Terms and Conditions

You may download, print, transfer, or copy this work, either electronically or mechanically, only under the following conditions.

If you are a student using this work for self-study, no payment is required.
If you are a teacher evaluating this work for use as a required or recommended text in a course, no payment is required.

Payment is required for any and all other uses of this work. In particular, but not exclusively, payment is required if
(1) you are a student and this is a required or recommended text for a course;
(2) you are a teacher and you are using this book as a reference, or as a required or recommended text, for a course.

Payment is made through the website http://www.trillia.com. For each individual using this book, payment of US$10 is required. A sitewide payment of US$300 allows the use of this book in perpetuity by all teachers, students, or employees of a single school or company at all sites that can be contained in a circle centered at the location of payment with a radius of 25 miles (40 kilometers). You may post this work to your own website or other server (FTP, etc.) only if a sitewide payment has been made and it is noted on your website (or other server) precisely which people have the right to download this work according to these terms and conditions.

Any copy you make of this work, by any means, in whole or in part, must contain this page, verbatim and in its entirety.

Mathematical Analysis II
© 1977 Elias Zakon
© 2009 Bradley J. Lucier and Tamara Zakon
ISBN 978-1-931705-03-8
Published by The Trillia Group, West Lafayette, Indiana, USA

The phrase “The Trillia Group” and the Trillia Group logo are trademarks of The Trillia Group and may not be used without permission.

This book was prepared by Bradley J. Lucier and Tamara Zakon from a manuscript prepared by Elias Zakon. We intend to correct and update this work as needed. If you notice any mistakes in this work, please send e-mail to lucier@math.purdue.edu and they will be corrected in a later version.

Contents

Preface ix
About the Author xi

Chapter 6. Differentiation on \(E^n\) and Other Normed Linear Spaces 1
1. Directional and Partial Derivatives ........................................ 1
   Problems on Directional and Partial Derivatives ....................... 6
2. Linear Maps and Functionals. Matrices ................................ 7
   Problems on Linear Maps and Matrices ............................... 14
3. Differentiable Functions ...................................................... 16
   Problems on Differentiable Functions .................................. 25
4. The Chain Rule. The Cauchy Invariant Rule ........................... 28
   Further Problems on Differentiable Functions ....................... 33
5. Repeated Differentiation. Taylor’s Theorem ........................... 35
   Problems on Repeated Differentiation and Taylor Expansions ...... 44
   Problems on Bijective Linear Maps and Jacobians ................. 55
7. Inverse and Implicit Functions. Open and Closed Maps .......... 57
   Problems on Inverse and Implicit Functions, Open and Closed Maps ................................................................. 67
8. Baire Categories. More on Linear Maps ................................. 70
   Problems on Baire Categories and Linear Maps ..................... 76
9. Local Extrema. Maxima and Minima ................................. 79
   Problems on Maxima and Minima ..................................... 84
10. More on Implicit Differentiation. Conditional Extrema ........... 87
    Further Problems on Maxima and Minima ............................ 94

Chapter 7. Volume and Measure 97
1. More on Intervals in \(E^n\), Semirings of Sets ................................ 97
   Problems on Intervals and Semirings ................................... 104
2. \(C^\alpha\)-Sets. Countable Additivity. Permutable Series ............ 104
   Problems on \(C^\alpha\)-Sets, \(\sigma\)-Additivity, and Permutable Series ...... 112

* “Starred” sections may be omitted by beginners.
Preface

This is a multipurpose text. When taken in full, including the “starred” sections, it is a graduate course covering differentiation on normed spaces and integration with respect to complex and vector-valued measures. The starred sections may be omitted without loss of continuity, however, for a junior or senior course. One also has the option of limiting all to $E^n$, or taking Riemann integration before Lebesgue theory (we call it the “limited approach”). The proofs and definitions are so chosen that they are as simple in the general case as in the more special cases. In a nutshell, the basic ideas of measure theory are given in Chapter 7, §§1 and 2. Not much more is needed for the “limited approach.”

In Chapter 6 (Differentiation), we have endeavored to present a modern theory, without losing contact with the classical terminology and notation. (Otherwise, the student is unable to read classical texts after having been taught the “elegant” modern theory.) This is why we prefer to define derivatives as in classical analysis, i.e., as numbers or vectors, not as linear mappings. The latter are used to define a modern version of differentials.

In Chapter 9, we single out those calculus topics (e.g., improper integrals) that are best treated in the context of Lebesgue theory.

Our principle is to keep the exposition more general whenever the general case can be handled as simply as the special ones (the degree of the desired specialization is left to the instructor). Often this even simplifies matters—for example, by considering normed spaces instead of $E^n$ only, one avoids cumbersome coordinate techniques. Doing so also makes the text more flexible.

Publisher’s Notes

Text passages in blue are hyperlinks to other parts of the text.
Several annotations are used throughout this book:
* This symbol marks material that can be omitted at first reading.
⇒ This symbol marks exercises that are of particular importance.
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 Zakon 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 Honours 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 Erdős, who was then banned from the United States for his political views. Erdős 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 last of these volumes, which was used in a two-semester class required of all Honours Mathematics students at Windsor.
Chapter 6
Differentiation on $E^n$ and Other
Normed Linear Spaces

§1. Directional and Partial Derivatives

In Chapter 5 we considered functions $f : E^1 \to E$ of one real variable.

Now we take up functions $f : E' \to E$ where both $E'$ and $E$ are normed spaces.\textsuperscript{1}

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

$$\vec{x} = \vec{p} + t\vec{u},$$

with $t \in E^1$ and $\vec{u}$ a unit vector. For if $\vec{x} \neq \vec{p}$, set

$$t = |\vec{x} - \vec{p}|$$

and if $\vec{x} = \vec{p}$, set $t = 0$, and any $\vec{u}$ will do. We often use the notation

$$\vec{t} = \Delta \vec{x} = \vec{x} - \vec{p} = t\vec{u} \quad (t \in E^1, \vec{t}, \vec{u} \in E').$$

First of all, we generalize Definition 1 in Chapter 5, §1.

Definition 1.

Given $f : E' \to E$ and $\vec{p}, \vec{u} \in E'$ ($\vec{u} \neq \vec{0})$, we define the directional derivative of $f$ along $\vec{u}$ (or $\vec{u}$-directed derivative of $f$) at $\vec{p}$ by

$$D_{\vec{u}} f(\vec{p}) = \lim_{t \to 0} \frac{1}{t}[f(\vec{p} + t\vec{u}) - f(\vec{p})],$$

\textsuperscript{1} We now presuppose §§9–12 of Chapter 3, including the “starred” parts.
Similarly in the case \( 2 \neq 0 \), this may fail even if \( 2 \) and \( 1 \) are constants.

For functions \( f \) such that \( f(x) = x^2 \), we can compute \( D_k f \) by the usual rules of differentiation, treating \( x_k \) as a variable.

Thus, the limit (1) is called the partial derivative of \( f \) at \( \bar{p} \), with respect to \( x_k \), denoted

\[
D_k f(\bar{p}), \quad \frac{\partial f}{\partial x_k}(\bar{p}), \quad \text{or } \frac{\partial f}{\partial x_k}(x_k = \bar{p}).
\]

If it exists for all \( \bar{p} \in B \), we call \( D_k f \) the partial derivative (briefly, partial) of \( f \) on \( B \), with respect to \( x_k \).

In any case, the derived functions \( D_k f \) \((k = 1, \ldots, n)\) are always defined on all of \( E^n \) (\( C^n \)).

If \( E' = E^1 \) \((C^1)\), we often write \( x, y, z \) for \( x_1, x_2, x_3 \), and

\[
\frac{\partial f}{\partial x_1}, \ \frac{\partial f}{\partial y}, \ \frac{\partial f}{\partial z} \text{ for } D_k f \quad (k = 1, 2, 3).^2
\]

**Note 1.** If \( E' = E^1 \), scalars are also “vectors,” and \( D_k f \) coincides with \( f' \) as defined in Chapter 5, §1 (except where \( f' = \pm \infty \)). Explain!

**Note 2.** As we have observed, the \( \bar{u} \)-directed derivative (1) is obtained by keeping \( \bar{x} \) on the line \( \bar{x} = \bar{p} + t\bar{u} \).

If \( \bar{u} = \bar{e}_k \), the line is parallel to the \( k \)th axis; so all coordinates of \( \bar{x} \), except \( x_k \), remain fixed \((x_i = p_i, i \neq k)\), and \( f \) behaves like a function of one variable, \( x_k \).

Thus we can compute \( D_k f \) 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 \( f(x, y) = x^2 y \). Then

\[
\frac{\partial f}{\partial x} = D_1 f(x, y) = 2xy \quad \text{and} \quad \frac{\partial f}{\partial y} = D_2 f(x, y) = x^2.
\]

**Note 3.** More generally, given \( \bar{p} \) and \( \bar{u} \neq \bar{0} \), set \( h(t) = f(\bar{p} + t\bar{u}), \quad t \in E^1 \).

Then \( h(0) = f(\bar{p}) \); so

\[
D_\bar{u} f(\bar{p}) = \lim_{t \to 0} \frac{h(t) - h(0)}{t - 0} = h'(0)
\]

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 \( D_\bar{u} f(\bar{p}) \) exists for all \( \bar{u} \neq \bar{0} \).

---

2 Similarly in the case \( E' = E^2 \) \((C^2)\).
Examples.
(a) Define \( f: E^2 \to E^1 \) by
\[
f(x, y) = \frac{x^2y}{x^4 + y^2}, \quad f(0, 0) = 0.
\]
Fix a unit vector \( \mathbf{u} = (u_1, u_2) \) in \( E^2 \). Let \( \mathbf{p} = (0, 0) \). To find \( D_2 f(p) \), use the \( h \) of Note 3:
\[
h(t) = f(p + t\mathbf{u}) = f(tu_1, tu_2) = \frac{t u_1^2 u_2}{t^4 u_1^4 + u_2^2} \quad \text{if} \quad u_2 \neq 0,
\]
and \( h = 0 \) if \( u_2 = 0 \). Hence
\[
D_2 f(p) = h'(0) = \frac{u_1^2}{u_2} \quad \text{if} \quad u_2 \neq 0,
\]
and \( h'(0) = 0 \) if \( u_2 = 0 \). Thus \( D_2 f(0) \) exists for all \( \mathbf{u} \). Yet \( f \) is discontinuous at \( 0 \) (see Problem 9 in Chapter 4, §3).
(b) Let
\[
f(x, y) = \begin{cases} 
x + y & \text{if} \ xy = 0, \\
1 & \text{otherwise}.
\end{cases}
\]
Then \( f(x, y) = x \) on the x-axis; so \( D_2 f(0, 0) = 1 \).

Similarly, \( D_2 f(0, 0) = 1 \). Thus both partials exist at \( 0 \).

Yet \( f \) is discontinuous at \( 0 \) (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 \( 0 \) 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 1. If \( f: E' \to E \) has a \( \mathbf{u} \)-directed derivative at \( p \in E' \), then \( f \) is relatively continuous at \( p \) over the line \( \bar{x} = \mathbf{p} + t\mathbf{u} \) \( (\mathbf{0} \neq \mathbf{u} \in \mathbb{R}^1) \).

Proof. Set \( h(t) = f(\bar{p} + t\mathbf{u}), \ t \in E^1 \).

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
\[
\lim_{t \to 0} h(t) = h(0) = f(\bar{p}),
\]
i.e.,
\[
\lim_{t \to 0} f(\bar{p} + t\mathbf{u}) = f(\bar{p}).
\]
But this means that \( f(\bar{x}) \to f(\bar{p}) \) as \( \bar{x} \to \bar{p} \) over the line \( \bar{x} = \mathbf{p} + t\mathbf{u} \), for, on that line, \( \bar{x} = \mathbf{p} + t\mathbf{u} \).

Thus, indeed, \( f \) is relatively continuous at \( \bar{p} \), as stated. \( \square \)

Note that we actually used the substitution \( \bar{x} = \mathbf{p} + t\mathbf{u} \). This is admissible since the dependence between \( x \) and \( t \) is one-to-one (Corollary 2(iii) of Chapter 4, §2). Why?

\( \textbf{Theorem 2.} \) Let \( E' \ni \mathbf{u} = \mathbf{q} - \mathbf{p}, \ \mathbf{u} \neq \mathbf{0} \).

If \( f: E' \to E \) is relatively continuous on the segment \( I = L[\mathbf{p}, \mathbf{q}] \) and has a \( \mathbf{u} \)-directed derivative on \( I - Q \) (\( Q \) countable), then
\[
|f(\mathbf{q}) - f(\mathbf{p})| \leq \sup_{x \in I - Q} |D_2 f(x)|, \quad \bar{x} \in I - Q.
\]

\( \textbf{Proof.} \) Set again \( h(t) = f(\bar{p} + t\mathbf{u}) \) and \( g(t) = \mathbf{p} + t\mathbf{u} \).

Then \( h = f \circ g \) and \( g \) is continuous on \( E^1 \). (Why?)

As \( f \) is relatively continuous on \( I = L[\mathbf{p}, \mathbf{q}] \), 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 \( \bar{x}_0 = \mathbf{p} + t_0\mathbf{u} \in I - Q \), our assumptions imply the existence of
\[
D_2 f(\bar{x}_0) = \lim_{t \to 0} \frac{f(\bar{x}_0 + t\mathbf{u}) - f(\bar{x}_0)}{t} = \lim_{t \to 0} \frac{f(\mathbf{p} + t_0\mathbf{u} + t\mathbf{u}) - f(\mathbf{p} + t_0\mathbf{u})}{t} = \lim_{t \to 0} \frac{h(t_0 + t) - h(t_0)}{t} = h'(t_0). \ (\text{Explain!})
\]

This can fail for at most a countable set \( Q' \) of points \( t_0 \in J \) (those for which \( \bar{x}_0 \in Q \)).

Thus \( h \) is differentiable on \( J - Q' \); and so, by Corollary 1 in Chapter 5, §4,
\[
|h(1) - h(0)| \leq \sup_{t \in J - Q'} |h'(t)| = \sup_{x \in I - Q} |D_2 f(x)|.
\]

Now as \( h(1) = f(\mathbf{p} + \mathbf{u}) = f(\mathbf{q}) \) and \( h(0) = f(\mathbf{p}) \), formula (2) follows. \( \square \)

\( \textbf{Theorem 3.} \) If in Theorem 2, \( E = E^1 \) and if \( f \) has a \( \mathbf{u} \)-directed derivative at least on the open line segment \( L[\mathbf{p}, \mathbf{q}] \), then
\[
|f(\mathbf{q}) - f(\mathbf{p})| = D_2 f(\bar{x}_0)
\]
for some \( \bar{x}_0 \in L[\mathbf{p}, \mathbf{q}] \).
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 2.** If in Theorems 2 and 3, we set
\[ r = |\vec{a}| = |\vec{q} - \vec{p}| \neq 0 \quad \text{and} \quad \vec{v} = \frac{1}{r} \vec{u}, \]
then formulas (2) and (3) can be written as
\[ |f(\vec{q}) - f(\vec{p})| \leq |\vec{q} - \vec{p}| \sup |D_\vec{v} f(\vec{x})|, \quad \vec{x} \in I - Q, \]
and
\[ f(\vec{q}) - f(\vec{p}) = |\vec{q} - \vec{p}| D_\vec{v} f(\vec{x}_0) \]
for some \( \vec{x}_0 \in L(\vec{p}, \vec{q}) \).

For by Corollary 1,
\[ D_\vec{v} f = r D_\vec{e} f = |\vec{q} - \vec{p}| D_\vec{e} f; \]
so (2') and (3') follow.

### Problems on Directional and Partial Derivatives

1. Complete all missing details in the proof of Theorems 1 to 3 and Corollaries 1 and 2.
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.
3. In Examples (a) and (b) describe \( D_\vec{v} f : E^2 \to E^1 \). Compute it for \( \vec{u} = (1, 1) = \vec{p} \).
   In (b), show that \( f \) has no directional derivatives \( D_\vec{v} f(\vec{p}) \) except if \( \vec{u} \parallel \vec{e}_1 \) or \( \vec{u} \parallel \vec{e}_2 \). Give two proofs: (i) use Theorem 1; (ii) use definitions only.
4. Prove that if \( f : E^n (C^n) \to E \) has a zero partial derivative, \( D_\vec{u} f = 0 \), 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.)
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 \( \overline{0} \), given that \( f(x, y) \) equals:
   (i) \( \frac{xy}{x^2 + y^2} \); (ii) the integral part of \( x + y \);
   (iii) \( \frac{xy}{|x|} + x \sin \frac{1}{y} \); (iv) \( \frac{x^2 - y^2}{x^2 + y^2} \);
   (v) \( \sin(y \cos x) \); (vi) \( x^y \).

### §1. Directional and Partial Derivatives

(\( \text{Set } f = 0 \text{ wherever the formula makes no sense.} \))

**⇒ 6.** Prove that if \( f : E' \to E^1 \) has a local maximum or minimum at \( \vec{p} \in E' \), then \( D_\vec{v} f(\vec{p}) = 0 \) for every vector \( \vec{u} \neq 0 \) in \( E' \).

[Hint: Use Note 3, then Corollary 1 in Chapter 5, §2.]

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, \( \varphi \) and \( \psi \).]

8. State and prove Theorems 4 and 5 of Chapter 5, §1, for directional derivatives.

### §2. Linear Maps and Functionals. Matrices

For an adequate definition of differentiability, we need the notion of a linear map. Below, \( E', E'', E^n \), and \( E \) denote normed spaces over the same scalar field, \( E^1 \) or \( C \).

**Definition 1.**

A function \( f : E' \to E \) is a linear map if and only if for all \( \vec{x}, \vec{y} \in E' \) and scalars \( a, b \)
\[
(1) \quad f(a\vec{x} + b\vec{y}) = af(\vec{x}) + bf(\vec{y});
\]
equivalently, iff for all such \( \vec{x}, \vec{y}, \) and \( a \)
\[
(2) \quad f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) \quad \text{and} \quad f(a\vec{x}) = af(\vec{x}). \quad \text{(Verify!)}
\]

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 map \( f \) is also called a (real or complex) linear functional on \( E' \).

**Note 1.** Induction extends formula (1) to any “linear combinations”:
\[
(3) \quad f\left(\sum_{i=1}^{m} a_i \vec{x}_i\right) = \sum_{i=1}^{m} a_i f(\vec{x}_i)
\]
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 \( f(\overline{0}) = 0 \) if \( f \) is linear.

**Examples.**
(a) Let \( E' = E^n (C^n) \). Fix a vector \( \vec{v} = (v_1, \ldots, v_n) \) in \( E' \) and set
\[
(\forall \vec{x} \in E') \quad f(\vec{x}) = \vec{x} \cdot \vec{v}
\]
This is the so-called uniform Lipschitz condition.

Then
\[ f(a\vec{x} + b\vec{y}) = (a\vec{x}) \cdot \vec{v} + (b\vec{y}) \cdot \vec{v} = a(\vec{x} \cdot \vec{v}) + b(\vec{y} \cdot \vec{v}) = af(\vec{x}) + bf(\vec{y}); \]
so \( f \) is linear. Note that if \( E' = E^n \), then by definition,
\[ f(\vec{x}) = \vec{x} \cdot \vec{v} = \sum_{k=1}^{n} x_k v_k = \sum_{k=1}^{n} v_k x_k. \]
If, however, \( E' = C^n \), then
\[ f(\vec{x}) = \vec{x} \cdot \vec{v} = \sum_{k=1}^{n} x_k v_k = \sum_{k=1}^{n} v_k x_k, \]
where \( v_k \) 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 \( u: I \to E \) that are of class \( CD^\infty \) (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
\[ \|u\| = \sup_{x \in I} |u(x)|. \]
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 \( CD^\infty \), so is \( u' \).

Thus \( D(u) = u' \in E' \), and so \( D: E' \to E' \) 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 \( u: I \to E \) 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 \( \phi: E' \to E \) by
\[ \phi(u) = \int_0^1 u, \]
with \( \int u \) as in Chapter 5, §5. (Recall that \( \int_0^1 u \) is an element of \( E \) if \( u: I \to E \).) By Corollary 1 in Chapter 5, §5, \( \phi \) is a linear map of \( E' \) into

§2. Linear Maps and Functionals. Matrices

E. (Why?)

(d) The zero map \( f = 0 \) on \( E' \) is always linear. (Why?)

Theorem 1. A linear map \( f: E' \to E \) is continuous (even uniformly so) on all of \( E' \) if it is continuous at \( \vec{0} \); equivalently, if there is a real \( c > 0 \) such that
\[ (\forall \vec{x} \in E') \quad |f(\vec{x})| \leq c|\vec{x}|. \]
We call this property linear boundedness.)

Proof. Assume that \( f \) is continuous at \( \vec{0} \). Then, given \( \varepsilon > 0 \), there is \( \delta > 0 \) such that
\[ |f(\vec{x}) - f(\vec{0})| = |f(\vec{x})| \leq \varepsilon \]
whenever \( |\vec{x} - \vec{0}| = |\vec{x}| < \delta \).

Now, for any \( \vec{x} \neq \vec{0} \), we surely have
\[ \frac{\delta|\vec{x}|}{2|\vec{x}|} = \frac{\delta}{2} < \delta. \]
Hence
\[ (\forall \vec{x} \neq \vec{0}) \quad |f(\vec{x})| \leq \varepsilon, \]
or, by linearity,
\[ \frac{\delta}{2|\vec{x}|}|f(\vec{x})| \leq \varepsilon, \]
i.e.,
\[ |f(\vec{x})| \leq \frac{2\varepsilon}{\delta}|\vec{x}|. \]
By Note 2, this also holds if \( \vec{x} \neq \vec{0} \).

Thus, taking \( c = 2\varepsilon/\delta \), we obtain
\[ (\forall \vec{x} \in E') \quad |f(\vec{x})| \leq c|\vec{x}| \]
linear boundedness).

Now assume (3). Then
\[ (\forall \vec{x}, \vec{y} \in E') \quad |f(\vec{x} - \vec{y})| \leq c|\vec{x} - \vec{y}|; \]
or, by linearity,
\[ (\forall \vec{x}, \vec{y} \in E') \quad |f(\vec{x}) - f(\vec{y})| \leq c|\vec{x} - \vec{y}|. \]
Hence \( f \) is uniformly continuous (given \( \varepsilon > 0 \), take \( \delta = \varepsilon/c \).) This, in turn, implies continuity at \( \vec{0} \); so all conditions are equivalent, as claimed. \( \square \)

\footnote{This is the so-called uniform Lipschitz condition.}
A linear map need not be continuous. But, for $E^n$ and $C^n$, we have the following result.

**Theorem 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

$$f(\vec{x}) = \vec{x} \cdot \vec{v} \quad \text{(dot product)}$$

for some unique vector $\vec{v} \in E^n$ ($C^n$), dependent on $f$ only.

**Proof.** Suppose $f : E^n \to E$ is linear; so $f$ preserves linear combinations.

But every $\vec{x} \in E^n$ is such a combination,

$$\vec{x} = \sum_{k=1}^n x_k \vec{e}_k \quad \text{(Theorem 2 in Chapter 3, §§1–3).}$$

Thus, by Note 1,

$$f(\vec{x}) = f \left( \sum_{k=1}^n x_k \vec{e}_k \right) = \sum_{k=1}^n x_k f(\vec{e}_k).$$

Here the function values $f(\vec{e}_k)$ are fixed vectors in the range space $E$, say,

$$f(\vec{e}_k) = \vec{v}_k \in E,$$

so that

$$f(\vec{x}) = \sum_{k=1}^n x_k f(\vec{e}_k) = \sum_{k=1}^n x_k \vec{v}_k, \quad \vec{v}_k \in E. \quad (5)$$

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 $\vec{v}_k$ in (5) are real numbers; so they form a vector

$$\vec{v} = (v_1, \ldots, v_n) \in E^n,$$

and (5) can be written as

$$f(\vec{x}) = \vec{x} \cdot \vec{v}. \quad (6)$$

The vector $\vec{v}$ is unique. For suppose there are two vectors, $\vec{v}$ and $\vec{u}$, such that

$$(\forall \vec{x} \in E^n) \quad f(\vec{x}) = \vec{x} \cdot \vec{v} = \vec{x} \cdot \vec{u}. \quad (\forall \vec{x} \in E^n) \quad \vec{x} \cdot (\vec{v} - \vec{u}) = 0.$$

\[\text{Note 3.} \quad \text{Formula (5) shows that a linear map } f : E^n (C^n) \to E \text{ is uniquely determined by the } n \text{ function values } v_k = f(\vec{e}_k). \]

If further $E = E^n (C^n)$, the vectors $v_k$ are $m$-tuples of scalars,

$$v_k = (v_{1k}, \ldots, v_{mk}).$$

We often write such vectors vertically, as the $n$ “columns” in an array of $m$ “rows” and $n$ “columns”:

$$\begin{bmatrix}
11 & 12 & \cdots & 1n \\
21 & 22 & \cdots & 2n \\
\vdots & \vdots & \ddots & \vdots \\
m1 & m2 & \cdots & mnn
\end{bmatrix}.$$  \( (6) \)

Formally, (6) is a double sequence of $mn$ terms, called an $m \times n$ matrix. We denote it by $[f] = (v_{ik})$, where for $k = 1, 2, \ldots, n$,

$$f(\vec{e}_k) = \vec{v}_k = (v_{1k}, \ldots, v_{mk}).$$

Thus linear maps $f : E^n \to E^n$ (or $f : C^n \to C^n$) correspond one-to-one to their matrices $[f]$.

The easy proof of Corollaries 1 to 3 below is left to the reader.

**Corollary 1.** If $f, g : E^r \to E$ are linear, so is

$$h = af + bg$$

for any scalars $a, b$.

If further $E^r = E^n (C^n)$ and $E = E^m (C^m)$, with $[f] = (v_{ik})$ and $[g] = (w_{ik})$, then

$$[h] = (av_{ik} + bw_{ik}).$$

**Corollary 2.** A map $f : E^n (C^n) \to E$ is linear iff

$$f(\vec{x}) = \sum_{k=1}^n v_k x_k,$$

where $v_k = f(\vec{e}_k)$.

Hint: For the “if,” use Corollary 1. For the “only if,” use formula (5) above.
Corollary 3. If \( f : E' \to E'' \) and \( g : E'' \to E \) 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 3.** Let \( f : E' \to E'' \) and \( g : E'' \to E \) be linear, with \( E' = E^n(C^m) \), \( E'' = E^m(C'^n) \), and \( E = E^o(C''^r) \).

If \( [f] = (v_{ik}) \) and \( [g] = (w_{jk}) \), then

\[
[ h ] = [ g \circ f ] = (z_{jk}),
\]

where

\[
z_{jk} = \sum_{i=1}^{m} w_{ji} v_{ik}, \quad j = 1, 2, \ldots, r, \ k = 1, 2, \ldots, n.
\]

**Proof.** Denote the basic unit vectors in \( E' \) by

\[ e'_1, \ldots, e'_m, \]

those in \( E'' \) by

\[ e_1'', \ldots, e_m'', \]

and those in \( E \) by

\[ e_1, \ldots, e_r. \]

Then for \( k = 1, 2, \ldots, n, \)

\[
f(e'_k) = v_k = \sum_{i=1}^{m} v_{ik} e'_i \quad \text{and} \quad h(e'_k) = \sum_{j=1}^{r} z_{jk} e_j,
\]

and for \( i = 1, \ldots, m, \)

\[
g(e''_i) = \sum_{j=1}^{r} w_{ji} e_j.
\]

Also,

\[
h(e'_k) = g(f(e'_k)) = g \left( \sum_{i=1}^{m} v_{ik} e''_i \right) = \sum_{i=1}^{m} v_{ik} g(e''_i) = \sum_{i=1}^{m} v_{ik} \left( \sum_{j=1}^{r} w_{ji} e_j \right).
\]

Thus

\[
h(e'_k) = \sum_{j=1}^{r} z_{jk} e_j = \sum_{j=1}^{r} \left( \sum_{i=1}^{m} w_{ji} v_{ik} \right) e_j.
\]

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 \( j \)th row of \([g]\) (an \( r \times m \) matrix) by the \( k \)th column of \([f]\) (an \( m \times n \) matrix).

It is natural to set

\[
[g] \ [f] = [g \circ f],
\]

or

\[
(w_{jk}) (v_{ik}) = (z_{jk}),
\]

with \( z_{jk} \) as in (7).

**Caution.** Matrix multiplication, so defined, is not commutative.

**Definition 2.**

The set of all continuous linear maps \( f : E' \to E \) (for fixed \( E' \) and \( E \)) is denoted \( L(E', E) \).

If \( E = E' \), we write \( L(E) \) instead.

For each \( f \) in \( L(E', E) \), we define its norm by

\[
\|f\| = \sup_{\|\xi\| \leq 1} |f(\xi)|. \tag{3}
\]

Note that \( \|f\| < +\infty \), by Theorem 1.

**Theorem 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,

\[
\|f + g\| \leq \|f\| + \|g\|,
\]

follows exactly as in Example (C) of Chapter 3, §10. (Verify!)

Also, by Problem 5 in Chapter 2, §§8–9, \( |a f(\vec{x})| = |a| |f(\vec{x})| \). 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

\[
\|f\| = \sup_{|\vec{x}| \leq 1} |f(\vec{x})| = 0,
\]

then \( |f(\vec{x})| = 0 \) when \( |\vec{x}| \leq 1 \). Hence, if \( \vec{x} \neq 0, \)

\[
(f(\vec{x})) = \frac{1}{|\vec{x}|} f(\vec{x}) = 0.
\]

As \( f(\vec{0}) = 0, \) 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. \( \square \)

\[\text{3} \text{Equivalently, } \|f\| = \sup_{\vec{x} \neq 0} |f(\vec{x})|/|\vec{x}|; \text{ see Note 5 below.}\]
Note 5. A similar proof, via $f(\vec{u})$ and properties of lub, shows that

$$\|f\| = \sup_{\|\vec{x}\| \leq 1} |f(\vec{x})|$$

and

$$(\forall \vec{x} \in E') \quad \|f(\vec{x})\| \leq \|f\| \|\vec{x}\|.$$

It also follows that $\|f\|$ is the least real $c$ such that

$$(\forall \vec{x} \in E') \quad |f(\vec{x})| \leq c\|\vec{x}\|.$$

Verify. (See Problem 3'.)

As in any normed space, we define distances in $L(E', E)$ by

$$\rho(f, g) = \|f - g\|,$$

making it a metric space; so we may speak of convergence, limits, etc., in it.

Corollary 4. If $f \in L(E', E'')$ and $g \in L(E'', E)$, then

$$\|g \circ f\| \leq \|g\| \|f\|.$$

Proof. By Note 5,

$$(\forall \vec{x} \in E') \quad |g(f(\vec{x}))| \leq \|g\| \|f(\vec{x})\| \leq \|g\| \|f\| \|\vec{x}\|.$$}

Hence

$$(\forall \vec{x} \neq \vec{0}) \quad \left|\frac{(g \circ f)(\vec{x})}{\|\vec{x}\|}\right| \leq \|g\| \|f\|,$$

and so

$$\|g\| \|f\| \geq \sup_{\vec{x} \neq \vec{0}} \left|\frac{(g \circ f)(\vec{x})}{\|\vec{x}\|}\right| = \|g \circ f\|. \quad \Box$$

Problems on Linear Maps and Matrices

1. Verify Note 1 and the equivalence of the two statements in Definition 1.

2. In Examples (b) and (c) show that $f_n \to f$ (uniformly) on $I$ iff $\|f_n - f\| \to 0$.

3. Prove Corollaries 1 to 3.

3'. Show that

$$\|f\| = \sup_{\|\vec{x}\| \leq 1} |f(\vec{x})| = \sup_{\|\vec{x}\| = 1} |f(\vec{x})| = \sup_{\vec{x} \neq \vec{0}} \left|\frac{f(\vec{x})}{\|\vec{x}\|}\right|$$

[Hint: From linearity of $f$ deduce that $|f(\vec{x})| \geq |f(c\vec{e})|, |c| < 1$. Hence one may disregard vectors of length $< 1$ when computing $\sup |f(\vec{x})|$. Why?]

4. Find the matrices $[f]$, $[g]$, $[h]$, $[k]$, and the defining formulas for the linear maps $f : E^2 \to E^3$, $g : E^3 \to E^3$, $h : E^3 \to E^3$, $k : E^1 \to E^3$ if

(i) $f(\vec{e}_1) = 3$, $f(\vec{e}_2) = -2$;

(ii) $g(\vec{e}_1) = (1, 0, -2, 4)$, $g(\vec{e}_2) = (0, 2, -1, 1)$, $g(\vec{e}_3) = (0, 1, 0, -1)$;

(iii) $h(\vec{e}_1) = (2, 2)$, $h(\vec{e}_2) = (0, -2)$, $h(\vec{e}_3) = (1, 0)$, $h(\vec{e}_4) = (-1, 1)$;

(iv) $k(1) = (0, 1, -1)$.

5. In Problem 4, use Note 4 to find the product matrices $[k][f]$, $[g][h]$, $[f][h]$, and $[g][g]$. Hence obtain the defining formulas for $k \circ f$, $g \circ k$, $f \circ g$, $g \circ g$.

6. For $m \times n$-matrices (with $m$ and $n$ fixed) define addition and multiplication by scalars as follows:

$$a[f] + b[g] = [af + bg] \quad \text{if } f, g \in L(E^m, E^n) \text{ (or } L(C^m, C^n)).$$

Show that these matrices form a vector space over $E^1$ (or $C$).

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 counterexamples!).

Which is the “unity” matrix?

8. Let $f : E' \to E$ be linear. Prove the following statements.

(i) The derivative $D_ef(\vec{u})$ exists and equals $f(\vec{u})$ for every $\vec{p}, \vec{u} \in E'$ ($\vec{u} \neq \vec{0}$);

(ii) $f$ is relatively continuous on any line in $E'$ (use Theorem 1 in §1);

(iii) $f$ carries any such line into a line in $E$.

9. Let $g : E'' \to E$ be linear. Prove that if some $f : E' \to E''$ has a $\vec{u}$-directed derivative at $\vec{p} \in E'$, so has $h = g \circ f$, and $D_fh(\vec{p}) = g(D_ef(\vec{p}))$.

[Hint: Use Problem 8.]
10. A set \( A \) in a vector space \( V \) (\( A \subseteq V \)) is said to be linear (or a linear subspace of \( V \)) if \( a\vec{x} + b\vec{y} \in A \) 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 \( f : V' \rightarrow E \) is a linear map and \( A \) is linear in \( E' \) (respectively, in \( E \), so is \( f[A] \) in \( E \) (respectively, so is \( f^{-1}[A] \) in \( E' \)).

11. A set \( A \) in a vector space \( V \) is called the span of a set \( B \subseteq A \) (\( A = \text{sp}(B) \)) if \( A \) consists of all linear combinations of vectors from \( B \). We then also say that \( B \) spans \( A \).

Prove the following:

(i) \( A = \text{sp}(B) \) is the smallest linear subspace of \( V \) that contains \( B \).
(ii) If \( f : V \rightarrow E \) is linear and \( A = \text{sp}(B) \), then \( f[A] = \text{sp}(f[B]) \) in \( E \).

12. A set \( B = \{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n\} \) in a vector space \( V \) is called a basis iff each \( \vec{v} \in V \) has a unique representation as

\[
\vec{v} = \sum_{i=1}^{n} a_i \vec{x}_i
\]

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 \( \mathbb{R}^n \) and \( \mathbb{C}^n \) (the \( e_i \) 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.,

\[
\sum_{i=1}^{n} a_i \vec{x}_i = \vec{0} \text{ iff all } a_i \text{ vanish}.
\]

(ii) If \( E' \) is finite-dimensional, all linear maps on \( E' \) are uniformly continuous. (See also Problems 3 and 4 of §6.)

13. Prove that if \( f : E^1 \rightarrow E \) is continuous and (\( \forall x, y \in E^1 \) )

\[
f(x + y) = f(x) + f(y),
\]

then \( f \) is linear; so, by Corollary 2, \( f(x) = vx \) where \( v = f(1) \).

[Hint: Show that \( f(ax) = af(x) \); for \( a = 1, 2, \ldots \) (note: \( ax = x + x + \cdots + x, n \) 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.]

§3. Differentiable Functions

As we know, a function \( f : E^1 \rightarrow E \) (on \( E^1 \)) is differentiable at \( p \in E^1 \) iff, with

\[
\Delta f = f(x) - f(p) \text{ and } \Delta x = x - p,
\]

\[
f'(p) = \lim_{x \to p} \frac{\Delta f}{\Delta x} \text{ exists (finite)}.
\]

Setting \( \Delta x = x - p = t, \Delta f = f(p + t) - f(p) \), and \( f'(p) = v \), we may write this equation as

\[
\lim_{t \to 0} \frac{\Delta f}{t} = v = 0,
\]

or

(1) \[ \lim_{t \to 0} \frac{1}{|t|}|f(p + t) - f(p) - vt| = 0. \]

Now define a map \( \phi : E' \rightarrow E \) by \( \phi(t) = tv, v = f'(p) \in E \).

Then \( \phi \) is linear and continuous, i.e., \( \phi \in L(E^1, E) \); so by Corollary 2 in §2, we may express (1) as follows: there is a map \( \varphi \in L(E^1, E) \) such that

\[
\lim_{t \to 0} \frac{1}{|t|} |\Delta f - \phi(t)| = 0.
\]

We adopt this as a definition in the general case, \( f : E' \rightarrow E \), as well.

Definition 1.

A function \( f : E' \rightarrow E \) (where \( E' \) and \( E \) are normed spaces over the same scalar field) is said to be differentiable at a point \( \vec{p} \in E' \) iff there is a map

\( \phi \in L(E', E) \)

such that

(2) \[ \lim_{t \to 0} \frac{1}{|t|} |f(\vec{p} + t\vec{e}) - f(\vec{p}) - \phi(t)| = 0. \]

As we show below, \( \phi \) is unique (for a fixed \( \vec{p} \)), if it exists.

We call \( \phi \) 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 \( df(t\vec{e}) \) and \( df(\vec{p}; \cdot) \) for \( df(\vec{p}; \cdot) \).

Some authors write \( f' \) 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 \( \{f'(\vec{p})\} \) for its matrix, as follows.

Definition 2.

If \( E' = \mathbb{R}^n \) (or \( \mathbb{C}^n \)) and \( E = \mathbb{R}^m \) (or \( \mathbb{C}^m \)), and \( f : E' \rightarrow E \) is differentiable at \( \vec{p} \), we set

\[
[f'(\vec{p})] = [df(\vec{p}; \cdot)]
\]
Note 1. In Chapter 5, §6, we did not define $df$ as a mapping. However, if $E' = E^3$, the function value

$$df(p; t) = vt = f'(p)\Delta x$$

is as in Chapter 5, §6.

Also, $[f'(p)]$ is a $1 \times 1$ matrix with single term $f'(p)$. (Why?) This motivated Definition 2.

Theorem 1 (uniqueness of $df$). If $f: E' \to E$ is differentiable at $\bar{p}$, then the map $\phi$ described in Definition 1 is unique (dependent on $f$ and $\bar{p}$ only).

Proof. Suppose there is another linear map $g: E' \to E$ such that

$$\lim_{t \to 0} \frac{1}{|t|} [f(\bar{p} + \bar{t}) - f(\bar{p}) - g(\bar{t})] = 0.$$  

Let $h = \phi - g$. By Corollary 1 in §2, $h$ is linear.

Also, by the triangle law,

$$|h(\bar{t})| = |\phi(\bar{t}) - g(\bar{t})| \leq |\Delta f - \phi(\bar{t})| + |\Delta f - g(\bar{t})|.$$  

Hence, dividing by $|\bar{t}|$,

$$\lim_{|\bar{t}| \to 0} \frac{1}{|\bar{t}|} |h(\bar{t})| \leq \frac{1}{|\bar{t}|} |\Delta f - \phi(\bar{t})| + \frac{1}{|\bar{t}|} |\Delta f - g(\bar{t})|.$$  

By (1) and (2), the right side expressions tend to 0 as $\bar{t} \to \bar{0}$. Thus

$$\lim_{\bar{t} \to \bar{0}} h \left( \frac{\bar{t}}{|\bar{t}|} \right) = 0.$$  

This remains valid also if $\bar{t} \to \bar{0}$ over any line through $\bar{0}$, so that $\bar{t}/|\bar{t}|$ remains constant, say $\bar{t}/|\bar{t}| = \bar{u}$, where $\bar{u}$ is an arbitrary (but fixed) unit vector.

Then

$$h(\bar{u}) = \lim_{\bar{t} \to \bar{0}} h \left( \frac{\bar{t}}{|\bar{t}|} \right) = 0.$$  

is constant; so it can tend to 0 only if it equals 0, so $h(\bar{u}) = 0$ for any unit vector $\bar{u}$.

Since any $\bar{x} \in E'$ can be written as $\bar{x} = |\bar{x}| \bar{u}$, linearity yields

$$h(\bar{x}) = |\bar{x}| h(\bar{u}) = 0.$$  

Thus $h = \phi - g = 0$ on $E'$, and so $\phi = g$ after all, proving the uniqueness of $\phi$. □

§3. Differentiable Functions

Theorem 2. If $f$ is differentiable at $\bar{p}$, then

(i) $f$ is continuous at $\bar{p}$;

(ii) for any $\bar{u} \neq \bar{0}$, $f$ has the $\bar{u}$-directed derivative

$$D_{\bar{u}} f(\bar{p}) = df(\bar{p}; \bar{u}).$$

Proof. By assumption, formula (2) holds for $\phi = df(\bar{p}; \cdot)$.

Thus, given $\varepsilon > 0$, there is $\delta > 0$ such that, setting $\Delta f = f(\bar{p} + \bar{t}) - f(\bar{p})$, we have

$$|\Delta f| < \varepsilon$$

or, by the triangle law,

$$|\Delta f| \leq |\Delta f - \phi(\bar{t})| + |\phi(\bar{t})| \to 0$$

as $|\bar{t}| < \delta$.

Now, by Definition 1, $\phi$ is linear and continuous; so

$$\lim_{|\bar{t}| \to 0} |\phi(\bar{t})| = 0.$$  

Thus, making $\bar{t} \to \bar{0}$ in (5), with $\varepsilon$ fixed, we get

$$\lim_{|\bar{t}| \to 0} |\Delta f| = 0.$$  

As $\bar{t}$ is just another notation for $\Delta \bar{x} = \bar{x} - \bar{p}$, this proves assertion (i).

Next, fix any $\bar{u} \neq \bar{0}$ in $E'$, and substitute $t\bar{u}$ for $\bar{t}$ in (4).

In other words, $t$ is a real variable, $0 < t < \delta/|\bar{u}|$, so that $\bar{t} = t\bar{u}$ satisfies

$$0 < |\bar{t}| < \delta.$$  

Multiplying by $|\bar{u}|$, we use the linearity of $\phi$ to get

$$\varepsilon|\bar{u}| > |\Delta f| = |\Delta f - \phi(\bar{u})|.$$  

As $\varepsilon$ is arbitrary, we have

$$\phi(\bar{u}) = \lim_{|\bar{t}| \to 0} \frac{1}{|t|} [f(\bar{p} + t\bar{u}) - f(\bar{p})].$$  

But this is simply $D_{\bar{u}} f(\bar{p})$, by Definition 1 in §1.

Thus $D_{\bar{u}} f(\bar{p}) = \phi(\bar{u}) = df(\bar{p}; \bar{u})$, proving (ii). □

Note 2. If $E' = E^n$ ($C^m$), Theorem 2(ii) shows that if $f$ is differentiable at $\bar{p}$, it has the $n$ partials

$$D_k f(\bar{p}) = df(\bar{p}; \bar{e}_k), \quad k = 1, \ldots, n.$$  


But the converse \textit{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 1.** If $E' = E^n (C^n)$ and if $f : E' \rightarrow E$ is differentiable at $\vec{p}$, then

\begin{equation}
\frac{df}{\partial \vec{x}} = \sum_{k=1}^{n} t_k D_k f(\vec{p}) = \sum_{k=1}^{n} t_k \frac{\partial f}{\partial x_k} (\vec{p}),
\end{equation}

where $\vec{t} = (t_1, \ldots, t_n)$.

**Proof.** By definition, $\phi = df(\vec{p}, \cdot)$ 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 $\phi$ and $\vec{t}$, and get

\[ \phi(\vec{t}) = df(\vec{p}; \vec{t}) = \sum_{k=1}^{n} t_k df(\vec{p}; \vec{e}_k) = \sum_{k=1}^{n} t_k D_k f(\vec{p}) \]

by Note 2.

**Note 3.** In classical notation, one writes $\Delta x_k$ or $dx_k$ for $t_k$ in (6). Thus, omitting $\vec{p}$ and $\vec{t}$, formula (6) is often written as

\begin{equation}
\frac{df}{\partial x_i} dx_1 + \frac{d f}{\partial x_2} dx_2 + \cdots + \frac{d f}{\partial x_n} dx_n.
\end{equation}

In particular, if $n = 3$, we write $x, y, z$ for $x_1, x_2, x_3$. This yields

\begin{equation}
\frac{df}{\partial x} dx + \frac{df}{\partial y} dy + \frac{df}{\partial z} dz
\end{equation}

(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

\[ \nabla f(\vec{p}) = (D_1 f(\vec{p}), \ldots, D_n f(\vec{p})) = \sum_{k=1}^{n} \vec{e}_k D_k f(\vec{p}). \]

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

\[ \nabla f(\vec{p}) = \sum_{k=1}^{n} \overline{e}_k D_k \overline{f(\vec{p})}. \]

The vector $\nabla f(\vec{p})$ is called the gradient of $f$ (“grad $f$”) at $\vec{p}$.

From (6) we obtain

\begin{equation}
\frac{df(\vec{p})}{\partial \vec{t}} = \sum_{k=1}^{n} t_k D_k f(\vec{p}) = \vec{t} \cdot \nabla f(\vec{p})
\end{equation}

(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 2.** A function $f : E^n \rightarrow E^1$ (or $f : C^n \rightarrow C$) is differentiable at $\vec{p}$ iff

\begin{equation}
\lim_{t \rightarrow 0} \frac{1}{|t|} [f(\vec{p} + \vec{t}) - f(\vec{p}) - \vec{t} \cdot \vec{v}] = 0
\end{equation}

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),

\[ \phi(\vec{t}) = df(\vec{p}; \vec{t}) = \vec{t} \cdot \vec{v}; \]

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 $\phi$ is linear and continuous.

Thus by definition, $f$ is differentiable at $\vec{p}$; so (7) holds.

Also, $\phi$ 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 $\phi$.

Thus by (7), $\vec{v} = \nabla f(\vec{p})$ necessarily. □

**Corollary 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

\begin{equation}
f(\vec{q}) - f(\vec{p}) = (\vec{q} - \vec{p}) \cdot \nabla f(\vec{x}_0)
\end{equation}

for some $\vec{x}_0 \in L(\vec{p}, \vec{q})$.

**Proof.** Let

\[ \vec{r} = |\vec{q} - \vec{p}|, \quad \vec{v} = \frac{1}{\vec{r}} (\vec{q} - \vec{p}), \quad \text{and} \quad r\vec{v} = (\vec{q} - \vec{p}). \]

By (7) and Theorem 2(ii),

\[ D_{\vec{e}} f(\vec{x}) = df(\vec{x}; \vec{v}) = \vec{v} \cdot \nabla f(\vec{x}) \]

for $\vec{x} \in L(\vec{p}, \vec{q})$. Thus by formula (3') of Corollary 2 in §1,

\[ f(\vec{q}) - f(\vec{p}) = r D_{\vec{e}} f(\vec{x}_0) = r \vec{v} \cdot \nabla f(\vec{x}_0) = (\vec{q} - \vec{p}) \cdot \nabla f(\vec{x}_0) \]

for some $\vec{x}_0 \in L(\vec{p}, \vec{q})$. □

As we know, the mere existence of partials does not imply differentiability. But the existence of \textit{continuous} partials does. Indeed, we have the following theorem.
Theorem 3. Let $E' = E^n (C^n)$.

If $f: E' \rightarrow E$ has the partial derivatives $D_k f$ $(k = 1, \ldots, n)$ on all of an open set $A \subseteq E'$, 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

$$\phi(\vec{t}) = \sum_{k=1}^{n} t_k D_k f(\vec{p})$$

Then $\phi$ is continuous (a polynomial!) and linear (Corollary 2 in §2).

Thus by Definition 1, it remains to show that

$$\lim_{\vec{t} \rightarrow \vec{0}} \frac{1}{|\vec{t}|} |\Delta f - \phi(\vec{t})| = 0;$$

that is,

$$\lim_{\vec{t} \rightarrow \vec{0}} \frac{1}{|\vec{t}|} |f(\vec{p} + \vec{t}) - f(\vec{p}) - \sum_{k=1}^{n} t_k D_k f(\vec{p})| = 0. \quad (10)$$

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 $G_\delta(\vec{p}) \subseteq A$ and simultaneously (explain this!)

$$(\forall \vec{x} \in G_\delta(\vec{p})) \quad |D_k f(\vec{x}) - D_k f(\vec{p})| < \frac{\varepsilon}{n}, \quad k = 1, \ldots, n.$$  

Hence for any set $I \subseteq G_\delta(\vec{p})$

$$\sup_{\vec{x} \in I} |D_k f(\vec{x}) - D_k f(\vec{p})| \leq \frac{\varepsilon}{n}. \quad (11)$$

Now fix any $\vec{i} \in E'$, $0 < |\vec{i}| < \delta$, and let $\vec{p}_0 = \vec{p}$,

$$\vec{p}_k = \vec{p} + \sum_{i=1}^{k} t_i \vec{e}_i, \quad k = 1, \ldots, n.$$  

Then

$$\vec{p}_n - \vec{p}_{k-1} = \vec{t}_k,$$

and all $\vec{p}_k$ lie in $G_\delta(\vec{p})$, for

$$|\vec{p}_k - \vec{p}_{k-1}| = |\vec{t}_k|, \quad k = 1, \ldots, n.$$  

as required.

\[\text{§3. Differentiable Functions}\]

As $G_\delta(\vec{p})$ is convex (Chapter 4, §9), the segments $I_k = L[\vec{p}_{k-1}, \vec{p}_k]$ all lie in $G_\delta(\vec{p}) \subseteq A$; 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

$$g_k(\vec{x}) = f(\vec{x}) - x_k D_k f(\vec{p}), \quad k = 1, \ldots, n. \quad (12)$$

(Why?) Here

$$D_k g_k(\vec{x}) = D_k f(\vec{x}) - D_k f(\vec{p}). \quad (12')$$

(Why?)

Thus by Corollary 2 in §1, and (11) above,

$$|g_k(\vec{p}_k) - g_k(\vec{p}_{k-1})| \leq |\vec{p}_k - \vec{p}_{k-1}| \sup_{\vec{e} \in I_k} |D_k f(\vec{x}) - D_k f(\vec{p})|$$

$$\leq \frac{\varepsilon}{n} |\vec{t}_k| \leq \frac{\varepsilon}{n} |\vec{i}|,$$

since

$$|\vec{p}_k - \vec{p}_{k-1}| = |\sum_{i=1}^{k} t_i \vec{e}_i| \leq |\vec{i}|,$$

by construction.

Combine with (12), recalling that the $k$th coordinates $x_k$, for $\vec{p}_k$ and $\vec{p}_{k-1}$, differ by $t_k$; so we obtain

$$|g_k(\vec{p}_k) - g_k(\vec{p}_{k-1})| = |f(\vec{p}_k) - f(\vec{p}_{k-1}) - t_k D_k f(\vec{p})|$$

$$\leq \frac{\varepsilon}{n} |\vec{i}|. \quad (13)$$

Also,

$$\sum_{k=1}^{n} |f(\vec{p}_k) - f(\vec{p}_{k-1})| = f(\vec{p}_n) - f(\vec{p}_0)$$

$$= f(\vec{p} + \vec{i}) - f(\vec{p}) = \Delta f \ (\text{see above}).$$

Thus

$$|\Delta f - \sum_{k=1}^{n} t_k D_k f(\vec{p})| = \left| \sum_{k=1}^{n} [f(\vec{p}_k) - f(\vec{p}_{k-1}) - t_k D_k f(\vec{p})] \right|$$

$$\leq n \cdot \frac{\varepsilon}{n} |\vec{i}| = \varepsilon |\vec{i}|.$$  

As $\varepsilon$ is arbitrary, (10) follows, and all is proved. \[\square\]

Theorem 4. If $f: E^m \rightarrow E^n$ (or $f: C^n \rightarrow C^m$) is differentiable at $\vec{p}$, with $f = (f_1, \ldots, f_m)$, then $f'(\vec{p})$ is an $m \times n$ matrix,

$$|f'(\vec{p})| = |D_k f_i(\vec{p})|, \quad i = 1, \ldots, m, \quad k = 1, \ldots, n. \quad (14)$$
Proof. By definition, \([f'(\bar{p})]\) is the matrix of the linear map \(\phi = df(\bar{p}, \cdot)\), \(\phi = (\phi_1, \ldots, \phi_m)\). Here
\[
\phi(\bar{t}) = \sum_{k=1}^{n} t_k D_k f(\bar{p})
\]
by Corollary 1.
As \(f = (f_1, \ldots, f_m)\), we can compute \(D_k f(\bar{p})\) componentwise by Theorem 5 of Chapter 5, §1, and Note 2 in §1 to get
\[
D_k f(\bar{p}) = (D_{k1} f_1(\bar{p}), \ldots, D_{km} f_m(\bar{p})) = \sum_{i=1}^{m} e'_i D_k f_i(\bar{p}), \quad k = 1, 2, \ldots, n,
\]
where the \(e'_i\) are the basic vectors in \(E^n (C^m)\). (Recall that the \(e_k\) are the basic vectors in \(E^n (C^n)\).
Thus
\[
\phi(\bar{t}) = \sum_{i=1}^{m} e'_i \phi_i(\bar{t}).
\]
Also,
\[
\phi(\bar{t}) = \sum_{k=1}^{n} t_k \sum_{i=1}^{m} e'_i D_k f_i(\bar{p}) = \sum_{i=1}^{m} e'_i \sum_{k=1}^{n} t_k D_k f_i(\bar{p}).
\]
The uniqueness of the decomposition (Theorem 2 in Chapter 3, §§1–3) now yields
\[
\phi_i(\bar{t}) = \sum_{k=1}^{n} t_k D_k f_i(\bar{p}), \quad i = 1, \ldots, m, \quad \bar{t} \in E^n (C^n).
\]
If here \(\bar{t} = e_k\), then \(t_k = 1, \) while \(t_j = 0\) for \(j \neq k\). Thus we obtain
\[
\phi_i(e_k) = D_k f_i(\bar{p}), \quad i = 1, \ldots, m, \quad k = 1, \ldots, n.
\]
Hence
\[
\phi(\bar{e}_k) = (v_{1k}, v_{2k}, \ldots, v_{mk}),
\]
where
\[
v_{ik} = \phi_i(e_k) = D_k f_i(\bar{p}).
\]
But by Note 3 of §2, \((v_{1k}, \ldots, v_{mk})\) (written vertically) is the \(k\)th column of the \(m \times n\) matrix \([\phi] = [f'(\bar{p})]\). Thus formula (14) results indeed. \(\square\)

In conclusion, let us stress again that while \(D_k f(\bar{p})\) is a constant, for a fixed \(\bar{p}\), \(df(\bar{p}, \cdot)\) is a mapping \(\phi \in L(E', E)\), especially “tailored” for \(\bar{p}\).

The reader should carefully study at least the “arrowed” problems below.

### Problems on Differentiable Functions

1. Complete the missing details in the proofs of this section.
2. Verify Note 1. Describe \([f'(\bar{p})]\) for \(f: E^1 \to E^m\), too. Give examples.
3. A map \(f: E' \to E\) is said to satisfy a Lipschitz condition (L) of order \(\alpha > 0\) at \(\bar{p}\) if
\[
(\exists \delta > 0) (\exists K \in E^1) (\forall \bar{x} \in G_{\bar{p}}(\delta)) \quad |f(\bar{x}) - f(\bar{p})| \leq K|\bar{x} - \bar{p}|^\alpha.
\]
Prove the following.
(i) This implies continuity at \(\bar{p}\) (but not conversely; see Problem 7 in Chapter 5, §1).
(ii) \(L\) of order \(\alpha > 1\) implies differentiability at \(\bar{p}\), with \(df(\bar{p}, \cdot) = 0\) on \(E'\).
(iii) Differentiability at \(\bar{p}\) implies \(L\) of order 1 (apply Theorem 1 in §2 to \(\phi = df\)).
(iv) If \(f \) and \(g\) are differentiable at \(\bar{p}\), then
\[
\lim_{\bar{x} \to \bar{p}} \frac{1}{|\Delta \bar{x}|} |\Delta f| |\Delta g| = 0.
\]
4. For the functions of Problem 5 in §1, find those \(\bar{p}\) at which \(f\) is differentiable. Find
\[
\nabla f(\bar{p}), \quad df(\bar{p}, \cdot), \quad \text{and} \quad [f'(\bar{p})].
\]
[Hint: Use Theorem 3 and Corollary 1.]
5. Prove the following statements.
(i) If \(f: E' \to E\) is constant on an open globe \(G \subset E'\), it is differentiable at each \(\bar{p} \in G\), and \(df(\bar{p}, \cdot) = 0\) on \(E'\).
(ii) If the latter holds for each \(\bar{p} \in G - Q\) (\(Q\) countable), then \(f\) is constant on \(G\) (even on \(\overline{G}\)) provided \(f\) is relatively continuous there.
[Hint: Given \(\bar{p}, \bar{q} \in G\), use Theorem 2 in §1 to get \(f(\bar{p}) = f(\bar{q})\).]
6. Do Problem 5 in case \(G\) is any open polygon-connected set in \(E'\). (See Chapter 4, §9.)
7. Prove the following.
(i) If \(f, g: E' \to E\) are differentiable at \(\bar{p}\), so is
\[
h = af + bg.
\]
for any scalars $a, b$ (if $f$ and $g$ are scalar valued, $a$ and $b$ may be vectors); moreover,
\[ d(af + bg) = adf + b dg, \]
i.e.,
\[ dh(\vec{p}; \vec{t}) = a df(\vec{p}; \vec{t}) + b dg(\vec{p}; \vec{t}), \]
$\vec{t} \in E'$.

(ii) In case $f, g: E^m \to E^1$ or $C^m \to C$, deduce also that
\[ \nabla h(\vec{p}) = a \nabla f(\vec{p}) + b \nabla g(\vec{p}). \]

**8.** Prove that if $f, g: E' \to E^1 (C)$ are differentiable at $\vec{p}$, then so are
\[ h = gf \text{ and } k = \frac{g}{f}. \]
(later, if $f(\vec{p}) \neq 0$). Moreover, with $a = f(\vec{p})$ and $b = g(\vec{p})$, show that
(i) $dh = adg + b df$ and
(ii) $dk = (a dg - b df)/a^2$.

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 $\phi = a dg + b df$, with $a$ and $b$ as above. Verify that
\[ \Delta h - \phi(\vec{t}) = g(\vec{p})(\Delta f - df(\vec{t})) + f(\vec{p})(\Delta g - dg(\vec{t})) + (\Delta f)(\Delta g). \]
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 $gF$.]

**9.** Let $f: E' \to E^m (C^m)$, $f = (f_1, \ldots, f_m)$. 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, $df = df_{re} + i \cdot df_{im}$.

**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.

**11.** Any rational function is differentiable in its domain.

[Hint: Use Problems 10(ii), 7, and 8. Proceed as in Theorem 3 in Chapter 4, §3.]

**12.** Do Problem 8(i) in case $g$ is only continuous at $\vec{p}$, and $f(\vec{p}) = 0$. Find $dh$.

**13.** Do Problem 8(i) for dot products $h = f \cdot g$ of functions $f, g: E' \to E^m (C^m)$.

**14.** Prove the following.

(i) If $\phi \in L(E^n, E^1)$ or $\phi \in L(C^n, C)$, then $||\phi|| = ||\vec{v}||$, with $\vec{v}$ as in §2,
Theorem 2(ii).

(ii) If $f: E^n \to E^1 (f: C^n \to C)$ is differentiable at $\vec{p}$, then
\[ ||df(\vec{p}; \cdot)|| = ||\nabla f(\vec{p})||. \]

Moreover, in case $f: E^n \to E^1$,
\[ ||\nabla f(\vec{p})|| \geq D_2 f(\vec{p}) \text{ if } |\vec{u}| = 1 \]
and
\[ ||\nabla f(\vec{p})|| = D_2 f(\vec{p}) \text{ when } \vec{u} = \frac{\nabla f(\vec{p})}{||\nabla f(\vec{p})||}, \]
thus
\[ ||\nabla f(\vec{p})|| = \max_{|\vec{u}| = 1} D_2 f(\vec{p}). \]

[Hints: Use the equality case in Theorem 4(c) of Chapter 3, §§1–3. Use formula (7),
Corollary 2, and Theorem 2(ii).]

