\int_A g=\sum_{n=1}^{\infty} \int_A\left|f_n\right|=q<\infty; \tag{8.6.E.31}$$

so $g<\infty(a.e.)$ on A. (Why?) Apply Theorem 5 and Note 1 to the maps

$$g_n=\sum_{k=1}^n f_k; \tag{8.6.E.32}$$

note that $\left|g_n\right| \leq g$.]

Exercise $8.6.E.12$

(Convergence in measure; see Problem 11(ii) of §3).
(i) Prove Riesz' theorem: If $f_n \rightarrow f$ in measure on $A \subseteq S$, there is a subsequence $\left\{f_{n_k}\right\}$ such that $f_{n_k} \rightarrow f$ (almost uniformly), hence (a.e.), on A.
[Outline: Taking

$$\sigma_k=\delta_k=2^{-k}, \tag{8.6.E.33}$$

pick, step by step, naturals

$$n_1<n_2<\cdots<n_k<\cdots \tag{8.6.E.34}$$

and sets $D_k \in \mathcal{M}$ such that $(\forall k)$

$$m D_k<2^{-k} \tag{8.6.E.35}$$

and

$$\rho^{\prime}\left(f_{n_k}, f\right)<2^{-k} \tag{8.6.E.36}$$

on $A-D_k$. (Explain!) Let

$$E_n=\bigcup_{k=n}^{\infty} D_k, \tag{8.6.E.37}$$

$m E_n<2^{1-n}$. (Why?) Show that

$$(\forall n)(\forall k>n) \quad \rho^{\prime}\left(f_{n_k}, f\right)<2^{1-n} \tag{8.6.E.38}$$

on $A-E_n$. Use Problem 11 in §3.]
(ii) For maps $f_n: S \rightarrow E$ and $g: S \rightarrow E^1$ deduce that if

$$f_n \rightarrow f \tag{8.6.E.39}$$

in measure on A and

$$(\forall n) \quad \left|f_n\right| \leq g(\text{ a.e. }) \text{ on } A, \tag{8.6.E.40}$$

then

$$|f| \leq g(\text{ a.e. }) \text{ on } A. \tag{8.6.E.41}$$

[Hint: $f_{n_k} \rightarrow f(a. e.)$ on A.]

Exercise $8.6.E.13$

Continuing Problem $12(\text{ ii })$, let

$$f_n \rightarrow f \tag{8.6.E.42}$$

in measure on $A \in \mathcal{M}\left(f_n: S \rightarrow E\right)$ and

$$(\forall n) \quad \left|f_n\right| \leq g(\text{a. e.}) \text{ on } A, \tag{8.6.E.43}$$

with

$$\overline{\int}_A g<\infty. \tag{8.6.E.44}$$

Prove that

$$\lim _{n \rightarrow \infty} \overline{\int}_A\left|f_n-f\right|=0. \tag{8.6.E.45}$$

Does

$$\overline{\int}_A f_n \rightarrow \overline{\int}_A f? \tag{8.6.E.46}$$

[Outline: From Corollary 1 of §5, infer that $g=0$ on $A-C$, where

$$C=\bigcup_{k=1}^{\infty} C_k(\text{disjoint}), \tag{8.6.E.47}$$

$m C_k<\infty$. (We may assume $g \mathcal{M}$-measurable on A. Why?) Also,

$$\infty>\int_A g=\int_{A-C} g+\int_C g=0+\sum_{k=1}^{\infty} \int_{C_k} g; \tag{8.6.E.48}$$

so the series converges. Hence

$$(\forall \varepsilon>0)(\exists p) \quad \int_A g-\varepsilon<\sum_{k=1}^p \int_{C_k} g=\int_H g, \tag{8.6.E.49}$$

where

$$H=\bigcup_{k=1}^{p} C_{k} \in \mathcal{M} \tag{8.6.E.50}$$

and $mH<\infty$. As $\left|f_{n}-f\right| \leq 2 g(\text { a.e. })$, we get

$$\text{(1)} \underline{\int}_{A}\left|f_{n}-f\right| \leq \overline{\int}_{A}\left|f_{n}-f\right| \leq \overline{\int}_{H}\left|f_{n}-f\right|+\int_{A-H} 2 g<\overline{\int_{H}}\left|f_{n}-f\right|+2 \varepsilon. \tag{8.6.E.51}$$

(Explain!)
As $mH<\infty$, we can fix $\sigma>0$ with

$$\sigma \cdot m H<\varepsilon. \tag{8.6.E.52}$$

Also, by Theorem 6, fix δ such that

$$2 \int_{X} g<\varepsilon \tag{8.6.E.53}$$

whenever $A \supseteq X, X \in \mathcal{M}$ and $mX<\delta$.
As $f_{n} \rightarrow f$ in measure on H, we find $\mathcal{M}$-sets $D_{n} \subseteq H$ such that

$$\left(\forall n>n_{0}\right) \quad m D_{n}<\delta \tag{8.6.E.54}$$

and

$$\left|f_{n}-f\right|<\sigma \text { on } A_{n}=H-D_{n}. \tag{8.6.E.55}$$

(We may use the standard metric, as $|f|$ and $\left|f_{n}\right|<\infty$ a.e. Why?) Thus from (1), we get

$$\begin{aligned} \overline{\int}_{A}\left|f_{n}-f\right| & \leq \overline{\int_{H}}\left|f_{n}-f\right|+2 \varepsilon \\ &=\overline{\int}_{A_{n}}\left|f_{n}-f\right|+\overline{\int}_{D_{n}}\left|f_{n}-f\right|+2 \varepsilon \\ &<\overline{\int}_{A_{n}}\left|f_{n}-f\right|+3 \varepsilon \\ & \leq \sigma \cdot m H+3 \varepsilon<4 \varepsilon \end{aligned}$$

for $n>n_{0}$. (Explain!) Hence

$$\lim \overline{\int_{A}}\left|f_{n}-f\right|=0. \tag{8.6.E.56}$$

See also Problem 7 in §5 and Note 1 of §6 (for measurable functions) as regards

$$\left.\lim \overline{\int_{A}} f_{n} \cdot\right] \tag{8.6.E.57}$$

Exercise 8.6.E.14

Do Problem 12 in §3 (Lebesgue-Egorov theorems) for $T=E,$ assuming

$$(\forall n) \quad \left|f_n\right| \leq g(a.e.) \text{ on } A, \tag{8.6.E.58}$$

with

$$\int_A g<\infty \tag{8.6.E.59}$$

(instead of $m A<\infty$).
[Hint: With $H_i(k)$ as before, it suffices that

$$\lim_{i \rightarrow \infty} m\left(A-H_i(k)\right)=0. \tag{8.6.E.60}$$

(Why?) Verify that

$$(\forall n) \quad \rho^{\prime}\left(f_n, f\right)=\left|f_n-f\right| \leq 2 g(a.e.) \text{ on } A, \tag{8.6.E.61}$$

and

$$(\forall i, k) \quad A-H_i(k) \subseteq A\left(2 g \geq \frac{1}{k}\right) \cup Q(m Q=0). \tag{8.6.E.62}$$

Infer that

$$(\forall i, k) \quad m\left(A-H_i(k)\right)<\infty. \tag{8.6.E.63}$$

Now, as $(\forall k) H_i(k) \searrow \emptyset$ (why?), right continuity applies.]

8.7: Integration of Complex and Vector-Valued Functions

I. First we consider functions $f : S \rightarrow E^n\left(C^n\right)$. For such functions, it is natural (and easy) to define integration "componentwise" as follows.

Definition

A function $f : S \rightarrow E^n$ is said to be integrable on $A \in \mathcal{M}$ iff its n (real) components, $f_1, \ldots, f_n$, are. In this case, we define

$$\int_A f=\int_A f d m=\left(\int_A f_1, \int_A f_2, \ldots, \int_A f_n\right)=\sum_{k=1}^n \bar{e}_k \cdot \int_A f_k \tag{8.7.1}$$

where the $\bar{e}_k$ are basic unit vectors (as in Chapter 3, §§1-3, Theorem 2).

In particular, a complex function f is integrable on A iff its real and imaginary parts $\left(f_{\text{re}} \text{ and} f_{\text{im}}\right)$ are. Then we also say that $\int_A f$ exists. By (1), we have

$$\int_A f=\left(\int_A f_{\mathrm{re}}, \int_A f_{\mathrm{im}}\right)=\int_A f_{\mathrm{re}}+i \int_A f_{\mathrm{im}}. \tag{8.7.2}$$

If $f : S \rightarrow C^n,$ we use (1), with complex components f_k

With this definition, integration of functions $f : S \rightarrow E^n\left(C^n\right)$ reduces to that of $f_k : S \rightarrow E^1(C)$, and one easily obtains the same theorems as in §§4-6, as far as they make sense for vectors.

Theorem 8.7.1

A function $f : S \rightarrow E^n\left(C^n\right)$ is integrable on $A \in \mathcal{M}$ iff it iscm-measurable on A and $\int_A|f|<\infty$.

(Alternate definition!)

Proof

Assume the range space is E^n.

By our definition, if f is integrable on A, then its components f_k are. Thus by Theorem 2 and Corollary 1, both in §6, for $k=1,2, \ldots, n$, the functions f_k^{+} and f_k^{-} are m-measurable; furthermore,

$$\int_A f_k^{+} \neq \pm \infty \text{ and } \int_A f_k^{-} \neq \pm \infty. \tag{8.7.3}$$

This implies

$$\infty>\int_A f_k^{+}+\int_A f_k^{-}=\int_A\left(f_k^{+}+f_k^{-}\right)=\int_A\left|f_k\right|, \quad k=1,2, \ldots, n. \tag{8.7.4}$$

Since $|f|$ is m-measurable by Problem 14 in §3 ($|\cdot|$ is a continuous mapping from E^n to E^1), and

$$|f|=\left|\sum_{k=1}^n \bar{e}_k f_k\right| \leq \sum_{k=1}^n\left|\bar{e}_k\right|\left|f_k\right|=\sum_{k=1}^n\left|f_k\right|, \tag{8.7.5}$$

we get

$$\int_A|f| \leq \int_A \sum_1^n\left|f_k\right|=\sum_1^n \int_A\left|f_k\right|<\infty. \tag{8.7.6}$$

Conversely, if f satisfies

$$\int_A|f|<\infty \tag{8.7.7}$$

then

$$(\forall k) \quad \left|\int_A f_k\right| < \infty. \tag{8.7.8}$$

Also, the f_k are m-measurable if f is (see Problem 2 in §3). Hence the f_k are integrable on A (by Theorem 2 of §6), and so is f.

The proof for C^n is analogous. □

Similarly for other theorems (see Problems 1 to 4 below). We have already noted that Theorem 5 of §6 holds for complex and vector-valued functions. So does Theorem 6 in §6. We prove another such proposition (Lemma 1) below.

II. Next we consider the general case, $f : S \rightarrow E$ (E complete). We now adopt Theorem 1 as a definition. (It agrees with Definition 1 of §4. Verify!) Even if $E = E^*$, we always assume $|f| < \infty$ a.e.; thus, dropping a null set, we can make f finite and use the standard metric on E^1.

First, we take up the case $mA < \infty$.

Lemma 8.7.1

If $f_n \rightarrow f$ (uniformly) on A ($mA < \infty$), then

$$\int_A |f_n - f| \rightarrow 0. \tag{8.7.9}$$

Proof

By assumption,

$$(\forall \varepsilon > 0)\,(\exists k)\,(\forall n > k) \quad |f_n - f| < \varepsilon \text{ on } A; \tag{8.7.10}$$

so

$$(\forall n > k) \quad \int_A |f_n - f| \leq \int_A (\varepsilon) = \varepsilon \cdot mA < \infty. \tag{8.7.11}$$

As ε is arbitrary, the result follows. □

Lemma 8.7.2

If

$$\int_A |f| < \infty \quad (mA < \infty) \tag{8.7.12}$$

and

$$f = \lim_{n \rightarrow \infty} f_n \text{ (uniformly) on } A - Q \ (mQ = 0) \tag{8.7.13}$$

for some elementary maps f_n on A, then all but finitely many f_n are elementary and integrable on A, and

$$\lim_{n \rightarrow \infty} \int_A f_n \tag{8.7.14}$$

exists in E; further, the latter limit does not depend on the sequence $\{f_n\}$.

Proof

By Lemma 1,

$$(\forall \varepsilon > 0)\,(\exists q)\,(\forall n, k > q) \quad \int_A |f_n - f| < \varepsilon \text{ and } \int_A |f_n - f_k| < \varepsilon. \tag{8.7.15}$$

(The latter can be achieved since

$$\lim_{k\to\infty}\int_A |f_n - f_k| = \int_A |f_n - f| < \varepsilon.) \tag{8.7.16}$$

Now, as

$$|f_n| \le |f_n - f| + |f|, \tag{8.7.17}$$

Problem 7 in §5 yields

$$(\forall n > k) \quad \int_A |f_n| \le \int_A |f_n - f| + \int_A |f| < \varepsilon + \int_A |f| < \infty. \tag{8.7.18}$$

Thus f_n is elementary and integrable for $n > k,$ as claimed. Also, by Theorem 2 and Corollary 1(ii), both in §4,

$$(\forall n, k > q) \quad \left|\int_A f_n - \int_A f_k\right| = \left|\int_A (f_n - f_k)\right| \le \int_A |f_n - f_k| < \varepsilon. \tag{8.7.19}$$

Thus $\left\{\int_A f_n\right\}$ is a Cauchy sequence. As E is complete,

$$\lim \int_A f_n \ne \pm\infty \tag{8.7.20}$$

exists in E, as asserted.

Finally, suppose $g_n \to f$ (uniformly) on $A - Q$ for some other elementary and integrable maps g_n. By what was shown above, $\lim \int_A g_n$ exists, and

$$\left|\lim \int_A g_n - \lim \int_A f_n\right| = \left|\lim \int_A (g_n - f_n)\right| \le \lim \int_A |g_n - f_n - 0| = 0 \tag{8.7.21}$$

by Lemma 1, as $g_n - f_n \to 0$ (uniformly) on A. Thus

$$\lim \int_A g_n = \lim \int_A f_n, \tag{8.7.22}$$

and all is proved. □

This leads us to the following definition.

Definition

If $f : S \to E$ is integrable on $A \in \mathcal{M}$ $(mA < \infty)$, we set

$$\int_A f = \int_A f\,dm = \lim_{n\to\infty}\int_A f_n \tag{8.7.23}$$

for any elementary and integrable maps f_n such that $f_n \to f$ (uniformly) on $A - Q, mQ = 0$.

Indeed, such maps exist by Theorem 3 of §1, and Lemma 2 excludes ambiguity.

***Note 1.** If f itself is elementary and integrable, Definition 2 agrees with that of §4. For, choosing $f_n = f (n = 1, 2, \ldots),$ we get

$$\int_A f = \int_A f_n \tag{8.7.24}$$

(the latter as in §4).

***Note 2.** We may neglect sets on which $f = 0$, along with null sets. For if $f = 0$ on $A - B$ $(A \supseteq B, B \in \mathcal{M}),$ we may choose $f_n = 0$ on $A - B$ in Definition 2. Then

$$\int_A f = \lim \int_A f_n = \lim \int_B f_n = \int_B f. \tag{8.7.25}$$

Thus we now define

$$\int_A f=\int_B f, \tag{8.7.26}$$

even if $mA=\infty$, provided $f=0$ on $A-B$, i.e.,

$$f=fC_B \text{ on } A \tag{8.7.27}$$

($C_B=$ characteristic function of B), with $A\supseteq B, B\in\mathcal{M},$ and $mB<\infty$.

If such a B exists, we say that f has m-finite support in A.

***Note 3.** By Corollary 1 in §5,

$$\int_A|f|<\infty \tag{8.7.28}$$

implies that $A(f\neq 0)$ is σ-finite. Neglecting $A(f=0)$, we may assume that

$$A=\bigcup B_n, mB_n<\infty, \text{ and } \{B_n\}\uparrow \tag{8.7.29}$$

(if not, replace B_n by $\cup_{k=1}^n B_k$); so $B_n\nearrow A$.

Lemma 8.7.3

Let $\phi:S\rightarrow E$ be integrable on A. Let $B_n\nearrow A, mB_n<\infty$ and set

$$f_n=\phi C_{B_n}, \quad n=1,2,\ldots. \tag{8.7.30}$$

Then $f_n\rightarrow\phi$ (pointwise) on A, all f_n are integrable on A, and

$$\lim_{n\rightarrow\infty}\int_A f_n \tag{8.7.31}$$

exists in E. Furthermore, this limit does not depend on the choice of $\{B_n\}$.

Proof

Fix any $x\in A$. As $B_n\nearrow A=\cup B_n$,

$$(\exists n_0)\ (\forall n>n_0)\quad x\in B_n. \tag{8.7.32}$$

By assumption, $f_n=\phi$ on B_n. Thus

$$(\forall n>n_0)\quad f_n(x)=\phi(x); \tag{8.7.33}$$

so $f_n\rightarrow\phi$ (pointwise) on A.

Moreover, $f_n=\phi C_{B_n}$ is m-measurable on A (as ϕ and C_{B_n} are); and

$$|f_n|=|\phi|C_{B_n} \tag{8.7.34}$$

implies

$$\int_A|f_n|\leq\int_A|\phi|<\infty. \tag{8.7.35}$$

Thus all f_n are integrable on A.

As $f_n=0$ on $A-B_n(mB<\infty)$,

$$\int_A f_n \tag{8.7.36}$$

is defined. Since $f_n\rightarrow\phi$ (pointwise) and $|f_n|\leq|\phi|$ on A, Theorem 5 in §6, with $g=|\phi|$, yields

$$\int_A|f_n-\phi|\rightarrow 0. \tag{8.7.37}$$

The rest is as in Lemma 2, with our present Theorem 2 below (assuming m-finite support of f and g), replacing Theorem 2 of §4. Thus all is proved. $\square$

Definition

If $\phi : S \rightarrow E$ is integrable on $A \in \mathcal{M},$ we set

$$\int_{A} \phi=\int_{A} \phi d m=\lim _{n \rightarrow \infty} \int_{A} f_{n}, \tag{8.7.38}$$

with the f_n as in Lemma 3 (even if ϕ has no m-finite support).

Theorem 8.7.2 (linearity)

If $f, g : S \rightarrow E$ are integrable on $A \in \mathcal{M},$ so is

$$p f+q g \tag{8.7.39}$$

for any scalars $p, q.$ Moreover,

$$\int_{A}(p f+q g)=p \int_{A} f+q \int_{A} g. \tag{8.7.40}$$

Furthermore if f and g are scalar valued, p and q may be vectors in E.

Proof

For the moment, f, g denotes mappings with m-finite support in A. Integrability is clear since $pf+qg$ is measurable on A (as f and g are), and

$$|p f+q g| \leq|p||f|+|q||g| \tag{8.7.41}$$

yields

$$\int_{A}|p f+q g| \leq|g| \int_{A}|f|+|q| \int_{A}|g|<\infty. \tag{8.7.42}$$

Now, as noted above, assume that

$$f=f C_{B_{1}} \text { and } g=g C_{B_{2}} \tag{8.7.43}$$

for some $B_{1}, B_{2} \subseteq A\left(m B_{1}+m B_{2}<\infty\right).$ Let $B=B_{1} \cup B_{2};$ so

$$f=g=p f+q g=0 \text { on } A-B; \tag{8.7.44}$$

additionally,

$$\int_{A} f=\int_{B} f, \int_{A} g=\int_{B} g, \text { and } \int_{A}(p f+q g)=\int_{B}(p f+q g). \tag{8.7.45}$$

Also, $m B<\infty;$ so by Definition 2,

$$\int_{B} f=\lim \int_{B} f_{n} \text { and } \int_{B} g=\lim \int_{B} g_{n} \tag{8.7.46}$$

for some elementary and integrable maps

$$f_{n} \rightarrow f \text { (uniformly) and } g_{n} \rightarrow g \text { (uniformly) on } B-Q, m Q=0. \tag{8.7.47}$$

Thus

$$p f_{n}+q g_{n} \rightarrow p f+q g \text { (uniformly) on } B-Q. \tag{8.7.48}$$

But by Theorem 2 and Corollary 1(vii), both of §4 (for elementary and integrable maps),

$$\int_B (pf_n + qg_n) = p\int_B f_n + q\int_B g_n. \tag{8.7.49}$$

Hence

$$\begin{aligned}\int_A (pf+qg) = \int_B (pf+qg) &= \lim \int_B (pf_n+qg_n)\\ &= \lim\left(p\int_B f_n + q\int_B g_n\right) = p\int_B f + q\int_B g = p\int_A f + q\int_A g.\end{aligned}$$

This proves the statement of the theorem, provided f and g have m-finite support in A. For the general case, we now resume the notation $f, g, \ldots$ for any functions, and extend the result to any integrable functions.

Using Definition 3, we set

$$A = \bigcup_{n=1}^{\infty} B_n, \{B_n\}\uparrow, mB_n < \infty, \tag{8.7.50}$$

and

$$f_n = fC_{B_n}, g_n = gC_{B_n}, \quad n = 1, 2, \ldots. \tag{8.7.51}$$

Then by definition,

$$\int_A f = \lim_{n\to\infty}\int_A f_n \text{ and } \int_A g = \lim_{n\to\infty}\int_A g_n, \tag{8.7.52}$$

and so

$$p\int_A f + q\int_A g = \lim_{n\to\infty}\left(p\int_A f_n + q\int_A g_n\right). \tag{8.7.53}$$

As f_n, g_n have m-finite supports, the first part of the proof yields

$$p\int_A f_n + q\int_A g_n = \int_A (pf_n + qg_n). \tag{8.7.54}$$

Thus as claimed,

$$p\int_A f + q\int_A g = \lim\int_A (pf_n + qg_n) = \int_A (pf+qg). \quad \square \tag{8.7.55}$$

Similarly, one extends Corollary 1(ii)(iii)(v) of §4 first to maps with m-finite support, and then to all integrable maps. The other parts of that corollary need no new proof. (Why?)

Theorem 8.7.3 (additivity)

(i) If $f : S \rightarrow E$ is integrable on each of n disjoint $\mathcal{M}$-sets A_k, it is so on their union

$$A = \bigcup_{k=1}^{n} A_k, \tag{8.7.56}$$

and

$$\int_A f = \sum_{k=1}^{n}\int_{A_k} f. \tag{8.7.57}$$

(ii) This holds for countable unions, too, if f is integrable on all of A.

Proof

Let f have m-finite support: $f = fC_B$ on $A, mB < \infty.$ Then

$$\int_{A} f=\int_{B} f \text{ and } \int_{A_{k}} f=\int_{B_{k}} f, \tag{8.7.58}$$

where

$$B_{k}=A_{k} \cap B, \quad k=1,2, \ldots, n. \tag{8.7.59}$$

By Definition 2, fix elementary and integrable maps f_{i} (on A) and a set Q $(mQ=0)$ such that $f_{i} \rightarrow f$ (uniformly) on $B-Q$ (hence also on $B_{k}-Q$), with

$$\int_{A} f=\int_{B} f=\lim _{i \rightarrow \infty} \int_{B} f_{i} \quad \text { and } \int_{A_{k}} f=\lim _{i \rightarrow \infty} \int_{B_{k}} f_{i}, \quad k=1,2, \ldots, n. \tag{8.7.60}$$

As the f_{i} are elementary and integrable, Theorem 1 in §4 yields

$$\int_{A} f_{i}=\int_{B} f_{i}=\sum_{k=1}^{n} \int_{B_{k}} f_{i}=\sum_{k=1}^{n} \int_{A_{k}} f_{i}. \tag{8.7.61}$$

Hence

$$\int_{A} f=\lim _{i \rightarrow \infty} \int_{B} f_{i}=\lim _{i \rightarrow \infty} \sum_{k=1}^{n} \int_{B_{k}} f_{i}=\sum_{k=1}^{n}\left(\lim _{i \rightarrow \infty} \int_{A_{k}} f_{i}\right)=\sum_{k=1}^{n} \int_{A_{k}} f. \tag{8.7.62}$$

Thus clause (i) holds for maps with m-finite support. For other functions, (i) now follows quite similarly, from Definition 3. (Verify!)

As for (ii), let f be integrable on

$$A=\bigcup_{k=1}^{\infty} A_{k} \text { (disjoint), } \quad A_{k} \in \mathcal{M}. \tag{8.7.63}$$

In this case, set $g_{n}=f C_{B_{n}}$, where $B_{n}=\bigcup_{k=1}^{n} A_{k}, n=1,2, \ldots$. By clause (i), we have

$$\int_{A} g_{n}=\int_{B_{n}} g_{n}=\sum_{k=1}^{n} \int_{A_{k}} g_{n}=\sum_{k=1}^{n} \int_{A_{k}} f, \tag{8.7.64}$$

since $g_{n}=f$ on each $A_{k} \subseteq B_{n}$.

Also, as is easily seen, $\left|g_{n}\right| \leq|f|$ on A and $g_{n} \rightarrow f$ (pointwise) on A (proof as in Lemma 3). Thus by Theorem 5 in §6,

$$\int_{A}\left|g_{n}-f\right| \rightarrow 0. \tag{8.7.65}$$

As

$$\left|\int_{A} g_{n}-\int_{A} f\right|=\left|\int_{A}\left(g_{n}-f\right)\right| \leq \int_{A}\left|g_{n}-f\right|, \tag{8.7.66}$$

we obtain

$$\int_{A} f=\lim _{n \rightarrow \infty} \int_{A} g_{n}, \tag{8.7.67}$$

and the result follows by (3). □

8.7.E: Problems on Integration of Complex and Vector-Valued Functions

Exercise $8.7.E.1$

Prove Corollary $1(\text{ iii })-$ (vii) in §4 componentwise for integrable maps $f: S \rightarrow E^{n}\left(C^{n}\right)$.

Exercise $8.7.E.2$

Prove Theorems 2 and 3 componentwise for $E=E^{n}\left(C^{n}\right)$.

Exercise $8.7.E.2'$

Do it for Corollary 3 in §6.

Exercise $8.7.E.3$

Prove Theorem 1 with

$$\int_{A}|f|<\infty \tag{8.7.E.1}$$

replaced by

$$\int_{A}\left|f_{k}\right|<\infty, \quad k=1, \ldots, n. \tag{8.7.E.2}$$

Exercise $8.7.E.4$

Prove that if $f: S \rightarrow E^{n}\left(C^{n}\right)$ is integrable on A, so is $|f|$. Disprove the converse.

Exercise $8.7.E.5$

Disprove Lemma 1 for $mA=\infty$.

Exercise $8.7.E.*6$

Complete the proof of Lemma 3.

Exercise $8.7.E.*7$

Complete the proof of Theorem 3.

Exercise $8.7.E.*8$

Do Problem 1 and $2'$ for $f: S \rightarrow E$.

Exercise $8.7.E.*9$

Prove formula (1) from definitions of Part II of this section.

Exercise $8.7.E.10$

$\Rightarrow$ 10. Show that

$$\left|\int_{A} f\right| \leq \int_{A}|f| \tag{8.7.E.3}$$

for integrable maps $f: S \rightarrow E$. See also Problem 14.
[Hint: If $mA<\infty$, use Corollary $1($ ii $)$ of §4 and Lemma $1.$ If $mA=\infty,$ us imitate" the proof of Lemma $3.$]

Exercise $8.7.E.11$

Do Problem 11 in §6 for $f_{n}: S \rightarrow E.$ Do it componentwise for $E=E^{n}\left(C^{n}\right).$

Exercise $8.7.E.12$

Show that if $f, g: S \rightarrow E^{1}(C)$ are integrable on A, then

$$\left|\int_{A} f g\right|^{2} \leq \int_{A}|f|^{2} \cdot \int_{A}|g|^{2}. \tag{8.7.E.4}$$

In what case does equality hold? Deduce Theorem $4\left(\mathrm{c}^{\prime}\right)$ in Chapter 3, §§1-3, from this result.
[Hint: Argue as in that theorem. Consider the case

$$\left.\left(\exists t \in E^{1}\right) \quad \int_{A}|f-t g|=0.\right] \tag{8.7.E.5}$$

Exercise $8.7.E.13$

Show that if $f: S \rightarrow E^{1}(C)$ is integrable on A and

$$\left|\int_{A} f\right|=\int_{A}|f|, \tag{8.7.E.6}$$

then

$$(\exists c \in C) \quad c f=|f| \quad \text { a.e. on } A. \tag{8.7.E.7}$$

[Hint: Let $a=\int_{A} f.$ The case $a=0$ is trivial. If $a \neq 0,$ let

$$c=\frac{|a|}{a} ;|c|=1 ; c a=|a|. \tag{8.7.E.8}$$

Let $r=(c f)_{\mathrm{re}}$. Show that $r \leq|c f|=|f|$,

$$\begin{gathered}\left|\int_{A} f\right|=\int_{A} c f=\int_{A} r \leq \int_{A}|f|=\left|\int_{A} f\right|, \\ \int_{A}|f|=\int_{A} r=\int_{A}(c f)_{\mathrm{re},}\end{gathered}$$

$(c f)_{\mathrm{re}}=|c f|(\mathrm{a.\,e.}),$ and $c f=|c f|=|f|$ a.e. on $A.$]

Exercise $8.7.E.14$

Do Problem 10 for $E=C$ using the method of Problem 13.

Exercise $8.7.E.15$

Show that if $f: S \rightarrow E$ is integrable on A, it is integrable on each $\mathcal{M}$-set $B \subseteq A$. If, in addition,

$$\int_B f=0 \tag{8.7.E.9}$$

for all such B, show that $f=0$ a.e. on A. Prove it for $E=E^n$ first.
[Hint for $E=E^*: A=A(f>0) \cup A(f \leq 0)$. Use Theorems $1(\mathrm{h})$ and 2 from $\S 5$.]

Exercise $8.7.E.16$

In Problem 15, show that

$$s=\int f \tag{8.7.E.10}$$

is a σ-additive set function on

$$\mathcal{M}_A=\{X \in \mathcal{M} | X \subseteq A\}. \tag{8.7.E.11}$$

(Note 4 in $\S 5$); s is called the indefinite integral of f in A.

8.8: Product Measures. Iterated Integrals

Let $(X, \mathcal{M}, m)$ and $(Y, \mathcal{N}, n)$ be measure spaces, with $X \in \mathcal{M}$ and $Y \in \mathcal{N}$. Let $\mathcal{C}$ be the family of all "rectangles," i.e., sets

$$A \times B, \tag{8.8.1}$$

with $A \in \mathcal{M}, B \in \mathcal{N}, mA < \infty,$ and $nB < \infty$.

Define a premeasure $s : \mathcal{C} \rightarrow E^1$ by

$$s(A \times B) = mA \cdot nB, \quad A \times B \in \mathcal{C}. \tag{8.8.2}$$

Let p^* be the s-induced outer measure in $X \times Y$ and

$$p : \mathcal{P}^* \rightarrow E^* \tag{8.8.3}$$

the p^*-induced measure ("product measure," $p = m \times n$) on the σ-field $\mathcal{P}^*$ of all p^*-measurable sets in $X \times Y$ (Chapter 7, §§5-6).

We consider functions $f : X \times Y \rightarrow E^*$ (extended-real).

I. We begin with some definitions.

Definitions

(1) Given a function $f : X \rightarrow Y \rightarrow E^*$ (of two variables x, y), let f_x or $f(x, \cdot)$ denote the function on Y given by

$$f_x(y) = f(x, y); \tag{8.8.4}$$

it arises from f by fixing x.

Similarly, f^y or $f(\cdot, y)$ is given by $f^y(x) = f(x, y)$.

(2) Define $g : X \rightarrow E^*$ by

$$g(x) = \int_Y f(x, \cdot) dn, \tag{8.8.5}$$

and set

$$\int_X \int_Y f dn dm = \int_X g dm, \tag{8.8.6}$$

also written

$$\int_X dm(x) \int_Y f(x, y) dn(y). \tag{8.8.7}$$

This is called the iterated integral of f on Y and X, in this order.

Similarly,

$$h(y) = \int_X f^y dm \tag{8.8.8}$$

and

$$\int_Y \int_X f dm dn = \int_Y h dn. \tag{8.8.9}$$

Note that by the rules of §5, these integrals are always defined.

(3) With f, g, h as above, we say that f is a Fubini map or has the Fubini properties (after the mathematician Fubini) iff

(a) g is m-measurable on X and h is n-measurable on Y;

(b) f_x is n-measurable on Y for almost all x (i.e., for $x \in X - Q$, $mQ = 0$); f^y is m-measurable on X for $y \in Y - Q', nQ' = 0;$ and

(c) the iterated integrals above satisfy

$$\int_X \int_Y f dndm = \int_Y \int_X f dmdn = \int_{X\times Y} f dp \tag{8.8.10}$$

(the main point).

For monotone sequences

$$f_k : X\times Y \rightarrow E^* \quad (k=1,2,\ldots), \tag{8.8.11}$$

we now obtain the following lemma.

Lemma 8.8.1

If $0 \leq f_k \nearrow f$ (pointwise) on $X\times Y$ and if each f_k has Fubini property (a), (b), or (c), then f has the same property.

Proof

For $k=1,2,\ldots,$ set

$$g_k(x) = \int_Y f_k(x,\cdot)dn \tag{8.8.12}$$

and

$$h_k(y) = \int_X f_k(\cdot,y)dm. \tag{8.8.13}$$

By assumpsion,

$$0 \leq f_k(x,\cdot) \nearrow f(x,\cdot) \tag{8.8.14}$$

pointwise on Y. Thus by Theorem 4 in §6,

$$\int_Y f_k(x,\cdot) \nearrow \int_Y f(x,\cdot)dn, \tag{8.8.15}$$

i.e., $g_k \nearrow g$ (pointwise) on X, with g as in Definition 2.

Again, by Theorem 4 of §6,

$$\int_X g_k dm \nearrow \int_X g dm; \tag{8.8.16}$$

or by Definition 2,

$$\int_X \int_Y f dndm = \lim_{k\rightarrow\infty} \int_X \int_Y f_k dndm. \tag{8.8.17}$$

Similarly for

$$\int_Y \int_X f dmdn \tag{8.8.18}$$

and

$$\int_{X\times Y} f dp. \tag{8.8.19}$$

Hence f satisfies (c) if all f_k do.

Next, let f_k have property (b); so $(\forall k) f_k(x,\cdot)$ is n-measurable on Y if $x\in X-Q_k$ $(mQ_k=0)$. Let

$$Q = \bigcup_{k=1}^{\infty} Q_k; \tag{8.8.20}$$

so $mQ=0$, and all $f_k(x,\cdot)$ are n-measurable on Y, for $x\in X-Q$. Hence so is

$$f(x,\cdot)=\lim_{k\to\infty} f_k(x,\cdot). \tag{8.8.21}$$

Similarly for $f(\cdot,y)$. Thus f satisfies (b).

Property (a) follows from $g_k\to g$ and $h_k\to h$. $\square$

Using Problems 9 and 10 from §6, the reader will also easily verify the following lemma.

Lemma 8.8.2

(i) If f_1 and f_2 are nonnegative, p-measurable Fubini maps, so is af_1+bf_2 for $a,b\geq 0$.

(ii) If, in addition,

$$\int_{X\times Y} f_1 dp<\infty \text{ or } \int_{X\times Y} f_2 dp<\infty, \tag{8.8.22}$$

then f_1-f_2 is a Fubini map, too

Lemma 8.8.3

Let $f=\sum_{i=1}^{\infty} f_i$ (pointwise), with $f_i\geq 0$ on $X\times Y$.

(i) If all f_i are p-measurable Fubini maps, so is f.

(ii) If the f_i have Fubini properties (a) and (b), then

$$\int_X\int_Y f\,dndm=\sum_{i=1}^{\infty}\int_X\int_Y f_i\,dndm \tag{8.8.23}$$

and

$$\int_Y\int_X f\,dmdn=\sum_{i=1}^{\infty}\int_Y\int_X f_i\,dmdn. \tag{8.8.24}$$

II. By Theorem 4 of Chapter 7, §3, the family $\mathcal{C}$ (see above) is a semiring, being the product of two rings,

$$\{A\in\mathcal{M}\,|\,mA<\infty\} \text{ and } \{B\in\mathcal{N}\,|\,nB<\infty\}. \tag{8.8.25}$$

(Verify!) Thus using Theorem 2 in Chapter 7, §6, we now show that p is an extension of $s:\mathcal{C}\to E^1$.

Theorem 8.8.1

The product premeasure s is σ-additive on the semiring $\mathcal{C}$. Hence

(i) $\mathcal{C}\subseteq\mathcal{P}^*$ and $p=s<\infty$ on $\mathcal{C}$, and

(ii) the characteristic function C_D of any set $D\in\mathcal{C}$ is a Fubini map.

Proof

Let $D=A\times B\in\mathcal{C}$; so

$$C_D(x,y)=C_A(x)\cdot C_B(y). \tag{8.8.26}$$

(Why?) Thus for a fixed x, $C_D(x,\cdot)$ is just a multiple of the $\mathcal{N}$-simple map C_B, hence n-measurable on Y. Also,

$$g(x)=\int_Y C_D(x,\cdot)dn=C_A(x)\cdot\int_Y C_B dn=C_A(x)\cdot nB; \tag{8.8.27}$$

so $g=C_A\cdot nB$ is $\mathcal{M}$-simple on X, with

$$\int_X \int_Y C_D dndm = \int_X gdm = nB \int_X C_A dm = nB \cdot mA = sD. \tag{8.8.28}$$

Similarly for $C_D(\cdot, y)$, and

$$h(y) = \int_X C_D(\cdot, y) dm. \tag{8.8.29}$$

Thus C_D has Fubini properties (a) and (b), and for every $D \in \mathcal{C}$

$$\int_X \int_Y C_D dndm = \int_Y \int_X C_D dmdn = sD. \tag{8.8.30}$$

To prove σ-additivity, let

$$D = \bigcup_{i=1}^{\infty} D_i \text{ (disjoint)}, D_i \in \mathcal{C}; \tag{8.8.31}$$

so

$$C_D = \sum_{i=1}^{\infty} C_{D_i}. \tag{8.8.32}$$

(Why?) As shown above, each C_{D_i} has Fubini properties (a) and (b); so by (1) and Lemma 3,

$$sD = \int_X \int_Y C_D dndm = \sum_{i=1}^{\infty} \int_X \int_Y C_{D_i} dndm = \sum_{i=1}^{\infty} sD_i, \tag{8.8.33}$$

as required.

Assertion (i) now follows by Theorem 2 in Chapter 7, §6. Hence

$$sD = pD = \int_{X \times Y} C_D dp; \tag{8.8.34}$$

so by formula (1), C_D also has Fubini property (c), and all is proved. □

Next, let $\mathcal{P}$ be the σ-ring generated by the semiring $\mathcal{C}$ (so $\mathcal{C} \subseteq \mathcal{P} \subseteq \mathcal{P}^*$).

Lemma 8.8.4

$\mathcal{P}$ is the least set family $\mathcal{R}$ such that

(i) $\mathcal{R} \supseteq \mathcal{C}$;

(ii) $\mathcal{R}$ is closed under countable disjoint unions; and

(iii) $H - D \in \mathcal{R}$ if $D \in \mathcal{R}$ and $D \subseteq H, H \in \mathcal{C}$.

This is simply Theorem 3 in Chapter 7, §3, with changed notation.

Lemma 8.8.5

If $D \in \mathcal{P}$ (σ-generated by $\mathcal{C}$), then C_D is a Fubini map.

Proof

Let $\mathcal{R}$ be the family of all $D \in \mathcal{P}$ such that C_D is a Fubini map. We shall show that $\mathcal{R}$ satisfies (i)-(iii) of Lemma 4, and so $\mathcal{P} \subseteq \mathcal{R}$.

(ii) Let

$$D=\bigcup_{i=1}^{\infty} D_{i} \text{ (disjoint)}, \quad D_{i} \in \mathcal{R}. \tag{8.8.35}$$

Then

$$C_{D}=\sum_{i=1}^{\infty} C_{D_{i}}, \tag{8.8.36}$$

and each C_{D_i} is a Fubini map. Hence so is C_D by Lemma 3. Thus $D \in \mathcal{R}$, proving (ii).

(iii) We must show that C_{H-D} is a Fubini map if C_D is and if $D \subseteq H, H \in \mathcal{C}$. Now, $D \subseteq H$ implies

$$C_{H-D}=C_{H}-C_{D}. \tag{8.8.37}$$

(Why?) Also, by Theorem 1, $H \in \mathcal{C}$ implies

$$\int_{X \times Y} C_{H} dp=pH=sH<\infty, \tag{8.8.38}$$

and C_H is a Fubini map. So is C_D by assumption. So also is

$$C_{H-D}=C_{H}-C_{D} \tag{8.8.39}$$

by Lemma 2(ii). Thus $H-D \in \mathcal{R}$, proving (iii).

By Lemma 4, then, $\mathcal{P} \subseteq \mathcal{R}$. Hence $(\forall D \in \mathcal{P}) C_D$ is a Fubini map. $\square$

We can now establish one of the main theorems, due to Fubini.

Theorem 8.8.2 (Fubini I)

Suppose $f: X \times Y \rightarrow E^{*}$ is $\mathcal{P}$-measurable on $X \times Y$ ($\mathcal{P}$ as above) rom. Then f is a Fubini map if either

(i) $f \geq 0$ on $X \times Y$, or

(ii) one of

$$\int_{X \times Y}|f| dp, \int_{X} \int_{Y}|f| dndm, or \int_{Y} \int_{X}|f| dmdn \tag{8.8.40}$$

is finite.

In both cases,

$$\int_{X} \int_{Y} f dndm=\int_{Y} \int_{X} f dmdn=\int_{X \times Y} f dp. \tag{8.8.41}$$

Proof

First, let

$$f=\sum_{i=1}^{\infty} a_{i} C_{D_{i}} \quad\left(a_{i} \geq 0, D_{i} \in \mathcal{P}\right), \tag{8.8.42}$$

i.e., f is $\mathcal{P}$-elementary, hence certainly p-measurable. (Why?) By Lemmas 5 and 2, each $a_i C_{D_i}$ is a Fubini map. Hence so is f (Lemma 3). Formula (2) is simply Fubini property (c).

Now take any $\mathcal{P}$-measurable $f \geq 0$. By Lemma 2 in §2,

$$f=\lim_{k \rightarrow \infty} f_{k} \text{ on } X \times Y \tag{8.8.43}$$

for some sequence $\{f_k\} \uparrow$ of $\mathcal{P}$-elementary maps, $f_k \geq 0$. As shown above, each f_k is a Fubini map. Hence so is f by Lemma 1. This settles case (i).

Next, assume (ii). As f is $\mathcal{P}$-measurable, so are f^+, f_-, and $|f|$ (Theorem 2 in §2). As they are nonnegative, they are Fubini maps by case (i).

So is $f = f^+ - f^-$ by Lemma 2(ii), since $f^+ \leq |f|$ implies

$$\int_{X \times Y} f^+ dp < \infty \tag{8.8.44}$$

by our assumption (ii). (The three integrals are equal, as $|f|$ is a Fubini map.)

Thus all is proved. □

III. We now want to replace $\mathcal{P}$ by $\mathcal{P}^*$ in Lemma 5 and Theorem 2. This works only under certain σ-finiteness conditions, as shown below.

Lemma 8.8.6

Let $D \in \mathcal{P}^*$ be σ-finite, i.e.,

$$D = \bigcup_{i=1}^{\infty} D_i \text{ (disjoint)} \tag{8.8.45}$$

for some $D_i \in \mathcal{P}^*$, with $pD_i < \infty$ $(i = 1, 2, \ldots)$.

Then there is a$K \in \mathcal{P}$ such that $p(K - D) = 0$ and $D \subseteq K$.

Proof

As $\mathcal{P}$ is a σ-ring containing $\mathcal{C}$, it also contains $\mathcal{C}_\sigma$. Thus by Theorem 3 of Chapter 7, §5, p^* is $\mathcal{P}$-regular.

For the rest, proceed as in Theorems 1 and 2 in Chapter 7, §7. □

Lemma 8.8.7

If $D \in \mathcal{P}^*$ is σ-finite (Lemma 6), then C_D is a Fubini map.

Proof

By Lemma 6,

$$(\exists K \in \mathcal{P}) \quad p(K - D) = 0, D \subseteq K. \tag{8.8.46}$$

Let $Q = K - D$, so $pQ = 0$, and $C_Q = C_K - C_D$; that is, $C_D = C_K - C_Q$ and

$$\int_{X \times Y} C_Q dp = pQ = 0. \tag{8.8.47}$$

As $K \in \mathcal{P}, C_K$ is a Fubini map. Thus by Lemma 2(ii), all reduces to proving the same for C_Q.

Now, as $pQ = 0, Q$ is certainly σ-finite; so by Lemma 6,

$$(\exists Z \in \mathcal{P}) \quad Q \subseteq Z, pZ = pQ = 0. \tag{8.8.48}$$

Again C_Z is a Fubini map; so

$$\int_X \int_Y C_Z dn dm = \int_{X \times Y} C_Z dp = pZ = 0. \tag{8.8.49}$$

As $Q \subseteq Z$, we have $C_Q \leq C_Z$, and so

$$\int_X \int_Y C_Q dndm = \int_X \left[\int_Y C_Q(x,\cdot)dn\right] dm$$
$$\leq \int_X \left[\int_Y C_Z(x,\cdot)dn\right] dm = \int_{X\times Y} C_Z dp = 0.$$

Similarly,

$$\int_Y \int_X C_Q dmdn = \int_Y \left[\int_X C_Q(\cdot,y)dm\right] dn = 0. \tag{8.8.50}$$

Thus setting

$$g(x) = \int_Y C_Q(x,\cdot)dn \text{ and } h(y) = \int_X C_Q(\cdot,y)dm, \tag{8.8.51}$$

we have

$$\int_X gdm = 0 = \int_Y hdn. \tag{8.8.52}$$

Hence by Theorem 1(h) in §5, $g=0$ a.e. on X, and $h=0$ a.e. on Y. So g and h are "almost" measurable (Definition 2 of §3); i.e., C_Q has Fubini property (a).

Similarly, one establishes (b), and (3) yields Fubini property (c), since

$$\int_X \int_Y C_Q dndm = \int_Y \int_X C_Q dmdn = \int_{X\times Y} C_Q dp = 0, \tag{8.8.53}$$

as required. □

Theorem 8.8.3 (Fubini II)

Suppose $f : X \times Y \rightarrow E^*$ is $\mathcal{P}^*$-measurable on $X \times Y$ and satisfies condition (i) or (ii) of Theorem 2.

Then f is a Fubini map, provided f has σ-finite support, i.e., f vanishes outside some σ-finite set $H \subseteq X \times Y$.

Proof

First, let

$$f = \sum_{i=1}^{\infty} a_i C_{D_i} \quad (a_i > 0, D_i \in \mathcal{P}^*), \tag{8.8.54}$$

with $f=0$ on $-H$ (as above).

As $f = a_i \neq 0$ on A_i, we must have $D_i \subseteq H$; so all D_i are σ-finite. (Why?) Thus by Lemma 7, each C_{D_i} is a Fubini map, and so is f. (Why?)

If f is $\mathcal{P}^*$-measurable and nonnegative, and $f=0$ on $-H$, we can proceed as in Theorem 2, making all f_k vanish on $-H$. Then the f_k and f are Fubini maps by what was shown above.

Finally, in case (ii), $f=0$ on $-H$ implies

$$f^+ = f^- = |f| = 0 \text{ on } -H. \tag{8.8.55}$$

Thus f^+, f^-, and f are Fubini maps by part (i) and the argument of Theorem 2. □

Note 1. The σ-finite support is automatic if f is p-integrable (Corollary 1 in §5), or if p or both m and n are σ-finite (see Problem 3). The condition is also redundant if f is $\mathcal{P}$-measurable (Theorem 2; see also Problem 4).

Note 2. By induction, our definitions and Theorems 2 and 3 extend to any finite number q of measure spaces

$$(X_i, \mathcal{M}_i, m_i), \quad i = 1, \ldots, q. \tag{8.8.56}$$

One writes

$$p=m_1\times m_2 \tag{8.8.57}$$

if $q=2$ and sets

$$m_1\times m_2\times\cdots\times m_{q+1}=(m_1\times\cdots\times m_q)\times m_{q+1}. \tag{8.8.58}$$

Theorems 2 and 3 with similar assumptions then state that the order of integrations is immaterial.

Note 3. Lebesgue measure in E^q can be treated as the product of q one- dimensional measures. Similarly for LS product measures (but this method is less general than that described in Problems 9 and 10 of Chapter 7, §9).

IV. Theorems 2(ii) and 3(ii) hold also for functions

$$f:X\times Y\rightarrow E^n\left(C^n\right) \tag{8.8.59}$$

if Definitions 2 and 3 are modified as follows (so that they make sense for such maps): In Definition 2, set

$$g(x)=\int_Y f_x dn \tag{8.8.60}$$

if f_x is n-integrable on Y, and $g(x)=0$ otherwise. Similarly for $h(y)$. In Definition 3, replace "measurable" by "integrable."

For the proof of the theorems, apply Theorems 2(i) and 3(i) to $|f|$. This yields

$$\int_Y\int_X|f|dmdn=\int_X\int_Y|f|dndm=\int_{X\times Y}|f|dp. \tag{8.8.61}$$

Hence if one of these integrals is finite, f is p-integrable on $X\times Y$, and so are its q components. The result then follows on noting that f is a Fubini map (in the modified sense) iff its components are. (Verify!) See also Problem 12 below.

V. In conclusion, note that integrals of the form

$$\int_D f dp \quad (D\in\mathcal{P}^*) \tag{8.8.62}$$

reduce to

$$\int_{X\times Y} f\cdot C_D dp. \tag{8.8.63}$$

Thus it suffices to consider integrals over $X\times Y$.

8.8.E: Problems on Product Measures and Fubini Theorems

Exercise $8.8.E.1$

Prove Lemmas 2 and 3.

Exercise $8.8.E.1'$

Show that $\{A \in \mathcal{M} | mA < \infty\}$ is a set ring.

Exercise $8.8.E.2$

Fill in all proof details in Theorems 1 to 3.

Exercise $8.8.E.2'$

Do the same for Lemmas 5 to 7.

Exercise $8.8.E.3$

Prove that if m and n are σ-finite, so is $p = m \times n$. Disprove the converse by an example.
[Hint: $(\cup_i A_i) \times (U_j B_j) = U_{i,j}(A_i \times B_j)$. Verify!]

Exercise $8.8.E.4$

Prove the following.
(i) Each $D \in \mathcal{P}$ (as in the text) is (p) σ-finite.
(ii) All $\mathcal{P}$-measurable maps $f : X \times Y \rightarrow E^*$ have σ-finite support.
[Hints: (i) Use Problem $14(\mathrm{b})$ from Chapter 7, §3. (ii) Use (i) for $\mathcal{P}$-elementary and nonnegative maps first.]

Exercise $8.8.E.5$

(i) Find $D \in \mathcal{P}^*$ and $x \in X$ such that $C_D(x, \cdot)$ is not n-measurable on Y. Does this contradict Lemma 7?
[Hint: Let $m = n =$ Lebesgue measure in $E^1; D = \{x\} \times Q,$ with Q non-measurable.]
(ii) Which $\mathcal{C}$-sets have nonzero measure if $X = Y = E^1, m^*$ is as in Problem $2(b)$ of Chapter $7, §5($ with $S = X),$ and n is Lebesgue measure?

Exercise $8.8.E.5'$

Let $m = n =$ Lebesgue measure in $[0, 1] = X = Y$. Let

$$f_k = \begin{cases} k(k+1) & \text{on } \left(\frac{1}{k+1}, \frac{1}{k}\right] \text{ and} \\ 0 & \text{elsewhere.} \end{cases} \tag{8.8.E.1}$$

Let

$$f(x, y) = \sum_{k=1}^{\infty} \left[f_k(x) - f_{k+1}(x)\right] f_k(y); \tag{8.8.E.2}$$

the series converges. (Why?) Show that
(i) $(\forall k) \int_X f_k = 1$;

(ii) $\int_X \int_Y f dndm = 1 \neq 0 = \int_Y \int_X f dmdn$.
What is wrong? Is f $\mathcal{P}$-measurable?
[Hint: Explore

$$\left.\int_X \int_Y |f| dndm.\right] \tag{8.8.E.3}$$

Exercise $8.8.E.6$

Let $X = Y = [0,1], m$ as in Example (c) of Chapter 7, §6, $(S = X)$ and $n =$ Lebesgue measure in Y.
(i) Show that $p = m \times n$ is a topological measure under the standard metric in E^2.
(ii) Prove that $D = \{(x,y) \in X \times Y | x = y\} \in \mathcal{P}^*$.
(iii) Describe $\mathcal{C}$.
[Hints: (i) Any subinterval of $X \times Y$ is in $\mathcal{P}^*$; (ii) D is closed. Verify!]

Exercise $8.8.E.7$

Continuing Problem 6, let $f = C_D$.
(i) Show that

$$\int_Y \int_X f dndm = 0 \neq 1 = \int_Y \int_X f dmdn. \tag{8.8.E.4}$$

What is wrong?
[Hint: D is not σ-finite; for if

$$D = \bigcup_{i=1}^{\infty} D_i, \tag{8.8.E.5}$$

at least one D_{i} is uncountable and has no finite basic covering values (why?), so $p^* D_{\mathrm{i}} = \infty.$]
(ii) Compute $p^*\{(x,0) | x \in X\}$ and $p^*\{(0,y) | y \in Y\}$.

Exercise $8.8.E.8$

Show that $D \in \mathcal{P}^*$ is σ-finite iff

$$D \subseteq \bigcup_{i=1}^{\infty} D_i (\text{disjoint}) \tag{8.8.E.6}$$

for some sets $D_i \in \mathcal{C}$.
[Hint: First let $p^* D < \infty.$ Use Corollary 1 from Chapter $7, §1.$]

Exercise $8.8.E.9$

Given $D \in \mathcal{P}, a \in X,$ and $b \in Y$, let

$$D_a = \{y \in Y | (a,y) \in D\} \tag{8.8.E.7}$$

and

$$D^b = \{x \in X | (x,b) \in D\}. \tag{8.8.E.8}$$

(See Figure 34 for $X=Y=E^1$.)
Prove that
(i) $D_a \in \mathcal{N}, D^b \in \mathcal{M}$;
(ii) $C_D(a, \cdot)=C_{D_a}, nD_a=\int_Y C_D(a, \cdot) dn, mD^b=\int_X C_D(\cdot, b) dm$.
[Hint: Let

$$H=\left\{(x, y) \in E^2 \mid 0 \leq y<f(x)\right\} \tag{8.8.E.9}$$

Show that $\mathcal{R}$ is a σ-ring $\supseteq C$. Hence $\mathcal{R} \supseteq \mathcal{P}$; $D \in \mathcal{R}$; $D_a \in \mathcal{N}$. Similariy for D^b.]

Exercise $8.8.E.10$

$\Rightarrow$ 10. Let $m=n=$ Lebesgue measure in $E^1=X=Y$. Let $f: E^1 \rightarrow[0, \infty)$ be m-mensurable on X. Let

$$H=\left\{(x, y) \in E^2 \mid 0 \leq y<f(x)\right\} \tag{8.8.E.10}$$

and

$$G=\left\{(x, y) \in E^2 \mid y=f(x, y)\right\} \tag{8.8.E.11}$$

(the "graph" of f). Prove that
(i) $H \in \mathcal{P}^*$ and

$$p H=\int_X f d m \tag{8.8.E.12}$$

(="the area under f")
(ii) $G \in \mathcal{P}^*$ and $p G=0$.
[Hints: (i) First take $f=C_D$, and elementary and nonnegative maps. Then use Lemma 2 in §2 (last clause). Fix elementary and nonnegative maps $f_k \nearrow f$, assuming $f_k<f$ (if not, replace f_k by $\left(1-\frac{1}{k}\right) f_k$). Let

$$H_k=\left\{(x, y) \mid 0 \leq y<f_k(x)\right\}. \tag{8.8.E.13}$$

Show that $H_k \nearrow H \in \mathcal{P}^*$.
(ii) Set

$$\phi(x, y)=y-f(x). \tag{8.8.E.14}$$

Using Corollary 4 of §1, show that ϕ is p-measurable on E^2; so $G=E^2(\phi=0) \in \mathcal{P}^*$. Dropping a null set (Lemma 6), assume $G \in \mathcal{P}$. By Problem 9 (ii),

$$\left(\forall x \in E^1\right) \quad \int_Y C_G(x, \cdot) d n=n G_x=0, \tag{8.8.E.15}$$

as $G_x=\{f(x)\}$, a singleton.]

Exercise $8.8.E.11$

Let

$$f(x, y)=\phi_1(x) \phi_2(y). \tag{8.8.E.16}$$

Prove that if ϕ_1 is m-integrable on X and ϕ_2 is n-integrable on Y, then f is p-integrable on $X \times Y$ and

$$\int_{X\times Y} f\,dp = \int_X \phi_1 \cdot \int_Y \phi_2. \tag{8.8.E.17}$$

Exercise 8.8.E. *12

Prove Theorem 3(ii) for $f : X \times Y \rightarrow E(E \text{ complete})$.
[Outline: If f is $\mathcal{P}^*$-simple, use Lemma 7 above and Theorem 2 in §7.
If

$$f = \sum_{k=1}^{\infty} a_k C_{D_k}, \quad D_k \in \mathcal{P}^*, \tag{8.8.E.18}$$

let

$$H_k = \bigcup_{i=1}^{k} D_i \tag{8.8.E.19}$$

and $f_k = fC_{H_k}$, so the f_k are $\mathcal{P}^*$-simple (hence Fubini maps), and $f_k \rightarrow f$ (point-wise) on $X \times Y$, with $|f_k| \leq |f|$ and

$$\int_{X\times Y} |f|\,dp < \infty \tag{8.8.E.20}$$

(by assumption). Now use Theorem 5 from §6.
Let now f be $\mathcal{P}^*$-measurable; so

$$f = \lim_{k\rightarrow\infty} f_k \text{ (uniformly)} \tag{8.8.E.21}$$

for some $\mathcal{P}^*$-elementary maps g_k (Theorem 3 in §1). By assumption, $f = fC_H(H\ \sigma$-finite); so we may assume $g_k = g_k C_H$. Then as shown above, all g_k are Fubini maps. So is f by Lemma 1 in §7 (verify!), provided $H \subseteq D$ for some $D \in \mathcal{C}$.
In the general case, by Problem 8 ,

$$H \subseteq \bigcup_i D_i \text{(disjoint)}, D_i \in \mathcal{C}. \tag{8.8.E.22}$$

Let $H_i = H \cap D_i$. By the previous step, each fC_{H_i} is a Fubini map; so is

$$f_k = \sum_{i=1}^{k} fC_{H_i} \tag{8.8.E.23}$$

(why?), hence so is $f = \lim_{k\rightarrow\infty} f_k$, by Theorem 5 of §6. (Verify!)]

Exercise 8.8.E. 13

Let $m =$ Lebesgue measure in E^1, $p =$ Lebesgue measure in E^s, $X = (0, \infty)$, and

$$Y = \{\bar{y} \in E^s \,|\, |\bar{y}| = 1\}. \tag{8.8.E.24}$$

Given $\bar{x} \in E^s - \{\bar{0}\}$, let

$$r = |\bar{x}| \text{ and } \bar{u} = \frac{\bar{x}}{r} \in Y. \tag{8.8.E.25}$$

Call r and $\bar{u}$ the polar coordinates of $\bar{x} \neq \bar{0}$.
If $D \subseteq Y$, set

$$n^* D = s \cdot p^* \{r\bar{u} | \bar{u} \in D, 0 < r \leq 1\}. \tag{8.8.E.26}$$

Show that n^* is an outer measure in Y; so it induces a measure n in Y.
Then prove that

$$\int_{E^s} f dp = \int_X r^{s-1} dm(r) \int_Y f(r\bar{u}) dn(\bar{u}) \tag{8.8.E.27}$$

if f is p-measurable and nonnegative on E^s.
[Hint: Start with $f = C_A$,

$$A = \{r\bar{u} | \bar{u} \in H, a < r < b\}, \tag{8.8.E.28}$$

for some open set $H \subseteq Y$ (subspace of E^s). Next, let $A \in \mathcal{B}($ Borel set in $Y)$; then $A \subseteq \mathcal{P}^*$. Then let f be p-elementary, and so on.]

8.9: Riemann Integration. Stieltjes Integrals

I. In this section, $\mathcal{C}$ is the family of all intervals in E^n, and m is an additive finite premeasure on $\mathcal{C}$ (or $\mathcal{C}_s$), such as the volume function v (Chapter 7, §§1-2).

By a $\mathcal{C}$-partition of $A \in \mathcal{C}$ (or $A \in \mathcal{C}_s$), we mean a finite family

$$\mathcal{P}=\{A_i\} \subset \mathcal{C} \tag{8.9.1}$$

such that

$$A=\bigcup_i A_i \text{ (disjoint). } \tag{8.9.2}$$

As we noted in §5, the Riemann integral,

$$R \int_A f=R \int_A f d m, \tag{8.9.3}$$

of $f: E^n \rightarrow E^1$ can be defined as its Lebesgue counterpart,

$$\int_A f, \tag{8.9.4}$$

with elementary maps replaced by simple step functions ("$\mathcal{C}$-simple" maps.) Equivalently, one can use the following construction, due to J. G. Darboux.

Definitions

(a) Given $f: E^n \rightarrow E^*$ and a $\mathcal{C}$-partition

$$\mathcal{P}=\{A_1, \ldots, A_q\} \tag{8.9.5}$$

of A, we define the lower and upper Darboux sums, $\underline{S}$ and $\overline{S}$, of f over $\mathcal{P}$ (with respect to m) by

$$\underline{S}(f, \mathcal{P})=\sum_{i=1}^q m A_i \cdot \inf f[A_i] \text{ and } \overline{S}(f, \mathcal{P})=\sum_{i=1}^q m A_i \cdot \sup f[A_i]. \tag{8.9.6}$$

(b) The lower and upper Riemann integrals ("R-integrals") of f on A (with respect to m) are

$$\left.\begin{array}{l} R \underline{\int}_A f=R \underline{\int}_A f d m=\sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) \text{ and } \\ R \overline{\int}_A f=R \overline{\int}_A f d m=\inf_{\mathcal{P}} \overline{S}(f, \mathcal{P}), \end{array}\right\} \tag{8.9.7}$$

where the "inf" and "sup" are taken over all $\mathcal{C}$-partitions $\mathcal{P}$ of A.

(c) We say that f is Riemann-integrable ("R-integrable") with respect to m on A iff f is bounded on A and

$$R \underline{\int}_A f=R \overline{\int}_A f. \tag{8.9.8}$$

We then set

$$R \int_A f=R \underline{\int}_A f=R \overline{\int}_A f d m=R \int_A f d m \tag{8.9.9}$$

and call it the Riemann integral ("R-integral") of f on A. "Classical" notation:

$$R \int_A f(\bar{x}) d m(\bar{x}). \tag{8.9.10}$$

If $A=[a, b] \subset E^1$, we also write

$$R \int_a^b f = R \int_a^b f(x) dm(x) \tag{8.9.11}$$

instead.

If m is Lebesgue measure (or premeasure) in E^1, we write "dx" for "$dm(x)$."

For Lebesgue integrals, we replace "R" by "L," or we simply omit "R."

If f is R-integrable on A, we also say that

$$R \int_A f \tag{8.9.12}$$

exists (note that this implies the boundedness of f); note that

$$R \underline{\int}_A f \text{ and } R \overline{\int}_A f \tag{8.9.13}$$

are always defined in E^*.

Below, we always restrict f to a fixed $A \in \mathcal{C}$ (or $A \in \mathcal{C}_s$); $\mathcal{P}, \mathcal{P}', \mathcal{P}'', \mathcal{P}^*$ and $\mathcal{P}_k$ denote $\mathcal{C}$-partitions of A.

We now obtain the following result for any additive $m : \mathcal{C} \rightarrow [0, \infty)$.

Corollary 8.9.1

If $\mathcal{P}$ refines $\mathcal{P}'$ (§1), then

$$\underline{S}\left(f, \mathcal{P}'\right) \leq \underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}) \leq \overline{S}\left(f, \mathcal{P}'\right). \tag{8.9.14}$$

Proof

Let $\mathcal{P}' = \left\{A_i\right\}, \mathcal{P} = \left\{B_{ik}\right\}$, and

$$(\forall i) \quad A_i = \bigcup_k B_{ik}. \tag{8.9.15}$$

By additivity,

$$m A_i = \sum_k m B_{ik}. \tag{8.9.16}$$

Also, $B_{ik} \subseteq A_i$ implies

$$\begin{aligned} f\left[B_{ik}\right] &\subseteq f\left[A_i\right]; \\ \sup f\left[B_{ik}\right] &\leq \sup f\left[A_i\right]; \text{ and} \\ \inf f\left[B_{ik}\right] &\geq \inf f\left[A_i\right]. \end{aligned}$$

So setting

$$a_i = \inf f\left[A_i\right] \text{ and } b_{ik} = \inf f\left[B_{ik}\right], \tag{8.9.17}$$

we get

$$\begin{aligned} \underline{S}\left(f, \mathcal{P}'\right) = \sum_i a_i m A_i &= \sum_i \sum_k a_i m B_{ik} \\ &\leq \sum_{i,k} b_{ik} m B_{ik} = \underline{S}(f, \mathcal{P}). \end{aligned}$$

Similarly,

$$\overline{S}\left(f, \mathcal{P}'\right) \leq \overline{S}(f, \mathcal{P}), \tag{8.9.18}$$

and

$$\underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}) \tag{8.9.19}$$

is obvious from (1). □

Corollary 8.9.2

For any $\mathcal{P}'$ and $\mathcal{P}''$,

$$\underline{S}(f, \mathcal{P}') \leq \overline{S}(f, \mathcal{P}''). \tag{8.9.20}$$

Hence

$$R \underline{\int}_A f \leq R \overline{\int}_A f. \tag{8.9.21}$$

Proof

Let $\mathcal{P} = \mathcal{P}' \cap \mathcal{P}''$ (see §1). As $\mathcal{P}$ refines both $\mathcal{P}'$ and $\mathcal{P}''$, Corollary 1 yields

$$\underline{S}(f, \mathcal{P}') \leq \underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}''). \tag{8.9.22}$$

Thus, indeed, no lower sum $\underline{S}(f, \mathcal{P}')$ exceeds any upper sum $\overline{S}(f, \mathcal{P}'')$.

Hence also,

$$\sup_{\mathcal{P}'} \underline{S}(f, \mathcal{P}') \leq \inf_{\mathcal{P}''} \overline{S}(f, \mathcal{P}''), \tag{8.9.23}$$

i.e.,

$$R \underline{\int}_A f \leq R \overline{\int}_A f, \tag{8.9.24}$$

as claimed. □

Lemma 8.9.1

A map $f : A \rightarrow E^1$ is R-integrable iff f is bounded and, moreover,

$$(\forall \varepsilon > 0)\ (\exists \mathcal{P}) \quad \overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) < \varepsilon. \tag{8.9.25}$$

Proof

By formulas (1) and (2),

$$\underline{S}(f, \mathcal{P}) \leq R \underline{\int}_A f \leq R \overline{\int}_A f \leq \overline{S}(f, \mathcal{P}). \tag{8.9.26}$$

Hence (3) implies

$$\left| R \overline{\int}_A f - R \underline{\int}_A f \right| < \varepsilon. \tag{8.9.27}$$

As ε is arbitrary, we get

$$R \overline{\int}_A f = R \underline{\int}_A f; \tag{8.9.28}$$

so f is R-integrable.

Conversely, if so, definitions (b) and (c) imply the existence of $\mathcal{P}'$ and $\mathcal{P}''$ such that

$$\underline{S}(f, \mathcal{P}') > R\int_A f - \frac{1}{2}\varepsilon \tag{8.9.29}$$

and

$$\overline{S}(f, \mathcal{P}'') < R\int_A f + \frac{1}{2}\varepsilon. \tag{8.9.30}$$

Let $\mathcal{P}$ refine both $\mathcal{P}'$ and $\mathcal{P}''$. Then by Corollary 1,

$$\begin{aligned}\overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) &\leq \overline{S}(f, \mathcal{P}'') - \underline{S}(f, \mathcal{P}') \\ &< \left(R\int_A f + \frac{1}{2}\varepsilon\right) - \left(R\int_A f - \frac{1}{2}\varepsilon\right) = \varepsilon,\end{aligned}$$

as required. □

Lemma 8.9.2

Let f be $\mathcal{C}$-simple; say, $f = a_i$ on A_i for some $\mathcal{C}$-partition $\mathcal{P}^* = \{A_i\}$ of A (we then write

$$f = \sum_i a_i C_{A_i} \tag{8.9.31}$$

on A; see Note 4 of §4).

Then

$$R\underline{\int}_A f = R\overline{\int}_A f = \underline{S}(f, \mathcal{P}^*) = \overline{S}(f, \mathcal{P}^*) = \sum_i a_i m A_i. \tag{8.9.32}$$

Hence any finite $\mathcal{C}$-simple function is R-integrable, with $R\int_A f$ as in (4).

Proof

Given any $\mathcal{C}$-partition $\mathcal{P} = \{B_k\}$ of A, consider

$$\mathcal{P}^* \backslash\text{acdot} \mathcal{P} = \{A_i \cap B_k\}. \tag{8.9.33}$$

As $f = a_i$ on $A_i \cap B_k$ (even on all of A_i),

$$a_i = \inf f[A_i \cap B_k] = \sup f[A_i \cap B_k]. \tag{8.9.34}$$

Also,

$$A = \bigcup_{i,k} (A_i \cap B_k) \text{ (disjoint)} \tag{8.9.35}$$

and

$$(\forall i) \quad A_i = \bigcup_k (A_i \cap B_k); \tag{8.9.36}$$

so

$$mA_i = \sum_k m(A_i \cap B_k) \tag{8.9.37}$$

and

$$\underline{S}(f, \mathcal{P}) = \sum_i \sum_k a_i m(A_i \cap B_k) = \sum_i a_i m A_i = \underline{S}(f, \mathcal{P}^*) \tag{8.9.38}$$

for any such $\mathcal{P}$.

Hence also

$$\sum_i a_i mA_i = \sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) = R \underline{\int}_A f. \tag{8.9.39}$$

Similarly for $R\overline{\int}_A f$. This proves (4).

If, further, f is finite, it is bounded (by max $|a_i|$) since there are only finitely many a_i; so f is R-integrable on A, and all is proved. $\square$

Note 1. Thus $\underline{S}$ and $\overline{S}$ are integrals of $\mathcal{C}$-simple maps, and definition (b) can be restated:

$$R \underline{\int}_A f = \sup_g R \int_A g \text{ and } R \overline{\int}_A f = \inf_h R \int_A h, \tag{8.9.40}$$

taking the sup and inf over all $\mathcal{C}$-simple maps g, h with

$$g \leq f \leq h \text{ on } A. \tag{8.9.41}$$

(Verify by properties of glb and lub!)

Therefore, we can now develop R-integration as in §§4-5, replacing elementary maps by $\mathcal{C}$-simple maps, with $S = E^n$. In particular, Problem 5 in §5 works out as before.

Hence linearity (Theorem 1 of §6) follows, with the same proof. One also obtains additivity (limited to $\mathcal{C}$-partitions). Moreover, the R-integrability of f and g implies that of $fg, f \vee g, f \wedge g,$ and $|f|$. (See the Problems.)

Theorem 8.9.1

If $f_i \to f$ (uniformly) on A and if the f_i are R-integrable on A, so also is f. Moreover,

$$\lim_{i\to\infty} R \int_A |f - f_i| = 0 \text{ and } \lim_{i\to\infty} R \int_A f_i = R \int_A f. \tag{8.9.42}$$

Proof

As all f_i are bounded (definition (c)), so is f, by Problem 10 of Chapter 4, §12.

Now, given $\varepsilon > 0,$ fix k such that

$$(\forall i \geq k) \quad |f - f_i| < \frac{\varepsilon}{mA} \quad \text{on } A. \tag{8.9.43}$$

Verify that

$$(\forall i \geq k)\ (\forall \mathcal{P}) \quad |\underline{S}(f - f_i, \mathcal{P})| < \varepsilon \text{ and } \left|\overline{S}(f - f_i, \mathcal{P})\right| < \varepsilon; \tag{8.9.44}$$

fix one such f_i and choose a $\mathcal{P}$ such that

$$\overline{S}(f_i, \mathcal{P}) - \underline{S}(f_i, \mathcal{P}) < \varepsilon, \tag{8.9.45}$$

which one can do by Lemma 1. Then for this $\mathcal{P}$,

$$\overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) < 3\varepsilon. \tag{8.9.46}$$

(Why?) By Lemma 1, then, f is R-integrable on A.

Finally,

$$\begin{aligned} \left|R\int_A f - R\int_A f_i\right| &\leq R\int_A |f - f_i| \\ &\leq R\int_A \left(\frac{\varepsilon}{mA}\right) = mA\left(\frac{\varepsilon}{mA}\right) = \varepsilon \end{aligned}$$

for all $i \geq k.$ Hence the second clause of our theorem follows, too. $\square$

> **Corollary 8.9.3**
>
> If $f: E^1 \rightarrow E^1$ is bounded and regulated (Chapter 5, §10) on $A=[a, b]$, then f is R-integrable on A.
>
> In particular, this applies if f is monotone, or of bounded variation, or relatively continuous, or a step function, on A.
>
> **Proof**
>
> By Lemma 2, this applies to $\mathcal{C}$-simple maps.
>
> Now, let f be regulated (e.g., of the kind specified above).
>
> Then by Lemma 2 of Chapter 5, §10,
>
> $$f=\lim_{i \rightarrow \infty} g_i \quad \text{(uniformly)} \tag{8.9.47}$$
>
> for finite $\mathcal{C}$-simple g_i.
>
> Thus f is R-integrable on A by Theorem 1. $\square$

II. Henceforth, we assume that m is a measure on a σ-ring $\mathcal{M} \supseteq \mathcal{C}$ in E^n, with $m<\infty$ on $\mathcal{C}$. (For a reader who took the "limited approach," it is now time to consider §§4-6 in full.) The measure m may, but need not, be Lebesgue measure in E^n.