**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 $\delta$ has
been chosen.]

**16.** Show that Theorem 3 and Problem 15 apply also to any $f: E' \to E$ where $E'$ is $n$-dimensional with basis $\{\vec{u}_1, \ldots, \vec{u}_n\}$ (see Problem 12 in
§2) if we write $D_1 f$ for $D_{\vec{u}} f$.

[Hints: Assume $||\vec{u}_k|| = 1$, $1 \leq k \leq n$ (if not, replace $\vec{u}_k$ by $\vec{u}_k/||\vec{u}_k||$; show that this
yields another basis). Modify the proof so that the $\vec{p}_k$ are still in $G_0(\delta)$. Caution: The
standard norm of $E^n$ does not apply here.]

**17.** Let $f_k: E^1 \to E^1$ be differentiable at $p_k$ ($k = 1, \ldots, n$). For $\vec{x} = (x_1, \ldots, x_n) \in E^n$, set
\[ F(\vec{x}) = \sum_{k=1}^n f_k(x_k) \quad \text{and} \quad G(\vec{x}) = \prod_{k=1}^n f_k(x_k). \]
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 $f_k'(p_k)$. 

§3. Differentiable Functions
### §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, \( f : E' \to E'' \) and \( g : E'' \to E \), with \( E', E'', \) and \( E \) as before.

**Theorem 1** (chain rule). If \( f : E' \to E'' \) and \( g : E'' \to E \) are differentiable at \( \vec{p} \) and \( \vec{q} = f(\vec{p}) \), respectively, then

\[ h = g \circ f \]

is differentiable at \( \vec{p} \), and

\[ dh(\vec{p}, \cdot) = dg(\vec{q}, \cdot) \circ df(\vec{p}, \cdot). \]

Briefly: “The differential of the composite is the composite of differentials.”

**Proof.** Let \( U = df(\vec{p}, \cdot), V = dg(\vec{q}, \cdot), \) and \( \phi = V \circ U \).

As \( U \) and \( V \) are linear continuous maps, so is \( \phi \). We must show that \( \phi = dh(\vec{p}, \cdot) \).

Here it is more convenient to write \( \Delta \vec{x} \) or \( \vec{x} - \vec{p} \) for the \( v^{th} \) of Definition 1 in §3. For brevity, we set (with \( \vec{q} = f(\vec{p}) \))

\[ w(\vec{x}) = \Delta h - \phi(\Delta \vec{x}) = h(\vec{x}) - h(\vec{p}) - \phi(\vec{x} - \vec{p}), \quad \vec{x} \in E', \]

\[ u(\vec{x}) = \Delta f - U(\Delta \vec{x}) = f(\vec{x}) - f(\vec{p}) - U(\vec{x} - \vec{p}), \quad \vec{x} \in E', \]

\[ v(\vec{y}) = \Delta g - V(\Delta \vec{y}) = g(\vec{y}) - g(\vec{q}) - V(\vec{y} - \vec{q}), \quad \vec{y} \in E''. \]

Then what we have to prove (see Definition 1 in §3) reduces to

\[ \lim_{\vec{x} \to \vec{p}} \frac{w(\vec{x})}{|\vec{x} - \vec{p}|} = 0, \]

while the assumed existence of \( df(\vec{p}, \cdot) = U \) and \( dg(\vec{q}, \cdot) = V \) can be expressed as

\[ \lim_{\vec{x} \to \vec{p}} \frac{u(\vec{x})}{|\vec{x} - \vec{p}|} = 0, \]

and

\[ \lim_{\vec{y} \to \vec{q}} \frac{v(\vec{y})}{|\vec{y} - \vec{q}|} = 0, \quad \vec{q} = f(\vec{p}). \]

**From (2) and (3), recalling that** \( h = g \circ f \) and \( \phi = V \circ U \), we obtain

\[ w(\vec{x}) = g(f(\vec{x})) - g(\vec{q}) - V(U(\vec{x} - \vec{p})) = g(f(\vec{x})) - g(\vec{q}) - V(f(\vec{x}) - f(\vec{p}) - u(\vec{x})). \]

Using (4), with \( \vec{y} = f(\vec{x}) \), and the linearity of \( V \), we rewrite (6) as

\[ w(\vec{x}) = g(f(\vec{x})) - g(\vec{q}) - V(f(\vec{x}) - f(\vec{p})) - V(u(\vec{x})) = v(f(\vec{x})) + V(u(\vec{x})). \]

(Verify!) Thus the desired formula (5) will be proved if we show that

\[ \lim_{\vec{x} \to \vec{p}} \frac{V(u(\vec{x}))}{|\vec{x} - \vec{p}|} = 0 \]

and

\[ \lim_{\vec{x} \to \vec{p}} \frac{v(f(\vec{x}))}{|\vec{x} - \vec{p}|} = 0, \]

Now, as \( V \) is linear and continuous, formula (5') yields (6'). Indeed,

\[ \lim_{\vec{x} \to \vec{p}} \frac{V(u(\vec{x}))}{|\vec{x} - \vec{p}|} = \lim_{\vec{x} \to \vec{p}} V\left( \frac{u(\vec{x})}{|\vec{x} - \vec{p}|} \right) = V(0) = 0 \]

by Corollary 2 in Chapter 4, §2. (Why?)

Similarly, (5') implies (6') by substituting \( \vec{y} = f(\vec{x}) \), since

\[ |f(\vec{x}) - f(\vec{p})| \leq K|\vec{x} - \vec{p}| \]

by Problem 3(iii) in §3. (Explain!) Thus all is proved. □

**Note 1** (Cauchy invariant rule). Under the same assumptions, we also have

\[ \lim_{\vec{x} \to \vec{p}} \frac{w(\vec{x})}{|\vec{x} - \vec{p}|} = 0, \]

if \( \vec{s} = df(\vec{p}; \vec{t}) \), \( \vec{t} \in E' \).

For with \( U \) and \( V \) as above,

\[ dh(\vec{p}, \cdot) = \phi = V \circ U. \]

Thus if

\[ \vec{s} = df(\vec{p}; \vec{t}) = U(\vec{t}), \]

we have

\[ dh(\vec{p}; \vec{t}) = \phi(\vec{t}) = V(U(\vec{t})) = V(\vec{s}) = dg(\vec{q}; \vec{s}), \]

proving (7).
Note 2. If
\[ E' = E^n(C^n), \quad E'' = E^m(C^m), \quad \text{and} \quad E = E'(C') \]
then by Theorem 3 of §2 and Definition 2 in §3, we can write (1) in matrix form,
\[ [h'(\vec{p})] = [g'(\vec{q})] \cdot [f'(\vec{p})]. \]
resembling Theorem 3 in Chapter 5, §1 with \( f \) and \( g \) interchanged. Moreover, we have the following theorem.

Theorem 2. With all as in Theorem 1, let
\[ E' = E^n(C^n), \quad E'' = E^m(C^m), \]
and
\[ f = (f_1, \ldots, f_n). \]

Then
\[ D_k h(\vec{p}) = \sum_{i=1}^{m} D_i g(\vec{q}) \cdot D_k f_i(\vec{p}); \]
or, in classical notation,
\[ \frac{\partial}{\partial x_k} h(\vec{p}) = \sum_{i=1}^{m} \frac{\partial}{\partial y_i} g(\vec{q}) \cdot \frac{\partial}{\partial x_k} f_i(\vec{p}), \quad k = 1, 2, \ldots, n. \]

Proof. Fix any basic vector \( \vec{e}_k \) in \( E' \) and set
\[ \vec{s} = df(\vec{p}; \vec{e}_k), \quad \vec{s} = (s_1, \ldots, s_m) \in E''. \]
As \( f \) is differentiable at \( \vec{p} \), so are its components \( f_i \) (Problem 9 in §3), and
\[ s_i = df_i(\vec{p}; \vec{e}_k) = D_k f_i(\vec{p}) \]
by Theorem 2(ii) in §3. Using also Corollary 1 in §3, we get
\[ dg(\vec{q}; \vec{s}) = \sum_{i=1}^{m} s_i D_i g(\vec{q}) = \sum_{i=1}^{m} D_k f_i(\vec{p}) \cdot D_i g(\vec{q}). \]
But as \( \vec{s} = df(\vec{p}; \vec{e}_k) \), formula (7) yields
\[ dg(\vec{q}; \vec{s}) = dh(\vec{p}; \vec{e}_k) = D_k h(\vec{p}) \]
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 \( m = n = 1 \).

§4. The Chain Rule. The Cauchy Invariant Rule

In classical calculus one often speaks of derivatives and differentials of variables \( y = f(x_1, \ldots, x_n) \) rather than those of mappings. Thus Theorem 2 is stated as follows.

Let \( u = g(y_1, \ldots, y_m) \) be differentiable. If, in turn, each
\[ y_i = f_i(x_1, \ldots, x_n) \]
is differentiable for \( i = 1, \ldots, m \), then \( u \) is also differentiable as a composite function of the \( n \) variables \( x_1, \ldots, x_n \), and (“simplifying” formula (8)) we have
\[ \left(8'\right) \quad \frac{\partial u}{\partial x_k} = \sum_{i=1}^{m} \frac{\partial u}{\partial y_i} \cdot \frac{\partial y_i}{\partial x_k}, \quad k = 1, 2, \ldots, n. \]

It is understood that the partials \( \frac{\partial u}{\partial x_k} \) and \( \frac{\partial y_i}{\partial x_k} \) are taken at some \( \vec{p} \in E' \), while the \( \frac{\partial u}{\partial y_i} \) are at \( \vec{q} = f(\vec{p}) \), where \( f = (f_1, \ldots, f_m) \). 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 \( y \):
\[ z = f_3(x, y). \]
Set \( f_1(x, y) = x, \quad f_2(x, y) = y, \quad f = (f_1, f_2, f_3), \quad \text{and} \quad h = g \circ f; \quad \text{so} \quad h(x, y) = g(x, y, z). \]

By \( \left(8'\right) \),
\[ \frac{\partial u}{\partial x} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial x} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial x} + \frac{\partial u}{\partial z} \cdot \frac{\partial z}{\partial x}. \]

Here
\[ \frac{\partial x}{\partial x} = \frac{\partial f_1}{\partial x} = 1 \quad \text{and} \quad \frac{\partial u}{\partial x} = 0, \]
for \( f_2 \) does not depend on \( x \). Thus we obtain
\[ \left(9\right) \quad \frac{\partial u}{\partial x} = \frac{\partial u}{\partial x} + \frac{\partial u}{\partial z} \cdot \frac{\partial z}{\partial x}. \]

(Question: Is \( (\partial u/\partial z) \cdot (\partial z/\partial x) = 0? \)

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
\[ g : E^3 \to E \text{ and } f : E^2 \to E^3, \]
with \( f_1, f_2, f_3 \) as indicated. Then if \( h = g \circ f \), Theorem 2 states (9) unambiguously as
\[ D_1 h(\bar{p}) = D_1 g(\bar{q}) + D_2 g(\bar{q}) \cdot D_1 f(\bar{p}), \]
where \( \bar{p} \in E^2 \) and
\[ \bar{q} = f(\bar{p}) = (p_1, p_2, f_3(\bar{p})). \]
(Why?) In classical notation,
\[ \frac{\partial h}{\partial x} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial z} \frac{\partial f_3}{\partial x}, \]
(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
\[ du = \sum_{i=1}^{m} \frac{\partial u}{\partial y_i} dy_i = \sum_{k=1}^{n} \frac{\partial u}{\partial x_k} dx_k, \]
where \( dx_k = t_k \) and \( dy_i = df_i(\bar{p}; \bar{t}) \).

Indeed, by Corollary 1 in §3,
\[ dh(\bar{p}; \bar{t}) = \sum_{k=1}^{n} D_k h(\bar{p}) \cdot t_k \text{ and } dg(\bar{q}; \bar{s}) = \sum_{i=1}^{m} D_i g(\bar{q}) \cdot s_i. \]
Now, in (7),
\[ \bar{s} = (s_1, \ldots, s_m) = df(\bar{p}; \bar{t}); \]
so by Problem 9 in §3,
\[ df_i(\bar{p}; \bar{t}) = s_i, \quad i = 1, \ldots, m. \]
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.

---

**Further Problems on Differentiable Functions**

1. For \( E = E^r (C^r) \) prove Theorem 2 directly.
   [Hint: Find \( D_j h_j(\bar{p}), \quad j = 1, \ldots, r. \)]
   from Theorem 4 of §3, and Theorem 3 of §2. Verify that
   \[ D_j h(\bar{p}) = \sum_{j=1}^{r} e_j D_j h_j(\bar{p}) \]
   and \( D_j g(\bar{q}) = \sum_{j=1}^{r} e_j D_j g_j(\bar{q}), \]
   where the \( e_j \) are the basic unit vectors in \( E^r \). Proceed.

2. Let \( g(x, y, z) = u = f_1(r, \theta), y = f_2(r, \theta), z = f_3(r, \theta), \) and
   \[ f = (f_1, f_2, f_3) : E^2 \to E^3, \]
   Assuming differentiability, verify (using “variables”) that
   \[ du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz = \sum_{i=1}^{2} \frac{\partial u}{\partial r} dr + \frac{\partial u}{\partial \theta} d\theta \]
   by computing derivatives from (8′). Then do all in the mapping notation for \( H = g \circ f, \) by \( dH(\bar{p}, \bar{t}) \).

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
   (a) \( k \circ f \); (b) \( g \circ k \); (c) \( f \circ h \); (d) \( h \circ g \).

4. For the functions of Problem 5 in §1, find the formulas for \( df(\bar{p}; \bar{t}) \). At which \( \bar{p} \) does \( df(\bar{p}, \cdot) \) exist in each given case? Describe it for a chosen \( \bar{p} \).

5. From Theorem 2, with \( E = E^1 (C) \), find
   \[ \nabla h(\bar{p}) = \sum_{k=1}^{n} D_k g(\bar{q}) \nabla f_k(\bar{p}). \]

6. Use Theorem 1 for a new solution of Problem 7 in §3 with \( E = E^1 (C) \).
   [Hint: Define \( F \) on \( E^r \) and \( G \) on \( E^r (C^r) \) by
   \[ F(\bar{z}) = (f(\bar{z}), g(\bar{z})) \]
   and \( G(\bar{y}) = a_1 y_1 + b y_2 \).
   Then \( h = a f + b g = G \circ F. \) (Why?) Use Problems 9 and 10(ii) of §3. Do all in “variable” notation, too.]

7. Use Theorem 1 for a new proof of the “only if” in Problem 9 in §3.
   [Hint: Set \( f_t = g \circ f, \) where \( g(\bar{z}) = x_t \) (the \( t \)th “projection map”) is a monomial. Verify!]

8. Do Problem 8(i) in §3 for the case \( E' = E^2 (C^2) \), with
   \[ f(\bar{z}) = x_1 \text{ and } g(\bar{z}) = x_2. \]
(Simplify!) Then do the general case as in Problem 6 above, with
\[ G(\tilde{y}) = y_1 y_2. \]

9. Use Theorem 2 for a new proof of Theorem 4 in Chapter 5, §1. (Proceed as in Problems 6 and 8, with \( E' = E \), so that \( D_i h \equiv h'. \)) Do it in the “variable” notation, too.

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) \( u = g(x^p, y^r, z^s). \)

Then redo all in the “mapping” terminology, too.

11. Let the map \( g: E^1 \to E^1 \) be differentiable on \( E^1 \). Find \(|\nabla h(\tilde{p})|\) if \( h = g \circ f \) and

(i) \( f(\tilde{x}) = \sum_{k=1}^{n} x_k, \tilde{x} \in E^n; \)

(ii) \( f = (f_1, f_2), f_1(\tilde{x}) = \sum_{k=1}^{n} x_k, f_2(\tilde{x}) = |\tilde{x}|^2, \tilde{x} \in E^n. \)

12. (Euler’s theorem.) A map \( f: E^n \to E^1 \) (or \( C^n \to C \)) is called homogeneous of degree \( m \) on \( G \) iff

\( (\forall t \in E^1 (C)) f(t\tilde{x}) = t^m f(\tilde{x}) \)

when \( \tilde{x}, t\tilde{x} \in G \). Prove the following statements.

(i) If so, and \( f \) is differentiable at \( \tilde{p} \in G \) (an open globe), then

\( \tilde{p} \cdot \nabla f(\tilde{p}) = m f(\tilde{p}). \)

* (ii) Conversely, if the latter holds for all \( \tilde{p} \in G \) and \( \tilde{b} \notin G \), then \( f \) is homogeneous of degree \( m \) on \( G. \)

(iii) What if \( \tilde{b} \in G? \)

[Hints: (i) Let \( g(t) = f(t\tilde{p}). \) Find \( g'(t). \) (ii) Take \( f(x, y) = x^2 y^2 \) if \( x \leq 0, \) \( f = 0 \) if \( x > 0, \) \( G = G_0(1). \)

13. Try Problem 12 for \( f: E' \to E \), replacing \( \tilde{p} \cdot \nabla f(\tilde{p}) \) by \( df(\tilde{p}) \).

14. With all as in Theorem 1, prove the following.

(i) If \( E' = E \) and \( s = f'(p) \neq 0, \) then \( h'(p) = D_s g(\tilde{q}). \)

(ii) If \( \tilde{u} \) and \( \tilde{v} \) are nonzero in \( E' \) and \( a\tilde{u} + b\tilde{v} \neq \tilde{0} \) for some scalars \( a, b, \) then

\( D_{a\tilde{u} + b\tilde{v}} f(\tilde{p}) = a D_{\tilde{v}} f(\tilde{p}) + b D_{\tilde{u}} f(\tilde{p}). \)

\[ \frac{\partial^2 f}{\partial x_1 \partial x_2}. \]

\[ D_{11} f = D_2^1 e_1 f = \frac{\partial^2 f}{\partial x_1 \partial x_2}, \]

\[ D_{12} f = D_2^1 e_2 f = \frac{\partial^2 f}{\partial x_1 \partial x_2}, \]

\[ D_{12} f = D_2^1 \tilde{e}_3 f = \frac{\partial^2 f}{\partial x_1 \partial x_2}, \]

\[ D_{12} f = D_2^1 \tilde{e}_4 f = \frac{\partial^2 f}{\partial x_1 \partial x_2}. \]

We also set \( D_0^0 f = f \) for any vector \( \tilde{u}. \)

Example.

(A) Define \( f: E^2 \to E^1 \) by

\[ f(0, 0) = 0, \]

\[ f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}. \]

\[ \frac{\partial^2 f}{\partial x_1 \partial x_2}, \]

\[ \frac{\partial^2 f}{\partial x_1 \partial x_2}. \]

\[ \frac{\partial^2 f}{\partial x_1 \partial x_2}, \]

\[ \frac{\partial^2 f}{\partial x_1 \partial x_2}. \]
Then
\[ \frac{\partial f}{\partial x} = D_1 f(x, y) = y \frac{x^4 + 4xy^2 - y^4}{(x^2 + y^2)^2}, \]
whence \( D_1 f(0, y) = -y \) if \( y \neq 0 \); and also
\[ D_1 f(0, 0) = \lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x} = 0. \] 
(Verify!)

Thus \( D_1 f(0, y) = -y \) always, and so \( D_{12} f(0, y) = -1; D_{12} f(0, 0) = -1. \)
Similary,
\[ D_2 f(x, y) = \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2} \]
if \( x \neq 0 \) and \( D_2 f(0, 0) = 0. \) Thus \( (\forall x) D_2 f(x, 0) = x \) and so
\[ D_{21} f(x, 0) = 1 \text{ and } D_{21} f(0, 0) = 1 \neq D_{12} f(0, 0) = -1. \]

The previous example shows that we may well have \( D_{12} f \neq D_{21} f \), or more generally, \( D_{12} f \neq D_{21} f \). However, we obtain the following theorem.

**Theorem 1.** Given nonzero vectors \( \vec{u} \) and \( \vec{v} \) in \( E' \), suppose \( f : E' \to E \) has the derivatives
\[ D_{12} f, D_{21} f, \text{ and } D_{22} f \]
on an open set \( A \subseteq E' \).

If \( D_{12} f \) is continuous at some \( \vec{p} \in A \), then the derivative \( D_{12} f(\vec{p}) \) also exists and equals \( D_{22} f(\vec{p}) \).

**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 \( G = G_0(\delta) \subseteq A \) and simultaneously
\[ \sup_{\vec{x} \in G} |D_{12} f(\vec{x}) - D_{12} f(\vec{p})| \leq \varepsilon \]
(by the continuity of \( D_{12} f \) at \( \vec{p} \)).

Now \( (\forall s, t \in E^1) \) define \( H_t : E^1 \to E \) by
\[ H_t(s) = D_{12} f(\vec{p} + t\vec{u} + s\vec{v}). \]
Let
\[ I = \left( -\frac{\delta}{2}, \frac{\delta}{2} \right). \]
If \( s, t \in I \), the point \( \vec{x} = \vec{p} + t\vec{u} + s\vec{v} \) is in \( G_0(\delta) \subseteq A \), since
\[ |\vec{x} - \vec{p}| = |t\vec{u} + s\vec{v}| < \frac{\delta}{2} + \frac{\delta}{2} = \delta. \]

Thus by assumption, the derivative \( D_{12} f(\vec{p}) \) exists. Also,
\[ H'_t(s) = \lim_{\Delta s \to 0} \frac{H_t(s + \Delta s) - H_t(s)}{\Delta s} \]
\[ = \lim_{\Delta s \to 0} \frac{D_{12} f(\vec{x} + \Delta s \cdot \vec{v}) - D_{12} f(\vec{x})}{|\Delta s|}. \]

But the last limit is \( D_{12} f(\vec{p}) \), by definition. Thus, setting
\[ h_t(s) = H_t(s) - s D_{12} f(\vec{p}), \]
we get
\[ h_t'(s) = H'_t(s) - D_{12} f(\vec{p}) = D_{12} f(\vec{x}) - D_{12} f(\vec{p}). \]
We see that \( h_t \) is differentiable on \( I \), and by (2),
\[ \sup_{s \in I} |h_t'(s)| \leq \sup_{\vec{x} \in G} |D_{12} f(\vec{x}) - D_{12} f(\vec{p})| \leq \varepsilon \]
for all \( t \in I \). Hence by Corollary 1 of Chapter 5, §4,
\[ |h_t(s) - h_t(0)| \leq |s| \sup_{s \in I} |h_t'(s)| \leq |s|\varepsilon. \]
But by definition,
\[ h_t(s) = D_{12} f(\vec{p} + t\vec{u} + s\vec{v}) - s D_{12} f(\vec{p}) \]
and
\[ h_t(0) = D_{12} f(\vec{p} + t\vec{u}). \]

Thus
\[ |D_{12} f(\vec{p} + t\vec{u} + s\vec{v}) - D_{12} f(\vec{p} + t\vec{u}) - s D_{12} f(\vec{p})| \leq |s|\varepsilon \]
for all \( s, t \in I \).
Next, set
\[ G_s(t) = f(\vec{p} + t\vec{u} + s\vec{v}) - f(\vec{p} + t\vec{u}) \]
and
\[ g_s(t) = G_s(t) - st \cdot D_{12} f(\vec{p}). \]
As before, one finds that \( (\forall s \in I) g_s \) is differentiable on \( I \) and that
\[ g'_s(t) = D_{12} f(\vec{p} + t\vec{u} + s\vec{v}) - D_{12} f(\vec{p} + t\vec{u}) - s D_{12} f(\vec{p}) \]
for \( s, t \in I \). (Verify!)

Hence by (3),
\[ \sup_{t \in I} |g'_s(t)| \leq |s|\varepsilon. \]
Again, by Corollary 1 of Chapter 5, §4, 
\[ |g_\epsilon (t) - g_\epsilon (0)| \leq |s||\epsilon|, \]
or by the definition of \( g_\epsilon \) (assuming \( s, t \in I \) \( \setminus \{0\} \) and dividing by \( st \)),
\[ \left| \frac{1}{st}[f(\tilde{p} + t\tilde{s}) - f(\tilde{p} + s\tilde{t})] - D_\tilde{s}^2 f(\tilde{p}) \right| \leq \epsilon. \] 
(Verify!) Making \( s \to 0 \) (with \( t \) fixed), we get, by the definition of \( D_\tilde{s} f \),
\[ \left| \frac{1}{t}D_\tilde{s} f(\tilde{p} + t\tilde{s}) - D_\tilde{s}^2 f(\tilde{p}) \right| \leq \epsilon \]
whenever \( 0 < |t| < \delta/2. \)
As \( \epsilon \) is arbitrary, we have
\[ D_\tilde{s}^2 f(\tilde{p}) = \lim_{t \to 0} \frac{1}{t}[D_\tilde{s} f(\tilde{p} + t\tilde{s}) - D_\tilde{s} f(\tilde{p})]. \]
But by definition, this limit is the derivative \( D_\tilde{s}^2 f(\tilde{p}) \). Thus all is proved. □

**Note 1.** By induction, the theorem extends to derivatives of order \( > 2 \). Thus the derivative \( D_{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_n} f \) is independent of the order in which the \( \tilde{s}_j \) follow each other if it exists and is continuous on an open set \( A \subseteq E' \), along with appropriate derivatives of order \( < k \).

If \( E' = E^n(C) \), this applies to **partials** as a special case.

For \( E^n \) and \( C^n \) only, we also formulate the following definition.

**Definition 1.**
Let \( E' = E^n(C) \). We say that \( f: E' \to E \) is \( m \) **times differentiable** at \( \tilde{p} \in E' \) iff \( f \) and all its partials of order \( < m \) are differentiable at \( \tilde{p} \).

If this holds for all \( \tilde{p} \) 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 \( \tilde{p} \) (on \( B \)), we say that \( f \) is **of class** \( CD^m \), or **continuously differentiable** \( m \) times there, and write \( f \in CD^m \) at \( \tilde{p} \) (on \( B \)).

Finally, if this holds for all natural \( m \), we write \( f \in CD^\infty \) at \( \tilde{p} \) (on \( B \), respectively).

**Definition 2.**
Given the space \( E' = E^n(C) \), the function \( f: E' \to E \), and a point \( \tilde{p} \in E' \), we define the mappings
\[ D^mf(\tilde{p}, \cdot), \ m = 1, 2, \ldots, \]
from \( E' \) to \( E \) by setting for every \( \tilde{t} = (t_1, \ldots, t_n) \)
\[ d^1 f(\tilde{p}, \tilde{t}) = \sum_{i=1}^n D_i f(\tilde{p}) \cdot t_i, \]
\[ d^2 f(\tilde{p}, \tilde{t}) = \sum_{i=1}^n \sum_{j=1}^n D_{ij} f(\tilde{p}) \cdot t_i t_j, \]
\[ d^m f(\tilde{p}, \tilde{t}) = \sum_{k=1}^n \sum_{s=1}^n \sum_{t=1}^n D_{skt} f(\tilde{p}) \cdot t_i t_j t_k, \]
and so on.

We call \( d^m (\tilde{p}, \cdot) \) the \( m \)th differential (or differential of order \( m \)) of \( f \) at \( \tilde{p} \). By our conventions, it is always defined on \( E^n(C) \) as are the partially derived functions involved.

If \( f \) is differentiable at \( \tilde{p} \) (but not otherwise), then \( d^1(\tilde{p}, \tilde{t}) = df(\tilde{p}, \tilde{t}) \) by Corollary 1 in §3; \( d^1(\tilde{p}, \cdot) \) 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.,
\[ d^2 f = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \, dx_i \, dx_j. \]

**Note 2.** Classical analysis tends to define differentials as above in terms of partials. Formula (4) for \( d^m f \) is often written **symbolically**:
\[ d^m f = \left( \frac{\partial}{\partial x_1} \, dx_1 + \frac{\partial}{\partial x_2} \, dx_2 + \cdots + \frac{\partial}{\partial x_n} \, dx_n \right)^m f, \quad m = 1, 2, \ldots. \]

Indeed, raising the bracketed expression to the \( m \)th “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 \( f: E^2 \to E^1 \) by
\[ f(x, y) = x \sin y. \]

Take any \( \tilde{p} = (x, y) \in E^2 \). Then
\[ D_1 f(x, y) = \sin y \text{ and } D_2 f(x, y) = x \cos y; \]
\[ D_{12} f(x, y) = D_{21} f(x, y) = \cos y, \]
\[ D_{11} f(x, y) = 0, \text{ and } D_{22} f(x, y) = -x \sin y; \]
\[ D_{111} f(x, y) = D_{112} f(x, y) = D_{121} f(x, y) = D_{211} f(x, y) = 0, \]
\[ D_{221} f(x, y) = D_{212} f(x, y) = D_{222} f(x, y) = -\sin y, \text{ and} \]
As is easily seen, \( f \) has continuous partials of all orders; so \( f \in CD^\infty \) on all of \( E^2 \). Also, 
\[
d f(\vec{p}, \vec{u}) = t_1 D_1 f(\vec{p}) + t_2 D_2 f(\vec{p})
\]
\[
= t_1 \sin y + t_2 x \cos y.
\]

In classical notation, 
\[
d f = d^1 f = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = \sin y dx + x \cos y dy;
\]
\[
d^2 f = \frac{\partial^2 f}{\partial x^2} dx^2 + 2 \frac{\partial^2 f}{\partial x \partial y} dx dy + \frac{\partial^2 f}{\partial y^2} dy^2 = 2 \cos y dx dy - x \sin y dy^2;
\]
\[
d^3 f = -3 \sin y dx dy^2 - x \cos y dy^3;
\]
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 2** (Taylor). Let \( \vec{u} = \vec{x} - \vec{p} \neq \vec{0} \) in \( E' = E^n \left( C^n \right) \).

If \( f : E' \to E \) is \( m + 1 \) times differentiable on the line segment 
\[ I = L[\vec{p}, \vec{x}] \subset E' \]
then
\[
f(\vec{x}) - f(\vec{p}) = \sum_{i=1}^m \frac{1}{i!} d^i f(\vec{p}, \vec{u}) + R_m,
\]
with
\[
|R_m| \leq \frac{K_m}{(m + 1)!}, \quad K_m \in E^1.
\]
and
\[
0 \leq K_m \leq \sup_{\vec{u} \in I} |d^{m+1} f(\vec{x}; \vec{u})|.
\]

**Proof.** Define \( g : E^3 \to E' \) and \( h : E^3 \to E \) 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 \( g \),
\[
g_k(t) = p_k + t u_k, \quad k \leq n.
\]
Clearly, \( g_k \) is differentiable, \( g_k'(t) = u_k \).

§5. Repeated Differentiation. Taylor’s Theorem

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,
\[
\vec{p} + t\vec{u} \in L[\vec{p}, \vec{x}] \iff t \in [0, 1].
\]

By **Theorem 2** in §4,
\[
h'(t) = \sum_{k=1}^n D_k f(\vec{p} + t\vec{u}) \cdot u_k = d f(\vec{p} + t\vec{u}; \vec{u}), \quad t \in J.
\]

(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
\[
h''(t) = \sum_{j=1}^n \sum_{k=1}^n D_{kj} f(\vec{p} + t\vec{u}) \cdot u_k u_j = d^2 f(\vec{p} + t\vec{u}; \vec{u}), \quad t \in J.
\]

By induction, \( h \) is \( m + 1 \) times differentiable on \( J \), and
\[
h^{(i)}(t) = d^i f(\vec{p} + t\vec{u}; \vec{u}), \quad t \in J, \quad i = 1, 2, \ldots, m + 1.
\]

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 \( Q = 0 \)); hence
\[
(9) \quad h(1) - h(0) = \sum_{i=1}^m \frac{h^{(i)}(0)}{i!} + R_m,
\]
\[
|R_m| \leq \frac{K_m}{(m + 1)!}, \quad K_m \in E^1,
\]
\[
K_m \leq \sup_{t \in J} |h^{(m+1)}(t)|.
\]

By construction,
\[
h(t) = f(g(t)) = f(\vec{p} + t\vec{u});
\]
so
\[
h(1) = f(\vec{p} + \vec{u}) = f(\vec{x}) \text{ and } h(0) = f(\vec{p}).
\]

Thus using (8) also, we see that (9) implies (6), indeed. □

**Note 3.** Formula (3’) of Chapter 5, §6, combined with (8), also yields
\[
R_m = \frac{1}{m!} \int_0^1 h^{(m+1)}(t) \cdot (1 - t)^m \, dt
\]
\[
= \frac{1}{m!} \int_0^1 d^{m+1} f(\vec{p} + t\vec{u}; \vec{u}) \cdot (1 - t)^m \, dt.
\]
Corollary 1 (the Lagrange form of \( R_m \)). If \( E = E^1 \) in Theorem 2, then

\[
R_m = \frac{1}{(m+1)!} d^{m+1} f(\xi; \vec{u})
\]

for some \( \xi \in L(\vec{p}, \vec{x}) \).

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! \( \square \)

Corollary 2. If \( f: E^n (C^n) \to E \) 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 \( D_m^m f(\vec{p}) \) 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,

\[
D_2 f(\vec{p}) = \lim_{s \to 0} \frac{1}{s} [f(\vec{p} + s\vec{u}) - f(\vec{p})]
\]

\[
= \lim_{s \to 0} \frac{1}{s} [h(s) - h(0)]
\]

\[
= h'(0) = df(\vec{p}; \vec{u})
\]

by (7). Induction yields

\[
D_m^m f(\vec{p}) = h^{(m)}(0) = d^m(\vec{p}; \vec{u})
\]

by (8). (See Problem 3.)

Example.

(C) Continuing Example (B), fix

\[
\vec{p} = (1, 0);
\]

thus replace \((x, y)\) by \((1, 0)\) there. Instead, write \((x, y)\) for \( \vec{x} \) in Theorem 2. Then

\[
\vec{u} = \vec{x} - \vec{p} = (x - 1, y);
\]

so

\[
u_1 = x - 1 = dx \text{ and } v_2 = y = dy,
\]

and we obtain

\[
df(\vec{p}; \vec{u}) = \langle \vec{s}, \vec{u} \rangle = D_1 f(1, 0) \cdot (x - 1) + D_2 f(1, 0) \cdot y
\]

\[
= (\sin 0) \cdot (x - 1) + (1 \cdot \cos 0) \cdot y
\]

\[
= y;
\]

\[
d^2 f(\vec{p}; \vec{u}) = D_{11} f(1, 0) \cdot (x - 1)^2 + 2 D_{12} f(1, 0) \cdot (x - 1)y + D_{22} f(1, 0) \cdot y^2
\]

\[
= (0) \cdot (x - 1)^2 + 2 (\cos 0) \cdot (x - 1)y - (1 \cdot \sin 0) \cdot y^2
\]

\[
= 2(x - 1)y;
\]

and for all \( \vec{s} = (s_1, s_2) \in I \),

\[
d^3 f(\vec{s}; \vec{u}) = D_{111} f(s_1, s_2) \cdot (x - 1)^3 + 3 D_{112} f(s_1, s_2) \cdot (x - 1)^2 y
\]

\[
+ 3 D_{122} f(s_1, s_2) \cdot (x - 1)y^2 + 2 D_{222} f(s_1, s_2) \cdot y^3
\]

\[
+ 3 \sin s_2 \cdot (x - 1)y^2 - s_1 \cos s_2 \cdot y^3.
\]

Hence by (6) and Corollary 1 (with \( m = 2 \)), noting that \( f(\vec{p}) = f(1, 0) = 0 \), we get

\[
f(x, y) = x \cdot \sin y
\]

\[
= y + (x - 1)y + R_2,
\]

where for some \( \vec{s} \in I \),

\[
R_2 = \frac{1}{3!} d^3 f(\vec{s}; \vec{u}) = \frac{1}{6} [-3 \sin s_2 \cdot (x - 1)y^2 - s_1 \cos s_2 \cdot y^3].
\]

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

\[
|s_1| \leq \max(|x|, 1) \leq |x| + 1.
\]

Finally, since \( |\sin s_2| \leq 1 \) and \( |\cos s_2| \leq 1 \), we obtain

\[
|R_2| \leq \frac{1}{6} \left[ 3|x - 1| + (|x| + 1)|y| \right] y^2.
\]

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

\[
\Delta f = \sum_{i=1}^{m} \frac{d^i f}{\vec{s}!} + R_2,
\]

generalizes formula (2) in Chapter 5, §6.

As in Chapter 5, §6, we set

\[
P_m(\vec{x}) = f(\vec{p}) + \sum_{i=1}^{m} \frac{1}{i!} d^i f(\vec{p}; \vec{x} - \vec{p})
\]

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

\[
u_k = x_k - p_k, \quad k = 1, \ldots, n,
\]

plus the remainder term \( R_m \).
If \( f \in CD^\infty \) on some \( G_\rho \) and if \( R_m \to 0 \) as \( m \to \infty \), we can express \( f(\bar{x}) \) as a convergent power series

\[
f(\bar{x}) = \lim_{m \to \infty} P_m(\bar{x}) = f(\bar{p}) + \sum_{i=1}^\infty \frac{1}{i!} d^i f(\bar{p}; \bar{x} - \bar{p}).
\]

We then say that \( f \) admits a Taylor series about \( \bar{p} \) on \( G_\rho \).

**Problems on Repeated Differentiation and Taylor Expansions**

1. Complete all details in the proof of Theorem 1. What is the motivation for introducing the auxiliary functions \( h_1 \) and \( g_k \) in this particular way?

2. Is symbolic “multiplication” in Note 2 always commutative? (See Example (A).) Why was it possible to collect “similar” terms

\[
\frac{\partial^2 f}{\partial x \partial y} \, dx \, dy \quad \text{and} \quad \frac{\partial^2 f}{\partial y \partial x} \, dy \, dx
\]

in Example (B)? Using (5), find the general formula for \( d^3 f \). Expand it!

3. Carry out the induction in Theorem 2 and Corollary 2. (Use a suitable notation for subscripts: \( k_1k_2 \ldots \) instead of \( jk \ldots \))

4. Do Example (C) with \( m = 3 \) (instead of \( m = 2 \)) and with \( \bar{p} = (0,0) \). Show that \( R_m \to 0 \), i.e., \( f \) admits a Taylor series about \( \bar{p} \).

Do it in the following two ways.

(i) Use Theorem 2.

(ii) Expand \( \sin y \) as in Problem 7(a) in Chapter 5, \( \S 6 \), and then multiply termwise by \( x \).

Give an estimate for \( R_4 \).

5. Use Theorem 2 to expand the following functions in powers of \( x - 3 \) and \( y + 2 \) exactly (choosing \( m \) so that \( R_m = 0 \)).

\[
\begin{align*}
(1) \quad f(x,y) &= 2xy^2 - 3y^3 + xy^2 - x^3; \\
(2) \quad f(x,y) &= x^4 - 3x^2y^2 + 2xy - 1; \\
(3) \quad f(x,y) &= x^5 - 3x^4y + axy^2 - x^3.
\end{align*}
\]

6. For the functions of Problem 15 in \( \S 4 \), give their Taylor expansions up to \( R_2 \), with

\[
\bar{p} = \left( \frac{1}{4}, \frac{3}{4}, 1 \right)
\]

in case (i) and

\[
\bar{p} = \left( e, \frac{\pi}{4}, e \right)
\]

in (ii). Bound \( R_2 \).

7. (Generalized Taylor theorem.) Let \( \bar{u} = \bar{x} - \bar{p} \neq 0 \) in \( E' \) (\( E' \) need not be \( E^n \) or \( C^m \)); let \( I = L[\bar{p}, \bar{x}] \). Prove the following statement:

If \( f: E' \to E \) and the derived functions \( D^i_k f \) (\( i \leq m \)) are relatively continuous on \( I \) and have \( \bar{u} \)-directed derivatives on \( I - Q \) (\( Q \) countable), then formula (6) and Note 3 hold, with \( d^i f(\bar{p}; \bar{u}) \) replaced by \( D^i_k f(\bar{p}) \).

[Hint: Proceed as in Theorem 2 without using the chain rule or any partials or components. Instead of (8), prove that \( h^{(i)}(t) = D^i_k f(\bar{p} + t\bar{u}) \) on \( J - Q' \), \( Q' = g^{-1}[Q] \).]

8. (i) Modify Problem 7 by setting

\[
\bar{u} = \frac{\bar{x} - \bar{p}}{\| \bar{x} - \bar{p} \|}.
\]

Thus expand \( f(\bar{x}) \) in powers of \( \| \bar{x} - \bar{p} \| \).

(ii) Deduce Theorem 2 from Problem 7, using Corollary 2.

9. Given \( f: E' \to E \), \( f \in CD^m \) on an open set \( A \), \( \bar{s} \in A \), prove that \( \bar{p} \in E' \) (\( C^2 \))

\[
d^i f(\bar{s}; \bar{u}) = \sum_{j=0}^i \binom{i}{j} u_1^{j} u_2^{i-j} D_{k_1 \ldots k_i} f(\bar{s}), \quad 1 \leq i \leq m,
\]

where the \( \binom{i}{j} \) are binomial coefficients, and in the \( j \)th term,

\[
k_1 = k_2 = \cdots = k_j = 2
\]

and

\[
k_{j+1} = \cdots = k_i = 1.
\]

Then restate formula (6) for \( n = 2 \). [Hint: Use induction, as in the binomial theorem.]

10. Given \( \bar{p} \in E' = E^n (C^m) \) and \( f: E' \to E \), prove that \( f \in CD^1 \) at \( \bar{p} \) iff \( f \) is differentiable at \( \bar{p} \) and

\[
(\forall \varepsilon > 0) \quad (\exists \delta > 0) \quad (\forall \bar{x} \in G_\rho(\delta)) \quad \| d^i f(\bar{p}; \bar{v}) - d^i f(\bar{x}; \bar{v}) \| < \varepsilon,
\]

with norm \( \| \| \) as in Definition 2 in \( \S 2 \). (Does it apply?)

[Hint: If \( f \in CD^1 \), use Theorem 2 in \( \S 3 \). For the converse, verify that

\[
\varepsilon > |d^i f(\bar{p}; \bar{v}) - d^i f(\bar{x}; \bar{v})| = \left| \sum_{k=1}^m |D_k f(\bar{p}) - D_k f(\bar{x})| v_k \right|
\]

if \( \bar{x} \in G_\rho(\delta) \) and \( |\bar{v}| \leq 1 \). Take \( \bar{v} = \bar{e}_k \), to prove continuity of \( D_k f \) at \( \bar{p} \).]

11. Prove the following.

(i) If \( \phi: E^n \to E^m \) is linear and \( [\phi] = (v_{ik}) \), then

\[
\| \phi \|^2 = \sum_{i,k} |v_{ik}|^2.
\]
(ii) If \( f : E^n \to E^m \) is differentiable at \( \bar{p} \), then
\[
\|df(\bar{p}, \cdot)\|^2 \leq \sum_{i,k} |D_{ik} f_i(\bar{p})|^2.
\]

(iii) Hence find a new converse proof in Problem 10 for \( f : E^n \to E^m \).
Consider \( f : C^n \to C^m \), too.

[Hints. (i) By the Cauchy–Schwarz inequality, \( |\phi(\bar{z})|^2 \leq |\bar{z}|^2 \sum_{i,k} |\phi_{ik}|^2 \). (Why?)
(ii) Use part (i) and Theorem 4 in §3.

12. (i) Find \( d^2 u \) for the functions of Problem 10 in §4, in the “variable” and “mapping” notations.

(ii) Do it also for
\[ u = f(x, y, z) = (x^2 + y^2 + z^2)^{-\frac{1}{2}} \]
and show that \( D_{11} f + D_{22} f + D_{33} f = 0 \).

(iii) Does the latter hold for \( u = \arctan \frac{y}{x} \)?

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 \( g \in C^2 \), express \( \frac{\partial^2 u}{\partial r \partial \theta} \) and \( \frac{\partial^2 u}{\partial \theta^2} \) as in (i).

14. Let \( f, g : E^1 \to E^1 \) be of class \( C^2 \) 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
\[ h(x, y) = f(ax + y) + g(y - ax). \]

(ii) \( x^2 D_{11} h(x, y) + 2xy D_{12} h(x, y) + y^2 D_{22} h(x, y) = 0 \) if
\[ h(x, y) = xf \left( \frac{y}{x} \right) + g \left( \frac{y}{x} \right). \]

(iii) \( D_1 h \cdot D_{21} h = D_2 h \cdot D_{11} h \) if
\[ h(x, y) = g(f(x) + y). \]
Find \( D_{12} h \), too.

§5. Repeated Differentiation. Taylor’s Theorem

15. Assume \( E' = E^n (C^m) \) and \( E'' = E^m (C^m) \). Let \( f : E' \to E'' \) and
\( g : E'' \to E \) be twice differentiable at \( \bar{p} \in E' \) and \( \bar{q} = f(\bar{p}) \in E'' \), respectively, and set \( h = g \circ f \).
Show that \( h \) is twice differentiable at \( \bar{p} \), and
\[ d^2 h(\bar{p}; \bar{r}) = d^2 g(\bar{q}; \bar{s}) + dg(\bar{q}; \bar{v}). \]
where \( \bar{r} \in E' \), \( \bar{s} = df(\bar{p}; \bar{r}) \), and \( \bar{v} = (v_1, \ldots, v_n) \in E'' \) satisfies
\[ v_i = df_i(\bar{p}; \bar{r}), \quad i = 1, \ldots, m. \]
Thus the second differential is not invariant in the sense of Note 4 in §4.

[Hint: Show that
\[ D_{ik} h(\bar{p}) = \sum_{j=1}^m \sum_{m} D_{ij} g(\bar{q}) D_{k} f_j(\bar{p}) + \sum_{m} D_{ig}(\bar{q}) D_{k} f_i(\bar{p}). \]
Proceed.]

16. Continuing Problem 15, prove the invariant rule:
\[ d^r h(\bar{p}; \bar{r}) = d^r g(\bar{q}; \bar{s}), \]
if \( f \) is a first-degree polynomial and \( g \) is \( r \) times differentiable at \( \bar{q} \).

[Hint: Here all higher-order partials of \( f \) vanish. Use induction.]


We assume the reader to be familiar with elements of linear algebra. Thus we only briefly recall some definitions and well-known rules.

Definition 1.

Given a linear operator \( \phi : E^n \to E^n \) (or \( \phi : C^n \to C^n \)), with matrix
\[ [\phi] = (v_{ik}) \quad i, k = 1, \ldots, n, \]
we define the determinant of \( [\phi] \) by
\[
\det[\phi] = \det(v_{ik}) = \begin{vmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\
v_{21} & v_{22} & \cdots & v_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
v_{n1} & v_{n2} & \cdots & v_{nn} \end{vmatrix}
= \sum (-1)^{i+k} v_{ik_1} v_{k_2} \cdots v_{kn_1}.
\]
where the sum is over all ordered \(n\)-tuples \((k_1, \ldots, k_n)\) of distinct integers \(k_j (1 \leq k_j \leq n)\), and

\[
\lambda = \begin{cases} 
0 & \text{if } \prod_{j<m} (k_m - k_j) > 0 \text{ and } \\
1 & \text{if } \prod_{j<m} (k_m - k_j) < 0.
\end{cases}
\]

Recall (Problem 12 in §2) that a set \(B = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}\) in a vector space \(E\) is a basis iff

(i) \(B\) spans \(E\), i.e., each \(\vec{v} \in E\) has the form

\[
\vec{v} = \sum_{i=1}^{n} a_i \vec{v}_i
\]

for some scalars \(a_i\), and

(ii) this representation is unique.

The latter is true iff the \(\vec{v}_i\) are independent, i.e.,

\[
\sum_{i=1}^{n} a_i \vec{v}_i = \vec{0} \iff a_i = 0, \; i = 1, \ldots, n.
\]

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 \(\phi, g: E^n \to E^n\) (or \(C^n \to C^n\)) are linear, then

\[
\det[\phi] \cdot \det[g] = \det([\phi] \cdot [g]) = \det[\phi \circ g]
\]

(see §2, Theorem 3 and Note 4).

(b) If \(\phi(\vec{x}) = \vec{x}\) (identity map), then \([\phi] = (v_{ik})\), where

\[
v_{ik} = \begin{cases} 
0 & \text{if } i \neq k \text{ and } \\
1 & \text{if } i = k;
\end{cases}
\]

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 2.

For any function \(f: E^n \to E^n\) (or \(f: C^n \to C^n\)), we define the \(f\)-induced Jacobian map \(J_f: E^n \to E^n\) (\(J_f: C^n \to C^n\)) by setting

\[
J_f(\vec{x}) = \det(v_{ik})
\]

where \(v_{ik} = D_k f_i(\vec{x})\), \(\vec{x} \in E^n (C^n)\), and \(f = (f_1, \ldots, f_n)\).

The determinant

\[
J_f(\vec{p}) = \det(D_k f_i(\vec{p}))
\]

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,

\[
J_f = \det(D_k f_i).
\]

By Definition 2 and Note 2 in §5,

\[
J_f(\vec{p}) = \det[d^1 f(\vec{p}), \cdot].
\]

If \(f\) is differentiable at \(\vec{p}\),

\[
J_f(\vec{p}) = \det[f'(\vec{p})].
\]

Note 1. More generally, given any functions \(v_{ik}: E' \to E^1(C)\), we can define a map \(f: E' \to E^1(C)\) by

\[
J_f(\vec{x}) = \det(v_{ik}(\vec{x}))
\]

briefly \(f = \det(v_{ik}), i, k = 1, \ldots, n\).

We then call \(f\) a functional determinant.

If \(E' = E^n (C^n)\) then \(f\) is a function of \(n\) variables, since \(\vec{x} = (x_1, x_2, \ldots, x_n)\).

If all \(v_{ik}\) are continuous or differentiable at some \(\vec{p} \in E'\), so is \(f\); for by (1), \(f\) is a finite sum of functions of the form

\[
(\vec{x}) = \sum_{j=1}^{n} a_i \vec{v}_j \vec{v}_{ik} \vec{v}_{jk} \vec{v}_{ik} \vec{v}_{ik},
\]

and each of these is continuous or differentiable if the \(v_{ik}\) 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 \(|J_f|\).

In classical notation, one writes

\[
\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \quad \text{or} \quad \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)}
\]

for \(J_f(\vec{x})\). Here \((y_1, \ldots, y_n) = f(x_1, \ldots, x_n)\).

The remarks made in §4 apply to this “variable” notation too. The chain rule easily yields the following corollary.

Corollary 1. If \(f: E^n \to E^n\) and \(g: E^n \to E^n\) (or \(f, g: C^n \to C^n\)) are differentiable at \(\vec{p}\) and \(\vec{q} = f(\vec{p})\), respectively, and if

\[
h = g \circ f,
\]

then

\[
J_h(\vec{p}) = J_f(\vec{p}) \cdot J_g(f(\vec{p})).
\]
Then 

\( J_h(\vec{p}) = J_g(\vec{q}) \cdot J_f(\vec{p}) = \det(z_{ik}), \)

where 

\[ z_{ik} = D_h h_i(\vec{p}), \quad i, k = 1, \ldots, n; \]

or, setting 

\[ (u_1, \ldots, u_n) = g(y_1, \ldots, y_n) \]\n\[ (y_1, \ldots, y_n) = f(x_1, \ldots, x_n) \] ("variables"),

we have 

\( \frac{\partial(u_1, \ldots, u_n)}{\partial(x_1, \ldots, x_n)} = \frac{\partial(y_1, \ldots, y_n)}{\partial(x_1, \ldots, x_n)} = \det(z_{ik}), \)

where 

\[ z_{ik} = \frac{\partial u_i}{\partial x_k} \quad i, k = 1, \ldots, n. \]

**Proof.** By Note 2 in §4,

\[ [h'(\vec{p})] = [g'(\vec{q})] \cdot [f'(\vec{p})]. \]

Thus by rule (a) above,

\[ \det[h'(\vec{p})] = \det[g'(\vec{q})] \cdot \det[f'(\vec{p})], \]

i.e.,

\[ J_h(\vec{p}) = J_g(\vec{q}) \cdot J_f(\vec{p}). \]

Also, if \([h'(\vec{p})] = (z_{ik}),\) Definition 2 yields \( z_{ik} = D_h h_i(\vec{p}). \)

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

\[ x = f_1(u, v) \quad \text{and} \quad y = f_2(u, v). \]

We then set \( f = (f_1, f_2) \) and obtain \( f: E^2 \to E^2, \)

\[ J_f = \det(D_k f_i), \quad k, i = 1, 2. \]

**Example** (passage to polar coordinates).

Let \( x = f_1(r, \theta) = r \cos \theta \) and \( y = f_2(r, \theta) = r \sin \theta. \)
where
\[ \vec{x} = \sum_{k=1}^{n} a_k \vec{e}_k \text{ (uniquely)}. \]

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
\[ \vec{x} = \sum_{k=1}^{n} x_k \vec{e}_k, \]
by Theorem 2 in Chapter 3, §1–3. Hence again
\[ \vec{y} = \sum_{k=1}^{n} x_k \phi (\vec{e}_k) = \sum_{k=1}^{n} x_k \vec{v}_k; \]
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
\[ \det [\phi] \cdot \det [\phi^{-1}] = \det [\phi \circ \phi^{-1}] = 1 \]
if \( \phi \) 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 \( \phi \) 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,
\[ \vec{v}_1 = \sum_{k=2}^{n} a_k \vec{v}_k. \]
But by linear algebra (Problem 13(iii)), \( \det [\phi] \) does not change if \( \vec{v}_1 \) is replaced by
\[ \vec{v}_1 - \sum_{k=2}^{n} a_k \vec{v}_k = \vec{0}, \]
Thus \( \det [\phi] = 0 \) (one column turning to \( \vec{0} \)). This completes the proof. \( \Box \)

**Note 3.** Maps that are both onto and one-to-one are called bijective. Such is \( \phi \) in Theorem 1. This means that the equation
\[ \phi (\vec{x}) = \vec{y} \]
has a unique solution
\[ \vec{x} = \phi^{-1} (\vec{y}) \]
for each \( \vec{y} \). Componentwise, by Theorem 1, the equations
\[ \sum_{k=1}^{n} x_k v_{ik} = y_i, \quad i = 1, \ldots, n, \]
have a unique solution for the \( x_k \) iff \( \det (v_{ik}) \neq 0 \).

**Corollary 2.** If \( \phi \in L(E', E) \) is bijective, with \( E' \) and \( E \) complete, then \( \phi^{-1} \in L(E, E') \).

**Proof for** \( E = E^n (C^n) \).\(^1\) The notation \( \phi \in L(E', E) \) means that \( \phi : E' \to E \) is linear and continuous.

As \( \phi \) is bijective, \( \phi^{-1} : E \to E' \) is linear (Problem 12).

If \( E = E^n (C^n) \), it is continuous, too (Theorem 2 in §2).

Thus \( \phi^{-1} \in L(E, E') \). \( \Box \)

**Note.** The case \( E = E^n (C^n) \) 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 2.** Let \( E, E' \) and \( \phi \) be as in Corollary 2. Set
\[ \| \phi^{-1} \| = \frac{1}{\varepsilon}. \]
Then any map \( \theta \in L(E', E) \) with \( \| \theta - \phi \| < \varepsilon \) is one-to-one, and \( \theta^{-1} \) is uniformly continuous.

**Proof.** By Corollary 2, \( \phi^{-1} \in L(E, E') \), so \( \| \phi^{-1} \| \) is defined and \( > 0 \) (for \( \phi^{-1} \) is not the zero map, being one-to-one).

Thus we may set
\[ \varepsilon = \frac{1}{\| \phi^{-1} \|} = \frac{1}{\| \phi^{-1} \|}. \]

Clearly \( \vec{x} = \phi^{-1} (\vec{y}) \) if \( \vec{y} = \phi (\vec{x}) \). Also,
\[ | \phi^{-1} (\vec{y}) | \leq \frac{1}{\varepsilon} | \vec{y} | \]
by Note 5 in §2. Hence
\[ | \vec{y} | \geq \varepsilon | \phi^{-1} (\vec{y}) |, \]
i.e.,
\[ (2) \quad | \phi (\vec{x}) | \geq \varepsilon | \vec{x} | \]
for all \( \vec{x} \in E' \) and \( \vec{y} \in E \).

Now suppose \( \phi \in L(E', E) \) and \( \| \theta - \phi \| = \sigma < \varepsilon \).

Obviously, \( \theta = \phi - (\phi - \theta) \), and by Note 5 in §2,
\[ | (\phi - \theta) (\vec{x}) | \leq \| \phi - \theta \| | \vec{x} | = \sigma | \vec{x} |. \]
\(^1\) See **§8** for the general case.
Thus for every \( \vec{x} \in E' \),
\[
|\theta(\vec{x})| \geq |\phi(\vec{x})| - |(\phi - \theta)(\vec{x})|
\]
(3)
\[
\geq |\phi(\vec{x})| - \sigma|\vec{x}|
\geq (\varepsilon - \sigma)|\vec{x}|
\]
by (2). Therefore, given \( \vec{p} \neq \vec{r} \) in \( E' \) and setting \( \vec{x} = \vec{p} - \vec{r} \neq 0 \), we obtain
(4)
\[
|\theta(\vec{p}) - \theta(\vec{r})| = |\theta(\vec{p} - \vec{r})| = |\theta(\vec{x})| \geq (\varepsilon - \sigma)|\vec{x}| > 0
\]
(since \( \sigma < \varepsilon \)).

We see that \( \vec{p} \neq \vec{r} \) implies \( \theta(\vec{p}) \neq \theta(\vec{r}) \); so \( \theta \) is one-to-one, indeed.

Also, setting \( \theta(\vec{x}) = \vec{z} \) and \( \vec{x} = \theta^{-1}(\vec{z}) \) in (3), we get
\[
|\vec{z}| \geq (\varepsilon - \sigma)|\theta^{-1}(\vec{z})|;
\]
that is,
(5)
\[
|\theta^{-1}(\vec{z})| \leq (\varepsilon - \sigma)^{-1}|\vec{z}|
\]
for all \( \vec{z} \) in the range of \( \theta \) (domain of \( \theta^{-1} \)).

Thus \( \theta^{-1} \) is linearly bounded (by Theorem 1 in §2), hence uniformly continuous, as claimed. \( \square \)

**Corollary 3.** If \( E' = E = E^n (C^n) \) in Theorem 2 above, then for given \( \phi \) and \( \delta > 0 \), there always is \( \delta' > 0 \) such that
\[
\|\theta - \phi\| < \delta' \iff \|\theta^{-1} - \phi^{-1}\| < \delta.
\]

In other words, the transformation \( \phi \rightarrow \phi^{-1} \) is continuous on \( L(E) \), \( E = E^n (C^n) \).

**Proof.** First, since \( E' = E = E^n (C^n) \), \( \theta \) 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
\[
\|\theta^{-1}\| \leq \frac{1}{\varepsilon - \sigma}
\]
Also,
\[
\phi^{-1} \circ (\theta - \phi) \circ \theta^{-1} = \phi^{-1} - \theta^{-1}
\]
(see Problem 11).

Hence by Corollary 4 in §2, recalling that \( \|\phi^{-1}\| = 1/\varepsilon \), we get
\[
\|\theta^{-1} - \phi^{-1}\| \leq \|\phi^{-1}\| \cdot \|\theta - \phi\| \cdot \|\theta^{-1}\| \leq \frac{\sigma}{\varepsilon(\varepsilon - \sigma)} \rightarrow 0 \text{ as } \sigma \rightarrow 0. \quad \square
\]

**Problems on Bijective Linear Maps and Jacobians**

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 \( J' \) be continuous or differentiable if \( f \) is?

   Give proofs or counterexamples.

   \( \Rightarrow 2. \) Prove rule (b) on determinants. More generally, show that if \( f(\vec{x}) = \vec{x} \) on an open set \( A \subseteq E^n (C^n) \), then \( J_f = 1 \) on \( A \).

2. Let \( f : E^n \rightarrow E^n \) (or \( C^n \rightarrow C^n \)), \( f = (f_1, \ldots, f_n) \).

   Suppose each \( f_k \) depends on \( x_k \) only, i.e.,
   \[
f_k(\vec{x}) = f_k(\vec{y}) \quad \text{if } x_k = y_k,
\]
   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 \( i \neq k \).]

3. In Corollary 1, show that
\[
J_h(p) = \prod_{k=1}^n D_k f_k(p) \cdot J_h(q)
\]
if \( f \) also has the property specified in Problem 3. Then do all in “variables,” with \( y_k = y_k(x_k) \) instead of \( f_k \).

4. Let \( E' = E^1 \) in Note 1. Prove that if all the \( v_{ik} \) are differentiable at \( p \), then \( f'(p) \) is the sum of \( n \) determinants, each arising from \( \det(v_{ik}) \), by replacing the terms of one column by their derivatives.

   [Hint: Use Problem 6 in Chapter 5, §1.]

5. Do Problem 5 for partials of \( f \), with \( E' = E^m (C^n) \), and for directionals \( D_k f \), in any normed space \( E' \). (First, prove formulas analogous to Problem 6 in Chapter 5, §1; use Note 3 in §1.) Finally, do it for the differential, \( df(\vec{p}, \cdot) \).

6. 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).


   55
We can always make \( \delta \) by reducing \( \frac{\lambda}{\| \varphi \|} \) (passage to cylindrical coordinates)?

(ii) What if \( x = r \cos \theta, \ y = r \sin \theta, \) and \( z = z \) remains unchanged (passage to cylindrical coordinates)?

(iii) Same for \( x = e^r \cos \theta, \ y = e^r \sin \theta, \) and \( z = z \).

9. If \( f = (f_1, f_2): E^2 \to E^2 \) 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 \( f_2(x, y) = 2xy \).

10. Define \( f: E^3 \to E^3 \) (or \( C^3 \to C^3 \)) by
    \[
    f(\vec{x}) = \frac{\vec{x}}{1 + \sum_{k=1}^{3} x_k^2}
    \]
    on
    \[
    A = \left\{ \vec{x} \mid 3 \sum_{k=1}^{3} x_k^2 \neq -1 \right\}
    \]
    and \( f = \vec{0} \) on \( -A \). Prove the following.
    (i) \( f \) is one-to-one on \( A \) (find \( f^{-1} \)).
    (ii) \( J_f(\vec{x}) = \frac{1}{\left( 1 + \sum_{k=1}^{3} \vec{x}_k \right)^2} \).
    (iii) Describe \( -A \) geometrically.

11. Given any sets \( A, B \) and maps \( f, g: A \to E', h: E' \to E \), and \( k: B \to A \), prove that
    (i) \( (f \pm g) \circ k = f \circ k \pm g \circ k \), and
    (ii) \( h \circ (f \pm g) = h \circ f + h \circ g \) if \( h \) is linear.

   Use these distributive laws to verify that
   \[
   \phi^{-1} \circ (\theta - \phi) \circ \theta^{-1} = \phi^{-1} - \theta^{-1}
   \]
   in Corollary 3.

   [Hint: First verify the associativity of mapping composition.]

12. Prove that if \( \phi: E' \to E \) is linear and one-to-one, so is \( \phi^{-1}: E'' \to E' \), where \( E'' = \phi[E'] \).

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 \( \lambda \) 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 \( \vec{v}_i \) is \( \vec{0} \), or if two of the \( \vec{v}_k \) are the same.

§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 \( f: E' \to E \) is continuously differentiable, or of class \( CD^1 \) (written \( f \in CD^1 \)), at \( \vec{p} \) iff the following statement is true:

Given any \( \varepsilon > 0 \), there is \( \delta > 0 \) such that \( f \) is differentiable on the globe \( \overline{C} = \overline{G}[\theta] \), with

\[
\| df(\vec{x}, \cdot) - df(\vec{p}, \cdot) \| < \varepsilon \quad \text{for all} \quad \vec{x} \in \overline{C}.
\]

By Problem 10 in §5, this definition agrees with Definition 1 of §5, but is no longer limited to the case \( E' = E^n (C^n) \). See also Problems 1 and 2 below.

We now obtain the following result.

Theorem 1. Let \( E' \) and \( E \) be complete. If \( f: E' \to E \) 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 \( \overline{C} = \overline{G}[\theta] \).

Thus \( f \) “locally” resembles \( df(\vec{p}, \cdot) \) in this respect.

Proof. Set \( \phi = df(\vec{p}, \cdot) \) and

\[
\| \phi^{-1} \| = \frac{1}{\varepsilon}
\]

(cf. Theorem 2 of §6).

By Definition 1, fix \( \delta > 0 \) so that for \( \vec{x} \in \overline{C} \), \( \| df(\vec{x}, \cdot) - \phi(\vec{u}) \| < \frac{1}{2} \varepsilon \).

Then by Note 5 in §2,

(1) \( \forall \vec{x} \in \overline{C} \) \( \forall \vec{u} \in E' \), \( \| df(\vec{x}, \vec{u}) - \phi(\vec{u}) \| < \frac{1}{2} \varepsilon \| \vec{u} \| \).

Now fix any \( \vec{r}, \vec{\bar{x}} \in \overline{C}, \vec{r} \neq \vec{\bar{x}}, \) and set \( \vec{u} = \vec{r} - \vec{\bar{x}} \neq \vec{0} \). Again, by Note 5 in §2,

\[
|\vec{u}| = |\phi^{-1}(\phi(\vec{u}))| \leq \| \phi^{-1} \| |\phi(\vec{u})| = \frac{1}{\varepsilon} |\phi(\vec{u})|.
\]

\footnote{We can always make \( \overline{C} \) closed by reducing \( \delta \).}
so
\[ 0 < \varepsilon |\bar{a}| \leq |\phi(\bar{u})|. \]

By convexity, \( \overline{G} \supseteq I = [\bar{s}, \bar{r}] \), so (1) holds for \( \bar{x} \in I \), \( \bar{x} = \bar{s} + t\bar{u} \), \( 0 \leq t \leq 1 \). Noting this, set
\[ h(t) = f(\bar{s} + t\bar{u}) - t\phi(\bar{u}), \quad t \in E^1. \]

Then for \( 0 \leq t \leq 1 \),
\[ h'(t) = D_{\bar{u}}f(\bar{s} + t\bar{u}) - \phi(\bar{u}). \]

(Verify!) Thus by (1) and (2),
\[ \sup_{0 \leq t \leq 1} |h'(t)| = \sup_{0 \leq t \leq 1} |D_{\bar{u}}f(\bar{s} + t\bar{u}) - \phi(\bar{u})| \leq \frac{\varepsilon}{2} |\bar{u}| \leq \frac{1}{2} |\phi(\bar{u})|. \]

(Explain!) Now, by Corollary 1 in Chapter 5, §4,
\[ |h(1) - h(0)| \leq (1 - 0) \cdot \sup_{0 \leq t \leq 1} |h'(t)| \leq \frac{1}{2} |\phi(\bar{u})|. \]

As \( h(0) = \bar{s} \) and
\[ h(1) = f(\bar{s} + \bar{u}) - \phi(\bar{u}) = f(\bar{r}) - \phi(\bar{u}), \]
we obtain (even if \( \bar{r} = \bar{s} \))
\[ |f(\bar{r}) - f(\bar{s})| \leq \frac{1}{2} |\phi(\bar{u})| \quad (\bar{r}, \bar{s} \in \overline{G}, \quad \bar{u} = \bar{r} - \bar{s}). \]

But by the triangle law,
\[ |\phi(\bar{u})| - |f(\bar{r}) - f(\bar{s})| \leq |f(\bar{r}) - f(\bar{s}) - \phi(\bar{u})|. \]

Thus
\[ |f(\bar{r}) - f(\bar{s})| \geq \frac{1}{2} |\phi(\bar{u})| \geq \frac{1}{2} \varepsilon |\bar{u}| = \frac{1}{2} \varepsilon |\bar{r} - \bar{s}| \]
by (2).

Hence \( f(\bar{r}) \neq f(\bar{s}) \) whenever \( \bar{r} \neq \bar{s} \) in \( \overline{G} \), so \( f \) is one-to-one on \( \overline{G} \), as claimed. \( \varnothing \)

**Corollary 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),
\[ |f(\bar{r}) - f(\bar{s})| \leq |\phi(\bar{u})| + \frac{1}{2} |\phi(\bar{u})| \leq 2|\phi| |\bar{r} - \bar{s}| \quad (\bar{r}, \bar{s} \in \overline{G}). \]

This implies uniform continuity for \( f \). (Why?)

Next, let \( g = f^{-1} \) on \( H = f[\overline{G}] \).

If \( \bar{x}, \bar{y} \in H \), let \( \bar{r} = g(\bar{x}) \) and \( \bar{s} = g(\bar{y}) \); so \( \bar{r}, \bar{s} \in \overline{G} \), with \( \bar{x} = f(\bar{r}) \) and \( \bar{y} = f(\bar{s}) \). Hence by (4),
\[ |\bar{x} - \bar{y}| \geq \frac{1}{2} \varepsilon |g(\bar{x}) - g(\bar{y})|, \]
proving all for \( g \), too. \( \varnothing \)

Again, \( f \) resembles \( \phi \) which is uniformly continuous, along with \( \phi^{-1} \).

**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 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** \( E' = E = E^n(\mathbb{C}^n) \) (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 \( X \in \text{Theorem 4 in Chapter 4, §6} \) and so is \( f[X] \) (Theorem 1 of Chapter 4, §8).

By Theorem 2 in Chapter 4, §6, \( f[X] \) is closed, as required. \( \varnothing \)

For the rest of this section, we shall set \( E' = E = E^n(\mathbb{C}^n) \).

**Theorem 2.** If \( E' = E = E^n(\mathbb{C}^n) \) in Theorem 1, with other assumptions unchanged, then \( f \) is open on the globe \( G = G_{\rho}(\delta) \), with \( \delta \) sufficiently small.\(^2\)

We first prove the following lemma.

\(^2\) Thus formula (1) still holds for \( \varepsilon = 1/\|\phi^{-1}\|, \phi = df(\bar{r}, \cdot) \).
Lemma. \( f[G] \) contains a globe \( G_M(\alpha) \) where \( q = f(\tilde{q}) \).

Proof. Indeed, let
\[
\alpha = \frac{1}{4} \delta,
\]
where \( \delta \) and \( \varepsilon \) are as in the proof of Theorem 1. (We continue the notation and formulas of that proof.)

Fix any \( \tilde{c} \in G_M(\alpha) \); so
\[
|\tilde{c} - \tilde{q}| < \alpha = \frac{1}{4} \delta.
\]
Set \( h = |f - \tilde{c}| \) on \( E' \). As \( f \) is uniformly continuous on \( \mathbb{H} \), so is \( h \).

Now, \( \mathbb{H} \) is compact in \( E^n(C^m) \); so Theorem 2(ii) in Chapter 4, §8, yields a point \( \tilde{r} \in \mathbb{H} \) such that
\[
(6) \quad h(\tilde{r}) = \min h(\mathbb{H}).
\]
We claim that \( \tilde{r} \) is in \( G \) (the interior of \( \mathbb{H} \)).

Otherwise, \( |\tilde{r} - \tilde{p}| = \delta \); for by (4),
\[
2\alpha = \frac{1}{2} \varepsilon \delta = \frac{1}{2} \varepsilon |\tilde{r} - \tilde{p}| \leq |f(\tilde{r}) - f(\tilde{p})| \leq |f(\tilde{r}) - \tilde{q}| + |\tilde{q} - f(\tilde{p})| = h(\tilde{r}) + h(\tilde{p}).
\]
But
\[
(7) \quad h(\tilde{p}) = |\tilde{c} - f(\tilde{p})| = |\tilde{c} - \tilde{q}| < \alpha;
\]
and so (7) yields
\[
h(\tilde{p}) < \alpha < h(\tilde{r}),
\]
contrary to the minimality of \( h(\tilde{r}) \) (see (6)). Thus \( |\tilde{r} - \tilde{p}| \) cannot equal \( \delta \).

We obtain \( |\tilde{r} - \tilde{p}| < \delta \), so \( \tilde{r} \in G_M(\delta) = \mathbb{H} \) and \( f(\tilde{r}) \in f[G] \). We shall now show that \( \tilde{c} = f(\tilde{r}) \).

To this end, we set \( \tilde{v} = \tilde{c} - f(\tilde{r}) \) and prove that \( \tilde{v} = 0 \). Let
\[
\tilde{u} = \phi^{-1}(\tilde{v}),
\]
where
\[
\phi = df(\tilde{p}, \cdot),
\]
as before. Then
\[
\tilde{v} = \phi(\tilde{u}) = df(\tilde{p}; \tilde{u}).
\]
With \( \tilde{r} \) as above, fix some
\[
\tilde{s} = \tilde{r} + t\tilde{u} \quad (0 < t < 1)
\]
with \( t \) so small that \( \tilde{s} \in \mathbb{H} \) also. Then by formula (3),
\[
|f(\tilde{s}) - f(\tilde{r}) - \phi(t\tilde{u})| \leq \frac{1}{2}|t\tilde{v}|
\]
also,
\[
|f(\tilde{r}) - \tilde{c} + \phi(t\tilde{u})| = (1 - t)|\tilde{v}| = (1 - t)h(\tilde{r})
\]
by our choice of \( \tilde{v}, \tilde{u} \) and \( h \). Hence by the triangle law,
\[
h(\tilde{s}) = |f(\tilde{s}) - \tilde{c}| \leq \left(1 - \frac{1}{2}\right) h(\tilde{r}).
\]
(Verify!)

As \( 0 < t < 1 \), this implies \( h(\tilde{r}) = 0 \) (otherwise, \( h(\tilde{s}) < h(\tilde{r}) \), violating (6)).

Thus, indeed,
\[
|\tilde{v}| = |f(\tilde{r}) - \tilde{c}| = 0,
\]
i.e.,
\[
\tilde{c} = f(\tilde{r}) \in f[G] \quad \text{for} \quad \tilde{r} \in \mathbb{H}.
\]
But \( \tilde{c} \) was an arbitrary point of \( G_M(\alpha) \). Hence
\[
G_M(\alpha) \subseteq f[G],
\]
proving the lemma. \( \square \)

Proof of Theorem 2. The lemma shows that \( f(\tilde{p}) \) is in the interior of \( f[G] \) if \( \tilde{p}, f, df(\tilde{p}, \cdot), \) and \( \delta \) are as in Theorem 1.

But Definition 1 implies that here \( f \in C D^1 \) on all of \( G \) (see Problem 1).

Also, \( df(\tilde{x}, \cdot) \) is bijective for any \( \tilde{x} \in \mathbb{H} \) by our choice of \( G \) and Theorems 1 and 2 in §6.

Thus \( f \) maps all \( \tilde{x} \in \mathbb{H} \) onto interior points of \( f[G] \); i.e., \( f \) maps any open set \( X \subseteq \mathbb{H} \) onto an open \( f[X] \), as required.

Note 1. A map
\[
f : (S, \rho) \to (T, \rho')
\]
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 \( \phi = df(\tilde{p}, \cdot) \) in Theorem 1 is "clopen" on all of \( E' \).

Again, \( f \) locally resembles \( df(\tilde{p}, \cdot) \).

III. The Inverse Function Theorem. We now further pursue these ideas.

Theorem 3 (inverse functions). Under the assumptions of Theorem 2, let \( g \) be the inverse of \( f \) (restricted to \( G = G_M(\delta) \)).

Then \( g \in C D^1 \) on \( f[G] \) and \( dg(\tilde{y}, \cdot) \) is the inverse of \( df(\tilde{x}, \cdot) \) whenever \( \tilde{x} = g(\tilde{y}), \tilde{x} \in \mathbb{H} \).
Briefly: “The differential of the inverse is the inverse of the differential.”

Proof. Fix any $\bar{y} \in f[G]$ and $\bar{x} = g(\bar{y})$; so $\bar{y} = f(\bar{x})$ and $\bar{x} \in G$. Let $U = df(\bar{x}, \cdot)$. As noted above, $U$ is bijective for every $\bar{x} \in G$ by Theorems 1 and 2 in §6; so we may set $V = U^{-1}$. We must show that $V = dg(\bar{y}, \cdot)$.

To do this, give $\bar{y}$ an arbitrary (variable) increment $\Delta \bar{y}$, so small that $\bar{y} + \Delta \bar{y}$ stays in $f[G]$ (an open set by Theorem 2).

As $g$ and $f_G$ are one-to-one, $\Delta \bar{y}$ uniquely determines

$$\Delta \bar{x} = g(\bar{y} + \Delta \bar{y}) - g(\bar{y}) = \bar{t},$$

and vice versa:

$$\Delta \bar{y} = f(\bar{x} + \bar{t}) - f(\bar{x}).$$

Here $\Delta \bar{y}$ and $\bar{t}$ are the mutually corresponding increments of $\bar{y} = f(\bar{x})$ and $\bar{x} = g(\bar{y})$. By continuity, $\bar{y} \rightarrow \bar{0}$ iff $\bar{t} \rightarrow \bar{0}$.

As $U = df(\bar{x}, \cdot)$,

$$\lim_{\bar{t} \rightarrow \bar{0}} \frac{1}{|\bar{t}|} |f(\bar{x} + \bar{t}) - f(\bar{t}) - U(\bar{t})| = 0,$$

or

$$\lim_{\bar{t} \rightarrow \bar{0}} \frac{1}{|\bar{t}|} |F(\bar{t})| = 0,$$

where

$$F(\bar{t}) = f(\bar{x} + \bar{t}) - f(\bar{t}) - U(\bar{t}).$$

As $V = U^{-1}$, we have

$$V(U(\bar{t})) = \bar{t} = g(\bar{y} + \Delta \bar{y}) - g(\bar{y}).$$

So from (9),

$$V(F(\bar{t})) = V(\Delta \bar{y}) = \bar{t} = g(\bar{y} + \Delta \bar{y}) - g(\bar{y});$$

that is,

$$\frac{1}{|\Delta \bar{y}|} |g(\bar{y} + \Delta \bar{y}) - g(\bar{y}) - V(\Delta \bar{y})| \leq \frac{|V(F(\bar{t}))|}{|\Delta \bar{y}|}, \quad |\Delta \bar{y}| \neq 0.$$

Now, formula (4), with $\bar{r} = \bar{x}$, $\bar{s} = \bar{x} + \bar{t}$, and $\bar{u} = \bar{t}$, shows that

$$|f(\bar{x} + \bar{t}) - f(\bar{x})| \geq \frac{1}{2} |\bar{t}|;$$

i.e., $|\Delta \bar{y}| \geq \frac{1}{2} |\bar{t}|$. Hence by (8),

$$|V(F(\bar{t}))| \leq \frac{|V(F(\bar{t}))|}{|\bar{t}|} = 2 \varepsilon |V \left( \frac{1}{|\bar{t}|} F(\bar{t}) \right)| \leq 2 \varepsilon |V\left( \frac{1}{|\bar{t}|} F(\bar{t}) \right)| \rightarrow 0 \text{ as } \bar{t} \rightarrow \bar{0}.$$

Since $\bar{t} \rightarrow \bar{0}$ as $\Delta \bar{y} \rightarrow \bar{0}$ (change of variables!), the expression (10) tends to 0 as $\Delta \bar{y} \rightarrow \bar{0}$.

By definition, then, $g$ is differentiable at $\bar{y}$, with $dg(\bar{y}, \cdot) = V = U^{-1}$.

Moreover, Corollary 3 in §6, applies here. Thus

$$(\forall \delta'' > 0) \left( \exists \delta' > 0 \right) \|U - W\| < \delta' \Rightarrow \|U^{-1} - W^{-1}\| < \delta'',$$

Taking here $U^{-1} = dg(\bar{y}, \cdot)$ and $W^{-1} = dg(\bar{y} + \Delta \bar{y})$, we see that $g \in C^1 \text{ near } \bar{y}$. This completes the proof. \(\square\)

Note 2. If $E' = E = E^n (C^n)$, the bijectionity of $\phi = df(\bar{x}, \cdot)$ is equivalent to $\det[\phi] = \det[f'(\bar{x})] \neq 0$ (Theorem 1 of §6).

In this case, the fact that $f$ is one-to-one on $G = G_{\bar{\delta}}(\bar{\delta})$ means, componentwise (see Note 3 in §6), that the system of $n$ equations

$$f_i(\bar{x}) = f(x_1, \ldots, x_n) = y_i, \quad i = 1, \ldots, n,$$

has a unique solution for the $n$ unknowns $x_k$ as long as

$$(y_1, \ldots, y_n) = \bar{y} \in f[G].$$

Theorem 3 shows that this solution has the form

$$x_k = g_k(\bar{y}), \quad k = 1, \ldots, n,$$

where the $g_k$ are of class $CD^1$ on $f[G]$ provided the $f_i$ are of class $CD^3$ near $\bar{p}$ and $\det[f'(\bar{p})] \neq 0$. Here

$$\det[f'(\bar{p})] = J_f(\bar{p}),$$

as in §6.

Thus again $f$ “locally” resembles a linear map, $\phi = df(\bar{x}, \cdot)$.

IV. The Implicit Function Theorem. Generalizing, we now ask, what about solving $n$ equations in $m + n$ unknowns $x_1, \ldots, x_n, y_1, \ldots, y_m$? Say, we want to solve

$$f_k(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0, \quad k = 1, 2, \ldots, n,$$

for the first $n$ unknowns (or variables) $x_k$, thus expressing them as

$$x_k = H_k(y_1, \ldots, y_m), \quad k = 1, \ldots, n,$$
with \( H_k : E^m \to E^1 \) or \( H_k : C^m \to C \).

Let us set \( \bar{x} = (x_1, \ldots, x_n), \, \bar{y} = (y_1, \ldots, y_m) \), and
\[
(\bar{x}, \bar{y}) = (x_1, \ldots, x_n, y_1, \ldots, y_m)
\]
so that \((\bar{x}, \bar{y}) \in E^{n+m} (C^{n+m})\).

Thus the system of equations \((11)\) simplifies to
\[
f_k(\bar{x}, \bar{y}) = 0, \quad k = 1, \ldots, n,
\]
or
\[
f(\bar{x}, \bar{y}) = \bar{0},
\]
where \( f = (f_1, \ldots, f_n) \) is a map of \( E^{n+m} (C^{n+m}) \) into \( E^n (C^n) \); \( f \) is a function of \( n + m \) variables, but it has \( n \) components \( f_k \); i.e.,
\[
f(\bar{x}, \bar{y}) = f(x_1, \ldots, x_n, y_1, \ldots, y_m)
\]
is a vector in \( E^n (C^n) \).

**Theorem 4** (implicit functions). Let \( E' = E^{n+m} (C^{n+m}), \, E = E^n (C^n) \), and let \( f : E' \to E \) be of class \( CD^1 \) near
\[
(\bar{p}, \bar{q}) = (p_1, \ldots, p_n, q_1, \ldots, q_m), \quad \bar{p} \in E^n (C^n), \quad \bar{q} \in E^m (C^m).
\]

Let \([\phi]\) be the \( n \times n \) matrix
\[
(D_j f_k)(\bar{p}, \bar{q}), \quad j, k = 1, \ldots, n.
\]

If \( \det[\phi] \neq 0 \) and if \( f(\bar{p}, \bar{q}) = \bar{0} \), then there are open sets
\[
P \subseteq E^n (C^n) \text{ and } Q \subseteq E^m (C^m),
\]
with \( \bar{p} \in P \) and \( \bar{q} \in Q \), for which there is a unique map
\[
H : Q \to P
\]
with
\[
f(H(\bar{y}), \bar{y}) = \bar{0}
\]
for all \( \bar{y} \in Q \); furthermore, \( H \in CD^1 \) on \( Q \).

Thus \( \bar{x} = H(\bar{y}) \) is a solution of \((11)\) in vector form.

**Proof.** With the above notation, set
\[
F(\bar{x}, \bar{y}) = (f(\bar{x}, \bar{y}), \bar{y}), \quad F : E' \to E'.
\]

Then
\[
F(\bar{p}, \bar{q}) = (f(\bar{p}, \bar{q}), \bar{q}) = (\bar{0}, \bar{q}),
\]
since \( f(\bar{p}, \bar{q}) = \bar{0} \).

As \( f \in CD^1 \) near \((\bar{p}, \bar{q})\), so is \( F \) (verify *componentwise* via **Problem 9**(ii) in §3 and **Definition 1** of §5).

By **Theorem 4**, §3, \( \det[F'(\bar{p}, \bar{q})] = \det[\phi] \neq 0 \) (explain!).

Thus \( H(\bar{y}) \) above shows that \( F \) is one-to-one on some globe \( G \) about \((\bar{p}, \bar{q})\).

Clearly \( G \) contains an open interval about \((\bar{p}, \bar{q})\). We denote it by \( P \times Q \) where \( \bar{p} \in P, \quad \bar{q} \in Q \); \( P \) is open in \( E^n (C^n) \) and \( Q \) is open in \( E^m (C^m) \).

By **Theorem 3**, \( F_{P \times Q} (F \text{ restricted to } P \times Q) \) has an inverse
\[
g : A \leftrightarrow P \times Q,
\]
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
\[
g(\bar{x}, \bar{y}) = (u(\bar{x}, \bar{y}), \bar{y})
\]
exactly as \( F(\bar{x}, \bar{y}) = (f(\bar{x}, \bar{y}), \bar{y}) \). Also, \( u : A \to P \) is of class \( CD^1 \) on \( A \), as \( g \) is (explain!).

Now set
\[
H(\bar{y}) = u(\bar{0}, \bar{y});
\]
here \( \bar{y} \in Q \), while
\[
(\bar{0}, \bar{y}) \in A = F[P \times Q],
\]
for \( F \) preserves \( \bar{y} \) (the last \( m \) coordinates). Also set
\[
\alpha(\bar{x}, \bar{y}) = \bar{x}.
\]

Then \( \overline{f} = \alpha \circ F \) (why?), and
\[
f(H(\bar{y}), \bar{y}) = f(u(\bar{0}, \bar{y}), \bar{y}) = f(g(\bar{0}, \bar{y})) = \alpha(F(g(\bar{0}, \bar{y})) = \alpha(\bar{0}, \bar{y}) = \bar{0}
\]
by our choice of \( \alpha \) and \( g \) (inverse to \( F \)). Thus
\[
f(H(\bar{y}), \bar{y}) = \bar{0}, \quad \bar{y} \in Q,
\]
as desired.

Moreover, as \( H(\bar{y}) = u(\bar{0}, \bar{y}) \), we have
\[
\frac{\partial}{\partial y_i} H(\bar{y}) = \frac{\partial}{\partial y_i} u(\bar{0}, \bar{y}), \quad \bar{y} \in Q, \, i \leq m.
\]

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 \( Q \).

---

4 This can be made more precise using the theory of *product spaces* (Chapter 4, *§11*).
Finally, \( H \) is unique for the given \( P, Q \); for
\[
\begin{align*}
f(\vec{x}, \vec{y}) &= 0 \\&\implies (f(\vec{x}, \vec{y}), \vec{y}) = (\vec{0}, \vec{y}) \\&\implies F(\vec{x}, \vec{y}) = (\vec{0}, \vec{y}) \\&\implies g(F(\vec{x}, \vec{y})) = g(\vec{0}, \vec{y}) \\&\implies (\vec{x}, \vec{y}) = g(\vec{0}, \vec{y}) = (u(\vec{0}, \vec{y}), \vec{y}) \\&\implies \vec{x} = u(\vec{0}, \vec{y}) = H(\vec{y}).
\end{align*}
\]
Thus \( f(\vec{x}, \vec{y}) = 0 \) 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 \( f(\vec{x}, \vec{y}) = 0 \). In this sense we say that \( H(\vec{y}) \) is an *implicit function*, given by \( f(\vec{x}, \vec{y}) = 0 \).

Similarly, under suitable assumptions, \( f(\vec{x}, \vec{y}) = 0 \) 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
\[
f(x, y) = x^2 + y^2 - 25
\]
(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:
\[
H_1(y) = +\sqrt{25 - y^2} \\
H_2(y) = -\sqrt{25 - y^2}.
\]
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.

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.\(^5\)

---

\(^5\)For more on implicit differentiation, see §10.

§7. Inverse and Implicit Functions. Open and Closed Maps

**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
\[
f(x, y, z) = x^2 + y^2 + z^2 - 1 = 0, \quad x, y, z \in E^3.
\]
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 \( \frac{\partial y}{\partial x} \) and \( \frac{\partial y}{\partial y} \).

**Problems on Inverse and Implicit Functions, Open and Closed Maps**

1. Discuss: In Definition 1, \( \overline{C} \) can equivalently be replaced by \( G = Gf(\delta) \) (an open globe).

2. Prove that if the set \( D \) is open (closed) in \((S, \rho)\), then the map \( f: S \to T \) is open (closed, respectively) on \( D \) if \( f[D] \) (\( f \) restricted to \( D \)) has this property as a map of \( D \) into \( f[D] \).

[Hint: Use Theorem 4 in Chapter 3, §12.]

3. Complete the missing details in the proofs of Theorems 1–4.

3'. Verify footnotes 2 and 3.

4. Show that a map \( f: E' \to E \) 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 \( E' = E = E^3 \).

[Hints: For the first part, take \( E' = C, f(x + iy) = e^x(\cos y + \sin y) \). For the second, use Theorem 1 in Chapter 5, §2.]

4'. (i) For maps \( f: E^1 \to E^1 \), prove that the existence of a bijective
\[
\frac{df}{d\vec{p}}(\vec{p}, \cdot) \text{ is equivalent to } f'(\vec{p}) \neq 0.
\]

(ii) Let
\[
f(x) = x + x^2 \sin \frac{1}{x}, \quad f(0) = 0.
\]

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?
5. Show that a map \( f : E^n (C^n) \to E^n (C^n) \), \( f \in CD^1 \), 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 \( \det[\frac{df}{dx}([\vec{q}])] = \det[\frac{df^{-1}}{dx}([\vec{q}])] = 1 \neq 0 \), since \( f \circ f^{-1} \) is the identity map.]

6. Prove Corollary 2 for the general case of complete \( E' \) and \( E \).

[Outline: Given a closed \( X \subseteq \overline{X} \), take any convergent sequence \( \{\vec{y}_n\} \subseteq f[X] \). By Problem 8 in Chapter 4, §8, \( f^{-1}(\vec{y}_n) = \vec{x}_n \) is a Cauchy sequence in \( X \) (why?) By the completeness of \( E' \), \( \exists \vec{x} \in X \) such that \( \vec{y}_n \to \vec{x} \) (Theorem 4 of Chapter 3, §16). Infer that \( \lim f(\vec{x}) = f(\vec{x}) \in f[X] \), so \( f[X] \) is closed.]

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.

8. Prove in detail that \( f : (S, \rho) \to (T, \rho') \) is open on \( D \subseteq S \) if \( f \) maps the interior of \( D \) into that of \( f[D] \); that is, \( f[D]^o \subseteq (f[D])^o \).

9. Verify by examples that \( f \) may be:
   (i) closed but not open;
   (ii) open but not closed.

[Hints: (i) Consider \( f = \text{constant} \). (ii) Define \( f : E^2 \to E^1 \) by \( f(x, y) = x \) and let \( D = \{(x, y) \in E^2 \mid y = \frac{1}{x}, x > 0\} \); 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 9. (Verify!)]

10. Continuing Problem 9(iii), define \( f : E^n \to E^1 \) (or \( C^n \to C \)) by \( f(\vec{x}) = x_k \) for a fixed \( k \leq n \) (the \( k \)-th projection map). Show that \( f \) is open, but not closed, on \( E^n (C^n) \).

   (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).

   (i) Treating \( z \) as a function of \( x \) and \( y \), given implicitly by

   \[
   f(x, y, z) = z^2 + xz^2 - yz = 0, \quad f : E^3 \to E^3,
   \]

   discuss the choices of \( P \) and \( Q \) that satisfy Theorem 4. Find \( \frac{\partial f}{\partial x} \) and \( \frac{\partial f}{\partial y} \).

   (ii) Do the same for \( f(x, y, z) = e^{xy} - 1 = 0 \).

13. Given \( f : E^n (C^n) \to E^m (C^m) \), \( n > m \), prove that if \( f \in CD^1 \) on a globe \( G \), \( f \) cannot be one-to-one.

[Hint for \( f : E^3 \to E^2 \): If, say, \( D_3 f \neq 0 \) on \( G \), set \( F(x, y) = (f(x, y)) \).]

§7. Inverse and Implicit Functions. Open and Closed Maps

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 \( g = f^{-1} \) (restrict \( f \) to \( A \) and take its inverse). Show that \( f \) and \( g \) are open and of class \( CD^1 \) on \( A \) and \( f[A] \), respectively.

15. Given \( \vec{v} \in E \) and a scalar \( c \neq 0 \), define \( T_{\vec{v}} : E \to E \) (“translation by \( \vec{v} \)”)

   \[
   T_{\vec{v}}(\vec{x}) = \vec{x} + \vec{v} \quad \text{and} \quad M_{\vec{v}, c} : E \to E \quad \text{("dilation by \( c\vec{v} \)}, \text{ by setting}
   \]

\[
T_{\vec{v}}(\vec{x}) = \vec{x} + \vec{v} \quad \text{and} \quad M_{\vec{v}, c} = c\vec{v}.
\]

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 \( M_{\vec{v}} \) and \( M_{\vec{v}, c}^{-1} = M_{\vec{v}, -c} \).

(ii) Similarly for the Lipschitz property on \( E \).

(iii) If \( G = G_{\vec{v}, c} \subseteq E \), then \( T_{\vec{v}}[G] = G_{\vec{v}, c} \) and \( M_{\vec{v}, c}[G] = G_{c\vec{v}}[\{\vec{v}\}] \).

(iv) If \( f : E' \to E \) 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 \( M_{\vec{v}, c} \circ f = f \circ M_{\vec{v}, c} \), where \( T_{\vec{p}} \) and \( M_{\vec{v}, c} \) 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 \( \vec{v} = f(\vec{x} - \vec{p}) \), \( g = T_{\vec{v}} \circ f \circ T_{\vec{v}, -\vec{x}} \). Verify that \( g = f : T_{\vec{v}, -\vec{x}} \circ (\vec{x} - \vec{p}) \) and \( g \) is continuous at \( \vec{x} \)).

16. Show that if \( f : E' \to E \) is linear and if \( f[G^*] \) is open in \( E \) for some \( G^* = G_{\vec{v}, c} \subseteq E' \), 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 \( G = G_{\vec{v}, c} \). Then use Problems 7 and 13(i)-(iv), with suitable \( \vec{v} \) and \( c \).

(ii) To prove \( f = f[E'] \), fix any \( \vec{y} \in E \). As \( f = G_{\vec{v}, c}(\vec{y}) \) is open, it contains a globe \( G^* = G_{\vec{v}, c}(\vec{y}) \). For small \( c, \vec{y} \in G^* \subseteq f[E'] \). Hence \( \vec{y} \in f[E'] \) (Problem 10 in §2).]

17. Continuing Problem 16, show that if \( f \) is also one-to-one on \( G^* \), then

\[
f : E' \to E\quad \text{onto}, \quad f \subseteq E(E), \quad f^{-1} \subseteq L(E, E') \quad \text{is open on} \quad E', \quad \text{and} \quad f^{-1} \quad \text{is so on} \quad E.
\]

[Hints: To prove that \( f \) is one-to-one on \( E' \), let \( f(\vec{x}) = f(\vec{x}) \) for some \( \vec{x}, \vec{x}' \in E' \). Show that

\[
(3c, \epsilon > 0) \quad \text{if } \epsilon \in G_{\vec{v}, c} \subseteq f[G_{\vec{v}, c}] \quad \text{and} \quad f(\vec{x} + \vec{p}) = f(\epsilon \vec{x} + \vec{p}) = f(G_{\vec{v}, c}) = f[G_{\vec{v}, c}].
\]

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.]

18. A map \( f : (S, \rho) \to (T, \rho') \)
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) \(x_n \to p\) in \(S\) iff \(f(x_n) \to f(p)\) 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) \(B = A\) in \(S\) iff \(f[B] = (f[A])\) in \(T\);
(v) \(A\) is dense in \(B\) (i.e., \(A \subseteq B \subseteq \overline{A} \subseteq S\)) in \((S, \rho)\) 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.]

19. Given \(A, B \subseteq E, \ \vec{v} \in E\) and a scalar \(c\), set
\[A + \vec{v} = \{ \vec{x} + \vec{v} \mid \vec{x} \in A \}\]and \(cA = \{ \vec{c}\vec{x} \mid \vec{x} \in A \}\).

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 \(cB + \vec{v} = c\overline{A} + \vec{v}\);
(iii) \(B = A\) iff \(cB + \vec{v} = (cA + \vec{v})\);
(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 \(M_{\vec{v}}\) of Problem 15, noting that \(A + \vec{v} = T_{\vec{v}}[A]\) and \(cA = M_{\vec{v}}[A]\).]

20. Prove Theorem 2, for a reduced \(\delta\), assuming that only one of \(E'\) and \(E\) is \(E^n\) (\(C^n\)), and the other is just complete.

[Hint: If, say, \(E = E^n\) (\(C^n\)), 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., \(f[G] \supseteq G_\alpha\).]

Now use the continuity of \(f\) to obtain a globe \(G' = G_\alpha \subseteq G\) such that \(f[G'] \supseteq G_\alpha\). Let \(g = f|G'\), further restricted to \(G_\alpha\). Apply Problem 15(v) in Chapter 4, §2, to \(g\), with \(S = G_\alpha\), \(T = E'\].]

**§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, \((S, \rho)\) 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 \(G\) disjoint from \(A\). (Why?)

A set \(A \subseteq (S, \rho)\) is meagre, or of Category I (in \(S\)), iff
\[A = \bigcup_{n=1}^{\infty} A_n\]
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 1** (Baire). In a complete metric space \((S, \rho)\), 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.,
\[G^* = \bigcup_{n=1}^{\infty} A_n\]
for some nowhere dense sets \(A_n\). Now, as \(A_1\) is nowhere dense, \(G^*\) contains a closed globe
\[\overline{G}_1 = G_1(\delta_1) \subseteq -A_1.\]

Again, as \(A_2\) is nowhere dense, \(G_1\) contains a globe
\[\overline{G}_2 = G_2(\delta_2) \subseteq -A_2, \quad \text{with } 0 < \delta_2 \leq \frac{1}{2}\delta_1.\]
By induction, we obtain a contracting sequence of closed globes
\[ \Gamma_n = \overline{G_{n,\delta_n}}, \] with \(0 < \delta_n \leq \frac{1}{9^n}\delta_0 \to 0\).

As \(S\) is complete, so are the \(\Gamma_n\) (Theorem 5 in Chapter 3, §17). Thus, by Cantor’s theorem (Theorem 5 of Chapter 4, §6), there is
\[ p \in \bigcap_{n=1}^{\infty} \Gamma_n. \]

As \(G^* \supseteq \Gamma_n\), we have \(p \in G^*\). But, as \(\Gamma_n \subseteq A_n\), we also have \((\forall n) p \notin A_n\); hence
\[ p \notin \bigcup_{n=1}^{\infty} A_n = G^* \]
(the desired contradiction!). □

We shall need a lemma based on Problems 15 and 19 in §7. (Review them!)

**Lemma.** Let \(f \in L(E', E')\), \(E'\) complete. Let \(G = G[0]\) be the unit globe in \(E'\). If \(f(G)\) (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 \(E'\) (e.g., \(\overrightarrow{0}\)), but not those from \(E\) (e.g., 0).

**Proof of lemma.** Let \(A = f[G] \cap G_0 \subseteq G_0\). We claim that \(A\) is dense in \(G_0\); i.e., \(G_0 \subseteq \overrightarrow{A}\). Indeed, by assumption, any \(q \in G_0\) is in \(\overline{f[G]}\). Thus by Theorem 3 in Chapter 3, §16, any \(G_q\) meets \(f[G] \cap G_0 = A\) if \(q \in G_0\). Hence
\[ (\forall q \in G_0) \quad q \in \overrightarrow{A}, \]

i.e., \(G_0 \subseteq \overrightarrow{A}\), 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 \(G_0\),
\[ A \cap G_{q_0}(cr) \neq \emptyset; \]
so let \(q_1 \in A \cap G_{q_0}(cr) \subseteq f[G]\). Then
\[ |q_1 - q_0| < cr, \quad q_0 \in G_{q_1}(cr). \]

As \(q_1 \in f[G]\), we can fix some \(\tilde{q}_1 \in G = G(1), \) with \(f(\tilde{q}_1) = q_1\). Also, by Problems 19(iv) and 15(iii) in §7, \(cA + q_1\) is dense in \(cG + q_1 = G_{q_1}(cr)\). But \(q_0 \in G_{q_1}(cr)\). Thus
\[ G_{q_0}(cr^2) \cap (cA + q_1) \neq \emptyset; \]
so let \(q_2 \in G_{q_0}(cr^2) \cap (cA + q_1)\), so \(q_0 \in G_{q_2}(cr^2)\), etc.

Inductively, we fix for each \(n > 1\) some \(q_n \in G_{q_0}(cr^n)\), with
\[ q_n \in c^{n-1}A + q_{n-1}, \]

*§8. Baire Categories. More on Linear Maps*

\[ q_n - q_{n-1} \in c^{n-1}A. \]
As \(A \subseteq f[G_0(1)]\), linearity yields
\[ q_n - q_{n-1} \in f[c^{n-1}G_0(1)] = f[G_0(c^{n-1})], \quad n > 1. \]

Thus for each \(n > 1\), there is \(\tilde{p}_n \in G_0(c^{n-1})\), i.e., \(|\tilde{p}_n| < c^{n-1}\) such that
\[ f(\tilde{p}_n) = q_n - q_{n-1}. \]
Now, as \(|\tilde{p}_n| < c^{n-1}\) and \(0 < c < 1\),
\[ \sum_{k=1}^{\infty} |\tilde{p}_n| < +\infty; \]
so by the completeness of \(E'\), \(\sum_{k=1}^{\infty} \tilde{p}_n\) converges in \(E'\) (Theorem 1 in Chapter 4, §13). Let \(\tilde{p} = \sum_{k=1}^{\infty} \tilde{p}_k\); then
\[ f(\tilde{p}) = f \left( \lim_{n \to \infty} \sum_{k=1}^{n} \tilde{p}_k \right) = \lim_{n \to \infty} f \left( \sum_{k=1}^{n} \tilde{p}_k \right) = \lim_{n \to \infty} \sum_{k=1}^{n} f(\tilde{p}_k) \quad \text{for } f \in L(E', E). \]

But \(f(\tilde{p}_k) = q_k - q_{k-1} (k > 1)\), and \(f(\tilde{p}_1) = q_1\); so
\[ \sum_{k=1}^{n} f(\tilde{p}_k) = q_1 + \sum_{k=2}^{n} (q_k - q_{k-1}) = q_n. \]

Thus
\[ f(\tilde{p}) = \lim_{n \to \infty} \sum_{k=1}^{n} f(\tilde{p}_k) = \lim_{n \to \infty} q_n = q_0. \]
Moreover, \(|\tilde{p}_n| < c^{k-1} (k \geq 1)\). Thus
\[ |\tilde{p}| \leq \sum_{k=1}^{\infty} |\tilde{p}_k| < \sum_{k=1}^{\infty} c^{k-1} = \frac{1}{1-c}. \]
i.e.,
\[ \tilde{p} \in G(c^{k-1}). \]

But \(q_0 = f(\tilde{p})\); so
\[ q_0 \in f \left[ G \left( \frac{1}{1-c} \right) \right]. \]

\[ \text{Note that } q_n \to q_0, \text{ since } q_n \in G_{q_0}(c^{n-1}) \text{ implies } |q_n - q_0| < c^{n-1}r \to 0, \text{ as } 0 < c < 1. \]
As \( q_0 \in G_0(r) \) was arbitrary, we have
\[
G_0(r) \subseteq \bigcup_{n=1}^{\infty} G_0(n) = f(G) = f(G).
\]
or by linearity,
\[
G_0(r(1 - c)) \subseteq \bigcup_{n=1}^{\infty} G_0(1) = f(G).
\]
This holds for any \( c \in (0, 1) \). Hence
\[
f(G) \supseteq \bigcup_{0 < c < 1} G_0(r(1 - c)) = G_0(r). \quad \text{(Verify!)}
\]
Thus all is proved. \( \Box \)

We can now establish an important result due to S. Banach.

**Theorem 2** (Banach). Let \( f \in L(E', E) \), with \( E' \) complete. Then \( f[E'] \) is meagre in \( E \) if and only if \( f[G_0(1)] \) is 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
\[
f[E'] = f \bigg[ \bigcup_{n=1}^{\infty} G_0(n) \bigg] = \bigcup_{n=1}^{\infty} f[G_0(n)]
\]
is a countable union of nowhere dense sets, hence meagre, by definition.

Now suppose \( f[G_0(1)] \) is not nowhere dense; so \( f[G_0(1)] \) contains some \( G_0(r) \subseteq E \). We may assume \( q \in f[G_0(1)] \) (if not, replace \( q \) by a close point from \( f[G_0(1)] \)). Then \( q = f(\tilde{p}) \) for some \( \tilde{p} \in G_0(1) \). The latter implies
\[
|q - \tilde{p}| = |\tilde{p}| = \rho(\tilde{p}, 0) < 1;
\]
so
\[
G_{-r}(\tilde{p}) \subseteq G_0(2).
\]
Also, as \( f[G_0(1)] \supseteq G_0(r) \), translation by \( -q = f(-\tilde{p}) \) yields
\[
f[G_0(1)] + f(-\tilde{p}) \supseteq G_0(r) - q = G_0(0),
\]
i.e.,
\[
G_0(0) \subseteq f[G_0(1)] \subseteq f(G_0(2)).
\]
Hence \( f[G_0(1)] \supseteq G_0(\frac{1}{2} r) \) (why?), so, by the Lemma
\[
f[G_0(1)] \supseteq G_0 \left( \frac{1}{2} r \right) \quad \text{in } E.
\]

This implies \( f[G_0(2n)] \supseteq G_0(nr) \), and so
\[
f[E'] \supseteq \bigcup_{n=1}^{\infty} G_0(nr) = E,
\]
i.e., \( f[E'] = E \), as required. Thus the theorem is proved. \( \Box \)

**Theorem 3** (Open map principle). Let \( f \in L(E', E) \), with \( E' \) and \( E \) complete. Then the map \( f \) is open on \( E' \). If \( f[E'] = E \), i.e., if \( f \) 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, \( f[G_0(1)] \) is not nowhere dense, and (1) follows as before. Hence by Problems 15(iii) and 19 in §7, \( f[\bar{G}_0] \supseteq G_0 \) whenever \( q = f(\tilde{p}) \). (Why?) Therefore, \( G_0(\tilde{p}) \subseteq \bar{A} \subseteq E' \)
\[
f[\bar{A}] \subseteq f[\bar{G}_0] \subseteq f[\bar{A}];
\]
i.e., \( f \) maps any interior point \( \tilde{p} \in A \) into such a point of \( f[A] \). By Problem 8 in §7, \( f \) is open on \( E' \).

Conversely, if \( f \) is open, then \( f[E'] \) is an open set \( \neq \emptyset \) in \( E \), a complete space; so by Theorems 1 and 2, \( f[E'] \) is nonmeagre and equals \( E \). (See also Problem 16(ii) in §7.) \( \Box \)

**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 \( E' \), and so \( f^{-1} \in L(E, E') \) by Note 1 in §7. (This is the promised general proof of Corollary 2 in §6.)

**Theorem 4** (Banach–Steinhaus uniform boundedness principle). Let \( E' \) be complete. Let \( \mathcal{N} \) be a family of maps \( f \in L(E', E) \) such that
\[
\begin{align*}
(2) & \quad (\forall x \in E') \quad (\exists k \in E^1) \quad (\forall f \in \mathcal{N}) \quad |f(x)| < k, \\
(\mathcal{N} \text{ is bounded at each } x).
\end{align*}
\]
Then \( \mathcal{N} \) is "norm-bounded," i.e.,
\[
(\exists K \in E^1) \quad (\forall f \in \mathcal{N}) \quad \|f\| < K,
\]
with \( \| \| \) as in §2.

**Proof.** It suffices to show that \( \mathcal{N} \) is "uniformly" bounded on some globe,
\[
(3) \quad (\exists \bar{x} \in E^1) \quad (\exists G = G_{\bar{x}}(r)) \quad (\forall f \in \mathcal{N}) \quad (\forall \bar{x} \in G) \quad |f(\bar{x})| < c.
\]
For then \( |\bar{x} - \bar{p}| \leq r \) implies
\[
2c > |f(\bar{x}) - f(\bar{p})| = |f(\bar{x}) - \bar{p}|,
\]
or (setting \( \bar{x} - \bar{p} = r\tilde{p} \)) \( |\tilde{p}| < 1 \) implies
\[
(\forall f \in \mathcal{N}) \quad |f(\tilde{p})| < \frac{2c}{r} \quad (\text{why}?)�
\]
\[2^c.
\]
\[2c.
\]
§8. Baire Categories. More on Linear Maps

3. Complete the missing details in the proof of Theorems 1 to 4.
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.
5. Prove that in a discrete space $(S, \rho)$, only $\emptyset$ is meagre.
   [Hint: Use Problem 8 in Chapter 3, §17, Example 7 in Chapter 3, §12, and our present Theorem 1.]
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.]
7. What is wrong about this “proof” that every closed set $F \neq \emptyset$ in a complete metric space $(S, \rho)$ is residual: “By Theorem 5 of Chapter 3, §17, $F$ is complete as a subspace. Thus by Theorem 1, $F$ is residual.” Give counterexamples!
8. We call $K$ a $G_δ$-set and write $K \in G_δ$ iff $K = \bigcap_{n=1}^{\infty} G_n$ for some open sets $G_n$.
    (i) Prove that if $K$ is a $G_δ$-set, and if $K$ is dense in a complete metric space $(S, \rho)$, 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 $\overline{K} = \bigcap_{n=1}^{\infty} G_n$ is meagre. Use Theorem 1.]
    (ii) Infer that $R$ (the rationals) is not a $G_δ$-set in $E^1$ (cf. Example (c)).
9. Show that, in a complete metric space $(S, \rho)$, a meagre set $A$ cannot have interior points.
   [Hint: Otherwise, $A$ would obtain a globe $G$. Use Theorem 1 and Problem 4(ii).]
10. (i) A singleton $\{p\} \subseteq (S, \rho)$ is nowhere dense if $S$ clusters at $p$; otherwise, it is nonmeagre in $(S, \rho)$ (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 $S$.

*§8. Baire Categories. More on Linear Maps
11. (i) Show that if \( \emptyset \neq A \subseteq G \) (see Problem 8) in a complete space \((S, \rho)\), 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 \((\bar{\mathcal{A}}, \rho)\) is complete (why?); so \( A \) is nonmeagre in \( \bar{\mathcal{A}} \), by Problem 8. Use Problem 10(ii). (ii) Use Footnote 3.]

12. If \( G \) is open in \((S, \rho)\), then \( G - G \) is nowhere dense in \( S \).

[Hint: \( \overline{G - G} = \overline{G} \cap (\overline{G})^c \) is closed; so
\[
\overline{G - G}^0 = \overline{G} - G^0 = \overline{G - G}^c = \emptyset
\]

by Problem 15 in Chapter 3, §12 and Problem 15 in Chapter 3, §16.]

13. (“Simplified” uniform boundedness theorem.) Let \( f_n : (S, \rho) \to (T, \rho') \) be continuous for \( n = 1, 2, \ldots \), with \( S \) complete. If \( \{f_n(x)\} \) is a bounded sequence in \( T \) for each \( x \in S \), then \( \{f_n\} \) is uniformly bounded on some open \( G \neq \emptyset \):

\[
(\forall p \in T) (\exists k) (\forall n) (\forall x \in G) \rho'(p, f_n(x)) \leq k.
\]

[Outline: Fix \( p \in T \) and \((\forall n)\) set
\[
F_n = \{x \in S | (\forall m) \rho \leq \rho(p, f_n(x))\}.
\]

Use the continuity of \( f_n \) and of \( \rho' \) 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 \( \overline{F}^0 \neq \emptyset \). Set \( G = \overline{F}^0 \) and show that \( G \) is as required.]

14. Let \( f_n : (S, \rho) \to (T, \rho') \) be continuous for \( n = 1, 2, \ldots \). Show that if \( f_n \to f \) (pointwise) on \( S \), then \( f \) is continuous on \( S - Q \), with \( Q \) meagre in \( S \).

[Outline: \( (\forall k, m) \) let
\[
A_{km} = \bigcup_{n=1}^{\infty} \{x \in S | \rho'(f_n(x), f_m(x)) > \frac{1}{k} \}
\]

By the continuity of \( \rho' \), \( f_n \) and \( f_m \), \( A_{km} \) is open in \( S \). (Why?) So by Problem 12, \( \bigcap_{m=1}^{\infty} A_{km} = \emptyset \) is meagre for \( k = 1, 2, \ldots \).

Also, as \( f_n \to f \) on \( S \), \( \bigcap_{m=1}^{\infty} A_{km} = \emptyset \). (Verify!) Thus

\[
(\forall k) \bigcup_{m=1}^{\infty} A_{km} \subseteq \bigcup_{m=1}^{\infty} (\overline{A_{km}} - A_{km}).
\]

(Why?) Hence the set \( Q = \bigcup_{m=1}^{\infty} \bigcap_{m=1}^{\infty} \overline{A_{km}} \) is meagre in \( S \).

Moreover, \( S - Q = \bigcap_{m=1}^{\infty} \bigcup_{m=1}^{\infty} \overline{A_{km}} \) by Problem 16 in Chapter 3, §16. Deduce that if \( p \in S - Q \), then

\[
(\forall \varepsilon > 0) (\exists m_0) (\exists G_m) (\forall n, m \geq m_0) (\forall x \in G_m) \rho'(f_m(x), f_n(x)) < \varepsilon.
\]

*§8. Baire Categories. More on Linear Maps

Keeping \( m \) fixed, let \( n \to \infty \) to get

\[
(\forall \varepsilon > 0) (\exists m_0) (\exists G_m) (\forall m \geq m_0) (\forall x \in G_m) \rho'(f_m(x), f(x)) < \varepsilon.
\]

Now modify the proof of Theorem 2 of Chapter 4, §12, to show that this implies the continuity of \( f \) at each \( p \in S - Q \).

§9. Local Extrema. Maxima and Minima

We say that \( f : E' \to E^1 \) has a local maximum (minimum) at \( \bar{p} \in E' \) if \( f(\bar{p}) \) is the largest (least) value of \( f \) on some globe \( G \) about \( \bar{p} \); more precisely, if

\[
(\forall x \in G) \Delta f = f(\bar{x}) - f(x) < 0 (> 0).
\]

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 \( \Delta f \).

From Problem 6 in §1, recall the following necessary condition.

Theorem 1. If \( f : E' \to E^1 \) has a local extremum at \( \bar{p} \) then \( D_{xy}f(\bar{p}) = 0 \) for all \( \bar{a} \neq \bar{0} \) in \( E' \).

In the case \( E' = E^n (C^n) \), this means that \( d^1f(\bar{p}, \cdot) = 0 \) on \( E' \).

(Recall that \( d^1f(\bar{p}, \bar{x}) = \sum_{k=1}^{n} D_k f(\bar{p}) \bar{x}_k \). It vanishes if the \( D_k f(\bar{p}) \) do.)

Note 1. This condition is only necessary, not sufficient. For example, if \( f(x, y) = xy \), then \( d^1f(0, \cdot) = 0 \); yet \( f \) has no extremum at \( \bar{0} \). (Verify!)

Sufficient conditions were given in Theorem 2 of §5, for \( E' = E^1 \). We now take up \( E' = E^2 \).

Theorem 2. Let \( f : E^2 \to E^1 \) be of class \( CD^2 \) on a globe \( G = G_\delta(\delta) \). Suppose \( d^1f(\bar{p}, \cdot) = 0 \) on \( E^2 \). Set \( A = D_{x_1}f(\bar{p}), B = D_{x_2}f(\bar{p}), \) and \( C = D_{x_2}f(\bar{p}) \).

Then the following statements are true.

(i) If \( AC > B^2 \), \( f \) has a maximum or minimum at \( \bar{p} \), according to whether \( A < 0 \) or \( A > 0 \).

(ii) If \( AC < B^2 \), \( f \) has no extremum at \( \bar{p} \).

The case \( AC = B \) is unresolved.

Proof. Let \( \bar{x} \in G \) and \( \bar{a} = \bar{x} - \bar{p} \neq \bar{0} \). As \( d^1f(\bar{p}, \cdot) = 0 \), Theorem 2 in §5, yields

\[
\Delta f = f(\bar{x}) - f(\bar{p}) = R_1 = \frac{1}{2}f^2(\bar{p}; \bar{a}),
\]

with \( \bar{s} \in L(\bar{p}, \bar{x}) \subseteq G \) (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,

\[
(1) \quad (\forall \varepsilon > 0) (\exists \varepsilon_0) (\exists \varepsilon_0 \subseteq G) (\forall x \in \varepsilon_0 \subseteq G) \rho'(f_m(x), f(x)) < \varepsilon.
\]
Now, as the partials involved are continuous, we can choose \( G = G_δ(\delta) \) so small that the sign of expression (1) will not change if \( \bar{s} \) is replaced by \( \bar{p} \). Then the crucial sign of \( \Delta f \) on \( G \) coincides with that of

\[
D = A \bar{u}^2 + 2Bu_1u_2 + Cu^2
\]

(with \( A, B, \) and \( C \) as stated in the theorem).

From (2) we obtain, by elementary algebra,

\[
AD = (A_1u_1 + Bu_2)^2 + (AC - B^2)u_2^2,
\]

\[
CD = (C_1u_1 + Bu_2)^2 + (AC - B^2)u_2^2.
\]

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 \( p \). If \( A > 0 \), then \( \Delta f > 0 \), and \( f \) has a minimum at \( p \).

Now let \( AC < B^2 \). We claim that no matter how small \( G = G_δ(\delta) \), \( \Delta f \) changes sign as \( \bar{x} \) varies in \( G \), and so \( f \) has no extremum at \( p \).

Indeed, we have \( \bar{x} = \bar{p} + \bar{u} = (u_1, u_2) \neq 0 \). If \( u_2 = 0 \), (3) shows that \( D \) and \( \Delta f \) 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 \( \Delta f \) have the sign opposite to that of \( A \); and \( \bar{x} \) 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 \( \Delta f \) change sign as \( u_1u_2 \) does; so \( f \) has no extremum at \( p \). Thus all is proved.

Briefly, the proof utilizes the fact that the trinomial (2) is sign-changing iff its discriminant \( B^2 - 4AC \) is positive, i.e., \( |A B C| < 0 \).

Note 2. Functions \( f: C \to E^1 \) (of one complex variable) are likewise covered by Theorem 2 if one treats them 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 \( P: E^n \to E^1 \) of the form

\[
P(\bar{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}u_iu_j,
\]

where \( \bar{u} = (u_1, \ldots, u_n) \in E^n \) 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.

§9. Local Extrema. Maxima and Minima

Let \( P: E^n \to E^1 \) be a symmetric quadratic form,

\[
P(\bar{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}u_iu_j.
\]

(i) \( P > 0 \) on all of \( E^n - \{ 0 \} \) iff the following \( n \) determinants \( A_k \) are positive:

\[
A_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}, \quad k = 1, 2, \ldots, n.
\]

(ii) We have \( P < 0 \) on \( E^n - \{ 0 \} \) iff \( (−1)^k A_k > 0 \) for \( k = 1, 2, \ldots, n \).

Now we can extend Theorem 2 to the case \( f: E^n \to E^1 \). (This will also cover \( f: E^n \to E^1 \), treated as \( f: E^{2n} \to E^1 \).) The proof resembles that of Theorem 2.

Theorem 3. Let \( f: E^n \to E^1 \) be of class \( CD^2 \) on some \( G = G_δ(\delta) \). Suppose \( df(\bar{p}, \cdot) = 0 \) on \( E^n \). Define the \( A_k \) as in (4), with \( a_{ij} = \frac{∂f}{∂u_j}(\bar{p}) \), \( i, j \leq \nu \). Then the following statements hold.

(i) \( f \) has a local minimum at \( p \) if \( A_k > 0 \) for \( k = 1, 2, \ldots, n \).

(ii) \( f \) has a local maximum at \( p \) if \( (−1)^k A_k > 0 \) for \( k = 1, \ldots, n \).

(iii) \( f \) has no extremum at \( p \) if the expression

\[
P(\bar{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}u_iu_j
\]

is \( > 0 \) for some \( \bar{u} \in E^n \) and \( < 0 \) for others (i.e., \( P \) changes sign on \( E^n \)).

Proof. Let again \( \bar{x} \in G, \bar{u} = \bar{x} - \bar{p} \neq 0 \), and use Taylor’s theorem to obtain

\[
\Delta f = f(\bar{x}) - f(\bar{p}) = R_1 + \frac{1}{2} f''(\bar{s}; \bar{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} D_{ij}f(\bar{s})u_iu_j,
\]

with \( \bar{s} \in L(\bar{x}, \bar{p}) \).

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 \( \bar{s} \) is replaced by \( \bar{p} \).

Hence the sign of \( \Delta f \) on \( G \) is the same as that of \( P(\bar{u}) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}u_iu_j \), 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.,
\[
\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} u_i u_j > 0 > \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} v_i v_j
\]
for some \( \vec{u}, \vec{v} \in \mathbb{E}^n - \{ \vec{0} \} \).

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
\[
P(t \vec{u}) = t^2 P(\vec{u}) > 0 > t^2 P(\vec{v}) = P(t \vec{v}).
\]
Now, for any \( t \in (0, |\vec{u}|) \), the point \( \vec{x} = \vec{y} + t \vec{v} \) lies on the \( \vec{u} \)-directed line through \( \vec{p} \), inside \( \mathcal{G} = \mathcal{G}_P(\delta) \). (Why?) Similarly for the point \( \vec{x}' = \vec{y}' + t \vec{v}' \).

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 \( \vec{s}' \in L(\vec{p}, \vec{x}') \) lying on the same two lines. It again follows that for small \( \delta \),
\[
f(\vec{x}) - f(\vec{y}) > 0 > f(\vec{x}') - f(\vec{y}).
\]
just as \( P(\vec{u}) > 0 > P(\vec{v}) \).