> **Theorem 8.9.2**
>
> If $f: E^n \rightarrow E^1$ is R-integrable on $A \in \mathcal{C}$, it is also Lebesgue integrable (with respect to m as above) on A, and
>
> $$L \int_A f=R \int_A f, \tag{8.9.48}$$
>
> **Proof**
>
> Given a $\mathcal{C}$-partition $\mathcal{P}=\{A_i\}$ of A, define the $\mathcal{C}$-simple maps
>
> $$g=\sum_i a_i C_{A_i} \text{ and } h=\sum_i b_i C_{A_i} \tag{8.9.49}$$
>
> with
>
> $$a_i=\inf f\left[A_i\right] \text{ and } b_i=\sup f\left[A_i\right]. \tag{8.9.50}$$
>
> Then $g \leq f \leq h$ on A with
>
> $$\underline{S}(f, \mathcal{P})=\sum_i a_i m A_i=L \int_A g \tag{8.9.51}$$
>
> and
>
> $$\bar{S}(f, \mathcal{P})=\sum_i b_i m A_i=L \int_A h. \tag{8.9.52}$$
>
> By Theorem 1(c) in §5,
>
> $$\underline{S}(f, \mathcal{P})=L \int_A g \leq L \underline{\int}_A f \leq L \overline{\int}_A f \leq L \int_A h=\bar{S}(f, \mathcal{P}). \tag{8.9.53}$$
>
> As this holds for any $\mathcal{P}$, we get
>
> $$R \underline{\int}_A f=\sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) \leq L \underline{\int}_A f \leq L \overline{\int}_A f=\inf_{\mathcal{P}} \bar{S}(f, \mathcal{P})=R \overline{\int}_A f. \tag{8.9.54}$$
>
> But by assumption,

$$R\underline{\int}_A f = R\overline{\int}_A f. \tag{8.9.55}$$

Thus these inequalities become equations:

$$R\int_A f = \underline{\int}_A f = \overline{\int}_A f = R\int_A f. \tag{8.9.56}$$

Also, by definition (c), f is bounded on A; so $|f| < K < \infty$ on A. Hence

$$\left|\int_A f\right| \leq \int_A |f| \leq K \cdot mA < \infty. \tag{8.9.57}$$

Thus

$$\underline{\int}_A f = \overline{\int}_A f \neq \pm\infty, \tag{8.9.58}$$

i.e., f is Lebesgue integrable, and

$$L\int_A f = R\int_A f, \tag{8.9.59}$$

as claimed. □

Note 2. The converse fails. For example, as shown in the example in §4 , $f = C_R$ ($R =$ rationals) is L-integrable on $A = [0, 1]$. Yet f is not R-integrable.

For $\mathcal{C}$-partitions involve intervals containing both rationals (on which $f = 1$) and irrationals (on which $f = 0$). Thus for any $\mathcal{P}$,

$$\underline{S}(f, \mathcal{P}) = 0 \text{ and } \overline{S}(f, \mathcal{P}) = 1 \cdot mA = 1. \tag{8.9.60}$$

(Why?) So

$$R\overline{\int}_A f = \inf \overline{S}(f, \mathcal{P}) = 1, \tag{8.9.61}$$

while

$$R\underline{\int}_A f = 0 \neq R\overline{\int}_A f. \tag{8.9.62}$$

Note 3. By Theorem 1, any $R\int_A f$ is also a Lebesgue integral. Thus the rules of §§5-6 apply to R-integrals, provided that the functions involved are R-integrable. For a deeper study, we need a few more ideas.

Definitions (continued)

(d) The mesh $|\mathcal{P}|$ of a $\mathcal{C}$-partition $\mathcal{P} = \{A_1, \ldots, A_q\}$ is the largest of the diagonals dA_i :

$$|\mathcal{P}| = \max\{dA_1, dA_2, \ldots, dA_q\}. \tag{8.9.63}$$

Note 4. For any $A \in \mathcal{C}$, there is a sequence of $\mathcal{C}$-partitions $\mathcal{P}_k$ such that

(i) each P_{k+1} refines P_k and

(ii) $\lim_{k\to\infty} |P_k| = 0$.

To construct such a sequence, bisect the edges of A so as to obtain 2^n subintervals of diagonal $\frac{1}{2}dA$ (Chapter 3, §7). Repeat this with each of the subintervals, and so on. Then

$$|P_k| = \frac{dA}{2^k} \to 0. \tag{8.9.64}$$

Lemma 8.9.3

Let $f : A \to E^1$ be bounded. Let $\{\mathcal{P}_k\}$ satisfy (i) of Note 4. If $P_k = \{A_1^k, \ldots, A_{q_k}^k\}$, put

$$g_k = \sum_{i=1}^{q_k} C_{A_i^k} \inf f\left[A_i^k\right] \tag{8.9.65}$$

and

$$h_k = \sum_{i=1}^{q_k} C_{A_i^k \sup} f\left[A_i^k\right]. \tag{8.9.66}$$

Then the functions

$$g = \sup_k g_k \text{ and } h = \inf_k h_k \tag{8.9.67}$$

are Lebesgue integrable on A, and

$$\int_A g = \lim_{k\to\infty} \underline{S}(f, \mathcal{P}_k) \leq R\underline{\int}_A f \leq R\overline{\int}_A f \leq \lim_{k\to\infty} \overline{S}(f, \mathcal{P}_k) = \int_A h. \tag{8.9.68}$$

Proof

As in Theorem 2, we obtain $g_k \leq f \leq h_k$ on A with

$$\int_A g_k = \underline{S}(f, \mathcal{P}_k) \tag{8.9.69}$$

and

$$\int_A h_k = \overline{S}(f, \mathcal{P}_k). \tag{8.9.70}$$

Since $\mathcal{P}_{k+1}$ refines $\mathcal{P}_k$, it also easily follows that

$$g_k \leq g_{k+1} \leq \sup_k g_k = g \leq f \leq h = \inf_k h_k \leq h_{k+1} \leq h_k. \tag{8.9.71}$$

(Verify!)

Thus $\{g_k\}\uparrow$ and $\{h_k\}\downarrow$, and so

$$g = \sup_k g_k = \lim_{k\to\infty} g_k \text{ and } h = \inf_k h_k = \lim_{k\to\infty} h_k. \tag{8.9.72}$$

Also, as f is bounded

$$\left(\exists K \in E^1\right) \quad |f| < K \text{ on } A. \tag{8.9.73}$$

The definition of g_k and h_k then implies

$$(\forall k) \quad |g_k| \leq K \text{ and } |h_k| \leq K \text{ (why?)}, \tag{8.9.74}$$

with

$$\int_A (K) = K \cdot mA < \infty. \tag{8.9.75}$$

The g_k and h_k are measurable (even simple) on A, with $g_k \to g$ and $h_k \to h$.

Thus by Theorem 5 and Note 1, both from §6, g and h are Lebesgue integrable, with

$$\int_A g = \lim_{k\to\infty} \int_A g_k \text{ and } \int_A h = \lim_{k\to\infty} \int_A h_k. \tag{8.9.76}$$

As

$$\int_A g_k = \underline{S}(f, \mathcal{P}_k) \leq R \underline{\int}_A f \tag{8.9.77}$$

and

$$\int_A h_k = \overline{S}(f, \mathcal{P}_k) \geq R \overline{\int}_A f, \tag{8.9.78}$$

passage to the limit in equalities yields (6). Thus the lemma is proved. □

Lemma 8.9.4

With all as in Lemma 3, let B be the union of the boundaries of all intervals from all $\mathcal{P}_k$. Let $|\mathcal{P}_k| \to 0$. Then we have the following.

(i) If f is continuous at $p \in A$, then $h(p) = g(p)$.

(ii) The converse holds if $p \in A - B$.

Proof

For each k, p is in one of the intervals in $\mathcal{P}_k$; call it A_{kp}.

If $p \in A - B, p$ is an interior point of A_{kp}; so there is a globe

$$G_p(\delta_k) \subseteq A_{kp}. \tag{8.9.79}$$

Also, by the definition of g_k and h_k,

$$g_k(p) = \inf f[A_{kp}] \text{ and } h_k = \sup f[A_{kp}]. \tag{8.9.80}$$

(Why?)

Now fix $\varepsilon > 0$. If $g(p) = h(p)$, then

$$0 = h(p) - g(p) = \lim_{k\to\infty} [h_k(p) - g_k(p)]; \tag{8.9.81}$$

so

$$(\exists k) \quad |h_k(p) - g_k(p)| = \sup f[A_{kp}] - \inf f[A_{kp}] < \varepsilon. \tag{8.9.82}$$

As $G_p(\delta_k) \subseteq A_{kp}$, we get

$$(\forall x \in G_p(\delta_k)) \quad |f(x) - f(p)| \leq \sup f[A_{kp}] - \inf f[A_{kp}] < \varepsilon, \tag{8.9.83}$$

proving continuity (clause (ii)).

For (i), given $\varepsilon > 0$, choose $\delta > 0$ so that

$$(\forall x, y \in A \cap G_p(\delta)) \quad |f(x) - f(y)| < \varepsilon. \tag{8.9.84}$$

Because

$$(\forall \delta > 0)(\exists k_0)(\forall k > k_0) \quad |\mathcal{P}_k| < \delta \tag{8.9.85}$$

for $k > k_0, A_{kp} \subseteq G_p(\delta)$. Deduce that

$$(\forall k > k_0) \quad |h_k(p) - g_k(p)| \leq \varepsilon. \quad \square \tag{8.9.86}$$

Note 5. The Lebesgue measure of B in Lemma 4 is zero; for B consists of countably many "faces" (degenerate intervals), each of measure zero.

Theorem 8.9.3

A map $f : A \rightarrow E^1$ is R-integrable on A (with $m =$ Lebesgue measure) iff f is bounded on A and continuous on $A - Q$ for some Q with $mQ = 0$.

Note that relative continuity on $A - Q$ is not enough-take $f = C_R$ of Note 2.

Proof

If these conditions hold, choose $\{\mathcal{P}_k\}$ as in Lemma 4.

Then by the assumed continuity, $g = h$ on $A - Q, mQ = 0$.

Thus

$$\int_A g = \int_A h \tag{8.9.87}$$

(Corollary 2 in §5).

Hence by formula (6), f is R-integrable on A.

Conversely, if so, use Lemma 1 with

$$\varepsilon = 1, \frac{1}{2}, \ldots, \frac{1}{k}, \ldots \tag{8.9.88}$$

to get for each k some $\mathcal{P}_k$ such that

$$\overline{S}(f, \mathcal{P}_k) - \underline{S}(f, \mathcal{P}_k) < \frac{1}{k} \rightarrow 0. \tag{8.9.89}$$

By Corollary 1, this will hold if we refine each $\mathcal{P}_k$, step by step, so as to achieve properties (i) and (ii) of Note 4 as well. Then Lemmas 3 and 4 apply.

As

$$\overline{S}(f, \mathcal{P}_k) - \underline{S}(f, \mathcal{P}_k) \rightarrow 0, \tag{8.9.90}$$

formula (6) show that

$$\int_A g = \lim_{k \rightarrow \infty} \underline{S}(f, \mathcal{P}_k) = \lim_{k \rightarrow \infty} \overline{S}(f, \mathcal{P}_k) = \int_A h. \tag{8.9.91}$$

As h and g are integrable on A,

$$\int_A (h - g) = \int_A h - \int_A g = 0. \tag{8.9.92}$$

Also $h - g \geq 0;$ so by Theorem 1(h) in §5, $h = g$ on $A - Q', mQ' = 0$ (under Lebesgue measure). Hence by Lemma 4, f is continuous on

$$A - Q' - B, \tag{8.9.93}$$

with $mB = 0$ (Note 5).

Let $Q = Q' \cup B$. Then $mQ = 0$ and

$$A - Q = A - Q' - B; \tag{8.9.94}$$

so f is continuous on $A - Q$. This completes the proof. $\square$

Note 6. The first part of the proof does not involve B and thus works even if m is not the Lebesgue measure. The second part requires that $mB=0$.

Theorem 3 shows that R-integrals are limited to a.e. continuous functions and hence are less flexible than L-integrals: Fewer functions are R-integrable, and convergence theorems (§6, Theorems 4 and 5) fail unless $R\int_A f$ exists.

III. Functions $f : E^n \to E^s\ (C^s)$. For such functions, R-integrals are defined componentwise (see §7). Thus $f=(f_1,\ldots,f_s)$ is R-integrable on A iff all $f_k\ (k \leq s)$ are, and then

$$R\int_A f=\sum_{k=1}^s \bar{e}_k R\int_A f_k. \tag{8.9.95}$$

A complex function f is R-integrable iff f_{re} and f_{im} are, and then

$$R\int_A f=R\int_A f_{re}+iR\int_A f_{im}. \tag{8.9.96}$$

Via components, Theorems 1 to 3, Corollaries 3 and 4, additivity, linearity, etc., apply.

IV. Stieltjes Integrals. Riemann used Lebesgue premeasure v only. But as we saw, his method admits other premeasures, too.

Thus in E^1, we may let m be the LS premeasure s_α or the LS measure m_α where $\alpha \uparrow$ (Chapter 7, §5, Example (b), and Chapter 7, §9).

Then

$$R\int_A f dm \tag{8.9.97}$$

is called the Riemann-Stieltjes (RS) integral of f with respect to α, also written

$$R\int_A f d\alpha \quad \text{or} \quad R\int_a^b f(x) d\alpha(x) \tag{8.9.98}$$

(the latter if $A=[a,b]$); f and α are called the integrand and integrator, respectively.

If $\alpha(x)=x$, m_α becomes the Lebesgue measure, and

$$R\int f(x) d\alpha(x) \tag{8.9.99}$$

turns into

$$R\int f(x) dx. \tag{8.9.100}$$

Our theory still remains valid; only Theorem 3 now reads as follows.

Corollary 8.9.4

If f is bounded and a.e. continuous on $A=[a,b]$ (under an LS measure m_α) then

$$R\int_a^b f d\alpha \tag{8.9.101}$$

exists. The converse holds if α is continuous on A.

For by Notes 5 and 6, the "only if" in Theorem 3 holds if $m_\alpha B=0$. Here consists of countably many endpoints of partition subintervals. But (see Chapter §9) $m_\alpha\{p\}=0$ if α is continuous at p. Thus the later implies $m_\alpha B=0$.

RS-integration has been used in many fields (e.g., probability theory, physics, etc.), but it is superseded by LS-integration, i.e., Lebesgue integration with respect to m_α, which is fully covered by the general theory of §§1-8.

Actually, Stieltjes himself used somewhat different definitions (see Problems 10-13), which amount to applying the set function σ_α of Problem 9 in Chapter 7, §4, instead of s_α or m_α. We reserve the name "Stieltjes integrals," denoted

$$S \int_a^b f d\alpha, \tag{8.9.102}$$

for such integrals, and "RS-integrals" for those based on m_α or s_α (this terminology is not standard).

Observe that σ_α need not be ≥ 0. Thus for the first time, we encounter integration with respect to sign-changing set functions. A much more general theory is presented in §10 (see Problem 10 there).

8.9.E: Problems on Riemann and Stieltjes Integrals

Exercise $8.9.E.1$

Replacing "$\mathcal{M}$" by "$\mathcal{C}$," and "elementary and integrable" or "elementary and nonnegative" by "$\mathcal{C}$-simple," price Corollary 1(ii)(iv)(vii) and Theorems 1(i) and 2(ii), all in §4, and do Problems 5-7 in §4, for R-integrals.

Exercise $8.9.E.2$

Verify Note 1.

Exercise $8.9.E.2'$

Do Problems $5-7$ in §5 for R-integrals.

Exercise $8.9.E.3$

Do the following for R-integrals.
(i) Prove Theorems $1(\mathrm{a})-(\mathrm{g})$ and 2, both in $\S 5(\mathcal{C}\text{-partitions only})$.
(ii) Prove Theorem 1 and Corollaries 1 and 2, all in §6.
(iii) Show that definition (b) can be replaced by formulas analogous to formulas $\left(1'\right)$, $\left(1''\right)$, and (1) of Definition 1 in §5.
[Hint: Use Problems 1 and $2'$.]

Exercise $8.9.E.4$

Fill in all details in the proof of Theorem 1, Lemmas 3 and 4, and Corollary 4.

Exercise $8.9.E.5$

For $f, g: E^{n} \rightarrow E^{s}\left(C^{s}\right)$, via components, prove the following.
(i) Theorems $1-3$ and
(ii) additivity and linearity of R-integrals.
Do also Problem 13 in §7 for R-integrals.

Exercise $8.9.E.6$

Prove that if $f: A \rightarrow E^{s}\left(C^{s}\right)$ is bounded and a.e. continuous on A, then

$$R \int_{A}|f| \geq\left|R \int_{A} f\right| . \tag{8.9.E.1}$$

For $m=$ Lebesgue measure, do it assuming R-integrability only.

Exercise $8.9.E.7$

Prove that if $f, g: A \rightarrow E^{1}$ are R-integrable, then
(i) so is f^{2}, and
(ii) so is $f g$.
[Hints: (i) Use Lemma 1. Let $h=|f| \leq K<\infty$ on A. Verify that

$$\left(\inf h\left[A_{i}\right]\right)^{2}=\inf f^{2}\left[A_{i}\right] \text{ and } \left(\sup h\left[A_{i}\right]\right)^{2}=\sup f^{2}\left[A_{i}\right] ; \tag{8.9.E.2}$$

so

$$\begin{aligned}\sup f^{2}\left[A_{i}\right]-\inf f^{2}\left[A_{i}\right] &=\left(\sup h\left[A_{i}\right]+\inf h\left[A_{i}\right]\right)\left(\sup h\left[A_{i}\right]-\inf h\left[A_{i}\right]\right) \\ & \leq\left(\sup h\left[A_{i}\right]-\inf h\left[A_{i}\right]\right) 2 K .\end{aligned}$$

(ii) Use

$$f g=\frac{1}{4}\left[(f+g)^{2}-(f-g)^{2}\right] . \tag{8.9.E.3}$$

(iii) For $m=$ Lebesgue measure, do it using Theorem 3.]

Exercise $8.9.E.8$

Prove that if $m=$ the volume function v (or LS function s_α for a continuous α), then in formulas (1) and (2), one may replace A_i by $\overline{A}_i$ (closure of A_i) .
[Hint: Show that here $mA=m\overline{A}$,

$$R \int_{A} f=R \int_{\overline{A}} f, \tag{8.9.E.4}$$

and additivity works even if the A_i have some common "faces" (only their interiors being disjoint).]

Exercise $8.9.E.9$

(Riemann sums.) Instead of $\underline{S}$ and $\bar{S}$, Riemann used sums

$$S(f, \mathcal{P})=\sum_{i} f\left(x_{i}\right) d m A_{i}, \tag{8.9.E.5}$$

where $m=v$ (see $\text{Problem } 8)$ and x_i is arbitrarily chosen from $\overline{A_i}$.
For a bounded f, prove that

$$r=R \int_{A} f d m \tag{8.9.E.6}$$

exists on $A=[a,b]$ iff for every $\varepsilon>0,$ there is $\mathcal{P}_\varepsilon$ such that

$$|S(f, \mathcal{P})-r|<\varepsilon \tag{8.9.E.7}$$

for every refinement

$$\mathcal{P}=\left\{A_{i}\right\} \tag{8.9.E.8}$$

of $\mathcal{P}_\varepsilon$ and any choice of $x_i \in \overline{A_i}$.
[Hint: Show that by Problem 8, this is equivalent to formula (3).]

Exercise $8.9.E.10$

Replacing m by the σ_α of Problem 9 of Chapter 7, §4, write $S(f, \mathcal{P}, \alpha)$ for $S(f, \mathcal{P})$ in Problem 9, treating Problem 9 as a definition of the Stieltjes integral,

$$S \int_a^b f d\alpha \quad \left(\text{or } S \int_a^b f d\sigma_\alpha\right). \tag{8.9.E.9}$$

Here $f, \alpha : E^1 \rightarrow E^1$ (monotone or not; even $f, \alpha : E^1 \rightarrow C$ will do).
Prove that if $\alpha : E^1 \rightarrow E^1$ is continuous and $\alpha \uparrow$, then

$$S \int_a^b f d\alpha = R \int_a^b f d\alpha, \tag{8.9.E.10}$$

the RS-integral.

Exercise $8.9.E.11$

(Integration by parts.) Continuing Problem 10, prove that

$$S \int_a^b f d\alpha \tag{8.9.E.11}$$

exists iff

$$S \int_a^b \alpha d f \tag{8.9.E.12}$$

does, and then

$$S \int_a^b f d\alpha + S \int_a^b \alpha d f = K, \tag{8.9.E.13}$$

where

$$K = f(b)\alpha(b) - f(a)\alpha(a). \tag{8.9.E.14}$$

[Hints: Take any $\mathcal{C}$-partition $\mathcal{P} = \{A_i\}$ of $[a, b]$, with

$$\overline{A_i} = [y_{i-1}, y_i], \tag{8.9.E.15}$$

say. For any $x_i \in \overline{A}_i$, verify that

$$S(f, \mathcal{P}, \alpha) = \sum f(x_i)[\alpha(y_i) - \alpha(y_{i-1})] = \sum f(x_i)\alpha(y_i) - \sum f(x_i)\alpha(y_{i-1}) \tag{8.9.E.16}$$

and

$$K = \sum f(x_i)\alpha(y_i) - \sum f(x_{i-1})\alpha(y_{i-1}). \tag{8.9.E.17}$$

Deduce that

$$K - S(f, \mathcal{P}, \alpha) = S(\alpha, \mathcal{P}', f) = \sum \alpha(x_i)[f(x_i) - f(y_i)] - \sum \alpha(x_{i-1})[f(y_i) - f(x_{i-1})]; \tag{8.9.E.18}$$

here $\mathcal{P}'$ results by combining the partition points x_i and y_i, so it refines $\mathcal{P}$.
Now, if $S\int_a^b \alpha df$ exists, fix $\mathcal{P}_\varepsilon$ as in Problem 9 and show that

$$\left|K - S(f,\mathcal{P},\alpha) - S\int_a^b \alpha df\right| < \varepsilon \tag{8.9.E.19}$$

whenever $\mathcal{P}$ refines $\mathcal{P}_\varepsilon$.]

Exercise 8.9.E. 12

If $\alpha : E^1 \to E^1$ is of class CD^1 on $[a,b]$ and if

$$S\int_a^b f d\alpha \tag{8.9.E.20}$$

exists (see Problem 10), it equals

$$R\int_a^b f(x)\alpha'(x)dx. \tag{8.9.E.21}$$

[Hints: Set $\phi = f\alpha', \mathcal{P} = \{A_i\}, \overline{A_i} = [a_{i-1}, a_i]$. Then

$$S(\phi,\mathcal{P}) = \sum f(x_i)\alpha'(x_i)(a_i - a_{i-1}), \quad x_i \in \overline{A_i} \tag{8.9.E.22}$$

and (Corollary 3 in Chapter 5, §2)

$$S(f,\mathcal{P},\alpha) = \sum f(x_i)[\alpha(a_i) - \alpha(a_{i-1})] = \sum f(x_i)\alpha'(q_i), \quad q_i \in A_i. \tag{8.9.E.23}$$

As f is bounded and α' is uniformly continuous on $[a,b]$ (why?), deduce that

$$(\forall \varepsilon > 0)(\exists \mathcal{P}_\varepsilon)(\forall \mathcal{P}_\varepsilon)(\forall \mathcal{P} \text{ refining } \mathcal{P}_\varepsilon)$$
$$|S(\phi,\mathcal{P}) - S(f,\mathcal{P},\alpha)| < \frac{1}{2}\varepsilon \text{ and } \left|S(f,\mathcal{P},\alpha) - S\int_a^b f d\alpha\right| < \frac{1}{2}\varepsilon.$$

Proceed. Use Problem 9.]

Exercise 8.9.E. 13

(Laws of the mean.) Let $f, g, \alpha : E^1 \to E^1; p \leq f \leq q$ on $A = [a,b]$; $p, q \in E^1$. Prove the following.
(i) If $\alpha \uparrow$ and if

$$s\int_a^b f d\alpha \tag{8.9.E.24}$$

exists, then $(\exists c \in [p,q])$ such that

$$S\int_a^b f d\alpha = c[\alpha(b) - \alpha(a)]. \tag{8.9.E.25}$$

Similarly, if

$$R \int_a^b f d\alpha \tag{8.9.E.26}$$

exists, then $(\exists c \in [p, q])$ such that

$$R \int_a^b f d\alpha = c[\alpha(b+) - \alpha(a-)]. \tag{8.9.E.27}$$

(i') If f also has the Darboux property on A, then $c = f(x_0)$ for some $x_0 \in A$.
(ii) If α is continuous, and $f \uparrow$ on A, then

$$S \int_a^b f d\alpha = [f(b)\alpha(b) - f(a)\alpha(a)] - S \int_a^b \alpha df \tag{8.9.E.28}$$

exists, and $(\exists z \in A)$ such that

$$\begin{aligned} S \int_a^b f d\alpha &= f(a) S \int_a^z d\alpha + f(b) S \int_z^b d\alpha \\ &= f(a)[\alpha(z) - \alpha(a)] + f(b)[\alpha(b) - \alpha(z)]. \end{aligned}$$

(ii') If g is continuous and $f \uparrow$ on A, then $(\exists z \in A)$ such that

$$R \int_a^b f(x) g(x) dx = p \cdot R \int_a^z g(x) dx + q \cdot R \int_z^b g(x) dx. \tag{8.9.E.29}$$

If $f \downarrow$, replace f by $-f$. (See also Corollary 5 in Chapter 9, §1.)
[Hints: (i) As $\alpha \uparrow$, we get

$$p[\alpha(b) - \alpha(a)] \leq S \int_a^b f d\alpha \leq q[\alpha(b) - \alpha(a)]. \tag{8.9.E.30}$$

(Why?) Now argue as in §6, Theorem 3 and Problem 2.
(ii) Use Problem 11, and apply (i) to $\int \alpha df$.
(ii') By Theorem 2 of Chapter 5, $10, g has a primitive $\beta \in CD^1$. Apply Problem 12 to $S \int_a^b f d\beta$.]

8.10: Integration in Generalized Measure Spaces

Let $(S, \mathcal{M}, s)$ be a generalized measure space. By Note 1 in §3, a map f is s-measurable iff it is v_s-measurable. This naturally leads us to the following definition.

Definition

A map $f : S \rightarrow E$ is s-integrable on a set A iff it is v_s-integrable on A. (Recall that v_s, the total variation of s, is a measure.)

Note 1. Here the range spaces of f and s are assumed complete and such that $f(x) s A$ is defined for $x \in S$ and $A \in \mathcal{M}$. Thus if s is vector valued, f must be scalar valued, and vice versa. Later, if a factor p occurs, it must be such that $p f(x) s A$ is defined, i.e., at least two of p, $f(x)$, and sA are scalars.

Note 2. If s is a measure (≥ 0), then $v_s = s^+ = s$ (Corollary 3 in Chapter 7, §11); so our present definition agrees with the previous ones (as in Theorem 1 of §7).

Lemma 8.10.1

If m' and m'' are measures, with $m' \geq m''$ on $\mathcal{M}$, then

$$\int_A |f| d m' \geq \int_A |f| d m'' \tag{8.10.1}$$

for all $A \in \mathcal{M}$ and any $f : S \rightarrow E$.

Proof

First, take any elementary and nonnegative map $g \geq |f|$,

$$g = \sum_i C_{A_i} a_i \text{ on } A. \tag{8.10.2}$$

Then (§4)

$$\int_A g d m' = \sum a_i m' A_i \geq \sum a_i m'' A_i = \int_A g d m''. \tag{8.10.3}$$

Hence by Definition 1 in §5,

$$\int_A |f| d m' = \inf_{g \geq |f|} \int_A g d m' \geq \inf_{g \geq |f|} \int_A g d m'' = \int_A |f| d m'', \tag{8.10.4}$$

as claimed. $\square$

Lemma 8.10.2

(i) If $s : \mathcal{M} \rightarrow E^n\left(C^n\right)$ with $s = \left(s_1, \ldots, s_n\right)$, and if f is s-integrable on $A \in \mathcal{M}$, then f is s_k-integrable on A for $k = 1, 2, \ldots, n$.

(ii) If s is a signed measure and f is s-integrable on A, then f is integrable on A with respect to both s^+ and s^- (with s^+ and s^- as in formula (3) in Chapter 7, §11).

Note 3. The converse statements hold if f is $\mathcal{M}$-measurable on A.

Proof

(i) If $s = \left(s_1, \ldots, s_n\right)$, then (Problem 4 of Chapter 7, §11)

$$v_s \geq v_{s_k}, \quad k = 1, \ldots, n. \tag{8.10.5}$$

Hence by Definition 1 and Lemma 1, the s-integrability of f implies

$$\infty > \int_A |f| dv_s \geq \int_A |f| dv_{s_k}. \tag{8.10.6}$$

Also, f is v_s-measurable, i.e., $\mathcal{M}$-measurable on $A - Q$, with

$$0 = v_s Q \geq v_{s_k} Q \geq 0. \tag{8.10.7}$$

Thus f is s_k -integrable on $A, k = 1, \ldots, n$, as claimed.

(ii) If $s = s^+ - s^-$, then by Theorem 4 in Chapter 7, §11, and Corollary 3 there, s^+ and s^- are measures (≥ 0) and $v_s = s^+ + s^-$, so that both

$$v_s \geq s^+ = v_{s^+} \text{ and } v_s \geq s^- = v_{s^-}. \tag{8.10.8}$$

Thus the desired result follows exactly as in part (i) of the proof. □

We leave Note 3 as an exercise.

Definition

If f is s-integrable on $A \in \mathcal{M}$, we set

(i) in the case $s : \mathcal{M} \to E^*$,

$$\int_A f ds = \int_A f ds^+ - \int_A f ds^-, \tag{8.10.9}$$

with s^+ and s^- as in formula (3) of Chapter 7, §11;

(ii) in the case $s : \mathcal{M} \to E^n\left(C^n\right)$,

$$\int_A f ds = \sum_{k=1}^n \vec{e}_k \int_A f ds_k, \tag{8.10.10}$$

with $\vec{e}_k$ as in Theorem 2 of Chapter 3, §§1-3;

(iii) if $s : \mathcal{M} \to C$,

$$\int_A f ds = \int_A f ds_{re} + i \cdot \int_A f ds_{im}. \tag{8.10.11}$$

(See also Problems 2 and 3.)

Note 4. If s is a measure, then

$$s = s^+ = s_{re} = s_1 \tag{8.10.12}$$

and

$$0 = s^- = s_{im} = s_2; \tag{8.10.13}$$

so Definition 2 agrees with our previous definitions. Similarly for $s : \mathcal{M} \to E^n\left(C^n\right)$.

Below, s, t, and u are generalized measures on $\mathcal{M}$ as in Definition 2, while $f, g : S \to E$ are functions, with E a complete normed space, as in Note 1.

Theorem 8.10.1

The linearity, additivity, and σ-additivity properties (as in §7, Theorems 2 and 3) also apply to integrals

$$\int_A f ds, \tag{8.10.14}$$

with s as in Definition 2.

Proof

(i) Linearity: Let $f, g : S \to E$ be s-integrable on $A \in \mathcal{M}$. Let p, q be suitable constants (see Note 1).

If s is a signed measure, then by Lemma 2(ii) and Definitions 1 and 2, f is integrable with respect to v_s, s^+, and s^-. As these are measures, Theorem 2 in §7 shows that $pf+qg$ is integrable with respect to v_s, s^+, and s^-, and by Definition 2,

$$\begin{aligned} \int_A (pf+qg)ds &= \int_A (pf+qg)ds^+ - \int_A (pf+qg)ds^- \\ &= p\int_A fds^+ + q\int_A gds^+ - p\int_A fds^- - q\int_A gds^- \\ &= p\int_A fds + q\int_A gds. \end{aligned}$$

Thus linearity holds for signed measures. Via components, it now follows for $s : \mathcal{M} \to E^n\left(C^n\right)$ as well. Verify!

(ii) Additivity and σ-additivity follow in a similar manner. □

Corollary 8.10.1

Assume f is s-integrable on A, with s as in Definition 2.

(i) If f is constant $(f=c)$ on A, we have

$$\int_A fds = c \cdot sA. \tag{8.10.15}$$

(ii) If

$$f = \sum_i a_i C_{A_i} \tag{8.10.16}$$

for an $\mathcal{M}$-partition $\{A_i\}$ of A, then

$$\int_A fds = \sum_i a_i sA_i \text{ and } \int_A |f| ds = \sum_i |a_i| sA_i \tag{8.10.17}$$

(both series absolutely convergent).

(iii) $|f| < \infty$ a.e. on A.

(iv) $\int_A |f| dv_s = 0$ iff $f=0$ a.e. on A.

(v) The set $A(f \neq 0)$ is (v_s) σ-finite (Definition 4 in Chapter 7, §5).

(vi) $\int_A fds = \int_{A-Q} fds$ if $v_sQ=0$ or $f=0$ on Q $(Q \in \mathcal{M})$.

(vii) f is s-integrable on any $\mathcal{M}$-set $B \subseteq A$.

Proof

(i) If $s = s^+ - s^-$ is a signed measure, we have by Definition 2 that

$$\int_A fds = \int_A fds^+ - \int_A fds^- = c\left(s^+A - s^-A\right) = c \cdot sA, \tag{8.10.18}$$

as required.

For $s : \mathcal{M} \to E^n\left(C^n\right),$ the result now follows via components. (Verify!)

(ii) As $f = a_i$ on A_i, clause (i) yields

$$\int_{A_i} fds = a_i sA_i, \quad i = 1, 2, \ldots. \tag{8.10.19}$$

Hence by σ-additivity,

$$\int_A f ds = \sum_i \int_{A_i} f ds = \sum_i a_i s A_i, \tag{8.10.20}$$

as claimed.

Clauses (iii), (iv), and (v) follow by Corollary 1 in §5 and Theorem 1(b)(h) there, as v_s is a measure; (vi) is proved as §5, Corollary 2. We leave (vii) as an exercise. □

Theorem 8.10.2 (dominated convergence)

If

$$f = \lim_{i\to\infty} f_i \text{ (pointwise)} \tag{8.10.21}$$

on $A - Q\ (v_s Q = 0)$ and if each f_i is s-integrable on A, so is f, and

$$\int_A f ds = \lim_{i\to\infty} \int_A f_i ds, \tag{8.10.22}$$

all provided that

$$(\forall i) \quad |f_i| \leq g \tag{8.10.23}$$

for some map g with $\int_A g dv_s < \infty$.

Proof

If s is a measure, this follows by Theorem 5 in §6. Thus as v_s is a measure, f is v_s-integrable (hence s-integrable) on A, as asserted.

Next, if $s = s^+ - s^-$ is a signed measure, Lemma 2 shows that f and the f_i are s^+ and s^--integrable as well, with

$$\int_A |f_i|\, ds^+ \leq \int_A |f_i|\, dv_s \leq \int_A g dv_s < \infty; \tag{8.10.24}$$

similarly for

$$\int_A |f_i|\, ds^-. \tag{8.10.25}$$

As s^+ and s^- are measures, Theorem 5 of §6 yields

$$\int_A f ds = \int_A f ds^+ - \int_A f ds^- = \lim \left(\int_A f_i ds^+ - \int_A f_i ds^- \right) = \lim \int_A f_i ds. \tag{8.10.26}$$

Thus all is proved for signed measures.

In the case $s : \mathcal{M} \to E^n\,(C^n)$, the result now easily follows by Definition 2(ii)(iii) via components.$quad$□

Theorem 8.10.3 (uniform convergence)

If $f_i \to f$ (uniformly) on $A - Q\ (v_s A < \infty, v_s Q = 0)$, and if each f_i is s-integrable on A, so is f, and

$$\int_A f ds = \lim_{i\to\infty} \int_A f_i ds. \tag{8.10.27}$$

Proof

Argue as in Theorem 2, replacing §6, Theorem 5, by §7, Lemma 1. □

Our next theorem shows that integrals behave linearly with respect to measures.

Theorem 8.10.4

Let $t, u : \mathcal{M} \rightarrow E^{*}\left(E^{n}, C^{n}\right)$, with $v_{t}<\infty$ on $\mathcal{M}$, and let

$$s=p t+q u \tag{8.10.28}$$

for finite constants p and q. Then the following statements are true.

(a) If t and u are generalized measures, so is s.

(b) If, further, f is $\mathcal{M}$-measurable on a set A and is both t-and u-integrable on A, it is also s-integrable on A, and

$$\int_{A} f d s=p \int_{A} f d t+q \int_{A} f d u. \tag{8.10.29}$$

Proof

We consider only assertion (b) for $s=t+u$; the rest is easy.

First, let f be $\mathcal{M}$-elementary on A. By Corollary 1(ii), we set

$$\int_{A} f d t=\sum_{i} a_{i} t A_{i} \text{ and } \int_{A} f d u=\sum_{i} a_{i} u A_{i}. \tag{8.10.30}$$

Also, by integrability,

$$\infty>\int_{A}|f| d v_{t}=\sum\left|a_{i}\right| v_{t} A_{i} \text{ and } \infty>\int_{A}|f| d v_{u}=\sum_{i}\left|a_{i}\right| v_{u} A_{i}. \tag{8.10.31}$$

Now, by Problem 4 in Chapter 7, §11,

$$v_{s}=v_{t+u} \leq v_{t}+v_{u}; \tag{8.10.32}$$

so

$$\begin{aligned}\int_{A}|f| d v_{s} &=\sum_{i}\left|a_{i}\right| v_{s} A_{i} \\ & \leq \sum_{i}\left|a_{i}\right|\left(v_{t} A_{i}+v_{u} A_{i}\right)=\int_{A}|f| d v_{t}+\int_{A}|f| d v_{u}<\infty.\end{aligned}$$

As f is also $\mathcal{M}$-measurable (even elementary), it is s-integrable on A (by Definition 1), and

$$\int_{A} f d s=\sum_{i} a_{i} s A_{i}=\sum_{i} a_{i}\left(t A_{i}+u A_{i}\right)=\int_{A} f d t+\int_{A} f d u, \tag{8.10.33}$$

as claimed.

Next, suppose f is $\mathcal{M}$-measurable on A and $v_{u} A<\infty$. By assumption, $v_{t} A<\infty$, too; so

$$v_{s} A \leq v_{t} A+v_{u} A<\infty. \tag{8.10.34}$$

Now, by Theorem 3 in §1,

$$f=\lim _{i \rightarrow \infty} f_{i} \text{ (uniformly)} \tag{8.10.35}$$

for some $\mathcal{M}$-elementary maps f_{i} on A. By Lemma 2 in §7, for large i, the f_{i} are integrable with respect to both v_{t} and v_{u} on A. By what was shown above, they are also s-integrable, with

$$\int_{A} f_{i} d s=\int_{A} f_{i} d t+\int_{A} f_{i} d u. \tag{8.10.36}$$

With $i \rightarrow \infty$, Theorem 3 yields the result.

Finally, let $v_{u} A=\infty$. By Corollary 1(v), we may assume (as in Lemma 3 of §7) that $A_{i} \nearrow A$, with $v_{u} A_{i}<\infty$, and $v_{t} A_{i}<\infty$ (since $v_{t}<\infty$, by assumption). Set

$$f_i = fC_{A_i} \rightarrow f \text{ (pointwise)} \tag{8.10.37}$$

on A, with $|f_i| \leq |f|$. (Why?)

As $f_i = f$ on A_i and $f_i = 0$ on $A - A_i$, all f_i are both t- and u-integrable on A (for f is). Since $v_t A_i < \infty$ and $v_u A_i < \infty$, the f_i are also s-integrable (as shown above), with

$$\int_A f_i ds = \int_{A_i} f_i ds = \int_{A_i} f_i dt + \int_{A_i} f_i du = \int_A f_i dt + \int_A f_i du. \tag{8.10.38}$$

With $i \rightarrow \infty$, Theorem 2 now yields the result.

To complete the proof of (b), it suffices to consider, along similar lines, the case $s = pt$ (or $s = qu$). We leave this to the reader.

For (a), see Chapter 7, §11. □

Theorem 8.10.5

If f is s-integrable on A, so is $|f|$, and

$$\left|\int_A f ds\right| \leq \int_A |f| dv_s. \tag{8.10.39}$$

Proof

By Definition 1, and Theorem 1 of §1, f and $|f|$ are $\mathcal{M}$-measurable on $A - Q, v_s Q = 0,$ and

$$\int_A |f| dv_s < \infty; \tag{8.10.40}$$

so $|f|$ is s-integrable on A.

The desired inequality is immediate by Corollary 1(ii) if f is elementary.

Next, exactly as in Theorem 4, one obtains it for the case $v_s A < \infty,$ and then for $v_s A = \infty.$ We omit the details. □

Definition

We write

$$\text{" } ds = gdt \text{ in } A \text{ "} \tag{8.10.41}$$

or

$$\text{" } s = \int gdt \text{ in } A \text{ "} \tag{8.10.42}$$

iff g is t-integrable on A, and

$$sX = \int_X gdt \tag{8.10.43}$$

for $A \supseteq X, X \in \mathcal{M}$.

We then call s the indefinite integral of g in A. ($\int_X gdt$ may be interpreted as in Problems 2-4 below.)

Lemma 8.10.3

If $A \in \mathcal{M}$ and

$$ds = gdt \text{ in } A, \tag{8.10.44}$$

then

$$dv_s = |g| dv_t \text{ in } A. \tag{8.10.45}$$

Proof

By assumption, g and $|g|$ are v_t-integrable on X, and

$$sX = \int_X g dt \tag{8.10.46}$$

for $A \supseteq X, X \in \mathcal{M}.$ We must show that

$$v_s X = \int_X |g| dv_t \tag{8.10.47}$$

for such X.

This is easy if $g = c$ (constant) on X. For by definition,

$$v_s X = \sup_{\mathcal{P}} \sum_i |sX_i|, \tag{8.10.48}$$

over all $\mathcal{M}$-partitions $\mathcal{P} = \{X_i\}$ of X. As

$$sX_i = \int_{X_i} g dt = c \cdot tX_i, \tag{8.10.49}$$

we have

$$v_s X = \sup_{\mathcal{P}} \sum_i |c| |tX_i| = |c| \sup_{\mathcal{P}} \sum_i |tX_i| = |c| v_t X; \tag{8.10.50}$$

so

$$v_s X = \int_X |g| dv_t. \tag{8.10.51}$$

Thus all is proved for constant g.

Hence by σ-additivity, the lemma holds for $\mathcal{M}$-elementary maps g. (Why?)

In the general case, g is t-integrable on X, hence $\mathcal{M}$-measurable and finite on $X - Q, v_t Q = 0.$ By Corollary 1(iii), we may assume g finite and measurable on X; so

$$g = \lim_{k \to \infty} g_k \text{ (uniformly)} \tag{8.10.52}$$

on X for some $\mathcal{M}$-elementary maps g_k, all integrable on X, with respect to v_t (and t).

Let

$$s_k = \int g_k dt \tag{8.10.53}$$

in X. By what we just proved for elementary and integrable maps,

$$v_{s_k} X = \int_X |g_k| dv_t, \quad k = 1, 2, \ldots. \tag{8.10.54}$$

Now, if $v_t X < \infty,$ Theorem 3 yields

$$\int_X |g| dv_t = \lim_{k \to \infty} \int_X |g_k| dv_t = \lim_{k \to \infty} v_{s_k} X = v_s X \tag{8.10.55}$$

(see Problem 6). Thus all is proved if $v_t X < \infty$.

If, however, $v_t X = \infty,$ argue as in Theorem 4 (the last step), using the left continuity of v_s and of

$$\int |g| dv_t. \tag{8.10.56}$$

Verify! $\square$

Theorem 8.10.6 (change of measure)

If f is s-integrable on $A \in \mathcal{M},$ with

$$ds = gdt \text{ in } A, \tag{8.10.57}$$

then (subject to Note 1) fg is t-integrable on A and

$$\int_A f ds = \int_A f g dt. \tag{8.10.58}$$

(Note the formal substitution of "gdt" for "ds.")

Proof

The proof is easy if f is constant or elementary on A (use Corollary 1(ii)). We leave this case to the reader, and next we assume g is bounded and $v_t A < \infty$.

By s-integrability, f is $\mathcal{M}$-measurable and finite on $A - Q,$ with

$$0 = v_s Q = \int_Q |g| dv_t \tag{8.10.59}$$

by Lemma 3. Hence $0 = g = fg$ on $Q - Z, v_t Z = 0.$ Therefore,

$$\int_Q f g dt = 0 = \int_Q f ds \tag{8.10.60}$$

for $v_s Q = 0.$ Thus we may neglect Q and assume that f is finite and $\mathcal{M}$-measurable on A.

As $ds = gdt,$ Definition 3 and Lemma 3 yield

$$v_s A = \int_A |g| dv_t < \infty. \tag{8.10.61}$$

Also (Theorem 3 in Chapter 8, §1),

$$f = \lim_{k \rightarrow \infty} f_k \quad \text{(uniformly)} \tag{8.10.62}$$

for elementary maps f_k, all v_s-integrable on A (Lemma 2 in §7). As g is bounded, we get on A

$$fg = \lim_{k \rightarrow \infty} f_k g \quad \text{(uniformly)}. \tag{8.10.63}$$

Moreover, as the theorem holds for elementary and integrable maps, $f_k g$ is t-integrable on A, and

$$\int_A f_k ds = \int_A f_k g dt, \quad k = 1, 2, \ldots. \tag{8.10.64}$$

Since $v_s A < \infty$ and $v_t A < \infty,$ Theorem 3 shows that fg is t-integrable on A, and

$$\int_A f ds = \lim_{k \rightarrow \infty} \int_A f_k ds = \lim_{k \rightarrow \infty} \int_A f_k g dt = \int_A f g dt. \tag{8.10.65}$$

Thus all is proved if $v_t A < \infty$ and g is bounded on A.

In the general case, we again drop a null set to make f and g finite and $\mathcal{M}$-measurable on A. By Corollary 1(v), we may again assume $A_i \nearrow A$, with $v_t A_i < \infty \ (\forall i)$.

Now for $i = 1, 2, \ldots$ set

$$g_i = \begin{cases} g & \text{on } A_i(|g| \leq i), \\ 0 & \text{elsewhere.} \end{cases} \tag{8.10.66}$$

Then each g_i is bounded,

$$g_i \rightarrow g \text{ (pointwise)}, \tag{8.10.67}$$

and

$$|g_i| \leq |g| \tag{8.10.68}$$

on A. We also set $f_i = fC_{A_i}$; so $f_i \rightarrow f$ (pointwise) and $|f_i| \leq |f|$ on A. Then

$$\int_A f_i ds = \int_{A_i} f_i ds = \int_{A_i} f_i g_i dt = \int_A f_i g_i dt. \tag{8.10.69}$$

(Why?) Since $|f_i g_i| \leq |fg|$ and $f_i g_i \rightarrow fg$, the result follows by Theorem 2. $\square$

8.10.E: Problems on Generalized Integration

Exercise $8.10.E.1$

Fill in the missing details in the proofs of this section. Prove Note 3.

Exercise $8.10.E.2$

Treat Corollary 1 (ii) as a definition of

$$\int_A f d s \tag{8.10.E.1}$$

for $s: \mathcal{M} \rightarrow E$ and elementary and integrable f, even if $E \neq E^n\left(C^n\right)$.
Hence deduce Corollary $1(\mathrm{i})(\mathrm{vi})$ for this more general case.

Exercise $8.10.E.3$

Using Lemma 2 in §7, with $m=v_s, s: \mathcal{M} \rightarrow E$, construct

$$\int_A f d s \tag{8.10.E.2}$$

as in Definition 2 of §7 for the case $v_s A \neq \infty$. Show that this agrees with Problem 2 if f is elementary and integrable. Then prove linearity for functions with v_s-finite support as in §7.

Exercise $8.10.E.4$

Define

$$\int_A f d s \quad(s: \mathcal{M} \rightarrow E) \tag{8.10.E.3}$$

also for $v_s A=\infty$.
[Hint: Set $m=v_s$ in Lemma 3 of $\S 7$.]

Exercise $8.10.E.5$

Prove Theorems 1 to 3 for the general case, $s: \mathcal{M} \rightarrow E$ (see Problem 4).
[Hint: Argue as in §7.]

Exercise $8.10.E.5'$

From Problems $2-4$, deduce Definition 2 as a theorem in the case $E=E^n\left(C^n\right)$.

Exercise $8.10.E.6$

Let $s, s_k: \mathcal{M} \rightarrow E(k=1,2, \ldots)$ be any set functions. Let $A \in \mathcal{M}$ and

$$\mathcal{M}_A=\{X \in \mathcal{M} \mid X \subseteq A\}. \tag{8.10.E.4}$$

Prove that if

$$(\forall X \in \mathcal{M}_A) \quad \lim_{k\to\infty} s_k X = sX, \tag{8.10.E.5}$$

then

$$\lim_{k\to\infty} v_{s_k} A = v_s A, \tag{8.10.E.6}$$

provided $\lim_{k\to\infty} v_{s_k}$ exists.
[Hint: Using Problem 2 in Chapter 7, §11, fix a finite disjoint sequence $\{X_i\} \subseteq \mathcal{M}_A$.
Then

$$\sum_i |sX_i| = \sum_i \lim_{k\to\infty} |s_k X_i| = \lim_{k\to\infty} \sum_i |s_k X_i| \leq \lim_{k\to\infty} v_{s_k} A. \tag{8.10.E.7}$$

Infer that

$$v_s A \leq \lim_{k\to\infty} v_{sk} A. \tag{8.10.E.8}$$

Also,

$$(\forall \varepsilon > 0)\,(\exists k_0)\,(\forall k > k_0) \quad \sum_i |s_k X_i| \leq \sum_i |sX_i| + \varepsilon \leq v_s A + \varepsilon. \tag{8.10.E.9}$$

Proceed.]

Exercise $8.10.E.7$

Let $(X, \mathcal{M}, m)$ and $(Y, \mathcal{N}, n)$ be two generalized measure spaces $(X \in M, Y \in \mathcal{N})$ such that mn is defined (Note 1).
Set

$$\mathcal{C} = \{A \times B | A \in \mathcal{M}, B \in \mathcal{N}, v_m A < \infty, v_n B < \infty\} \tag{8.10.E.10}$$

and $s(A \times B) = mA \cdot nB$ for $A \times B \in \mathcal{C}$.
Define a Fubini map as in §8, Part IV, dropping, however, $\int_{X\times Y} f dp$ from Fubini property (c) temporarily.
Then prove Theorem 1 in §8, including formula (1), for Fubini maps so modified.
[Hint: For σ -additivity, use our present Theorem 2 twice. Omit $\mathcal{P}^*$.]

Exercise $8.10.E.8$

Continuing Problem 7, let $\mathcal{P}$ be the σ-ring generated by $\mathcal{C}$ in $X \times Y$. Prove that $(\forall D \in \mathcal{P}) C_D$ is a Fubini map (as modified).
[Outline: Proceed as in Lemma 5 of §8. For step (ii), use Theorem 2 in §10 twice.]

Exercise $8.10.E.9$

Further continuing Problems 7 and 8, define

$$(\forall D \in \mathcal{P}) \quad pD = \int_X \int_Y C_D dn dm. \tag{8.10.E.11}$$

Show that p is a generalized measure, with $p = s$ on $\mathcal{C}$, and that

$$(\forall D \in \mathcal{P}) \quad pD = \int_{X\times Y} C_D dp, \tag{8.10.E.12}$$

with the following convention: If $X \times Y \notin \mathcal{P}$, we set

$$\int_{X\times Y} f\,dp = \int_H f\,dp \tag{8.10.E.13}$$

whenever $H \in \mathcal{P}$, f is p-integrable on H, and $f=0$ on $-H$.

Verify that this is unambiguous, i.e.,

$$\int_{X\times Y} f\,dp \tag{8.10.E.14}$$

so defined is independent of the choice of H.

Finally, let $\bar{p}$ be the completion of p (Chapter 7, §6, Problem 15); let $\mathcal{P}^*$ be its domain.

Develop the rest of Fubini theory "imitating" Problem 12 in §8.

Exercise $8.10.E.10$

Signed Lebesgue-Stieltjes (LS) measures in E^1 are defined as shown in Chapter 7, §11, Part V. Using the notation of that section, prove the following:

(i) Given a Borel-Stieltjes measure σ_α^* in an interval $I \subseteq E^1$ (or an LS measure $s_\alpha = \bar{\sigma}_\alpha^*$ in I), there are two monotone functions $g \uparrow$ and $h \uparrow$ $(\alpha = g - h)$ such that

$$m_g = s_\alpha^+ \text{and } m_h = s_\alpha^-, \tag{8.10.E.15}$$

both satisfying formula (3) of Chapter 7, §11, inside I.

(ii) If f is s_α-integrable on $A \subseteq I$, then

$$\int_A f\,ds_\alpha = \int_A f\,dm_g - \int_A f\,dm_h \tag{8.10.E.16}$$

for any $g \uparrow$ and $h \uparrow$ (finite) such that $\alpha = g - h$.

[Hints: (i) Define s_α^+ and s_α^- by formula (3) of Chapter 7, §7. Then arguing as in Theorem 2 in Chapter 7, §9, find $g \uparrow$ and $h \uparrow$ with $m_g = s_\alpha^+$ and $m_h = s_\alpha^-$.

(ii) First let $A = (a, b] \subseteq I$, then $A \in B$. Proceed.]

8.11: The Radon–Nikodym Theorem. Lebesgue Decomposition

I. As you know, the indefinite integral

$$\int f d m \tag{8.11.1}$$

is a generalized measure. We now seek conditions under which a given generalized measure μ can be represented as

$$\mu=\int f d m \tag{8.11.2}$$

for some f (to be found). We start with two lemmas.

Lemma 8.11.1

Let $m, \mu: \mathcal{M} \rightarrow[0, \infty)$ be finite measures in S. Suppose $S \in \mathcal{M}, \mu S>0$ (i.e., $\mu \not\equiv 0$) and μ is m-continuous (Chapter 7, §11).

Then there is $\delta>0$ and a set $P \in \mathcal{M}$ such that $m P>0$ and

$$(\forall X \in \mathcal{M}) \quad \mu X \geq \delta \cdot m(X \cap P). \tag{8.11.3}$$

Proof

As $m<\infty$ and $\mu S>0$, there is $\delta>0$ such that

$$\mu S-\delta \cdot m S>0. \tag{8.11.4}$$

Fix such a δ and define a signed measure (Lemma 2 of Chapter 7, §11)

$$\Phi=\mu-\delta m, \tag{8.11.5}$$

so that

$$(\forall Y \in \mathcal{M}) \quad \Phi Y=\mu Y-\delta \cdot m Y; \tag{8.11.6}$$

hence

$$\Phi S=\mu S-\delta \cdot m S>0. \tag{8.11.7}$$

By Theorem 3 in Chapter 7, §11 (Hahn decomposition), there is a Φ-positive set $P \in \mathcal{M}$ with a Φ-negative complement $-P=S-P \in \mathcal{M}$.

Clearly, $m P>0$; for if $m P=0$, the m-continuity of μ would imply $\mu P=0$, hence

$$\Phi P=\mu P-\delta \cdot m P=0, \tag{8.11.8}$$

contrary to $\Phi P \geq \Phi S>0$.

Also, $P \supseteq Y$ and $Y \in \mathcal{M}$ implies $\Phi Y \geq 0$; so by (1),

$$0 \leq \mu Y-\delta \cdot m Y. \tag{8.11.9}$$

Taking $Y=X \cap P,$ we get

$$\delta \cdot m(X \cap P) \leq \mu(X \cap P) \leq \mu X, \tag{8.11.10}$$

as required. $\square$

Lemma 8.11.2

With m, μ, and S as in Lemma 1, let $\mathcal{H}$ be the set of all maps $g: S \rightarrow E^{*}, \mathcal{M}$ -measurable and nonnegative on S, such that

$$\int_{X} g d m \leq \mu X \tag{8.11.11}$$

for every set X from $\mathcal{M}$.

Then there is $f \in \mathcal{H}$ with

$$\int_S f dm = \max_{g\in\mathcal{H}} \int_S g dm. \tag{8.11.12}$$

Proof

$\mathcal{H}$ is not empty; e.g., $g = 0$ is in $\mathcal{H}$. We now show that

$$(\forall g, h \in \mathcal{H}) \quad g \vee h = \max(g, h) \in \mathcal{H}. \tag{8.11.13}$$

Indeed, $g \vee h$ is ≥ 0 and $\mathcal{M}$-measurable on S, as g and h are.

Now, given $X \in \mathcal{M}$, let $Y = X(g > h)$ and $Z = X(g \leq h)$. Dropping "dm" for brevity, we have

$$\int_X (g \vee h) = \int_Y (g \vee h) + \int_Z (g \vee h) = \int_Y g + \int_Z h \leq \mu Y + \mu Z = \mu X, \tag{8.11.14}$$

proving (2).

Let

$$k = \sup_{g\in\mathcal{H}} \int_S g dm \in E^*. \tag{8.11.15}$$

Proceeding as in Problem 13 of Chapter 7, §6, and using (2), one easily finds a sequence $\{g_n\} \uparrow, g_n \in \mathcal{H}$, such that

$$\lim_{n\to\infty} \int_S g_n dm = k. \tag{8.11.16}$$

(Verify!) Set

$$f = \lim_{n\to\infty} g_n. \tag{8.11.17}$$

(It exists since $\{g_n\} \uparrow$.) By Theorem 4 in §6,

$$k = \lim_{n\to\infty} \int_S g_n = \int_S f. \tag{8.11.18}$$

Also, f is $\mathcal{M}$-measurable and ≥ 0 on S, as all g_n are; and if $X \in \mathcal{M}$, then

$$(\forall n) \quad \int_X g_n \leq \mu X; \tag{8.11.19}$$

hence

$$\int_X f = \lim_{n\to\infty} \int_X g_n \leq \mu X. \tag{8.11.20}$$

Thus $f \in \mathcal{H}$ and

$$\int_S f = k = \sup_{g\in H} \int_S g, \tag{8.11.21}$$

i.e.,

$$\int_S f = \max_{g\in\mathcal{H}} \int_S g \leq \mu S < \infty. \tag{8.11.22}$$

This completes the proof. □

Note 1. As $\mu < \infty$ and $f \geq 0$, Corollary 1 in §5 shows that f can be made finite on all of S. Also, f is m-integrable on S.

Theorem 8.11.1 (Radon-Nikodym)

If $(S, \mathcal{M}, m)$ is a σ-finite measure space, if $S \in \mathcal{M}$, and if

$$\mu : \mathcal{M} \rightarrow E^n\left(C^n\right) \tag{8.11.23}$$

is a generalized m-continuous measure, then

$$\mu = \int f dm \text{ on } \mathcal{M} \tag{8.11.24}$$

for at least one map

$$f : S \rightarrow E^n\left(C^n\right), \tag{8.11.25}$$

$\mathcal{M}$-measurable on S.

Moreover, if h is another such map, then $mS(f \neq h) = 0$

The last part of Theorem 1 means that f is "essentially unique." We call f the Radon-Nikodym (RN) derivative of μ, with respect to m.

Proof

Via components (Theorem 5 in Chapter 7, §11), all reduces to the case

$$\mu : \mathcal{M} \rightarrow E^1. \tag{8.11.26}$$

Then Theorem 4 (Jordan decomposition) in Chapter 7, §11, yields

$$\mu = \mu^{+} - \mu^{-}, \tag{8.11.27}$$

where μ^{+} and μ^{-} are finite measures (≥ 0), both m-continuous (Corollary 3 from Chapter 7, §11). Therefore, all reduces to the case $0 \leq \mu < \infty$.

Suppose first that m, too, is finite. Then if $\mu = 0$, just take $f = 0$.

If, however, $\mu S > 0$, take $f \in \mathcal{H}$ as in Lemma 2 and Note 1; f is nonnegative, bounded, and $\mathcal{M}$-measurable on S,

$$\int f \leq \mu < \infty, \tag{8.11.28}$$

and

$$\int_S f dm = k = \sup_{g \in \mathcal{H}} \int_S g dm. \tag{8.11.29}$$

We claim that f is the required map.

Indeed, let

$$\nu = \mu - \int f dm; \tag{8.11.30}$$

so ν is a finite m-continuous measure (≥ 0) on $\mathcal{M}$. (Why?) We must show that $\nu = 0$.

Seeking a contradiction, suppose $\nu S > 0$. Then by Lemma 1, there are $P \in \mathcal{M}$ and $\delta > 0$ such that $mP > 0$ and

$$(\forall X \in \mathcal{M}) \quad \nu X \geq \delta \cdot m(X \cap P). \tag{8.11.31}$$

Now let

$$g = f + \delta \cdot C_P; \tag{8.11.32}$$

so g is $\mathcal{M}$-measurable and ≥ 0. Also,

$$(\forall X\in\mathcal{M})\quad \int_X g=\int_X f+\delta\int_X C_P \begin{aligned}[t] &=\int_X f+\delta\cdot m(X\cap P)\\ &\leq\int_X f+\nu(X\cap P)\\ &\leq\int_X f+\nu X=\mu X\end{aligned}$$

by our choice of δ and ν. Thus $g\in\mathcal{H}$. On the other hand,

$$\int_S g=\int_S f+\delta\int_S C_P=k+\delta mP>k, \tag{8.11.33}$$

contrary to

$$k=\sup_{g\in\mathcal{H}}\int_S g. \tag{8.11.34}$$

This proves that $\int f=\mu$, indeed.

Now suppose there is another map $h\in\mathcal{H}$ with

$$\mu=\int hdm=\int fdm\neq\infty; \tag{8.11.35}$$

so

$$\int(f-h)dm=0. \tag{8.11.36}$$

(Why?) Let

$$Y=S(f\geq h)\text{ and }Z=S(f<h); \tag{8.11.37}$$

so $Y,Z\in\mathcal{M}$ (Theorem 3 of §2) and $f-h$ is sign-constant on Y and Z. Also, by construction,

$$\int_Y(f-h)dm=0=\int_Z(f-h)dm. \tag{8.11.38}$$

Thus by Theorem 1(h) in §5, $f-h=0$ a.e. on Y, on Z, and hence on $S=Y\cup Z$ that is,

$$mS(f\neq h)=0. \tag{8.11.39}$$

Thus all is proved for the case $mS<\infty$.

Next, let m be σ-finite:

$$S=\bigcup_{k=1}^{\infty}S_k\text{ (disjoint)} \tag{8.11.40}$$

for some sets $S_k\in\mathcal{M}$ with $mS_k<\infty$.

By what was shown above, on each S_k there is an $\mathcal{M}$-measurable map $f_k\geq 0$ such that

$$\int_X f_k dm=\mu X \tag{8.11.41}$$

for all $\mathcal{M}$-sets $X\subseteq S_k$. Fixing such an f_k for each k, define $f:S\rightarrow E^1$ by

$$f=f_k\quad\text{ on }S_k,\quad k=1,2,\ldots. \tag{8.11.42}$$

Then (Corollary 3 in §1) f is $\mathcal{M}$-measurable and ≥ 0 on S.

Taking any $X\in\mathcal{M}$, set $X_k=X\cap S_k$. Then

$$X=\bigcup_{k=1}^{\infty}X_k\text{ (disjoint)} \tag{8.11.43}$$

and $X_k \in \mathcal{M}$. Also,

$$(\forall k) \quad \int_{X_k} f dm = \int_{X_k} f_k dm = \mu X_k. \tag{8.11.44}$$

Thus by σ-additivity (Theorem 2 in §5),

$$\int_X f dm = \sum_{k=1}^{\infty} \int_{X_k} f dm = \sum_k \mu X_k = \mu X < \infty \quad (\mu \text{ is finite!}). \tag{8.11.45}$$

Thus f is as required, and its "uniqueness" follows as before. □

Note 2. By Definition 3 in §10, we may write

$$\text{"} d\mu = f dm \text{"} \tag{8.11.46}$$

for

$$\text{"} \int f dm = \mu. \text{"} \tag{8.11.47}$$

Note 3. Using Definition 2 in §10 and an easy "componentwise" proof, one shows that Theorem 1 holds also with m replaced by a generalized measure s. The formulas

$$\mu = \int f dm \text{ and } mS(f \neq h) = 0 \tag{8.11.48}$$

then are replaced by

$$\mu = \int f ds \text{ and } v_s S(f \neq h) = 0. \tag{8.11.49}$$

II. Theorem 1 requires μ to be m-continuous $(\mu \ll m)$. We want to generalize Theorem 1 so as to lift this restriction. First, we introduce a new concept.

Definition

Given two set functions $s, t : \mathcal{M} \rightarrow E \left(\mathcal{M} \subseteq 2^S\right)$, we say that s is t-singular $(s \perp t)$ iff there is a set $P \in \mathcal{M}$ such that $v_t P = 0$ and

$$(\forall X \in \mathcal{M} | X \subseteq -P) \quad sX = 0. \tag{8.11.50}$$

(We then briefly say "s resides in P.")

For generalized measures, this means that

$$(\forall X \in \mathcal{M}) \quad sX = s(X \cap P). \tag{8.11.51}$$

Why?

Corollary 8.11.1

If the generalized measures $s, u : \mathcal{M} \rightarrow E$ are t-singular, so is ks for any scalar k (if s is scalar valued, k may be a vector).

So also are $s \pm u$, provided t is additive.

Proof

(Exercise! See Problem 3 below.)

Corollary 8.11.2

If a generalized measure $s: \mathcal{M} \rightarrow E$ is t-continuous $(s \ll t)$ and also t-singular $(s \perp t)$, then $s=0$ on $\mathcal{M}$.

Proof

As $s \perp t$, formula (3) holds for some $P \in \mathcal{M}, v_t P=0$. Hence for all $X \in \mathcal{M}$,

$$s(X-P)=0 \text{ (for } X-P \subseteq -P) \tag{8.11.52}$$

and

$$v_t(X \cap P)=0 \text{ (for } X \cap P \subseteq P). \tag{8.11.53}$$

As $s \ll t$, we also have $s(X \cap P)=0$ by Definition 3(i) in Chapter 7, §11. Thus by additivity,

$$sX=s(X \cap P)+s(X-P)=0, \tag{8.11.54}$$

as claimed. □

Theorem 8.11.2 (Lebesgue decomposition)

Let $s, t: \mathcal{M} \rightarrow E$ be generalized measures.

If v_s is t-finite (Definition 3(iii) in Chapter 7, §11), there are generalized measures $s', s'': \mathcal{M} \rightarrow E$ such that

$$s' \ll t \text{ and } s'' \perp t \tag{8.11.55}$$

and

$$s=s'+s''. \tag{8.11.56}$$

Proof

Let v_0 be the restriction of v_s to

$$\mathcal{M}_o=\{X \in \mathcal{M} | v_t X=0\}. \tag{8.11.57}$$

As v_s is a measure (Theorem 1 of Chapter 7, §11), so is v_0 (for $\mathcal{M}_0$ is a σ-ring; verify!).

Thus by Problem 13 in Chapter 7, §6, we fix $P \in \mathcal{M}_0$, with

$$v_s P=v_0 P=\max\{v_s X | X \in \mathcal{M}_0\}. \tag{8.11.58}$$

As $P \in \mathcal{M}_0$, we have $v_t P=0$; hence

$$|sP| \leq v_s P<\infty \tag{8.11.59}$$

(for v_s is t-finite).

Now define s', s'', v', and v'' by setting, for each $X \in \mathcal{M}$,

$$\begin{aligned} s'X &= s(X-P); \\ s''X &= s(X \cap P); \\ v'X &= v_s(X-P); \\ v''X &= v_s(X \cap P). \end{aligned}$$

As s and v_s are σ-additive, so are s', s'', v', and v''. (Verify!) Thus $s', s'': \mathcal{M} \rightarrow E$ are generalized measures, while v' and v'' are measures (≥ 0).

We have

$$(\forall X \in \mathcal{M}) \quad sX=s(X-P)+s(X \cap P)=s'X+s''X; \tag{8.11.60}$$

i.e.,

$$s=s'+s''. \tag{8.11.61}$$

Similarly one obtains $v_s = v' + v''$.

Also, by (5), since $X \cap P = \emptyset$,

$$-P \supseteq X \text{ and } X \in \mathcal{M} \Longrightarrow s''X = 0, \tag{8.11.62}$$

while $v_t P = 0$ (see above). Thus s'' is t-singular, residing in P.

To prove $s' \ll t$, it suffices to show that $v' \ll t$ (for by (4) and (6), $v'X = 0$ implies $|s'X| = 0$).

Assume the opposite. Then

$$(\exists Y \in \mathcal{M}) \quad v_t Y = 0 \tag{8.11.63}$$

(i.e., $Y \in \mathcal{M}_0$), but

$$0 < v'Y = v_s(Y - P). \tag{8.11.64}$$

So by additivity,

$$v_s(Y \cup P) = v_s P + v_s(Y - P) > v_s P, \tag{8.11.65}$$

with $Y \cup P \in \mathcal{M}_0$, contrary to

$$v_s P = \max\{v_s X | X \in \mathcal{M}_0\}. \tag{8.11.66}$$

This contradiction completes the proof. $\square$

Note 4. The set function s'' in Theorem 2 is bounded on $\mathcal{M}$. Indeed, $s'' \perp t$ yields a set $P \in \mathcal{M}$ such that

$$(\forall X \in \mathcal{M}) \quad s''(X - P) = 0; \tag{8.11.67}$$

and $v_t P = 0$ implies $v_s P < \infty$. (Why?) Hence

$$s''X = s''(X \cap P) + s''(X - P) = s''(X \cap P). \tag{8.11.68}$$

As $s = s' + s''$, we have

$$|s''| \leq |s| + |s'| \leq v_s + v_{s'}; \tag{8.11.69}$$

so

$$|s''X| = |s''(X \cap P)| \leq v_s P + v_{s'} P. \tag{8.11.70}$$

But $v_{s'} P = 0$ by t-continuity (Theorem 2 of Chapter 7, §11). Thus $|s''| \leq v_s P < \infty$ on $\mathcal{M}$.

Note 5. The Lebesgue decomposition $s = s' + s''$ in Theorem 2 is unique. For if also

$$u' \ll t \text{ and } u'' \perp t \tag{8.11.71}$$

and

$$u' + u'' = s = s' + s'', \tag{8.11.72}$$

then with P as in Problem 3, $(\forall X \in \mathcal{M})$

$$s'(X \cap P) + s''(X \cap P) = u'(X \cap P) + u''(X \cap P) \tag{8.11.73}$$

and $v_t(X \cap P) = 0$. But

$$s'(X \cap P) = 0 = u'(X \cap P) \tag{8.11.74}$$

by t-continuity; so (8) reduces to

$$s''(X \cap P) = u''(X \cap P), \tag{8.11.75}$$

or $s''X = u''X$ (for s'' and u'' reside in P). Thus $s'' = u''$ on $\mathcal{M}$.

By Note 4, we may cancel s'' and u'' in

$$s' + s'' = u' + u'' \tag{8.11.76}$$

to obtain $s' = u'$ also.

Note 6. If $E = E^n\left(C^n\right)$, the t-finiteness of v_s in Theorem 2 is redundant, for v_s is even bounded (Theorem 6 in Chapter 7, §11).

We now obtain the desired generalization of Theorem 1.

Corollary 8.11.3

If $(S, \mathcal{M}, m)$ is a σ-finite measure space $(S \in \mathcal{M})$, then for any generalized measure

$$\mu : \mathcal{M} \rightarrow E^n\left(C^n\right), \tag{8.11.77}$$

there is a unique m-singular generalized measure

$$s'' : \mathcal{M} \rightarrow E^n\left(C^n\right) \tag{8.11.78}$$

and a ("essentially" unique) map

$$f : S \rightarrow E^n\left(C^n\right), \tag{8.11.79}$$

$\mathcal{M}$-measurable and m-integrable on S, with

$$\mu = \int f dm + s''. \tag{8.11.80}$$

(Note 3 applies here.)

Proof

By Theorem 2 and Note 5, $\mu = s' + s''$ for some (unique) generalized measures $s', s'' : \mathcal{M} \rightarrow E^n\left(C^n\right)$, with $s' \ll m$ and $s'' \perp m$.

Now use Theorem 1 to represent s' as $\int f dm$, with f as stated. This yields the result. $\square$

8.11.E: Problems on Radon-Nikodym Derivatives and Lebesgue Decomposition

Exercise $8.11.E.1$

Fill in all proof details in Lemma 2 and Theorem 1.

Exercise $8.11.E.2$

Verify the statement following formula (3). Also prove the following:
(i) If $P \in \mathcal{M}$ along with $-P \in \mathcal{M}$, then $s \perp t$ implies $t \perp s$;
(ii) $s \perp t$ iff $v_s \perp t$.

Exercise $8.11.E.3$

Prove Corollary 1.
[Hints: Here $\mathcal{M}$ is a σ-ring. Suppose s and u reside in P' and P'', respectively, and $v_t P' = 0 = v_t P''$. Let $P = P' \cup P'' \in \mathcal{M}$. Verify that $v_t P = 0$ (use Problem 8 in Chapter $7, §11$). Then show that both s and u reside in P.]

Exercise $8.11.E.4$

Show that if $s : \mathcal{M} \to E^*$ is a signed measure in $S \in \mathcal{M}$, then $s^+ \perp s^-$ and $s^- \perp s^+$.

Exercise $8.11.E.5$

Fill in all details in the proof of Theorem 2. Also prove the following:
(i) s' and $v_{s'}$ are absolutely t-continuous.
[Hint: Use Theorem 2 in Chapter 7, §11.]
(ii) $v_s = v_{s'} + v_{s''}, v_{s''} \perp t$.
(iii) If s is a measure (≥ 0), so are s' and s''.

Exercise $8.11.E.6$

Verify Note 3 for Theorem 1 and Corollary 3. State and prove both generalized propositions precisely.

8.12: Integration and Differentiation

I. We shall now link RN-derivatives (§11) to those of Chapter 7, §12.

Below, we use the notation of Definition 3 in Chapter 7, §10 and Definition 1 of Chapter 7, §12. (Review them!) In particular,

$$m : \mathcal{M}^* \rightarrow E^* \tag{8.12.1}$$

is Lebesgue measure in E^n (presupposed in such terms as "a.e.," etc.); s is an arbitrary set function. For convenience, we set

$$s^{\prime}(\overline{p})=0 \tag{8.12.2}$$

and

$$\int_X f dm=0, \tag{8.12.3}$$

unless defined otherwise; thus $s^{\prime}$ and $\int_X f$ exist always.

We start with several lemmas that go back to Lebesgue.