Thus \( \Delta f \) changes sign on \( \mathcal{G}_P(\delta) \), and (iii) is proved. \( \square \)

Note 3. Still unresolved cases in which \( P(\vec{u}) \) vanishes for some \( \vec{u} \neq \vec{0} \), 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 \( d^2 f(\vec{p}, \cdot) = d^2 f(\vec{\tilde{p}}, \cdot) = 0 \) on \( \mathbb{E}^n \), then \( \Delta f = R_2 = \frac{1}{3} d^3 f(\vec{\tilde{p}}, \vec{s}) \), etc.

Note 4. The largest or least value of \( f \) on a set \( A \) (sometimes called the absolute maximum or minimum) may occur at some noninterior (e.g., boundary) point \( \vec{\tilde{p}} \in A \), and then fails to be among the local extrema (where, by definition, a globe \( \mathcal{G}_\vec{p} \subseteq A \) is presupposed). Thus to find absolute extrema, one must also explore the behavior of \( f \) at noninterior points of \( A \).

By Theorem 1, local extrema can occur only at so-called critical points \( \vec{\tilde{p}} \), i.e., those at which all directional derivatives vanish (or fail to exist, in which case \( D_k f(\vec{\tilde{p}}) = 0 \) by convention).

In practice, to find such points in \( \mathbb{E}^n (C^n) \), 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
\[
f(x, y) = \sin x + \sin y - \sin(x + y)
\]
on the set \( A \subseteq \mathbb{E}^2 \) bounded by the lines \( x = 0, y = 0 \) and \( x + y = 2\pi \).

We have
\[
D_1 f(x, y) = \cos x - \cos(x + y) \quad \text{and} \quad D_2 f(x, y) = \cos y - \cos(x + y).
\]

Inside the triangle \( A \), both partials vanish only at the point \( (\frac{\pi}{3}, \frac{\pi}{3}) \) at which \( f = \frac{2\pi}{3} \). 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,
\[
f(\frac{2\pi}{3}, \frac{2\pi}{3}) = \frac{3}{2} \sqrt{3},
\]
at this unique critical point.

(B) Find the largest and least value of
\[
f(x, y, z) = a^2 x^2 + b^2 y^2 + c^2 z^2 - (ax^2 + by^2 + cz^2)^2,
\]
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 \( F : \mathbb{E}^2 \to \mathbb{E}^1 \):
\[
F(x, y) = (a^2 - c^2)x^2 + (b^2 - c^2)y^2 + c^2 - ((a - c)x^2 + (b - c)y^2 + c^2)^2.
\]
(Explain!) For \( F \), we seek the extrema on the disc \( \mathcal{C} = \mathcal{C}_0(1) \subseteq \mathbb{E}^2 \), 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
\[
D_1 F(x, y) = 2x(a - c) - 2((a - c)x^2 + (b - c)y^2 + c^2)^2 = 0,
\]
\[
D_2 F(x, y) = 2y(b - c) - 2((a - c)x^2 + (b - c)y^2 + c^2)^2 = 0
\]
and solving this system of equations, we find these critical points inside \( G \):

1. \( x = y = 0 \) (\( F = 0 \));
2. \( x = 0, y = \pm 2\sqrt{\frac{2}{3}} \) (\( F = \frac{1}{4}(b - c)^2 \)); and
3. \( x = \pm 2\sqrt{\frac{2}{3}}, y = 0 \) (\( F = \frac{1}{4}(a - c)^2 \)).

(Verify!) Now, for the boundary of \( \mathcal{C} \), 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
\[
h(x) = (a^2 - b^2)x^2 + b^2 - ((a - b)x^2 + b^2)^2, \quad h : \mathbb{E}^1 \to \mathbb{E}^1
\]
on the interval \( [-1, 1] \subseteq \mathbb{E}^1 \). In \( (-1, 1) \) the derivative
\[
h'(x) = 2(a - b)x(1 - 2x^2)
\]
vanishes only when
4. \( x = 0 \) (\( h = 0 \)), and
5. \( x = \pm 2\sqrt{\frac{2}{3}} \) (\( h = \frac{1}{4}(a - b)^2 \)).
Finally, at the endpoints of \([-1, 1]\), we have
\[ (6) \quad x = \pm 1 \quad (h = 0). \]

Comparing the resulting function values in all six cases, we conclude that the least of them is 0, while the largest is \(\frac{1}{2}(a - c)^2\). These are the desired least and largest values of \(f\), subject to the conditions stated. They are attained, respectively, at the points
\[ (0, 0, \pm 1), \quad (0, \pm 1, 0), \quad (\pm 1, 0, 0), \quad \text{and} \quad (\pm 2^{\frac{1}{2}}, 0, \pm 2^{\frac{1}{2}}). \]

Again, the use of Theorems 2 and 3 was redundant.\(^1\) 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
\[ x^2 + y^2 + z^2 = 1 \]
(which geometrically limits all to the surface of \(G_0(1)\) in \(E^3\)), so that one of them, \(z\), could be eliminated. Only then can Theorems 1 to 3 be used.

**Problems on Maxima and Minima**

1. Verify Note 1.

2. \(^\dagger\) Complete the missing details in the proof of Theorems 2 and 3.

3. Test \(f\) for extrema in \(E^2\) if \(f(x, y)\) is

   \[ (i) \quad \frac{x^2}{2p} + \frac{y^2}{2q} \quad (p > 0, \quad q > 0); \]

   \[ (ii) \quad \frac{x^2}{2p} - \frac{y^2}{2q} \quad (p > 0, \quad q > 0); \]

   \[ (iii) y^2 + x^4; \]

   \[ (iv) y^2 + x^3. \]

4. \(\dagger\) 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: \(x + y + z = a\).
   
   (i) Do the same in \(E^4\) and in \(E^n\); show that \(A\) is a cube.

\(^1\) Indeed, by Theorem 2(ii) in Chapter 4, §8, absolute extrema must exist here, as all is limited to the compact sphere, \(x^2 + y^2 + z^2 = 1\).

\[ \text{§9. Local Extrema. Maxima and Minima} \]

(iii) Hence deduce that
\[ \sqrt[3]{x_1x_2 \cdots x_n} \leq \frac{1}{n} \sum_{i=1}^{n} x_i \quad (x_k \geq 0), \]

i.e., the geometric mean of \(n\) nonnegative numbers is \(\leq\) their arithmetic mean.

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: \(x = y = z = t = c; \quad f_{\text{max}} = 4c\).]

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 \(z = 2\pi - (x + y)\).]

7. Among all intervals \(A \subset E^3\) inscribed in the ellipsoid
\[ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \]
find the one of largest volume.

[Answer: the edge lengths are \(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}, \frac{c}{\sqrt{2}}\).]

8. Let \(P_i = (a_i, b_i), \quad i = 1, 2, 3\), be 3 points in \(E^2\) forming a triangle in which one angle (say, \(\angle P_1\)) is \(\geq 2\pi/3\).

Find a point \(P = (x, y)\) for which the sum of the distances,
\[ PP_1 + PP_2 + PP_3 = \sqrt{(x - a_1)^2 + (y - b_1)^2} + \sqrt{(x - a_2)^2 + (y - b_2)^2} + \sqrt{(x - a_3)^2 + (y - b_3)^2}, \]
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 \(P_1, P_2, P_3\) (and so \(P_1, P_2, P_3\) are critical points at which an extremum may occur), while at other points \(P\) parts do exist but never vanish simultaneously, so that there are no other critical points. Indeed, prove that \(D_1 f(P) = 0 = D_2 f(P)\) would imply that
\[ \sum_{i=1}^{3} \cos \theta_i = 0 = \sum_{i=1}^{3} \sin \theta_i, \]
where \(\theta_i\) is the angle between \(PP_i\) and the \(x\)-axis; hence
\[ \sin(\theta_1 - \theta_2) = \sin(\theta_2 - \theta_3) = \sin(\theta_3 - \theta_1) \quad (\text{why?}), \]
and so \(\theta_1 - \theta_2 = \theta_2 - \theta_3 = \theta_3 - \theta_1 = 2\pi/3\), 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.)]
9. Continuing Problem 8, show that if none of \( \mathcal{L}P_1, \mathcal{L}P_2, \) and \( \mathcal{L}P_3 \) is \( \geq 2\pi/3 \), then \( f \) attains its least value at some \( P \) (inside the triangle) such that \( \mathcal{L}P_1P_2P = \mathcal{L}P_2P_3P = \mathcal{L}P_3P_1P = 2\pi/3 \).

[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 \( P_1P_2 + P_1P_3 > PP_1 + PP_1 + PP_1, \) 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.]

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
\[
A = \frac{1}{2}R^2 \sum_{k=1}^{n+1} \sin x_k,
\]
with \( x_{n+1} = 2\pi - \sum_{k=1}^{n} x_k \). (Why?) Thus all reduces to minimizing
\[f(x_1, \ldots, x_n) = \sum_{k=1}^{n} \sin x_k + \sin\left(2\pi - \sum_{k=1}^{n} x_k\right),\]
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 \( D \subset \mathbb{E}^n \) (called a simplex). Equating all partials of \( f \) to 0, show that the only critical point interior to \( D \) is \( \bar{x} = (x_1, \ldots, x_n) \), with \( x_k = 2\pi/n \) \( k \leq n \) (implying that \( x_{n+1} = \frac{2\pi}{n} \)). For that \( \bar{x} \), we get
\[f(\bar{x}) = (n + 1)\sin(2\pi/n + 1)).\]
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 \( f(\bar{x}) \) is indeed the largest when all \( x_k \) equal \( \frac{2\pi}{n} \). 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 \( n \sin(2\pi/n) \leq (n + 1) \sin(2\pi/n + 1)) \), the induction is complete.
Thus, the area \( A \) is the largest when the polygon is regular, for which
\[A = \frac{1}{2}R^2(n + 1)\sin\left(\frac{2\pi}{n + 1}\right)\]

11. Among all triangles of a prescribed perimeter \( 2p \), find the one of maximum area.

[Hint: Maximize \( p(x - y)(y - z) \) on the condition that \( x + y + z = 2p \).]

12. Among all triangles of area \( A \), find the one of smallest perimeter.

13. Find the shortest distance from a given point \( \bar{p} \in \mathbb{E}^n \) to a given plane \( \bar{a} \cdot \bar{x} = c \) (Chapter 3, \S 4-6). Answer:
\[d \cdot \bar{p} - c \]

\[
\pm \frac{d \cdot \bar{p} - c}{|d|},
\]

§9. Local Extrema. Maxima and Minima

14. Differentiate \( f(x,y) = \sqrt{x^2 + y^2} \) on the “faces” of the simplex \( D \) (see Note 4).

[Hint: First do it in \( \mathbb{E}^3 \), writing \( (x, y, z) \) for \( \bar{x} \).]

§10. More on Implicit Differentiation. Conditional Extrema

I. Implicit differentiation was sketched in \S 7. Under suitable assumptions (Theorem 4 in \S 7), one can differentiate a given system of equations,
\[g_k(x_1, \ldots, x_n, y_1, \ldots, y_m) = 0, \quad k = 1, 2, \ldots, n,
\]
treating the \( x_i \) as implicit functions of the \( y_i \) without seeking an explicit solution of the form
\[x_j = H_j(y_1, \ldots, y_m).
\]
This yields a new system of equations from which the partials \( D_iH_j = \frac{\partial x_j}{\partial y_i} \) can be found directly.

We now supplement Theorem 4 in \S 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 \S 7, we set
\[
(\bar{x}, \bar{y}) = (x_1, \ldots, x_n, y_1, \ldots, y_m) \text{ and } g = (g_1, \ldots, g_n),
\]
replacing the \( f \) of \S 7 by \( g \). Then equations (1) simplify to
\[g(\bar{x}, \bar{y}) = 0,
\]
where \( g : \mathbb{E}^{n+m} \to \mathbb{E}^n \) (or \( g : C^{n+m} \to C^n \)).

Theorem 1 (implicit differentiation). Adopt all assumptions of Theorem 4 in \S 7, replacing \( f \) by \( g \) and setting \( H = (H_1, \ldots, H_n) \),
\[D_i g_j(\bar{p}, \bar{q}) = a_{ij}, \quad j \leq n + m, \quad k \leq n.
\]
Then for each \( i = 1, \ldots, m \), we have \( n \) linear equations,
\[\sum_{j=1}^{n} a_{jk}D_iH_j(\bar{q}) = -a_{n+i,k}, \quad k \leq n,
\]
with
\[\det(a_{ij}) \neq 0, \quad (j, k \leq n),
\]
that uniquely determine the partials \( D_iH_j(\bar{q}) \) for \( j = 1, 2, \ldots, n \).

Proof. As usual, extend the map \( H : Q \to P \) of Theorem 4 in \S 7 to \( H : E^m \to E^n \) (or \( C^m \to C^n \)) by setting \( H = 0 \) on \( -Q \).
Also, define \( \sigma : E^m \to E^{n+m} \) (or \( C^m \to C^{n+m} \)) by
\[\sigma(\bar{y}) = (H(\bar{y}), \bar{y}) = (H_1(\bar{y}), \ldots, H_n(\bar{y}), y_1, \ldots, y_m).
\]
\[\bar{y} \in E^m(C^n).
\]
Then $\sigma$ 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,

$$g(\sigma(\vec{q})) = g(H(\vec{y}), \vec{y}) = \vec{0} \text{ for } \vec{y} \in Q;$$

i.e., $g \circ \sigma = \vec{0}$ on $Q$.

Now, by assumption, $g \in \mathcal{C}^1$ 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 $h = g \circ \sigma = \vec{0}$ on $Q$, an open set, the partials of $h$ vanish on $Q$. So by Theorem 2 of §4, writing $\sigma_j$ for the $j$th component of $\sigma$,

$$\vec{0} = \sum_{j=1}^{n+m} D_j g(\vec{p}, \vec{q}) \cdot D_j \sigma_j(\vec{q}), \quad i \leq m.$$

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),

$$\vec{0} = \sum_{j=1}^n D_j g(\vec{p}, \vec{q}) \cdot D_j H_j(\vec{q}) + D_{n+i} g(\vec{p}, \vec{q}), \quad i \leq 1, 2, \ldots, m.$$

As $g = (g_1, \ldots, g_n)$, each of these vector equations splits into $n$ scalar ones:

$$0 = \sum_{j=1}^n D_j g_k(\vec{p}, \vec{q}) \cdot D_j H_j(\vec{q}) + D_{n+i} g_k(\vec{p}, \vec{q}), \quad i \leq m, \quad k \leq n.$$

With $D_j g_k(\vec{p}, \vec{q}) = a_{jk}$, this yields (3), where $\det(a_{jk}) = \det(D_j g_k(\vec{p}, \vec{q})) \neq 0$, by hypothesis (see Theorem 4 in §7).

Thus all is proved. $\square$

**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 $f: \mathcal{E}^{n+m} \to \mathcal{E}^1$ has a local conditional maximum (minimum) at $\vec{p} \in \mathcal{E}^{n+m}$, with constraints

$$g = (g_1, \ldots, g_n) = \vec{0}$$

§10. More on Implicit Differentiation. Conditional Extrema

$$(g : \mathcal{E}^{n+m} \to \mathcal{E}^n) \text{ iff in some neighborhood } G \text{ of } \vec{p} \text{ we have}$$

$$\Delta f = f(\vec{x}) - f(\vec{p}) \leq 0 \quad (\geq 0, \text{ respectively})$$

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** $f: \mathcal{E}^{n+m} \to \mathcal{E}^1$, set

$$F = f + \sum_{k=1}^n c_k g_k,$$

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

$$D_j F(\vec{x}) = 0, \quad j \leq n + m, \quad \text{and} \quad g_k(\vec{\vec{x}}) = 0, \quad k \leq n,$$

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 again we write $(\vec{p}, \vec{q})$ for $\vec{p}$ and $(\vec{x}, \vec{y})$ for $\vec{x}$ (we call it “double notation”).

**Theorem 2** (Lagrange multipliers). Suppose $f: \mathcal{E}^{n+m} \to \mathcal{E}^1$ is differentiable at

$$(\vec{p}, \vec{q}) = (p_1, \ldots, p_n, q_1, \ldots, q_m)$$

and has a local extremum at $(\vec{p}, \vec{q})$ subject to the constraints

$$g = (g_1, \ldots, g_n) = \vec{0},$$

with $g$ as in Theorem 1, $g: \mathcal{E}^{n+m} \to \mathcal{E}^n$. Then

$$\sum_{k=1}^n c_k D_j g_k(\vec{p}, \vec{q}) = -D_j f(\vec{p}, \vec{q}), \quad j = 1, 2, \ldots, n + m,$$

for certain multipliers $c_k$ (determined by the first $n$ equations in (9)).

$^1$ That is, $D_j F(\vec{p}, \vec{q}) = 0$, with $F$ as in (7).
Proof. These \( n \) equations admit a unique solution for the \( c_k \), as they are linear, and

\[
\det(D_{jk}g_k(\vec{p}, \vec{q})) \neq 0 \quad (j, k \leq n)
\]

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

\[
f(\vec{x}, \vec{y}) - f(\vec{p}, \vec{q}) \leq 0 \quad (\text{or} \quad \geq 0)
\]

for all \((\vec{x}, \vec{y}) \in P \times Q\) with \( g(\vec{x}, \vec{y}) = 0\), provided we make the neighborhood \( P \times Q \) small enough.

Define \( H \) and \( \sigma \) as in the previous proof (see (4)); so \( \vec{x} = H(\vec{y}) \) is equivalent to \( g(\vec{x}, \vec{y}) = 0 \) 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 \( g(\vec{x}, \vec{y}) = 0 \) and also

\[
f(\vec{x}, \vec{y}) = f(H(\vec{y}), \vec{y}) = f(\sigma(\vec{y})).
\]

Set \( h = f \circ \sigma, h: E^n \to E^1 \). Then (10) reduces to

\[
h(\vec{y}) - h(\vec{q}) \leq 0 \quad (\text{or} \quad \geq 0) \quad \text{for all} \ \vec{y} \in Q.
\]

This means that \( h \) has an unconditional extremum at \( \vec{q} \), an interior point of \( Q \). Thus, by Theorem 1 in §9,

\[
D_i h(\vec{q}) = 0, \quad i = 1, \ldots, m.
\]

Hence, applying the chain rule (Theorem 2 of §4) to \( h = f \circ \sigma \), we get, much as in the previous proof,

\[
0 = \sum_{j=1}^{n+m} D_j f(\vec{p}, \vec{q}) D_j \sigma(\vec{q})
\]

for all \((\vec{x}, \vec{y}) \in P \times Q\) with \( g(\vec{x}, \vec{y}) = 0\), provided we make the neighborhood \( P \times Q \) small enough.

Then, for all such \((\vec{x}, \vec{y})\), with \( \vec{x} = H(\vec{y}) \), we surely have \( g(\vec{x}, \vec{y}) = 0 \) and also

\[
f(\vec{x}, \vec{y}) = f(H(\vec{y}), \vec{y}) = f(\sigma(\vec{y})).
\]

Set \( h = f \circ \sigma, h: E^n \to E^1 \). Then (10) reduces to

\[
h(\vec{y}) - h(\vec{q}) \leq 0 \quad (\text{or} \quad \geq 0) \quad \text{for all} \ \vec{y} \in Q.
\]

(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

\[
\sum_{k=1}^{n} c_k D_{n+k} g_k(\vec{p}, \vec{q}) = -D_{n+k} f(\vec{p}, \vec{q}), \quad i = 1, 2, \ldots, m.
\]

Hence (9) holds for \( n < j \leq n + m \), too, and all is proved. \( \square \)

Remarks. Lagrange’s method has the advantage that all variables (the \( x_k \) and \( y_j \)) 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 3 (sufficient conditions). Let

\[
F = f + \sum_{k=1}^{n} c_k g_k,
\]

with \( f: E^{n+m} \to E^1 \), \( g: E^{n+m} \to E^m \), and \( c_k \) as in Theorem 2.

Then \( f \) has a maximum (minimum) at \( \vec{p} = (p_1, \ldots, p_{n+m}) \) (with constraints \( g = (g_1, \ldots, g_n) = 0 \)) 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 \( g = 0 \). Then

\[
0 \geq F(\vec{x}) - F(\vec{p}) = f(\vec{x}) - f(\vec{p}) + \sum_{k=1}^{n} c_k [g_k(\vec{x}) - g_k(\vec{p})]
\]

for those \( \vec{x} \) near \( \vec{p} \) (including \( \vec{x} = \vec{p} \)) for which \( g(\vec{x}) = 0 \).

But for such \( \vec{x} \), \( g_k(\vec{x}) = g_k(\vec{p}) = 0 \), \( c_k [g_k(\vec{x}) - g_k(\vec{p})] = 0 \), and so

\[
0 \geq F(\vec{x}) - F(\vec{p}) = f(\vec{x}) - f(\vec{p})
\]

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 \( \vec{p} \). \( \square \)

Example 1.

Find the local extrema of

\[
f(x, y, z, t) = x + y + z + t
\]
on the condition that
\[ g(x, y, z, t) = xyzt - a^4 = 0, \]
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
\[ 0 = 1 + cxyzt = 1 + czyt = 1 + cxyz, \quad xyzt - a^4 = 0. \]

Solving for \( x, z \), and \( t \), we get \( c = -a^{-3}, x = y = z = t = a \).

Thus \( F(x, y, z, t) = x + y + z + t - xyzt/a^4 \), and the only critical point is \( \vec{p} = (a, a, a, a) \). (Verify!)

By Theorem 3, one can now explore the sign of \( F(x) - F(\vec{p}) \), where \( x = (x, y, z, t) \). For \( \vec{x} \) near \( \vec{p} \), it agrees with the sign of \( d^2F(\vec{p}, \cdot) \). (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 \( d^2F(\vec{p}, \cdot) = 0 \) on \( E^{n+m} \) (briefly \( dF \equiv 0 \)).

Conversely, if \( d^2f(\vec{p}, \cdot) = 0 \) on a globe \( G_{\vec{p}} \), for some function \( f \) on \( n \) independent variables, then
\[ D_k f(\vec{p}) = 0, \quad k = 1, 2, \ldots, n, \]
since \( d^2 f(\vec{p}, \cdot) \) (a polynomial!) vanishes at infinitely many points if its coefficients \( D_k f(\vec{p}) \) vanish. (The latter fails, however, if the variables are independent.)

Thus, instead of working with the partials, one can equate to 0 the differential of \( F \) 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 \( x, y, z \),
\[ t = a^4/(xyz) = H(x, y, z), \]
and differentiate the identity \( xyzt - a^4 = 0 \) to obtain
\[ 0 = yzt \, dx + xzt \, dy + xyt \, dz + xyz \, dt; \]

so
\[ dt = -t \left( \frac{dx}{x} + \frac{dy}{y} + \frac{dz}{z} \right). \]

Substituting this value of \( dt \) in \( df = dx + dy + dz + dt = 0 \) (the equation for critical points), we eliminate \( dt \) and find:
\[ (1 - \frac{t}{x}) \, dx + (1 - \frac{t}{y}) \, dy + (1 - \frac{t}{z}) \, dz = 0. \]

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
\[ 1 - \frac{t}{x} = 1 - \frac{t}{y} = 1 - \frac{t}{z} = 0. \]

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
\[ d^2F = -\frac{2}{a} (dx \, dy + dx \, dz + dy \, dt + dx \, dt + dy \, dz + dy \, dt). \]

(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
\[ dt = -t(a^4 + dy + dz), \]
and (13) turns into
\[ -\frac{2}{a} \left[ dx \, dy + dx \, dz + dy \, dz - (dx \, dy + dy \, dz)^2 \right] \]
\[ = \frac{1}{a} \left[ dx^2 + dy^2 + dz^2 + (dx \, dy + dy \, dz)^2 \right] = d^2F. \]

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 is mandatory.

**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 \( (D_g(x, \vec{q})) \) 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.

**Further Problems on Maxima and Minima**

1. Fill in all details in Examples 1 and 2 and the proofs of all theorems in this section.

2. Redo Example (B) in §9 by Lagrange’s method.

3. An ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

is cut by a plane $ux + vy + wz = 0$. Find the semiaxes of the section-ellipse, i.e., the extrema of

$$\rho^2 = \left| f(x, y, z) \right|^2 = x^2 + y^2 + z^2$$

under the constraints $g = (g_1, g_2) = 0$, where

$$g_1(x, y, z) = ux + vy + wz \quad \text{and} \quad g_2(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1.$$ 

Assume that $a > b > c > 0$ and that not all $a, v, w = 0$.

[Outline: By Note 2, explore the rank of the matrix (14)]

$$\begin{pmatrix} x/a^2 & y/b^2 & z/c^2 \\ u & v & w \end{pmatrix}.$$ 

(Why this particular matrix?)

Seeking a contradiction, suppose all its $2 \times 2$ 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

$$F(x, y, z) = x^2 + y^2 + z^2 + r\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) + 2s(ux + vy + wz).$$

Equate $dF$ to 0:

$$x + \frac{fr}{a^2} + su = 0, \quad y + \frac{fr}{b^2} + sv = 0, \quad z + \frac{fr}{c^2} + sw = 0.$$  

Multiplying by $x, y, z$, respectively, adding, and combining with $g = 0$, obtain $r = -\rho^2$; so, by (15), for $a, b, c \neq \rho$,

$$x = \frac{-suv^2}{\rho^2 - \rho^2}, \quad y = \frac{-sv^2}{\rho^2 - \rho^2}, \quad z = \frac{-suv^2}{\rho^2 - \rho^2}.$$ 

Find $s, x, y, z$, then compare the $\rho$-values at critical points.

---

2 This suffices here, since the equation $g = 0$ defines a compact set $S$; see §9.

---

4. Find the least and the largest values of the quadratic form

$$f(\vec{x}) = \sum_{i,k=1}^{n} a_{ik} x_i x_k \quad (a_{ik} = a_{ki})$$

on the condition that $g(\vec{x}) = |\vec{x}|^2 - 1 = 0$ ($f, g: E^n \to E^1$).

[Outline: Let $F(\vec{x}) = f(\vec{x}) - t(x_1^2 + x_2^2 + \ldots + x_n^2)$. Equating $dF$ to 0, obtain

$$(a_{11} - t) x_1 + a_{12} x_2 + \ldots + a_{1n} x_n = 0,$$

$$(a_{21} + (a_{22} - t) x_2 + \ldots + a_{2n} x_n = 0,$$

$$\ldots \ldots \ldots$$

(16)

Using Theorem 1(iv) in §6, derive the so-called characteristic equation of $f$,

$$|a_{11} - t | a_{12} \ldots a_{1n} | a_{21} a_{22} - t \ldots a_{2n} | a_{n1} a_{n2} \ldots a_{nn} - t | = 0.$$ 

(17)

of degree $n$ in $t$. If $t$ is one of its $n$ roots (known to be real$^3$), then equations (16) admit a nonzero solution for $\vec{x} = (x_1, \ldots, x_n)$; by replacing $\vec{x}$ by $\vec{\rho}/|\vec{\rho}|$ if necessary, $\vec{x}$ satisfies also the constraint equation $g(\vec{x}) = |\vec{x}|^2 - 1 = 0$. Thus each root $t$ of (17) yields a critical point $\vec{x}_1 = (x_1, \ldots, x_n)$.

Now, to find $f(\vec{x}_1)$, multiply the $t$th equation in (16) by $x_k$, $k = 1, \ldots, n$, and add to get

$$0 = \sum_{i,k=1}^{n} a_{ik} x_i x_k - t \sum_{k=1}^{n} x_k^2 = f(\vec{x}_1) - t.$$ 

Hence $f(\vec{x}_1) = t$.

Thus the values of $f$ at the critical points $\vec{x}_1$ are simply the roots of (17). The largest (smallest) root is also the largest (least) value of $f$ on $S = \{ \vec{x} \in E^n \mid |\vec{x}| = 1 \}$. (Explain!)]

5. Use the method of Problem 4 to find the semiaxes of

(i) the quadric curve in $E^2$, centered at $\vec{0}$, 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 $\vec{0}$.

Assume $a_{ik} = a_{ki}$.

[Hint: Explore the extrema of $f(\vec{x}) = |\vec{x}|^2$ on the condition that $g(\vec{x}) = \sum_{i,k=1}^{n} a_{ik} x_i x_k - 1 = 0$.]


7. In $E^2$, find the shortest distance from $\vec{0}$ to the parabola $y^2 = 2(x + a)$. 

---

8. In $E^3$, find the shortest distance from $\vec{0}$ 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 $\vec{0}$. (Rewrite all in coordinate form!)

9. In $E^n$, find the largest value of $|\vec{a} \cdot \vec{x}|$ if $|\vec{x}| = 1$. Use Lagrange’s method.

10. (Hadamard’s theorem.) If $A = \det(x_{ik})$ ($i, k \leq n$), then

$$|A| \leq \prod_{i=1}^{n} |\vec{x}_i|,$$

where $\vec{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in})$.

[Hints: Set $a_i = |\vec{x}_i|$. Treat $\det$ 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 $\det(y_{ik})$, with $y_{ik} = 0$ (if $i \neq k$) and $y_{ii} = a_i^2$.]

Chapter 7
Volume and Measure

Our intuitive idea of “volume” is rather vague. We just tend to assume that “bodies” in space (i.e., in $E^3$) somehow have numerically expressed “volumes,” but it remains unclear which sets in $E^3$ are “bodies” and how volume is defined.

We also intuitively assume that volumes behave “additively.” That is, if a body is split into disjoint parts, then the volume of the whole equals the sum of the volumes of the parts. Similarly for “areas” in $E^2$. In elementary calculus, that is often just taken for granted.

The famous mathematician Henri Lebesgue (1875–1941) extended the idea of “volume” to a large, strictly defined family of sets in $E^n$, called Lebesgue-measurable sets, thus giving rise to what is called measure theory. Its basic idea remains that of additivity, precisely formulated and proved. Modern theory has still more generalized these ideas. In this text, we have so far defined “volumes” for intervals in $E^n$ only. Thus it is natural to take intervals as our starting point. This will also lead to the important idea of a semiring of sets and its extension: a ring of sets.

§1. More on Intervals in $E^n$. Semirings of Sets

1. As a prologue, we turn to intervals in $E^n$ (Chapter 3, §7).

Theorem 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

$$A = X \times Y$$
$$B = X' \times Y',$$

where $X, Y, X'$, and $Y'$ are intervals in $E^1$. Then (see Figure 29)

$$A \cap B = (X \times Y) \cap (X' \times Y') = (X \cap X') \times (Y \cap Y')$$
and
\[ A - B = [(X - X') \times Y] \cup [(X \cap X') \times (Y - Y')]; \]

see Problem 8 in Chapter 1, §§1-3.

As the theorem holds in \( E^1 \),
\[ X \cap X' \text{ and } Y \cap Y' \]
are intervals in \( E^1 \), while
\[ X - Y' \text{ and } Y - Y' \]
are finite unions of disjoint line intervals.
(In Figure 29 they are 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 an element), 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 if it satisfies Theorem 1.

Note that here \( \mathcal{C} \) is not just a set, but a whole family of sets. Recall (Chapter 1, §§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:

\[ \mathcal{M} = \{ X_i \mid i \in I \}, \]

briefly
\[ \mathcal{M} = \{ X_i \}, \]

where the \( X_i \) are \( \mathcal{M} \)-sets distinguished from each other by the subscripts \( i \) varying over some index set \( I \).

§1. More on Intervals in \( E^n \). Semirings of Sets

A set family \( \mathcal{M} = \{ X_i \} \) and its union

\[ \bigcup_i X_i \]

are said to be **disjoint** iff

\[ X_i \cap X_j = \emptyset \text{ whenever } i \neq j. \]

Notation:

\[ \bigcup_i X_i \text{ (disjoint)}. \]

In our case, \( A \in \mathcal{C} \) means that \( A \) is a \( \mathcal{C} \)-set (a member of the semiring \( \mathcal{C} \)).

The formula

\[(\forall A, B \in \mathcal{C}) \quad A \cap B \in \mathcal{C} \]

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** if and write

\[ A \in \mathcal{C}_s \]

if \( A \) is a finite union of disjoint \( \mathcal{C} \)-sets (such as \( A - B \) in Theorem 1).

Thus \( \mathcal{C}_s \) is the family of all \( \mathcal{C} \)-simple sets.

Every \( \mathcal{C} \)-set is also a \( \mathcal{C}_s \)-set, i.e., a \( \mathcal{C} \)-simple one. (Why?) Briefly:

\[ \mathcal{C} \subseteq \mathcal{C}_s. \]

If \( \mathcal{C} \) is the set of all intervals, a \( \mathcal{C} \)-simple set may look as in Figure 30.

**Theorem 2.** If \( \mathcal{C} \) is a semiring, and if \( A \) and \( B \) are \( \mathcal{C} \)-simple, so also are

\[ A \cap B, \ A - B, \text{ and } A \cup B. \]

In symbols,

\[(\forall A, B \in \mathcal{C}_s) \quad A \cap B \in \mathcal{C}_s, \ A - B \in \mathcal{C}_s, \text{ and } A \cup B \in \mathcal{C}_s.\]
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 $A \cap B \in C'_s$, let
   \[ A = \bigcup_{i=1}^{m} A_i \text{ (disjoint)} \quad \text{and} \quad B = \bigcup_{k=1}^{n} B_k \text{ (disjoint)}, \]
   with $A_i, B_k \in C$. Verify that
   \[ A \cap B = \bigcup_{i=1}^{m} \bigcap_{k=1}^{n} (A_i \cap B_k) \text{ (disjoint)}, \]
   and so $A \cap B \in C'_s$.

2. Next prove that $A - B \in C'_s$ if $A \in C'_s$ and $B \in C$.
   Indeed, if
   \[ A = \bigcup_{i=1}^{m} A_i \text{ (disjoint)}, \]
   then
   \[ A - B = \bigcup_{i=1}^{m} A_i - B = \bigcup_{i=1}^{m} (A_i - B) \text{ (disjoint)}. \]
   Verify and use Definition 2.

3. Prove that
   \[ (\forall A, B \in C'_s) \quad A - B \in C'_s, \]
   we suggest the following argument.
   Let
   \[ B = \bigcup_{k=1}^{n} B_k, \quad B_k \in C. \]
   Then
   \[ A - B = A - \bigcap_{k=1}^{n} B_k = \bigcap_{k=1}^{n} (A - B_k) \]
   by duality laws. But $A - B_k$ is $C$-simple by step (2). Hence so is
   \[ A - B = \bigcap_{k=1}^{n} (A - B_k) \]
   by step (1) plus induction.

4. To prove $A \cup B \in C'_s$, verify that
   \[ A \cup B = A \cup (B - A), \]
   where $B - A \in C'_s$, by (3).

Note 2. By induction, Theorem 2 extends to any finite number of $C'_s$-sets. It is a kind of “closure law.”

We thus briefly say that $C'_s$ 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 $C$ is a semiring, then $C'_s$ is a ring.

Caution. An infinite union of $C$-simple sets need not be $C$-simple. Yet we may consider such unions, as we do next.

In Corollary 1 below, $C'_s$ may be replaced by any set ring $M$.

Corollary 1. If $\{A_n\}$ is a finite or infinite sequence of sets from a semiring $C$ (or from a set ring $M$, such as $C'_s$), then there is a disjoint sequence of $C$-simple sets (or $M$-sets) $B_n \subseteq A_n$ such that
   \[ \bigcup_{n} A_n = \bigcup_{n} B_n. \]

Proof. Let $B_1 = A_1$ and for $n = 1, 2, \ldots$,
   \[ B_{n+1} = A_{n+1} - \bigcup_{k=1}^{n} A_k, \quad A_k \in C. \]
   By Theorem 2, the $B_n$ are $C$-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 $\bigcup_{n=1}^{n} A_n = \bigcup_{n=1}^{n} B_n$. If $x \in \bigcup_{n=1}^{n} A_n$, take the least $n$ for which $x \in A_n$. Then $n > 1$ and
   \[ x \in A_n - \bigcup_{k=1}^{n-1} A_k = B_n, \]
   or $n = 1$ and $x \in A_1 = B_1$. \(\square\)

Note 3. In Corollary 1, $B_n \in C'_s$, i.e., $B_n = \bigcup_{m=1}^{m} C_{ni}$ for some disjoint sets $C_{ni} \in C$. Thus
   \[ \bigcup_{n=1}^{n} A_n = \bigcup_{n=1}^{m} C_{ni} \]
   is also a countable disjoint union of $C$-sets.

II. Recall that the volume of intervals is additive (Problem 9 in Chapter 3, §7). That is, if $A \in C$ 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 $X_1, X_2, \ldots, X_m \in C$ (intervals in $E^n$). If the $X_i$ are mutually disjoint, then

(i) $\bigcup_{i=1}^{m} X_i \subseteq Y \in C$ 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 $Y_k \in C$) implies $\sum_{i=1}^{m} vX_i \leq \sum_{k=1}^{p} vY_k$.

**Proof.** (i) By Theorem 2, the set

$$Y - \bigcup_{i=1}^{m} X_i$$

is $C$-simple; so

$$Y - \bigcup_{i=1}^{m} X_i = \bigcup_{j=1}^{q} C_j$$

for some disjoint intervals $C_j$. Hence

$$Y = \bigcup_{i=1}^{m} X_i \cup \bigcup_{j=1}^{q} C_j \text{ (all disjoint)}.$$

Thus by additivity,

$$vY = \sum_{i=1}^{m} vX_i + \sum_{j=1}^{q} vC_j \geq \sum_{i=1}^{m} vX_i,$$

as claimed.

(ii) By set theory (Problem 9 in Chapter 1, §1–3),

$$X_i \subseteq \bigcup_{k=1}^{p} Y_k$$

implies

$$X_i = X_i \cap \bigcup_{k=1}^{p} Y_k = \bigcup_{k=1}^{p} (X_i \cap Y_k).$$

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,

$$vX_i = \sum_{k=1}^{p} v(X_i \cap Y_k).$$

Thus we can (and do) unambiguously define $vA$ to be either of these sums.

Hence

$$\sum_{i=1}^{m} vX_i = \sum_{i=1}^{m} \sum_{k=1}^{p} v(X_i \cap Y_k) = \sum_{k=1}^{p} \sum_{i=1}^{m} v(X_i \cap Y_k).$$

But by (i),

$$\sum_{i=1}^{m} v(X_i \cap Y_k) \leq vY_k \text{ (why?)};$$

so

$$\sum_{i=1}^{m} vX_i \leq \sum_{k=1}^{p} vY_k,$$

as required.

If, however, the $Y_k$ are not disjoint, Corollary 1 yields

$$\bigcup_{k=1}^{p} Y_k = \bigcup_{j=1}^{m_k} B_k \text{ (disjoint)},$$

with

$$Y_k \supseteq B_k = \bigcup_{j=1}^{m_k} C_{kj} \text{ (disjoint)}, \quad C_{kj} \in C.$$

By (i),

$$\sum_{j=1}^{m_k} vC_{kj} \leq vY_k.$$

As

$$\bigcup_{i=1}^{m} X_i \subseteq \bigcup_{k=1}^{p} Y_k = \bigcup_{k=1}^{p} B_k = \bigcup_{j=1}^{m_k} C_{kj} \text{ (disjoint)},$$

all reduces to the previous disjoint case. \(\square\)

**Corollary 2.** Let $A \in C_0' \subseteq C$ (intervals in $E^n$). If

$$A = \bigcup_{i=1}^{m} X_i \text{ (disjoint)} = \bigcup_{k=1}^{p} Y_k \text{ (disjoint)}$$

with $X_i, Y_k \in C$, then

$$\sum_{i=1}^{m} vX_i \leq \sum_{k=1}^{p} vY_k.$$

(Use part (ii) of the lemma twice.)
Problems on Intervals and Semirings

1. Complete the proof of Theorem 1 and Note 1.
2. Prove Theorem 2 in detail.
3. Fill in the details in the proof of Corollary 1.

1'. Prove Corollary 2.

2. Show that, in the definition of a semiring, the condition \( \emptyset \in \mathcal{C} \) is equivalent to \( \mathcal{C} \neq \emptyset \).
   [Hint: Consider \( \emptyset = A - A = \bigcup_{i=1}^{m} A_i \), \( (A, A_i \in \mathcal{C}) \) to get \( \emptyset = A_i \in \mathcal{C} \).]

3. Given a set \( S \), show that the following are semirings or rings.
   - (a) \( \mathcal{C} = \{ \text{all subsets of } S \} \);
   - (b) \( \mathcal{C} = \{ \text{all finite subsets of } S \} \);
   - (c) \( \mathcal{C} = \{ \emptyset \} \);
   - (d) \( \mathcal{C} = \{ \emptyset \text{ and all singletons in } S \} \).
   Disprove it for \( \mathcal{C} = \{ \emptyset \text{ and all two-point sets in } S \} \), \( S = \{1, 2, 3, \ldots \} \).
   In (a)-(c), show that \( \mathcal{C}'_n = \mathcal{C} \). Disprove it for (d).

4. Show that the cubes in \( E^n \) (\( n > 1 \)) do not form a semiring.

5. Using Corollary 2 and the definition thereafter, show that volume is additive for \( \mathcal{C} \)-simple sets. That is,
   
   \[
   \text{if } A = \bigcup_{i=1}^{m} A_i \text{ (disjoint) then } vA = \sum_{i=1}^{m} vA_i \quad (A, A_i \in \mathcal{C}'_n).
   \]

6. Prove the lemma for \( \mathcal{C} \)-simple sets.
   [Hint: Use Problem 6 and argue as before.]

7. Prove that if \( \mathcal{C} \) is a semiring, then \( \mathcal{C}'_n (\mathcal{C} \text{-simple sets}) = \mathcal{C}_n \), the family of all finite unions of \( \mathcal{C} \)-sets (disjoint or not).
   [Hint: Use Theorem 2.]

8. Problems on Intervals and Semirings

   §2. \( \mathcal{C}_n \)-Sets. Countable Additivity. Permutable Series

   We now want to further extend the definition of volume by considering countable unions of intervals, called \( \mathcal{C}_n \)-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 \( \sigma \)-additivity (the \( \sigma \) is used whenever countable unions are involved).

   We need two lemmas in addition to that of \( \S1 \).

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

   \[
   A \subseteq B \subseteq C
   \]

   and

   \[
   vC - \varepsilon < vB < vA + \varepsilon.
   \]

Proof. Let the endpoints of \( B \) be \( \bar{a} = (a_1, \ldots, a_n) \) and \( \bar{b} = (b_1, \ldots, b_n) \).

   For each natural number \( i \), consider the open interval \( C_i \), with endpoints

   \[
   (a_1 - \frac{1}{i}, a_2 - \frac{1}{i}, \ldots, a_n - \frac{1}{i})\quad \text{and} \quad (b_1 + \frac{1}{i}, b_2 + \frac{1}{i}, \ldots, b_n + \frac{1}{i}).
   \]

   Then \( B \subseteq C_i \) and

   \[
   vC_i = \prod_{k=1}^{n} \left[ b_k + \frac{1}{i} - \left( a_k - \frac{1}{i} \right) \right] = \prod_{k=1}^{n} \left( b_k - a_k + \frac{2}{i} \right).
   \]

   Making \( i \to \infty \), we get

   \[
   \lim_{i \to \infty} vC_i = \prod_{k=1}^{n} (b_k - a_k) = vB.
   \]

(Why?) Hence by the sequential limit definition, given \( \varepsilon > 0 \), there is a natural \( i \) such that

   \[
   vC_i - vB < \varepsilon,
   \]

   or

   \[
   vC_i - \varepsilon < vB.
   \]

   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

   \[
   C_p \subseteq G,
   \]

   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
\[ G \subseteq \bigcup_{p \in G} C_p. \]
But by construction, $G$ contains all $C_p$, so that
\[ G = \bigcup_{p \in G} C_p. \]

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 \{$A_k$\}:
\[ G = \bigcup_{k=1}^{\infty} A_k. \]

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!). \(\square\)

Now let $C_\sigma$ be the family of all possible countable unions of intervals in $E^n$, such as $G$ in Lemma 2 (we use $C_\sigma$ for all finite unions). Thus $A \in C_\sigma$ means that $A$ is a $C_\sigma$-set, i.e.,
\[ A = \bigcup_{i=1}^{\infty} A_i \]
for some sequence of intervals \{$A_i$\}. Such are all open sets in $E^n$, but there also are many other $C_\sigma$-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 $C_\sigma$-set $A$ into countably many disjoint intervals. This can be done in many ways. However, we have the following result.

**Theorem 1.** If
\[ A = \bigcup_{i=1}^{\infty} A_i \text{ (disjoint)} = \bigcup_{k=1}^{\infty} B_k \text{ (disjoint)} \]
for some intervals $A_i, B_k$ in $E^n$, then
\[ \sum_{i=1}^{\infty} vA_i = \sum_{k=1}^{\infty} vB_k. \]
\[^1\text{Recall that a positive series always has a (possibly infinite) sum.}\]

Thus we can (and do) unambiguously define either of these sums to be the volume $vA$ of the $C_\sigma$-set $A$.

**Proof.** We shall use the Heine–Borel theorem (Problem 10 in Chapter 4, §6; review it!).

Seeking a contradiction, let (say)
\[ \sum_{i=1}^{\infty} vA_i > \sum_{k=1}^{\infty} vB_k, \]
so, in particular,
\[ \sum_{k=1}^{\infty} vB_k < +\infty. \]

As
\[ \sum_{i=1}^{\infty} vA_i = \lim_{m \to \infty} \sum_{i=1}^{m} vA_i, \]
there is an integer $m$ for which
\[ \sum_{i=1}^{m} vA_i > \sum_{k=1}^{\infty} vB_k. \]

We fix that $m$ and set
\[ 2\varepsilon = \sum_{i=1}^{m} vA_i - \sum_{k=1}^{\infty} vB_k > 0. \]

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
\[ vB_k > vY_k - \frac{\varepsilon}{2^k}, \quad k = 1, 2, \ldots, \]
and closed ones $X_i \subseteq A_i$, with
\[ vX_i + \frac{\varepsilon}{m} > vA_i; \]
so
\[ 2\varepsilon = \sum_{i=1}^{m} vA_i - \sum_{k=1}^{\infty} vB_k < \sum_{i=1}^{m} \left( vX_i + \frac{\varepsilon}{m} \right) - \sum_{k=1}^{\infty} \left( vY_k - \frac{\varepsilon}{2^k} \right) = \sum_{i=1}^{m} vX_i - \sum_{k=1}^{\infty} vY_k + 2\varepsilon. \]

Thus
\[ (1) \quad \sum_{i=1}^{m} vX_i > \sum_{k=1}^{\infty} vY_k. \]
(Explain in detail!)

Now, as

$$X_i \subseteq A_i \subseteq A = \bigcup_{k=1}^{\infty} B_k \subseteq \bigcup_{k=1}^{\infty} Y_k,$$

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,

$$\bigcup_{i=1}^{m} X_i \subseteq \bigcup_{k=1}^{p} Y_k.$$

The $X_i$ are disjoint, for even the larger sets $A_i$ are. Thus by Lemma 1(ii) in §1,

$$\sum_{i=1}^{m} vX_i \leq \sum_{k=1}^{p} vY_k \leq \sum_{k=1}^{\infty} vY_k,$$

contrary to (1). This contradiction completes the proof. \(\square\)

**Corollary 1.** If

$$A = \bigcup_{k=1}^{\infty} B_k \text{ (disjoint)}$$

for some intervals $B_k$, then

$$vA = \sum_{k=1}^{\infty} vB_k.$$

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 $\sigma$-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 2.** If $A_i, B_k$ are intervals in $E^n$, with

$$\bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{k=1}^{\infty} B_k,$$

then provided the $A_i$ are mutually disjoint,

$$\sum_{i=1}^{\infty} vA_i \leq \sum_{k=1}^{\infty} vB_k,$$

(2)

by Corollary 1.

2 This notion is treated in more detail in §5.

The proof is as in Theorem 1 (but the $B_k$ need not be disjoint here).

**Corollary 3** ("$\sigma$-subadditivity" of the volume). If

$$A \subseteq \bigcup_{k=1}^{\infty} B_k,$$

where $A \in \mathcal{C}_\sigma$ and the $B_k$ are intervals in $E^n$, then

$$vA \leq \sum_{k=1}^{\infty} vB_k.$$

**Proof.** Set

$$A = \bigcup_{i=1}^{\infty} A_i \text{ (disjoint), } A_i \in \mathcal{C},$$

and use Corollary 2. \(\square\)

**Corollary 4** ("monotonicity"). If $A, B \in \mathcal{C}_\sigma$, with

$$A \subseteq B,$$

then

$$vA \leq vB.$$

(“Larger sets have larger volumes.”)

This is simply Corollary 3, with $\bigcup_k B_k = B$.

**Corollary 5.** The volume of all of $E^n$ is $\infty$ (we write $\infty$ 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. \(\square\)

**Corollary 6.** For any countable set $A \subseteq E^n$, $vA = 0$. 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

$$A = \bigcup_k \{\bar{a}_k\};$$

so

$$vA = \sum_k v\{\bar{a}_k\} = 0$$

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 $\sigma$-additivity involves countable sums, it appears useful to generalize the notion of a series.

We say that a series of constants, \[ \sum a_n, \]

is **permutable** iff it has a definite (possibly infinite) sum obeying the general commutative law:

*Given any one–one map \( u: N \leftrightarrow N \) (\( N = \) the naturals), we have \[ \sum_n a_n = \sum_n a_{u_n}, \]

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 \( u: N \leftrightarrow J \) (where \( J \) is a countable index set) and a set \( \{a_i \mid i \in J\} \subseteq E \)

(where \( E \) is \( E^* \) or a normed space), we can define \[ \sum_{i \in J} a_i = \sum_{n=1}^{\infty} a_{u_n}, \]

if \( \sum_n a_{u_n} \) is permutable.

In particular, if \( J = N \times N \)

(a countable set, by Theorem 1 in Chapter 1, §9), we call \[ \sum_{i \in J} a_i \]
a **double series**, denoted by symbols like \[ \sum_{n,k} a_{kn} \quad (k,n \in N). \]

Note that \[ \sum_{i \in J} |a_i| \]
is always defined (being a positive series).

If \[ \sum_{i \in J} |a_i| < \infty, \]
we say that \( \sum_{i \in J} a_i \) converges absolutely.

For a positive series, we obtain the following result.

**Theorem 2.**

(i) All positive series in \( E^* \) are permutable.

(ii) For positive double series in \( E^* \), we have

\[ \sum_{n,k=1}^{\infty} a_{nk} = \sum_{k=1}^{\infty} \left( \sum_{n=1}^{\infty} a_{nk} \right) = \sum_{n=1}^{\infty} \left( \sum_{k=1}^{\infty} a_{nk} \right). \]

**Proof.** (i) Let

\[ s = \sum_{n=1}^{\infty} a_n \quad \text{and} \quad s_m = \sum_{n=1}^{m} a_n \quad (a_n \geq 0). \]

Then clearly

\[ s = \lim_{m \to \infty} s_m = \sup s_m \]

by Theorem 3 in Chapter 3, §15.

Hence \( s \) certainly does not exceed the lub of all possible sums of the form \[ \sum_{i \in I} a_i, \]

where \( I \) is a finite subset of \( N \) (the partial sums \( s_m \) are among them). Thus

\[ s \leq \sup_{i \in I} a_i, \]

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

\[ s = \sup_{i \in I} a_i. \]
But $\sup \sum_{i \in J} a_i$ 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 $\sum_{i \in J} a_i$, where $J$ 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).

**Problems on $\mathcal{C}_\sigma$-Sets, $\sigma$-Additivity, and Permutable Series**

1. Fill in the missing details in the proofs of this section.

1'. Prove Note 3.

2. Show that every open set $A \neq \emptyset$ in $\mathbb{E}^n$ is a countable union of disjoint half-open cubes.

[Outline: For each natural $m$, show that $\mathbb{E}^n$ is split into such cubes of edge length $2^{-m}$ by the hyperplanes

$$x_k = \frac{i}{2^m} \quad i = 0, \pm 1, \pm 2, \ldots; \quad k = 1, 2, \ldots, n,$$

and that the family $\mathcal{C}_m$ of such cubes is countable.

For $m > 1$, let $\mathcal{C}_{m1}, \mathcal{C}_{m2}, \ldots$ be the sequence of those cubes from $\mathcal{C}_m$ (if any) that lie in $A$ but not in any cube $\mathcal{C}_{mj}$ with $s < m$.

As $A$ is open, $x \in A$ if $x \in$ some $\mathcal{C}_{mj}.$]

3. Prove that any open set $A \subseteq \mathbb{E}^3$ is a countable union of disjoint (possibly infinite) open intervals.

[Hint: By Lemma 2, $A = \bigcup_i (a_i, b_i)$. If, say, $(a_1, b_1)$ overlaps with some $(a_m, b_m)$, replace both by their union. Continue inductively.]

4. Prove that $\mathcal{C}_\sigma$ is closed under finite intersections and countable unions.

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 $A = \mathbb{E}^3, B = \mathbb{R}$ (the rationals).]

**Note.** In the following problems, $J$ is countably infinite, $a_i \in E$ ($E$ complete).

6. Prove that

$$\sum_{i \in J} |a_i| < \infty,$$

iff for every $\varepsilon > 0$, there is a finite set $F \subseteq J$ ($F \neq \emptyset$) such that

$$\sum_{i \in J} |a_i| < \varepsilon$$

for every finite $I \subseteq J - F$.

7. Prove that if

$$\sum_{i \in J} |a_i| < \infty,$$

then for every $\varepsilon > 0$, there is a finite $F \subseteq J$ ($F \neq \emptyset$) such that

$$\sum_{i \in K} |a_i| < \varepsilon$$

for each finite $K \supseteq F$ ($K \subseteq J$).

[Hint: Proceed as in Problem 6, with $I = K - F$ and $q$ so large that $\left|\sum_{i \in J} a_i - \sum_{i \in F} a_i\right| < \frac{1}{2} \varepsilon$, and $\left|\sum_{i \in F} a_i\right| < \frac{1}{2} \varepsilon$.]

8. Show that if $J = \bigcup_{n=1}^\infty I_n$ (disjoint), then

$$\sum_{i \in J} |a_i| = \sum_{n=1}^\infty b_n,$$

where $b_n = \sum_{i \in I_n} |a_i|$. (Use Problem 8' below.)

8'. Show that

$$\sum_{i \in J} |a_i| = \sup_{F} \sum_{i \in F} |a_i|$$

over all finite sets $F \subseteq J$ ($F \neq \emptyset$).
9. Show that if $\emptyset \neq J \subseteq I$, then
\[ \sum_{i \in J} |a_i| = \sum_{i \in J} |a_i| \leq \sum_{i \in I} |a_i| . \]

[Hint: as in Theorem 2.]

10. Continuing Problem 8, prove that if
\[ \sum_{i \in J} |a_i| = \sum_{n=1}^k b_n < \infty , \]
then
\[ \sum_{i \in J} a_i = \sum_{n=1}^k c_n \text{ with } c_n = \sum_{i \in I_n} a_i , \]

[Outline: By Problem 9, $(\forall n) \sum_{i \in I_n} |a_i| < \infty$; so
\[ c_n = \sum_{i \in I_n} a_i , \]
and
\[ \sum_{n=1}^\infty c_n \]
converge absolutely.

Fix $\varepsilon$ and $F$ as in Problem 7. Choose the largest $q \in N$ with
\[ F \cap I_q \neq \emptyset \]
(why does it exist?), and fix any $n > q$. By Problem 7, $(\forall k \leq n)$
\[ (\forall k \leq n) \left( \exists \text{ finite } F_k \mid J \supseteq F_k \supseteq F \cap I_k \right) \]
\[ (\forall k \leq n) \left( \exists \text{ finite } H_k \mid I_n \supseteq H_k \supseteq F_k \right) \]
\[ \left( \sum_{i \in H_k} a_i - \sum_{k=1}^n c_k \right) < \frac{1}{2} \varepsilon . \]
(Explain!) Let
\[ K = \bigcup_{k=1}^n H_k , \]
so
\[ \left| \sum_{k=1}^n c_k - \sum_{i \in J} a_i \right| < \varepsilon \]
and $K \supseteq F$. By Problem 7,
\[ \left| \sum_{i \in K} a_i - \sum_{i \in J} a_i \right| < \varepsilon . \]

\section*{§3. More on Set Families$^1$}

Lebesgue extended his theory far beyond $\mathcal{C}_r$-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.

\begin{definition}
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:
\[ (\forall X, Y \in \mathcal{M}) \quad X \cup Y \in \mathcal{M} \text{ and } X - Y \in \mathcal{M} . \]
\end{definition}

\footnote{For a limited approach (see the preface), this topic may be omitted.}
(For intersections, see Theorem 1 below.)

If $\mathcal{M}$ is also closed under countable unions, we call it a $\sigma$-ring (in $S$). Then

$$\bigcup_{i=1}^{\infty} X_i \in \mathcal{M}$$

whenever

$$X_i \in \mathcal{M} \text{ for } i = 1, 2, \ldots.$$

If $S$ itself is a member of a ring ($\sigma$-ring) $\mathcal{M}$, we call $\mathcal{M}$ a set field ($\sigma$-field), or a set algebra ($\sigma$-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 $\sigma$-field. (Why?)

(b) The family $\{\emptyset\}$, consisting of $\emptyset$ alone, is a $\sigma$-ring; $\{\emptyset, S\}$ is a $\sigma$-field in $S$. (Why?)

(c) The family of all finite (countable) subsets of $S$ is a ring ($\sigma$-ring) in $S$.

(d) For any semiring $\mathcal{C}$, $\mathcal{C}'$ is a ring (Theorem 2 in §1). Not so for $\mathcal{C}_\sigma$ (Problem 5 in §2).

**Theorem 1.** Any set ring is closed under finite intersections. A $\sigma$-ring is closed under countable intersections.

**Proof.** Let $\mathcal{M}$ be a $\sigma$-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

$$U = \bigcup_n A_n.$$ 

By Definition 1,

$$U \in \mathcal{M} \text{ and } U - A_n \in \mathcal{M},$$

as $\mathcal{M}$ is closed under these operations. Hence

$$\bigcup_n (U - A_n) \in \mathcal{M}$$

and

$$U - \bigcup_n (U - A_n) \in \mathcal{M},$$

or, by duality,

$$\bigcap_n [U - (U - A_n)] \in \mathcal{M},$$

i.e.,

$$\bigcap_n A_n \in \mathcal{M}. \ □$$

**Corollary 1.** Any set ring (field, $\sigma$-ring, $\sigma$-field) is also a semiring.

Indeed, by Theorem 1 and Definition 1, if $\mathcal{M}$ is a ring, then $\emptyset \in \mathcal{M}$ and $(\forall A, B \in \mathcal{M}) \ A \cap B \in \mathcal{M}$ and $A - B \in \mathcal{M}$.

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 $\sigma$-rings, fields, etc.

In §1 we saw that any semiring $\mathcal{C}$ can be enlarged to become a ring, $\mathcal{C}'$. More generally, we obtain the following result.

**Theorem 2.** For any set family $\mathcal{M}$ in a space $S$ ($\mathcal{M} \subseteq 2^S$), there is a unique "smallest" set ring $\mathcal{R}$ such that

$$\mathcal{R} \supseteq \mathcal{M}$$

("smallest" in the sense that

$$\mathcal{R} \subseteq \mathcal{R}'$$

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 $\sigma$-rings, fields, and $\sigma$-fields in $S$.

**Proof.** We give the proof for $\sigma$-fields; it is similar in the other cases.

There surely are $\sigma$-fields in $S$ that contain $\mathcal{M}$; e.g., take $2^S$. Let $\{\mathcal{R}_i\}$ be the family of all possible $\sigma$-fields in $S$ such that $\mathcal{R}_i \supseteq \mathcal{M}$. Let

$$\mathcal{R} = \bigcap_i \mathcal{R}_i.$$ 

We shall show that this $\mathcal{R}$ is the required "smallest" $\sigma$-field containing $\mathcal{M}$.

Indeed, by assumption,

$$\mathcal{M} \subseteq \bigcap_i \mathcal{R}_i = \mathcal{R}.$$ 

We now verify the $\sigma$-field properties for $\mathcal{R}$.

(1) We have that

$$(\forall i) \ \emptyset \in \mathcal{R}_i \text{ and } S \in \mathcal{R}_i.$$
(for $\mathcal{R}_i$ is a $\sigma$-field, by assumption). Hence
\[ \emptyset \in \bigcap_i \mathcal{R}_i = \mathcal{R}. \]

Similarly, $S \in \mathcal{R}$. Thus
\[ \emptyset, S \in \mathcal{R}. \]

(2) Suppose
\[ X, Y \in \mathcal{R} = \bigcap_i \mathcal{R}_i. \]

Then $X, Y$ are in every $\mathcal{R}_i$, and so is $X - Y$. Hence $X - Y$ is in
\[ \bigcap_i \mathcal{R}_i = \mathcal{R}. \]

Thus $\mathcal{R}$ is closed under differences.

(3) Take any sequence
\[ \{A_n\} \subseteq \mathcal{R} = \bigcap_i \mathcal{R}_i. \]

Then all $A_n$ are in each $\mathcal{R}_i$. $\bigcup_n A_n$ is in each $\mathcal{R}_i$; so
\[ \bigcap_n A_n \in \mathcal{R}. \]

Thus $\mathcal{R}$ is closed under countable unions.