Lemma 8.12.1

With the notation of Definition 3 of Chapter 7, §10, the functions

$$\overline{D} s, \underline{D} s, \text{ and } s^{\prime} \tag{8.12.4}$$

are Lebesgue measurable on E^n for any set function

$$s : \mathcal{M}^{\prime} \rightarrow E^* \quad\left(\mathcal{M}^{\prime} \supseteq \overline{\mathcal{K}}\right). \tag{8.12.5}$$

Proof

By definition,

$$\overline{D} s(\overline{p})=\inf_r h_r(\overline{p}), \tag{8.12.6}$$

where

$$h_r(\overline{p})=\sup \left\{\frac{s I}{m I} | I \in \mathcal{K}_{\overline{p}}^r\right\} \tag{8.12.7}$$

and

$$\mathcal{K}_{\overline{p}}^r=\left\{I \in \overline{\mathcal{K}} | \overline{p} \in I, d I<\frac{1}{r}\right\}, \quad r=1,2, \ldots . \tag{8.12.8}$$

As is easily seen (verify!),

$$E^n\left(h_r>a\right)=\bigcup\left\{I \in \overline{\mathcal{K}} | a<\frac{s I}{m I}, d I<\frac{1}{r}\right\}, \quad a \in E^*. \tag{8.12.9}$$

The right-side union is Lebesgue measurable by Problem 2 in Chapter 7, §10. Thus by Theorem 1 of §2, the function h_r is measurable on E^n for $r=1,2, \ldots$ and so is

$$\overline{D} s=\inf_r h_r \tag{8.12.10}$$

by Lemma 1 of §2 and Definition 3 in Chapter 7, §10. Similarly for $\underline{D} s$.

Hence by Corollary 1 in §2, the set

$$A=E^n(\underline{D} s=\overline{D} s) \tag{8.12.11}$$

is measurable. As $s'=\overline{D}s$ on A, s' is measurable on A and also on $-A$ (by convention, $s'=0$ on $-A$), hence on all of E^n. □

Lemma 8.12.2

With the same notation, let $s:\mathcal{M}'\rightarrow E^*\left(\mathcal{M}'\supseteq\overline{\mathcal{K}}\right)$ be a regular measure in E^n. Let $A\in\mathcal{M}^*$ and $B\in\mathcal{M}'$ with $A\subseteq B$, and $a\in E^1$.

If

$$\overline{D}s>a \quad \text{on } A, \tag{8.12.12}$$

then

$$a\cdot mA\leq sB. \tag{8.12.13}$$

Proof

Fix $\varepsilon>0$. By regularity (Definition 4 in Chapter 7, §7), there is an open set $G\supseteq B$, with

$$sB+\varepsilon\geq sG. \tag{8.12.14}$$

Now let

$$\mathcal{K}^{\varepsilon}=\{I\in\overline{\mathcal{K}}\mid I\subseteq G, sI\geq(a-\varepsilon)mI\}. \tag{8.12.15}$$

As $\overline{D}s>a$, the definition of $\overline{D}s$ implies that $\mathcal{K}^{\varepsilon}$ is a Vitali covering of A. (Verify!)

Thus Theorem 1 in Chapter 7, §10, yields a disjoint sequence $\{I_k\}\subseteq\mathcal{K}^{\varepsilon}$, with

$$m\left(A-\bigcup_k I_k\right)=0 \tag{8.12.16}$$

and

$$mA\leq m\left(A-\bigcup I_k\right)+m\bigcup I_k=0+m\bigcup I_k=\sum_k mI_k. \tag{8.12.17}$$

As

$$\bigcup I_k\subseteq G \text{ and } sB+\varepsilon\geq sG \tag{8.12.18}$$

(by our choice of $\mathcal{K}^{\varepsilon}$ and G, we obtain

$$sB+\varepsilon\geq s\bigcup_k I_k=\sum_k sI_k\geq(a-\varepsilon)\sum_k mI_k\geq(a-\varepsilon)mA. \tag{8.12.19}$$

Thus

$$(a-\varepsilon)mA\leq sB+\varepsilon. \tag{8.12.20}$$

Making $\varepsilon\rightarrow 0$, we obtain the result. □

Lemma 8.12.3

If

$$t=s\pm u, \tag{8.12.21}$$

with $s,t,u:\mathcal{M}'\rightarrow E^*$ and $\mathcal{M}'\supseteq\overline{\mathcal{K}}$, and if u is differentiable at a point $\overline{p}\in E^n$, then

$$\overline{D}t=\overline{D}s\pm u' \text{ and } \underline{D}t=\underline{D}s\pm u' \text{ at } \overline{p}. \tag{8.12.22}$$

Proof

The proof, from definitions, is left to the reader (Chapter 7, §12, Problem 7).

Lemma 8.12.4

Any m-continuous measure $s : \mathcal{M}^* \rightarrow E^1$ is strongly regular.

Proof

By Corollary 3 of Chapter 7, §11, $v_s = s < \infty$ (s is finite!). Thus v_s is certainly m-finite.

Hence by Theorem 2 in Chapter 7, §11, s is absolutely m-continuous. So given $\varepsilon > 0$, there is $\delta > 0$ such that

$$(\forall X \in \mathcal{M}^* | mX < \delta) \quad sX < \varepsilon. \tag{8.12.23}$$

Now, let $A \in \mathcal{M}^*$. By the strong regularity of Lebesgue measure m (Chapter 7, §8, Theorem 3(b)), there is an open set $G \supseteq A$ and a closed $F \subseteq A$ such that

$$m(A-F) < \delta \text{ and } m(G-A) < \delta. \tag{8.12.24}$$

Thus by our choice of δ,

$$s(A-F) < \varepsilon \text{ and } s(G-A) < \varepsilon, \tag{8.12.25}$$

as required. $\square$

Lemma 8.12.5

Let $s, s_k (k = 1, 2, \ldots)$ be finite m-continuous measures, with $s_k \nearrow s$ or $s_k \searrow s$ on $\mathcal{M}^*$.

If the s_k are a.e. differentiable, then

$$\overline{D}s = \underline{D}s = \lim_{k\rightarrow\infty} s_k' \text{ a.e.} \tag{8.12.26}$$

Proof

Let first $s_k \nearrow s$. Set

$$t_k = s - s_k. \tag{8.12.27}$$

By Corollary 2 in Chapter 7, §11, all t_k are m-continuous, hence strongly regular (Lemma 4). Also, $t_k \searrow 0$ (since $s_k \nearrow s$). Hence

$$t_k I \geq t_{k+1} I \geq 0 \tag{8.12.28}$$

for each cube I; and the definition of $\overline{D}t_k$ implies that

$$\overline{D}t_k \geq \overline{D}t_{k+1} \geq \underline{D}t_{k+1} \geq 0. \tag{8.12.29}$$

As $\{\overline{D}t_k\} \downarrow, \lim_{k\rightarrow\infty} \overline{D}t_k$ exists (pointwise). Now set

$$A_r = E^n\left(\lim_{k\rightarrow\infty} \overline{D}t_k \geq \frac{1}{r}\right), \quad r = 1, 2, \ldots. \tag{8.12.30}$$

By Lemma 1 (and Lemma 1 in §2), $A_r \in \mathcal{M}^*$. Since

$$\overline{D}t_k \geq \lim_{i\rightarrow\infty} \overline{D}t_i \geq \frac{1}{r} \tag{8.12.31}$$

on A_r, Lemma 2 yields

$$\frac{1}{r} mA_r \leq t_k A_r. \tag{8.12.32}$$

As $t_k \searrow 0$, we have

$$\frac{1}{r} m A_r \leq \lim_{k\to\infty} t_k A_r = 0. \tag{8.12.33}$$

Thus

$$mA_r = 0, \quad r = 1, 2, \ldots. \tag{8.12.34}$$

Also, as is easily seen

$$E^n\left(\lim_{k\to\infty} \overline{D} t_k > 0\right) = \bigcup_{r=1}^{\infty} E^n\left(\lim_{k\to\infty} \overline{D} t_k \geq \frac{1}{r}\right) = \bigcup_{r=1}^{\infty} A_r \tag{8.12.35}$$

and

$$m \bigcup_{r=1}^{\infty} A_r = 0. \tag{8.12.36}$$

Hence

$$\lim_{k\to\infty} \overline{D} t_k \leq 0 \quad \text{a.e.} \tag{8.12.37}$$

As

$$\overline{D} t_k \geq \underline{D} t_k \geq 0 \tag{8.12.38}$$

(see above), we get

$$\lim_{k\to\infty} \overline{D} t_k = 0 = \lim_{k\to\infty} D t_k \quad \text{ a.e. on } E^n. \tag{8.12.39}$$

Now, as $t_k = s - s_k$ and as the s_k are differentiable, Lemma 3 yields

$$\overline{D} t_k = \overline{D} s - s_k' \text{ and } \underline{D} t_k = \underline{D} s - s_k' \quad \text{a.e.} \tag{8.12.40}$$

Thus

$$\lim_{k\to\infty}\left(\overline{D} s - s_k'\right) = 0 = \lim\left(\underline{D} s - s_k'\right), \tag{8.12.41}$$

i.e.,

$$\overline{D} s = \lim_{k\to\infty} s_k' = \underline{D} s \quad \text{a.e.} \tag{8.12.42}$$

This settles the case $s_k \nearrow s$.

In the case $s_k \searrow s$, one only has to set $t_k = s_k - s$ and proceed as before. (Verify!) □

Lemma 8.12.6

Given $A \in \mathcal{M}^*, mA < \infty$, let

$$s = \int C_A dm \tag{8.12.43}$$

on $\mathcal{M}^*$. Then s is a.e. differentiable, and

$$s' = C_A \text{ a.e. on } E^n. \tag{8.12.44}$$

$(C_A =$ characteristic function of $A.)$

Proof

First, let A be open and let $\overline{p} \in A$.

Then A contains some $G_{\overline{p}}(\delta)$ and hence also all cubes $I \in \overline{\mathcal{K}}$ with $dI < \delta$ and $\overline{p} \in I$.

Thus for such $I \in \overline{\mathcal{K}}$,

$$sI = \int_I C_A dm = \int_I (1) dm = mI; \tag{8.12.45}$$

i.e.,

$$\frac{sI}{mI} = 1 = C_A(\overline{p}), \quad \overline{p} \in A. \tag{8.12.46}$$

Hence by Definition 1 of Chapter 7, §12,

$$s'(\overline{p}) = 1 = C_A(\overline{p}) \tag{8.12.47}$$

if $\overline{p} \in A$; i.e., $s' = C_A$ on A.

We calim that

$$\overline{D}s = s' = 0 \quad \text{a.e. on } -A. \tag{8.12.48}$$

To prove it, note that

$$s = \int C_A dm \tag{8.12.49}$$

is a finite (why?) m-continuous measure on $\mathcal{M}^*$. By Lemma 4, s is strongly regular. Also, as $sI \geq 0$ for any $I \in \overline{\mathcal{K}}$, we certainly have

$$\overline{D}s \geq \underline{D}s \geq 0. \tag{8.12.50}$$

(Why?) Now let

$$B = E^n(\overline{D}s > 0) = \bigcup_{r=1}^{\infty} B_r, \tag{8.12.51}$$

where

$$B_r = E^n\left(\overline{D}s \geq \frac{1}{r}\right), \quad r = 1, 2, \ldots . \tag{8.12.52}$$

We have to show that $m(B - A) = 0$.

Suppose

$$m(B - A) > 0. \tag{8.12.53}$$

Then by (2), we must have $m\left(B_r - A\right) > 0$ for at least one B_r; we fix this B_r Also, by (3),

$$\overline{D}s \geq \frac{1}{r} \text{ on } B_r - A \tag{8.12.54}$$

(even on all of B_r). Thus by Lemma 2,

$$0 < \frac{1}{r} m\left(B_r - A\right) \leq s\left(B_r - A\right) = \int_{B_r - A} C_A dm. \tag{8.12.55}$$

But this is impossible. Indeed, as $C_A = 0$ on $-A$ (hence on $B_r - A$),the integral in (4) cannot be > 0. This refutes the assumption $m(B - A) > 0$; so by (2),

$$m\left(E^n(\overline{D}s > 0) - A\right) = 0; \tag{8.12.56}$$

i.e.,

$$\overline{D}s = 0 = \underline{D}s \quad \text{a.e. on } -A. \tag{8.12.57}$$

We see that

$$s' = 0 = C_A \quad \text{a.e. on } -A, \tag{8.12.58}$$

and

$$s' = 1 = C_A \quad \text{on } A, \tag{8.12.59}$$

proving the lemma for open sets A.

Now take any $A \in \mathcal{M}^*, mA < \infty$. As Lebesgue measure is regular (Chapter 7, §8, Theorem 3(b)), we find for each $k \in N$ an open set $G_k \supseteq A$, with

$$m(G_k - A) < \frac{1}{k} \text{ and } G_k \supseteq G_{k+1}. \tag{8.12.60}$$

Let

$$s_k = \int C_{G_k} dm. \tag{8.12.61}$$

Then $s_k \searrow s$ on $\mathcal{M}^*$ (see Problem 5 (ii) in §6). Also, by what was shown above, the s_k are differentiable, with $s_k' = C_{G_k}$ a.e.

Hence by Lemma 5,

$$\overline{D}s = \underline{D}s = \lim_{k\to\infty} C_{G_k} = C_A \text{ (a.e.)}. \tag{8.12.62}$$

The lemma is proved. □

Theorem 8.12.1

Let $f : E^n \to E^* \left(E^r, C^r\right)$ be m-integrable, at least on each cube in E^n. Then the set function

$$s = \int f dm \tag{8.12.63}$$

is differentiable, with $s' = f$, a.e. on E^n.

Thus s' is the RN-derivative of s with respect to Lebesgue measure m (Theorem 1 in §11).

Proof

As E^n is a countable union of cubes (Lemma 2 in Chapter 7, §2), it suffices to show that $s' = f$ a.e. on each open cube J, with s differentiable a.e. on J.

Thus fix such a $J \neq \emptyset$ and restrict s and m to

$$\mathcal{M}_0 = \{X \in \mathcal{M}^* | X \subseteq J\}. \tag{8.12.64}$$

This does not affect s' on J; for as J is open, any sequence of cubes

$$I_k \to \bar{p} \in J \tag{8.12.65}$$

terminates inside J anyway.

When so restricted,

$$s = \int f \tag{8.12.66}$$

is a generalized measure in J; for $\mathcal{M}_0$ is a σ-ring (verify!), and f is integrable on J. Also, m is strongly regular, and s is m-continuous.

First, suppose f is $\mathcal{M}_0$-simple on J, say,

$$f=\sum_{i=1}^{q} a_i C_{A_i}, \tag{8.12.67}$$

say, with $0<a_i<\infty$, $A_i \in \mathcal{M}^*$, and

$$J=\bigcup_{i=1}^{q} A_i \text{ (disjoint).} \tag{8.12.68}$$

Then

$$s=\int f=\sum_{i=1}^{q} a_i \int C_{A_i}. \tag{8.12.69}$$

Hence by Lemma 6 above and by Theorem 1 in Chapter 7, §12, s is differentiable a.e. (as each $\int C_{A_i}$ is), and

$$s^{\prime}=\sum_{i=1}^{q} a_i\left(\int C_{A_i}\right)^{\prime}=\sum_{i=1}^{q} a_i C_{A_i}=f \text{ (a.e.)}, \tag{8.12.70}$$

as required.

The general case reduces (via components and the formula $f=f^{+}-f^{-}$) to the case $f \geq 0$, with f measurable (even integrable) on J.

By Problem 6 in §2, then, we have $f_k \nearrow f$ for some simple maps $f_k \geq 0$. Let

$$s_k=\int f_k \text{ on } M_0, k=1,2, \ldots. \tag{8.12.71}$$

Then all s_k and $s=\int f$ are finite measures and $s_k \nearrow s$, by Theorem 4 in §6. Also, by what was shown above, each s_k is differentiable a.e. on J, with $s_k^{\prime}=f_k$ (a.e.). Thus as in Lemma 5,

$$\overline{D} s=\underline{D} s=s^{\prime}=\lim _{k \rightarrow \infty} s_k^{\prime}=\lim f_k=f \text{ (a.e.) on } J, \tag{8.12.72}$$

with $s^{\prime}=f \neq \pm \infty$ (a.e.), as f is integrable on J. Thus all is proved. □

II. So far we have considered Lebesgue $(\overline{\mathcal{K}})$ differentiation. However, our results easily extend to Ω-differentiation (Definition 2 in Chapter 7, §12).

The proof is even simpler. Thus in Lemma 1, the union in formula (1) is countable (as $\overline{\mathcal{K}}$ is replaced by the countable set family Ω); hence it is μ-measurable. In Lemma 2, the use of the Vitali theorem is replaced by Theorem 3 in Chapter 7, §12. Otherwise, one only has to replace Lebesgue measure m by μ on $\mathcal{M}$. Once the lemmas are established (reread the proofs!), we obtain the following.

Theorem 8.12.2

Let S, ρ, Ω, and $\mu: \mathcal{M} \rightarrow E^*$ be as in Definition 2 of Chapter 7, §12. Let $f: S \rightarrow E^*\left(E^r, C^r\right)$ be mu-integrable on each $A \in \mathcal{M}$ with $\mu A<\infty$.

Then the set function

$$s=\int f d \mu \tag{8.12.73}$$

is Ω-differentiable, with $s^{\prime}=f$, (a.e.) on S.

Proof

Recall that S is a countable union of sets $U_n^i \in \Omega$ with $0<\mu U_n^i<\infty$. As μ^* is $\mathcal{G}$-regular, each U_n^i lies in an open set $J_n^i \in \mathcal{M}$ with

$$\mu J_n^i < \mu U_n^i + \varepsilon_n^i < \infty. \tag{8.12.74}$$

Also, f is μ-measurable (even integrable) on J_n^i. Dropping a null set, assume that f is $\mathcal{M}$-measurable on $J = J_n^i$.

From here, proceed exactly as in Theorem 1, replacing m by μ. □

Both theorems combined yield the following result.

Corollary 8.12.1

If $s : \mathcal{M}' \to E^* \left(E^r, C^r\right)$ is an m-continuous and m-finite generalized measure in E^n, then s is $\overline{\mathcal{K}}$-differentiable a.e. on E^n, and $ds = s' dm$ (see Definition 3 in §10) in any $A \in \mathcal{M}^*(mA < \infty)$.

Similarly for Ω-differentiation.

Proof

Given $A \in \mathcal{M}^*(mA < \infty)$, there is an open set $J \supseteq A$ such that

$$mJ < mA + \varepsilon < \infty. \tag{8.12.75}$$

As before, restrict s and m to

$$\mathcal{M}_0 = \{X \in \mathcal{M}^* | X \subseteq J\}. \tag{8.12.76}$$

Then by assumption, s is finite and m-continuous on $\mathcal{M}_0$ (a σ-ring); so by Theorem 1 in §11,

$$s = \int f dm \tag{8.12.77}$$

on $\mathcal{M}_0$ for some m-integrable map f on J.

Hence by our present Theorem 1, s is differentiable, with $s' = f$ a.e. on J and so

$$s = \int f = \int s' \text{ on } \mathcal{M}_0. \tag{8.12.78}$$

This implies $ds = s' dm$ in A.

For Ω-differentiation, use Theorem 2. □

Corollary 8.12.2 (change of measure)

Let s be as in Corollary 1. Subject to Note 1 in §10, if f is s-integrable on $A \in \mathcal{M}^*(mA < \infty)$, then fs' is m-integrable on A and

$$\int_A f ds = \int_A f s' dm. \tag{8.12.79}$$

Similarly for Ω-derivatives, with m replaced by μ.

Proof

By Corollary 1, $ds = s' dm$ in A. Thus Theorem 6 of §10 yields the result. □

Note 1. In particular, Corollary 2 applies to m-continuous signed LS measures $s = s_\alpha$ in E^1 (see end of §11). If $A = [a, b]$, then s_α is surely finite on s_α-measurable subsets of A; so Corollaries 1 and 2 show that

$$\int_A f ds_\alpha = \int_A f s_\alpha' dm = \int_A f \alpha' dm, \tag{8.12.80}$$

since $s_\alpha' = \alpha'$. (See Problem 9 in Chapter 7, §12.)

Note 2. Moreover, $s=s_\alpha$ (see Note 1) is absolutely m-continuous iff α is absolutely continuous in the stronger sense (Problem 2 in Chapter 4, §8).

Indeed, assuming the latter, fix $\varepsilon>0$ and choose δ as in Definition 3 of Chapter 7, §11. Then if $mX<\delta$, we have

$$X \subseteq \bigcup I_k \text{ (disjoint)} \tag{8.12.81}$$

for some intervals $I_k=(a_k, b_k]$, with

$$\delta>\sum m I_k=\sum\left(b_k-a_k\right). \tag{8.12.82}$$

Hence

$$|s X| \leq \sum\left|s I_k\right|<\varepsilon. \tag{8.12.83}$$

(Why?) Similarly for the converse.

8.12.E: Problems on Differentiation and Related Topics

Exercise $8.12.E.1$

Fill in all proof details in this section. Verify footnote 4 and Note 2.

Exercise $8.12.E.2$

Given a measure $s : \mathcal{M}' \rightarrow E^* \left(\mathcal{M}' \supseteq \overline{\mathcal{K}}\right)$, prove that

(i) s is topological;
(ii) its Borel restriction σ is strongly regular; and
(iii) $\underline{D}s, \overline{D}s$, and s' do not change if s or m are restricted to the Borel field $\mathcal{B}$ in E^n; neither does this affect the propositions on $\overline{\mathcal{K}}$-differentiation proved here.
[Hints: (i) Use Lemma 2 of Chapter 7, §2. (ii) Use also Problem 10 in Chapter 7, §7. (iii) All depends on $\overline{\mathcal{K}.]}$

Exercise $8.12.E.3$

What analogues to $2(\mathrm{i})-(\text{ iii })$ apply to Ω-differentiation in $E^n? \operatorname{In}(S, \rho)$?

Exercise $8.12.E.4$

(i) Show that any m-singular measure s in E^n, finite on $\overline{\mathcal{K}}$, has a zero derivative (a.e.).
(ii) For Ω-derivatives, prove that this holds if s is also regular.
[Hint for (i): By Problem 2, we may assume s regular (if not, replace it by σ).
Suppose

$$m E^n(\overline{D} s>0)>a>0 \tag{8.12.E.1}$$

and find a contradiction to Lemma 2.]

Exercise $8.12.E.5$

Give another proof for Theorem 4 in Chapter $7, 812$.
[Hint: Fix an open cube $J \in \overline{\mathcal{K}}$. By Problem 2(iii), restrict s and m to

$$\mathcal{M}_0=\{X \in \mathcal{B} | X \subseteq J\} \tag{8.12.E.2}$$

to make them finite. Apply Corollary 2 in §11 to s. Then use Problem 4, Theorem 1 of the present section, and Theorem 1 of Chapter 7, §12.
For Ω-differentiation, assume s regular; argue as in Corollary 1, using Corollary 2
of 11.]

Exercise $8.12.E.6$

Prove that if

$$F(x)=L \int_a^x f d m \quad(a \leq x \leq b), \tag{8.12.E.3}$$

with $f : E^1 \rightarrow E^*\left(E^n, C^n\right) m$ -integrable on $A=[a, b]$, then F is differentiable, with $F'=f$, a.e. on A.

[Hint: Via components, reduce all to the case $f \geq 0, F \uparrow$ on A.
Let

$$s=\int f dm \tag{8.12.E.4}$$

on $\mathcal{M}^{*}$. Let $t=m_{F}$ be the F-induced LS measure. Show that $s=t$ on intervals in A; so $s^{\prime}=t^{\prime}=F^{\prime}$ a.e. on A (Problem 9 in Chapter 7, §11). Use Theorem 1.]

CHAPTER OVERVIEW

9: Calculus Using Lebesgue Theory

9.1: L-Integrals and Antiderivatives

I. Lebesgue theory makes it possible to strengthen many calculus theorems. We shall start with functions on E^1, $f : E^1 \rightarrow E$. (A reader who has omitted the "starred" part of Chapter 8, §7, will have to set $E = E^* \ (E^n, C^n)$ throughout.)

By L-integrals of such functions, we mean integrals with respect to Lebesgue measure m in E^1. Notation:

$$L \int_a^b f = L \int_a^b f(x) dx = L \int_{[a,b]} f \tag{9.1.1}$$

and

$$L \int_b^a f = -L \int_a^b f. \tag{9.1.2}$$

For Riemann integrals, we replace "L" by "R." We compare such integrals with antiderivatives (Chapter 5, §5), denoted

$$\int_a^b f, \tag{9.1.3}$$

without the "L" or "R." Note that

$$L \int_{[a,b]} f = L \int_{(a,b)} f, \tag{9.1.4}$$

etc., since $m\{a\} = m\{b\} = 0$ here.

Theorem 9.1.1

Let $f : E^1 \rightarrow E$ be L-integrable on $A = [a, b]$. Set

$$H(x) = L \int_a^x f, \quad x \in A. \tag{9.1.5}$$

Then the following are true.

(i) The function f is the derivative of H at any $p \in A$ at which f is finite and continuous. (At a and b, continuity and derivatives may be one-sided from within.)

(ii) The function H is absolutely continuous on A; hence $V_H[A] < \infty$.

Proof

(i) Let $p \in (a, b], q = f(p) \neq \pm\infty$. Let f be left continuous at p; so, given $\varepsilon > 0$, we can fix $c \in (a, p)$ such that

$$|f(x) - q| < \varepsilon \text{ for } x \in (c, p). \tag{9.1.6}$$

Then

$$\begin{aligned}(\forall x \in (c, p)) \left| L \int_x^p (f - q) \right| &\leq L \int_x^p |f - q| \\ &\leq L \int_x^p (\varepsilon) = \varepsilon \cdot m[x, p] = \varepsilon(p - x).\end{aligned}$$

But

$$\begin{aligned} L \int_x^p (f-q) &= L \int_x^p f - L \int_x^p q \\ L \int_x^p q &= q(p-x), \quad \text{and} \\ L \int_x^p f &= L \int_a^p f - L \int_a^x f \\ &= H(p) - H(x). \end{aligned}$$

Thus

$$|H(p)-H(x)-q(p-x)| \leq \varepsilon(p-x); \tag{9.1.7}$$

i.e.,

$$\left|\frac{H(p)-H(x)}{p-x}-q\right| \leq \varepsilon \quad (c<x<p). \tag{9.1.8}$$

Hence

$$f(p)=q=\lim_{x \rightarrow p^{-}} \frac{\Delta H}{\Delta x}=H_{-}^{\prime}(p). \tag{9.1.9}$$

If f is right continuous at $p \in [a, b)$, a similar formula results for $H_{+}^{\prime}(p)$. This proves clause (i).

(ii) Let $\varepsilon > 0$ be given. Then Theorem 6 in Chapter 8, §6, yields a $\delta > 0$ such that

$$\left|L \int_X f\right| \leq L \int_X |f| < \varepsilon \tag{9.1.10}$$

whenever

$$mX < \delta \text{ and } A \supseteq X, X \in \mathcal{M}. \tag{9.1.11}$$

Here we may set

$$X=\bigcup_{i=1}^r A_i \text{ (disjoint)} \tag{9.1.12}$$

for some intervals

$$A_i=(a_i, b_i) \subseteq A \tag{9.1.13}$$

so that

$$mX=\sum_i mA_i=\sum_i (b_i-a_i)<\delta. \tag{9.1.14}$$

Then (1) implies that

$$\varepsilon > L \int_X |f| = \sum_i L \int_{A_i} |f| \geq \sum_i \left| L \int_{a_i}^{b_i} f \right| = \sum_i \left|H(b_i)-H(a_i)\right|. \tag{9.1.15}$$

Thus

$$\sum_i \left|H(b_i)-H(a_i)\right| < \varepsilon \tag{9.1.16}$$

whenever

$$\sum_i (b_i-a_i) < \delta \tag{9.1.17}$$

and

$$A \supseteq \bigcup_i (a_i, b_i) \text{ (disjoint).} \tag{9.1.18}$$

(This is what we call "absolute continuity in the stronger sense.") By Problem 2 in Chapter 5, §8, this implies "absolute continuity" in the sense of Chapter 5, §8, hence $V_H[A] < \infty$. $\square$

Note 1. The converse to (i) fails: the differentiability of H at p does not imply the continuity of its derivative f at p (Problem 6 in Chapter 5, §2).

Note 2. If f is continuous on $A - Q$ (Q countable), Theorem 1 shows that H is a primitive (antiderivative): $H = \int f$ on A. Recall that "Q countable" implies $mQ = 0$, but not conversely. Observe that we may always assume $a, b \in Q$.

We can now prove a generalized version of the so-called fundamental theorem of calculus, widely used for computing integrals via antiderivatives.

Theorem 9.1.2

If $f : E^1 \rightarrow E$ has a primitive F on $A = [a, b]$, and if f is bounded on $A - P$ for some P with $mP = 0$, then f is L-integrable on A, and

$$L \int_a^x f = F(x) - F(a) \quad \text{for all } x \in A. \tag{9.1.19}$$

Proof

By Definition 1 of Chapter 5, §5, F is relatively continuous and finite on $A = [a, b]$, hence bounded on A (Theorem 2 in Chapter 4, §8).

It is also differentiable, with $F' = f$, on $A - Q$ for a countable set $Q \subseteq A$, with $a, b \in Q$. We fix this Q along with P.

As we deal with A only, we surely may redefine F and f on $-A$:

$$F(x) = \begin{cases} F(a) & \text{if } x < a, \\ F(b) & \text{if } x > b, \end{cases} \tag{9.1.20}$$

and $f = 0$ on $-A$. Then f is bounded on $-P$, while F is bounded and continuous on E^1, and $F' = f$ on $-Q$; so $F = \int f$ on E^1.

Also, for $n = 1, 2, \ldots$ and $t \in E^1$, set

$$f_n(t) = n\left[F\left(t + \frac{1}{n}\right) - F(t)\right] = \frac{F(t + 1/n) - F(t)}{1/n}. \tag{9.1.21}$$

Then

$$f_n \rightarrow F' = f \quad \text{on } -Q; \tag{9.1.22}$$

i.e., $f_n \rightarrow f$ (a.e.) on E^1 (as $mQ = 0$).

By (3), each f_n is bounded and continuous (as F is). Thus by Theorem 1 of Chapter 8, §3, F and all f_n are m-measurable on A (even on E^1). So is f by Corollary 1 of Chapter 8, §3.

Moreover, by boundedness, F and f_n are L-integrable on finite intervals. So is f. For example, let

$$|f| \leq K < \infty \text{ on } A - P; \tag{9.1.23}$$

as $mP = 0$,

$$\int_A |f| \leq \int_A (K) = K \cdot mA < \infty, \tag{9.1.24}$$

proving integrability. Now, as

$$F=\int f \text{ on any interval } \left[t, t+\frac{1}{n}\right], \tag{9.1.25}$$

Corollary 1 in Chapter 5, §4 yields

$$\left(\forall t \in E^1\right) \quad \left|F\left(t+\frac{1}{n}\right)-F(t)\right| \leq \sup_{t \in-Q}\left|F^{\prime}(t)\right| \frac{1}{n} \leq \frac{K}{n}. \tag{9.1.26}$$

Hence

$$\left|f_n(t)\right|=n\left|F\left(t+\frac{1}{n}\right)-F(t)\right| \leq K; \tag{9.1.27}$$

i.e., $\left|f_n\right| \leq K$ for all n.

Thus f and f_n satisfy Theorem 5 of Chapter 8, §6, with $g=K$. By Note 1 there,

$$\lim_{n \rightarrow \infty} L \int_a^x f_n=L \int_a^x f. \tag{9.1.28}$$

In the next lemma, we show that also

$$\lim_{n \rightarrow \infty} L \int_a^x f_n=F(x)-F(a), \tag{9.1.29}$$

which will complete the proof. □

Lemma 9.1.1

Given a finite continuous $F: E^1 \rightarrow E$ and given f_n as in (3), we have

$$\lim_{n \rightarrow \infty} L \int_a^x f_n=F(x)-F(a) \quad \text{for all } x \in E^1. \tag{9.1.30}$$

Proof

As before, F and f_n are bounded, continuous, and L-integrable on any $[a, x]$ or $[x, a]$. Fixing a, let

$$H(x)=L \int_a^x F, \quad x \in E^1. \tag{9.1.31}$$

By Theorem 1 and Note 2, $H=\int F$ also in the sense of Chapter 5, §5, with $F=H^{\prime}$ (derivative of H) on E^1.

Hence by Definition 2 the same section,

$$\int_a^x F=H(x)-H(a)=H(x)-0=L \int_a^x F; \tag{9.1.32}$$

i.e.,

$$L \int_a^x F=\int_a^x F, \tag{9.1.33}$$

and so

$$\begin{aligned} L \int_a^x f_n(t) d t & =n \int_a^x F\left(t+\frac{1}{n}\right) d t-n \int_a^x F(t) d t \\ & =n \int_{a+1 / n}^{b+1 / n} F(t) d t-n \int_a^x F(t) d t. \end{aligned}$$

(We computed

$$\int F(t+1 / n) d t \tag{9.1.34}$$

by Theorem 2 in Chapter 5, §5, with $g(t)=t+1/n$.) Thus by additivity,

$$L \int_{a}^{x} f_{n}=n \int_{a+1/n}^{x+1/n} F-n \int_{a}^{x} F=n \int_{x}^{x+1/n} F-n \int_{a}^{a+1/n} F. \tag{9.1.35}$$

But

$$n \int_{x}^{x+1/n} F=\frac{H\left(x+\frac{1}{n}\right)-H(x)}{\frac{1}{n}} \rightarrow H^{\prime}(x)=F(x). \tag{9.1.36}$$

Similarly,

$$\lim_{n \rightarrow \infty} n \int_{a}^{a+1/n} F=F(a). \tag{9.1.37}$$

This combined with (5) proves (4), and hence Theorem 2, too. □

We also have the following corollary.

Corollary 9.1.1

If $f: E^{1} \rightarrow E^{*}\left(E^{n}, C^{n}\right)$ is R-integrable on $A=[a, b]$, then

$$(\forall x \in A) \quad R \int_{a}^{x} f=L \int_{a}^{b} f=F(x)-F(a), \tag{9.1.38}$$

provided F is primitive to f on A.

Proof

This follows from Theorem 2 by Definition (c) and Theorem 2 of Chapter 8, §9.

Caution. Formulas (2) and (6) may fail if f is unbounded, or if F is not a primitive in the sense of Definition 1 of Chapter 5, §5: We need $F^{\prime}=f$ on $A-Q, Q$ countable ($mQ=0$ is not enough!). Even R-integrability (which makes f bounded and a.e. continuous) does not suffice if

$$F \neq \int f. \tag{9.1.39}$$

For examples, see Problems 2-5.

Corollary 9.1.2

If f is relatively continuous and finite on $A=[a, b]$ and has a bounded derivative on $A-Q$ (Q countable), then $f^{\prime}$ is L-integrable on A and

$$L \int_{a}^{x} f^{\prime}=f(x)-f(a) \quad \text{for } x \in A. \tag{9.1.40}$$

This is simply Theorem 2 with F, f, P replaced by $f, f^{\prime}, Q$, respectively

Corollary 9.1.3

If in Theorem 2 the primitive

$$F=\int f \tag{9.1.41}$$

is exact on some $B \subseteq A$, then

$$f(x)=\frac{d}{dx}L\int_a^x f, \quad x\in B. \tag{9.1.42}$$

(Recall that $\frac{d}{dx}F(x)$ is classical notation for $F'(x)$.)

Proof

By (2), this holds on $B\subseteq A$ if $F'=f$ there. $\square$

II. Note that under the assumptions of Theorem 2,

$$L\int_a^x f=F(x)-F(a)=\int_a^x f. \tag{9.1.43}$$

Thus all laws governing the primitive $\int f$ apply to $L\int f$. For example, Theorem 2 of Chapter 5, §5, yields the following corollary.

Corollary 9.1.4 (change of variable)

Let $g:E^1\to E^1$ be relatively continuous on $A=[a,b]$ and have a bounded derivative on $A-Q$ (Q countable).

Suppose that $f:E^1\to E$ (real or not) has a primitive on $g[A]$, exact on $g[A-Q]$, and that f is bounded on $g[A-Q]$.

Then f is L-integrable on $g[A]$, the function

$$(f\circ g)g' \tag{9.1.44}$$

is L-integrable on A, and

$$L\int_a^b f(g(x))g'(x)dx=L\int_p^q f(y)dy, \tag{9.1.45}$$

where $p=g(a)$ and $q=g(b)$.

For this and other applications of primitives, see Problem 9. However, often a direct approach is stronger (though not simpler), as we illustrate next.

Lemma 9.1.2 (Bonnet)

Suppose $f:E^1\to E^1$ is ≥ 0 and monotonically decreasing on $A=[a,b]$. Then, if $g:E^1\to E^1$ is L-integrable on A, so also is fg, and

$$L\int_a^b fg=f(a)\cdot L\int_a^c g \quad \text{for some } c\in A. \tag{9.1.46}$$

Proof

The L-integrability of fg follows by Theorem 3 in Chapter 8, §6, as f is monotone and bounded, hence even R-integrable (Corollary 3 in Chapter 8, §9).

Using this and Lemma 1 of the same section, fix for each n a $\mathcal{C}$-partition

$$\mathcal{P}_n=\{A_{ni}\} \quad (i=1,2,\ldots,q_n) \tag{9.1.47}$$

of A so that

$$(\forall n)\quad \frac{1}{n}>\overline{S}(f,\mathcal{P}_n)-\underline{S}(f,\mathcal{P}_n)=\sum_{i=1}^{q_n} w_{ni}mA_{ni}, \tag{9.1.48}$$

where we have set

$$w_{ni}=\sup f[A_{ni}]-\inf f[A_{ni}]. \tag{9.1.49}$$

Consider any such $\mathcal{P}=\{A_i\}$, $i=1,\ldots,q$ (we drop the "n" for brevity). If $A_i=[a_{i-1},a_i]$, then since $f\downarrow$,

$$w_i=f\left(a_{i-1}\right)-f\left(a_i\right) \geq\left|f(x)-f\left(a_{i-1}\right)\right|, \quad x \in A_i . \tag{9.1.50}$$

Under Lebesgue measure (Problem 8 of Chapter 8, §9), we may set

$$A_i=\left[a_{i-1}, a_i\right] \quad(\forall i) \tag{9.1.51}$$

and still get

$$\begin{aligned} L \int_A f g= & \sum_{i=1}^q f\left(a_{i-1}\right) L \int_{A_i} g(x) d x \\ & +\sum_{i=1}^q L \int_{A_i}\left[f(x)-f\left(a_{i-1}\right)\right] g(x) d x . \end{aligned}$$

(Verify!) Here $a_0=a$ and $a_q=b$.

Now, set

$$G(x)=L \int_a^x g \tag{9.1.52}$$

and rewrite the first sum (call it r or r_n) as

$$\begin{aligned} r & =\sum_{i=1}^q f\left(a_{i-1}\right)\left[G\left(a_i\right)-G\left(a_{i-1}\right)\right] \\ & =\sum_{i=1}^{q-1} G\left(a_i\right)\left[f\left(a_{i-1}\right)-f\left(a_i\right)\right]+G(b) f\left(a_{q-1}\right), \end{aligned}$$

or

$$r=\sum_{i=1}^{q-1} G\left(a_i\right) w_i+G(b) f\left(a_{q-1}\right), \tag{9.1.53}$$

because $f\left(a_{i-1}\right)-f\left(a_i\right)=w_i$ and $G(a)=0$.

Now, by Theorem 1 (with H, f replaced by G, g), G is continuous on $A=[a, b]$; so G attains a largest value K and a least value k on A.

As $f \downarrow$ and $f \geq 0$ on A, we have

$$w_i \geq 0 \text { and } f\left(a_{q-1}\right) \geq 0 . \tag{9.1.54}$$

Thus, replacing $G(b)$ and $G\left(a_i\right)$ by $K(\text { or } k)$ in (13) and noting that

$$\sum_{i=1}^{q-1} w_i=f(a)-f\left(a_{q-1}\right), \tag{9.1.55}$$

we obtain

$$k f(a) \leq r \leq K f(a) ; \tag{9.1.56}$$

more fully, with $k=\min G[A]$ and $K=\max G[A]$,

$$(\forall n) \quad k f(a) \leq r_n \leq K f(a) . \tag{9.1.57}$$

Next, let s (or rather s_n be the second sum in (12). Noting that

$$w_i \geq\left|f(x)-f\left(a_{i-1}\right)\right|, \tag{9.1.58}$$

suppose first that $|g| \leq B$ (bounded) on A.

Then for all n,

$$|s_n| \leq \sum_{i=1}^{q_n} L\int_{A_{ni}} (w_{ni}B) = B\sum_{i=1}^{q_n} w_{ni} m A_{ni} < \frac{B}{n} \to 0 \quad \text{(by (11))}. \tag{9.1.59}$$

But by (12),

$$L\int_A fg = r_n + s_n \quad (\forall n). \tag{9.1.60}$$

As $s_n \to 0$,

$$L\int_A fg = \lim_{n\to\infty} r_n, \tag{9.1.61}$$

and so by (14),

$$kf(a) \leq L\int_A fg \leq Kf(a). \tag{9.1.62}$$

By continuity, $f(a)G(x)$ takes on the intermediate value $L\int_A fg$ at some $c \in A$; so

$$L\int_A fg = f(a)G(c) = f(a)L\int_a^c g, \tag{9.1.63}$$

since

$$G(x) = L\int_a^x f. \tag{9.1.64}$$

Thus all is proved for a bounded g.

The passage to an unbounded g is achieved by the so-called truncation method described in Problems 12 and 13. (Verify!) □

Corollary9.1.5 (second law of the mean)

Let $f: E^1 \to E^1$ be monotone on $A = [a, b]$. Then if $g: E^1 \to E^1$ is L-integrable on A, so also is fg, and

$$L\int_a^b fg = f(a)L\int_a^c g + f(b)L\int_c^b g \quad \text{for some } c \in A. \tag{9.1.65}$$

Proof

If, say, $f\downarrow$ on A, set

$$h(x) = f(x) - f(b). \tag{9.1.66}$$

Then $h \geq 0$ and $h\downarrow$ on A; so by Lemma 2,

$$\int_a^b gh = h(a)L\int_a^c g \quad \text{for some } c \in A. \tag{9.1.67}$$

As

$$h(a) = f(a) - f(b), \tag{9.1.68}$$

this easily implies (15).

If $f\uparrow$, apply this result to $-f$ to obtain (15) again. □

Note 3. We may restate (15) as

$$(\exists c \in A) \quad L\int_a^b fg = pL\int_a^c g + qL\int_c^b g, \tag{9.1.69}$$

provided either

(i) $f \uparrow$ and $p \leq f(a+) \leq f(b-) \leq q$, or

(ii) $f \downarrow$ and $p \geq f(a+) \geq f(b-) \geq q$.

This statement slightly strengthens (15).

To prove clause (i), redefine

$$f(a)=p \text{ and } f(b)=q. \tag{9.1.70}$$

Then still $f \uparrow$; so (15) applies and yields the desired result. Similarly for (ii). For a continuous g, see also Problem 13(ii') in Chapter 8, §9, based on Stieltjes theory.

III. We now give a useful analogue to the notion of a primitive.

Definition

A map $F: E^{1} \rightarrow E$ is called an L-primitive or an indefinite L-integral of $f: E^{1} \rightarrow E,$ on $A=[a, b]$ iff f is L-integrable on A and

$$F(x)=c+L \int_{a}^{x} f \tag{9.1.71}$$

for all $x \in A$ and some fixed finite $c \in E$.

Notation:

$$F=L \int f \quad\left(\text {not } F=\int f\right) \tag{9.1.72}$$

or

$$F(x)=L \int f(x) d x \quad \text { on } A. \tag{9.1.73}$$

By (16), all L-primitives of f on A differ by finite constants only.

If $E=E^{*}\left(E^{n}, C^{n}\right)$, one can use this concept to lift the boundedness restriction on f in Theorem 2 and the corollaries of this section. The proof will be given in §2. However, for comparison, we state the main theorems already now.

Theorem 9.1.3

Let

$$F=L \int f \quad \text { on } A=[a, b] \tag{9.1.74}$$

for some $f: E^{1} \rightarrow E^{*}\left(E^{n}, C^{n}\right)$.

Then F is differentiable, with

$$F^{\prime}=f \quad \text { a.e. on } A. \tag{9.1.75}$$

In classical notation,

$$f(x)=\frac{d}{d x} L \int_{a}^{x} f(t) d t \quad \text { for almost all } x \in A. \tag{9.1.76}$$

A proof was sketched in Problem 6 of Chapter 8, §12. (It is brief but requires more "starred" material than used in §2.)

Theorem 9.1.4

Let $F: E^{1} \rightarrow E^{n}\left(C^{n}\right)$ be differentiable on $A=[a, b]$ (at a and b differentiability may be one sided). Let $F^{\prime}=f$ be L-integrable on A.

Then

$$L \int_{a}^{x} f=F(x)-F(a) \quad \text { for all } x \in A. \tag{9.1.77}$$

9.1.E: Problems on L-Integrals and Antiderivatives

Exercise $9.1.E.1$

Fill in proof details in Theorems 1 and 2, Lemma 1, and Corollaries 1-3.

Exercise $9.1.E.1'$

Verify Note 2.

Exercise $9.1.E.2$

Let F be Cantor's function (Problem 6 in Chapter 4, §5). Let

$$G=\bigcup_{k,i} G_{kt} \tag{9.1.E.1}$$

$(G_{kt} \text{ as in that problem }).$ So $[0,1]-G=P(\text{ Cantor's set }); mP=0$ (Problem 10 in Chapter 7, §8).
Show that F is differentiable $(F'=0)$ on G. By Theorems 2 and 3 of Chapter 8, §9,

$$R\int_0^1 F'=L\int_0^1 F'=L\int_G F'=0 \tag{9.1.E.2}$$

exists, yet $F(1)-F(0)=1-0\neq 0$.
Does this contradict Corollary 1? Is F a genuine antiderivative of f? If not, find one.

Exercise $9.1.E.3$

Let

$$F=\begin{cases}0 & \text{on } \left[0,\frac{1}{2}\right), \text{ and} \\ 1 & \text{on } \left[\frac{1}{2},1\right].\end{cases} \tag{9.1.E.3}$$

Show that

$$R\int_0^1 F'=0 \tag{9.1.E.4}$$

exists, yet

$$F(1)-F(0)=1-0=1. \tag{9.1.E.5}$$

What is wrong?
[Hint: A genuine primitive of F' (call it ϕ) has to be relatively continuous on $[0,1]$; find ϕ and show that $\phi(1)-\phi(0)=0.$]

Exercise $9.1.E.4$

What is wrong with the following computations?
(i) $L\int_{-1}^{\frac{1}{2}}\frac{dx}{x^2}=-\frac{1}{x}\Big|_{-1}^{\frac{1}{2}}=-1$.
(ii) $L\int_{-1}^{1}\frac{dx}{x}=\ln|x|\Big|_{-1}^{1}=0.$ Is there a primitive on the whole interval?

[Hint: See hint to Problem 3.]
(iii) How about $L \int_{-1}^{1} \frac{|x|}{x} dx$ (cf. examples (a) and (b) of Chapter 5, §5)?

Exercise $9.1.E.5$

Let

$$F(x)=x^{2} \cos \frac{\pi}{x^{2}}, \quad F(0)=1. \tag{9.1.E.6}$$

Prove the following:
(i) F is differentiable on $A=[0,1]$.
(ii) $f=F'$ is bounded on any $[a, b] \subset(0,1)$, but not on A.
(iii) Let

$$a_{n}=\sqrt{\frac{2}{4 n+1}} \text { and } b_{n}=\frac{1}{\sqrt{2 n}} \text { for } n=1,2, \ldots \tag{9.1.E.7}$$

Show that

$$A \supseteq \bigcup_{n=1}^{\infty}\left[a_{n}, b_{n}\right] \text { (disjoint) } \tag{9.1.E.8}$$

and

$$L \int_{a_{n}}^{b_{n}} f=\frac{1}{2 n}; \tag{9.1.E.9}$$

so

$$L \int_{a}^{b} f \geq L \int_{\cup_{n=1}^{\infty}\left[a_{n}, b_{n}\right]} f \geq \sum_{n=1}^{\infty} \frac{1}{2 n}=\infty, \tag{9.1.E.10}$$

and $f=F'$ is not L-integrable on A.
What is wrong? Is there a contradiction to Theorem 2?

Exercise $9.1.E.6$

Consider both
(a) $f(x)=\frac{\sin x}{x}, f(0)=1$, and
(b) $f(x)=\frac{1-e^{-x}}{x}, f(0)=1$.
In each case, show that f is continuous on $A=[0,1]$ and

$$R \int_{A} f \leq 1 \tag{9.1.E.11}$$

exists, yet it does not "work out" via primitives. What is wrong? Does a primitive exist?
To use Corollary 1, first expand $\sin x$ and e^{-x} in a Taylor series and find the series for

$$\int f \tag{9.1.E.12}$$

by Theorem 3 of Chapter 5, §9.
Find

$$R \int_A f \tag{9.1.E.13}$$

approximately, to within $1/10$, using the remainder term of the series to estimate accuracy.
[Hint: Primitives exist, by Theorem 2 of Chapter 5, §11, even though they are none of the known "calculus functions."]

Exercise $9.1.E.7$

Take $A, G_n = (a_n, b_n)$, and $P(mP > 0)$ as in Problem 17(iii) of Chapter 7, §8.
Define $F = 0$ on P and

$$F(x) = (x - a_n)^2 (x - b_n)^2 \sin \frac{1}{(b_n - a_n)(x - a_n)(x - b_n)} \quad \text{if } x \notin P. \tag{9.1.E.14}$$

Prove that F has a bounded derivative f, yet f is not R-integrable on A; so Theorem 2 applies, but Corollary 1 does not.
[Hints: If $p \notin P$, compute $F'(p)$ as in calculus.
If $p \in P$ and $x \rightarrow p+$ over $A - P$, then x is always in some $(a_n, b_n), p \leq a_n < x$. (Why?) Deduce that $\Delta x = x - p > x - a_n$ and

$$\left|\frac{\Delta F}{\Delta x}\right| \leq (x - a_n)(b - a)^2 \leq |\Delta x|(b - a)^2; \tag{9.1.E.15}$$

so $F'_+(p) = 0$. (What if $x \rightarrow p+$ over P ?) Similarly, show that $F'_- = 0$ on P.
Prove however that $F'(x)$ oscillates from 1 to -1 as $x \rightarrow a_n+$ or $x \rightarrow b_n-$, hence also as $x \rightarrow p \in P$ (why?); so F' is discontinuous on all of P, with $mP > 0$. Now use Theorem 3 in Chapter 8, §9.]

Exercise $9.1.E.8$

$\Rightarrow 8$. If

$$Q \subseteq A = [a, b] \tag{9.1.E.16}$$

and $mQ = 0$, find a continuous map $g : A \rightarrow E^1, g \geq 0, g \uparrow$, with

$$g' = +\infty \quad \text{on } Q. \tag{9.1.E.17}$$

[Hints: By Theorem 2 of Chapter 7, §8, fix ($\forall n$) an open $G_n \supseteq Q$, with

$$mG_n < 2^{-n}. \tag{9.1.E.18}$$

Set

$$g_n(x) = m(G_n \cap [a, x]) \tag{9.1.E.19}$$

and

$$g = \sum_{n=1}^{\infty} g_n \tag{9.1.E.20}$$

on A; $\sum g_n$ converges uniformly on A. (Why?)
By Problem 4 in Chapter 7, §9, and Theorem 2 of Chapter 7, §4, each g_n (hence g) is continuous. (Why?) If $[p, x] \subseteq G_n$, show that

$$g_n(x) = g_n(p) + (x - p), \tag{9.1.E.21}$$

so

$$\frac{\Delta g_n}{\Delta x} = 1 \tag{9.1.E.22}$$

and

$$\left. \frac{\Delta g}{\Delta x} = \sum_{n=1}^{\infty} \frac{\Delta g_n}{\Delta x} \rightarrow \infty. \right] \tag{9.1.E.23}$$

Exercise $9.1.E.9$

(i) Prove Corollary 4.
(ii) State and prove earlier analogues for Corollary 5 of Chapter 5, §5, and Theorems 3 and 4 from Chapter 5, §10.
[Hint for (i): For primitives, this is Problem 3 in Chapter 5, §5. As $g[Q]$ is countable (Problem 2 in Chapter 1, §9) and f is bounded on

$$g[A] - g[Q] \subseteq g[A - Q], \tag{9.1.E.24}$$

f satisfies Theorem 2 on $g[A]$, with $P = g[Q]$, while $(f \circ g)g'$ satisfies it on A.]

Exercise $9.1.E.10$

$\Rightarrow$ 10. Show that if $h : E^1 \rightarrow E^*$ is L-integrable on $A = [a, b]$, and

$$(\forall x \in A) \quad L\int_a^x h = 0, \tag{9.1.E.25}$$

then $h = 0$ a.e. on A.
[Hints: Let $K = A(h > 0)$ and $H = A - K$, with, say, $mK = \varepsilon > 0$.
Then by Corollary 1 in Chapter 7, §1 and Definition 2 of Chapter 7, §5,

$$H \subseteq \bigcup_n B_n (disjoint) \tag{9.1.E.26}$$

for some intervals $B_n \subseteq A$, with

$$\sum_n mB_n < mH + \varepsilon = mH + mK = mA. \tag{9.1.E.27}$$

(Why?) Set $B = \cup_n B_n$; so

$$\int_B h = \sum_n \int_{B_n} h = 0 \tag{9.1.E.28}$$

(for $L \int h = 0$ on intervals B_n). Thus

$$\int_{A-B} h=\int_A h-\int_B h=0. \tag{9.1.E.29}$$

But $B \supseteq H;$ so

$$A-B \subseteq A-H=K, \tag{9.1.E.30}$$

where $h>0,$ even though $m(A-B)>0.$ (Why?)
Hence find a contradiction to Theorem $1(\mathrm{h})$ of Chapter 8, §5. Similarly, disprove that $mA(h<0)=\varepsilon>0.]$

? Exercise $9.1.E.11$

$\Rightarrow 11.$ Let $F \uparrow$ on $A=[a, b],|F|<\infty,$ with derived function $F^{\prime}=f.$ Taking Theorem 3 from Chapter 7, §10, for granted, prove that

$$L \int_a^x f \leq F(x)-F(a), \quad x \in A, \tag{9.1.E.31}$$

[Hints: With f_n as in $(3), F$ and f_n are bounded on A and measurable by Theorem 1 of Chapter 8, §2. (Why?) Deduce that $f_n \rightarrow f$ (a.e.) on $A.$ Argue as in Lemma 1 using Fatou's lemma (Chapter 8, §6, Lemma 2).]

? Exercise $9.1.E.12$

("Truncation.") Prove that if $g: S \rightarrow E$ is m-integrable on $A \in \mathcal{M}$ in a measure space $(S, \mathcal{M}, m),$ then for any $\varepsilon>0,$ there is a bounded, M-measurable and integrable on A map $g_0: S \rightarrow E$ such that

$$\int_A\left|g-g_0\right| d m<\varepsilon. \tag{9.1.E.32}$$

[Outline: Redefine $g=0$ on a null set, to make $g \mathcal{M}$-measurable on $A.$ Then for $n=1,2, \ldots$ set

$$g_n=\left\{\begin{array}{ll}g & \text { on } A(|g|<n), \text { and } \\ 0 & \text { elsewhere. }\end{array}\right. \tag{9.1.E.33}$$

(The function g_n is called the nth truncate of g.)
Each g_n is bounded and $\mathcal{M}$ -measurable on A (why?), and

$$\int_A|g| d m<\infty \tag{9.1.E.34}$$

by integrability. Also, $\left|g_n\right| \leq|g|$ and $g_n \rightarrow g(\text{pointwise})$ on $A.$ (Why?)
Now use Theorem 5 from Chapter 8, §6, to show that one of the g_n may serve as the desired $g_o.]$

? Exercise $9.1.E.13$

Fill in all proof details in Lemma $2.$ Prove it for unbounded $g.$
[Hints: By Problem $12,$ fix a bounded $g_o\left(\left|g_o\right| \leq B\right),$ with

$$L \int_A\left|g-g_o\right|<\frac{1}{2} \frac{\varepsilon}{f(a)-f(b)}. \tag{9.1.E.35}$$

Verify that

$$|s_n| \leq \sum_{i=1}^{q_n} \int_{A_{ni}} w_{ni}|g| \leq \sum_i \int_{A_{ni}} w_{ni} |g_o| + \sum_i \int_{A_{ni}} w_{ni} |g - g_o|$$

$$\leq B \sum_i w_{ni} m A_{ni} + \sum_i \int_{A_{ni}} [f(a) - f(b)] |g - g_o|$$

$$< \frac{1}{n} + \int_A [f(a) - f(b)] |g - g_o| < \frac{1}{n} + \frac{1}{2}\varepsilon.$$

For all $n > 2/\varepsilon$, we get $|s_n| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$. Hence $s_n \rightarrow 0$. Now finish as in the text.]

Exercise $9.1.E.14$

Show that Theorem 4 fails if F is not differentiable at some $p \in A$.
[Hint: See Problems 2 and 3.]

9.2: More on L-Integrals and Absolute Continuity

I. In this section, we presuppose the "starred" §10 in Chapter 7. First, however, we add some new ideas that do not require any starred material. The notation is as in §1.

Definition

Given $F : E^1 \rightarrow E, p \in E^1$, and $q \in E$, we write

$$q \sim DF(p) \tag{9.2.1}$$

and call q an F-derivate at p iff

$$q=\lim_{k \rightarrow \infty} \frac{F\left(x_k\right)-F(p)}{x_k-p} \tag{9.2.2}$$

for at least one sequence $x_k \rightarrow p\left(x_k \neq p\right)$.

If F has a derivative at p, it is the only F-derivate at p; otherwise, there may be many derivates at p (finite or not).

Such derivates must exist if $E=E^1\left(E^*\right)$. Indeed, given any $p \in E^1$, let

$$x_k=p+\frac{1}{k} \rightarrow p; \tag{9.2.3}$$

let

$$y_k=\frac{F\left(x_k\right)-F(p)}{x_k-p}, \quad k=1,2, \ldots \tag{9.2.4}$$

By the compactness of E^* (Chapter 4, §6, example (d)), $\left\{y_k\right\}$ must have a subsequence $\left\{y_{k_i}\right\}$ with a limit $q \in E^*$ (e.g., take $q=\underline{\lim}\, y_k$), and so $q \sim DF(p)$.

We also obtain the following lemma.

Lemma 9.2.1

If $F : E^1 \rightarrow E^*$ has no negative derivates on $A-Q$, where $A=[a, b]$ and $mQ=0$, and if no derivate of F on A equals $-\infty$, then $F \uparrow$ on A.

Proof

First, suppose F has no negative derivates on A at all. Fix $\varepsilon>0$ and set

$$G(x)=F(x)+\varepsilon x. \tag{9.2.5}$$

Seeking a contradiction, suppose $a \leq p<q \leq b$, yet $G(q)<G(p)$. Then if

$$r=\frac{1}{2}(p+q), \tag{9.2.6}$$

one of the intervals $[p, r]$ and $[r, q]$ (call it $\left[p_1, q_1\right]$) satisfies $G\left(q_1\right)<G\left(p_1\right)$.

Let

$$r_1=\frac{1}{2}\left(p_1+q_1\right). \tag{9.2.7}$$

Again, one of $\left[p_1, r_1\right]$ and $\left[r_1, q_1\right]$ (call it $\left[p_2, q_2\right]$) satisfies $G\left(q_2\right)<G\left(p_2\right)$. Let

$$r_2=\frac{1}{2}\left(p_2+q_2\right), \tag{9.2.8}$$

and so on.

Thus obtain contracting intervals $\left[p_n, q_n\right]$, with

$$G\left(q_{n}\right)<G\left(p_{n}\right), \quad n=1,2, \ldots . \tag{9.2.9}$$

Now, by Theorem 5 of Chapter 4, §6, let

$$p_{o} \in \bigcap_{n=1}^{\infty}\left[p_{n}, q_{n}\right] . \tag{9.2.10}$$

Then set $x_{n}=q_{n}$ if $G\left(q_{n}\right)<G\left(p_{o}\right)$, and $x_{n}=p_{n}$ otherwise. Then

$$\frac{G\left(x_{n}\right)-G\left(p_{o}\right)}{x_{n}-p_{o}}<0 \tag{9.2.11}$$

and $x_{n} \rightarrow p_{o}$. By the compactness of E^{*}, fix a subsequence

$$\frac{G\left(x_{n_{k}}\right)-G\left(p_{o}\right)}{x_{n_{k}}-p_{o}} \rightarrow c \in E^{*}, \tag{9.2.12}$$

say. Then $c \leq 0$ is a G-derivate at $p_{o} \in A$.

But this is impossible; for by our choice of G and our assumption, all derivates of G are >0. (Why?)

This contradiction shows that $a \leq p<q \leq b$ implies $G(p) \leq G(q)$, i.e.,

$$F(p)+\varepsilon p \leq F(q)+\varepsilon q. \tag{9.2.13}$$

Making $\varepsilon \rightarrow 0$, we obtain $F(p) \leq F(q)$ when $a \leq p<q \leq b$, i.e., $F \uparrow$ on A.

Now, for the general case, let Q be the set of all $p \in A$ that have at least one $DF(p)<0$; so $mQ=0$.

Let g be as in Problem 8 of §1; so $g^{\prime}=\infty$ on Q. Given $\varepsilon>0$, set

$$G=F+\varepsilon g. \tag{9.2.14}$$

As $g \uparrow$, we have

$$(\forall x, p \in A) \quad \frac{G(x)-G(p)}{x-p} \geq \frac{F(x)-F(p)}{x-p}. \tag{9.2.15}$$

Hence $DG(p) \geq 0$ if $p \notin Q$.

If, however, $p \in Q$, then $g^{\prime}(p)=\infty$ implies $DG(p) \geq 0$. (Why?) Thus all $DG(p)$ are ≥ 0; so by what was proved above, $G \uparrow$ on A. It follows, as before, that $F \uparrow$ on A, also. The lemma is proved. □

We now proceed to prove Theorems 3 and 4 of §1. To do this, we shall need only one "starred" theorem (Theorem 3 of Chapter 7, §10).

Proof of Theorem 3 of §1. (1) First, let f be bounded:

$$|f| \leq K \quad \text{on } A. \tag{9.2.16}$$

Via components and by Corollary 1 of Chapter 8, §6, all reduces to the real positive case $f \geq 0$ on A. (Explain!)

Then (Theorem 1(f) of Chapter 8, §5) $a \leq x<y \leq b$ implies

$$L \int_{a}^{x} f \leq L \int_{a}^{y} f, \tag{9.2.17}$$

i.e., $F(x) \leq F(y)$; so $F \uparrow$ and $F^{\prime} \geq 0$ on A.

Now, by Theorem 3 of Chapter 7, §10, F is a.e. differentiable on A. Thus exactly as in Theorem 2 in §1, we set

$$f_{n}(t)=\frac{F\left(t+\frac{1}{n}\right)-F(t)}{\frac{1}{n}} \rightarrow F^{\prime}(t) \text{ a.e.} \tag{9.2.18}$$

Since all f_{n} are m-measurable on A (why?), so is $F^{\prime}$. Moreover, as $|f| \leq K$, we obtain (as in Lemma 1 of §1)

$$|f_n(x)| = n\left(L\int_x^{x+1/n} f\right) \leq n\cdot\frac{K}{n} = K. \tag{9.2.19}$$

Thus by Theorem 5 from Chapter 8, §6 (with $g = K$),

$$L\int_a^x F' = \lim_{n\to\infty} L\int_a^x f_n = L\int_a^x f \tag{9.2.20}$$

(Lemma 1 of §1). Hence

$$L\int_a^x (F'-f) = 0, \quad x\in A, \tag{9.2.21}$$

and so (Problem 10 in §1) $F' = f$ (a.e.) as claimed.

(2) If f is not bounded, we still can reduce all to the case $f\geq 0, f: E^1 \to E^*$ so that $F\uparrow$ and $F'\geq 0$ on A.

If so, we use "truncation": For $n = 1, 2, \ldots,$ set

$$g_n = \begin{cases} f & \text{on } A(f\leq n), \text{ and} \\ 0 & \text{elsewhere.} \end{cases} \tag{9.2.22}$$

Then (see Problem 12 in §1) the g_n are L-measurable and bounded, hence L-integrable on A, with $g_n \to f$ and

$$0\leq g_n \leq f \tag{9.2.23}$$

on A. By the first part of the proof, then,

$$\frac{d}{dx} L\int_a^x g_n = g_n \quad \text{a.e. on } A, n = 1, 2, \ldots. \tag{9.2.24}$$

Also, set $(\forall n)$

$$F_n(x) = L\int_a^x (f-g_n) \geq 0; \tag{9.2.25}$$

so F_n is monotone $(\uparrow)$ on A. (Why?)

Thus by Theorem 3 in Chapter 7, §10, each F_n has a derivative at almost every $x\in A$,

$$F_n'(x) = \frac{d}{dx}\left(L\int_a^x f - L\int_a^x g_n\right) = F'(x) - g_n(x) \geq 0 \quad \text{a.e. on } A. \tag{9.2.26}$$

Making $n\to\infty$ and recalling that $g_n\to f$ on A, we obtain

$$F'(x) - f(x) \geq 0 \quad \text{a.e. on } A. \tag{9.2.27}$$

Thus

$$L\int_a^x (F'-f) \geq 0. \tag{9.2.28}$$

But as $F\uparrow$ (see above), Problem 11 of §1 yields

$$L\int_a^x F' \leq F(x) - F(a) = L\int_a^x f; \tag{9.2.29}$$

so

$$L\int_a^x (F'-f) = L\int_a^x F' - L\int_a^x f \leq 0. \tag{9.2.30}$$

Combining, we get

$$(\forall x\in A) \quad L\int_a^x (F'-f) = 0; \tag{9.2.31}$$

so by Problem 10 of §1, $F' = f$ a.e. on A, as required. $\square$

Proof of Theorem 4 of §1. Via components, all again reduces to a real f.

Let $(\forall n)$

$$g_n = \begin{cases} f & \text{on } A(f \leq n), \\ 0 & \text{on } A(f > n); \end{cases} \tag{9.2.32}$$

so $g_n \rightarrow f$ (pointwise), $g_n \leq f, g_n \leq n,$ and $|g_n| \leq |f|$.

This makes each g_n L-integrable on A. Thus as before, by Theorem 5 of Chapter 8, §6,

$$\lim_{n \rightarrow \infty} L \int_a^x g_n = L \int_a^x f, \quad x \in A. \tag{9.2.33}$$

Now, set

$$F_n(x) = F(x) - L \int_a^x g_n. \tag{9.2.34}$$

Then by Theorem 3 of §1 (already proved),

$$F_n'(x) = F'(x) - \frac{d}{dx} L \int_a^x g_n = f(x) - g_n(x) \geq 0 \quad \text{a.e. on } A \tag{9.2.35}$$

(since $g_n \leq f$).

Thus F_n has solely nonnegative derivates on $A - Q(mQ = 0)$. Also, as $g_n \leq n,$ we get

$$\frac{1}{x-p} L \int_a^x g_n \leq n, \tag{9.2.36}$$

even if $x < p$. (Why?) Hence

$$\frac{\Delta F_n}{\Delta x} \geq \frac{\Delta F}{\Delta x} - n, \tag{9.2.37}$$

as

$$F_n(x) = F(x) - L \int_a^x g_n. \tag{9.2.38}$$

Thus none of the F_n-derivates on A can be $-\infty$.

By Lemma 1, then, F_n is monotone $(\uparrow)$ on A; so $F_n(x) \geq F_n(a),$ i.e.,

$$F(x) - L \int_a^x g_n \geq F(a) - L \int_a^a g_n = F(a), \tag{9.2.39}$$

or

$$F(x) - F(a) \geq L \int_a^x g_n, \quad x \in A, n = 1, 2, \ldots. \tag{9.2.40}$$

Hence by (1),

$$F(x) - F(a) \geq L \int_a^x f, \quad x \in A. \tag{9.2.41}$$

For the reverse inequality, apply the same formula to $-f$. Thus we obtain the desired result:

$$F(x) = F(a) + L \int_a^x f \quad \text{for } x \in A. \quad \square \tag{9.2.42}$$

Note 1. Formula (2) is equivalent to $F = L \int f$ on A (see the last part of §1). For if (2) holds, then

$$F(x)=c+L\int_{a}^{x} f, \tag{9.2.43}$$

with $c=F(a)$; so $F=L\int f$ by definition.

Conversely, if

$$F(x)=c+L\int_{a}^{x} f, \tag{9.2.44}$$

set $x=a$ to find $c=F(a)$.

II. A bsolute continuity redefined.

Definition

A map $f: E^{1} \rightarrow E$ is absolutely continuous on an interval $I \subseteq E^{1}$ iff for every $\varepsilon>0$, there is $\delta>0$ such that

$$\sum_{i=1}^{r}\left(b_{i}-a_{i}\right)<\delta \text { implies } \sum_{i=1}^{r}\left|f\left(b_{i}\right)-f\left(a_{i}\right)\right|<\varepsilon \tag{9.2.45}$$

for any disjoint intervals $\left(a_{i}, b_{i}\right)$, with $a_{i}, b_{i} \in I$.

From now on, this replaces the "weaker" definition given in Chapter 5, §8. The reader will easily verify the next three "routine" propositions.

Theorem 9.2.1

If $f, g, h: E^{1} \rightarrow E^{*}(C)$ are absolutely continuous on $A=[a, b]$ so are

$$f \pm g, h f, \text { and }|f|. \tag{9.2.46}$$

So also is f / h if

$$(\exists \varepsilon>0) \quad|h| \geq \varepsilon \text { on } A. \tag{9.2.47}$$

All this also holds if $f, g: E^{1} \rightarrow E$ are vector valued and h is scalar valued. Finally, if $E \subseteq E^{*}$, then

$$f \vee g, f \wedge g, f^{+}, \text {and } f^{-} \tag{9.2.48}$$

are absolutely continuous along with f and g.

Corollary 9.2.1

A function $F: E^{1} \rightarrow E^{n}\left(C^{n}\right)$ is absolutely continuous on $A=[a, b]$ iff all its components $F_{1}, \ldots, F_{n}$ are.

Hence a complex function $F: E^{1} \rightarrow C$ is absolutely continuous iff its real and imaginary parts, F_{re} and F_{im}, are.

Corollary 9.2.2

If $f: E^{1} \rightarrow E$ is absolutely continuous on $A=[a, b]$, it is bounded, is uniformly continuous, and has bounded variation, $V_{f}[a, b]<\infty$ all on A.

Lemma 9.2.2

If $F: E^{1} \rightarrow E^{n}\left(C^{n}\right)$ is of bounded variation on $A=[a, b]$, then

(i) F is a.e. differentiable on A, and

(ii) F' is L-integrable on A.

Proof

Via components (Theorem 4 of Chapter 5, §7), all reduces to the real case, $F: E^{1} \rightarrow E^{1}$.

Then since $V_{F}[A]<\infty$, we have

$$F=g-h \tag{9.2.49}$$

for some nondecreasing g and h (Theorem 3 in Chapter 5, §7).

Now, by Theorem 3 from Chapter 7, §10, g and h are a.e. differentiable on A. Hence so is

$$g-h=F. \tag{9.2.50}$$

Moreover, $g^{\prime} \geq 0$ and $h^{\prime} \geq 0$ since $g \uparrow$ and $h \uparrow$.

Thus for the L-integrability of $F^{\prime}$, proceed as in Problem 11 in §1, i.e., show that $F^{\prime}$ is measurable on A and that

$$L \int_{a}^{b} F^{\prime}=L \int_{a}^{b} g^{\prime}-L \int_{a}^{b} h^{\prime} \tag{9.2.51}$$

is finite. This yields the result. $\square$

Theorem 9.2.2 (Lebesgue)

If $F: E^{1} \rightarrow E^{n}\left(C^{n}\right)$ is absolutely continuous on $A=[a, b]$, then the following are true:

(i*) F is a.e. differentiable, and $F^{\prime}$ is L-integrable, on A.

(ii*) If, in addition, $F^{\prime}=0$ a.e. on A, then F is constant on A.

Proof

Assertion (i*) is immediate from Lemma 2, since any absolutely continuous function is of bounded variation by Corollary 2.

(ii*) Now let $F^{\prime}=0$ a.e. on A. Fix any

$$B=[a, c] \subseteq A \tag{9.2.52}$$

and let Z consist of all $p \in B$ at which the derivative $F^{\prime}=0$.

Given $\varepsilon>0$, let $\mathcal{K}$ be the set of all closed intervals $[p, x], p<x$, such that

$$\left|\frac{\Delta F}{\Delta x}\right|=\left|\frac{F(x)-F(p)}{x-p}\right|<\varepsilon. \tag{9.2.53}$$

By assumption,

$$\lim _{x \rightarrow p} \frac{\Delta F}{\Delta x}=0 \quad(p \in Z), \tag{9.2.54}$$

and $m(B-Z)=0 ; B=[a, c] \in \mathcal{M}^{*}$. If $p \in Z$, and $x-p$ is small enough, then

$$\left|\frac{\Delta F}{\Delta x}\right|<\varepsilon, \tag{9.2.55}$$

i.e., $[p, x] \in \mathcal{K}$.