We see that $\mathcal{R}$ is indeed a $\sigma$-field in $S$, with $\mathcal{M} \subseteq \mathcal{R}$. As $\mathcal{R}$ is the intersection of all $\mathcal{R}_i$ (i.e., all $\sigma$-fields $\supseteq \mathcal{M}$), we have
\[ (\forall i) \quad \mathcal{R} \subseteq \mathcal{R}_i; \]
so $\mathcal{R}$ is the smallest of such $\sigma$-fields.

It is unique; for if $\mathcal{R}'$ is another such $\sigma$-field, then
\[ \mathcal{R} \subseteq \mathcal{R}' \subseteq \mathcal{R} \]
(as both $\mathcal{R}$ and $\mathcal{R}'$ are “smallest”); so
\[ \mathcal{R} = \mathcal{R}'. \]

Note 1. This proof also shows that the intersection of any family $\{\mathcal{R}_i\}$ of $\sigma$-fields is a $\sigma$-field. Similarly for $\sigma$-rings, fields, and rings.

Corollary 2. The ring $\mathcal{R}$ generated by a semiring $\mathcal{C}$ coincides with
\[ \mathcal{C}_s = \{ \text{all finite unions of } \mathcal{C}-\text{sets} \} \]
and with
\[ \mathcal{C}'_s = \{ \text{disjoint finite unions of } \mathcal{C}-\text{sets} \}. \]

§3. More on Set Families

Proof. By Theorem 2 in §1, $\mathcal{C}'_s$ is a ring $\supseteq \mathcal{C}$; and
\[ \mathcal{C}'_s \subseteq \mathcal{C}_s \subseteq \mathcal{R} \]
(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
\[ \mathcal{R} \subseteq \mathcal{C}'_s \subseteq \mathcal{C}_s \subseteq \mathcal{R}. \]

Hence
\[ \mathcal{R} = \mathcal{C}'_s = \mathcal{C}_s, \]
as claimed. \(\Box\)

It is much harder to characterize the $\sigma$-ring generated by a semiring. The following characterization proves useful in theory and as an exercise.\(^2\)

Theorem 3. The $\sigma$-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
\[ (A - B)C = A - (AC' \cup BC). \]

(3) For each $I \in \mathcal{D}$, set
\[ \mathcal{D}_I = \{ A \in \mathcal{D} \mid AI \in \mathcal{D}, A - I \in \mathcal{D} \}. \]

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,
\[ (\forall A \in \mathcal{D}) \quad (\forall I \in \mathcal{C}) \quad AI \in \mathcal{D} \text{ and } A - I \in \mathcal{D}. \]

(4) Using this, show that $\mathcal{D}_I$ satisfies (i)–(iii) for any $I \in \mathcal{D}$.

Deduce
\[ \mathcal{D} \subseteq \mathcal{D}_I; \]
so $\mathcal{D}$ is closed under finite intersections and differences.

Combining with property (ii), show that $\mathcal{D}$ is a $\sigma$-ring (see Problem 12 below).

\(^2\) It may be deferred until Chapter 8, §8, though.
By its minimality, \( D \) is the smallest \( \sigma \)-ring \( \supseteq C \) (for any other such \( \sigma \)-ring

Thus \( D = \mathcal{R} \), as claimed. \( \square \)

**Definition 2.**

Given a set family \( \mathcal{M} \), we define (following Hausdorff)

(a) \( \mathcal{M}_s = \{ \text{all countable unions of } \mathcal{M}\text{-sets} \} \) (cf. \( C_s \) in §2);

(b) \( \mathcal{M}_d = \{ \text{all countable intersections of } \mathcal{M}\text{-sets} \} \).

We use \( \mathcal{M}_s \) and \( \mathcal{M}_d \) for similar notions, with “countable” replaced by “finite.”

Clearly,

\[
\mathcal{M}_s \supseteq \mathcal{M}_d \supseteq \mathcal{M}
\]

and

\[
\mathcal{M}_d \supseteq \mathcal{M}_s \supseteq \mathcal{M}
\]

Why?

**Note 2.** Observe that \( \mathcal{M} \) is closed under finite (countable) unions iff

\[
\mathcal{M} = \mathcal{M}_s \ (\mathcal{M} = \mathcal{M}_d).
\]

Verify! Interpret \( \mathcal{M} = \mathcal{M}_d \ (\mathcal{M} = \mathcal{M}_s) \) similarly.

In conclusion, we generalize **Theorem 1** in §1.

**Definition 3.**

The product \( \mathcal{M} \times \mathcal{N} \) of two set families \( \mathcal{M} \) and \( \mathcal{N} \) is the family of all sets of the form

\[
A \times B,
\]

with \( A \in \mathcal{M} \) and \( B \in \mathcal{N} \).

(The dot in \( \times \) is to stress that \( \mathcal{M} \times \mathcal{N} \) is not really a Cartesian product.)

**Theorem 4.** If \( \mathcal{M} \) and \( \mathcal{N} \) are semirings, so is \( \mathcal{M} \times \mathcal{N} \).

The proof runs along the same lines as that of **Theorem 1** in §1, via the set identities

\[
(X \times Y) \cap (X' \times Y') = (X \cap X') \times (Y \cap Y')
\]

and

\[
(X \times Y) - (X' \times Y') = [(X - X') \times Y] \cup [(X \cap X') \times (Y - Y')].
\]

The details are left to the reader.

**Note 3.** As every ring is a semiring (Corollary 1), the product of two rings (fields, \( \sigma \)-rings, \( \sigma \)-fields) is a semiring. However, see **Problem 6** below.

§3. More on Set Families

**Problems on Set Families**

1. Verify Examples (a), (b), and (c).

1’. Prove **Theorem 1** for rings.

2. Show that in **Definition 1** \( \{ \emptyset \in \mathcal{M} \} \) may be replaced by \( \{ \mathcal{M} \neq \emptyset \} \).

3. Prove that \( \mathcal{M} \) is a field (\( \sigma \)-field) iff \( \mathcal{M} \neq \emptyset \), \( \mathcal{M} \) is closed under finite (countable) unions, and

\[
(\forall \, A \in \mathcal{M} \) \quad A \in \mathcal{M}.
\]

4. Prove **Theorem 2** for set fields.

4’. Does **Note 1** apply to semirings?

5. Prove **Note 2**.

5’. Prove **Theorem 3** in detail.

6. Prove **Theorem 4** and show that the product \( \mathcal{M} \times \mathcal{N} \) of two rings need not be a ring.

7. Let \( \mathcal{R}, \mathcal{R}' \) be the rings (\( \sigma \)-rings, fields, \( \sigma \)-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 \( \mathcal{M} = \{ \text{open intervals in } E^n \} \) and \( \mathcal{N} = \{ \text{all open sets in } E^n \} \),

then \( \mathcal{R} = \mathcal{R}' \).

8. Is any of the following a semiring, ring, \( \sigma \)-ring, field, or \( \sigma \)-field? Why?

(a) All infinite intervals in \( E^1 \).

(b) All open sets in a metric space \( (S, \rho) \).

(c) All closed sets in \( (S, \rho) \).

(d) All “clopen” sets in \( (S, \rho) \).

(e) \( \{ \, X \in 2^S \mid -X \text{ finite} \} \).

(f) \( \{ \, X \in 2^S \mid -X \text{ countable} \} \).
9. Prove that for any sequence \( \{A_n\} \) in a ring \( R \), there is
(a) an expanding sequence \( \{B_n\} \subseteq R \) such that
\[
(\forall n) \quad B_n \supseteq A_n
\]
and
\[
\bigcup_n B_n = \bigcup_n A_n; \quad \text{and}
\]
(b) a contracting sequence \( C_n \subseteq A_n \), with
\[
\bigcap_n C_n = \bigcap_n A_n.
\]
(The latter holds in semirings, too.)
[Hint: Set \( B_n = \bigcap_n A_k, C_n = \bigcap_n A_k \).

10. The symmetric difference, \( A \triangle B \), of two sets is defined as
\[
A \triangle B = (A - B) \cup (B - A).
\]
Inductively, we also set
\[
\Delta_{k=1}^1 A_k = A_1
\]
and
\[
\Delta_{k=1}^{n+1} A_k = \left( \bigcup_{k=1}^n \Delta A_k \right) \Delta A_{n+1}.
\]
Show that symmetric differences
(i) are commutative,
(ii) are associative, and
(iii) satisfy the distributive law:
\[
(A \triangle B) \cap C = (A \cap C) \triangle (B \cap C).
\]
[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 \( C \).]

11. Prove that \( M \) is a ring if
(i) \( 0 \in M; \)
(ii) \( (\forall A, B \in M) \ A \triangle B \in M \) and \( A \cap B \in M \) (see Problem 10); equivalently,
(ii') \( A \triangle B \in M \) and \( A \cup B \in M \).
[Hint: Verify that
\[
A \cup B = (A \triangle B) \triangle (A \cap B)
\]
and
\[
A - B = (A \cup B) \triangle B,
\]
while
\[
A \cap B = (A \cup B) \triangle (A \triangle B).
\]

12. Show that a set family \( M \neq \emptyset \) is a \( \sigma \)-ring iff one of the following conditions holds.
(a) \( M \) is closed under countable unions and proper differences \( X - Y \) with \( X \supseteq Y \);
(b) \( M \) is closed under countable disjoint unions, proper differences, and finite intersections; or
(c) \( M \) 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
\[
X \cup Y = (X - Y) \cup (Y - X) \cup (X \cap Y)
\]
shows that \( M \) is closed under all finite unions; so \( M \) is a ring. Now use Corollary 1 in §1 for countable unions.
(c) Use Problem 11.]

13. From Problem 10, treating \( \triangle \) as addition and \( \cap \) as multiplication, show that any set ring \( M \) is an algebraic ring with unity, i.e., satisfies the six field axioms (Chapter 2, §§1–4), except V(b) (existence of multiplicative inverses).

14. A set family \( H \) is said to be hereditary iff
\[
(\forall X \in H) \ (\forall Y \subseteq X) \ Y \in H.
\]
Prove the following.
(a) For every family \( M \subseteq 2^S \), there is a “smallest” hereditary ring \( H \supseteq M \) (\( H \) is said to be generated by \( M \)). Similarly for \( \sigma \)-rings, fields, and \( \sigma \)-fields.
(b) The hereditary \( \sigma \)-ring generated by \( M \) consists of those sets which can be covered by countably many \( M \)-sets.

15. Prove that the field \( \sigma \)-field in \( S \), generated by a ring \( \sigma \)-ring \( R \), consists exactly of all \( R \)-sets and their complements in \( S \).

16. Show that the ring \( R \) generated by a set family \( C \neq \emptyset \) consists of all sets of the form
\[
\bigtriangleup_{k=1}^n A_k
\]
(see Problem 10), where each \( A_k \in C \) (finite intersection of \( C \)-sets).
[Outline: By Problem 11, \( R \) must contain the family (call it \( M \)) of all such \( \bigtriangleup_{k=1}^n A_k \).
(Why?) It remains to show that \( M \) is a ring \( \supseteq C \).]
Write \( A + B \) for \( A \Delta B \) and \( AB \) for \( A \cap B \); so each \( M \)-set is a "sum" of finitely many "products"

\[
A_1 A_2 \cdots A_n,
\]

By algebra, the "sum" and "product" of two such "polynomials" is such a polynomial itself. Thus

\[
(\forall X, Y \in M) \quad X \Delta Y \text{ and } X \cap Y \in M.
\]

Now use Problem 11.

17. Use Problem 16 to obtain a new proof of Theorem 2 in §1 and Corollary 2 in the present section.

[Hints: For semirings, \( C = C_d \). (Why?) Thus in Problem 16, \( A_k \in C \).

Also,

\[
(\forall A, B \in C) \quad A \triangle B = (A - B) \cup (B - A)
\]

where \( A - B \) and \( B - A \) are finite disjoint unions of \( C \)-sets. (Why?)

Deduce that \( A \Delta B \in C \) and, by induction,

\[
\bigtriangleup_{k=1}^n A_k \in C;
\]

so \( R \subseteq C \subseteq R \). (Why?)

18. Given a set \( A \) and a set family \( M \), let

\[
A \cap M
\]

be the family of all sets \( A \cap X \), with \( X \in M \); similarly,

\[
\mathcal{N} = \{ \text{all sets } Y \cup (X - A), \text{ with } Y \in \mathcal{N}, X \in M \}, \text{ etc.}
\]

Show that if \( M \) generates the ring \( R \) then \( A \in M \) generates the ring

\[
\mathcal{R}' = A \cap R.
\]

Similarly for \( \sigma \)-rings, fields, \( \sigma \)-fields.

[Hint for rings: Prove the following.

(i) \( A \cap R \) is a ring.

(ii) \( M \subseteq R' \cup (R - A) \), with \( R' \) as above.

(iii) \( R \cup (R - A) \) is a ring (call it \( \mathcal{N} \)).

(iv) By (ii) \( R \subseteq \mathcal{N} \), so \( A \cap R \subseteq A \cap \mathcal{N} \subseteq R' \).

(v) \( A \in \mathcal{N} \subseteq R' \) (for \( A \cap \mathcal{N} \subseteq A \cap M \).

Hence \( R' = A \cap R \).]

§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 \( C \)-simple sets, and even for \( C_\sigma \)-sets in \( E^n \).

Mathematically this means that the volume function \( v \) has been defined first on \( C \) (the intervals), then on \( C_\sigma \) (\( C \)-simple sets), and finally on \( C_\sigma \).

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 \( (C, C_\sigma, \text{ or } C_\sigma) \).

The "volumes" assigned to such sets are the function values (for \( C \) and \( C_\sigma \)-sets they are real numbers; for \( C_\sigma \)-sets they may reach \(+\infty\)). This is symbolized by

\[
v : C \rightarrow E^1
\]
or

\[
v : C_\sigma \rightarrow E^*;
\]

more precisely,

\[
v : C_\sigma \rightarrow [0, \infty),
\]

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.,

\[
v = \left( \begin{array}{cccc}
A & B & C & \ldots \\
vA & vB & vC & \ldots 
\end{array} \right).
\]

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

\[
s : M \rightarrow E
\]

whose domain is a set family \( M \).

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 \( X \in M \), 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 \( N \subseteq M \) iff

\[
(\forall X \in N) \quad |sX| < \infty;
\]

briefly, \(|s| < \infty \) on \( N \). (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.,
\[
s : \mathcal{M} \to [0, \infty].
\]
(The symbol \( \infty \) stands for \( +\infty \) throughout.)

**Definition 2.**

A set function
\[
s : \mathcal{M} \to E
\]
is called **additive** (or **finitely additive**) on \( \mathcal{N} \subseteq \mathcal{M} \) iff for any finite disjoint union \( \bigcup_k A_k \), we have
\[
\sum_k sA_k = s\left(\bigcup_k A_k\right).
\]
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 **\( \sigma \)-additive** (or **countably additive** or **completely additive**) on \( \mathcal{N} \).

If \( \mathcal{N} = \mathcal{M} \) here, we simply say that \( s \) is **additive** (\( \sigma \)-additive, respectively).

**Note 1.** As \( \bigcup_k A_k \) is independent of the order of the \( A_k \), \( \sigma \)-additivity presupposes and implies that the series
\[
\sum sA_k
\]
is **permutable** (§2) for any disjoint sequence
\[
\{A_k\} \subseteq \mathcal{N}.
\]
(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 \( \sigma \)-additive (Corollary 1 in §2).

Examples (b)–(d) show that set functions may arise from ordinary “point functions.”

**Examples.**

(a) The volume function \( v : \mathcal{C} \to E^3 \) 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 \( f : E^3 \to E \), set
\[
(\forall I \in \mathcal{M}) \quad sI = V_f[\mathcal{I}],
\]
the total variation of \( f \) on the **closure** of \( I \) (Chapter 5, §7).

Then \( s : \mathcal{M} \to [0, \infty) \) 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 \( E^3 \). For each interval \( X \) with endpoints \( a, b \in E^1 \) \((a \leq b)\), set
\[
sX = \int_a^b f.
\]

This yields a set function \( s : \mathcal{M} \to E \) (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
\[
\alpha : E^1 \to E^1
\]
has finite one-sided limits
\[
\alpha(p+) \text{ and } \alpha(p-)
\]
at each \( p \in E^1 \). The **Lebesgue–Stieltjes (LS) function**
\[
s_\alpha : \mathcal{C} \to E^1
\]
(is important for **Lebesgue–Stieltjes integration**) is defined as follows.

Set \( s_\alpha \theta = 0 \). For **nonvoid** intervals, including \( [a, a] = \{ a \} \), set
\[
\begin{align*}
\sigma_{\alpha}[a, b] &= \alpha(b+) - \alpha(a-), \\
\sigma_{\alpha}(a, b] &= \alpha(b) - \alpha(a+), \\
\sigma_{\alpha}[a, b) &= \alpha(b-) - \alpha(a-), \quad \text{and} \\
\sigma_{\alpha}(a, b) &= \alpha(b-) - \alpha(a+).
\end{align*}
\]

For the properties of \( s_\alpha \) 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
\[
2^S = \{ \text{all subsets } X \subset S \} \quad (\text{§3}).
\]

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 \( s : \mathcal{M} \to E \) be additive on \( \mathcal{N} \subseteq \mathcal{M} \). Let
\[
A, B \in \mathcal{N}, \ A \subseteq B.
\]

Then we have the following.

(1) If \( |sA| < \infty \) and \( B - A \in \mathcal{N} \), then
\[
s(B - A) = sB - sA \quad (\text{"subtractivity"}).
\]

\( \leftarrow \)
(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

$$|sB| < \infty \Rightarrow |sA| < \infty.$$  

If further $s$ is semifinite then

$$sA = \pm \infty \Rightarrow sB = \pm \infty$$  

(same sign).

Proof.

(1) As $B \supseteq A$, we have

$$B = (B - A) \cup A \ (\text{disjoint});$$  

so by additivity,

$$sB = s(B - A) + sA.$$  

If $|sA| < \infty$, we may transpose to get

$$sB - sA = s(B - A),$$  

as claimed.

(2) Hence

$$s\emptyset = s(X - X) = sX - sX = 0$$  

if $X, \emptyset \in \mathcal{N}$, and $|sX| < \infty$.

(3) If $\mathcal{N}$ is a semiring, then

$$B - A = \bigcup_{k=1}^{n} A_k \ (\text{disjoint})$$  

for some $\mathcal{N}$-sets $A_k$; so

$$B = A \cup \bigcup_{k=1}^{n} A_k \ (\text{disjoint}).$$  

By additivity,

$$sB = sA + \sum_{k=1}^{n} sA_k;$$  

so by our conventions,

$$|sA| = \infty \Rightarrow |sB| = \infty.$$  

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. □

§4. Set Functions. Additivity. Continuity

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 1. If

$$s : \mathcal{C} \to E$$

is additive on $\mathcal{C}$, an arbitrary semiring, there is a unique set function

$$\bar{s} : \mathcal{C}_s \to E,$$

additive on $\mathcal{C}_s$, with $\bar{s} = s$ on $\mathcal{C}$, i.e.,

$$\bar{s}X = sX \text{ for } X \in \mathcal{C}.$$  

We call $\bar{s}$ the additive extension of $s$ to $\mathcal{C}_s = \mathcal{C}'_s$ (Corollary 2 in §3).

Proof. If $s \geq 0$ ($s : \mathcal{C} \to [0, \infty]$), proceed as in Lemma 1 and Corollary 2, all of §1.

The general proof (which may be omitted or deferred) is as follows.

Each $X \in \mathcal{C}_s$ has the form

$$X = \bigcup_{i=1}^{m} X_i \ (\text{disjoint}), \ X_i \in \mathcal{C}.$$  

Thus if $\bar{s}$ is to be additive, the only way to define it is to set

$$\bar{s}X = \sum_{i=1}^{m} sX_i.$$  

This already makes $\bar{s}$ unique, provided we show that

$$\sum_{i=1}^{m} sX_i$$  

does not depend on the particular decomposition

$$X = \bigcup_{i=1}^{m} X_i$$  

(otherwise, all is ambiguous).

Then take any other decomposition

$$X = \bigcup_{k=1}^{n} Y_k \ (\text{disjoint}), \ Y_k \in \mathcal{C}.$$  

Additivity implies

$$(\forall i, k) \ sX_i = \sum_{k=1}^{n} s(X_i \cap Y_k) \text{ and } sY_k = \sum_{i=1}^{m} s(X_i \cap Y_k).$$
Chapter 7. Volume and Measure

(Verify!) Hence
\[ \sum_{i=1}^{m} sX_i = \sum_{i,k} s(X_i \cap Y_k) = \sum_{k=1}^{n} sY_k. \]

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)
\[ X = \bigcup_{i=1}^{1} X_i, \quad X_1 = X; \]
so
\[ \bar{s}X = sX_1 = sX; \]
i.e., \( \bar{s} = s \) on \( \mathcal{C} \), as required.

Finally, for the additivity of \( \bar{s} \), let
\[ A = \bigcup_{k=1}^{m} B_k \text{ (disjoint)}, \quad A, B_k \in \mathcal{C}_s. \]

Here we may set
\[ B_k = \bigcup_{i=1}^{\alpha_k} C_{ki} \text{ (disjoint)}, \quad C_{ki} \in \mathcal{C}. \]

Then
\[ A = \bigcup_{k,i} C_{ki} \text{ (disjoint)}; \]
so by our definition of \( \bar{s} \),
\[ \bar{s}A = \sum_{k,i} sC_{ki} = \sum_{k=1}^{m} \left( \sum_{i=1}^{\alpha_k} sC_{ki} \right) = \sum_{k=1}^{m} sB_k, \]
as required. \( \square \)

**Continuity.** We write \( X_n \nearrow X \) to mean that
\[ X = \bigcup_{n=1}^{\infty} X_n \]
and \( \{X_n\} \uparrow \), i.e.,
\[ X_n \subseteq X_{n+1}, \quad n = 1, 2, \ldots. \]
Similarly, \( X_n \searrow X \) iff
\[ X = \bigcap_{n=1}^{\infty} X_n \]

|4. Set Functions. Additivity. Continuity |

and \( \{X_n\} \downarrow \), i.e.,
\[ X_n \supseteq X_{n+1}, \quad n = 1, 2, \ldots. \]

In both cases, we set
\[ X = \lim_{n \to \infty} X_n. \]

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
\[ sX = \lim_{n \to \infty} sX_n \]
whenever \( X_n \nearrow X \) and \( X, X_n \in \mathcal{M} \);
(ii) right continuous iff
\[ sX = \lim_{n \to \infty} sX_n \]
whenever \( X_n \searrow X \), with \( X, X_n \in \mathcal{M} \) and \( |sX| < \infty \).

Thus in case (i),
\[ \lim_{n \to \infty} sX_n = s \bigcup_{n=1}^{\infty} X_n \]
if all \( X_n \) and \( \bigcup_{n=1}^{\infty} X_n \) are \( \mathcal{M} \)-sets.

In case (ii),
\[ \lim_{n \to \infty} sX_n = s \bigcap_{n=1}^{\infty} X_n \]
if all \( X_n \) and \( \bigcap_{n=1}^{\infty} X_n \) are in \( \mathcal{M} \), and \( |sX| < \infty \).

**Note 2.** The last restriction applies to right continuity only. (We choose simply to exclude from consideration sequences \( \{X_n\} \downarrow \), with \( |sX| = \infty \); see Problem 4.)

**Theorem 2.** If \( s: \mathcal{C} \to E \) is \( \sigma \)-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
\[ X = \bigcup_{n=1}^{\infty} X_n. \]

If \( sX_n = \pm \infty \) for some \( n \), then (Lemma 3)
\[ sX = sX_m = \pm \infty \] for \( m \geq n, \)
Chapter 7. Volume and Measure

since $X \supseteq X_m \supseteq X_n$; so
\[ \lim sX_m = \pm\infty = sX, \]
as claimed.

Thus assume all $sX_n$ finite; so $s\emptyset = 0$, by Lemma 2.

Set $X_0 = \emptyset$. As is easily seen,
\[ X = \bigcup_{n=1}^{\infty} X_n = \bigcup_{n=1}^{\infty} (X_n - X_{n-1}) \text{ (disjoint)}, \]
and
\[ (\forall n) \ X_n - X_{n-1} \in \mathcal{C} \text{ (a ring)}. \]
Also,
\[ (\forall m \geq n) \ X_m = \bigcup_{n=1}^{m} (X_n - X_{n-1}) \text{ (disjoint)}. \]

(Verify!) Thus by additivity,
\[ sX_m = \sum_{n=1}^{m} s(X_n - X_{n-1}), \]
and by the assumed $\sigma$-additivity,
\[ sX = s \bigcup_{n=1}^{\infty} (X_n - X_{n-1}) = \sum_{n=1}^{\infty} s(X_n - X_{n-1}) \]
\[ = \lim_{m \to \infty} \sum_{n=1}^{m} s(X_n - X_{n-1}) = \lim_{m \to \infty} sX_m, \]
as claimed.

Right continuity. Let $X_n \searrow X$ with $X, X_n \in \mathcal{C}$,
\[ X = \bigcap_{n=1}^{\infty} X_n, \]
and
\[ |sX_1| < \infty. \]
As $X \subseteq X_n \subseteq X_1$, Lemma 3 yields that
\[ (\forall n) \ |sX_n| < \infty \]
and $|sX| < \infty$.

As
\[ X = \bigcap_{k=1}^{\infty} X_k, \]
we have
\[ (\forall n) \ X_n = X \cup \bigcup_{k=n+1}^{\infty} (X_{k-1} - X_k) \text{ (disjoint)}. \]

(Verify!) Thus by $\sigma$-additivity,
\[ (\forall n) \ sX_n = sX + \sum_{k=n+1}^{\infty} s(X_{k-1} - X_k), \]
with $|sX| < \infty$, $|sX_n| < \infty$ (see above).

Hence the sum
\[ \sum_{k=n+1}^{\infty} s(X_{k-1} - X_k) = sX_n - sX \]
is finite. Therefore, it tends to 0 as $n \to \infty$ (being the “remainder term” of a convergent series). Thus $n \to \infty$ yields
\[ \lim_{n \to \infty} sX_n = sX + \lim_{n \to \infty} \sum_{k=n+1}^{\infty} s(X_{k-1} - X_k) = sX, \]
as claimed. \(\square\)

### Problems on Set Functions

1. Prove Theorem 2 in detail for semirings.
   [Hint: We know that $X_n - X_{n-1} = \bigcup_{i=1}^{m_n} Y_{ni} \text{ (disjoint)}$ for some $Y_{ni} \in \mathcal{C}$, so $s(X_n - X_{n-1}) = \sum_{i=1}^{m_n} sY_{ni}$, with $s$ as in Theorem 1.]

2. Let $s$ be additive on $\mathcal{M}$, a ring. Prove that $s$ is also $\sigma$-additive provided $s$ is either
   (i) left continuous, or
   (ii) finite on $\mathcal{M}$ and right-continuous at $\emptyset$; i.e.,
   \[ \lim_{n \to \infty} sX_n = 0 \]
   when $X_n \searrow \emptyset$ ($X_n \in \mathcal{M}$).
   [Hint: Let $A = \bigcup_{n} A_n \text{ (disjoint), } A, A_n \in \mathcal{M}$.]
Set

\[ X_n = \bigcup_{k=1}^n A_k, \ Y_n = A - X_n. \]

Verify that \( X_n, Y_n \in \mathcal{M}, \ X_n \uparrow A, \ Y_n \downarrow \emptyset. \)

In case (i),

\[ sA = \lim sX_n = \sum_{k=1}^\infty sA_k. \]

(Why?)

For (ii), use the \( Y_n. \]

3. Let \( \mathcal{M} = \{ \text{all intervals in the rational field } R \in \mathbb{E}^1 \}. \)

Let

\[ sX = b - a \]

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 set field;

(ii) \( s \) is continuous;

(iii) \( s \) is additive but not \( \sigma \)-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 \( \sigma \)-additive.]

3'. Let \( N = \{ \text{natural numbers} \}. \) Let

\( \mathcal{M} = \{ \text{all finite subsets of } N \text{ and their complements in } N \}. \)

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 \( \sigma \)-additive.

Thus Problem 2(ii) fails if \( s \) is not finite.

4. Let

\( \mathcal{C} = \{ \text{finite and infinite intervals in } E^1 \}. \)

If \( a, b \) are the endpoints of an interval \( X \) \( a, b \in \mathbb{E}^*, \ a < b. \), set

\[ sX = \begin{cases} b - a, & a < b, \\ 0, & a = b. \end{cases} \]

Show that \( s \) is \( \sigma \)-additive on \( \mathcal{C}, \) a semiring.

Let

\[ X_n = (n, \infty); \]

so \( sX_n = \infty - n = \infty \) and \( X_n \downarrow \emptyset. \) (Verify!) Yet

\[ \lim sX_n = \infty \neq s\emptyset. \]

\section*{§4. Set Functions. Additivity. Continuity}

Does this contradict Theorem 2?

5. Fill in the missing proof details in Theorem 1.

6. Let \( s \) be additive on \( \mathcal{M}. \) Prove the following.

(i) If \( \mathcal{M} \) is a ring or semiring, then

\[ \mathcal{N} = \{ X \in \mathcal{M} \mid |sX| < \infty \} \]

if \( \mathcal{N} \neq \emptyset. \)

(ii) If \( \mathcal{M} \) is generated by a set family \( \mathcal{C}, \) with \( |s| < \infty \) on \( \mathcal{C}, \) then

\[ |s| < \infty \text{ on } \mathcal{M}. \]

[Hint: Use Problem 16 in §4.]

\( \Rightarrow 7. \) (Lebesgue–Stieltjes set functions.) Let \( \alpha \) and \( s_\alpha \) be as in Example (d). Prove the following.

(i) \( s_\alpha \geq 0 \) on \( \mathcal{C} \) if \( \alpha \uparrow \) on \( E^1 \) (see Theorem 2 in Chapter 4, §5).

(ii) \( s_\alpha[p] = s_\alpha[p, p] = 0 \) if \( \alpha \) is continuous at \( p. \)

(iii) \( s_\alpha \) is additive.

[Hint: If \( A = \bigcup_{i=1}^{\infty} A_i \) (disjoint),

the intervals \( A_{i-1}, A_i \) must be adjacent. For two such intervals, consider all cases like

\( (a, b) \cup (b, c), (a, b) \cup [b, c], \) etc.

Then use induction on \( n. \)]

(iv) If \( \alpha \) is right continuous at \( a \) and \( b \), then

\[ s_\alpha(a, b] = \alpha(a) - \alpha(b). \]

If \( \alpha \) is continuous at \( a \) and \( b \), then

\[ s_\alpha[a, b] = s_\alpha(a, b] = s_\alpha[a, b) = s_\alpha(a, b). \]

(v) If \( \alpha \uparrow \) on \( E^1, \) then \( s_\alpha \) satisfies Lemma 1 and Corollary 2 in §1 (same proof), as well as Lemma 1, Theorem 1, Corollaries 1–4, and Note 3 §2 (every except Corollaries 5 and 6).

[Hint: Use (i) and (ii). For Lemma 1 §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

\[ (\forall \varepsilon > 0) \ (\exists \delta > 0) \ 0 \leq \alpha(\varepsilon-) - \alpha(b+) < \varepsilon. \]

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 \( E^1. \)
8. Construct LS set functions (Example (d)), with \( \alpha \uparrow \) (see Problem 7(v)), so that

\begin{enumerate}
    \item \( s_n[0, 1] \neq s_n[1, 2] \);
    \item \( s_n E^3 = 1 \) (after extending \( s_n \) to \( C_\sigma \)-sets in \( E^3 \));
    \item \( s_n E^3 = c \) for a fixed \( c \in (0, \infty) \);
    \item \( s_n \{0\} = 1 \) and \( s_n[0, 1] > s_n(0, 1] \).
\end{enumerate}

Describe \( s_\alpha \) if \( \alpha(x) = \lfloor x \rfloor \) (the integral part of \( x \)).

\[ \text{[Hint: See Figure 16 in Chapter 4, \S1.]} \]

9. For an arbitrary \( \alpha : E^1 \to E^1 \), define \( \sigma_\alpha : C \to E^1 \) by

\[
\sigma_\alpha[a, b] = \sigma_\alpha(a, b) = \sigma_\alpha(a, b) = \alpha(b) - \alpha(a)
\]

(the original Stieltjes method). Prove that \( \sigma_\alpha \) is additive but not \( \sigma \)-additive unless \( \alpha \) is continuous (for Theorem 2 fails).

\section{Nonnegative Set Functions. Premeasures. Outer Measures}

We now concentrate on nonnegative set functions

\[ m : \mathcal{M} \to [0, \infty] \]

(we mostly denote them by \( m \) or \( \mu \)). Such functions have the advantage that

\[ \sum_{n=1}^{\infty} mX_n \]

exists and is permutative (Theorem 2 in \S2) for any sets \( X_n \in \mathcal{M} \), since \( mX_n \geq 0 \). Several important notions apply to such functions (only). They “mimic” \S1 and 2.

\begin{definition}
A set function

\[ m : \mathcal{M} \to [0, \infty] \]

is said to be

\begin{enumerate}
    \item \textit{monotone} (on \( \mathcal{M} \)) iff
        \[ mX \leq mY \]
        whenever \( X \subseteq Y \) and \( X, Y \in \mathcal{M} \);
\end{enumerate}

\end{definition}

\begin{corollary}
Subadditivity implies monotonicity.
\end{corollary}

Take \( n = 1 \) in formula (1).

\begin{corollary}
If \( m : \mathcal{C} \to [0, \infty] \) is additive (\( \sigma \)-additive) on \( \mathcal{C} \), a semiring, then \( m \) is also subadditive (\( \sigma \)-subadditive, respectively), hence monotone, on \( \mathcal{C} \).
\end{corollary}

The proof is a mere repetition of the argument used in Lemma 1 in \S1. Taking \( n = 1 \) in formula (ii) there, we obtain finite subadditivity.

For \( \sigma \)-subadditivity, one only has to use countable unions instead of finite ones.

\begin{note}
The converse fails: subadditivity does not imply additivity.
\end{note}

\begin{note}
Of course, Corollary 2 applies to rings, too (see Corollary 1 in \S3).
\end{note}

\begin{definition}
A \textit{premeasure} is a set function

\[ \mu : \mathcal{C} \to [0, \infty] \]

such that

\[ \emptyset \in \mathcal{C} \text{ and } \mu \emptyset = 0. \]

(\( \mathcal{C} \) may, but need not, be a semiring.)
A premeasure space is a triple 
\[(S, C, \mu),\]
where \(C\) is a family of subsets of \(S\) (briefly, \(C \subseteq 2^S\)) and 
\[\mu : C \to [0, \infty]\]
is a premeasure. In this case, \(C\)-sets are also called basic sets.

If 
\[A \subseteq \bigcup_n B_n,\]
with \(B_n \in C\), the sequence \(\{B_n\}\) is called a basic covering of \(A\), and 
\[\sum_n \mu B_n\]
is a basic covering value of \(A\); \(\{B_n\}\) may be finite or infinite.

Examples.
(a) The volume function \(v\) on \(C\) (intervals in \(E^n\)) is a premeasure, as \(v \geq 0\) and \(\emptyset = 0\). \((E^n, C, v)\) is the Lebesgue premeasure space.
(b) The LS set function \(s\), is a premeasure if \(\sigma^*\) (see Problem 7 in §4). We call it the \(\alpha\)-induced Lebesgue–Stieltjes (LS) premeasure in \(E^1\).

We now develop a method for constructing \(\sigma\)-subadditive premeasures. (This is a first step toward achieving \(\sigma\)-additivity; see §4.)

Definition 3.

For any premeasure space \((S, C, \mu)\), we define the \(\mu\)-induced outer measure \(m^*\) on \(2^S\) (= all subsets of \(S\)) by setting, for each \(A \subseteq S\),
\[m^* A = \inf \left\{ \sum_n \mu B_n \mid A \subseteq \bigcup_n B_n, B_n \in C \right\},\]
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 
\[(\forall A \subseteq S) \quad 0 \leq m^* A.\]

If \(A \in C\), then \(\{A\}\) is a basic covering; so
\[m^* A \leq \mu A.\]

\(\S\) Nonnegative Set Functions. Premeasures. Outer Measures

In particular, \(m^0 = \mu \emptyset = 0\).

Theorem 1.\(^1\) The set function \(m^*\) so defined is \(\sigma\)-subadditive on \(2^S\).

Proof. Given 
\[A \subseteq \bigcup_n A_n \subseteq S,\]
we must show that 
\[m^* A \leq \sum_n m^* A_n.\]

This is trivial if \(m^* A_n = \infty\) for some \(n\). Thus assume 
\[(\forall n) \quad m^* A_n < \infty\]
and fix \(\varepsilon > 0\).

By Note 3, each \(A_n\) has a basic covering 
\[\{B_{nk}\}, \quad k = 1, 2, \ldots\]
(otherwise, \(m^* A_n = \infty\)). By properties of the glb, we can choose the \(B_{nk}\) so that 
\[(\forall n) \quad \sum_k \mu B_{nk} < m^* A_n + \frac{\varepsilon}{2^n}.\]
(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,
\[m^* A \leq \sum_n \left( \sum_k \mu B_{nk} \right) \leq \sum_n \left( m^* A_n + \frac{\varepsilon}{2^n} \right) \leq \sum_n m^* A_n + \varepsilon.\]

As \(\varepsilon\) is arbitrary, we can let \(\varepsilon \to 0\) to obtain the desired result. \(\square\)

Note 5. In view of Theorem 1, we now generalize the notion of an outer measure in \(S\) to mean any \(\sigma\)-subadditive premeasure defined on all of \(2^S\).

By Note 4, \(m^* \leq \mu\) on \(C\), not \(m^* = \mu\) in general. However, we obtain the following result.

Theorem 2. With \(m^*\) as in Definition 3, we have \(m^* = \mu\) on \(C\) iff \(\mu\) is \(\sigma\)-subadditive on \(C\). Hence, in this case, \(m^*\) is an extension of \(\mu\).

Proof. Suppose \(\mu\) is \(\sigma\)-subadditive and fix any \(A \in C\). By Note 4,
\[m^* A \leq \mu A.\]

We shall show that 
\[\mu A \leq m^* A,\]

\(^1\) Theorems 1–3 are redundant for a “limited approach” (see the preface). Pass to Chapter 8, §1.
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:

\[
A \subseteq \bigcup_{n} B_n, \quad B_n \in \mathcal{C}.
\]

As \( \mu \) is \( \sigma \)-subadditive,

\[
\mu A \leq \sum_{n} \mu B_n.
\]

Thus \( \mu A \) 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 \( \sigma \)-subadditivity of \( m^* \) (Theorem 1) implies that of \( \mu \) (on \( \mathcal{C} \)). Thus all is proved. \( \square \)

**Note 6.** If, in (2), we allow only finite basic coverings, then the \( \mu \)-induced set function is called the \( \mu \)-induced outer content, \( \varepsilon^* \). 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 \( s: \mathcal{M} \to E \) (\( \mathcal{M} \subseteq 2^S \)) is called \( \sigma \)-finite iff every \( X \in \mathcal{M} \) can be covered by a sequence of \( \mathcal{M} \)-sets \( X_n \), with

\[
|sX_n| < \infty \quad (\forall n).
\]

Any set \( A \subseteq S \) which can be so covered is said to be \( \sigma \)-finite with respect to \( s \) (briefly, \( (s) \)-finite).

If the whole space \( S \) can be so covered, we say that \( s \) is totally \( \sigma \)-finite.

For example, the Lebesgue premeasure \( v \) on \( E^n \) is totally \( \sigma \)-finite.

**Definition 5.**

A set function \( s: \mathcal{M} \to E^* \) is said to be regular with respect to a set family \( A \) (briefly, \( A \)-regular) iff for each \( A \in \mathcal{M} \),

\[
sA = \inf \{sX | A \subseteq X, X \in \mathcal{A}\};
\]

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 3.** For any premeasure space \((S, \mathcal{C}, \mu)\), the \( \mu \)-induced outer measure \( m^* \) is \( A \)-regular whenever

\[
C_n \subseteq A \subseteq 2^S.
\]

Thus in this case,

\[
(\forall A \subseteq S) \quad m^*A = \inf \{m^*X | A \subseteq X, X \in \mathcal{A}\}.
\]

**Proof.** As \( m^* \) is monotone, \( m^*A \) is surely a lower bound of

\[
\{m^*X | A \subseteq X, X \in \mathcal{A}\}.
\]

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

\[
A \subseteq \bigcup_{n} B_n
\]

and

\[
m^*A + \varepsilon > \sum_{n} \mu B_n \geq \sum_{n} m^*B_n \geq m^* \bigcup_{n} B_n.
\]

\((m^* \text{ is } \sigma \text{-subadditive!})\)

Let

\[
X = \bigcup_{n} B_n.
\]

Then \( X \) is in \( C_n \), hence in \( A \), and \( A \subseteq X \). Also,

\[
m^*A + \varepsilon > m^*X.
\]

Thus \( m^*A + \varepsilon \) is not a lower bound of

\[
\{m^*X | A \subseteq X, X \in \mathcal{A}\}.
\]

This proves (4). \( \square \)

**Problems on Premeasures and Related Topics**

1. Fill in the missing details in the proofs, notes, and examples of this section.
2. Describe \( m^* \) on \( 2^S \) induced by a premeasure \( \mu: \mathcal{C} \to E^* \) 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.
3. Show that the premeasures

\[ v': C' \to [0, \infty) \]

induce one and the same (Lebesgue) outer measure \( m^* \) in \( E^n \), with \( v' = v \) (volume, as in \( \S 2 \)):

(a) \( C' = \{ \text{open intervals} \} \);
(b) \( C' = \{ \text{half-open intervals} \} \);
(c) \( C' = \{ \text{closed intervals} \} \);
(d) \( C' = C'_\sigma \);
(e) \( C' = \{ \text{open sets} \} \);
(f) \( C' = \{ \text{half-open cubes} \} \).

[Hints: (a) Let \( m' \) be the \( v' \)-induced outer measure; let \( C = \{ \text{all intervals} \} \). As \( C' \subseteq C \), \( m'A \geq m^*A \) (Why?) Also,

\[
(\forall \varepsilon > 0) (\exists \{B_k\} \subseteq C) \quad A \subseteq \bigcup_k B_k \text{ and } \sum vB_k \leq m^*A + \varepsilon.
\]

(Why?) By Lemma 1 in \( \S 2 \),

\[
(\exists \{C_k\} \subseteq C') \quad B_k \subseteq C_k \text{ and } vB_k + \frac{\varepsilon}{2^n} > v'C_k.
\]

Deduce that \( m^*A \geq m'A, \) \( m^* = m' \). Similarly for (b) and (c). For (d), use Corollary 1 and Note 3 in \( \S 1 \). For (e), use Lemma 2 in \( \S 2 \). For (f), use Problem 2 in \( \S 2 \).]

3’. Do Problem 3(a)–(c), with \( m^* \) replaced by the Jordan outer content \( c^* \) (Note 6).

4. Do Problem 3, with \( v \) and \( m^* \) replaced by the LS premeasure and outer measure. (Use Problem 7 in \( \S 4 \)).

5. Show that a set \( A \subseteq E^n \) is bounded iff its outer Jordan content is finite.

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 \( \S 2 \)).

7. Let

\[ \mu_1, \mu_2: C \to [0, \infty] \]

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 \( C \), then \( m_1^* = m_2^* \) on all of \( 2^S \).

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

\[ \{X \subseteq S \mid c^*A = 0\} \]

is a hereditary set ring, i.e., a ring \( R \) such that

\[ (\forall B \in R) (\forall A \subseteq B) \quad A \in R. \]

(iii) The set family

\[ \{X \subseteq S \mid m^*X = 0\} \]

is a hereditary \( \sigma \)-ring.

(iv) So also is

\[ \mathcal{H} = \{ \text{those } X \subseteq S \text{ that have basic coverings} \}; \]

thus \( \mathcal{H} \) is the hereditary \( \sigma \)-ring generated by \( C \) (see Problem 14 in \( \S 3 \)).

9. Continuing Problem 8(iv), prove that if \( \mu \) is \( \sigma \)-finite (Definition 4), so is \( m^* \) when restricted to \( \mathcal{H} \).

Show, moreover, that if \( C \) 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

\[ X = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} B_{nk} \]

for some sets \( B_{nk} \in C \), with \( \mu B_{nk} < \infty \). Then use Note 4 in \( \S 5 \) and Corollary 1 of \( \S 1 \).]

10. Show that if

\[ s: C \to E^n \]

is \( \sigma \)-finite and additive on \( C \), a semiring, then the \( \sigma \)-ring \( R \) generated by \( C \) equals the \( \sigma \)-ring \( R' \) generated by

\[ C' = \{X \in C \mid |sX| < \infty\} \]

(cf. Problem 6 in \( \S 4 \)).

[Hint: By \( \sigma \)-finiteness,

\[ (\forall X \in C) (\exists \{A_n\} \subseteq C \mid |sA_n| < \infty) \quad X \subseteq \bigcup_n A_n; \]

so

\[ X = \bigcup_n (X \cap A_n), \quad X \cap A_n \in C'. \]

(Use Lemma 3 in \( \S 4 \)).]
11. With all as in Theorem 3, prove that if \( A \) has basic coverings, then
\[
(\exists B \in \mathcal{A}_k) \quad A \subseteq B \quad \text{and} \quad m^* A = m^* B.
\]
[Hint: By formula (4),
\[
(\forall n \in N) \quad (\exists X_n \in \mathcal{A} \mid A \subseteq X_n) \quad m^* A \leq m X_n \leq m^* A + \frac{1}{n}.
\]
(Explain!) Set
\[
B = \bigcap_{n=1}^{\infty} X_n \in \mathcal{A}_k.
\]
Proceed. For \( \mathcal{A}_k \), see Definition 2(b) in §3.

12. Let \((\mathcal{S}, \mathcal{C}, \mu)\) and \(m^*\) be as in Definition 3. Show that if \( \mathcal{C} \) is a \( \sigma \)-field in \( S \), then
\[
(\forall A \subseteq S) \quad (\exists B \in \mathcal{C}) \quad A \subseteq B \quad \text{and} \quad m^* A = \mu B.
\]
[Hint: Use Problem 11 and Note 3.]

13. Show that if
\[
s : \mathcal{C} \to E
\]
is \( \sigma \)-finite and \( \sigma \)-additive on \( \mathcal{C} \), a semiring, then \( s \) has at most one \( \sigma \)-additive extension to the \( \sigma \)-ring \( \mathcal{R} \) generated by \( \mathcal{C} \).

(Nota note: \( s \) is automatically \( \sigma \)-finite if it is finite, e.g., complex or vector valued.)

[Outline: Let \( s', s'' : \mathcal{R} \to E \) be two \( \sigma \)-additive extensions of \( s \). By Problem 10, \( \mathcal{R} \) is also generated by \( \mathcal{C}' \),
\[
\mathcal{C}' = \{ X \in \mathcal{C} \mid |s| X < \infty \}.
\]
Now set
\[
\mathcal{R}^* = \{ X \in \mathcal{R} \mid s' X = s'' X \}.
\]
Show that \( \mathcal{R}^* \) satisfies properties (i)-(iii) of Theorem 3 in §3, with \( \mathcal{C} \) replaced by \( \mathcal{C}' \); so \( \mathcal{R} = \mathcal{R}^* \).]

14. Let \( m_n^* \) \((n = 1, 2, \ldots)\) be outer measures in \( S \) such that
\[
(\forall X \subseteq S) \quad (\forall n) \quad m_n^* X \leq m_{n+1}^* X.
\]
Set
\[
m^* = \lim_{n \to \infty} m_n^*.
\]
Show that \( m^* \) is an outer measure in \( S \) (see Note 5).

15. An outer measure \( m^* \) in a metric space \((S, \rho)\) is said to have the Carathéodory property (CP) iff
\[
m^*(X \cup Y) \geq m^* X + m^* Y
\]
whenever \( \rho(X, Y) > 0 \), where
\[
\rho(X, Y) = \inf \{ \rho(x, y) \mid x \in X, y \in Y \}.
\]
For such \( m^* \), prove that
\[
m^*(\bigcup_k X_k) = \sum_k m^* X_k
\]
if \( \{X_k\} \subseteq 2^S \) and
\[
\rho(X_i, X_k) > 0 \quad (i \neq k).
\]
[Hint: For finite unions, use the CP, subadditivity, and induction. Deduce that
\[
(\forall n) \quad \sum_{k=1}^{n} m^* X_k \leq m^* \bigcup_{k=1}^{\infty} X_k.
\]
Let \( n \to \infty \). Proceed.]

16. Let \((\mathcal{S}, \mathcal{C}, \mu)\) and \(m^*\) be as in Definition 3, with \( \rho \) a metric for \( S \). Let \( \mu_n \) be the restriction of \( \mu \) to the family \( \mathcal{C}_n \) of all \( X \in \mathcal{C} \) of diameter
\[
dX \leq \frac{1}{n}.
\]
Let \( m_n^* \) be the \( \mu_n \)-induced outer measure in \( S \).

Prove that
(i) \( \{m_n^*\}^+ \) as in Problem 14;
(ii) the outer measure
\[
\mu^* = \lim_{n \to \infty} m_n^*
\]
has the CP (see Problem 15), and
\[
\mu^* \geq m^* \quad \text{on} \quad 2^S.
\]
[Outline: Let \( \rho(X, Y) > \varepsilon > 0 \quad (X, Y \subseteq S) \).
If for some \( n \), \( X \cup Y \) has no basic covering from \( \mathcal{C}_n \), then
\[
\mu^*(X \cup Y) \geq m^*_n (X \cup Y) = \infty \geq \mu^* X + \mu^* Y,
\]
and the CP follows. (Explain!) Thus assume
\[
\left( \forall n \right) \quad \left( \forall k \right) \quad (\exists B_{nk} \in \mathcal{C}_n) \quad X \cup Y \subseteq \bigcup_{k=1}^{\infty} B_{nk}.
\]
1. In §5, we considered premeasure spaces, stressing mainly the idea of \( \sigma \)-subadditivity (Note 5 in §5). Now we shall emphasize \( \sigma \)-additivity.

**Definition 1.**

A premeasure

\[ m : \mathcal{M} \to [0, \infty] \]

is called a measure (in \( S \)) iff \( \mathcal{M} \) is a \( \sigma \)-ring (in \( S \)), and \( m \) is \( \sigma \)-additive on \( \mathcal{M} \).

If so, the system

\[ (S, \mathcal{M}, m) \]

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 1.** Measures are \( \sigma \)-additive, \( \sigma \)-subadditive, monotone, and continuous.

**Proof.** Use Corollary 2 in §5 and Theorem 2 in §4, noting that \( \mathcal{M} \) is a \( \sigma \)-ring. \( \square \)

**Corollary 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 \( X \setminus Y \) if \( X, Y \in \mathcal{M} \).

This is obvious since \( \mathcal{M} \) is a \( \sigma \)-ring.

As measures and other premeasures are understood to be \( \geq 0 \), we often write

\[ m : \mathcal{M} \to E^* \]

for

\[ m : \mathcal{M} \to [0, \infty]. \]

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 \( \sigma \)-additive premeasure, but not a measure since its domain (the intervals) is not a \( \sigma \)-ring.

(b) Let \( \mathcal{M} = 2^S \). Define

\[ (\forall X \subseteq S) \quad mX = 0. \]

---

1 Sections 6–12 are not needed for a “limited approach.” (Pass to Chapter 8, §1.)
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

$$mX = \begin{cases} 
1 & \text{if } p \in X, \\
0 & \text{otherwise.}
\end{cases}$$

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

$$S \in \mathcal{M} \text{ and } mS = 1.$$  

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),

$$\mathcal{M} = 2^S \text{ (all subsets of } S).$$

More often, however,

$$\mathcal{M} \neq 2^S,$$

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 *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 $m : \mathcal{M} \to E^*$ 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 $\sigma$-subadditive premeasure defined on all of $2^S$ (even if it is not derived via **Definition 3** in §5).² In Examples (b), (c), and (d), $m$ is both a measure and an outer measure. (Why?)

An outer measure

$$m^* : 2^S \to E^*$$

need not be additive; but consider this fact:

Any set $A \subseteq S$ splits $S$ into two parts: $A$ itself and $-A$.

It also splits any other set $X$ into $X \cap A$ and $X - A$; indeed,

$$X = (X \cap A) \cup (X - A) \text{ (disjoint).}$$

We want to single out those sets $A$ for which $m^*$ behaves *additively," i.e.,

$$m^*X = m^*(X \cap A) + m^*(X - A).$$

This motivates our next definition.

**Definition 3.**

Given an outer measure $m^* : 2^S \to E^*$ 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,

$$\forall X \subseteq S \quad m^*(X \cap A) + m^*(X - A).$$

As is easily seen (see Problem 1), this is equivalent to

$$\forall X \subseteq A \quad (\forall Y \subseteq -A) \quad m^*(X \cup Y) = m^*X + m^*Y.$$  

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

$$\{X \cap A, X - A\},$$

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 1.** In any outer measure space

$$(S, \mathcal{M}^*, m^*),$$

the family $\mathcal{M}^*$ of all $m^*$-measurable sets is a $\sigma$-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 $\mathcal{M}^*$).

We split the proof into several steps (lemmas).
Lemma 1. $\mathcal{M}^*$ is closed under complementation:

$$(\forall A \in \mathcal{M}^*) \quad -A \in \mathcal{M}^*.$$  

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,

$$m^*X \leq m^*A = 0$$

and

$$m^*Y \leq m^*(X \cup Y).$$

Thus

$$m^*X + m^*Y = 0 + m^*Y \leq m^*(X \cup Y),$$

as required.

In particular, as $m^*\emptyset = 0$, $\emptyset$ is $m^*$-measurable ($\emptyset \in \mathcal{M}^*$).

So is $S$ (the complement of $\emptyset$) by Lemma 1. □

Lemma 3. $\mathcal{M}^*$ is closed under finite unions:

$$(\forall A, B \in \mathcal{M}^*) \quad A \cup B \in \mathcal{M}^*.$$  

Proof. This time we shall use formula (1). By Note 2, it suffices to show that

$$(\forall X \subseteq S) \quad m^*X \geq m^*(X \cap (A \cup B)) + m^*(X - (A \cup B)).$$

Fix any $X \subseteq S$; as $A \in \mathcal{M}^*$, we have

$$m^*X = m^*(X \cap A) + m^*(X - A).$$

Similarly, as $B \in \mathcal{M}^*$, we have (replacing $X$ by $X - A$ in (1))

$$m^*(X - A) = m^*((X - A) \cap B) + m^*(X - (A - B))$$

$$= m^*(X \cap -A \cap B) + m^*(X - (A \cup B)),$$

since

$$X - A = X \cap -A$$

and

$$X - A - B = X - (A \cup B).$$

Combining (4) with (3), we get

$$m^*X = m^*(X \cap A) + m^*(X \cap -A \cap B) + m^*(X - (A \cup B)).$$

Now verify that

$$(X \cap A) \cup (X \cap -A \cap B) \supseteq X \cap (A \cup B),$$

As $m$ is subadditive, this yields

$$m^*(X \cap A) + m^*(X \cap -A \cap B) \geq m^*(X \cap (A \cup B)).$$

Combining with (5), we get

$$m^*X \geq m^*(X \cap (A \cup B)) + m^*(X - (A \cup B)),$$

so that $A \cup B \in \mathcal{M}^*$, indeed. □

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

$$X_k \subseteq A_k \subseteq S, \quad k = 0, 1, 2, \ldots,$$

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

$$m^*\left(\bigcup_{k=0}^{\infty} X_k\right) = \sum_{k=0}^{\infty} m^*X_k.$$  

Proof. We start with two sets, $A_0$ and $A_1$; so

$$A_1 \in \mathcal{M}^*, \quad A_0 \cap A_1 = \emptyset, \quad X_0 \subseteq A_0, \quad \text{and} \quad X_1 \subseteq A_1.$$  

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

$$X = X_1 \subseteq A_1 \quad \text{and} \quad Y = X_0 \subseteq -A,$$

to obtain

$$m^*(X_0 \cup X_1) = m^*X_0 + m^*X_1.$$  

Thus (6) holds for two sets.

Induction now easily yields

$$(\forall n) \quad \sum_{k=0}^{n} m^*X_k = m^*\left(\bigcup_{k=0}^{n} X_k\right) \leq m^*\left(\bigcup_{k=0}^{\infty} X_k\right)$$

by monotonicity of $m^*$. Now let $n \to \infty$ and pass to the limit to get

$$\sum_{k=0}^{\infty} m^*X_k \leq m^*\left(\bigcup_{k=0}^{\infty} X_k\right).$$

As $\bigcup X_k$ is covered by the $X_k$, the $\sigma$-subadditivity of $m^*$ yields the reverse inequality as well. Thus (6) is proved. □
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

$$U = \bigcup_{k=1}^{\infty} A_k, \quad A_k \in \mathcal{M}^*.$$  

We have to prove that $U \in \mathcal{M}^*$; or by (2) and Note 2,

$$m^*(\forall X \subseteq U \quad \forall Y \subseteq -U) \quad m^*(X \cup Y) \geq m^*X + m^*Y.$$  

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 of §1.)

To prove (7), fix any $X \subseteq U$ and $Y \subseteq -U$, and let

$$X_k = X \cap A_k \subseteq A_k, \quad A_0 = -U, \quad Y = \bigcup_{k=1}^{\infty} X_k,$$

satisfying all assumptions of Lemma 4. Thus by (6), writing the first term separately, we have

$$m^*\left(\bigcup_{k=1}^{\infty} X_k\right) = m^*Y + \sum_{k=1}^{\infty} m^*X_k.$$  

But

$$\bigcup_{k=1}^{\infty} X_k = \bigcup_{k=1}^{\infty} (X \cap A_k) = X \cap \bigcup_{k=1}^{\infty} A_k = X \cap U = X$$

(as $X \subseteq U$). Also, by σ-subadditivity,

$$\sum_{k=1}^{\infty} m^*X_k \geq m^*\bigcup_{k=1}^{\infty} X_k = m^*X.$$  

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

$$X_k = A_k \in \mathcal{M}^*, \quad A_0 = \emptyset.$$  

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.

Then all is proved. □

We thus have a standard method for constructing measures: From a premeasure

$$\mu : \mathcal{C} \to E^*$$

in $S$, we obtain the $\mu$-induced outer measure

$$m^* : 2^S \to E^* (\S 5);$$

this, in turn, induces a complete measure

$$m : \mathcal{M}^* \to E^*.$$  

But we need more: We want $m$ to be an extension of $\mu$, i.e.,

$$m = \mu$$

on $\mathcal{C}$, with $\mathcal{C} \subseteq \mathcal{M}^*$ (meaning that all $\mathcal{C}$-sets are $m^*$-measurable). We now explore this question.

Lemma 5. Let $(\mathcal{S}, \mathcal{C}, \mu)$ and $m^*$ be as in Definition 3 of §5. Then for a set $A \subseteq S$ to be $m^*$-measurable, it suffices that

$$m^*X \geq m^*(X \cap A) + m^*(X - A)$$

for all $X \in \mathcal{C}$.

Proof. Assume (9). 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 $\{B_n\} \subseteq \mathcal{C}$ so that

$$X \subseteq \bigcup_{n=1}^{\infty} B_n$$

and

$$m^*X + \varepsilon > \sum_{k=1}^{\infty} \mu B_n \geq \sum_{k=1}^{\infty} m^*B_n.$$  

(Explain!)

Now, as $X \subseteq \bigcup B_n$, we have

$$X \cap A \subseteq \bigcup B_n \cap A = \bigcup (B_n \cap A).$$

Similarly,

$$X - A = X \cap -A \subseteq \bigcup (B_n - A).$$

Hence, as $m^*$ is σ-subadditive and monotone, we get

$$m^*(X \cap A) + m^*(X - A) \leq m^*\left(\bigcup (B_n \cap A)\right) + m^*\left(\bigcup (B_n - A)\right)$$

$$\leq \sum [m^*(B_n \cap A) + m^*(B_n - A)].$$

But by assumption, (9) holds for any $\mathcal{C}$-set, hence for each $B_n$. Thus

$$m^*(B_n \cap A) + m^*(B_n - A) \leq m^*B_n,$$

and (11) yields

$$m^*(X \cap A) + m^*(X - A) \leq \sum [m^*(B_n \cap A) + m^*(B_n - A)] \leq \sum m^*B_n.$$  

Therefore, by (10),
\[ m^*(X \cap A) + m^*(X - A) \leq m^*X + \varepsilon. \]
Making \( \varepsilon \to 0 \), we prove (10) for any \( X \subseteq S \), so that \( A \in \mathcal{M}^* \), as required. \( \Box \)

**Theorem 2.** Let the premeasure

\[ \mu: \mathcal{C} \to E^* \]

be \( \sigma \)-additive on \( \mathcal{C} \), a semiring in \( S \). Let \( m^* \) be the \( \mu \)-induced outer measure, and

\[ m: \mathcal{M}^* \to E^* \]

be the \( m^* \)-induced measure. Then

(i) \( \mathcal{C} \subseteq \mathcal{M}^* \) and

(ii) \( \mu = m^* = m \) on \( \mathcal{C} \).

Thus \( m \) is a \( \sigma \)-additive extension of \( \mu \) (called its Lebesgue extension) to \( \mathcal{M}^* \).

**Proof.** By Corollary 2 in \( \S 5 \), \( \mu \) is also \( \sigma \)-subadditive on the semiring \( \mathcal{C} \). Thus by Theorem 2 in \( \S 5 \), \( \mu = m^* \) 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

\[ X - A = \bigcup_{k=1}^{n} A_k \text{ (disjoint)} \]

for some sets \( A_k \in \mathcal{C} \). Hence

\[ m^*(X \cap A) + m^*(X - A) = m^*(X \cap A) + m^* \bigcup_{k=1}^{n} A_k \]

\[ \leq m^*(X \cap A) + \sum_{k=1}^{n} m^*A_k. \] (12)

As \( X = (X \cap A) \cup (X - A) = (X \cap A) \cup \bigcup_{k=1}^{n} A_k \text{ (disjoint)} \),

the additivity of \( \mu \) and the equality \( \mu = m^* \) on \( \mathcal{C} \) yield

\[ m^*X = m^*(X \cap A) + \sum_{k=1}^{n} m^*A_k. \]

Hence by (12),

\[ m^*X \geq m^*(X \cap A) + m^*(X - A); \]

so by Lemma 5, \( A \in \mathcal{M}^* \), as required.

Also, by definition, \( m = m^* \) on \( \mathcal{M}^* \), hence on \( \mathcal{C} \). Thus

\[ \mu = m^* = m \] on \( \mathcal{C} \),

as claimed. \( \Box \)

**Note 3.** In particular, Theorem 2 applies if

\[ \mu: \mathcal{M} \to E^* \]

is a measure (so that \( \mathcal{C} = \mathcal{M} \) is even a \( \sigma \)-ring).

Thus any such \( \mu \) can be extended to a complete measure \( m \) (its Lebesgue extension) on a \( \sigma \)-field

\[ \mathcal{M}^* \supseteq \mathcal{M} \]

via the \( \mu \)-induced outer measure (call it \( \mu^* \) this time), with

\[ \mu^* = m = \mu \] on \( \mathcal{M} \).

Moreover,

\[ \mathcal{M}^* \supseteq \mathcal{M} \supseteq \mathcal{M}_\sigma \]

(see Note 2 in \( \S 3 \); so \( \mu^* \) is \( \mathcal{M} \)-regular and \( \mathcal{M}^* \)-regular (Theorem 3 of \( \S 5 \)).

**Note 4.** A reapplication of this process to \( m \) does not change \( m \) (Problem 16).

---

**Problems on Measures and Outer Measures**

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).]

2. Given an outer measure space \( (S, \mathcal{M}^*, m^*) \) and \( A \subseteq S \), set

   \[ A \cap \mathcal{M}^* = \{ A \cap X \mid X \in \mathcal{M}^* \} \]

   (all sets of the form \( A \cap X \) with \( X \in \mathcal{M}^* \)).

   Prove that \( A \cap \mathcal{M}^* \) is a \( \sigma \)-field in \( A \), and \( m^* \) is \( \sigma \)-additive on it.

   [Hint: Use Lemma 4, with \( X_k = A \cap A_k \in A \cap \mathcal{M}^* \).]

3. Prove Lemmas 1 and 2, using formula (1).

3'. Prove Corollary 1.

4. Verify Examples (b), (c), and (d). Why is \( m \) an outer measure as well?

   [Hint: Use Corollary 2 in \( \S 5 \).

5. Fill in all details (induction, etc.) in the proofs of this section.
6. Verify that \( m^* \) is an outer measure and describe \( \mathcal{M}^* \) under each of the following conditions.
   (a) \( m^*A = 1 \) if \( \emptyset \subset A \subset S; \) \( m^*\emptyset = 0. \)
   (b) \( m^*A = 1 \) if \( \emptyset \subset A \subset S; \) \( m^*S = 2; \) \( m^*\emptyset = 0. \)
   (c) \( m^*A = 0 \) if \( A \subset S \) is countable; \( m^*A = 1 \) otherwise (\( S \) is uncountable).
   (d) \( S = N \) (naturals); \( m^*A = 1 \) if \( A \) is infinite; \( m^*A = \frac{1}{n+1} \) if \( A \) has \( n \) elements.

7. Prove the following.
   (i) An outer measure \( m^* \) is \( \mathcal{M}^* \)-regular (Definition 5 in §5) iff
   \[
   (\forall A \subseteq S) \ (\exists B \in \mathcal{M}^*) \ A \subseteq B \text{ and } m^*A = mB.
   \]
   \[
   \text{[Hint: If } m^*A = \inf\{mX \mid A \subseteq X \in \mathcal{M}^*\}, \text{ then}
   \]
   \[
   (\forall n) \ (\exists \ X_n \in \mathcal{M}^*) \ A \subseteq X_n \text{ and } mX_n \leq m^*A + \frac{1}{n}.
   \]
   Set \( B = \bigcap_{n=1}^{\infty} X_n \).
   (ii) If \( m^* \) is as in Definition 3 of §5, with \( C \subseteq \mathcal{M}^* \), then \( m^* \) is \( \mathcal{M}^* \)-regular.

8. Show that if \( m^* \) is \( \mathcal{M}^* \)-regular (Problem 7), it is left continuous.
   [Hints: Let \( \{A_n\} \uparrow \); let \( B_n \) be a measurable cover of \( A_n \); set
   \[
   C_n = \bigcap_{k=n}^{\infty} B_k.
   \]
   Verify that \( \{C_n\} \uparrow, B_n \supseteq C_n \supseteq A_n \), and \( mC_n = m^*A_n \).
   By the left continuity of \( m \) (Theorem 2 in §4),
   \[
   \lim m^*A_n = \lim mC_n = m \bigcup_{n=1}^{\infty} C_n \geq m^* \bigcup_{n=1}^{\infty} A_n.
   \]
   Prove the reverse inequality as well.]

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.)

10. In Problem 2, let \( m^* \) be the restriction of \( m^* \) to \( 2^A \). Prove the following.
    (a) \( m^* \) is an outer measure in \( A \).
    (b) \( A \cap \mathcal{M}^* \subseteq \mathcal{M}^* = \{m^*-measurable sets\}. \)
    (c) \( A \cap \mathcal{M}^* = \mathcal{M}^* \) if \( A \in \mathcal{M}^* \), or if \( m^* \) is \( \mathcal{M}^* \)-regular (see Problem 7) and finite.
    (d) \( m^* \) is \( \mathcal{M}^* \)-regular.

11. Show that if \( m^* \) is \( \mathcal{M}^* \)-regular and finite, then \( A \subseteq S \) is \( m^* \)-measurable iff
    \[
    mS = m^*A + m^*(-A).
    \]
    [Hint: Assume the latter. By Problem 7,
    \[
    (\forall X \subseteq S) \ (\exists B \in \mathcal{M}^*, B \supseteq X) \ m^*X = mB;
    \]
    so
    \[
    m^*A = m^*(A \cap B) + m^*(A - B).
    \]
    Similarly for \( -A \). Deduce that
    \[
    m^*(A \cap B) + m^*(A - B) + m^*(B - A) + m^*(-A - B) = mS = mB + m(-B);
    \]
    hence
    \[
    m^*X = mB \geq m^*(B \cap A) + m^*(B - A) \geq m^*(X \cap A) + m^*(X - A);
    \]
    so \( A \in \mathcal{M}^* \).

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
    \[
    (\forall X \subseteq G) \ (\forall Y \subseteq -G) \ m^*(X \cup Y) \geq m^*X + m^*Y,
    \]
    assuming \( m^*X < \infty \). (Why?) Set
    \[
    D_0 = \{x \in X \mid \rho(x, -G) \geq 1\}
    \]
    and
    \[
    D_k = \{x \in X \mid \frac{1}{k+1} \leq \rho(x, -G) < \frac{1}{k}\}; \quad k \geq 1.
    \]
    Prove that
    (i) \( X = \bigcup_{k=0}^{\infty} D_k \)
    and
    (ii) \( \rho(D_k, D_{k+1}) > 0 \);
    so by Problem 15 in §5,
    \[
    \sum_{n=0}^{\infty} m^*D_{2n} = m^* \bigcup_{n=0}^{\infty} D_{2n} \leq m^* \bigcup_{n=0}^{\infty} D_{2n} = m^*X < \infty.
    \]
Similarly, 
\[ \sum_{n=0}^{\infty} m^*D_{2n+1} \leq m^*X < \infty. \]
Hence 
\[ \sum_{n=0}^{\infty} m^*D_n < \infty; \]
so 
\[ \lim_{n \to \infty} \sum_{k=0}^{\infty} m^*D_k = 0. \]
(Why?) Thus 
\[ (\forall \varepsilon > 0) \left( \exists n \right) \sum_{k=0}^{n} m^*D_k < \varepsilon. \]
Also, 
\[ X = \bigcup_{k=0}^{\infty} D_k = \bigcup_{k=0}^{n-1} D_k \cup \bigcup_{k=n}^{\infty} D_k; \]
so 
\[ m^*X \leq m^* \bigcup_{k=0}^{n-1} D_k + \sum_{k=n}^{\infty} m^*D_k < m^* \bigcup_{k=0}^{n-1} D_k + \varepsilon. \]
Adding \( m^*Y \) on both sides, get 
\[ (iii) \quad m^*X + m^*Y \leq m^* \bigcup_{k=0}^{n-1} D_k + m^*Y + \varepsilon. \]
Moreover, 
\[ \rho \left( \bigcup_{k=0}^{n-1} D_k, Y \right) > 0, \]
for \( Y \subseteq G \) and 
\[ \rho(D_k, -G) \geq \frac{1}{k+1}. \]
Hence by the CP, 
\[ m^*Y + \sum_{k=0}^{n-1} m^*D_k = m^* \left( Y \cup \bigcup_{k=0}^{n-1} D_k \right) < m^*(Y \cup X). \]
(Why?) Combining with (iii), obtain 
\[ m^*X + m^*Y \leq m^*(X \cup Y) + \varepsilon. \]
Now let \( \varepsilon \to 0. \]
\[ \Rightarrow 14. \quad \text{Given a measure } m : \mathcal{M} \to E^*, \text{ let } \overline{\mathcal{M}} = \{ \text{all sets of the form } X \cup Z \text{ where } X \in \mathcal{M} \text{ and } Z \text{ is } m\text{-null} \}. \]
Prove that \( \overline{\mathcal{M}} \) is a \( \sigma \)-ring \( \supseteq \mathcal{M}. \)
[Hint: To prove that \( (\forall A, B \in \overline{\mathcal{M}}) \ A - B \in \overline{\mathcal{M}}, \)]
suppose first \( A \in \mathcal{M} \) and \( B \) is "null," i.e., \( B \subseteq U \subseteq \mathcal{M}, mU = 0. \)
Show that 
\[ A - B = X \cup Z, \]
with \( X = A - U \in \mathcal{M} \) and \( Z = A \cap U - B \text{-null} \) \( (Z \text{ is shaded in Figure 31}) \).
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 
\[ A - B = (X \cup Z) - B = (X - B) \cup (Z - B) = (X - B) \cup Z', \]
where 
\[ Z'' = Z - B \]
is m-null. Also, \( B = X' \cup Z' \) implies 
\[ X - B = (X - X') - Z' \in \overline{\mathcal{M}}, \]
by the first part of the proof.
Deduce that 
\[ A - B = (X - B) \cup Z'' \in \overline{\mathcal{M}}, \]
(after checking closure under unions.).
\[ \Rightarrow 15. \quad \text{Continuing Problem 14, define } \overline{m} : \overline{\mathcal{M}} \to E^* \text{ by setting } \overline{m}A = mX \]
whenever \( A = X \cup Z \), with \( X \in \mathcal{M} \) and \( Z \text{-null} \). (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 \( \overline{m} = \overline{m} \) on \( \mathcal{M} \).
(ii) $\mathfrak{m}$ is the least complete extension of $m$; that is, if $n: N \to E^*$ is another complete measure, with $\mathcal{M} \subseteq \mathcal{N}$ and $n = m$ on $\mathcal{M}$, then $\overline{\mathcal{M}} \subseteq \mathcal{N}$ and $n = \mathfrak{m}$ on $\overline{\mathcal{M}}$.

(iii) $m = \mathfrak{m}$ if and only if $m$ is complete.

**16.** Show that if $m: \mathcal{M}^* \to E^*$ is induced by an $\mathcal{M}^*$-regular outer measure $\mu^*$, then $m$ equals its Lebesgue extension $\mu$ and completion $\mathfrak{m}$ (see Problem 15).

[Hint: By Definition 3 in §5, $m$ induces an outer measure $m^*$. By Theorem 3 in §5, $m^* A = \inf \{m X | A \subseteq X \subseteq \mathcal{M}^*\} = \mu^* A$

(for $\mu^*$ is $\mathcal{M}^*$-regular).

As $m^* = \mu^*$, we get $m^* = m$. Also, $m = \mathfrak{m}$, by Problem 15(iii).]

**17.** Prove that if a measure $\mu: \mathcal{M} \to E^*$ is $\sigma$-finite (Definition 4 in §5), with $S \in \mathcal{M}$, then its Lebesgue extension $m: \mathcal{M}^* \to E^*$ equals its completion $\mathfrak{m}$ (see Problem 15).

[Outline: It suffices to prove $\mathcal{M} \subseteq \overline{\mathcal{M}}$. (Why?)

To start with, let $A \in \mathcal{M}^*$, $m A < \infty$. By Problem 12 in §5,

$$\exists (B \in \mathcal{M}) \quad A \subseteq B \text{ and } m^* A = m A = m B < \infty;$$

so

$$m(B - A) = m B - m A = 0.$$  

Also,

$$\exists (B \in \mathcal{M}) \quad B - A \subseteq B \text{ and } m B = m(B - A) = 0.$$  

Thus $B - A$ is $\mu$-null; so $B - A \in \overline{\mathcal{M}}$. (Why?) Deduce that

$$A = B - (B - A) \in \overline{\mathcal{M}}.$$  

Thus $\overline{\mathcal{M}}$ contains any $A \in \mathcal{M}^*$ with $m A < \infty$. Use the $\sigma$-finiteness of $\mu$ to show

$$(\forall x \in \mathcal{M}^*) \exists \{A_n\} \subseteq \mathcal{M}^* \quad m A_n < \infty \text{ and } X = \bigcup A_n \in \overline{\mathcal{M}}.$$]


I. Our theory of set families leads quite naturally to a generalization of metric spaces. As we know, in any such space $(S, \rho)$, 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, $\emptyset \in \mathcal{G}$ and $S \in \mathcal{G}$.

(ii) $\mathcal{F}$ has these properties, with “unions” and “intersections” interchanged (Chapter 3, §12, Theorem 3). Moreover, by definition,

$$A \in \mathcal{F} \iff -A \in \mathcal{G}.$$  

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

$$f: (S, \rho) \to (T, \rho')$$

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,

$$\mathcal{G} = \mathcal{G}_\sigma.$$  

Also, $\mathcal{G} = \mathcal{G}_d$ and

$$\mathcal{F}_d = \mathcal{F} = \mathcal{F}_s,$$

but not

$$\mathcal{G} = \mathcal{G}_d \text{ or } \mathcal{F} = \mathcal{F}_s$$

in general.

$\mathcal{G}$ and $\mathcal{F}$ need not be rings or $\sigma$-rings (closure fails for differences). But by Theorem 2 in §3, $\mathcal{G}$ and $\mathcal{F}$ can be “embedded” in a smallest $\sigma$-ring. We name it in the following definition.

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

$$f: (S, \rho) \to (T, \rho')$$

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,

$$\mathcal{G} = \mathcal{G}_\sigma.$$  

Also, $\mathcal{G} = \mathcal{G}_d$ and

$$\mathcal{F}_d = \mathcal{F} = \mathcal{F}_s,$$

but not

$$\mathcal{G} = \mathcal{G}_d \text{ or } \mathcal{F} = \mathcal{F}_s$$

in general.

$\mathcal{G}$ and $\mathcal{F}$ need not be rings or $\sigma$-rings (closure fails for differences). But by Theorem 2 in §3, $\mathcal{G}$ and $\mathcal{F}$ can be “embedded” in a smallest $\sigma$-ring. We name it in the following definition.
Definition 2.
The $\sigma$-ring $\mathcal{B}$ generated by a topology $\mathcal{G}$ in $S$ is called the Borel field in $S$. (It is a $\sigma$-field, as $S \in \mathcal{G} \subseteq \mathcal{B}$.) Equivalently, $\mathcal{B}$ is the least $\sigma$-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 $\mathcal{B} \supseteq \mathcal{G}$ and $\mathcal{B} \supseteq \mathcal{F}$, but also $\mathcal{B} \supseteq \mathcal{G}_d, \mathcal{B} \supseteq \mathcal{F}_d, \mathcal{B} \supseteq \mathcal{G}_{ds}, \mathcal{B} \supseteq \mathcal{F}_{ds}$, etc.

Note that $\mathcal{G}_{ds} = \mathcal{G}_d$, $\mathcal{F}_{ds} = \mathcal{F}_d$, etc. (Why?)

II. Special notions apply to measures in metric and topological spaces.

Definition 3.
A measure $m : \mathcal{M} \to E^*$ 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 $\sigma$-ring), then also $\mathcal{B} \subseteq \mathcal{M}$ since $\mathcal{B}$ is, by definition, the least $\sigma$-ring $\supseteq \mathcal{G}$.

Thus $m$ is topological iff $\mathcal{B} \subseteq \mathcal{M}$ (hence surely $\mathcal{F} \subseteq \mathcal{M}, \mathcal{G}_d \subseteq \mathcal{M}, \mathcal{F}_d \subseteq \mathcal{M}$, 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 $m : \mathcal{M} \to E^*$ 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.,

$$\forall A \in \mathcal{M} \quad mA = \inf\{mX \mid A \subseteq X \in \mathcal{M} \cap \mathcal{G}\}.$$ 

If $m$ is topological ($\mathcal{G} \subseteq \mathcal{M}$), this simplifies to

$$mA = \inf\{mX \mid A \subseteq X \in \mathcal{G}\},$$

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

$$F \subseteq A \subseteq G, \text{ with } m(A - F) < \varepsilon \text{ and } m(G - A) < \varepsilon;$$

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 1. If a measure $m : \mathcal{M} \to E^*$ in $(S, \mathcal{G})$ is regular and $\sigma$-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,

$$mA = \inf\{mX \mid A \subseteq X \in \mathcal{M} \cap \mathcal{G}\};$$

so there is a set $X \in \mathcal{M} \cap \mathcal{G}$ (measurable and open), with

$$A \subseteq X \text{ and } mX < mA + \varepsilon.$$

Then

$$m(X - A) = mX - mA < \varepsilon,$$

and $X$ is the open set $G$ required in (2).

If, however, $mA = \infty$, use $\sigma$-finiteness to obtain

$$A \subseteq \bigcup_{k=1}^{\infty} X_k$$

for some sets $X_k \in \mathcal{M}, mX_k < \infty$; so

$$A = \bigcup_{k}(A \cap X_k).$$

Put

$$A_k = A \cap X_k \in \mathcal{M}.$$ (Why?) Then

$$A = \bigcup_{k} A_k,$$

and

$$mA_k \leq mX_k < \infty.$$ 

Now, by what was proved above, for each $A_k$ there is an open measurable $G_k \supseteq A_k$, with

$$m(G_k - A_k) < \frac{\varepsilon}{2^k}.$$ 

Set

$$G = \bigcup_{k=1}^{\infty} G_k.$$
Then \( G \in \mathcal{M} \cap \mathcal{G} \) and \( G \supseteq A \). Moreover,
\[
G - A = \bigcup_k G_k - \bigcup_k A_k \subseteq \bigcup_k (G_k - A_k).
\]
(Verify!) Thus by \( \sigma \text{-subadditivity} \),
\[
m(G - A) = \sum_k m(G_k - A_k) < \sum_k \frac{\varepsilon}{2^k} = \varepsilon,
\]
as required.

To find also the closed set \( F \), consider
\[
-\ A = S - A \in \mathcal{M}.
\]
As shown above, there is an open measurable set \( G' \supseteq -A \), with
\[
\varepsilon > m(G' - (-A)) = m(G' \cap A) = m(-(-G')).
\]
Then
\[
F = -G' \subseteq A
\]
is the desired closed set, with \( m(A - F) < \varepsilon \). \( \square \)

**Theorem 2.** If \( m : \mathcal{M} \to E^* \) 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
\[
(3) \quad H \subseteq A \subseteq K \quad \text{and} \quad m(A - H) = 0 = m(K - A);
\]
hence
\[
ma = mH = mK.
\]

**Proof.** Let \( A \in \mathcal{M} \). By strong regularity, given \( \varepsilon_n = 1/n \), one finds measurable sets
\[
G_n \in \mathcal{G} \quad \text{and} \quad F_n \in \mathcal{F}, \quad n = 1, 2, \ldots,
\]
such that
\[
F_n \subseteq A \subseteq G_n
\]
and
\[
(4) \quad m(A - F_n) < \frac{1}{n} \quad \text{and} \quad m(G_n - A) < \frac{1}{n}, \quad n = 1, 2, \ldots.
\]
Let
\[
H = \bigcup_{n=1}^\infty F_n \quad \text{and} \quad K = \bigcap_{n=1}^\infty G_n.
\]
Then \( H, K \in \mathcal{M}, H \in \mathcal{F}_\sigma, K \in \mathcal{G}_\delta \), and
\[
H \subseteq A \subseteq K.
\]

Also, \( F_n \subseteq H \) and \( G_n \subseteq K \).

Hence
\[
A - H \subseteq A - F_n \quad \text{and} \quad K - A \subseteq G_n - A;
\]
so by (4),
\[
m(A - H) < \frac{1}{n} \to 0 \quad \text{and} \quad m(K - A) < \frac{1}{n} \to 0.
\]
Finally,
\[
mA = m(A - H) + mH = mH,
\]
and similarly \( mA = mK \).

Thus all is proved. \( \square \)

**Problems on Topologies, Borel Sets, and Regular Measures**

1. Show that \( \mathcal{G} \) is a topology in \( S \) (in (a)–(c), describe \( \mathcal{B} \) also), given
   \( G = 2^S; \)
   \( G = \{ \emptyset, S \}; \)
   \( G = \{ \emptyset \ \text{and} \ \text{all sets in} \ \mathcal{S}, \ \text{containing a fixed point} \ p \}; \)
   \( G \) consists of all possible unions of sets of the form \( (a, b), \ (a, \infty), \ \text{and} \ [-\infty, b], \ \text{with} \ a, b \in E^* \).

2. \( (S, \rho) \) is called a **pseudometric space** (and \( \rho \) is a **pseudometric**) iff the metric laws (i)–(iii) of Chapter 3, \( \S 11 \) hold, but (i') is weakened to
   \[ \rho(x, x) = 0 \]
   (so that \( \rho(x, y) \) may be 0 even if \( x \neq y \)).
   \( (a) \) Define “globs,” “interiors,” “open sets” (i.e., \( \mathcal{G} \)) as in Chapter 3, \( \S 12 \); then show that \( \mathcal{G} \) is a topology for \( S \).
   \( (b) \) Let \( S = E^2 \) and
   \[ \rho(\bar{x}, \bar{y}) = ||x_1 - y_1||, \]
   where \( \bar{x} = (x_1, x_2) \) and \( \bar{y} = (y_1, y_2) \). Show that \( \rho \) is a pseudometric but not a metric (the Haussdorff properly fails!).

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).

4. In a topological space \( (S, \mathcal{G}) \), define
   \[
   \mathcal{G}^0 = \mathcal{G}, \ \mathcal{G}^1 = \mathcal{G}_\delta, \ \mathcal{G}^2 = \mathcal{G}_{\delta_\delta}, \ldots
   \]
and
\[ F^0 = \mathcal{F}, \quad F^1 = \mathcal{F}_a, \quad F^2 = \mathcal{F}_{abs}, \quad F^3 = \mathcal{F}_{alls}, \text{ etc.} \]

(Give an inductive definition.) Then prove by induction that
(a) \( G^n \subseteq \mathcal{B}, \quad F^n \subseteq \mathcal{B}; \)
(b) \( G^{n-1} \subseteq G^n, \quad F^{n-1} \subseteq F^n; \)
(c) \( \forall X \subseteq S \) \( X \in F^n \) if \( -X \in G^n; \)
(d) \( \forall X \in F^n \) \( X \cap Y \in F^n, \quad X \cup Y \in F^n; \) same for \( G^n; \)
(e) \( \forall X \in G^n \) \( \forall Y \in G^n \) \( X - Y \in G^n \) and \( Y - X \in F^n. \)

Hence, by Theorem 3 in Chapter 3, §16. Hence deduce that
\( (\forall F \in F) \ F \subseteq G_n \in \mathcal{G}. \)

Also, \( \bigcap_n G_n = \mathcal{F} = F \)

So
\( (\forall n) \ F \subseteq G_n \in \mathcal{G}. \)

Hence, by Problem 4(c), \( F \subseteq G_n \in \mathcal{G}. \)

5. For metric and pseudometric spaces (see Problem 2) prove that \( F^n \subseteq G^{n+1} \) and \( G^n \subseteq F^{n+1} \)

(cf. Problem 4).

6. If \( m \) is as in Definition 5, then prove the following.
(i) \( m \) is regular.
(ii) \( \forall \ A \in \mathcal{M} \) \( mA = \text{sup}\{mX \mid A \subseteq X \in \mathcal{M} \cap F\}. \)
(iii) The latter implies strong regularity if \( m < \infty \) and \( S \in \mathcal{M}. \)

7. Let \( \mu: \mathcal{B} \to E^* \) be a Borel measure in a metric space \((S, \rho)\). Set
\( (\forall A \subseteq S) \quad n^*A = \inf\{\mu X \mid A \subseteq X \in \mathcal{G}\}. \)

Prove that
(i) \( n^* \) is an outer measure in \( S; \)
(ii) \( n^* \mu \) on \( \mathcal{G}; \)
(iii) the \( n^* \)-induced measure, \( n: N^* \to E^* \), is topological (so \( \mathcal{B} \subseteq N^*));
(iv) \( n \geq \mu \) on \( B; \)
(v) \( \forall A \subseteq S \) \( (\exists \ H \in \mathcal{G}) \ A \subseteq H \text{ and } \mu H = n^*A. \)

8. From Problem 7 with \( m = \mu \), prove that if
\( A \subseteq G \subseteq \mathcal{G}, \)
with \( mG < \infty \) and \( A \in \mathcal{B} \), then \( mA = nA. \)

9. Let \( m, n, \text{ and } n^* \) be as in Problems 7 and 8. Suppose
\( S = \bigcup_{n=1}^{\infty} G_n, \)
with \( G_n \in \mathcal{G} \) and \( mG_n < \infty \) (this is called \( \sigma^n \)-finiteness).

Prove that
(i) \( m = n \) on \( \mathcal{B}, \)
(ii) \( m \) and \( n \) are strongly regular.
[Hints: Fix $A \in \mathcal{B}$. Show that
$$A = \bigcup A_n \text{ (disjoint)}$$
for some Borel sets $A_n \subseteq G_n$ (use Corollary 1 in §1). By Problem 8, $mA_n = nA_n$ since
$$A_n \subseteq G_n \in \mathcal{G}$$
and $mG_n < \infty$. Now use $\sigma$-additivity to find $mA = nA$.
(ii) Use $\mathcal{G}$-regularity, part (i), and Theorem 1.]

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 $\sigma^0$-finite Borel measure $m$ in $(S, \rho)$ 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),
$$(\exists H \in \mathcal{B}) A \subseteq H \text{ and } n^* A = mH = m^* H.$$ Also,
$$(\exists K \in \mathcal{B}) A \subseteq K \text{ and } m^* A = mK$$
(Problem 12 in §5). Deduce that
$$n^* A \leq n(H \cap K) = m(H \cap K) \leq mH = n^* A$$
and
$$n^* A = m(H \cap K) = m^* A.$$

§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 $v: \mathcal{C} \to E^1$ (“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 $\sigma$-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.

§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 $v: \mathcal{C} \to E^1$ (“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 $\sigma$-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 1. Lebesgue premeasure $v$ is $\sigma$-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:
$$v = m^* = m \text{ on } \mathcal{C}.$$

This follows by Corollary 1 in §2 and Theorem 2 of §6.

Note 1. As $\mathcal{M}^*$ is a $\sigma$-field (§6), it is closed under countable unions, countable intersections, and differences. Thus
$$\mathcal{C} \subseteq \mathcal{M}^* \text{ implies } \mathcal{C}_0 \subseteq \mathcal{M}^*;$$
i.e., any countable union of intervals is L-measurable. Also, $E^n \in \mathcal{M}^*$.

Corollary 1. Any countable set $A \subset E^n$ is L-measurable, with $mA = 0$.

The proof is as in Corollary 6 of §2.

Corollary 2. The Lebesgue measure of $E^n$ is $\infty$.

Prove as in Corollary 5 of §2.

Examples.
(a) Let
$$R = \{\text{rationals in } E^1\};$$
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 $E^3$ is its length, $b - a$. Let
$$R_o = \{\text{all rationals in } [a, b]\};$$
so $mR_o = 0$. As $[a, b]$ and $R_o$ are in $\mathcal{M}^*$ (a $\sigma$-field), so is
$$[a, b] - R_o,$$
the irrationals in $[a, b]$. By Lemma 1 in §4, if $b > a$, then
$$m([a, b] - R_o) = m([a, b]) - mR_o = m([a, b]) = b - a = 0 = mR_o.$$ 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 2. Lebesgue measure in $E^n$ is complete, topological, and totally $\sigma$-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, $\mathcal{M}^* \supseteq \mathcal{F}$, $\mathcal{M}^* \supseteq \mathcal{G}_0$, $\mathcal{M}^* \supseteq \mathcal{F}_\omega$, $\mathcal{M}^* \supseteq \mathcal{F}_{\omega_1}$, 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 \( \mathcal{G} \), hence in \( \mathcal{M}^* \) (Note 1). Thus \( \mathcal{M}^* \supseteq \mathcal{G} \). But by definition, the Borel field \( \mathcal{B} \) is the least \( \sigma \)-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,

\[
E^n = \bigcup_{k=1}^{\infty} A_k \text{ (disjoint)},
\]

with \( mA_k < \infty \) (Lemma 2 in §2). Hence

\[(\forall A \subseteq E^n) \quad A \subseteq \bigcup_{k} A_k; \]

so

\[A = \bigcup_{k}(A \cap A_k) \text{ (disjoint)}.\]

If, further, \( A \in \mathcal{M}^* \), then \( A \cap A_k \in \mathcal{M}^* \), and

\[m(A \cap A_k) \leq mA_k < \infty. \quad \text{(Why?) \quad \Box}\]

**Note 2.** More generally, a \( \sigma \)-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 3.**

(a) *Lebesgue outer measure* \( m^* \) in \( E^n \) is \( \mathcal{G} \)-regular; that is,

\[\forall A \subseteq E^n \quad m^* A = \inf \{mX \mid A \subseteq X \in \mathcal{G}\}\]

(\( \mathcal{G} \) is 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 basic 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

\[A \subseteq \bigcup_{k} B_k \text{ and } m^* A + \frac{1}{2} \varepsilon \geq \sum_{k} vB_k.\]

(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

\[vC_k - \frac{\varepsilon}{2k+1} < vB_k.\]

\[m^* A + \frac{1}{2} \varepsilon \geq \sum_k (vC_k - \frac{\varepsilon}{2k+1}) = \sum_k vC_k - \frac{1}{2} \varepsilon;\]

so by \( \sigma \)-subadditivity,

\[\sum_k mC_k \leq \sum_k vC_k \leq m^* A + \varepsilon.\]

Let

\[X = \bigcup_k C_k.\]

Then \( X \) is open (as the \( C_k \) are). Also, \( A \subseteq X \), and by (3),

\[mX \leq m^* A + \varepsilon.\]

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 \( \sigma \)-finite, and \( E^n \in \mathcal{M}^* \); so (b) follows by Theorem 1 in §7. \( \Box \)

**Definition.**

Given \( A \subseteq E^n \) and \( \bar{p} \in E^n \), let \( \bar{p} + A \) or \( A + \bar{p} \) denote the set of all points of the form

\[\bar{x} + \bar{p}, \quad \bar{x} \in A.\]

We call \( A + \bar{p} \) the translate of \( A \) by \( \bar{p} \).

**Theorem 4.** *Lebesgue outer measure* \( m^* \) and *Lebesgue measure* \( m \) in \( E^n \) are translation invariant. That is,

(i) \( \forall A \subseteq E^n \) \( \forall \bar{p} \in E^n \) \( m^* (A + \bar{p}) = m^* A \);

(ii) if \( A \) is L-measurable, so is \( A + \bar{p} \), and \( mA = mA + \bar{p} \).

See also Problem 7 in §10.

**Proof.** (i) If \( A \) is an interval with endpoints \( \bar{a} \) and \( \bar{b} \), then \( A + \bar{p} \) is the interval with endpoints \( \bar{a} + \bar{p} \) and \( \bar{b} + \bar{p} \). (Verify!)

Hence the edge lengths of \( A \) and \( A + \bar{p} \) are the same,

\[\ell_k = b_k - a_k = (b_k + p_k) - (a_k + p_k), \quad k = 1, 2, \ldots, n.\]

Thus

\[mA = vA = \prod_{k=1}^{n} \ell_k = mA + \bar{p};\]

so the theorem holds for intervals.
In the general case, $m*A$ is the glb of all basic covering values of $A$. But a basic covering consists of intervals that, when translated by $\bar{p}$, cover $A + \bar{p}$ and retain the same volumes, as was shown above.

Hence any covering value for $A$ is also one for $A + \bar{p}$, and conversely (since $A$, in turn, is a translate of $A + \bar{p}$ by $-\bar{p}$).

Thus the basic covering values of $A$ and of $A + \bar{p}$ are the same, with one and the same glb. Hence

$$m^*A = m^*(A + \bar{p}),$$

as claimed.

(ii) Now let $A \in M^*$. We must show that $A + \bar{p} \in M^*$, i.e., that

$$(\forall X \subseteq A + \bar{p}) (\forall Y \subseteq -(A + \bar{p})) \ m^*X + m^*Y = m^*(X \cup Y).$$

Thus fix $X \subseteq A + \bar{p}$ and $Y \subseteq -(A + \bar{p})$.

As is easily seen, $X - \bar{p} \subseteq A$ and $Y - \bar{p} \subseteq -A$ (translate all by $-\bar{p}$). Since $A \in M^*$, we get

$$m^*(X - \bar{p}) + m^*(Y - \bar{p}) = m^*((X \cup Y) - \bar{p}).$$

(Why?) But by (i), $m^*X = m^*(X - \bar{p})$, $m^*Y = m^*(Y - \bar{p})$, and

$$m^*(X \cup Y) = m^*((X \cup Y) - \bar{p}).$$

Hence

$$m^*X + m^*Y = m^*(X \cup Y),$$

and so $A + \bar{p} \in M^*$.

Now, as $m^* = m$ on $M^*$, (i) yields $mA = m(A + \bar{p})$, proving (ii) also. □

Problems on Lebesgue Measure

1. Fill in all details in the proof of Theorems 3 and 4.
2. Prove Note 2.
3. From Theorem 3 deduce that

$$(\forall A \subseteq E^n) (\exists B \in G^m) \ A \subseteq B \text{ and } m^*A = mB.$$ 

[Hint: See the hint to Problem 7 in §5.]
4. Review Problem 3 in §5.
5. Consider all translates

$$R + p \quad (p \in E^n)$$

of

$$R = \{\text{rationals in } E^1\}.$$ 

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 $y + p \in [0, 1]$]

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 $y \in (Q + r) \cap (Q + r')$.

Then

$$y = q + r = q' + r' \quad (q, q' \in Q, \ r, r' \in R);$$

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 $Q + r = Q + r'$.]

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).]

7. Show that if $f : S \to T$ 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

$$R = \{X \subseteq T \mid f^{-1}[X] \in \mathcal{B} \text{ in } S\}$$

is a $\sigma$-ring in $T$. As $\mathcal{B}$ is the least $\sigma$-ring $\supseteq \mathcal{G}, \mathcal{R} \supseteq \mathcal{B}$ (the Borel field in $T$).]

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 $a = (0, 0), b = (0, 1)$. Define $f : E^1 \to E^2$ by $f(x) = (0, x)$.

Show that $f$ is one-to-one and that $[a, b]$ is the $f$-image of $[0, 1]$. Use Problem 2 of Chapter 1, §9.]
9. Show that not all $L$-measurable sets are Borel sets in $E^n$.
[Hint for $E^2$: With $[a, b]$ and $f$ as in Problem 8, show that $f$ is continuous (use the sequential criterion). As $m[a, b] = 0$, all subsets of $[a, b]$ 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 $E^n$ ($n > 1$).]

10. Show that Cantor’s set $P$ (Problem 17 in Chapter 3, §14) has Lebesgue measure zero, even though it is uncountable.
[Outline: Let $U = [0, 1] - P$; so $U$ is the union of open intervals removed from $[0,1]$. Show that
$$mU = \frac{1}{2} \sum_{n=1}^{\infty} \left( \frac{2}{3} \right)^n = 1$$
and use Lemma 1 in §4.]

11. Let $\mu: \mathcal{B} \to E^*$ be the Borel restriction of Lebesgue measure $m$ in $E^n$ (§7). Prove that
(i) $\mu$ is incomplete;
(ii) $m$ is the Lebesgue extension (*and completion, as in Problem 15 of §6) of $\mu$.
[Hints: (i) By Problem 9, some $\mu$-null sets are not in $\mathcal{B}$. (ii) See the proof (end) of Theorem 2 in §9 (the next section).]

12. Prove the following.
(i) All intervals in $E^n$ are Borel sets.
(ii) The $\sigma$-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.]

13. Show that if a measure $m': \mathcal{M}' \to E^n$ agrees on intervals with Lebesgue measure $m: \mathcal{M}^* \to E^n$, then the following are true.
(i) $m' = m$ on $\mathcal{B}$, the Borel field in $E^n$.
(ii) If $m'$ is also complete, then $m' = m$ on $\mathcal{M}^*$.
[Hint: (i) Use Problem 13 of §5 and Problem 12 above.]

14. Show that globes of equal radius have the same Lebesgue measure.
[Hint: Use Theorem 4.]

15. Let $f: E^n \to E^n$, with
$$f(\bar{x}) = c\bar{x} \quad (0 < c < \infty).$$

Prove the following.
(i) $f(A \subseteq E^n)$ $m^*f[A] = c^n m^*A$ ($m^*$ is Lebesgue outer measure).
(ii) $A \in \mathcal{M}^*$ iff $f[A] \in \mathcal{M}^*$.
[Hint: If, say, $A = (a, b]$, then $f[A] = (ca, cb]$. (Why?) Proceed as in Theorem 4, using $f^{-1}$ also.]

16. From Problems 14 and 15 show that
(i) $mG(b) = c^n \cdot mG(a)$;
(ii) $mG(b) = \overline{mG}(a)$;
(iii) $mG(b) = a \cdot mI$, where $I$ is the cube inscribed in $G(b)$ and
$$a = \left(\frac{1}{2}\sqrt{n}\right)^n \cdot mG_0(1).$$
[Hints: (i) $f(G(a)) = G(b)$. (ii) Prove that $mG \leq \overline{mG} \leq c^n mG$ if $c > 1$. Let $c \to 1$.]

17. Given $a < b$ in $E^1$, let $\{r_n\}$ be the sequence of all rationals in $A = [a, b]$. Set
$$\delta_n = \frac{b - a}{2^n + 1}$$
and
$$G_n = (a_n, b_n) = (a, b) \cap \left( r_n - \frac{1}{2} \delta_n, r_n + \frac{1}{2} \delta_n \right).$$

Let
$$P = A - \bigcup_{n=1}^{\infty} G_n.$$

Prove the following.
(i) $\sum_{n=1}^{\infty} \delta_n = \frac{1}{2}(b - a) = \frac{1}{4}mA$.
(ii) $P$ is closed; $P^n = \emptyset$, yet $mP > 0$.
(iii) The $G_n$ can be made disjoint (see Problem 3 in §2), with $mP$ still > 0.
(iv) Construct such a $P \subseteq A$ ($P = \overline{P}$, $P^n = \emptyset$) of prescribed measure $mP = \varepsilon > 0$.

18. Find an open set $G \subset E^1$, with $mG < \overline{mG} < \infty$.
[Hint: $G = \bigcup_{n=1}^{\infty} G_n$ with $G_n$ as in Problem 17.]

19. If $A \subseteq E^n$ is open and convex, then $mA = \overline{mA}$.
[Hint: Let first $\emptyset \in A$. Argue as in Problem 16.]
§9. Lebesgue–Stieltjes Measures

Let
\[ \alpha : E^1 \to E^1 \]
be a nondecreasing function (\( \alpha^\uparrow \)). Consider the Lebesgue–Stieltjes set function \( s_\alpha \) (Example (d) in §4).

As we noted in Problem 7 of §4, \( s_\alpha \geq 0 \) when \( \alpha^\uparrow \); for then
\[ s_\alpha(a, b) = \alpha(b) - \alpha(a) + \alpha(a+) \geq 0. \]
Similarly for other intervals. Also, \( \emptyset \in \mathcal{C} \) and \( s_\alpha \emptyset = 0 \) by definition.

Thus \( s_\alpha \) is a premeasure on \( \mathcal{C} \) (finite intervals in \( E^1 \)), called the \( \alpha \)-induced Lebesgue–Stieltjes (LS) premeasure in \( E^1 \).

The outer measure \( m_\alpha^* \) induced by \( s_\alpha \) (§5) is called the \( \alpha \)-induced LS outer measure; its restriction to the family \( \mathcal{M}_\alpha^* \) of \( m_\alpha^* \)-measurable (or LS-measurable) sets is the \( \alpha \)-induced LS measure on \( E^1 \), denoted \( m_\alpha \).

Recall that, by our definitions, premeasures, outer measures, and measures are all nonnegative.

Note 1. No generality is lost by assuming that \( \alpha \) is right continuous (if not, replace it by the right-continuous function \( \beta^\uparrow \), with \( \beta(x) = \alpha(x+) \)). Similarly, one achieves left continuity by setting \( \beta(x) = \alpha(x-) \).

Note 2. If \( \alpha \) is right continuous, one often restricts \( s_\alpha \) to the family \( \mathcal{C}^* \) of all half-open intervals (for motivation, see Problem 7(iv) in §4). This does not affect \( m_\alpha^* \) or \( m_\alpha \) (Problem 3′ in §5), and simplifies the proof of additivity
\[ s_\alpha(a, b) + s_\alpha(b, c) = \alpha(b) - \alpha(a) + \alpha(c) - \alpha(b) = \alpha(c) - \alpha(a) = s_\alpha(a, c). \]

Recall that both \( \mathcal{C} \) and \( \mathcal{C}^* \) are semirings (Note 1 in §1).

**Theorem 1.** The LS premeasure \( s_\alpha \) is \( \sigma \)-additive on the semiring \( \mathcal{C} \) of all finite intervals in \( E^1 \).

Hence (by Theorem 2 in §6) all such intervals are LS-measurable (\( \mathcal{C} \subseteq \mathcal{M}_\alpha^* \)), and
\[ m_\alpha A = s_\alpha A \]
for any such interval \( A \).

**Proof.** As is easily seen, \( s_\alpha \) is additive (Problem 7 of §4).

It also satisfies Lemma 1 of §1 and Lemma 1 in §2 (Problem 7(v) in §4).

The proof of \( \sigma \)-additivity is then quite analogous to that of Theorem 1 of §2; we omit its repetition.

The rest is immediate by Theorem 2 of §6. □

Similarly, the proofs of Theorems 2 and 3 (but not 4) of §8 carry over to LS measures. Thus LS measures are complete, topological, totally \( \sigma \)-finite and strongly regular.

As in §8, it follows that singletons and countable sets are measurable, but their LS measure need not be 0 (Problem 8(iii) in §4).

Also, \( E^1 \in \mathcal{M}_\alpha^* \), but \( m_\alpha E^1 \) may be finite (Problem 8(ii)(ii′) in §4).

Since the proofs are the same as in §8, we omit them.

Note, however, the following facts.

(i) For singletons, \( m_\alpha \{ p \} = 0 \) iff \( \alpha \) is continuous at \( p \) (Problem 7(ii) in §4).

(ii) Hence
\[ m_\alpha[a, b] = m_\alpha(a, b) = m_\alpha(a, b) = m_\alpha(a, b) = \alpha(b) - \alpha(a) \]
if \( \alpha \) is continuous at \( a \) and \( b \) (Problem 7(iv) in §4).

(iii) LS measures need not be translation invariant (Problem 8(i) of §4).

(iv) If \( \alpha(x) = x \) on \( E^1 \), then \( m_\alpha^* = m^* \) (= Lebesgue outer measure in \( E^1 \)).

Thus Lebesgue measure is a special case of LS measure.

The latter is a kind of “weighted length.” Imagine that mass is distributed along the line, with \( \alpha(x) \) equal to the mass of
\[ (-\infty, x]. \]

For simplicity, assume that \( \alpha \) is right-continuous (cf. Notes 1 and 2). Then the mass of \( (a, b) \) is
\[ \alpha(b) - \alpha(a) \]
and \( p \) is a “point mass” iff \( m_\alpha \{ p \} > 0 \).

Our next theorem shows that LS measures practically exhaust all topological measures in \( E^1 \) of any importance. We shall use Notes 1 and 2 above.

**Theorem 2.** Let \( m : \mathcal{M} \to E^1 \) be a topological measure in \( E^1 \), finite on \( \mathcal{C}^* \) (half-open intervals). Then there is an LS measure \( m_\alpha \) such that
\[ m_\alpha = m \]
on the Borel field \( \mathcal{B} \) in \( E^1 \).

If \( m \) is also complete, then \( m_\alpha = m \) on all of \( \mathcal{M}_\alpha^* \).

**Proof.** Define \( \alpha \) as follows:
\[ \alpha(x) = \begin{cases} m(0, x) & \text{if } x \geq 0, \\ -m(x, 0) & \text{if } x < 0. \end{cases} \]
Clearly, \( \alpha^\uparrow \) on \( E^1 \). Also, the right continuity of \( m \) (Theorem 2 of §4) implies that of \( \alpha \). (Verify!)
Thus \( \alpha \) induces an LS measure \( m_\alpha \), with
\[
m_\alpha(a,b) = s_\alpha(a,b) = \alpha(b) - \alpha(a)
\]
(Problem 7(iv) in §4). We claim that \( m_\alpha = m \) on \( \mathcal{B} \).

First, consider any \( (a,b) \in \mathcal{C}^* \). If \( 0 \leq a \leq b \), then
\[
m(a,b) = m(0,b] - m(0,a] = \alpha(b) - \alpha(a) = m_\alpha(a,b).
\]
Similarly in the cases \( a < 0 \leq b \) and \( a \leq b < 0 \). Thus
\[
m_\alpha = m \text{ (finite) on } \mathcal{C}^*.
\]

By Problem 13 in §5,
\[
m_\alpha = m \text{ on } \mathcal{B},
\]
the \( \sigma \)-ring generated by \( \mathcal{C}^* \) (Problem 12 of §8). Thus \( m \) and \( m_\alpha \) have the same restriction to \( \mathcal{B} \) (call it \( \mu \)).

Now, by Note 3 in §6, \( \mu \) induces an outer measure \( \mu^* \).

As \( \mathcal{B} \supseteq \mathcal{C}^*_\mu \), both \( \mu^* \) and \( m_\mu^* \) are \( \mathcal{B} \)-regular, by Theorem 3 in §5. Thus
\[
(\forall A \subseteq E^1) \quad m_\mu^*(A) = \inf \{\mu X \mid A \subseteq X \subseteq B\} = \mu^* A,
\]
i.e., \( m_\mu^* = \mu^* \), and so \( m_\alpha \) is the restriction of both \( m_\mu^* \) and \( \mu^* \) to measurable sets. Hence \( m_\alpha \) is the Lebesgue extension of \( \mu \), by definition.

By Problem 17 in §6, \( m_\alpha = \overline{\mu} \) is the “least” complete extension of \( \mu \). Thus if \( m \) is complete, it is an extension of \( m_\alpha \); so \( m = m_\alpha \) on \( M_{\mu^*}^\alpha \), as claimed. \( \square \)

**Problems on Lebesgue–Stieltjes Measures**

1. Do Problems 7 and 8 in §4 and Problem 3’ in §5, if not done before.

2. Prove in detail Theorems 1 to 3 in §8 for LS measures and outer measures.

3. Do Problem 2 in §8 for LS-outter measures in \( E^1 \).

4. Prove that \( f : E^1 \to (S, \rho) \) is right (left) continuous at \( p \) iff
\[
\lim_{n \to \infty} f(x_n) = f(p) \text{ as } x_n \searrow (x_n \nearrow p).
\]
[Hint: Modify the proof of Theorem 1 in Chapter 4, §2.]

5. Fill in all proof details in Theorem 2.
[Hint: Use Problem 4.]

6. In Problem 8(iv) of §4, describe \( m_\mu^* \) and \( M_\mu^* \).

7. Show that if \( \alpha = c \) (constant) on an open interval \( I \subseteq E^1 \) then
\[
(\forall A \subseteq I) \quad m_\mu^*(A) = 0.
\]
Disprove it for nonopen intervals \( I \) (give a counterexample).

---

8. Let \( m' : \mathcal{M} \to E^* \) 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 \( m' : \mathcal{M}^* \to E^* \) 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 \( Q \subseteq [0,1] \) 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,
\[
m'(0,r] = cm(0,r)
\]
for rational \( r = \frac{a}{k}, \quad n,k \in \mathbb{N}\)
(first take \( r = \frac{1}{n} \), then \( r = \frac{1}{n} \)).

By right continuity (Theorem 2 in §4), prove it for real \( r > 0 \) (take rationals \( r_i \downarrow r \)).

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 \( m' = m_\mu \) (a translation-invariant LS measure). As \( m_\mu = cm \) on half-open intervals, Lemma 2 in §2 yields \( m_\mu = cm \) on \( \mathcal{G} \) (open sets). Use \( \mathcal{G} \)-regularity to prove \( m_\mu^* = cm^* \) and \( M_\mu^* = M^* \).]

*(9. LS measures in \( E^n \).)* Let
\[
\mathcal{C}^* = \{\text{half-open intervals in } E^n\}.
\]

For any map \( G : E^n \to E^1 \) and any \( \{a,b\} \in \mathcal{C}^* \), set
\[
\Delta_k G(a,b] = G(x_1, \ldots, x_{k-1},b, x_{k+1}, \ldots, x_n)
\]
\[
- G(x_1, \ldots, x_{k-1},a, x_{k+1}, \ldots, x_n), \quad 1 \leq k \leq n.
\]

Given \( \alpha : E^n \to E^1 \), set
\[
s_\alpha(a,b] = \Delta_1(\Delta_2(\cdots(\Delta_n \alpha(a,b]) \cdots)).
\]

For example, in \( E^2 \),
\[
s_\alpha(a,b] = \alpha(b_1, b_2) - \alpha(a_1, b_2) - \alpha(a_1, a_2) + \alpha(a, a_2).
\]

Show that \( s_\alpha \) is additive on \( \mathcal{C}^* \). Check that the order in which the \( \Delta_k \) are applied is immaterial. Set up a formula for \( s_\alpha \) in \( E^3 \).
[Hint: First take two disjoint intervals
\[
(a,q] \cup (p,b] = (a,b],
\]
as in Figure 2 in Chapter 3, §7. Then use induction, as in Problem 9 of Chapter 3, §7.]
*10. If $s_\alpha$ in Problem 9 is nonnegative, and $\alpha$ is right continuous in each variable $x_k$ separately, we call $\alpha$ a distribution function, and $s_\alpha$ is called the $\alpha$-induced LS premeasure in $\mathbb{E}_n^\alpha$; the LS outer measure $m^*_\alpha$ and measure $m_\alpha : \mathcal{M}_\alpha^* \to \mathbb{E}_n^*$

in $\mathbb{E}_n$ (obtained from $s_\alpha$ as shown in §§5 and 6) are said to be induced by $\alpha$.

For $s_\alpha$, $m^*_\alpha$, and $m_\alpha$ so defined, redo Problems 1–3 above.

*§10. Vitali Coverings

Lebesgue measure $m$ leads to an interesting analogue of the Heine–Borel theorem. Below, $m^*$ is Lebesgue outer measure in $\mathbb{E}_n$. We start with some definitions.

**Definition 1.**

A sequence $\{I_k\}$ of sets in a metric space $(S, \rho)$ converges to a point $p$ $(I_k \to p)$ if

$$p \in \bigcap_{k=1}^{\infty} I_k$$

and

$$\lim_{k \to \infty} d I_k = 0,$$

where $d I_k$ = diameter of $I_k$.

**Definition 2.**

A family $K$ of nonempty sets in $(S, \rho)$ is a Vitali covering (V-covering) of a set $A \subseteq (S, \rho)$ iff for each $p \in A$ there is a sequence $\{I_k\} \subseteq K$, with $I_k \to p$.

We then also say that $K$ covers $A$ in the Vitali sense (V-sense).

**Theorem 1** (Vitali). If a set $K$ of nondegenerate cubes (or globes) in $\mathbb{E}_n$ covers $A$ in the V-sense, then

$$m^*(A - \bigcup_k I_k) = 0$$

for some disjoint sequence $\{I_k\} \subseteq K$.

**Proof.** We give the proof for cubes (it is similar for globes).

First, suppose $A \subseteq \mathbb{I}^n$ for some open cube $\mathbb{I}^n$. Then $A$ is also covered in the $V$-sense by the subfamily $K^\alpha \subseteq K$ of those cubes that lie in $\mathbb{I}^*$. (Why?) We also assume that $A \not\subseteq \bigcup I_j$ for any disjoint finite sequence $\{I_j\} \subseteq K$ (otherwise, all is trivial). Finally, we assume that all cubes in $K$ are closed; for other kinds of cubes, the theorem then easily follows (see Problem 3 below).

We claim that

$$\forall \text{disjoint cubes } I_1, \ldots, I_h \in K^\alpha \quad (\exists I \in K^\alpha) \quad I \cap \bigcup_{j=1}^h I_j = \emptyset.$$

Indeed, as

$$A \subseteq \bigcup_{j=1}^h I_j,$$

there is some

$$\bar{p} \in A - \bigcup_{j=1}^h I_j.$$ 

By assumption, all $I_j$ are closed; so

$$\bar{p} = \bigcap_{j=1}^h I_j$$

is open. Hence there is a globe

$$G_{\bar{p}}(\delta) \subseteq \bigcup_{j=1}^h I_j.$$ 

As $K^\alpha$ is a $V$-covering, it contains a sequence $I_i \to \bar{p}$, $d I_i \to 0$; so there is $I_i = I_i^* \in K^\alpha$ with $\bar{p} \in I_i$ and $d I_i < \delta$. Therefore,

$$I \subseteq G_{\bar{p}}(\delta) \subseteq \bigcup_{j=1}^h I_j;$$

so

$$I \cap \bigcup_{j=1}^h I_j = \emptyset,$$

as claimed.

Now, using induction, suppose we have already fixed $k$ disjoint cubes $I_j$ in $K^\alpha$. By (1), there are cubes $I \in K^\alpha$ with

$$I \cap \bigcup_{j=1}^k I_j = \emptyset.$$
Let $\delta_k$ be the lub of their diameters. As all $I \in K^o$ lie in $I^o$, 
\[ \delta_k = \sup \left\{ dl \bigg| I \in K^o, I \subseteq \bigcup_{j=1}^{k} I_j \right\} \leq dI^o < \infty. \]
Hence by properties of the lub, we find a cube $I_{k+1} \in K^o$ such that 
\[ I_k \subseteq -I_{k+1}. \]
and $dI_{k+1} > \frac{1}{2} \delta_k$.
In this way, taking $k = 1, 2, \ldots$, we select a disjoint sequence $\{I_k\} \subseteq K^o$ with $dI_{k+1} > \frac{1}{2} \delta_k$ for all $k$. We shall show that 
\[ m^*(A - \bigcup_{k=1}^{\infty} I_k) = 0 \]
in four steps.

(I) Let $\ell_k$ be the edge length of $I_k$; so $dI_k = \ell_k \sqrt{n}$. (Why?) Enclose each $I_k$ in a cube $J_k$ with the same center and with edge length 
\[ (4n + 1) \ell_k. \]
Then 
\[ (\forall \bar{x} \in I_k) \ (\forall \bar{y} \notin J_k) \]
\[ \rho(\bar{x}, \bar{y}) > 2n \ell_k \geq 2 \ell_k \sqrt{n} = 2dI_k > \delta_{k-1}. \]
(See Figure 32, where $n = 2$.) Also, 
\[ mJ_k = (4n + 1)^n mI_k. \]

(II) As the $I_k$ lie in $I^o$, the $\sigma$-additivity of $m$ yields
\[ \sum_{k=1}^{\infty} mI_k = (4n + 1)^n \sum_{k=1}^{\infty} mI_k \]
\[ = (4n + 1)^n m \bigcup_{k=1}^{\infty} I_k \]
\[ \leq (4n + 1)^n mI^o < \infty. \]
Thus the series $\sum mJ_k$ converges; so its “remainder” tends to 0:
\[ \lim_{r \to \infty} \sum_{k=r}^{\infty} mJ_k = 0. \]

Also, $mJ_k \to 0$. But by definition, 
\[ \delta_k < 2dI_{k+1} < 2dJ_{k+1} = 2\sqrt{n}(mJ_{k+1})^{1/n} \quad (n \text{ fixed}). \]
Hence $\lim_{k \to \infty} \delta_k = 0$, too.

(III) Now, seeking a contradiction, suppose
\[ m^*(A - \bigcup_{k=1}^{\infty} I_k) > 0. \]
Then as 
\[ \lim_{r \to \infty} \sum_{k=r}^{\infty} mJ_k = 0, \]
there is $r$ such that 
\[ m \bigcup_{k=r}^{\infty} J_k \leq \sum_{k=r}^{\infty} mJ_k < m^*(A - \bigcup_{k=1}^{\infty} I_k). \]
Hence
\[ A - \bigcup_{k=1}^{\infty} I_k \not\subseteq \bigcup_{k=r}^{\infty} J_k. \]
(Why?) Thus there is 
\[ \bar{p} \in A - \bigcup_{k=1}^{\infty} I_k \]
not in 
\[ \bigcup_{k=r}^{\infty} J_k, \]
so that 
\[ (\forall k \geq r) \quad \bar{p} \neq J_k, \quad \bar{p} \in A, \quad \text{and} \quad \bar{p} \in -\bigcup_{k=1}^{k=r} I_k \subseteq -\bigcup_{k=r}^{\infty} I_k. \]
As 
\[ -\bigcup_{k=1}^{r} I_k \in \mathcal{G}, \]
we find (as before) a cube $K \in K^o$ such that $\bar{p} \in K$ and 
\[ K \cap \bigcup_{k=1}^{r} I_k = \emptyset. \]
Also, as \( \delta_k \to 0 \), we have \( \delta_k < dK \) for large \( k \). But by our choice of the \( \delta_k \), this implies
\[
K \cap \bigcup_{j=1}^{k} I_j \neq \emptyset
\]
for large \( k \) (why?), whereas
\[
K \cap \bigcup_{j=1}^{r} I_j = \emptyset,
\]
as shown above.

Thus there is a least \( k > r \), call it \( q \), such that
\[
K \cap I_q \neq \emptyset,
\]
and \( \delta_q < dK \leq \delta_{q-1} \).

By (3), \( \bar{p} \notin J_q \). As
\[
K \cap I_q \neq \emptyset,
\]
let \( \bar{x} \in K \cap I_q \). Since \( \bar{x}, \bar{p} \in K \),
\[
\rho(\bar{x}, \bar{p}) \leq dK < \delta_{q-1}.
\]
But as \( \bar{x} \in I_q \) and \( \bar{p} \notin J_q \), we have
\[
\rho(\bar{x}, \bar{p}) > \delta_{q-1}
\]
by (2).

This contradiction proves the theorem for bounded sets \( A \).

(IV) If \( A \) is not bounded, use Lemma 2 in §2 to find a sequence \( \{K_i\} \) of disjoint half-open intervals with
\[
\bigcup K_i = E^n \supseteq A.
\]
Let
\[
A_i = A \cap K_i^\circ,
\]
where \( K_i^\circ \) is the open interval with the same endpoints; so \( mK_i = mK_i^\circ \) and \( m(K_i - K_i^\circ) = 0 \).