It easily follows that $\mathcal{K}$ covers Z in the Vitali sense (verify!); so for any $\delta>0$, Theorem 2 of Chapter 7, §10 yields disjoint intervals

$$I_{k}=\left[p_{k}, x_{k}\right] \in \mathcal{K}, I_{k} \subseteq B, \tag{9.2.56}$$

with

$$m^{*}\left(Z-\bigcup_{k=1}^{q} I_{k}\right)<\delta, \tag{9.2.57}$$

hence also

$$m\left(B-\bigcup_{k=1}^{q} I_{k}\right)<\delta \tag{9.2.58}$$

(for $m(B-Z)=0$). But

$$\begin{aligned} B-\bigcup_{k=1}^{q} I_{k} &=[a, c]-\bigcup_{k=1}^{q-1}\left[p_{k}, x_{k}\right] \\ &=\left[a, p_{1}\right) \cup \bigcup_{k=1}^{q-1}\left[x_{k}, p_{k+1}\right) \cup\left[x_{q}, c\right] \quad\left(\text { if } x_{k}<p_{k}<x_{k+1}\right) ; \end{aligned}$$

so

$$m\left(B-\bigcup_{k=1}^{q} I_{k}\right)=\left(p_{1}-a\right)+\sum_{k=1}^{q-1}\left(p_{k+1}-x_{k}\right)+\left(c-x_{q}\right)<\delta . \tag{9.2.59}$$

Now, as F is absolutely continuous, we can choose $\delta>0$ so that (3) implies

$$\left|F\left(p_{1}\right)-F(a)\right|+\sum_{k=1}^{q-1}\left|F\left(p_{k+1}\right)-F\left(x_{k}\right)\right|+\left|F(c)-F\left(x_{q}\right)\right|<\varepsilon . \tag{9.2.60}$$

But $I_{k} \in \mathcal{K}$ also implies

$$\left|F\left(x_{k}\right)-F\left(p_{k}\right)\right|<\varepsilon\left(x_{k}-p_{k}\right)=\varepsilon \cdot m I_{k} . \tag{9.2.61}$$

Hence

$$\left|\sum_{k=1}^{q}\left[F\left(x_{k}\right)-F\left(p_{k}\right)\right]\right|<\varepsilon \sum_{k=1}^{q} m I_{k} \leq \varepsilon \cdot m B=\varepsilon(c-p). \tag{9.2.62}$$

Combining with (4), we get

$$|F(c)-F(a)| \leq \varepsilon(1+c-a) \rightarrow 0 \text { as } \varepsilon \rightarrow 0; \tag{9.2.63}$$

so $F(c)=F(a)$. As $c \in A$ was arbitrary, F is constant on A, as claimed. $\square$

Note 2. This shows that Cantor's function (Problem 6 of Chapter 4, §5) is not absolutely continuous, even though it is continuous and monotone, hence of bounded variation on $[0,1]$. Indeed (see Problem 2 in §1), it has a zero derivative a.e. on $[0,1]$ but is not constant there. Thus absolute continuity, as now defined, differs from its "weak" counterpart (Chapter 5, §8).

Theorem 9.2.3

A map $F: E^{1} \rightarrow E^{1}\left(C^{n}\right)$ is absolutely continuous on $A=[a, b]$ iff

$$F=L \int f \quad \text { on } A \tag{9.2.64}$$

for some function f; and then

$$F(x)=F(a)+L \int_{a}^{x} f, \quad x \in A. \tag{9.2.65}$$

Briefly: Absolutely continuous maps are exactly all L-primitives.

Proof

If $F=L \int f$, then by Theorem 1 of §1, F is absolutely continuous on A, and by Note 1,

$$F(x)=F(a)+L\int_{a}^{x} f, \quad x\in A. \tag{9.2.66}$$

Conversely, if F is absolutely continuous, then by Theorem 2, it is a.e. differentiable and $F'=f$ is L-integrable (all on A). Let

$$H(x)=L\int_{a}^{x} f, \quad x\in A. \tag{9.2.67}$$

Then H, too, is absolutely continuous and so is $F-H$. Also, by Theorem 3 of §1,

$$H'=f=F', \tag{9.2.68}$$

and so

$$(F-H)'=0 \quad \text{a.e. on } A. \tag{9.2.69}$$

By Theorem 2, $F-H=c$; i.e.,

$$F(x)=c+H(x)=c+L\int_{a}^{x} f, \tag{9.2.70}$$

and so $F=L\int f$ on A, as claimed. □

Corollary 9.2.3

If $f, F: E^{1}\rightarrow E^{*}\left(E^{n}, C^{n}\right)$, we have

$$F=L\int f \tag{9.2.71}$$

on an interval $I\subseteq E^{1}$ iff F is absolutely continuous on I and $F'=f$ a.e. on I.

(Use Problem 3 in §1 and Theorem 3.)

Note 3. This (or Theorem 3) could serve as a definition. Comparing ordinary primitives

$$F=\int f \tag{9.2.72}$$

with L-primitives

$$F=L\int f, \tag{9.2.73}$$

we see that the former require F to be just relatively continuous but allow only a countable "exceptional" set Q, while the latter require absolute continuity but allow Q to even be uncountable, provided $mQ=0$.

The simplest and "strongest" kind of absolutely continuous functions are so-called Lipschitz maps (see Problem 6). See also Problems 7 and 10.

III. We conclude with another important idea, due to Lebesgue.

Definition

We call $p\in E^{1}$ a Lebesgue point ("L-point") of $f: E^{1}\rightarrow E$ iff

(i) f is L-integrable on some $G_{p}(\delta)$;

(ii) $q=f(p)$ is finite; and

(iii) $\lim_{x\rightarrow p}\frac{1}{x-p} L\int_{p}^{x}|f-q|=0$.

The Lebesgue set of f consists of all such p.

Corollary 9.2.4

Let

$$F=L\int f \quad \text{on } A=[a,b]. \tag{9.2.74}$$

If $p\in A$ is an L-point of f, then $f(p)$ is the derivative of F at p (but the converse fails).

Proof

By assumption,

$$F(x)=c+L\int_p^x f, \quad x\in G_p(\delta), \tag{9.2.75}$$

and

$$\frac{1}{|\Delta x|}\left|L\int_p^x(f-q)\right|\leq\frac{1}{|\Delta x|}L\int_p^x|f-q|\rightarrow 0 \tag{9.2.76}$$

as $x\rightarrow p$. (Here $q=f(p)$ and $\Delta x=x-p$.)

Thus with $x\rightarrow p$, we get

$$\begin{aligned}\left|\frac{F(x)-F(p)}{x-p}-q\right| &= \frac{1}{|x-p|}\left|L\int_p^x f-(x-p)q\right| \\ &= \frac{1}{|x-p|}\left|L\int_p^x f-L\int_p^x(q)\right|\rightarrow 0,\end{aligned}$$

as required. □

Corollary 9.2.5

Let $f:E^1\rightarrow E^n\left(C^n\right)$. Then p is an L-point of f iff it is an L-point for each of the n components, $f_1,\ldots,f_n$, of f.

Proof

(Exercise!)

Theorem 9.2.4

If $f:E^1\rightarrow E^*\left(E^n,C^n\right)$ is L-integrable on $A=[a,b]$, then almost all $p\in A$ are Lebesgue points of f.

Note that this strengthens Theorem 3 of §1.

Proof

By Corollary 5, we need only consider the case $f:E^1\rightarrow E^*$.

For any $r\in E^1, |f-r|$ is L-integrable on A; so by Theorem 3 of §1, setting

$$F_r(x)=L\int_a^x|f-r|, \tag{9.2.77}$$

we get

$$F_r'(p)=\lim_{x\rightarrow p}\frac{1}{|x-p|}L\int_p^x|f-r|=|f(p)-r| \tag{9.2.78}$$

for almost all $p\in A$.

Now, for each r, let A_r be the set of those $p \in A$ for which (5) fails; so $mA_r = 0$. Let $\{r_k\}$ be the sequence of all rationals in E^1. Let

$$Q = \bigcup_{k=1}^{\infty} A_{r_k} \cup \{a, b\} \cup A_{\infty}, \tag{9.2.79}$$

where

$$A_{\infty} = A(|f| = \infty); \tag{9.2.80}$$

so $mQ = 0$. (Why?)

To finish, we show that all $p \in A - Q$ are L-points of f. Indeed, fix any $p \in A - Q$ and any $\varepsilon > 0$. Let $q = f(p)$. Fix a rational r such that

$$|q - r| < \frac{\varepsilon}{3}. \tag{9.2.81}$$

Then

$$||f - r| - |f - q|| \leq |(f - r) - (f - q)| = |q - r| < \frac{\varepsilon}{3} \text{ on } A - A_{\infty}. \tag{9.2.82}$$

Hence as $mA_{\infty} = 0$, we have

$$\left| L \int_p^x \left| f - r \right| - L \int_p^x \left| f - q \right| \right| \leq L \int_p^x \left(\frac{\varepsilon}{3}\right) = \frac{\varepsilon}{3} |x - p|. \tag{9.2.83}$$

Since

$$p \notin Q \supseteq \bigcup_k A_{r_k}, \tag{9.2.84}$$

formula (5) applies. So there is $\delta > 0$ such that $|x - p| < \delta$ implies

$$\left| \left(\frac{1}{|x - p|} L \int_p^x |f - r| \right) - \left| f(p) - r \right| \right| < \frac{\varepsilon}{3}. \tag{9.2.85}$$

As

$$|f(p) - r| = |q - r| < \frac{\varepsilon}{3}, \tag{9.2.86}$$

we get

$$\begin{aligned} \frac{1}{|x - p|} L \int_p^x |f - r| & \leq \left| \left(\frac{1}{|x - p|} L \int_p^x |f - r| \right) - \left| q - r \right| \right| + |q - r| \\ & < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3}. \end{aligned}$$

Hence

$$L \int_p^x |f - r| < \frac{2\varepsilon}{3} |x - p|. \tag{9.2.87}$$

Combining with (6), we have

$$\frac{1}{|x - p|} L \int_p^x |f - q| < \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} = \varepsilon \tag{9.2.88}$$

whenever $|x - p| < \delta$. Thus

$$\lim_{x \to p} \frac{1}{|x - p|} L \int_p^x |f - q| = 0, \tag{9.2.89}$$

as required. $\square$

9.2.E: Problems on L-Integrals and Absolute Continuity

Exercise $9.2.E.1$

Fill in all details in the proof of Lemma 1 and of Theorems 3 and 4 from §1.

Exercise $9.2.E.2$

Prove Theorem 1 and Corollaries 1, 2, and 5.

Exercise $9.2.E.2'$

Disprove the converse to Corollary 4. (Give an example!)

Exercise $9.2.E.3$

$\Rightarrow 3$. Show that if $F : E^1 \rightarrow E$ is L-integrable on $A = [a, b]$ and continuous at $p \in A$, then p is an L-point of F.
[Hint: Use the ε, δ definition of continuity.]

Exercise $9.2.E.4$

Complete all proof details for Lemma 2, Theorems 3 and 4, and Corollary 3.

Exercise $9.2.E.5$

Let $F = 1$ on $R(=$ rationals $)$ and $F = 0$ on $E^1 - R$ (Dirichlet function).
Show that F has exactly three derivates $(0, +\infty,$ and $-\infty)$ at every $p \in E^1$.

Exercise $9.2.E.6$

$\Rightarrow 6$. We say that F is a Lipschitz map, or has the uniform Lipschitz property on A, iff

$$\left(\exists K \in E^1\right)(\forall x, y \in A) \quad |F(x) - F(y)| \leq K|x - y|. \tag{9.2.E.1}$$

Prove the following:
(i) Any such F is absolutely continuous on $A = [a, b]$.
(ii) If all derivates of f satisfy

$$|Df(x)| \leq k < \infty, \quad x \in A = [a, b], \tag{9.2.E.2}$$

then f is a Lipschitz map on A.

Exercise $9.2.E.7$

$\Rightarrow 7$. Let $g : E^1 \rightarrow E^1$ and $f : E^1 \rightarrow E$ (real or not) be absolutely continuous on $A = [a, b]$ and $g[A]$, respectively.
Prove that $h = f \circ g$ is absolutely continuous on A, provided that either f is as in Problem 6, or g is strictly monotone on A.

Exercise $9.2.E.8$

Prove that if $F : E^1 \rightarrow E^1$ is absolutely continuous on $A = [a, b]$, if $Q \subseteq A$, and if $mQ = 0$, then $m^* F[Q] = 0 (m =$ Lebesgue measure).

[Outline: We may assume $Q \subseteq (a, b)$. (Why?)

Fix $\varepsilon > 0$ and take δ as in Definition 2. As m is regular, there is an open G,

$$Q \subseteq G \subseteq (a, b), \tag{9.2.E.3}$$

with $mG < \delta$. By Lemma 2 of Chapter 7, §2,

$$G = \bigcup_{k=1}^{\infty} I_k (\text{disjoint}) \tag{9.2.E.4}$$

for some $I_k = (a_k, b_k]$.

Let $u_k = \inf F[I_k], v_k = \sup F[I_k]$; so

$$F[I_k] \subseteq [u_k, v_k] \tag{9.2.E.5}$$

and

$$m^* F[I_k] \leq v_k - u_k. \tag{9.2.E.6}$$

Also,

$$\sum (b_k - a_k) = \sum m I_k = mG < \delta. \tag{9.2.E.7}$$

From Definition 2, show that

$$\sum_{k=1}^{\infty} (v_k - u_k) \leq \varepsilon \tag{9.2.E.8}$$

(first consider partial sums). As

$$F[Q] \subseteq F[G] \subseteq \bigcup_k F[I_k], \tag{9.2.E.9}$$

get

$$m^* F[Q] \leq \sum_k m^* F[I_k] = \sum_k (v_k - u_k) \leq \varepsilon \rightarrow 0. \Bigg] \tag{9.2.E.10}$$

? Exercise $9.2.E.9$

Show that if F is as in Problem 8 and if

$$A = [a, b] \supseteq B, \quad B \in \mathcal{M}^* \tag{9.2.E.11}$$

(L-measurable sets), then

$$F[B] \in \mathcal{M}^*. \tag{9.2.E.12}$$

(" F preserves $\mathcal{M}^*$-sets.")

[Outline: (i) If B is closed, it is compact, and so is $F[B]$ (Theorems 1 and 4 of Chapter 4, §6).

(ii) If $B \in \mathcal{F}_\sigma$, then

$$B=\bigcup_i B_i, \quad B_i \in \mathcal{F}; \tag{9.2.E.13}$$

so by (i),

$$F[B]=\bigcup_i F\left[B_i\right] \in \mathcal{F}_\sigma \subseteq \mathcal{M}^*. \tag{9.2.E.14}$$

(iii) If $B \in \mathcal{M}^*$, then by Theorem 2 of Chapter 7, §8,

$$\left(\exists K \in \mathcal{F}_\sigma\right) \quad K \subseteq B, m(B-K)=0. \tag{9.2.E.15}$$

Now use $\text{Problem } 8, \text{ with } Q=B-K.]$

Exercise $9.2.E.10$

$\Rightarrow 10.$ (Change of variable.) Suppose $g: E^1 \rightarrow E^1$ is absolutely continuous and one-to-one on $A=[a, b]$, while $f: E^1 \rightarrow E^*\left(E^n, C^n\right)$ is L-integrable on $g[A]$.
Prove that $(f \circ g) g'$ is L-integrable on A and

$$L \int_a^b(f \circ g) g'=L \int_p^q f, \tag{9.2.E.16}$$

where $p=g(a)$ and $q=g(b)$.
[Hints: Let $F=L \int f$ and $H=F \circ g$ on A.
By Theorems 2 and 3 and Problem 7 (end), F and H are absolutely continuous on $g[A]$ and A, respectively; and H' is L-integrable on A. So by Theorem 3

$$H=L \int H'=L \int(f \circ g) g', \tag{9.2.E.17}$$

as $H'=(f \circ g) g'$ a.e. on $A.]$

Exercise $9.2.E.11$

Setting $f(x)=0$ if not defined otherwise, find the intervals (if any) on which f is absolutely continuous if $f(x)$ is defined by
(a) $\sin x$;
(b) $\cos 2x$;
(c) $1/x$;
(d) $\tan x$;
(e) x^x;
(f) $x \sin(1/x)$;
(g) $x^2 \sin x^{-2}$ (Problem 5 in §1);
(h) $\sqrt{x^3} \cdot \sin(1/x)$ $\left(\text{verify that } |f'(x)| \leq \frac{3}{2}+x^{-\frac{1}{2}}\right)$;
[Hint: Use Problems 6 and 7.]

9.3: Improper (Cauchy) Integrals

Cauchy extended R-integration to unbounded sets and functions as follows.

Given $f : E^1 \rightarrow E$ and assuming that the right-hand side R-integrals and limits exist, define (first for unbounded sets, then for unbounded functions)

(i) $\int_a^\infty f = \int_{[a,\infty)} f = \lim_{x\to\infty} R\int_a^x f$;

(ii) $\int_{-\infty}^a f = \int_{(-\infty,a]} f = \lim_{x\to-\infty} R\int_x^a f$.

If both

$$\int_0^\infty f \text{ and } \int_{-\infty}^0 f \tag{9.3.1}$$

exists, define

$$\int_{-\infty}^\infty f = \int_{(-\infty,0)} f + \int_{[0,\infty)} f. \tag{9.3.2}$$

Now, suppose f is unbounded near some $p \in A = [a,b]$, i.e., unbounded on $A \cap G_{\neg p}$ for every deleted globe $G_{\neg p}$ about p (such points p are called **singularities**).

Then (again assuming existence of the R-integrals and limits), we define

1. in case of a singularity $p = a$,

$$\int_{a+}^b f = \int_{(a,b]} f = \lim_{x\to a+} R\int_x^b f; \tag{9.3.3}$$

2. if $p = b$, then

$$\int_a^{b-} f = \int_{[a,b)} f = \lim_{x\to b-} R\int_a^x f; \tag{9.3.4}$$

3. if $a < p < b$ and if

$$\int_a^{p-} f \text{ and } \int_{p+}^b f \tag{9.3.5}$$

exist, then

$$\int_a^b f = \int_a^{p-} f + \int_p^p f + \int_{p+}^b f. \tag{9.3.6}$$

The term

$$\int_p^p f = \int_{[p,p]} f \tag{9.3.7}$$

is necessary if RS- or LS-integrals are used.

Finally, if A contains several singularities, it must be split into subintervals, each with at most one endpoint singularity; and $\int_a^b f$ is split accordingly. We call all such integrals improper or Cauchy (C) integrals. A C-integral is said to converge iff it exists and is finite.

This theory is greatly enriched if in the above definitions, one replaces R-integrals by Lebesgue integrals, using Lebesgue or LS measure in E^1. (This makes sense even when a Lebesgue integral (proper) does exist; see Theorem 1.) Below, m shall denote such a measure unless stated otherwise.

C-integrals with respect to m will be denoted by

$$C\int_a^\infty f dm,\quad C\int_{[a,b)} f,\quad \text{etc.} \tag{9.3.8}$$

"Classical" notation:

$$C\int f(x)dm(x) \text{ or } C\int f(x)dx \tag{9.3.9}$$

(the latter if m is Lebesgue measure). We omit the "C" if confusion with proper integrals $\int_a^x f$ is unlikely.

Note 1. C-integrals are limits of integrals, not integrals proper. Yet they may equal the latter (Theorem 1 below) and then may be used to compute them.

Caution. "Singularities" in $[a,b]$ may affect the primitive used in computations (cf. Problem 4 in §1). Then $[a,b]$ must be split (see above), and $C\int_a^b f$ splits accordingly. (Additivity applies to C-integrals; see Problem 9, below.)

Examples

(A) The integral

$$L\int_{-1}^{1/2}\frac{dx}{x^2} \tag{9.3.10}$$

has a singularity at 0. By Theorem 1 below, we get

$$\begin{aligned} L\int_{-1}^{1/2}\frac{dx}{x^2} &= \int_{-1}^{0-}\frac{dx}{x^2}+\int_{0+}^{1/2}\frac{dx}{x^2} \\ &= \lim_{x\to 0^-}\left(-\frac{1}{x}-1\right)+\lim_{x\to 0+}\left(-2+\frac{1}{x}\right)=\infty+\infty=\infty. \end{aligned}$$

(B) We have

$$C\int_{1/2}^{\infty}\frac{dx}{x^2}=\lim_{x\to\infty}\left(-\frac{1}{x}+2\right)=2. \tag{9.3.11}$$

Hence

$$C\int_{-1}^{\infty}\frac{dx}{x^2}=C\int_{-1}^{1/2}\frac{dx}{x^2}+C\int_{1/2}^{\infty}\frac{dx}{x^2}=\infty+2=\infty. \tag{9.3.12}$$

(C) The integral

$$L\int_{-1}^{1}\frac{|x|}{x}dx \tag{9.3.13}$$

has no singularities (consider deleted globes about 0). The primitive $F(x)=|x|$ exists (example (b) in Chapter 5, §5); so

$$L\int_{-1}^{1}\frac{|x|}{x}dx=|x|\,\Big|_{-1}^{1}=0. \tag{9.3.14}$$

In the rest of this section, we state our theorems mainly for

$$C\int_a^\infty f, \tag{9.3.15}$$

but they apply, with similar proofs, to

$$C\int_{-\infty}^{\infty} f,\quad C\int_a^{b-} f,\quad \text{etc.} \tag{9.3.16}$$

The measure m is as explained above.

Theorem 9.3.1

Let $A=[a, \infty), f: E^{1} \rightarrow E \quad(E$ complete).

(i) If $f \geq 0$ on A, then

$$C \int_{a}^{\infty} f d m \tag{9.3.17}$$

exists $(\leq \infty)$ and equals

$$\int_{A} f d m . \tag{9.3.18}$$

(ii) The map f is m-integrable on A iff

$$C \int_{a}^{\infty}|f|<\infty \tag{9.3.19}$$

and f is m-measurable on A; then again,

$$C \int_{a}^{\infty} f d m=\int_{A} f d m . \tag{9.3.20}$$

Proof

(i) Let $f \geq 0$ on A. By the rules of Chapter 8, §5, $\int_{A} f$ is always defined for such f; so we may set

$$F(x)=\int_{a}^{x} f d m, \quad x \geq a . \tag{9.3.21}$$

Then by Theorem 1(f) in Chapter 8, §5, $F \uparrow$ on A; for $a \leq x \leq y$ implies

$$F(x)=\int_{a}^{x} f \leq \int_{a}^{y} f=F(y) . \tag{9.3.22}$$

Now, by the properties of monotone limits,

$$\lim _{x \rightarrow \infty} F(x)=\lim _{x \rightarrow \infty} \int_{a}^{x} f=C \int_{a}^{\infty} f \tag{9.3.23}$$

exists in E^{*}; so by Theorem 1 of Chapter 4, §2, it can be found by making x run over some sequence $x_{k} \rightarrow \infty$, say, $x_{k}=k$.

Thus set

$$A_{k}=[a, k], \quad k=1,2, \ldots . \tag{9.3.24}$$

Then $\left\{A_{k}\right\} \uparrow$ and

$$\bigcup A_{k}=A=[a, \infty), \tag{9.3.25}$$

i.e., $A_{k} \nearrow A$.

Moreover, by Note 4 in Chapter 8, §5, the set function $s=\int f$ is σ-additive and semifinite (≥ 0). Thus by Theorem 2 of Chapter 7, §4 (left continuity)

$$\int_{A} f d m=\lim _{k \rightarrow \infty} \int_{A_{k}} f=\lim _{k \rightarrow \infty} \int_{a}^{k} f=C \int_{a}^{\infty} f, \tag{9.3.26}$$

proving (i).

(ii) By clause (i),

$$C \int_{a}^{\infty}|f|=\int_{A}|f| d m \tag{9.3.27}$$

exists, as $|f| \geq 0$. Hence

$$C\int_a^\infty |f| < \infty \tag{9.3.28}$$

plus measurability amounts to integrability (Theorem 2 of Chapter 8, §6).

Moreover,

$$C\int_a^\infty |f| < \infty \tag{9.3.29}$$

implies the convergence of $C\int_a^\infty f$ (see Corollary 1 below). Thus as

$$\lim_{x\to\infty}\int_a^x f \tag{9.3.30}$$

exists, we proceed exactly as before (here $s = \int f$ is finite), proving (ii) also. □

Note 2. If $E \subseteq E^*$, formula (1) results even if f is not m-measurable.

Note 3. While f cannot be integrable unless $|f|$ is (Corollary 2 of Chapter 8, §6), it can happen that

$$C\int f \tag{9.3.31}$$

converges even if

$$C\int |f| = \infty \tag{9.3.32}$$

(this is called conditional convergence). A case in point is

$$C\int_0^\infty \frac{\sin x}{x} dx; \tag{9.3.33}$$

see Problem 8.

Thus C-integrals may be finite where proper integrals are ∞ or fail to exist (a great advantage!). Yet they are deficient in other respects (see Problem 9(c)).

For our next theorem, we need the previously "starred" Theorem 2 in Chapter 4, (Review it!) As we shall see, C-integrals resemble infinite series.

Theorem 9.3.2 (Cauchy criterion)

Let $A = [a, \infty), f : E^1 \to E, E$ complete.
Suppose

$$\int_a^x f dm \tag{9.3.34}$$

exists for each $x \in A$. (This is automatic if $E \subseteq E^*$; see Chapter 8, §5.)

Then

$$C\int_a^\infty f \tag{9.3.35}$$

converges iff for every $\varepsilon > 0$, there is $b \in A$ such that

$$\left|\int_v^x f dm\right| < \varepsilon \quad \text{whenever } b \leq v \leq x < \infty, \tag{9.3.36}$$

and

$$\left|\int_{a}^{b} f dm\right|<\infty. \tag{9.3.37}$$

Proof

By additivity (Chapter 8, §5, Theorem 2; Chapter 8, §7, Theorem 3),

$$\int_{a}^{x} f=\int_{a}^{v} f+\int_{v}^{x} f \tag{9.3.38}$$

if $a \leq v \leq x<\infty.$ (In case $E \subseteq E^{*}$, this holds even if f is not integrable; see Theorem 2, of Chapter 8, §5.)

Now, if

$$C \int_{a}^{\infty} f \tag{9.3.39}$$

converges, let

$$r=\lim _{x \rightarrow \infty} \int_{a}^{x} f dm \neq \pm \infty. \tag{9.3.40}$$

Then for any $\varepsilon>0,$ there is some

$$b \in[a, \infty)=A \tag{9.3.41}$$

such that

$$\left|\int_{a}^{x} f dm-r\right|<\frac{1}{2} \varepsilon \quad \text { for } x \geq b. \tag{9.3.42}$$

(Why may we use the standard metric here?)

Taking $x=b,$ we get (2'). Also, if $a \leq b \leq v \leq x,$ we have

$$\left|\int_{a}^{x} f dm-r\right|<\frac{1}{2} \varepsilon \tag{9.3.43}$$

and

$$\left|r-\int_{a}^{\nu} f dm\right|<\frac{1}{2} \varepsilon. \tag{9.3.44}$$

Hence by the triangle law, (2) follows also. Thus this b satisfies (2).

Conversely, suppose such a b exists for every given $\varepsilon>0$. Fixing b, we thus have (2) and (2'). Now, with $A=[a, \infty)$, define $F: A \rightarrow E$ by

$$F(x)=\int_{a}^{x} f dm, \tag{9.3.45}$$

so

$$C \int_{a}^{\infty} f=\lim _{x \rightarrow \infty} F(x) \tag{9.3.46}$$

if this limit exists. By (2),

$$|F(x)|=\left|\int_{a}^{x} f dm\right| \leq\left|\int_{a}^{b} f dm\right|+\left|\int_{b}^{x} f dm\right|<\left|\int_{a}^{b} f dm\right|+\varepsilon \tag{9.3.47}$$

if $x \geq b.$ Thus F is finite on $[b, \infty)$, and so we may again use the standard metric

$$\rho(F(x), F(v))=|F(x)-F(v)|=\left|\int_{a}^{x} f dm-\int_{a}^{v} f dm\right| \leq\left|\int_{v}^{x} f dm\right|<\varepsilon \tag{9.3.48}$$

if $x, v \geq b$. The existence of

$$C \int_a^\infty f dm = \lim_{x \to \infty} F(x) \neq \pm\infty \tag{9.3.49}$$

now follows by Theorem 2 of Chapter 4, §2. (We shall henceforth presuppose this "starred" theorem.)

Thus all is proved. □

Corollary 9.3.1

Under the same assumptions as in Theorem 2, the convergence of

$$C \int_a^\infty |f| dm \tag{9.3.50}$$

implies that of

$$C \int_a^\infty f dm. \tag{9.3.51}$$

Indeed,

$$\left| \int_v^x f \right| \leq \int_v^x |f| \tag{9.3.52}$$

(Theorem 1(g) of Chapter 8, §5, and Problem 10 in Chapter 8, §7).

Note 4. We say that $C \int f$ converges absolutely iff $C \int |f|$ converges.

Corollary 9.3.2 (comparison test)

If $|f| \leq |g|$ a.e. on $A = [a, \infty)$ for some $f, g : E^1 \rightarrow E,$ then

$$C \int_a^\infty |f| \leq C \int_a^\infty |g|; \tag{9.3.53}$$

so the convergence of

$$C \int_a^\infty |g| \tag{9.3.54}$$

implies that of

$$C \int_a^\infty |f|. \tag{9.3.55}$$

For as $|f|, |g| \geq 0$, Theorem 1 reduces all to Theorem 1(c) of Chapter 8, §5.

Note 5. As we see, absolutely convergent C-integrals coincide with proper (finite) Lebesgue integrals of nonnegative or m-measurable maps. For conditional (i.e., nonabsolute) convergence, see Problems 6-9, 13, and 14.

Iterated C-Integrals. Let the product space $X \times Y$ of Chapter 8, §8 be

$$E^1 \times E^1 = E^2, \tag{9.3.56}$$

and let $p = m \times n,$ where m and n are Lebesgue measure or LS measures in E^1. Let

$$A = [a, b], B = [c, d], \text{ and } D = A \times B. \tag{9.3.57}$$

Then the integral

$$\int_B \int_A f dm dn = \int_Y \int_X f C_D dm dn \tag{9.3.58}$$

is also written

$$\int_c^d \int_a^b f dm dn \tag{9.3.59}$$

or

$$\int_c^d \int_a^b f(x, y) dm(x) dn(y). \tag{9.3.60}$$

As usual, we write "dx" for "$dm(x)$" if m is Lebesgue measure in E^1; similarly for n.

We now define

$$\begin{aligned} C \int_a^\infty \int_c^\infty f dn dm &= \lim_{b \rightarrow \infty} \int_a^b \left(\lim_{d \rightarrow \infty} \int_c^d f(x, y) dn(y) \right) dm(x) \\ &= C \int_a^\infty \int_c^\infty f(x, y) dn(y) dm(x), \end{aligned}$$

provided the limits and integrals involved exist.

If the integral (3) is finite, we say that it converges. Again, convergence is absolute if it holds also with f replaced by $|f|$, and conditional otherwise. Similar definitions apply to

$$C \int_c^\infty \int_a^\infty f dm dn, C \int_{-\infty}^b \int_c^\infty f dn dm, \text{ etc.} \tag{9.3.61}$$

Theorem 9.3.3

Let $f: E^2 \rightarrow E^*$ be p-measurable on E^2 (p, m, n as above). Then we have the following.

(i*) The Cauchy integrals

$$C \int_{-\infty}^\infty \int_{-\infty}^\infty |f| dn dm \text{ and } C \int_{-\infty}^\infty \int_{-\infty}^\infty |f| dm dn \tag{9.3.62}$$

exist $(\leq \infty)$, and both equal

$$\int_{E^2} |f| dp. \tag{9.3.63}$$

(ii*) If one of these three integrals is finite, then

$$C \int_{-\infty}^\infty \int_{-\infty}^\infty f dn dm \text{ and } C \int_{-\infty}^\infty \int_{-\infty}^\infty f dm dn \tag{9.3.64}$$

converge, and both equal

$$\int_{E^2} f dp. \tag{9.3.65}$$

(Similarly for $C \int_a^\infty \int_{-\infty}^b f dn dm$, etc.)

Proof

As m and n are σ-finite (finite on intervals!), f surely has σ-finite support.

As $|f| \geq 0$, clause (i*) easily follows from our present Theorem 1(i) and Theorem 3(i) of Chapter 8, §8.

Similarly, clause (ii*) follows from Theorem 3(ii) of the same section. □

Theorem 9.3.4 (passage to polars)

Let $p=$ Lebesgue measure in E^2. Suppose $f: E^2 \rightarrow E^*$ is p-measurable on E^2. Set

$$F(r, \theta)=f(r \cos \theta, r \sin \theta), \quad r>0. \tag{9.3.66}$$

Then

(a) $C \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f d x d y=C \int_0^{\infty} r d r \int_0^{2 \pi} F d \theta$, and

(b) $C \int_0^{\infty} \int_0^{\infty} f d x d y=C \int_0^{\infty} r d r \int_0^{\pi / 2} F d \theta$,

provided f is nonnegative or p-integrable on E^2 (for (a)) or on $(0, \infty) \times(0, \infty)$ (for (b)).

Proof Outline

First let $f=C_D$, with D a "curved rectangle"

$$\{(r, \theta) | r_1<r \leq r_2, \theta_1<\theta \leq \theta_2\} \tag{9.3.67}$$

for some $r_1<r_2$ in $X=(0, \infty)$ and $\theta_1<\theta_2$ in $Y=[0,2 \pi)$. By elementary geometry (or calculus), the area

$$p D=\frac{1}{2}\left(r_2^2-r_1^2\right)\left(\theta_2-\theta_1\right) \tag{9.3.68}$$

(the difference between two circular sectors).

For $f=C_D$, formulas (a) and (b) easily follow from

$$p D=L \int_{E^2} C_D d p. \tag{9.3.69}$$

(Verify!) Now, curved rectangles behave like half-open intervals

$$\left(r_1, r_2\right] \times\left(\theta_1, \theta_2\right] \tag{9.3.70}$$

in E^2, since Theorem 1 in Chapter 7, §1, and Lemma 2 of Chapter 7, §2, apply with the same proof. Thus they form a semiring generating the Borel field in E^2.

Hence show (as in Chapter 8, §8 that Theorem 4 holds for $f=C_D(D \in \mathcal{B})$. Then take $D \in \mathcal{M}^*$. Next let f be elementary and nonnegative, and so on, as in Theorems 2 and 3 in Chapter 8, §8. $\square$

Examples (continued)

(D) Let

$$J=L \int_0^{\infty} e^{-x^2} d x; \tag{9.3.71}$$

so

$$\begin{aligned} J^2 &=\left(C \int_0^{\infty} e^{-x^2} d x\right)\left(C \int_0^{\infty} e^{-y^2} d y\right) \\ &=C \int_0^{\infty} \int_0^{\infty} e^{-\left(x^2+y^2\right)} d x d y. \quad \text{(Why?)} \end{aligned}$$

Set

$$f(x, y)=e^{-\left(x^2+y^2\right)} \tag{9.3.72}$$

in Theorem 4(b). Then $F(r, \theta)=e^{-r^2}$; hence

$$J^2 = C\int_0^\infty rdr\left(\int_0^{\frac{\pi}{2}} e^{-r^2}\,d\theta\right)$$
$$= C\int_0^\infty re^{-r^2}\,dr\cdot\frac{\pi}{2} = -\frac{1}{4}\pi e^{-t}\Big|_0^\infty = \frac{1}{4}\pi.$$

(Here we computed

$$\int re^{-r^2}\,dr \tag{9.3.73}$$

by substituting $r^2 = t$.) Thus

$$C\int_0^\infty e^{-x^2}\,dx = L\int_0^\infty e^{-x^2}\,dx = \sqrt{\frac{1}{4}\pi} = \frac{1}{2}\sqrt{\pi}. \tag{9.3.74}$$

9.3.E: Problems on Cauchy Integrals

Exercise $9.3.E.1$

Fill in all proof details in Theorems $1-3$. Verify also at least some of the cases other than $\int_a^\infty f$. Check the validity for LS-integrals (footnote 6).

Exercise $9.3.E.2$

Prove Theorem 4 in detail.

Exercise $9.3.E.2'$

Verify Notes 2 and 3 and examples (A)-(D).

Exercise $9.3.E.3$

Assuming $a>0,$ verify the following:
(i) $\int_1^\infty \frac{1}{t} e^{-t} dt \leq \int_1^\infty e^{-t} dt = \frac{1}{e}$.
[Hint: Use Corollary 2.]
(ii) $\int_1^\infty e^{-at} dt = \frac{e^{-a}}{a}$.
(iii) $\int_0^\infty e^{-at} dt = \frac{1}{a}$.
(iv) $\int_0^\infty e^{-at} \sin bt dt = \frac{b}{a^2+b^2}$.

Exercise $9.3.E.4$

Verify the following:
(i) $\int_1^\infty \int_1^\infty e^{-xy} dy dx = \int_1^\infty \frac{1}{x} e^{-x} dx \leq \frac{1}{e}$(converges, by 3(i)) .
(ii) $\int_0^\infty \int_0^\infty e^{-xy} dy dx \geq \int_1^\infty \int_0^\infty e^{-xy} dy dx = \int_1^\infty \frac{1}{x}(1-e^{-x}) dx \geq \int_1^\infty \left(\frac{1}{x}-e^{-x}\right) dx = \infty$.
Does this contradict formula (4) in the text, or Problem 5, which follows?

Exercise $9.3.E.5$

Let $f(x,y)=e^{-xy}$ and

$$g(x)=L\int_0^1 e^{-xy} dy; \tag{9.3.E.1}$$

so $g(0)=1.$ (Why?)
(i) Is g R-integrable on $A=[0,1]$? Is f so on $A\times A$?
(ii) Find $g(x)$ using Corollary 1 in §1.
(iii) Find the value of

$$R\int_0^1 \int_0^1 e^{-xy} dy dx = R\int_0^1 g \tag{9.3.E.2}$$

to within $1/10$.
[Hint: Reduce it to Problem 6(b) in §1.]