Set
\[
Z = \bigcup_{i=1}^{\infty} (K_i - K_i^\circ);
\]
so \( mZ = 0 \) and
\[
\bigcup_{i=1}^{\infty} K_i^\circ = E^n - Z.
\]

*§10. Vitali Coverings

(Why?) As \( A_i = A \cap K_i^\circ \), we have

\[
\bigcup_{i=1}^{\infty} A_i = A \cap \bigcup_{i=1}^{\infty} K_i^\circ = A \cap (E^n - Z) = A - Z.
\]

Clearly, each \( A_i \) is covered in the \( V \)-sense by those \( K \)-cubes that lie in \( K_i^\circ \). Thus as shown above,
\[
(\forall i) \quad m^*\left(A - \bigcup_{j} I_{ij}\right) = 0
\]
for disjoint cubes \( I_{ij} \subseteq K_i^\circ \). That is,
\[
(\forall i) \quad \bigcup_{j} I_{ij} \cup Z_i \supseteq A_i,
\]
where
\[
Z_i = A_i - \bigcup_{j} I_{ij}
\]
and \( mZ_i = 0 \). Hence by (4),
\[
\bigcup_{i,j} I_{ij} \cup \bigcup_{i} Z_i \supseteq \bigcup_{i} A_i = A - Z,
\]
so that
\[
m^*\left(A - \bigcup_{i,j} I_{ij}\right) = 0.
\]

Rearranging the \( I_{ij} \) in a single sequence \( \{I_k\} \), we complete the proof. □

**Theorem 2.** If \( m^*A < \infty \) in Theorem 1, then for every \( \varepsilon > 0 \) there is a finite disjoint sequence \( \{I_k\} \subseteq K \) such that
\[
m^*(A - \bigcup_{k} I_k) < \varepsilon.
\]

**Proof.** Fix \( \varepsilon > 0 \). As \( m^*A < \infty \), the \( G \)-regularity of \( m^* \) (Theorem 3 of §8) yields an open \( G \supseteq A \) such that
\[
mG < m^*A + \varepsilon.
\]

Clearly, \( A \) is covered in the \( V \)-sense by the subfamily \( K^\circ \) of those \( K \)-sets that lie in \( G \). Thus by Theorem 1,
\[
m^*\left(A - \bigcup_{k} I_k\right) = 0
\]
for a \textit{disjoint} sequence \( \{I_k\} \subseteq \mathcal{K}^n \). Also,
\[
\bigcup I_k \subseteq G,
\]
and so
\[
\sum m I_k = m \bigcup I_k \leq m G < \infty.
\]
Thus \( \sum m I_k \) converges; so
\[
\sum_{k=r}^{\infty} m I_k < \varepsilon
\]
for large \( r \).

On the other hand,
\[
A - \bigcup_{k=1}^{r} I_k \subseteq \left( A - \bigcup_{k=1}^{\infty} I_k \right) \cup \bigcup_{k=r}^{\infty} I_k.
\]
Hence
\[
m^*(A - \bigcup_{k=1}^{r} I_k) \leq m^*(A - \bigcup_{k=1}^{\infty} I_k) + m^* \bigcup_{k=r}^{\infty} I_k \leq 0 + \sum_{k=r}^{\infty} m I_k < \varepsilon,
\]
as required. \( \square \)

As an application, we obtain the following important theorem.

**Theorem 3 (Lebesgue).** If \( f : E^1 \rightarrow E^1 \) is monotone, it is differentiable almost everywhere ("a.e."), \( i.e., \) on \( E^1 - Z \) for some \( Z \) of Lebesgue measure zero.

We sketch the proof in a few steps (lemmas). These lemmas anticipate a more general approach to be taken in §12, with the notation in the following definition.

**Definition 3.**

Let \( m = \text{Lebesgue measure} \) and
\[
\overline{E} = \{ \text{all cubes } I \subset E^n \text{ with } m I > 0 \}.
\]

Let
\[
s : \mathcal{M}' \rightarrow [0, \infty], \quad \mathcal{M}' \supseteq \overline{E},
\]
be another measure in \( E^n \), finite on \( \overline{E} \).

For any natural \( r > 0 \), and \( \bar{p} \in E^n \), we set
\[
g_r(\bar{p}) = \inf \left\{ \frac{s I}{m I} \mid \bar{p} \in I \in \overline{E}, \ dI < \frac{1}{r} \right\}
\]
and
\[
h_r(\bar{p}) = \sup \left\{ \frac{s I}{m I} \mid \bar{p} \in I \in \overline{E}, \ dI < \frac{1}{r} \right\};
\]
furthermore, we denote
\[
D_s(\bar{p}) = \sup_{r} g_r(\bar{p}) \quad \text{and} \quad \overline{D}_s(\bar{p}) = \inf_{r} h_r(\bar{p}).
\]
Clearly, \( \{g_r\} \uparrow, \{h_r\} \downarrow, \) and
\[
0 \leq D_s = \lim_{r \to \infty} g_r \leq \lim_{r \to \infty} h_r = \overline{D}_s
\]
at each \( \bar{p} \in E^n \). (Why?)

We also write \( J(D_s > i) \) for
\[
\{ \bar{x} \in J \mid D_s(\bar{x}) > i \},
\]
\( J(D_s = a) \) for
\[
\{ \bar{x} \in J \mid D_s(\bar{x}) = a \},
\]
etc.

**Lemma 1.** With the above notation, \( 0 \leq D_s \leq \overline{D}_s < \infty \) a.e. on \( E^n \).

**Proof Outline.** Fix any open set \( J \subset E^n \), with \( m J < \infty \) and \( s J < \infty \) (\( e.g., \) an open cube in \( \overline{E} \)).

For \( i = 1, 2, \ldots \) set
\[
A_i = J(D_s > i)
\]
and
\[
\mathcal{K}_i = \left\{ I \in \overline{E} \mid I \subseteq J, \quad \frac{s I}{m I} > i \right\}.
\]
Verify that \( \mathcal{K}_i \) is a \( V \)-covering of \( A_i \); so there is a \textit{disjoint} sequence \( \{I_k\} \subseteq \mathcal{K}_i \), with
\[
m^*(A_i - \bigcup I_k) = 0
\]
and
\[
\bigcup I_k \subseteq J.
\]
Hence (cf. Problem 3 below)
\[
m^* A_i \leq m \bigcup I_k = \sum m I_k \leq \frac{1}{i} \sum s I_k = \frac{1}{i} s \bigcup I_k \leq \frac{s J}{i}, \quad i = 1, 2, \ldots.
\]
It follows that
\[
m^* \bigcap_{i=1}^{\infty} A_i = 0.
\]
(Why?) But

\[ \bigcap_{i=1}^{\infty} A_i = J(\overline{D}_s = \infty). \]

(Why?) This implies that

\[ m^* J(\overline{D}_s = \infty) = 0, \]

and so \( \overline{D}_s < \infty \) on \( J \), except for a null set.

But by Lemma 2 in §2, all of \( E^n \) is a countable union of such sets \( J \) (open cubes). Thus \( \overline{D}_s < \infty \) on \( E^n - Z \), where \( Z \) is a countable union of null sets: \( mZ = 0 \).

As \( 0 \leq D_s \leq \overline{D}_s \) on all of \( E^n \), we have

\[ 0 \leq D_s \leq \overline{D}_s < \infty \quad \text{a.e. on } E^n, \]

as claimed. \( \Box \)

**Lemma 2.** With the same notation, \( D_s = \overline{D}_s \) a.e. on \( E^n \).

**Proof Outline.** With \( J \) as in the previous proof, let

\[ H = J(\overline{D}_s > D_s). \]

Then \( H \) is a countable union of sets

\[ H_{uv} = J(\overline{D}_s > v > u > D_s) \]

over rational \( u, v \). Thus it suffices to show that all such \( H_{uv} \) are \( m \)-null.

Let \( Q \) be one of them; so \( Q \subseteq J \) and

\[ m^* Q \leq mJ < \infty. \]

Hence given \( \varepsilon > 0 \), there is an open set \( G \subseteq J \) with \( G \supseteq Q \) and

\[ mG < m^* Q + \varepsilon. \]

(Why?) We fix this \( G \) and set

\[ K = \left\{ I \in \overline{K} \mid I \subseteq G, \frac{sI}{mG} < u \right\}. \]

By the definition of \( D_s \), \( K \) is a \( V \)-covering of \( Q \) (verify!); so by Problem 3,

\[ m^* \left( Q \cap \bigcup I_k \right) = m^* Q \]

for a disjoint sequence

\[ \{ I_k \} \subseteq K, \quad \bigcup I_k \subseteq G. \]

\[ G' = \bigcup_{k=1}^{\infty} I_k^o \]

(an open set), and \( Q_o = Q \cap G' \); so

\[ m^* Q = m^* Q_o \leq mG^* \leq mG < m^* Q + \varepsilon. \]

(Explain!)

Next, let

\[ K' = \left\{ I \in K \mid I \subseteq G', \frac{sI}{mI} > v \right\}. \]

It is a \( V \)-covering of \( Q_o \) (why?); so

\[ m\left( Q_o - \bigcup I_k \right) = 0 \]

for a disjoint sequence \( \{ I_k \} \subseteq K' \). Verify that

\[ u \cdot (m^* Q + \varepsilon) > u \cdot mG^* = u \cdot \sum mI_k^o \]

\[ \geq \sum sI_k^o = sG^* \]

\[ \geq \sum sI_k' \]

\[ \geq v \cdot \sum mI_k' = v \cdot m \bigcup I_k' \]

\[ \geq v \cdot m^* \left( Q_o \cap \bigcup I_k' \right) = v \cdot m^* Q_o = v \cdot m^* Q. \]

Thus

\[ (\forall \varepsilon > 0) \quad u \cdot (m^* Q + \varepsilon) \geq v \cdot m^* Q. \]

Making \( \varepsilon \to 0 \), we get

\[ u \cdot m^* Q \geq v \cdot m^* Q. \]

As \( u < v \), \( m^* A \) must be 0. This is the desired result. \( \Box \)

**Proof of Theorem 3.** To fix ideas, let \( f^\uparrow \).

Let \( s = m_f \) be the \( f \)-induced LS measure in \( E^1 \) (§9) so that

\[ s[p, x] = f(x+) - f(p-). \]

By Lemmas 1 and 2, it suffices to show that \( f \) is differentiable at every \( p \in E^1 \), with

\[ D_s(p) = \overline{D}_s(p) \neq \infty. \]

Fix any such \( p \) and set

\[ q = D_s(p) = \overline{D}_s(p) \neq \infty. \]
Then \( f \) is continuous at \( p \); for otherwise,
\[
f(p+) - f(p-) > 0,
\]
whence
\[
\mathcal{D}s(p) = \infty.
\]
(Why?) Also, by Definition 3, given \( \varepsilon > 0 \), there is a natural \( r \) such that
\[
q - \varepsilon < g_r(p) \leq h_r(p) < q + \varepsilon.
\]
Let
\[
x \in G_{-p}(\frac{1}{r}).
\]
If \( x > p \), then
\[
\Delta x = x - p = m[p, x],
\]
and by continuity,
\[
\Delta f = f(x) - f(p) \leq f(x+) - f(p) = f(x+) - f(p-) = s[p, x]
\leq \Delta x \cdot h_r(p) < \Delta x(q + \varepsilon).
\]
Also, if \( x > y > p \), then
\[
\Delta f \geq f(y+) - f(p-) = s[p, y] \geq \Delta y \cdot g_r(p) > \Delta y(q - \varepsilon),
\]
where
\[
\Delta y = y - p = m[p, y].
\]
Making \( y \nearrow x \), with \( x \) fixed, we get
\[
(q - \varepsilon)\Delta x \leq \Delta f < (q + \varepsilon)\Delta x.
\]
Similarly in the case \( x < p \).
Thus with \( \varepsilon \to 0 \), we obtain
\[
f'(p) = \lim_{x \to p} \frac{\Delta f}{\Delta x} = q \neq \infty. \quad \Box
\]

**Problems on Vitali Coverings**

1. Prove Theorem 1 for *globes*, filling in all details.
   [Hint: Use Problem 16 in §8.]

2. Show that any (even uncountable) union of globes or nondegenerate cubes \( J_i \subset E^n \) is \( L \)-measurable.
   [Hint: Include in \( K \) each globe (cube) that lies in some \( J_i \). Then Theorem 1 represents \( \bigcup J_i \) as a countable union plus a null set.]

3. Supplement Theorem 1 by proving that
\[
m^* \left( A - \bigcup I_k^n \right) = 0
\]
and
\[
m^* A = m^* \left( A \cap \bigcup I_k^n \right);
\]
here \( I^o \) = interior of \( I \).

4. Fill in all proof details in Lemmas 1 and 2. Do it also for \( \overline{K} = \{ \text{globes} \} \).

5. Given \( mZ = 0 \) and \( \varepsilon > 0 \), prove that there are open globes
\[
G_k^* \subset E^n,
\]
with
\[
Z \subset \bigcup_{k=1}^{\infty} G_k^*
\]
and
\[
\sum_{k=1}^{\infty} mG_k^* < \varepsilon.
\]
   [Hint: Use Problem 3(i) in §5 and Problem 16(iii) from §8.]

6. Do Problem 3 in §5 for
   (i) \( C^r \) = \{ open globes \}, and
   (ii) \( C^r = \{ \text{all globes in } E^n \} \).
   [Hints for (i): Let \( m' = \) outer measure induced by \( \nu': C^r \to E^1 \). From Problem 3(e) in §5, show that
   \( \forall A \subseteq E^n \) \( m'A \geq m^* A \).
   To prove \( m'A \leq m^* A \) also, fix \( \varepsilon > 0 \) and an open set \( G \supset A \) with
   \( m^* A + \varepsilon \geq mG \) (Theorem 3 of §8).
   Globes inside \( G \) cover \( A \) in the \( V \)-sense (why?); so
   \[
   A \subseteq Z \cup \bigcup G_k \quad \text{(disjoint)}
   \]
   for some globes \( G_k \) and null set \( Z \). With \( G_k^* \) as in Problem 5,
   \[m'A \leq \sum (mG_k + mG_k^*) \leq mG + \varepsilon \leq m^* A + 2\varepsilon.\]

7. Suppose \( f: E^n \to E^n \) is an isometry, i.e., satisfies
\[
|f(x) - f(y)| = |x - y| \quad \text{for } \bar{x}, \bar{y} \in E^n.
\]
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.]

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 \( f(\bar{p}) = \bar{p} \).

8. A \( V \)-covering \( K \) of \( A \subseteq E^n \) is called normal iff
   (i) \((\forall I \in K) 0 < mI = mI^o < \infty \), and
   (ii) for every \( \bar{p} \in A \), there is some \( c \in (0, \infty) \) and a sequence
   \[ I_k \to \bar{p} \quad ([I_k] \subseteq K) \]
   such that
   \[ (\forall k) (\exists J_k \supseteq I_k) \quad c \cdot m^* I_k \geq mJ_k. \]

(We then say that \( \bar{p} \) and \( [I_k] \) are normal; specifically, \( c \)-normal.)

Prove Theorems 1 and 2 for any normal \( K \).

[Hints: By Problem 21 of Chapter 3, \( dl = d\mathcal{T} \).
   First, suppose \( K \) is uniformly normal, i.e., all \( \bar{p} \in A \) are \( c \)-normal for the same \( c \).
   In the general case, let
   \[ A_i = \{ \bar{x} \in A \mid \bar{x} - \bar{p} = i \cdot \bar{n} \} \]
   so \( K \) is uniform for \( A_i \). Verify that \( A_i \nsubseteq A \).
   Then select, step by step, a disjoint sequence \( [I_k] \subseteq K \) and naturals \( n_1 < n_2 < \cdots < n_i < \cdots \) such that
   \[ (\forall i) \quad m^* \left( A_i - \bigcup_{k=1}^{n_i} I_k \right) < \frac{1}{i}. \]

Let
   \[ U = \bigcup_{k=1}^{\infty} I_k. \]

Then
   \[ (\forall i) \quad m^* (A_i - U) < \frac{1}{i}, \]

and
   \[ A_i - U \nsubseteq A - U. \]
   (Why?) Thus by Problems 7 and 8 in §6,
   \[ m^*(A - U) \leq \lim_{i \to \infty} \frac{1}{i} = 0. \]

9. A \( V \)-covering \( \mathcal{K} \) of \( E^n \) is called universal iff
   (i) \((\forall I \in \mathcal{K}) 0 < mI = mI^o \to \infty \), and
   (ii) whenever a subfamily \( \mathcal{K} \subseteq \mathcal{K} \) covers a set \( A \subseteq E^n \) in the \( V \)-sense,
   we have
   \[ m^* \left( A - \bigcup I_k \right) = 0 \]
   for a disjoint sequence
   \[ \{I_k\} \subseteq \mathcal{K}. \]

Show the following.
   (a) \( \mathcal{K} \subseteq \mathcal{M}^* \).
   (b) Lemmas 1 and 2 are true with \( \mathcal{K} \) replaced by any universal \( \mathcal{K}' \). (In this case, write \( D^* s \) and \( D^s \) for the analogues of \( Ds \) and \( Ds \).)
   (c) \( D_k = D^* s = D^s = D s \) a.e.

[Hints: (a) By (i), \( I = T \) minus a null set \( Z \subseteq T - I^o \).
   (c) Argue as in Lemma 2, but set
   \[ Q = J \left( \bigcup s > u > v > D_k \right) \]
   and
   \[ \mathcal{K}' = \left\{ I \in \mathcal{K} \mid I \subseteq G', \frac{1}{mJ} > v \right\} \]
   to prove a.e. that \( D^* s \leq D_k \); similarly for \( Ds \leq D^k \).
   Throughout assume that \( s: \mathcal{M}^* \to E^n \) (\( \mathcal{M}^* \subseteq K \cap K' \) is a measure in \( E^n \), finite on \( K \cup K' \)).

10. Continuing Problems 8 and 9, verify that
   (a) \( \mathcal{K} = \{ \text{nondegenerate cubes} \} \) is a normal and universal \( V \)-covering of \( E^n \);
   (b) so also is \( \mathcal{K}' = \{ \text{all globes in} \; E^n \} \);
   (c) \( \mathcal{U} = \{ \text{nondegenerate intervals} \} \) is normal.
   Note that \( \mathcal{U} \) is not universal.\( ^1 \)

11. Continuing Definition 3, we call \( q \) a derive of \( s \), and write \( q \sim D s(\bar{p}) \), iff
   \[ q = \lim_{k \to \infty} \frac{sI_k}{mI_k} \]
   for some sequence \( I_k \to \bar{p} \), with \( I_k \in \mathcal{K} \).

   Set
   \[ D_{\bar{p}} = \{ q \in E^n \mid q \sim D s(\bar{p}) \} \]
   and prove that
   \[ D s(\bar{p}) = \min D_{\bar{p}} \text{ and } D\bar{s}(\bar{p}) = \max D_{\bar{p}}. \]

\( ^1 \) See M. E. Munroe, Measure and Integration, Addison–Wesley (1971), pp. 173–175.
12. Let $\mathcal{K}^*$ be a normal $V$-covering of $E^n$ (see Problem 8). Given a measure $s$ in $E^n$, finite on $K^* \cup \overline{\mathbb{R}}$, write
\[
q \sim D^* s(\bar{p})
\]
iff
\[
q = \lim_{k \to \infty} \frac{sl_k}{mI_k}
\]
for some normal sequence $I_k \to \bar{p}$, with $I_k \in \mathcal{K}^*$.

Set
\[
D^*_p = \{q \in E^* | q \sim D^* s(\bar{p})\},
\]
and then
\[
D^* s(\bar{p}) = \inf D^*_p \quad \text{and} \quad \overline{D}^* s(\bar{p}) = \sup D^*_p.
\]

Prove that
\[
\overline{D}^* s = D^* s = I\overline{D}s \text{ a.e. on } E^n.
\]

[Hint: $E^n = \bigcup_{i=1}^{\infty} E_i$, where $E_i = \{x \in E^n | x \text{ is } i\text{-normal}\}$.

On each $E_i$, $\mathcal{K}^*$ is uniformly normal. To prove $\overline{D}^* s = D^* s$ a.e. on $E_i$, “imitate” Problem 9(c). Proceed.]

*§11. Generalized Measures. Absolute Continuity

I. We now return to general set functions $s: \mathcal{M} \to E$, with $E$ as in Definition 1 of §4.

Definition 1.

A set function $s: \mathcal{M} \to E$ is a generalized measure in a set $S$, and $(S, \mathcal{M}, s)$ is a generalized measure space, iff $s$ is $\sigma$-additive and semi-finite (i.e., $s \neq +\infty$ or $s \neq -\infty$) on $\mathcal{M}$, a $\sigma$-ring in $S$, and $s\emptyset = 0$.

We call $s$ a signed measure iff $E \subseteq E^*$ (i.e., $s$ is real or extended real); if $s \geq 0$ then $s$ is a measure; $s$ may also be complex ($E = C$) or vector valued.

Definition 2.

Given a set function $s: \mathcal{M} \to E$, we define its total variation
\[
v_s: \mathcal{M} \to [0, \infty]
\]
by
\[
(\forall A \in \mathcal{M}) \quad v_s A = \sup \sum_i |sX_i|,
\]
taking the sup over all countable disjoint subfamilies $\{X_i\} \subseteq \mathcal{M}$ with $\bigcup_i X_i \subseteq A$.

Note 1. If $\mathcal{M}$ is a $\sigma$-ring, we may equivalently require that
\[
\bigcup X_i = A
\]
with $\{X_i\}$ a disjoint sequence in $\mathcal{M}$ (add the term $X_0 = A - \bigcup X_i$ if necessary).\(^1\)

Corollary 1. If $s$ and $v_s$ are as in Definition 2, then
(i) $v_s$ is monotone on $\mathcal{M}$, and
(ii) $|sA| \leq v_s A$ for every $A \in \mathcal{M}$.

Proof. For (i), let $A \subseteq B$, $A, B \in \mathcal{M}$. Take any disjoint sequence $\{X_i\} \subseteq \mathcal{M}$, with
\[
\bigcup X_i \subseteq A \subseteq B.
\]

By definition,
\[
\sum_i |sX_i| \leq v_s B.
\]

Thus $v_s B$ is an upper bound of all such sums, with $\bigcup X_i \subseteq A$. Hence
\[
v_s A = \text{lub} \sum_i |sX_i| \leq v_s B,
\]
proving (i).

To prove (ii), just let $\{X_i\}$ consist of $A$ alone. \(\Box\)

Theorem 1. If $s: \mathcal{M} \to E$ is a generalized measure, then $v_s$ is a measure on $\mathcal{M}$.

Proof. By definition, $v_s \geq 0$ on $\mathcal{M}$, a $\sigma$-ring, and $v_s \emptyset = 0$. (Why?) It remains to prove $\sigma$-additivity.

Thus let
\[
A = \bigcup_{i=1}^{\infty} A_n \quad (\text{disjoint}),
\]
with $A, A_n \in \mathcal{M}$. To show that
\[
v_s A = \sum_{i=1}^{\infty} v_s A_n,
\]
take any $\mathcal{M}$-partition $\{X_i\}$ of $A$. Then
\[
(\forall i) \quad X_i = X_i \cap A = X_i \cap \bigcup_{i=1}^{\infty} A_n = \bigcup_{i=1}^{\infty} (X_i \cap A_n) \quad (\text{disjoint}).
\]

\(^1\) Any such $\{X_i\}$ is called an $\mathcal{M}$-partition of $A$ (Chapter 8, §11); it may consist of $A$ alone.
Similarly,

\[(\forall n) \quad A_n = \bigcup_i (A_n \cap X_i);\]

so by definition,

\[(\forall n) \quad \sum_i |s(A_n \cap X_i)| \leq v_A A_n.\]

Hence as

\[X_i = \bigcup_n (X_i \cap A_n),\]

we get

\[\sum_i |sX_i| = |s\bigcup_i (A_n \cap X_i)| = \sum_i |s(A_n \cap X_i)| \leq \sum_i |s(A_n \cap X_i)| \leq \sum_i |s(A_n \cap X_i)| \leq \sum_v A_n.\]

As \(\{X_i\}\) was an arbitrary \(M\)-partition of \(A\),

\[v_A A = \sup_n \sum_i |sX_i| \leq \sum v_A A_n.\]

It remains to show that

\[\sum v_A A_n \leq v_A A.\]

This is trivial if \(v_A A = \infty\).

Thus let \(v_A A < \infty\). Then

\[(\forall n) \quad v_A A_n \leq v_A A < \infty\]

by Corollary 1(i). Now fix \(\varepsilon > 0\). By properties of lub, each \(A_n\) has an \(M\)-partition,

\[A_n = \bigcup_k X_{nk},\]

such that

\[v_A A_n - \frac{\varepsilon}{2^n} < \sum_k |sX_{nk}|.\]

All \(X_{nk}\) combined (for all \(n\) and \(k\)) form an \(M\)-partition of \(A\). Thus by definition,

\[v_A A \geq \sum_n k |sX_{nk}| \geq \sum v_A A_n - \frac{\varepsilon}{2^n} \geq \sum v_A A_n - \varepsilon.\]

With \(\varepsilon \to 0\), we get

\[\sum v_A A_n \leq v_A A,\]

as required. \(\Box\)

**Definition 3.**

Given

\[s: \mathcal{M} \to E \text{ and } t: \mathcal{M}' \to E',\]

we say that \(s\) is

(i) \(t\)-continuous (written \(s \ll t\)) iff

\[v_t X = 0 \implies |sX| = 0 \quad (X \in \mathcal{M}');\]

(ii) absolutely \(t\)-continuous (or absolutely continuous with respect to \(t\)) iff

\[v_t X \to 0 \implies sX \to 0,\]

i.e.,

\[(\forall \varepsilon > 0) \ (\exists \delta > 0) \ (\forall X \in \mathcal{M}') \quad v_t X < \delta \implies |sX| < \varepsilon;\]

(iii) \(t\)-finite iff

\[v_t X < \infty \implies |sX| < \infty \quad (X \in \mathcal{M}').\]

**Corollary 2.** If two set functions \(s, u: \mathcal{M} \to E\) are \(t\)-continuous (absolutely \(t\)-continuous) so are \(s \pm u\), and so is \(ks\) for any \(k\) from the scalar field of \(E\).\(^3\)

The proof is left to the reader. (Use Definition 3(i)(ii), quantified formula.)

**Theorem 2.** Let \(s: \mathcal{M} \to E \text{ and } t: \mathcal{M}' \to E',\)

(i) If \(s \ll t\), then \(v_s \ll t\).

(ii) If, in addition, \(s\) and \(t\) are generalized measures and \(v_s\) is \(t\)-finite, then both \(v_s\) and \(s\) are absolutely \(t\)-continuous.

(iii) \(v_s \ll t\) implies \(s \ll t\) (which is obvious).

**Proof.** Fix \(A \in \mathcal{M}\) and any disjoint sequence \(X_i \in \mathcal{M}\), with

\[\bigcup X_i \subseteq A.\]

If \(v_A A = 0\), then (Corollary 1)

\[(\forall i) \quad v_A X_i = 0;\]

\(^2\) For the rest of this section, we assume that \(\mathcal{M}\) and \(\mathcal{M}'\) satisfy \(X \in \mathcal{M}\) whenever \(X \in \mathcal{M}'\) and \(v_t X < \infty\).

\(^3\) If \(E = E^*\), we assume \(k \in E^1\). If \(s\) is scalar valued, \(k\) may be a vector in \(E\).
so by the $t$-continuity of $s$, $|sX_t| = 0$, and hence $\sum |sX_t| = 0$. As this holds for any such sum, we also have

$$v_t A = \sup \sum |sX_t| = 0$$

whenever $v_t A = 0$. This proves assertion (i).

Now, let $s$ and $t$ be as in (ii); so $v_s$ and $v_t$ are measures by Theorem 1. Suppose $v_t$ is not absolutely $t$-continuous. Then

$$(\exists \varepsilon > 0) (\forall \delta > 0) (\exists X \in M') \ v_t X < \delta \text{ and } v_s X \geq \varepsilon.$$  

(Why?) Taking $\delta_n = 2^{-n}$, fix $(\forall n)$ a set $X_n \in M'$, with $v_t X_n < 2^{-n}$ and $v_s X_n \geq \varepsilon$.

Let

$$Y_n = \bigcup_{k=n}^\infty X_k \text{ and } Y = \bigcap_{n=1}^\infty Y_n;$$

so $Y, Y_n \in M'$, $Y_n \uparrow Y$, and

$$v_t Y_n \leq \sum_{k=n}^\infty v_t X_k < \sum_{k=n}^\infty 2^{-k} \leq 2^{1-n}.$$  

Thus by Theorem 2 in §4 (right continuity),

$$v_t Y = \lim_{n \to \infty} v_t Y_n \leq \lim_{n \to \infty} 2^{1-n} = 0.$$  

Hence by the $t$-continuity of $v_t$ (see (i)),

$$v_t Y = 0 < \varepsilon.$$  

On the other hand, as $Y_n \supseteq X_n$, we have

$$v_t Y_n \geq v_s X_n \geq \varepsilon.$$  

Also, $v_t Y_n \leq 2^{1-n}$ implies $v_t Y_n < \infty$ ($v_t$ is $t$-finite). Hence

$$v_s Y = \lim_{n \to \infty} v_s Y_n \geq \varepsilon,$$

a contradiction. Thus $v_s$ is absolutely $t$-continuous.

So is $s$; for by Corollary 1(ii), we have

$$(\forall \varepsilon > 0) (\exists \delta > 0) (\forall X \in M') \ v_t X < \delta \implies |sX| \leq v_s X < \varepsilon,$$

proving (ii). \[\square\]
Again, if  
\[ A - \bigcup_{i=1}^{2} A_i \]
is positive, put  
\[ Q = A - \bigcup_{i=1}^{2} A_i. \]
If not, let  \( n_3 \) be the least natural for which there is  \( A_3 \in M \), with  
\[ A_3 \subseteq A - \bigcup_{i=1}^{2} A_i \]
and  
\[ sA_3 < -\frac{1}{n_3}. \]
Continuing, we either find the desired  \( Q \) at some step or obtain a sequence  \( \{A_k\} \subseteq M \) such that
\[
(1) \quad (\forall k \in N) \quad sA_k < -\frac{1}{n_k} \quad \text{and} \quad A_{k+1} \subseteq A - \bigcup_{i=1}^{k} A_i
\]
(so the  \( A_k \) are disjoint). In the latter case, let  
\[ Q = \bigcup_{k=1}^{\infty} A_k; \]
so  
\[ A = Q \cup \bigcup_{k=1}^{\infty} A_k \quad \text{(disjoint)}, \]
and  
\[ sQ + \sum_{k} sA_k = sA. \]
As  \(|sA| < \infty\) (by assumption),  \( \sum sA_k \) converges. By (1), then,  
\[ \frac{1}{n_k} \leq \sum_{k} (-sA_k) < \infty. \]
Therefore,  
\[ \lim_{k \to \infty} \frac{1}{n_k} = 0, \]
i.e.,  
\[ \lim_{k \to \infty} n_k = \infty. \]
By additivity,
\[s(Y \cup P) = sY + sP > sP,\]
contrary to the maximality of \(sP\). This contradiction settles the case \(s \neq +\infty\).

In case \(s \neq -\infty\), consider \(-s\), which by Lemma 2 is likewise a signed measure, with \(-s \neq +\infty\). By what was proved above, there is a set \(P' \in \mathcal{M}\) that is positive for \(-s\) (hence negative for \(s\)), and whose complement is positive for \(s\).

Finally, if \(S \in \mathcal{M}\), then \(P \in \mathcal{M}\) implies
\[S - P = -P \in \mathcal{M};\]
so both \(P\) and \(-P\) are in \(\mathcal{M}\). Thus all is proved. □

**Note 4.** The set \(P\) in Theorem 3 is not unique. However, if \(P' \in \mathcal{M}^+\) is another such set, then
\[s(P - P') = 0 = s(P' - P),\]
i.e., any two such sets can differ by a set of measure \(0\) only. Indeed,
\[P - P' \subseteq P\]
and \(P - P' \subseteq -P';\)
so \(s(P - P')\) is both \(\geq 0\) and \(\leq 0\). Thus \(s(P - P') = 0\). Similarly for \(P' - P\).

**Theorem 4** (Jordan decomposition). Every signed measure \(s: \mathcal{M} \to E^*\) is the difference of two measures,
\[s = s^+ - s^- \quad (s^+, s^- \geq 0),\]
with \(s^+\) or \(s^-\) bounded on \(\mathcal{M}\).

**Proof.** Suppose \(s \neq +\infty\) on \(\mathcal{M}\). Then by Theorem 3, there is a set \(P \in \mathcal{M}^+\) such that \(-P\) is negative and \(sP < \infty\). Now define, for all sets \(A \in \mathcal{M}\),
\[s^+ A = s(A \cap P) \quad \text{and} \quad s^- A = -s(A - P).\]

By additivity,
\[sA = s(A \cap P) + s(A - P) = s^+ A - s^- A;\]
so \(s = s^++s^-\) on \(\mathcal{M}\), as required. Moreover,
\[s^+ A = s(A \cap P) \geq 0,\]
since \(A \cap P \subseteq P\) and \(P\) is positive. Similarly,
\[s^- A = -s(A - P) \geq 0,\]
since \(A - P \subseteq -P\) and \(-P\) is negative. Thus \(s^+, s^- \geq 0\) on \(\mathcal{M}\), a \(\sigma\)-ring.

The \(\sigma\)-additivity of \(s^+\) and \(s^-\) easily follows from that of \(s\) (we leave the proof to the reader). Thus \(s^+\) and \(s^-\) are measures.

Finally, by (2),
\[s^+ A = s(A \cap P) \leq sP < \infty\]
for all \(A \in \mathcal{M}\) (for \(sP = \max\{sX \mid X \in \mathcal{M}\}\);
see the proof of Theorem 3). Thus \(s^+\) is bounded, and all is proved.

The case \(s \neq -\infty\) is similar. □

**Note 5.** For any set \(X \subseteq A (X \in \mathcal{M})\), we have
\[sX = s^+X - s^-X \leq s^+A,\]
for \(s^+\) and \(s^-\) are \(\geq 0\) and monotone. Thus \(s^+ A\) is an upper bound of
\[\{sX \mid A \supseteq X \in \mathcal{M}\}.\]
By (2), this bound is reached when \(X = A \cap P\); so it is a maximum. Similarly for \(s^-\); thus
\[(3) \quad s^+ A = \max\{sX \mid A \supseteq X \in \mathcal{M}\}\]
and \(s^- A = \max\{-sX \mid A \supseteq X \in \mathcal{M}\}\).

**Note 6.** The decomposition is not unique, for we also have
\[s = (s^+ + m) - (s^- + m)\]
for any finite measure \(m\) on \(\mathcal{M}\). However, it becomes unique if we add condition (3). When so defined, \(s^+\) and \(s^-\) are called the Jordan components of \(s\).

**Note 7.** Formula (2) shows that
\[(-s)^+ = s^- \quad \text{and} \quad (-s)^- = s^+.\]

**Corollary 3.** With \(s, s^+,\) and \(s^-\) as in (3), we have the following.
\[(i) \quad v_+ = s^+ + s^-; \quad \text{hence if } s \text{ is a measure (} s = 0\text{), then} \quad s = v_+ = s^+.\]
\[(ii) \quad v_+ \text{ is finite (} t\text{-finite, } t\text{-continuous, absolutely } t\text{-continuous) iff } s^+ \text{ and } s^- \text{ are, i.e., if } s \text{ is.}\]

**Proof.** We give only an outline here.
\[(i) \quad \text{Take any } \mathcal{M}\text{-partition} \quad A = \bigcup X_i \text{ (disjoint).} \]
Setting
\[m = s^+ + s^-; \]
verify that
\[|sX_i| \leq mX_i\]
and
\[ \sum |sX_i| \leq \sum mX_i = m \bigcup X_i = mA. \]
Thus \( mA \) is an upper bound of sums
\[ \sum |sX_i|. \]
This bound is reached when \( X_1 = A \cap P, X_2 = A - P \) (as in (2)).

(ii) Use Theorem 2, Corollary 2, and Definition 3. Note that \( v_\ast \geq |s|, s^+, \) and \( s^-. \)

**Corollary 4.** A \( t \)-finite signed measure \( s \) is absolutely \( t \)-continuous iff it is \( t \)-continuous.

In particular, this applies to \( \text{finite} \) measures. Corollary 4 follows from Theorem 2 and Note 2, by Corollary 3.

**III.** If \( E = E^n(C^n) \), the function
\[ s: \mathcal{M} \to E \]
has \( n \) real (complex) components
\[ s_1, \ldots, s_n, \]
as defined in Chapter 4, §3. As in Theorem 2 of Chapter 4, §3, one easily obtains the following.

**Theorem 5.** A set function \( s: \mathcal{M} \to E^n(C^n) \) is \( t \)-continuous (absolutely \( t \)-continuous, additive, \( \sigma \)-additive) iff its \( n \) components are. Hence a complex set function \( s \) is \( t \)-continuous (etc.) iff its real and imaginary parts are.

For \( \sigma \)-additivity, one can argue as follows. Let
\[ A = \bigcup_{i=1}^{\infty} A_i \text{ (disjoint)}, \]
with \( A, A_i \in \mathcal{M} \). Use Theorem 2 in Chapter 3, §15, with \( \bar{p} = sA \) and
\[ \bar{x}_m = \sum_{i=1}^{m} sA_i, \]
to get \( p_k = s_k A \), and
\[ x_{mk} = \sum_{i=1}^{m} s_k A_i, \quad k = 1, \ldots, n. \]

**Theorem 6.** A generalized measure \( s: \mathcal{M} \to E^n(C^n) \) is \( t \)-continuous iff it is absolutely \( t \)-continuous. It is always bounded on \( \mathcal{M} \), as is \( v_\ast \).

**Proof.** As \( s: \mathcal{M} \to E^n \) is \( \sigma \)-additive, so is each of its components \( s_k \), by Theorem 5. Thus each \( s_k \) is a \( \text{finite (real) signed measure, with} \)
\[ s_k = \bar{s}_k^+ - \bar{s}_k^-, \]
as in Theorem 4. Here the measures \( \bar{s}_k^+ \) and \( \bar{s}_k^- \) are both finite (as \( s \) is).

Thus by Problem 13 in §6, they are \( \text{bounded} \), say, \( \bar{s}_k^+ \leq K_1 \) and \( \bar{s}_k^- \leq K_2 \) on \( \mathcal{M} \). Hence by Corollaries 1 and 3,
\[ |s_k| \leq v_{sk} = \bar{s}_k^+ + \bar{s}_k^- \leq K_1 + K_2; \]
that is, \( v_{sk} \) is bounded on \( \mathcal{M} \) (\( k = 1, 2, \ldots, n \)). Hence so are \( s \) and \( v_\ast \), for
\[ |s| \leq v_\ast \leq \sum v_{sk} \]
(see Problem 4(iii)).

Now, as \( v_\ast \) is finite, it is certainly \( t \)-finite. Thus by Theorem 2 and Note 2, \( s \) is \( t \)-continuous iff it is absolutely \( t \)-continuous.

This settles the case \( E = E^n \), hence also \( E = C = E^2 \). The case \( E = C^n \) is analogous. \( \square \)

**IV. Completion of a Generalized Measure.** From Problems 14 and 15 of §6, recall that every \textit{measure} \( m \) has a completion \( \overline{\mathcal{M}} \). A similar construction, which we now describe, applies to \textit{generalized measures} \( s: \mathcal{M} \to E \).

Given such an \( s \), let \( \overline{\mathcal{M}} \) be the family of all sets \( X \cup Z \), where \( X \in \mathcal{M} \) and \( Z \) is \( v_\ast \)-null, i.e., \( Z \subseteq U \) for some \( U \in \mathcal{M} \), \( v_\ast U = 0 \) (note that \( v_\ast \) is a \textit{measure} here, by Theorem 2). That is,
\[ \overline{\mathcal{M}} = \{ X \cup Z \mid X \in \mathcal{M}, Z \subseteq U, U \in \mathcal{M}, v_\ast U = 0 \}. \]

We now define \( \bar{s}: \overline{\mathcal{M}} \to E \) by setting
\[ \bar{s}A = sX \]
whenever \( A = X \cup Z \), with \( X \) and \( Z \) as above.

As in Problems 14 and 15 of §6, it follows that \( \overline{\mathcal{M}} \) is a \( \sigma \)-ring \( \supseteq \mathcal{M} \), and that \( \bar{s} \) is a \( \sigma \)-additive extension of \( s \), hence a generalized measure. We call \( \bar{s} \) the completion of \( s \). It is \textit{complete} in the sense that \( \overline{\mathcal{M}} \) contains all \( v_\ast \)-null sets (but it may miss some subsets of \( X \) with \( sX = 0 \)). If \( s \geq 0 \) (a \textit{measure}), then \( s = v_\ast \); so our present definitions agree with Problem 15 in §6. We use these ideas in the following part.
V. Signed Lebesgue–Stieltjes (LS) Measures. Motivated by Theorem 3 in Chapter 5, §7, we shall say that a function

\[ \alpha : E^1 \to E^1 \]

is of bounded variation on \( E^1 \) iff

\[ \alpha = g - h, \]

with \( g^+ \) and \( h^+ \) on all of \( E^1 \).

Then \( g \) and \( h \) induce two LS measures \( m_g \) and \( m_h \) in \( E^1 \).

Let \( m_g \) and \( m_h \) be their restrictions to the Borel field \( \mathcal{B} \) in \( E^1 \). Then

\[ \sigma^*_\alpha = m_g - m_h \]

is finite for sets \( X \in \mathcal{B} \) inside any finite interval \( I \subset E^1 \) (as \( m_g \) and \( m_h \) are finite on intervals).

By Lemma 2, \( \sigma^*_\alpha \) is a signed measure on the \( \mathcal{B} \)-sets in \( I \). Moreover, \( \sigma^*_\alpha \) does not depend on the particular choice of \( g^+ \) and \( h^+ \) \((g - h = \alpha)\) on \( I \). For if also \( \alpha = u - v \) \((u^+, v^+)\) on \( E^1 \), set

\[ \sigma'_\alpha = m_u - m_v. \]

Then for any \((x, y) \subseteq I, \)

\[ \sigma'_\alpha(x, y) = \alpha(y) - \alpha(x) + \sigma^*_\alpha(x, y) \quad \text{(verify!)}; \]

so by Problem 13 in §5, \( \sigma'_\alpha \) on \( \mathcal{B} \)-sets in \( I \).

Thus \( \sigma^*_\alpha \) is uniquely determined by \( \alpha \). Its completion

\[ s_\alpha = \sigma^*_\alpha \]

is the \( \alpha \)-induced Lebesgue–Stieltjes (LS) signed measure in \( I \).

If further \( m_g \) or \( m_h \) is finite on all of \( \mathcal{B} \), the same process defines a signed LS measure in all of \( E^1 \).

Problems on Generalized Measures

1. Complete the proofs of Theorems 1, 4, and 5.

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 \( s : \mathcal{M} \to E \) is additive on \( \mathcal{M} \), a semiring, so is \( v_s \).

   [Hint: Use Theorem 1 from §4.]

3. For any set functions \( s, t \) on \( \mathcal{M} \), prove that

   (i) \( v_s| = v_s \), and

   (ii) \( v_{st} \leq av_t \), provided \( st \) is defined and

   \[ a = \sup \{|x| \mid X \in \mathcal{M} \}. \]

4. Given \( s, t : \mathcal{M} \to E \), show that

   (i) \( v_{s+t} \leq v_s + v_t; \)

   (ii) \( v_{sk} = |k|v_k \) \((k \text{ as in Corollary 2); and} \)

   (iii) if \( E = E^n \) \((C^n) \) and

   \[ s = \sum_{k=1}^{n} s_k \bar{e}_k, \]

   then

   \[ v_{sk} \leq v_k \leq \sum_{k=1}^{n} v_{sk}. \]

   [Hints: (i) If \( A \supseteq \bigcup X_i \) (disjoint),

   with \( A_i, X_i \in \mathcal{M} \), verify that

   \[ |(s + t)X_i| \leq |sX_i| + |tX_i|, \]

   \[ \sum |(s + t)X_i| \leq v_sA + v_tA, \text{ etc.;} \]

   (ii) is analogous.

   (iii) Use (ii) and (i), with \( |\bar{e}_k| = 1 \).]

5. If \( g^+, h^+ \), and \( \alpha = g - h \) on \( E^1 \), can one define the signed LS measure \( s_\alpha \) by simply setting \( s_\alpha = m_g - m_h \) \((\text{assuming } m_h < \infty ) \)?

   [Hint: the domains of \( m_g \) and \( m_h \) may be different. Give an example. How about taking their intersection?]?

6. Find an LS measure \( m_\alpha \) such that \( \alpha \) is continuous and one-to-one, but \( m_\alpha \) is not \( m \)-finite \((m = \text{Lebesgue measure}) \).

   [Hint: Take

   \[ \alpha(x) = \begin{cases} \frac{x^2}{4}, & x \neq 0, \\ 0, & x = 0, \end{cases} \]

   and

   \[ A = \bigcup_{n=1}^{\infty} \left( n + \frac{1}{n^2} \right) \]

   7. Construct complex and vector-valued LS measures \( s_\alpha : \mathcal{M}_\alpha \to E^n \) \((C^n) \) in \( E^1 \).
8. Show that if $s: \mathcal{M} \to E^n (C^n)$ 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 $U^+ (U^-)$ be the union of those $X_i$ for which $sX_i \geq 0$ ($sX_i < 0$, respectively). Show that

$$\sum sX_i = sU^+ - sU^- \leq 2 \sup |s| < \infty$$

9. For any $s: \mathcal{M} \to \mathbb{E}^+$ and $A \in \mathcal{M}$, set

$$s^+ A = \sup\{sX \mid A \supseteq X \in \mathcal{M}\}$$

and

$$s^- A = \sup\{-sX \mid A \supseteq X \in \mathcal{M}\}.$$

Prove that if $s$ is additive and bounded on $\mathcal{M}$, a ring, so are $s^+$ and $s^-$; furthermore,

$$s^+ = \frac{1}{2} (v_s + s) \geq 0,$$

$$s^- = \frac{1}{2} (v_s - s) \geq 0,$$

$$s = s^+ - s^-,$$ and

$$v_s = s^+ + s^-.$$ 

[Hints: Use Problem 8. Set

$$s' = \frac{1}{2} (v_s + s).$$

Then $(\forall X \in \mathcal{M}) (|X| \subseteq A)$

$$2sX = sA + sX - s(A - X) \leq sA + (|sX| + |s(A - X)|) \leq sA + v_s A = 2s'A.$$ 

Deduces that $s^+ A \leq s'A$.

To prove also that $s' \leq s^+$, let $\varepsilon > 0$. By Problems 2 and 8, fix $\{X_i\} \subseteq \mathcal{M}$, with

$$A = \bigcup_{i=1}^{\infty} X_i \text{ (disjoint)}$$

and

$$v_s A - \varepsilon < \sum_{i=1}^{\infty} |sX_i|.$$ 

Show that

$$2s'A - \varepsilon = v_s A + sA - \varepsilon \leq sU^+ - sU^- + \sum_{i=1}^{\infty} X_i = 2sU^+$$ 

and

$$2s^+ A \geq 2sU^+ \geq 2s'A - \varepsilon.$$ 

10. Let

$$\mathcal{K} = \{\text{compact sets in a topological space} (S, \mathcal{G})\}$$

(adopt Theorem 2 in Chapter 4, §7, as a definition). Given

$$s: \mathcal{M} \to E, \quad \mathcal{M} \subseteq 2^S,$$

we call $s$ compact regular (CR) iff

$$(\forall \varepsilon > 0) \ (\forall A \in \mathcal{M}) \ (\exists F \in \mathcal{K}) \ (\exists G \in \mathcal{G})$$

$$F, G \in \mathcal{M}, \quad F \subseteq A \subseteq G, \quad \text{and} \quad v_s G - \varepsilon \leq v_s A \leq v_s F + \varepsilon.$$ 

Prove the following.

(i) If $s, t: \mathcal{M} \to E$ 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 $\mathcal{M}_v$ (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 $v_s (G - F)$.

11. (Aleksandrov.) Show that if $s: \mathcal{M} \to E$ 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 $\sigma$-additive, and $v_s$ has a unique $\sigma$-additive extension $\theta_v$ to the $\sigma$-ring $\mathcal{N}$ generated by $\mathcal{M}$.

The latter holds for $s$, too, if $S \in \mathcal{M}$ and $E = E^n (C^n)$.

[Proof outline: The $\sigma$-additivity of $v_s$ results as in Theorem 1 of §2 (first check Lemma 1 in §1 for $v_s$).

For the $\sigma$-additivity of $s$, let

$$A = \bigcup_{i=1}^{\infty} A_i \text{ (disjoint), } A, A_i \in \mathcal{M};$$

then

$$\left| sA - \sum_{i=1}^{r-1} sA_i \right| \leq \sum_{i=r}^{\infty} v_s A_i \to 0$$

as $r \to \infty$, for

$$\sum_{i=1}^{\infty} v_s A_i = v_s \bigcup_{i=1}^{\infty} A_i < \infty.$$ 

(Explain!) Now, Theorem 2 of §6 extends $v_s$ to a measure on a $\sigma$-field

$$\mathcal{M}^* \supseteq \mathcal{N} \supseteq \mathcal{M}$$

(use the minimality of $\mathcal{N}$). Its restriction to $\mathcal{N}$ is the desired $\theta_v$ (unique by Problem 15 in §6).
A similar proof holds for $s$, too, if $s: \mathcal{M} \to [0, \infty)$. The case $s: \mathcal{M} \to E^n (C^m)$ results via Theorem 5 and Problem 10(iii) provided $S \in \mathcal{M}$; for then by Corollary 1, \( v_\ast S \leq \infty \) ensures the finiteness of $v_\ast, s^\ast$, and $s^\ast$ even on $\mathcal{N}'$.

12. Do Problem 11 for semirings $\mathcal{M}$.

[Hint: Use Problem 10(ii).]

### §12. Differentiation of Set Functions

In the proof of Theorem 3 in §10 and the lemmas of that section, we saw the connection between quotients of the form

$$\frac{\Delta f}{\Delta x} = \frac{f(x) - f(p)}{x - p}$$

and those of the form

$$\frac{sI}{mI},$$

where $m$ is Lebesgue measure and $s$ is another suitable measure. With this in mind, we now use quotients $sI/mI$ for forming derivatives of set functions. Below, $m$ is Lebesgue measure in $E^n$;

$$\overline{K} = \{ \text{nondegenerate cubes} \}.$$

**Definition 1.**

Assume the set function

$s: \mathcal{M} \to E^n (\mathcal{M} \supseteq \overline{K})$

in $E^n$ and that $q \in E$.

(i) We say that $q$ is the derivative of $s$ at a point $\bar{p} \in E^n$ iff

$$q = \lim_{k \to \infty} \frac{sI_k}{mI_k}$$

for all sequences $\{I_k\} \subseteq \overline{K}$, with $I_k \to \bar{p}$ (see Definition 1 in §10).

Notation:

$$q = s'(\bar{p}) = \frac{d}{dm} s(\bar{p}).$$

If, in addition, $|q| < \infty$, we say that $s$ is differentiable at $\bar{p}$. If

$$q = \lim_{k \to \infty} \frac{sI_k}{mI_k}$$

for at least one such sequence $I_k \to \bar{p}$, we call $q$ a derivate of $s$ at $\bar{p}$ and write

$q \sim Ds(\bar{p}).$

If $s'(\bar{p})$ exists, it is the unique derivate of $s$ at $\bar{p}$.

(ii) In case $E$ is $E^*$ or $E^1$, we admit infinite derivate and derivatives.

For any set function

$s: \mathcal{M} \to E^*$

(measure or not) with

$$\mathcal{M} \supseteq \overline{K},$$

we also define

$$\overline{D}s(\bar{p}) \text{ and } \overline{D}s(\bar{p})$$

exactly as in Definition 3 of §10.

Equivalently, $\overline{D}s(\bar{p})$ is the least and $\overline{D}s(\bar{p})$ is the largest derivate of $s$ at $\bar{p}$ (Problem 11 in §10). This shows that if $E = E^*$ or $E = E^1$, derivate exist at every $\bar{p}$.

**Note 1.** Hence $q = s'(\bar{p})$ in $E^*$ iff

$$q = \overline{D}s(\bar{p}) = \overline{D}s(\bar{p}).$$

**Note 2.** We treat $D$s, $\overline{D}s$, and $s'$ as functions on points of $E^n$. Thus they are point functions, even though $s$ is a set function.

The easy proofs of Theorems 1 and 2 (with $\mathcal{M}$ and $\mathcal{M} \supseteq \overline{K}$ as above) are left to the reader.

**Theorem 1.** If $s, t: \mathcal{M} \to E$ are differentiable at $\bar{p}$, so are $s \pm t$ and $ks$ for any scalar $k$. (If $s, t$ are scalar valued, $k$ may be a vector.) Moreover,

$$(s \pm t)' = s' \pm t' \text{ and } (ks)' = ks' \text{ at } \bar{p}.$$  

(See also Problem 7.)

**Theorem 2.** A set function $s: \mathcal{M} \to E^*$ ($C^m$) is differentiable at $\bar{p}$ if its components $s_1, s_2, \ldots, s_r$ are; and then

$$s' = (s'_1, \ldots, s'_r) = \sum_{i=1}^r c_i s'_i \text{ at } \bar{p}.$$  

In particular, for complex functions,

$$s' = s'_{re} + i \cdot s'_{im} \text{ at } \bar{p}.$$  

The process described in Definition 1 will be called Lebesgue differentiation or $\overline{K}$-differentiation, as opposed to “$\Omega$-differentiation,” defined next.\(^\dagger\)

\(^\dagger\) We follow some ideas by E. Munroe here.
Definition 2.
Let $\mu^*$ be a $\mathcal{G}$-regular (§5) outer measure in a metric space $(S, \rho)$; recall that
\[ \mathcal{G} = \{ \text{all open sets in } S \}. \]
Let $\mu: \mathcal{M} \to E^*$ be the $\mu^*$-induced (§6) measure in $S$.
A countable (two-indexed) set family
\[ \Omega = \{ U_n^i \} \subseteq \mathcal{M} \quad (i, n = 1, 2, \ldots) \]
is called a network in $S$ (with respect to $\mu$ and $\rho$) iff
(i*) the space
\[ S = \bigcup_{n=1}^{\infty} U_n^i \quad (\text{disjoint}), \quad i = 1, 2, \ldots, \]
with
\[ 0 < \mu U_n^i < \infty, \quad i, n = 1, 2, \ldots; \]
(ii*) each $U_n^{i+1}$ is a subset of some $U_n^i$ (the $U_n^i$ decrease as $i$ increases);
(iii*) for each $p \in S$, there is a sequence
\[ \{ I_k \} \subseteq \Omega, \]
with $I_k \to p$; that is,
\[ p \in \bigcap_{k=1}^{\infty} I_k \]
and $dI_k \to 0$ ($dI_k =$ diameter of $I_k$ in $(S, \rho)$).

Now, given any set function
\[ s: \mathcal{M}' \to E \quad (\mathcal{M}' \supseteq \Omega), \]
we define derivatives, derivates (also $D s$ and $D s$ if $E \subseteq E^*$), and differentiability exactly as in Definition 1, replacing $\mathcal{M}$ by $\Omega$, and Lebesgue measure $m$ by $\mu$.

Note that these derivates and derivatives depend not only on $\mu$ and $\rho$ but also on the choice of $\Omega$. To stress this, one might write $s'_{\Omega}$ and $D_{\Omega}s$ for $s'$ and $Ds$, respectively. Mostly, however, no confusion is caused by simply writing $s'$ and $Ds$ (and we shall do so).

A network for $E^\alpha$ is suggested in the “hint” to Problem 2 of §2. See also Note 3.

---

2Thus for each fixed $i$, the $U_n^i$ are disjoint. Also, $\mu$ is $\sigma$-finite, and $S \in \mathcal{M}$. 

---

Theorems 1 and 2 carry over to $\Omega$-differentiation, with the same proofs. We shall also need a substitute for the Vitali theorem (Theorem 1 of §10). It is quite simple.

Definition 3.
Let $\Omega$ be as in Definition 2. A set family $\mathcal{N} \subseteq \Omega$ is called an $\Omega$-covering of $A \subseteq (S, \rho)$ iff
\[ A \subseteq \bigcup \mathcal{N}, \]
where $\bigcup \mathcal{N}$ is defined to be $\bigcup_{X \in \mathcal{N}} X$.

Theorem 3. Let $\mathcal{N}$ be an $\Omega$-covering of $A \subseteq S$. Then there is a disjoint sequence
\[ \{ I_k \} \subseteq \mathcal{N} \]
with
\[ A \subseteq \bigcup_k I_k \]
so that
\[ \mu^* \left( A - \bigcup_k I_k \right) = 0 \]
and
\[ \mu^* A = \mu^* \left( A \cap \bigcup_k I_k \right). \]

Proof. As $\mathcal{N} \subseteq \Omega$, $\mathcal{N}$ consists of some of the $U_n^i$. For each $i$, let
\[ \mathcal{N}^i = \{ U_n^i \in \mathcal{N} \mid n = 1, \ldots \}, \]
i.e., $\mathcal{N}^i$ consists of all $U_n^i \in \mathcal{N}$ with that particular index $i$.

Now, by Definition 2(ii*) and (i*) each $U_n^i$ is either disjoint, or one contains the other. (Why?) Thus to construct $\{ I_k \}$, start with all the (disjoint) $\mathcal{N}^1$-sets (if $\mathcal{N}^1 \neq \emptyset$). Then add those $U_n^1 \in \mathcal{N}^2$ that are not subsets of any set from $\mathcal{N}^1$ and hence are disjoint from such sets. Next, add those $U_n^2 \in \mathcal{N}^3$ that are not subsets of any set chosen from $\mathcal{N}^1$ or $\mathcal{N}^2$, and so on.

All $U_n^i$ so chosen form a disjoint subfamily $\mathcal{K} \subseteq \mathcal{N}$ that covers all of $A$, as
\[ A \subseteq \bigcup \mathcal{N} = \bigcup \mathcal{K}. \]
(Why?)

$\mathcal{K}$ is countable (as $\Omega$ is); so we can put it in a sequence $\{ I_k \}$, with
\[ A \subseteq \bigcup_k I_k \quad \text{(disjoint)}, \]
as required. □
We can now prove our main result for $\overline{\mathcal{K}}$ and $\Omega$-differentiation alike.

**Theorem 4.**

(i) If $s: \mathcal{M}^\prime \to E^n (E^r, C^r)$ is a generalized measure in $E^n$, finite on $\overline{\mathcal{K}}$, then $s$ is differentiable a.e. on $E^n$ (under Lebesgue measure $\mathfrak{m}$).

(ii) Similarly for $\Omega$-differentiation in $(S, \rho)$, provided $s$ is finite on $\Omega$ and regular.$^3$

**Proof.** Via components and the Jordan decomposition (Theorem 4 of §11), all reduces to the case where $s$ is a measure ($\geq 0$). Then the proof for $\overline{\mathcal{K}}$-differentiation is as in Lemmas 1 and 2 in §10. (Verify!) For $\Omega$-differentiation, the proof of Lemma 1 in §10 still works, with $\overline{\mathcal{K}}$-coverings replaced by $\Omega$-coverings.

In the proof of Lemma 2, after choosing rationals $v > u$, we choose $Q$, $G \supseteq Q$, the $\Omega$-covering

$$K = \left\{ I \in \Omega \mid I \subseteq G, \frac{sI}{\mu I} < u \right\}$$

of $Q$, and the sequence $\{I_k\} \subseteq K$, as before. (In selecting $G$, we use the $\mathcal{G}$-regularity of $\mu^\prime$; the $I_k$ need not be cubes here, of course.)

Then, however, instead of forming the set $Q_0$, we use the regularity of $s$ to select an open set $G' \in \mathcal{M}^\prime$ with

$$G' \supseteq \bigcup_k I_k \supseteq G$$

and

$$sG' - \varepsilon \leq s \bigcup I_k \leq \sum sI_k.$$

The set family

$$K' = \left\{ I \in \Omega \mid I \subseteq G', \frac{sI}{\mu I} \geq v \right\}$$

is then an $\Omega$-covering of $Q$ (why?); so we find a disjoint sequence $\{I'_k\} \subseteq K'$ with

$$Q \subseteq \bigcup I'_k \subseteq G' \subseteq G$$

and obtain

$$u \cdot (\mu^\prime Q + \varepsilon) \geq u \cdot sG' \geq u \cdot \sum_k \mu I_k \geq \sum_k sI_k \geq sG' - \varepsilon \geq \sum_k sI'_k - \varepsilon$$

$$\geq v \cdot \sum_k \mu I'_k - \varepsilon = v \cdot \mu \bigcup I'_k - \varepsilon \geq v \cdot \mu^\prime Q - \varepsilon.$$

Thus

$$(\forall \varepsilon > 0) \quad u \cdot (\mu^\prime Q + \varepsilon) \geq v \cdot \mu^\prime Q - \varepsilon.$$

The rest is as in Lemma 2 of §10. $\Box$

**Note 3.** If $\mu^\prime = \mu^*\cdot \overline{\mathcal{K}}$-derivatives equal $\Omega$-derivatives a.e. for a regular $s$ (Problem 6). One may use $\Omega$ in $E^n$, thus avoiding Theorem 1 of §10 (Problem 13).