Exercise 9.3.E.6

$\Rightarrow 6$. Let $f, g : E^1 \rightarrow E^*$ be m-measurable on $A=[a, b), b \leq \infty$. Prove the following:
(i) If

$$C \int_a^{b-} f^+ < \infty \text{ or } C \int_a^{b-} f^- < \infty, \tag{9.3.E.3}$$

then $C \int_a^{b-} f$ exists and equals

$$C \int_a^{b-} f^+ - C \int_a^{b-} f^- = \int_A f dm (\text{ proper }). \tag{9.3.E.4}$$

(ii) If $\int_a^{b-} f$ converges conditionally only, then

$$\int_a^{b-} f^+ = \int_a^{b-} f^- = +\infty. \tag{9.3.E.5}$$

(iii) In case $C \int_a^{b-} |f| < \infty$, we have

$$C \int_a^{b-} |f \pm g| = \infty \tag{9.3.E.6}$$

iff $C \int_a^{b-} |g| = \infty$; also,

$$C \int_a^{b-} (f \pm g) = C \int_a^{b-} f \pm C \int_a^{b-} g \tag{9.3.E.7}$$

if $C \int_a^{b-} g$ exists (finite or not).

Exercise 9.3.E.7

$\Rightarrow 7$. Suppose $f : E^1 \rightarrow E^*$ is m-integrable and sign-constant on each

$$A_n = [a_n, a_{n+1}), \quad n = 1, 2, \ldots \tag{9.3.E.8}$$

but changes sign from A_n to A_{n+1}, with

$$\bigcup_{n=1}^{\infty} A_n = [a, \infty) \tag{9.3.E.9}$$

and $\{a_n\} \uparrow$ fixed.
Prove that if

$$\left| \int_{A_n} f dm \right| \searrow 0 \tag{9.3.E.10}$$

as $n \rightarrow \infty$, then

$$c \int_a^{\infty} f \tag{9.3.E.11}$$

converges.
[Hint: Use Problem 10 in Chapter 4, §13.]

Exercise $9.3.E.8$

$\Rightarrow 8$. Let

$$f(x)=\frac{\sin x}{x}, \quad f(0)=1. \tag{9.3.E.12}$$

Prove that

$$C \int_{0}^{\infty} f(x) d x \tag{9.3.E.13}$$

converges conditionally only.
[Hints: Use Problem 7. Show that

$$C \int_{0}^{\infty}|f|=L \int_{(0, \infty)}|f|=L \int_{0}^{\infty} f^{+}=L \int_{0}^{\infty} f^{-}=\infty .] \tag{9.3.E.14}$$

Exercise $9.3.E.9$

$\Rightarrow 9$. (Additivity.) Given $f: E^{1} \rightarrow E(E$ complete$)$ and $a<b<c \leq \infty$, suppose that

$$\int_{a}^{x} f d m \neq \pm \infty \tag{9.3.E.15}$$

(proper) exists for each $x \in[a, c)$. Prove the following:
(a) $C \int_{a}^{b-} f$ and $C \int_{a+}^{b} f$ converge.
(b) If

$$C \int_{b}^{c-} f \tag{9.3.E.16}$$

converges, so does

$$C \int_{a}^{c-} f=C \int_{a}^{b-} f+C \int_{b}^{c-} f. \tag{9.3.E.17}$$

(c) Countable additivity does not necessarily hold for C-integrals.
[Hint: Use Problem 8 suitably splitting $[0, \infty)$.]

Exercise $9.3.E.10$

(Refined comparison test.) Given $f, g: E^{1} \rightarrow E(E$ complete$)$ and $b \leq \infty$, prove the following:
(i) If for some $a<b$ and $k \in E^{1}$,

$$|f| \leq|k g| \quad \text { on }[a, b) \tag{9.3.E.18}$$

then

$$\int_a^{b-}|g|<\infty \text{ implies } \int_a^{b-}|f|<\infty. \tag{9.3.E.19}$$

(ii) Such $a, k \in E^1$ do exist if

$$\lim_{t \rightarrow b-} \frac{|f(t)|}{|g(t)|}<\infty \tag{9.3.E.20}$$

exists.
(iii) If this limit is not zero, then

$$\int_a^{b-}|g|<\infty \text{ iff } \int_a^{b-}|f|<\infty. \tag{9.3.E.21}$$

(Similarly in the case of $\int_{a+}^b$ with $a \geq-\infty.\Big)$

Exercise $9.3.E.11$

Prove that
(I) $\int_1^{\infty} t^p d t<\infty$ iff $p<-1$;
(ii) $\int_{0+}^1 t^p d t<\infty$ iff $p>-1$;
(iii) $\int_{0+}^{\infty} t^p d t=\infty$.

Exercise $9.3.E.12$

Use Problems 10 and 11 to test for convergence of the following:
(a) $\int_0^{\infty} \frac{t^{3/2} d t}{1+t^2}$;
(b) $\int_1^{\infty} \frac{d t}{t \sqrt{1+t^2}}$;
(c) $\int_a^{\infty} \frac{P(t)}{Q(t)} d t$
$(Q, P$ polynomials of degree s and $r, s>r; Q \neq 0$ for $t \geq a)$;
(d) $\int_0^{1-} \frac{d t}{\sqrt{1-t^4}}$;
(e) $\int_{0+}^1 t^p \ln t d t$;
(f) $\int_0^{1-} \frac{d t}{\ln t}$;
(g) $\int_{0+}^{\frac{\pi}{2}-} \tan^p t d t$.

Exercise $9.3.E.13$

$\Rightarrow 13$. (The Abel-Dirichlet test.) Given $f, g: E^1 \rightarrow E^1$, suppose that
(a) $f \downarrow$, with $\lim_{t \rightarrow \infty} f(t)=0$;
(b) g is L-measurable on $A=[a, \infty)$; and;
(c) $\left(\exists K \in E^1\right)(\forall x \in A) \quad\left|L \int_a^x g\right|<K$.
Then $C \int_a^{\infty} f(x) g(x) d x$ converges.
[Outline: Set

$$G(x)=\int_a^x g; \tag{9.3.E.22}$$

so $|G|<K$ on A. By Lemma 2 of §1, fg is L-integrable on each $[u, v] \subset A$, and $(\exists c \in[u, v])$ such that

$$\left|L\int_u^v fg\right| = \left|f(u)\int_u^c g\right| = |f(u)[G(c)-G(u)]| < 2Kf(u). \tag{9.3.E.23}$$

Now, by (a),

$$(\forall \varepsilon > 0)(\exists k \in A)(\forall u \geq k) \quad |f(u)| < \frac{\varepsilon}{2K}; \tag{9.3.E.24}$$

so

$$(\forall v \geq u \geq k) \quad \left|L\int_u^v fg\right| < \varepsilon. \tag{9.3.E.25}$$

Now use Theorem 2.

Now extend this to $g: E^1 \rightarrow E^n\ (C^n)$.]

Exercise $9.3.E.14$

$\Rightarrow$ 14. Do Problem 13, replacing assumptions (a) and (c) by

(a') f is monotone and bounded on $[a, \infty) = A,$ and

(c') $C\int_a^\infty g(x)dx$ converges.

[Hint: If $f \uparrow$, say, set $q = \lim_{t\rightarrow\infty} f(t)$ and $F = q - f;$ so

$$fg = qg - Fg. \tag{9.3.E.26}$$

Apply Problem 13 to

$$C\int_a^\infty F(x)g(x)dx.\Big] \tag{9.3.E.27}$$

Exercise $9.3.E.15$

Use Problems 13 and 14 to test the convergence of the following:

(a) $\int_0^\infty t^p \sin t dt$.

[Hint: The integral converges iff $p < 0$.]

(b) $\int_{0+}^\infty \frac{\cos t}{\sqrt{t}} dt$.

$\left[\text{Hint: Integrate } \int_u^v \frac{\cos t}{\sqrt{t}} dt \text{ by parts; then let } u \rightarrow 0 \text{ and } v \rightarrow \infty.\right]$

(c) $\int_1^\infty \frac{\cos t}{t^p} dt$.

(d) $\int_0^\infty \sin t^2 dt$.

$\left[\text{Hint: Substitute } t^2 = u; \text{ then use (a).}\right]$

Exercise $9.3.E.16$

The Cauchy principal value (CPV) of $C\int_{-\infty}^\infty f(t)dt$ is defined by

$$(\mathrm{CPV})\int_{-\infty}^\infty f = \lim_{x\rightarrow\infty}\int_{-x}^x f(t)dt \tag{9.3.E.28}$$

(if it exists). Prove the following:

(i) If $C\int f(t)dt$ exists, so does $(\mathrm{CPV})\int f$, and the two are equal.

Disprove the converse.

[Hint: Take $f(t)=\operatorname{sign}(t)/\sqrt{|t|}$.]

(ii) Do the same for

$$(\mathrm{CPV})\int_{a}^{b} f=\lim_{\delta\rightarrow 0+}\left(\int_{a}^{p-\delta} f+\int_{p+\delta}^{b} f\right), \tag{9.3.E.29}$$

p being the only singularity in (a, b).

9.4: Convergence of Parametrized Integrals and Functions

I. We now consider C-integrals of the form

$$C \int f(t, u) dm(t), \tag{9.4.1}$$

where m is Lebesgue or LS measure in E^1. Here the variable u, called a parameter, remains fixed in the process of integration; but the end result depends on u, of course.

We assume $f : E^2 \rightarrow E$ (E complete) even if not stated explicitly. As before, we give our definitions and theorems for the case

$$C \int_a^\infty . \tag{9.4.2}$$

The other cases $\left(C \int_{-\infty}^a, C \int_a^{b-}, \text{ etc.}\right)$ are analogous; they are treated in Problems 2 and 3. We assume

$$a, b, c, x, t, u, v \in E^1 \tag{9.4.3}$$

throughout, and write "dt" for "$dm(t)$" iff m is Lebesgue measure.

If

$$C \int_a^\infty f(t, u) dm(t) \tag{9.4.4}$$

converges for each u in a set $B \subseteq E^1,$ we can define a map $F : B \rightarrow E$ by

$$F(u) = C \int_a^\infty f(t, u) dm(t) = \lim_{x \rightarrow \infty} \int_a^x f(t, u) dm(t). \tag{9.4.5}$$

This means that

$$(\forall u \in B)(\forall \varepsilon > 0)(\exists b > a)(\forall x \geq b) \quad \left| \int_a^x f(t, u) dm(t) - F(u) \right| < \varepsilon, \tag{9.4.6}$$

so $|F| < \infty$ on B.

Here b depends on both ε and u (convergence is "pointwise"). However, it may occur that one and the same b fits all $u \in B$, so that b depends on ε alone. We then say that

$$C \int_a^\infty f(t, u) dm(t) \tag{9.4.7}$$

converges uniformly on B (i.e., for $u \in B$), and write

$$F(u) = C \int_a^\infty f(t, u) dm(t) \text{ (uniformly) on } B. \tag{9.4.8}$$

Explicitly, this means that

$$(\forall \varepsilon > 0)(\exists b > a)(\forall u \in B)(\forall x \geq b) \quad \left| \int_a^x f(t, u) dm(t) - F(u) \right| < \varepsilon. \tag{9.4.9}$$

Clearly, this implies (1), but not conversely. We now obtain the following.

Theorem 9.4.1 (Cauchy criterion)

Suppose

$$\int_a^x f(t, u) dm(t) \tag{9.4.10}$$

exists for $x \geq a$ and $u \in B \subseteq E^1.$ (This is automatic if $E \subseteq E^*;$ see Chapter 8, §5.)

Then

$$C\int_a^\infty f(t,u)dm(t) \tag{9.4.11}$$

converges uniformly on B iff for every $\varepsilon>0,$ there is $b>a$ such that

$$(\forall v,x\in[b,\infty))(\forall u\in B)\quad\left|\int_v^x f(t,u)dm(t)\right|<\varepsilon, \tag{9.4.12}$$

and

$$\left|\int_a^b f(t,u)dm(t)\right|<\infty. \tag{9.4.13}$$

Proof

The necessity of (3) follows as in Theorem 2 of §3. (Verify!)

To prove sufficiency, suppose the desired b exists for every $\varepsilon>0.$ Then for each (fixed) $u\in B,$

$$C\int_a^\infty f(t,u)dm(t) \tag{9.4.14}$$

satisfies Theorem 2 of §3. Hence

$$F(u)=\lim_{x\to\infty}\int_a^x f(t,u)dm(t)\neq\pm\infty \tag{9.4.15}$$

exists for every $u\in B$ (pointwise). Now, from (3), writing briefly $\int f$ for $\int f(t,u)dm(t)$, we obtain

$$\left|\int_v^x f\right|=\left|\int_a^x f-\int_a^v f\right|<\varepsilon \tag{9.4.16}$$

for all $u\in B$ and all $x>v\geq b$.

Making $x\to\infty$ (with u and v temporarily fixed), we have by (4) that

$$\left|F(u)-\int_a^v f\right|\leq\varepsilon \tag{9.4.17}$$

whenever $v\geq b$.

But by our assumption, b depends on ε alone (not on u). Thus unfixing $u,$ we see that (5) establishes the uniform convergence of

$$\int_a^\infty f, \tag{9.4.18}$$

as required. □

Corollary 9.4.1

Under the assumptions of Theorem 1,

$$C\int_a^\infty f(t,u)dm(t) \tag{9.4.19}$$

converges uniformly on B if

$$C\int_a^\infty |f(t,u)|dm(t) \tag{9.4.20}$$

does.

Indeed,

$$\left|\int_v^x f\right| \leq \int_v^x |f| < \varepsilon. \tag{9.4.21}$$

Corollary 9.4.2 (comparison test)

Let $f : E^2 \rightarrow E$ and $M : E^2 \rightarrow E^*$ satisfy

$$|f(t, u)| \leq M(t, u) \tag{9.4.22}$$

for $u \in B \subseteq E^1$ and $t \geq a$.

Then

$$C \int_a^\infty |f(t, u)| dm(t) \tag{9.4.23}$$

converges uniformly on B if

$$C \int_a^\infty M(t, u) dm(t) \tag{9.4.24}$$

does.

Indeed, Theorem 1 applies, with

$$\left|\int_v^x f\right| \leq \int_v^x M < \varepsilon. \tag{9.4.25}$$

Hence we have the following corollary.

Corollary 9.4.3 ("M-test")

Let $f : E^2 \rightarrow E$ and $M : E^1 \rightarrow E^*$ satisfy

$$|f(t, u)| \leq M(t) \tag{9.4.26}$$

for $u \in B \subseteq E^1$ and $t \geq a$. Suppose

$$C \int_a^\infty M(t) dm(t) \tag{9.4.27}$$

converges. Then

$$C \int_a^\infty |f(t, u)| dm(t) \tag{9.4.28}$$

converges (uniformly) on B. So does

$$C \int_a^\infty f(t, u) dm(t) \tag{9.4.29}$$

by Corollary 1.

Proof

Set

$$h(t, u) = M(t) \geq |f(t, u)|. \tag{9.4.30}$$

Then Corollary 2 applies (with M replaced by h there). Indeed, the convergence of

$$C \int h = C \int M \tag{9.4.31}$$

is trivially "uniform" for $u \in B$, since M does not depend on u at all. $\square$

Note 1. Observe also that, if $h(t, u)$ does not depend on u, then the (pointwise) and (uniform) convergence of $C \int h$ are trivially equivalent.

We also have the following result.

Corollary 9.4.4

Suppose

$$C \int_a^\infty f(t, u) dm(t) \tag{9.4.32}$$

converges (pointwise) on $B \subseteq E^1$. Then this convergence is uniform iff

$$\lim_{\nu \to \infty} C \int_v^\infty f(t, u) dm(t) = 0 \text{ (uniformly) on } B, \tag{9.4.33}$$

i.e., iff

$$(\forall \varepsilon > 0)(\exists b > a)(\forall u \in B)(\forall v \geq b) \quad \left| C \int_v^\infty f(t, u) dm(t) \right| < \varepsilon. \tag{9.4.34}$$

Proof

The proof (based on Theorem 1) is left to the reader, along with that of the following corollary.

Corollary 9.4.5

Suppose

$$\int_a^b f(t, u) dm(t) \neq \pm \infty \tag{9.4.35}$$

exists for each $u \in B \subseteq E^1$.

Then

$$C \int_a^\infty f(t, u) dm(t) \tag{9.4.36}$$

converges (uniformly) on B iff

$$C \int_b^\infty f(t, u) dm(t) \tag{9.4.37}$$

does.

II. The Abel-Dirichlet tests for uniform convergence of series (Problems 9 and 11 in Chapter 4, §13) have various analogues for C-integrals. We give two of them, using the second law of the mean (Corollary 5 in §1).

First, however, we generalize our definitions, "unstarring" some ideas of Chapter 4, §11. Specifically, given

$$H : E^2 \rightarrow E \text{ (} E \text{ complete)}, \tag{9.4.38}$$

we say that $H(x, y)$ converges to $F(y)$, uniformly on B, as $x \rightarrow q \, (q \in E^*)$, and write

$$\lim_{x \rightarrow q} H(x, y) = F(y) \text{ (uniformly) on } B \tag{9.4.39}$$

iff we have

$$(\forall \varepsilon > 0) \left(\exists G_{\neg q}\right) (\forall y \in B) \left(\forall x \in G_{\neg q}\right) \quad |H(x, y) - F(y)| < \varepsilon; \tag{9.4.40}$$

hence $|F|<\infty$ on B.

If here $q=\infty$, the deleted globe $G_{\neg q}$ has the form (b,∞). Thus if

$$H(x,u)=\int_a^x f(t,u)dt, \tag{9.4.41}$$

(6) turns into (2) as a special case. If (6) holds with " $(\exists G_{\neg q})$ " and " $(\forall y\in B)$ " interchanged, as in (1), convergence is pointwise only.

As in Chapter 8, §8, we denote by $f(\cdot,y)$, or f^y, the function of x alone (on E^1) given by

$$f^y(x)=f(x,y). \tag{9.4.42}$$

Similarly,

$$f_x(y)=f(x,y). \tag{9.4.43}$$

Of course, we may replace $f(x,y)$ by $f(t,u)$ or $H(t,u)$, etc.

We use Lebesgue measure in Theorems 2 and 3 below.

Theorem 9.4.2

Assume $f,g:E^2\rightarrow E^1$ satisfy

(i) $C\int_a^\infty g(t,u)dt$ converges (uniformly) on B;

(ii) each $g^u(u\in B)$ is L-measurable on $A=[a,\infty)$;

(iii) each $f^u(u\in B)$ is monotone ($\downarrow$ or $\uparrow$) on A; and

(iv) $|f|<K\in E^1$ (bounded) on $A\times B$.

Then

$$C\int_a^\infty f(t,u)g(t,u)dt \tag{9.4.44}$$

converges uniformly on B.

Proof

Given $\varepsilon>0$, use assumption (i) and Theorem 1 to choose $b>a$ so that

$$\left|L\int_v^x g(t,u)dt\right|<\frac{\varepsilon}{2K}, \tag{9.4.45}$$

written briefly as

$$\left|L\int_v^x g^u\right|<\frac{\varepsilon}{2K}, \tag{9.4.46}$$

for all $u\in B$ and $x>v\geq b$, with K as in (iv).

Hence by (ii), each $g^u(u\in B)$ is L-integrable on any interval $[v,x]\subset A$, with $x>v\geq b$. Thus given such u and $[v,x]$, we can use (iii) and Corollary 5 from §1 to find that

$$L\int_v^x f^u g^u=f^u(v)L\int_v^c g^u+f^u(x)L\int_c^x g^u \tag{9.4.47}$$

for some $c\in[v,x]$.

Combining with (7) and using (iv), we easily obtain

$$\left|L\int_v^x f(t,u)g(t,u)dt\right|<\varepsilon \tag{9.4.48}$$

whenever $u \in B$ and $x > v \geq b$. (Verify!)

Our assertion now follows by Theorem 1. □

Theorem 9.4.3 (Abel-Dirichlet test)

Let $f, g : E^2 \rightarrow E^*$ satisfy

(a) $\lim_{t\rightarrow\infty} f(t,u) = 0$ (uniformly) for $u \in B$;

(b) each $f^u (u \in B)$ is nonincreasing ($\downarrow$) on $A = [0, \infty)$;

(c) each $g^u (u \in B)$ is L-measurable on A; and

(d) $\left(\exists K \in E^1\right)(\forall x \in A)(\forall u \in B)\left|L \int_a^x g(t,u) dt\right| < K$.

Then

$$C \int_a^\infty f(t,u) g(t,u) dt \tag{9.4.49}$$

converges uniformly on B.

Proof Outline

Argue as in Problem 13 of §3, replacing Theorem 2 in §3 by Theorem 1 of the present section.

By Lemma 2 in §1, obtain

$$\left|L \int_v^x f^u g^u\right| = \left|f^u(v) L \int_a^x g^u\right| \leq K f(v,u) \tag{9.4.50}$$

for $u \in B$ and $x > v \geq a$.

Then use assumption (a) to fix k so that

$$|f(t,u)| < \frac{\varepsilon}{2K} \tag{9.4.51}$$

for $t > k$ and $u \in B$. □

Note 2. Via components, Theorems 2 and 3 extend to the case $g : E^2 \rightarrow E^n \left(C^n\right)$.

Note 3. While Corollaries 2 and 3 apply to absolute convergence only, Theorems 2 and 3 cover conditional convergence, too (a great advantage!). The theorems also apply if f or g is independent of u (see Note 1). This supersedes Problems 13 and 14 in §3.

Examples

(A) The integral

$$\int_0^\infty \frac{\sin tu}{t} dt \tag{9.4.52}$$

converges uniformly on $B_\delta = [\delta, \infty)$ if $\delta > 0$, and pointwise on $B = [0, \infty)$.

Indeed, we can use Theorem 3, with

$$g(t,u) = \sin tu \tag{9.4.53}$$

and

$$f(t,u) = \frac{1}{t}, f(0,u) = 1, \tag{9.4.54}$$

say. Then the limit

$$\lim_{t\to\infty}\frac{1}{t}=0 \tag{9.4.55}$$

is trivially uniform for $u\in B_\delta$, as f is independent of u. Thus assumption (a) is satisfied. So is (d) because

$$\left|\int_0^x \sin tu\,dt\right|=\left|\frac{1}{u}\int_0^{xu}\sin\theta\,d\theta\right|\leq\frac{1}{\delta}\cdot 2. \tag{9.4.56}$$

(Explain!) The rest is easy.

Note that Theorem 2 fails here since assumption (i) is not satisfied.

(B) The integral

$$\int_0^\infty \frac{1}{t}e^{-tu}\sin at\,dt \tag{9.4.57}$$

converges uniformly on $B=[0,\infty)$. It does so absolutely on $B_\delta=[\delta,\infty)$, if $\delta>0$.

Here we shall use Theorem 2 (though Theorem 3 works, too). Set

$$f(t,u)=e^{-tu} \tag{9.4.58}$$

and

$$g(t,u)=\frac{\sin at}{t}, g(0,u)=a. \tag{9.4.59}$$

Then

$$\int_0^\infty g(t,u)\,dt \tag{9.4.60}$$

converges (substitute $x=at$ in Problem 8 or 15 in §3). Convergence is trivially uniform, by Note 1. Thus assumption (i) holds, and so do the other assumptions. Hence the result.

For absolute convergence on B_δ, use Corollary 3 with

$$M(t)=e^{-\delta t}, \tag{9.4.61}$$

so $M\geq|fg|$.

Note that, quite similarly, one treats C-integrals of the form

$$\int_a^\infty e^{-tu}g(t)\,dt, \int_a^\infty e^{-t^2u}g(t)\,dt, \text{ etc.,} \tag{9.4.62}$$

provided

$$\int_a^\infty g(t)\,dt \tag{9.4.63}$$

converges $(a\geq 0)$.

In fact, Theorem 2 states (roughly) that the uniform convergence of $C\int g$ implies that of $C\int fg$, provided f is monotone (in t) and bounded.

III. We conclude with some theorems on uniform convergence of functions $H:E^2\to E$ (see (6)). In Theorem 4, m is again an LS (or Lebesgue) measure in E^1; the deleted globe $G^*_{\neg q}$ is fixed.

Theorem 9.4.4

Suppose

$$\lim_{x\to q}H(x,y)=F(y)\text{ (uniformly)} \tag{9.4.64}$$

for $y \in B \subseteq E^{1}$. Then we have the following:

(i) If all H_x $(x \in G^*_{\neg q})$ are continuous or m-measurable on B, so also is F.

(ii) The same applies to m-integrability on B, provided $mB < \infty$; and then

$$\lim_{x \rightarrow q} \int_{B} |H_x - F| = 0; \tag{9.4.65}$$

hence

$$\lim_{x \rightarrow q} \int_{B} H_x = \int_{B} F = \int_{B} \left(\lim_{x \rightarrow q} H_x \right). \tag{9.4.66}$$

Formula (8') is known as the rule of passage to the limit under the integral sign.

Proof

(i) Fix a sequence $x_k \rightarrow q$ $(x_k$ in the deleted globe $G^*_{\neg q})$, and set

$$H_k = H_{x_k} \quad (k = 1, 2, \ldots). \tag{9.4.67}$$

The uniform convergence

$$H(x, y) \rightarrow F(y) \tag{9.4.68}$$

is preserved as x runs over that sequence (see Problem 4). Hence if all H_k are continuous or measurable, so is F (Theorem 2 in Chapter 4, §12 and Theorem 4 in Chapter 8, §1. Thus clause (i) is proved.

(ii) Now let all H_x be m-integrable on B; let

$$mB < \infty. \tag{9.4.69}$$

Then the H_k are m-measurable on B, and so is F, by (i). Also, by (6),

$$(\forall \varepsilon > 0)\left(\exists G_{\neg q}\right)\left(\forall x \in G_{\neg q}\right) \quad \int_{B} |H_x - F| \leq \int_{B} (\varepsilon) = \varepsilon m B < \infty, \tag{9.4.70}$$

proving (8). Moreover, as

$$\int_{B} |H_x - F| < \infty, \tag{9.4.71}$$

$H_x - F$ is m-integrable on B, and so is

$$F = H_x - (H_x - F). \tag{9.4.72}$$

Hence

$$\left|\int_{B} H_x - \int_{B} F\right| = \left|\int_{B} (H_x - F)\right| \leq \int_{B} |H_x - F| \rightarrow 0, \tag{9.4.73}$$

as $x \rightarrow q$, by (8). Thus (8') is proved, too. $\square$

Quite similarly (keeping E complete and using sequences), we obtain the following result.

Theorem 9.4.5

Suppose that

(i) all H_x $\left(x \in G^*_{\neg q}\right)$ are continuous and finite on a finite interval $B \subset E^{1}$, and differentiable on $B - Q$, for a fixed countable set Q;

(ii) $\lim_{x \rightarrow q} H\left(x, y_0\right) \neq \pm \infty$ exists for some $y_0 \in B$; and

(iii) $\lim_{x \rightarrow q} D_2 H(x, y) = f(y)$ (uniformly) exists on $B - Q$.

Then f, so defined, has a primitive F on B, exact on $B-Q$ (so $F'=f$ on $B-Q$); moreover,

$$F(y)=\lim_{x\rightarrow y} H(x,y) \text{ (uniformly) for } y\in B. \tag{9.4.74}$$

Outline of Proof

Note that

$$D_2 H(x,y)=\frac{d}{dy} H_x(y). \tag{9.4.75}$$

Use Theorem 1 of Chapter 5, §9, with $F_n=H_{x_n}, x_n\rightarrow q$. $\square$

Note 4. If $x\rightarrow q$ over a path P (clustering at q), one must replace $G_{\neg q}$ and $G^*_{\neg q}$ by $P\cap G_{\neg q}$ and $P\cap G^*_{\neg q}$ in (6) and in Theorems 4 and 5.

9.4.E: Problems on Uniform Convergence of Functions and C-Integrals

Exercise $9.4.E.1$

Fill in all proof details in Theorems $1-5$, Corollaries 4 and 5, and examples (A) and (B).

Exercise $9.4.E.1'$

Using (6), prove that

$$\lim_{x \rightarrow q} H(x, y) \text{ (uniformly)} \tag{9.4.E.1}$$

exists on $B \subseteq E^{1}$ iff

$$(\forall \varepsilon>0)\left(\exists G_{\neg q}\right)(\forall y \in B)\left(\forall x, x^{\prime} \in G_{\neg q}\right) \quad\left|H(x, y)-H\left(x^{\prime}, y\right)\right|<\varepsilon. \tag{9.4.E.2}$$

Assume E complete and $|H|<\infty$ on $G_{\neg q} \times B$.
[Hint: "Imitate" the proof of Theorem 1, using Theorem 2 of Chapter 4, §2.]

Exercise $9.4.E.2$

State formulas analogous to (1) and (2) for $\int_{-\infty}^{a}, \int_{a}^{b-}$, and $\int_{a+}^{b}$.

Exercise $9.4.E.3$

State and prove Theorems 1 to 3 and Corollaries 1 to 3 for

$$\int_{-\infty}^{a}, \int_{a}^{b-}, \text{ and } \int_{a+}^{b}. \tag{9.4.E.3}$$

In Theorems 2 and 3 explore absolute convergence for

$$\int_{a}^{b-} \text{ and } \int_{a+}^{b}. \tag{9.4.E.4}$$

Do at least some of the cases involved.
[Hint: Use Theorem 1 of §3 and Problem 1', if already solved.]

Exercise $9.4.E.4$

Prove that

$$\lim_{x \rightarrow q} H(x, y)=F(y) \text{ (uniformly)} \tag{9.4.E.5}$$

on B iff

$$\lim_{n \rightarrow \infty} H\left(x_{n}, \cdot\right)=F(\text{ uniformly }) \tag{9.4.E.6}$$

on B for all sequences $x_{n} \rightarrow q\left(x_{n} \neq q\right)$.
[Hint: "Imitate" Theorem 1 in Chapter 4, §2. Use Definition 1 of Chapter 4, §12.]

Exercise 9.4.E.5

Prove that if

$$\lim_{x \rightarrow q} H(x, y)=F(y) \text { (uniformly) } \tag{9.4.E.7}$$

on A and on B, then this convergence holds on $A \cup B$. Hence deduce similar propositions on C-integrals.

Exercise 9.4.E.6

Show that the integrals listed below violate Corollary 4 and hence do not converge uniformly on $P=(0, \delta)$ though proper L-integrals exist for each $u \in P$. Thus show that Theorem 1 (ii) does not apply to uniform convergence.

(a) $\int_{0+}^{1} \frac{u d t}{t^{2}-u^{2}}$;

(b) $\int_{0+}^{1} \frac{u^{2}-t^{2}}{\left(t^{2}+u^{2}\right)^{2}} d t$;

(c) $\int_{0+}^{1} \frac{t u\left(t^{2}-u^{2}\right)}{\left(t^{2}+u^{2}\right)^{2}} d t$.

$[$ Hint for (b) : To disprove uniform convergence, fix any $\varepsilon, v>0$. Then

$$\int_{0}^{v} \frac{u^{2}-t^{2}}{\left(t^{2}+u^{2}\right)^{2}} d t=\frac{v}{v^{2}+u^{2}} \rightarrow \frac{1}{v} \tag{9.4.E.8}$$

as $u \rightarrow 0$. Thus if $v<\frac{1}{2 \varepsilon}$,

$$(\exists u \in P) \quad \int_{0}^{v} \frac{u^{2}-t^{2}}{\left(t^{2}+u^{2}\right)^{2}} d t>\frac{1}{2 v}>\varepsilon .\left.\right] \tag{9.4.E.9}$$

Exercise 9.4.E.7

Using Corollaries 3 to 5, show that the following integrals converge (uniformly) on U (as listed) but only pointwise on P (for the latter, proceed as in Problem 6). Specify P and $M(t)$ in each case where they are not given.

(a) $\int_{0}^{\infty} e^{-u t^{2}} d t ; U=[\delta, \infty) ; P=(0, \delta)$.

$\left[\right.$ Hint: Set $M(t)=e^{-\delta t}$ for $t \geq 1$ (Corollaries 3 and 5$\left.)\right.$.]

(b) $\int_{0}^{\infty} e^{-u t} t^{a} \cos t d t(a \geq 0) ; U=[\delta, \infty)$.

(c) $\int_{0+}^{1} t^{u-1} d t ; U=[\delta, \infty)$.

(d) $\int_{0+}^{1} t^{-u} \sin t d t ; U=[0, \delta], 0<\delta<2 ; P=[\delta, 2) ; M(t)=t^{1-\delta}$.

[Hint: Fix v so small that

$$(\forall t \in(0, v)) \quad \frac{\sin t}{t}>\frac{1}{2} . \tag{9.4.E.10}$$

Then, if $u \rightarrow 2$,

$$\int_{0}^{v} t^{-u} \sin t d t \geq \frac{1}{2} \int_{0}^{v} \frac{d t}{t^{u-1}} \rightarrow \infty .] \tag{9.4.E.11}$$

Exercise 9.4.E.8

In example (A), disprove uniform convergence on $P=(0, \infty)$.

[Hint: Proceed as in Problem 6.]

Exercise 9.4.E.9

Do example (B) using Theorem 3 and Corollary 5. Disprove uniform convergence on B.

Exercise 9.4.E.10

Show that

$$\int_{0+}^{\infty} \frac{\sin tu}{t} \cos t dt \tag{9.4.E.12}$$

converges uniformly on any closed interval U, with $\pm 1 \notin U$.
[Hint: Transform into

$$\left.\frac{1}{2} \int_{0+}^{\infty} \frac{1}{t}\{\sin[(u+1)t]+\sin[(u-1)t]\} dt.\right] \tag{9.4.E.13}$$

Exercise 9.4.E.11

Show that

$$\int_{0}^{\infty} t \sin t^{3} \sin tu dt \tag{9.4.E.14}$$

converges (uniformly) on any finite interval U.
[Hint: Integrate

$$\int_{x}^{y} t \sin t^{3} \sin tu dt \tag{9.4.E.15}$$

by parts twice. Then let $y \rightarrow \infty$ and $x \rightarrow 0$.]

Exercise 9.4.E.12

Show that

$$\int_{0+}^{\infty} e^{-tu} \frac{\cos t}{t^{a}} dt \quad (0<a<1) \tag{9.4.E.16}$$

converges (uniformly) for $u \geq 0$.
[Hints: For $t \rightarrow 0+,$ use $M(t)=t^{-a}.$ For $t \rightarrow \infty,$ use example (B) and Theorem 2.]

Exercise 9.4.E.13

Prove that

$$\int_{0+}^{\infty} \frac{\cos tu}{t^{a}} dt \quad (0<a<1) \tag{9.4.E.17}$$

converges (uniformly) for $u \geq \delta>0$, but (pointwise) for $u>0$.
[Hint: Use Theorem 3 with $g(t, u)=\cos tu$ and

$$\left|\int_{0}^{x} g\right|=\left|\frac{\sin xu}{u}\right| \leq \frac{1}{\delta}. \tag{9.4.E.18}$$

For $u > 0$,

$$\int_v^\infty \frac{\cos tu}{t^a} dt = u^{a-1} \int_{vu}^\infty \frac{\cos z}{z} dz \rightarrow \infty \tag{9.4.E.19}$$

if $v = 1/u$ and $u \rightarrow 0$. Use Corollary 4.]

Exercise $9.4.E.14$

Given $A, B \subseteq E^1 (mA < \infty)$ and $f : E^2 \rightarrow E$, suppose that
(i) each $f(x, \cdot) = f_x (x \in A)$ is relatively (or uniformly) continuous on B; and
(ii) each $f(\cdot, y) = f^y (y \in B)$ is m-integrable on A.
Set

$$F(y) = \int_A f(x, y) dm(x), \quad y \in B. \tag{9.4.E.20}$$

Then show that F is relatively (or uniformly) continuous on B.
[Hint: We have

$$\begin{aligned} &(\forall x \in A)(\forall \varepsilon > 0) (\forall y_0 \in B)(\exists \delta > 0) \left(\forall y \in B \cap G_{y_0}(\delta)\right) \\ &\qquad |F(y) - F(y_0)| \leq \int_A |f(x, y) - f(x, y_0)| \, dm(x) \leq \int_A \left(\frac{\varepsilon}{mA}\right) dm = \varepsilon. \end{aligned}$$

Similarly for uniform continuity.]

Exercise $9.4.E.15$

Suppose that
(a) $C \int_a^\infty f(t, y) dm(t) = F(y)$(uniformly) on $B = [b, d] \subseteq E^1$;
(b) each $f(x, \cdot) = f_x (x \geq a)$ is relatively continuous on B; and
(c) each $f(\cdot, y) = f^y (y \in B)$ is m -integrable on every $[a, x] \subset E^1$, $x \geq a$.
Then show that F is relatively continuous, hence integrable, on B and that

$$\int_B F = \lim_{x \rightarrow \infty} \int_B H_x, \tag{9.4.E.21}$$

where

$$H(x, y) = \int_a^x f(t, y) dm(t). \tag{9.4.E.22}$$

(Passage to the limit under the $\int$-sign.)
[Hint: Use Problem 14 and Theorem 4; note that

$$C \int_0^\infty f(t, y) dm(t) = \lim_{x \rightarrow \infty} H(x, y) (\text{uniformly}).] \tag{9.4.E.23}$$

Index

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Book: Mathematical Analysis (Zakon)

Webpages: 185

All licenses found:

- CC BY 1.0: 51.9% (96 pages)
- CC BY 3.0: 46.5% (86 pages)
- Undeclared: 1.6% (3 pages)

By Page