**Problems on Differentiation of Set Functions**

1. Complete the proofs of Theorems 1 to 4 in detail. Verify Note 1.
2. Verify that the hint for Problem 2 in §2 describes a network for $E^n$ (see Note 3).
3. Show that the measure $\mu$ in Definition 2 is necessarily topological. [Hint: Any $G \in \mathcal{G}$ is a countable union of $\Omega$-sets. Why?]  
4. (i) Show that the derivatives of $s$ at $\bar{p}$ form exactly the set $D^\prime_p$ of all cluster points of sequences $sI_k/mI_k$ with $I_k \to \bar{p}$ and $\{I_k\} \subseteq K$. Do the same considering sequences $sI_k/\mu I_k$ with $I_k \to \bar{p}$ and $\{I_k\} \subseteq \Omega$.

(ii) Do Problem 11 in §10 for $\Omega$-differentiation. Must $s$ be regular here?
5. Verify that if

$$(\forall I \in \Omega) \quad \mu I = \mu I^o,$$

then Theorem 4 holds for $\Omega$-differentiation even if $s$ is not regular. [Hint: The proof of Lemma 2 of §10 holds unchanged.]

6. Prove Note 3 assuming that (i) $s$ is regular, or (ii) $(\forall I \in \Omega) \mu I = \mu I^o$ (see Problem 5).

[Hint: Imitate Problem 9(b) in §10 and the “$\Omega$” part in the proof of Theorem 4.]

7. Prove for $\overline{\mathcal{K}}$- and $\Omega$-differentiation that if

$$s = t + u \quad (s, t, u: \mathcal{M}^\prime \to E^n)$$

and if $u$ is differentiable at $p$, then $\overline{\mathcal{D}} s = \overline{\mathcal{D}} t \pm u'$ and $\overline{\mathcal{D}} s = \overline{\mathcal{D}} t \pm u'$ at $p$.

8. In Theorem 4 show that $\overline{\mathcal{D}} s = \overline{\mathcal{D}} s$ a.e. even if $s$ is not finite on all of $\overline{\mathcal{K}}(\Omega)$.

[Hint: For $s \geq 0$, Lemma 1 in §10 still holds. For signed measures, use Problem 7, noting that $s^+$ or $s^-$ is finite, hence differentiable a.e.]

9. Prove that if $f$ and $s = mf$ are as in the proof of Theorem 3 in §10, then $s$ and $f$ are differentiable at the same points in $E^1$, and $s' = f'$ there.

---

$^3$A signed measure $s$ is called regular iff $s^+$ and $s^-$ are regular (Definition 4 in §7). A complex measure $s$ is regular iff $s_{\text{re}}$ and $s_{\text{im}}$ are. Finally, $s: \mathcal{M}^\prime \to E^n (C^n)$ is regular iff all its components $s_i$ are.
10. Given a universal $V$-covering $\mathcal{K}$ (see Problem 9 in §10), develop $\mathcal{K}$-differentiation as in Definition 1, replacing $\mathcal{K}$ by $K^*$ and writing $s^{**}$, $D^{**} s$, ... for $s'$, $D s$, etc.

Extend Theorems 1–4 and Problem 7 to $K^*$-differentiation. Under the assumptions of Theorem 4, show that $s^{**} = s'$ a.e. on $E^n$ (use Problem 9 in §10).

11. Given a normal $V$-covering $K^*$ of $E^n$ (Problem 8 in §10), develop $K^*$-differentiation along the lines of Problem 12 in §10 (admitting normal sequences $\{I_k\}$ only). Do the same questions as in Problem 10, for $K^*$-differentiation.

12. Describe what changes if, in Problem 11, we drop the normality restriction on sequences $I_k \to \bar{p}$ (call it strong $K^*$-differentiation; write $D^{**} s$, $s^{**}$, etc.).

Show that
\[ D^{**} s \leq D' s \leq D^* s \leq D^{**} s \]
on $E^n$, and so the existence of $s^{**}$ implies that of $s^{*}$.

However the proof of Lemmas 1 and 2 in §10 fails for $D^{**} s$ and $D^{**} s$ (at what step?). So does the proof of Theorem 4. What about Theorems 1 and 2?

---

Chapter 8

Measurable Functions. Integration

§1. Elementary and Measurable Functions

From set functions, we now return to point functions
\[ f: S \to (T, \rho') \]
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:
\[ "A(P)" \text{ for } \{x \in A \mid P(x)\}. \]

Thus
\[ A(f \neq a) = \{x \in A \mid f(x) \neq a\}, \]
\[ A(f = g) = \{x \in A \mid f(x) = g(x)\}, \]
\[ A(f > g) = \{x \in A \mid f(x) > g(x)\}, \text{ etc.} \]

Definition 1.

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 2.

An $\mathcal{M}$-partition of a set $A$ is a countable set family $\mathcal{P} = \{A_i\}$ such that
\[ A = \bigcup_i A_i \text{ (disjoint)}, \]
with $A, A_i \in \mathcal{M}$.\(^1\)

We briefly say “the partition $A = \bigcup A_i$.”

\(^1\) $\mathcal{P}$ may be finite; it may even consist of $A$ alone.
An $\mathcal{M}$-partition $\mathcal{P}' = \{B_k\}$ is a refinement of $\mathcal{P} = \{A_i\}$ (or $\mathcal{P}'$ refines $\mathcal{P}$, or $\mathcal{P}'$ is finer than $\mathcal{P}$) iff
\[
(\forall i) \ A_i = \bigcup_k B_k;
\]
i.e., each $B_k$ is contained in some $A_i$.

The intersection $\mathcal{P}' \cap \mathcal{P}'' = \{A_i\}$ and $\mathcal{P}'' = \{B_k\}$ is understood to be the family of all sets of the form
\[
A_i \cap B_k, \quad i, k = 1, 2, \ldots .
\]
It is an $\mathcal{M}$-partition that refines both $\mathcal{P}'$ and $\mathcal{P}''$.

**Definition 3.**
A map (function) $f: S \to T$ is elementary, or $\mathcal{M}$-elementary, on a set $A \in \mathcal{M}$ iff there is an $\mathcal{M}$-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.\(^2\)

The function values $a_i$ are elements of $T$ (possibly vectors). They may be infinite if $T = E^n$. Any simple map is also elementary, of course.

**Definition 4.**
A map $f: S \to (T, \rho')$ is said to be measurable (or $\mathcal{M}$-measurable) on a set $A \in (S, \mathcal{M})$ iff
\[
f = \lim_{m \to \infty} f_m \quad \text{(pointwise on } A)\]
for some sequence of functions $f_m: S \to T$, 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 1.** If $f: S \to (T, \rho')$ is elementary on $A$, it is measurable on $A$.

**Proof.** Set $f_m = f$, $m = 1, 2, \ldots$, in Definition 4. Then clearly $f_m \to f$ on $A$. \(\square\)

**Corollary 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 $B \in \mathcal{M}$.

\(^2\) Only simple step functions are needed for a “limited approach.” (One may proceed from here to §4, treating $m$ as an additive premeasure.)

---

**§1. Elementary and Measurable Functions**

**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
\[
\{B \cap A_i\}, \quad i = 1, 2, \ldots , n,
\]
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.,
\[
f = \lim_{m \to \infty} f_m
\]
for some elementary maps $f_m$ on $A$. As shown above, the $f_m$ are elementary on $B$, too, and $f_m \to f$ on $B$; so $f$ is measurable on $B$. \(\square\)

**Corollary 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,
\[
A = \bigcup_n A_n = \bigcup_n B_n
\]
for some disjoint sets $B_n \subseteq A_n$ ($B_n \in \mathcal{M}$).

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 $A$,
\[
A = \bigcup_n B_{ni} = \bigcup_{n,i} B_{ni}.
\]
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 4.** If $f: S \to (T, \rho')$ is measurable on $A$ in $(S, \mathcal{M})$, so is the composite map $g \circ f$, provided $g: T \to (U, \rho'')$ is relatively continuous on $f[A]$.

**Proof.** By assumption,
\[
f = \lim_{m \to \infty} f_m \quad \text{(pointwise)}
\]
for some elementary maps $f_m$ on $A$.

Hence by the continuity of $g$,
\[
g(f_m(x)) \to g(f(x)),
\]
i.e., $g \circ f_m \to g \circ f$ (pointwise) on $A$.\(\square\)
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. ☐

**Theorem 1.** If the maps \( f, g, h : S \to E^1(C) \) 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,

\[
A = \bigcup A_i = \bigcup B_k,
\]

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 \( A \); hence so is \( f \pm g \) and \( f \pm g \) is elementary on \( A \). Similarly for simple functions.

Next, let \( f \) and \( g \) be measurable on \( A \); so \( f = \lim f_m \) and \( g = \lim g_m \) (pointwise) on \( A \)

for some elementary maps \( f_m, g_m \).

By what was shown above, \( f_m \pm g_m \) is elementary for each \( m \). Also,

\[
f_m \pm g_m \to f \pm g \text{ (pointwise) on } A.
\]

Thus \( f \pm g \) is measurable on \( A \).

The rest of the theorem follows quite similarly. ☐

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 2.** A function \( f : S \to E^n(C^n) \) 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 \( f : S \to E^2 \), \( f = (f_1, f_2) \).

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 \( f = (f_1, f_2) \), too, is constant on each \( A_i \cap B_k \), as required.

Conversely, let

\[
f = c_i \pm (a_i, b_i) \text{ on } C_i,
\]

for some \( \mathcal{M} \)-partition

\[
A = \bigcup C_i.
\]

Then by definition, \( f_1 = a_i \) and \( f_2 = b_i \) on \( C_i \); so both are elementary (or simple) on \( A \).

§1. Elementary and Measurable Functions

In the general case \((E^n \text{ or } C^n)\), 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 \( f : S \to C \) 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 \( M \) is a \( \sigma \)-ring, one can make the limit uniform. Indeed, we have the following theorem.

**Theorem 3.** If \( M \) is a \( \sigma \)-ring, and \( f : S \to (T, \rho') \) is \( \mathcal{M} \)-measurable on \( A \), then

\[
f = \lim_{m \to \infty} g_m \text{ (uniformly) on } A
\]

for some finite elementary maps \( g_m \).

Thus given \( \varepsilon > 0 \), there is a finite elementary map \( g \) such that \( \rho'(f, g) < \varepsilon \) on \( A \).

The proof will be given in §2 for \( T = E^* \). The general case is sketched in Problem 7 of §2. Meanwhile, we take the theorem for granted and use it below.

**Theorem 4.** If \( M \) is a \( \sigma \)-ring in \( S \), if

\[
f_m \to f \text{ (pointwise) on } A
\]

\((f_m : S \to (T, \rho'))\), 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 \( \varepsilon = 1/m, m = 1, 2, \ldots \),

\[
\rho'(f_m(x), g_m(x)) < \frac{1}{m} \quad \text{for all } x \in A \text{ and all } m.
\]

Fixing such a \( g_m \) for each \( m \), we show that \( g_m \to f \) (pointwise) on \( A \), as required in Definition 4.

Indeed, fix any \( x \in A \). By assumption, \( f_m(x) \to f(x) \). Hence, given \( \delta > 0 \),

\[
(\exists k) \ (\forall m > k) \quad \rho'(f(x), f_m(x)) < \delta.
\]

Take \( k \) so large that, in addition,

\[
(\forall m > k) \quad \frac{1}{m} < \delta.
\]

\footnote{We briefly write \( \rho'(f, g) \) for \( \sup_{x \in X} \rho'(f(x), g(x)) \).}

\footnote{The theorem holds also for \( T = E^* \), with \( \rho' \) as in Problem 5 of Chapter 3, §11.}
Then by the triangle law and by (1), we obtain for $m > k$ that
\[
\rho'(f(x), g_m(x)) \leq \rho'(f(x), f_m(x)) + \rho'(f_m(x), g_m(x)) < \delta + \frac{1}{m} < 2\delta.
\]
As $\delta$ is arbitrary, this implies $\rho'(f(x), g_m(x)) \to 0$, i.e., $g_m(x) \to f(x)$ for any (fixed) $x \in A$, thus proving the measurability of $f$. \[\square\]

**Note 3.** If $M = B$ (= Borel field in $S$), we often say “Borel measurable” for $M$-measurable. If $M = \{\text{Lebesgue measurable sets in } E^n\}$, we say “Lebesgue (L) measurable” instead. Similarly for “Lebesgue–Stieltjes (LS) measurable.”

**Problems on Measurable and Elementary Functions in $(S, M)$**

1. Fill in all proof details in Corollaries 2 and 3 and Theorems 1 and 2.
2. Show that $P' \cap P''$ as stated at the end of Definition 2.
3. Given $A \subseteq S$ and $f, f_m : S \to (T, \rho')$, $m = 1, 2, \ldots$, let
\[
H = A(f_m \to f)
\]
and
\[
A_{mn} = A\left(\rho'(f_m, f) < \frac{1}{n}\right).
\]
Prove that
\[
\text{(i) } H = \bigcap_{n=1}^{\infty} \bigcap_{k=1}^{\infty} \bigcap_{m=k}^{\infty} A_{mn};
\]
\[
\text{(iii) } H \in M \text{ if all } A_{mn} \text{ are in } M \text{ and } M \text{ is a } \sigma\text{-ring.}
\]
[Hint: $x \in H$ iff $\forall n \exists k \forall m \geq k \ x \in A_{mn}$. Why?]

4. Do Problem 3 for $T = E^*$ and $f = \pm \infty$ on $H$.
[Hint: If $f = \infty, A_{mn} = A(f_m > n)$.]

5. Let $f : S \to T$ be $M$-elementary on $A$, with $M \text{ a } \sigma\text{-ring in } S$. Show the following.
\[
\text{(i) } A(f = a) \in M, A(f \neq a) \in M.
\]
\[\square\]

6. Given $f, g : S \to (T, \rho')$, let $h = \rho'(f, g)$, i.e.,
\[
h(x) = \rho'(f(x), g(x)).
\]
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.]

7. A set $B \subseteq (T, \rho')$ is called separable (in $T$) iff $B \subseteq \overline{D}$ (closure of $D$) for a countable set $D \subseteq T$.

8. Continuing Problem 7, prove that if $B \subseteq \overline{D}$ and $D = \{q_1, q_2, \ldots\}$, then
\[
\forall n \ B \subseteq \bigcup_{i=1}^{\infty} G_{\rho_i}(\frac{1}{n}).
\]
9. Prove Corollaries 2 and 3 and Theorems 1 and 2, assuming that $\mathcal{M}$ is a semiring only.

10. Do Problem 4 for $\mathcal{M}$-simple maps, assuming that $\mathcal{M}$ is a ring only.

§2. Measurability of Extended-Real Functions

Henceforth we presuppose a measurable space $(S, \mathcal{M})$, where $\mathcal{M}$ is a $\sigma$-ring in $S$. Our aim is to prove the following basic theorem, which is often used as a definition, for extended-real functions $f : S \to E^*$.

**Theorem 1.** A function $f : S \to E^*$ is measurable on a set $A \in \mathcal{M}$ if it satisfies one of the following equivalent conditions (hence all of them):

(i) $(\forall a \in E^*)$ $A(f > a) \in \mathcal{M}$;  
(ii) $(\forall a \in E^*)$ $A(f \geq a) \in \mathcal{M}$;  
(iii) $(\forall a \in E^*)$ $A(f < a) \in \mathcal{M}$;  
(iv) $(\forall a \in E^*)$ $A(f \leq a) \in \mathcal{M}$.

We first prove the equivalence of these conditions by showing that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i), closing the “circle.”

(i) $\Rightarrow$ (ii). Assume (i). If $a = -\infty$,  
\[ A(f \geq a) = A \in \mathcal{M} \]
by assumption. If $a = +\infty$,  
\[ A(f \geq a) = A(f = \infty) = \bigcap_{n=1}^{\infty} A(f > a) \in \mathcal{M} \]
by (i). And if $a \in E$,
\[ A(f \geq a) = \bigcap_{n=1}^{\infty} A(f > a - \frac{1}{n}) \].
(Verify!) By (i),  
\[ A(f > a - \frac{1}{n}) \in \mathcal{M}; \]
so $A(f \geq a) \in \mathcal{M}$ (a $\sigma$-ring!).

(ii) $\Rightarrow$ (iii). For (ii) and $A \in \mathcal{M}$ imply  
\[ A(f < a) = A - A(f \geq a) \in \mathcal{M}. \]

(iii) $\Rightarrow$ (iv). If $a \in E$,  
\[ A(f \leq a) = \bigcap_{n=1}^{\infty} A(f < a + \frac{1}{n}) \in \mathcal{M}. \]

What if $a = \pm \infty$?

(iv) $\Rightarrow$ (i). Indeed, (iv) and $A \in \mathcal{M}$ imply  
\[ A(f > a) = A - A(f \leq a) \in \mathcal{M}. \]

Thus, indeed, each of (i) to (iv) implies the others. To finish, we need two lemmas that are of interest in their own right.

**Lemma 1.** If the maps $f_m : S \to E^*$ $(m = 1, 2, \ldots)$ satisfy conditions (i)–(iv), so also do the functions
\[ \sup f_m, \inf f_m, \liminf f_m, \text{ and } \limsup f_m, \]
defined pointwise, i.e.,  
\[ (\sup f_m)(x) = \sup f_m(x), \]
and similarly for the others.

**Proof.** Let $f = \sup f_m$. Then  
\[ A(f \leq a) = \bigcap_{n=1}^{\infty} A(f_m \leq a). \] (Why?)

But by assumption,  
\[ A(f_m \leq a) \in \mathcal{M} \]
($f_m$ satisfies (iv$^*$)). Hence $A(f \leq a) \in \mathcal{M}$ (for $\mathcal{M}$ is a $\sigma$-ring).

Thus $\sup f_m$ satisfies (i$^*$)–(iv$^*$).

So does $\inf f_m$; for  
\[ A(\inf f_m \geq a) = \bigcap_{m=1}^{\infty} A(f_m \geq a) \in \mathcal{M}. \]
(Explain!)

So also do $\liminf f_m$ and $\limsup f_m$; for by definition,  
\[ \liminf f_m = \sup_{k} g_k, \]
where  
\[ g_k = \inf_{m \geq k} f_m, \]
satisfies (i∗)–(iv∗), as was shown above; hence so does sup $g_k = \lim f_m$.

Similarly for $\lim f_m$. □

**Lemma 2.** If $f$ satisfies (i∗)–(iv∗), then
\[
f = \lim_{m \to \infty} f_m \text{ (uniformly) on } A
\]
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 $\{f_m\}$ on $A$.

**Proof.** Let $H = A(f = +\infty)$, $K = A(f = -\infty)$, and
\[
A_m = A\left(\frac{k + 1}{2^m} \leq f < \frac{k}{2^m}\right)
\]
for $m = 1, 2, \ldots$ and $k = 0, \pm 1, \pm 2, \ldots, \pm n, \ldots$.

By (i∗)–(iv∗),
\[
H = A(f = +\infty) = A(f \geq +\infty) \in \mathcal{M},
\]
$K \in \mathcal{M}$, and
\[
A_m = A\left(\frac{k + 1}{2^m} \leq f < \frac{k}{2^m}\right) \cap A(f < k) \in \mathcal{M}.
\]

Now define
\[
(\forall m) \quad f_m = \frac{k + 1}{2^m} \text{ on } A_m.
\]

$f_m = m$ on $H$, and $f_m = -m$ on $K$. Then each $f_m$ is finite and elementary on $A$ since
\[
(\forall m) \quad A = H \cup K \cup \bigcup_{k = -\infty}^{\infty} A_m \text{ (disjoint)}
\]
and $f_m$ is constant on $H, K,$ and $A_m$.

We now show that $f_m \to f$ (uniformly) on $H, K,$ and $A_m$.

Indeed, on $H$ we have
\[
\lim f_m = \lim m = +\infty = f,
\]
and the limit is uniform since the $f_m$ are constant on $H$.

Similarly,
\[
f_m = -m \to -\infty = f \text{ on } K.
\]

Finally, on $A_m$ we have
\[
(k - 1)2^{-m} \leq f < k2^{-m}
\]
and $f_m = (k - 1)2^{-m}$; hence
\[
|f_m - f| < k2^{-m} - (k - 1)2^{-m} = 2^{-m}.
\]

Thus
\[
|f_m - f| < 2^{-m} \to 0
\]
on each $A_m$, hence on
\[
J = \bigcup_{k = -\infty}^{\infty} A_m.
\]

By Theorem 1 of Chapter 4, §12, it follows that $f_m \to f$ (uniformly) on $J$. Thus, indeed, $f_m \to f$ (uniformly) on $A$.

If, further, $f \geq 0$ on $A$, then $K = \emptyset$ and $A_m = \emptyset$ for $k \leq 0$. Moreover, on passage from $m$ to $m + 1$, each $A_m$ ($k > 0$) splits into two sets. On one, $f_{m+1} = f_m$; on the other, $f_{m+1} > f_m$. (Why?)

Thus $0 \leq f_m \not\to f$ (uniformly) on $A$, and all is proved. □

**Proof of Theorem 1.** 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 $f = \lim f_m = \lim f_m$.

The converse follows by Lemma 2. This completes the proof. □

**Note 1.** Lemmas 1 and 2 prove Theorems 3 and 4 of §1, for $f : S \to E^*$. By using also Theorem 2 in §1, one easily extends this to $f : S \to E^n(C^m)$. Verify!

**Corollary 1.** If $f : S \to E^*$ is measurable on $A$, then
\[
(\forall a \in E^*) \quad A(f = a) \in \mathcal{M} \quad \text{and} \quad A(f \neq a) \in \mathcal{M}.
\]

Indeed,
\[
A(f = a) = A(f \geq a) \cap A(f \leq a) \in \mathcal{M}
\]
and
\[
A(f \neq a) = A - A(f = a) \in \mathcal{M}.
\]

**Corollary 2.** If $f : S \to (T, \rho')$ is measurable on $A$ in $(S, \mathcal{M})$, then
\[
A \cap f^{-1}[G] \in \mathcal{M}
\]
for every globe $G = G_\rho(\delta)$ in $(T, \rho')$.

**Proof.** Define $h : S \to E^1$ by
\[
h(x) = \rho'(f(x), q).
\]
Then \( h \) is measurable on \( A \) by Problem 6 in §1. Thus by Theorem 1,
\[ A(h < \delta) \in \mathcal{M}. \]

But as is easily seen,
\[ A(h < \delta) = \{ x \in A \mid \rho'(f(x), q) < \delta \} = A \cap f^{-1}[G_q(\delta)]. \]

Hence the result. \( \square \)

**Definition.**

Given \( f, g : S \to E^* \), we define the maps \( f \lor g \) and \( f \land g \) on \( S \) by
\[ (f \lor g)(x) = \max\{f(x), g(x)\} \]
and
\[ (f \land g)(x) = \min\{f(x), g(x)\}; \]
similarly for \( f \lor g \lor h, f \land g \land h, \) etc.

We also set
\[ f^+ = f \lor 0 \text{ and } f^- = - f \lor 0. \]

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 2.** If the functions \( f, g : S \to E^* \) are simple, elementary, or measurable on \( A \), so also are \( f \pm g, f, g, f \lor g, f \land g, f^+, f^- \), and \( |f|^a \) (\( a \neq 0 \)).

**Proof.** If \( f \) and \( g \) are finite, this follows by Theorem 1 of §1 on verifying that
\[ f \lor g = \frac{1}{2}(f + g + \lvert f - g \rvert) \]
and
\[ f \land g = \frac{1}{2}(f + g - \lvert f - g \rvert) \]
on \( S \). (Check it!)

Otherwise, consider
\[ A(f = +\infty), A(f = -\infty), A(g = +\infty), \text{ and } A(g = -\infty). \]

By Theorem 1, these are \( \mathcal{M} \)-sets; hence so is their union \( U \).

On each of them \( f \lor g \) and \( f \land g \) equal \( f \lor g \) or \( f \land g \); so by Corollary 3 in §1, \( f \lor g \) and \( f \land g \) have the desired properties on \( U \). So also have \( f^+ = f \lor 0 \) and \( f^- = - f \lor 0 \) (take \( g = 0 \)).

We claim that the maps \( f \pm g \) and \( fg \) are simple (hence elementary and measurable) on each of the four sets mentioned above, hence on \( U \).

For example, on \( A(f = +\infty) \),
\[ f \pm g = +\infty \text{ (constant)} \]

by our conventions (2") in Chapter 4, §4. For \( fg \), 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 \( fg = +\infty \) on \( A_1 \), \( fg = -\infty \) on \( A_2 \), and \( fg = 0 \) on \( A_3 \). Hence \( fg \) 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. \( \square \)

**Note 2.** Induction extends Theorem 2 to any finite number of functions.

**Note 3.** Combining Theorem 2 with \( f = f^+ - f^- \), we see that \( f : S \to E^* \)
is simple (elementary, measurable) if \( f^+ \) and \( f^- \) are. We also obtain the following result.

**Theorem 3.** If the functions \( f, g : S \to E^* \) 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 \( A(f \neq g) \in \mathcal{M} \).

(See Problem 4 below.)

**Further Problems on Measurable Functions in \((S, \mathcal{M})\)**

1. In Theorem 1, give the details in proving the equivalence of \((i') - (iv')\).

2. Prove Note 1.

2'. Prove that \( f = f^+ - f^- \) and \( |f| = f^+ + f^- \).

3. Complete the proof of Theorem 2, in detail.

\( \Rightarrow 4. \) 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 \( A(f < g) = A(f \geq g) \in \mathcal{M} \). Proceed.]

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 \( Q \notin \mathcal{M} \) \( (Q \subseteq A) \); e.g., use \( Q \) of Problem 6 in Chapter 7, §8.]

\( \Rightarrow 6. \) Show that a function \( f \geq 0 \) is measurable on \( A \) if \( f \ni \neq f \) (pointwise) on \( A \) for some finite simple maps \( f_m \geq 0 \), \( \{f_m\} \).\( ^* \)

[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 \( m \geq k \leq m^2 \) only.]

\( \Rightarrow 7. \) Prove Theorem 3 in §1.

[Outline: By Problems 7 and 8 in §1, there are \( q_i \in T \) such that
\[ (\forall n) \quad f[A] \subseteq \bigcup_{i=1}^{m} G_{q_i}\left(\frac{1}{n}\right). \]
Set
\[ A_{m_i} = A \cap f^{-1}\left[ G_{q_i}\left(\frac{1}{n}\right) \right] \in \mathcal{M} \]
by Corollary 2, so \( \rho'(f(x), q_i) < \frac{1}{n} \) on \( A_{m_i} \).]
By Corollary 1 in Chapter 7, §1,

\[ A = \bigcup_{i=1}^{\infty} A_{n_i} = \bigcup_{i=1}^{\infty} B_{n_i} \text{ (disjoint)} \]

for some sets \( B_{n_i} \subseteq M \), \( M \subseteq A_{n_i} \). Now define \( g_n = g_{0i} \) on \( B_{n_i} \); so \( \rho'(f, g_n) < \frac{1}{n} \) on each \( B_{n_i} \), hence on \( A \). By Theorem 1 in Chapter 4, §12, \( g_n \to f \) (uniformly) on \( A \).

\[ \Rightarrow 8. \] Prove that \( f : S \to E^1 \) is \( M \)-measurable on \( A \) iff \( A \cap f^{-1}[B] \in M \) for every Borel set \( B \) (equivalently, for every open set \( B \)) in \( E^1 \). (In the case \( f : S \to E^* \), add: “and for \( B = \{ \pm \infty \} \).”)

[Outline: Let \( R = \{ X \subseteq E^1 \mid A \cap f^{-1}[X] \in M \} \).

Show that \( R \subseteq \mathcal{S} \), and \( R \subseteq 2^S \).

Now, by Theorem 1, if \( f \) is measurable on \( A \), \( R \) contains all open intervals; for

\[ A \cap f^{-1}[(a, b)] = A(f > a) \cap A(f < b). \]

Then by Lemma 2 of Chapter 7, §2, \( R \supseteq \mathcal{G} \), hence \( R \subseteq 2^S \). (Why?)

Conversely, if so,

\[ (a, \infty) \in R \Rightarrow A \cap f^{-1}[(a, \infty)] \in M \Rightarrow A(f > a) \in M. \]

\[ \Rightarrow 9. \] Do Problem 8 for \( f : S \to E^n \).

[Hint: If \( f = (f_1, \ldots, f_n) \) and \( B = (a, b) \subseteq E^n \), with \( a = (a_1, \ldots, a_n) \) and \( b = (b_1, \ldots, b_n) \), show that

\[ f^{-1}[B] = \bigcap_{k=1}^{n} F_k^{-1}[(a_k, b_k)]. \]

Apply Problem 8 to each \( f_k : S \to E^1 \) and use Theorem 2 in §1. Proceed as in Problem 8.]

10. Do Problem 8 for \( f : S \to C^n \), treating \( C^n \) as \( E^{2n} \).

11. Prove that \( f : S \to (T, \rho) \) is measurable on \( A \) in \((S, M)\) iff

\[ (i) \quad A \cap f^{-1}[G] \in M \text{ for every open globe } G \subseteq T, \text{ and} \]

\[ (ii) \quad f[A] \text{ is separable in } T \text{ (Problem 7 in §1).} \]

[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.]

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 \( G_k \) of Problem 8 in §1, putting them in one sequence.

(ii) Take \( D = R^n \subseteq E^n \) in Problem 7 of §1.]

\[ \S 2. \] Measurability of Extended-Real-Functions

13. Do Problem 11 with “globe \( G \subseteq T \)” replaced by “Borel set \( B \subseteq T \).”

[Hints: Treat \( f \) as \( f : A \to T^*, T^* = f[A], \) noting that

\[ A \cap f^{-1}[B] = A \cap f^{-1}[B \cap T^*]. \]

By Problem 12, if \( B \neq \emptyset \) is open in \( T \), then \( B \cap T^* \) is an open set of “globes” \( G_k \cap T^* \) in \( T^*(\rho) \); see Theorem 4 in Chapter 3, §12. Proceed as in Problem 8, replacing \( E^1 \) by \( T \).]

14. A map \( g : (T, \rho^*) \to (U, \rho^u) \) is said to be of Baire class \( 0 \) \((g \in \mathcal{B}_0) \) iff \( g \) is relatively continuous on \( D \subseteq T \) and \( g \) is of Baire class \( n \) \((g \in \mathcal{B}_n) \) iff \( g = \lim g_n \) (pointwise) on \( D \) for some maps \( g_n \in \mathcal{B}_{n-1} \). Show by induction that Corollary 4 in §1 holds also if \( n \in \mathcal{B}_n \).

\[ \S 3. \] Measurable Functions in \((S, \mathcal{M}, m)\)

1. Henceforth we shall presuppose not just a measurable space \((\mathcal{S}, \mathcal{M}, m)\), where \( m : \mathcal{M} \to E^* \) is a measure on a \( \sigma \)-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 1.**

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 \( A \) if \( f \) holds on \( A - Q \) for some \( Q \in \mathcal{M} \) with \( mQ = 0 \).

Thus we write

\[ f_n \to f \text{ (a.e.) or } f = \lim f_n \text{ (a.e.(m)) on } A \]

if \( f_n \to f \) (pointwise) on \( A - Q, mQ = 0 \). Of course, “pointwise” implies “a.e.” (take \( Q = \emptyset \)), but the converse fails.

**Definition 2.**

We say that \( f : S \to (T, \rho) \) is almost measurable on \( A \) if \( f \) is \( M \)-measurable on \( A - Q \), \( mQ = 0 \).

We then also say that \( f \) is \( m \)-measurable \((m \text{ being the measure involved})\) as opposed to \( M \)-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 \( v_mQ = 0 \) \((v_m = \text{total variation of } m)\) in Definitions 1 and 2 and in the following proofs.
Corollary 1. If the functions
\[ f_n : S \to (T, \rho'), \quad n = 1, 2, \ldots, \]
are \( m \)-measurable on \( A \), and if
\[ f_n \to f \text{ (a.e.)(m))} \]
on \( A \), then \( f \) is \( m \)-measurable on \( A \).

Proof. By assumption, \( f_n \to f \) (pointwise) on \( A - Q_0 \), \( mQ_0 = 0 \). Also, \( f_n \) is \( \mathcal{M} \)-measurable on
\[ A - Q_n, \quad mQ_n = 0, \quad n = 1, 2, \ldots. \]
(The \( Q_n \) need not be the same.)

Let
\[ Q = \bigcup_{n=0}^{\infty} Q_n; \]
so
\[ mQ = \sum_{n=0}^{\infty} mQ_n = 0. \]

By Corollary 2 in §1, all \( f_n \) are \( \mathcal{M} \)-measurable on \( A - Q \) (why?), and \( f_n \to f \) (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. \( \square \)

Corollary 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. \( \square \)

Corollary 3. If \( f : S \to (T, \rho') \) is \( m \)-measurable on \( A \), then
\[ f = \lim_{n \to \infty} f_n \quad (\text{uniformly}) \quad \text{on } A - Q \quad (mQ = 0), \]
for some maps \( f_n \), all elementary on \( A - Q \).

(Compare Corollary 3 with Theorem 3 in §1).

\section{Measurable Functions in \((S, \mathcal{M}, m)\)}

Corollary 4. If the functions
\[ f_n : S \to E^* \quad (n = 1, 2, \ldots) \]
are \( m \)-measurable on a set \( A \), so also are
\[ \sup f_n, \inf f_n, \lim f_n, \text{ and } \liminf f_n. \]

(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 \( f : S \to E^n (C^n), m \) 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 1. A map \( f : S \to E^n (C^n) \) is \( \mathcal{M} \)-measurable on \( A \) iff
\[ A \cap f^{-1} [B] \in \mathcal{M} \]
for each Borel set (equivalently, open set) \( B \) in \( E^n (C^n) \).

See Problems 8–10 in §2 for a sketch of the proof.

Lemma 2. A map \( f : (S, \rho) \to (T, \rho') \) 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 \((A, \rho)\) as a subspace of \((S, \rho)\).

(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 \( g : A \to T \) is continuous in the ordinary sense.

Now, by Problem 15(iv)(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 open in \((T, \rho')\). But
\[ g^{-1} [B] = \{ x \in A \mid f(x) \in B \} = A \cap f^{-1} [B]. \]

(Why?) Hence the result follows. \( \square \)

Theorem 1. Let \( m : \mathcal{M} \to E^* \) be a topological measure in \((S, \rho)\). If \( f : S \to E^n (C^n) \) is relatively continuous on a set \( A \subseteq \mathcal{M} \), it is \( \mathcal{M} \)-measurable on \( A \).

Proof. Let \( B \) be open in \( E^n (C^n) \). By Lemma 2,
\[ A \cap f^{-1} [B] \]
is open in \((A, \rho)\). Hence by Theorem 4 of Chapter 3, §12,
\[ A \cap f^{-1} [B] = A \cap U \]
for some open set \( U \) in \((S, \rho)\).
Now, by assumption, \( A \) is in \( \mathcal{M} \). So is \( U \), as \( \mathcal{M} \) is topological (\( \mathcal{M} \supseteq \mathcal{G} \)). Hence 
\[
A \cap f^{-1}[B] = A \cap U \in \mathcal{M}
\]
for any open \( B \subseteq E^n \) (\( C^n \)). The result follows by Lemma 1. \( \square \)

**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.,
\[
f[A] \subseteq \mathcal{T}
\]
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 2** (Luzin). Let \( m : \mathcal{M} \rightarrow E^* \) be a strongly regular measure in \((S, \rho)\).

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 
\[
m(A - F) < \varepsilon
\]
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 
\[
H = A - Q, \ mQ = 0;
\]
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
\[
(1) \quad \text{each open set } B \neq \emptyset \text{ in } T \text{ is the union of a subsequence of } \{G_k\}.
\]

Now let \( \varepsilon > 0 \), and set
\[
S_k = S \cap f^{-1}[G_k] = f^{-1}[G_k], \quad k = 1, 2, \ldots,
\]
By Corollary 2 in §2, \( S_k \in \mathcal{M} \). As \( m \) is strongly regular, we find for each \( S_k \) an open set 
\[
U_k \supseteq S_k,
\]
with \( U_k \in \mathcal{M} \) and
\[
m(U_k - S_k) < \frac{\varepsilon}{2^{k+1}}.
\]
Let \( B_k = U_k - S_k, D = \bigcup_k B_k; \) so \( D \in \mathcal{M} \) and
\[
(2) \quad mD \leq \sum_k mB_k \leq \sum_k \frac{\varepsilon}{2^{k+1}} \leq \frac{1}{2}\varepsilon
\]
and
\[
(2') \quad U_k - B_k = S_k = f^{-1}[G_k].
\]
As \( D = \bigcup B_k \), we have
\[
(\forall k) \quad B_k - D = B_k \cap (-D) = \emptyset.
\]
Hence by \((2')\),
\[
(\forall k) \quad f^{-1}[G_k] \cap (-D) = (U_k - B_k) \cap (-D) = (U_k \cap (-D)) - (B_k \cap (-D)) = U_k \cap (-D).
\]
Combining this with \((1)\), we have, for each open set \( B = \bigcup G_k \) in \( T \),
\[
(3) \quad f^{-1}[B] \cap (-D) = \bigcup_i f^{-1}[G_k] \cap (-D) = \bigcup_i U_k \cap (-D).
\]
Since the \( U_k \) 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 
\[
H - D = A - Q - D
\]
(since we actually substituted \( S \) for \( H \) in the course of the proof). As \( mQ = 0 \) and \( mD < \frac{1}{2}\varepsilon \) by \((2)\),
\[
m(H - D) < mA - \frac{1}{2}\varepsilon.
\]
Finally, as \( m \) is strongly regular and \( H - D \in \mathcal{M} \), there is a closed \( \mathcal{M} \)-set 
\[
F \subseteq H - D \subseteq A
\]
such that
\[
m(H - D - F) < \frac{1}{2}\varepsilon.
\]
Since \( f \) is relatively continuous on \( H - D \), it is surely so on \( F \). Moreover, 
\[
A - F = (A - (H - D)) \cup (H - D - F);
\]
so
\[
m(A - F) \leq m(A - (H - D)) + m(H - D - F) < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.
\]
\[\text{§3. Measurable Functions in } (S, \mathcal{M}, m)\]
This completes the proof. □

*Lemma 3. Given \([a, b] \subseteq E^1\) and disjoint closed sets \(A, B \subseteq (S, \rho)\), there always is a continuous map \(g: S \to [a, b]\) such that \(g = a\) on \(A\) and \(g = b\) on \(B\).

**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

\[
g(x) = a + \frac{(b-a)\rho(x, A)}{\rho(x, A) + \rho(x, B)}
\]

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). □

*Lemma 4 (Tietze). If \(f: (S, \rho) \to E^*\) is relatively continuous on a closed set \(F \subseteq S\), there is a function \(g: S \to E^*\) such that \(g = f\) on \(F\),

\[
\inf g[S] = \inf f[F], \quad \sup g[S] = \sup f[F],
\]

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

\[
A = F \left( f \leq \frac{1}{3} \right) = F \cap f^{-1} \left( \left( 0, \frac{1}{3} \right) \right)
\]

and

\[
B = F \left( f \geq \frac{2}{3} \right) = F \cap f^{-1} \left( \left( \frac{2}{3}, 1 \right) \right).
\]

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 \(g_1: S \to [0, \frac{1}{3}]\), with \(g_1 = 0\) on \(A\), and \(g_1 = \frac{1}{3}\) on \(B\). Set \(f_1 = f - g_1\) on \(F\); so \(|f_1| \leq \frac{2}{3}\), 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 \(0 \leq g_2 \leq \frac{1}{3} \cdot \frac{2}{3}\) 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 \(\{g_n\}\) and \(\{f_n\}\) of real functions such that each \(g_n\) is continuous on \(S\),

\[
0 \leq g_n \leq \frac{1}{3} \left( \frac{2}{3} \right)^{n-1},
\]

and \(f_n = f_{n-1} - g_n\) is defined and continuous on \(F\), with

\[
0 \leq f_n \leq \left( \frac{2}{3} \right)^n
\]

there \((f_0 = f)\).

We claim that

\[
g = \sum_{n=1}^{\infty} g_n
\]

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,

\[
\left| f - \sum_{n=1}^{\infty} g_n \right| \leq \left( \frac{2}{3} \right)^n \to 0
\]

on \(F\) (why?); so \(f = g\) on \(F\). Moreover,

\[
0 \leq g_1 \leq g \leq \sum_{n=1}^{\infty} \frac{1}{3} \left( \frac{2}{3} \right)^n = 1 \quad \text{on} \quad S.
\]

Hence \(\inf g[S] = 0\) and \(\sup g[S] = 1\), as required.

Now assume

\[
\inf f[F] = a < \sup f[F] = b \quad (a, b \in E^1).
\]

Set

\[
h(x) = \frac{f(x) - a}{b - a}
\]

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

\[
g_0 = h = \frac{f - a}{b - a}
\]

on \(F\), \(\inf g_0[S] = 0\), and \(\sup g_0[S] = 1\). Set

\[
a + (b - a)g_0 = g.
\]

Then \(g\) is the required function. (Verify!)

Finally, if \(a, b \in E^* \quad (a < b)\), all reduces to the *bounded* case by considering \(H(x) = \arctan(f(x))\). □

*Theorem 3 (Fréchet).* Let \(m: \mathcal{M} \to E^*\) be a strongly regular measure in \((S, \rho)\). If \(f: S \to E^*\) (\(E^n, C^n\)) is \(m\)-measurable on \(A\), then

\[
f = \lim_{i \to \infty} f_i \quad (a.e.(m)) \quad \text{on} \quad A
\]
for some sequence of maps \( f_i \), continuous on \( S \). (We assume \( E^* \) to be metrized as in Lemma 4.)

**Proof.** We consider \( f : S \to E^* \) (the other cases reduce to \( E^1 \) via components).

Taking \( \varepsilon = \frac{1}{i} \left( i = 1, 2, \ldots \right) \) in Theorem 2, we obtain for each \( i \) a closed \( \mathcal{M} \)-set \( F_i \subseteq A \) such that

\[
m(A - F_i) < \frac{1}{i}
\]

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 \( \bigcup_{k=i}^{\infty} F_k \)).

Now, Lemma 4 yields for each \( i \) a continuous map \( f_i : S \to E^* \) such that \( f_i = f \) on \( F_i \). We complete the proof by showing that \( f_i \to f \) (pointwise) on the set

\[
B = \bigcup_{i=1}^{\infty} F_i
\]

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

\[
(\forall i > i_0) \quad f_i(x) = f(x),
\]

and so \( f_i(x) \to f(x) \) for \( x \in B \). As \( F_i \subseteq B \), we get

\[
m(A - B) \leq m(A - F_i) < \frac{1}{i}
\]

for all \( i \). Hence \( m(A - B) = 0 \), and all is proved.

\( \square \)

**Problems on Measurable Functions in \( (S, \mathcal{M}, m) \)**

1. Fill in all proof details in Corollaries 1 to 4.

2. Prove Theorems 1 and 2 in §1 and Theorem 2 in §2, for almost measurable functions.

3. Prove Note 2.

   [Hint: If \( f : S \to E^* \) is \( \mathcal{M} \)-measurable on \( B = A - Q \) (\( mQ = 0 \), \( Q \subseteq A \)), then \( A = B \cup Q \) and

   \[
   (\forall a \in E^*) \quad A(f > a) = B(f > a) \cup Q(f > a).
   \]

   Here \( B(f > a) \in \mathcal{M} \) by Theorem 1 in §2, and \( Q(f > a) \in \mathcal{M} \) if \( m \) is complete. For \( f : S \to E^* \) (\( C^* \)), use Theorem 2 of §1.)

4. Show that if \( m \) is complete and \( f : S \to (T, \rho') \) 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.]

5. Prove Theorem 1 for \( f : S \to (T, \rho') \), assuming that \( f[A] \) is separable in \( T \).

6. Given \( f_n \to f \) (a.e.) on \( A \), prove that \( f_n \to g \) (a.e.) on \( A \) if \( f = g \) (a.e.) on \( A \).

7. Given \( A \in \mathcal{M} \) in \( (S, \mathcal{M}, m) \), let \( m_A \) be the restriction of \( m \) to \( \mathcal{M}_A = \{X \in \mathcal{M} | X \subseteq A\} \).

   Prove that

   (i) \((A, \mathcal{M}_A, m_A)\) is a measure space (called a *subspace* of \((S, \mathcal{M}, m)\));

   (ii) if \( m \) is complete, topological, \( \sigma \)-finite or (strongly) regular, so is \( m_A \).

8. (i) Show that if \( D \subseteq K \subseteq (T, \rho') \), then the closure of \( D \) in the subspace \((K, \rho')\) is \( K \cap D' \), where \( D' \) is the closure of \( D \) in \((T, \rho')\).

   [Hint: Use Problem 11 in Chapter 3, §16.]

   (ii) Prove that if \( B \subseteq K \) and if \( B \) is separable in \((T, \rho')\), it is so in \((K, \rho')\).

   [Hint: Use Problem 7 from §1.]


10. Simplify the proof of Theorem 2 for the case \( mA < \infty \).

    [Outline: (i) First, let \( f \) be elementary, with \( f = a_1 \) on \( A_1 \in \mathcal{M} \), \( A = \bigcup \{A_i\} \) (disjoint), \( \sum a_i \), \( mA = \infty \).

    Given \( \varepsilon > 0 \),

    \[
    (\exists \delta > 0) \quad mA - \sum_{i=1}^{n} a_i < \frac{1}{2} \varepsilon.
    \]

    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 §1, \( f_n \to f \) (uniformly) on \( H \) for some elementary maps \( f_n \). By (i), each \( f_n \) is relatively continuous on a closed \( \mathcal{M} \)-set \( F \), 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.]

11. Given \( f_n : S \to (T, \rho') \), \( n = 1, 2, \ldots \), we say that

   (i) \( f_n \to f \) almost uniformly on \( A \subseteq S \) iff

   \[
   (\forall \delta > 0) \exists D \in \mathcal{M} | mD < \delta \quad f_n \to f \text{ (uniformly)} \quad \text{on } A - D;
   \]

   (ii) \( f_n \to f \) in measure on \( A \) iff

   \[
   (\forall \delta, \sigma > 0) \exists k \quad (\forall n > k) \exists D_n \in \mathcal{M} | mD_n < \delta \quad \rho'(f, f_n) < \sigma \quad \text{on } A - D_n.
   \]
Prove the following.
(a) $f_n \to f$ (uniformly) implies $f_n \to f$ (almost uniformly), and the latter implies both $f_n \to f$ (in measure) and $f_n \to f$ (a.e.).
(b) Given $f_n \to f$ (almost uniformly), we have $f_n \to g$ (almost uniformly) iff $f = g$ (a.e.); similarly for convergence in measure.
(c) If $f$ and $f_n$ are $\mathcal{M}$-measurable on $A$, then $f_n \to f$ in measure on $A$ iff
$$(\forall \sigma > 0) \lim_{n \to \infty} \mu(A(f_n, f, f_n) \geq \sigma) = 0.$$

12. Assuming that $f_n \colon S \to (T, \rho')$ is $m$-measurable on $A$ for $n = 1, 2, \ldots$, that $m A < \infty$, and that $f_n \to f$ (a.e.) on $A$, prove the following.
(i) Lebesgue’s theorem: $f_n \to f$ (in measure) on $A$ (see Problem 11).
(ii) Egorov’s theorem: $f_n \to f$ (almost uniformly) on $A$.

[Outline: (i) $f_n$ and $f$ are $\mathcal{M}$-measurable on $H = A - Q$, $m Q = 0$ (Corollary 1), with $f_n \to f$ (pointwise) on $H$. For all $i, k$, set $H_i(k) = \bigcap_{n=i}^{\infty} H_1(f_n, f) < \frac{1}{k} \in \mathcal{M}$
by Problem 6 in §1. Show that $(\forall k) H_i(k) \subset H$; hence
$$\lim_{k \to \infty} m H_i(k) = m H = m A < \infty;$$
so
$$(\forall \delta > 0) (\forall k) (\exists i_k) m(A - H_{i_k}(k)) < \frac{\delta}{2^k}.$$"
(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
\[
\sum |a_i| \, mA_i < \infty
\]
implies
\[
(\forall i) \quad |a_i| \, mA_i < \infty;
\]
so \(a_i = 0\) if \(mA_i = \infty\), and \(mA_i = 0\) if \(|a_i| = \infty\). Thus by our conventions, all “bad” terms \(a_i \, mA_i\) *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,
\[
(\forall i) \quad A_i = \bigcup_k (A_i \cap B_k) \quad (disjoint);
\]
so
\[
(\forall i) \quad a_i \, mA_i = \sum_k a_i \, m(A_i \cap B_k),
\]
and hence (see [Theorem 2 of Chapter 7, §2, and Problem 11 there])
\[
\sum a_i \, mA_i = \sum_k \sum_i a_i \, m(A_i \cap B_k) = \sum_k \sum_i b_k \, m(A_i \cap B_k) = \sum_k b_k \, mB_k.
\]
(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 1.** Let \(f: S \to E\) 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
\[
\left| \int_B f \right| \leq \int_B |f| \leq \int_A |f|.
\]

(iii) The set \(B = A(f \neq 0)\) is \(\sigma\)-finite (Definition 4 in Chapter 7, §5), and
\[
\int_A f = \int_B f.
\]

(iv) If \(f = a\) (constant) on \(A\),
\[
\int_A f = a \cdot mA.
\]

(v) \(\int_A |f| = 0\) iff \(f = 0\) a.e. on \(A\).

(vi) If \(mQ = 0\), then
\[
\int_A f = \int_{A-Q} f
\]
(so we may neglect sets of measure 0 in integrals).

(vii) For any \(k\) in the scalar field of \(E\), \(kf\) is elementary and integrable, and
\[
\int_A kf = k \int_A f.
\]

Note that if \(f\) is scalar valued, \(k\) may be a vector. If \(E = E^\ast\), we assume \(k \in E^1\).

**Proof.**

(i) By Note 1, \(|f| = |a_i| = \infty\) only on those \(A_i\) with \(mA_i = 0\). Let \(Q\) be the union of all such \(A_i\). Then \(mQ = 0\) 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 \(|f| = |a_i|\) on \(B \cap A_i \subseteq A_i\).

Also,
\[
\sum |a_i| \, m(B \cap A_i) \leq \sum |a_i| \, mA_i < \infty.
\]
(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 \(mA_i = \infty\). Thus \(f \neq 0\) on \(A_i\) only if \(mA_i < \infty\). Let \(\{A_{i_k}\}\) be the subsequence of those \(A_i\) on which \(f \neq 0\); so
\[
(\forall k) \quad mA_{i_k} < \infty.
\]
Also,
\[
B = A(f \neq 0) = \bigcup_k A_{i_k} \in \mathcal{M} \quad (\sigma\text{-finite}!).
\]

By (ii), \(f\) is elementary and integrable on \(B\). Also,
\[
\int_B f = \sum_k a_{i_k} \, mA_{i_k},
\]
while
\[
\int_A f = \sum_i a_i \, mA_i.
\]
These sums differ only by terms with $a_i = 0$. Thus (iii) follows.

The proof of (iv)–(vii) is left to the reader. $\square$

**Note 3.** If $f: S \to E^*$ is elementary and sign-constant on $A$, we also allow that

$$\int_A f = \sum_i a_i m A_i = \pm \infty.$$ 

Thus here $\int_A f$ exists even if $f$ is not integrable. Apart from claims of integrability and $\sigma$-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 $E^1 \setminus R$; see Chapter 4, §1, Example (c). Let $A = [0, 1]$.

By Corollary 1 in Chapter 7, §8, $A \cap R \in M^*$ and $m(A \cap R) = 0$. Also, $A - R \in M^*$.

Thus $\{A \cap R, A - R\}$ is an $M^*$-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

$$\int_A f = 1 \cdot m(A \cap R) + 0 \cdot m(A - R) = 0.$$ 

Thus $f$ is L-integrable (even though it is nowhere continuous).

**Theorem 1** (additivity).

(i) If $f: S \to E$ is elementary and integrable or elementary and nonnegative on $A \in M$, then

$$\int_A f = \sum_k \int_{B_k} f$$ 

for any $M$-partition $\{B_k\}$ of $A$.

(ii) If $f$ is elementary and integrable on each set $B_k$ of a finite $M$-partition

$$A = \bigcup_k B_k,$$

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.

§4. Integration of Elementary Functions

Thus for each $k$, we can fix an $M$-partition $B_k = \bigcup_i A_{ki}$, with $f$ constant $(f = a_k)$ on $A_{ki}$, $i = 1, 2, \ldots$. Then

$$A = \bigcup_k B_k = \bigcup_i A_{ki}$$

is an $M$-partition of $A$ into the disjoint sets $A_{ki} \in M$.

Now, by definition,

$$\int_{B_k} f = \sum_i a_{ki} m A_{ki}$$

and

$$\int_A f = \sum_k \sum_i a_{ki} m A_{ki} = \sum_k \left( \sum_i a_{ki} m A_{ki} \right) = \sum_k \int_{B_k} f$$

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

$$\sum_i |a_{ki}| m A_{ki} < \infty$$

(by integrability); hence

$$\sum_{k=1}^n \sum_i |a_{ki}| m A_{ki} < \infty.$$ 

This means, however, that $f$ is elementary and integrable on $A$, and so clause (ii) follows. $\square$

**Caution.** Clause (ii) fails if the partition $\{B_k\}$ is infinite.

**Theorem 2.**

(i) If $f, g: S \to E^*$ are elementary and nonnegative on $A$, then

$$\int_A (f + g) = \int_A f + \int_A g.$$ 

(ii) If $f, g: S \to E$ are elementary and integrable on $A$, so is $f \pm g$, and

$$\int_A (f \pm g) = \int_A f \pm \int_A g.$$ 

**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 $M$-partition of $A$, say, $f = a_i$ and $g = b_i$ on $A_i \in M$; so

$$f \pm g = a_i \pm b_i \text{ on } A_i, \quad i = 1, 2, \ldots$$
In case (i), \( f, g \geq 0 \); so integrability is irrelevant by Note 3, and formula (1) yields
\[
\int_A (f + g) = \sum_i (a_i + b_i) mA_i = \sum_i a_i mA_i + \sum_i b_i mA_i = \int_A f + \int_A g.
\]

In (ii), we similarly obtain
\[
\sum_i |a_i \pm b_i| mA_i \leq \sum_i |a_i| mA_i + \sum_i |b_i| mA_i < \infty.
\]
(Why?) Thus \( f \pm g \) is elementary and integrable on \( A \). As before, we also get
\[
\int_A (f \pm g) = \int_A f \pm \int_A g,
\]
simply by rules for addition of convergent series. (Verify!) \( \square \)

**Note 4.** As we know, the *characteristic function* \( C_B \) of a set \( B \subseteq S \) is defined
\[
C_B(x) = \begin{cases} 1, & x \in B, \\ 0, & x \in S - B. \end{cases}
\]
If \( g: S \to E \) is elementary on \( A \), so that
\[
g = a_i \text{ on } A_i, \ 1, 2, \ldots,
\]
for some \( \mathcal{M} \)-partition
\[
A = \bigcup A_i,
\]
then
\[
g = \sum_i a_i C_{A_i} \text{ on } A.
\]
(This sum always exists for *disjoint* sets \( A_i \). Why?) We shall often use this notation.

If \( m \) is Lebesgue measure in \( E^1 \), the integral
\[
\int_A g = \sum_i a_i mA_i
\]
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 \( a_i mA_i \) is the area of a rectangle with base \( A_i \) and altitude \( a_i \). (We assume the \( A_i \) to be

\[\text{Figure 33}\]

\[\int_A g = \sum_i a_i mA_i,\]
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 2.**

If
\[
f = \sum_i a_i C_{A_i}, \ (a_i \in E^*)
\]
on
\[
A = \bigcup_i A_i, \ (A_i \in \mathcal{M}),
\]
we set
\[
(3) \quad \int_A f = \int_A f^+ - \int_A f^-,
\]
with
\[
f^+ = f \vee 0 \geq 0 \text{ and } f^- = (-f) \vee 0 \geq 0;
\]

see §2.

By Theorem 2 in §2, \( f^+ \) and \( f^- \) are elementary and nonnegative on \( A \); so
\[
\int_A f^+ \text{ and } \int_A f^-
\]
are defined by Note 3, and so is
\[
\int_A f = \int_A f^+ - \int_A f^-
\]
by our conventions \((2*)\) in Chapter 4, §4.

We shall have use for formula (3), even if
\[
\int_A f^+ = \int_A f^- = \infty;
\]
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
\[
\int_A f^+ = a_i^+ \quad \text{and} \quad \int_A f^- = a_i^- \quad \text{on} \quad A_i,
\]
where
\[
a_i^+ = \max(a_i, 0) \quad \text{and} \quad a_i^- = \max(-a_i, 0).
\]
Thus
\[
\int_A f^+ = \sum a_i^+ \cdot m_A_i \quad \text{and} \quad \int_A f^- = \sum a_i^- \cdot m_A_i,
\]
so that
\[
\int_A f = \int_A f^+ - \int_A f^- = \sum a_i^+ \cdot m_A_i - \sum a_i^- \cdot m_A_i.
\]
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
\[
\int_A f = \sum (a_i^+ - a_i^-) m_A_i = \sum a_i m_A_i
\]
for \( a_i^+ - a_i^- = a_i \). Thus formulas (3) and (4) agree with our previous definitions.\(^3\)

Problems on Integration of Elementary Functions

1. Verify Note 2.

1' Prove Corollary 1(iv)–(vii).

2. Prove that \( \int_A f = 0 \) if \( mA = 0 \) or \( f = 0 \) on \( A \). Disprove the converse by examples.

3. Find a primitive \( F \) for \( f = C_R \) in our example. Show that
\[
\int_{[0,1]} f \, dm = F(1) - F(0).
\]

4. Fill in the proof details in Theorem 2.

[Hint: Use comparison test for series.]

\( \Rightarrow \) 5. Show that if \( f \) and \( g \) are elementary and nonnegative with \( f \geq g \) on \( A \), then
\[
\int_A f \geq \int_A g \geq 0.
\]

[Hint: As in Theorem 2, let
\[
f = \sum a_i C_{A_i} \quad \text{and} \quad g = \sum b_i C_{A_i}.
\]

Then \( f \geq g \geq 0 \) implies \( a_i \geq b_i \geq 0 \).

\( \Rightarrow \) 6. Prove that if \( f \) and \( g \) are elementary and (extended) real on \( A \), then
\[
\int_A (f \pm g) = \int_A f \pm \int_A g,
\]
provided
\( \quad (i) \ \int_A f \) or \( \int_A g \) is finite, or
\( \quad (ii) \ \int_A f, \int_A g, \) and \( \int_A f \pm \int_A g \) are all orthodox.

[Outline: As in Theorem 2, let
\[
f = \sum a_i C_{A_i} \quad \text{and} \quad g = \sum b_i C_{A_i},
\]
so
\[
f \pm g = a_i \pm b_i \quad \text{on} \quad A_i.
\]

Now, if
\[
\left| \int_A f \right| < \infty,
\]
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,
\[
\sum (a_i \pm b_i)^+ m_A_i = \infty
\]
iff
\[
\sum b_i^+ m_A_i = \infty.
\]

Thus if
\[
\int_A g = \pm \infty,
\]
then
\[
\int_A (f \pm g) = \int_A g = \pm \infty = \int_A f \pm \int_A g.
\]

If both
\[
\int_A f, \int_A g \neq \pm \infty,
\]
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
\[
\int_A f \pm \int_A g
\]
is orthodox, too. (Verify!)

7. Show that if \( f \) is elementary and nonnegative on \( A \) and
\[
\int_A f > p \in E^*,
\]
then there is an elementary and nonnegative map \( g \) on \( A \) such that
\[
\int_A f \geq \int_A g > p,
\]
g = 0 on \( A(f = 0) \), and
\[
f > g \quad \text{on} \quad A - A(f = 0).
\]

[Hints: Let
\[
B = A(f = \infty)
\]
and
\[
C = A - B;
\]
so \( B, C \in \mathcal{M} \) (Corollary 2 in §2). For all \( n > 0 \), define
\[
g_n = n \quad \text{on} \quad B
\]
and
\[
g_n = \left(1 - \frac{1}{n}\right)f \quad \text{on} \quad C;
\]
so \( g_n \) is elementary and nonnegative on \( A \) and
\[
f > g_n \quad \text{on} \quad A - A(f = 0). \quad \text{(Why?)}
\]

By Theorem 1 and Corollary 1(iv)(vii),
\[
\int_A g_n = \int_B g_n + \int_C g_n = \int_B n + \int_C \left(1 - \frac{1}{n}\right)f = n \cdot m_B + \left(1 - \frac{1}{n}\right)\int_C f.
\]
Deduce that
\[
\lim_{n \to \infty} \int_A g_n = \int_B f + \int_C f = \int_A f > p;
\]
so
\[
\exists \, n \quad \int_A g_n > p.
\]

Take \( g = g_n \) for that \( n \).

8. Show that if \( E = E^* \), Theorem 1(i) holds also if \( \int_A f \) is infinite but orthodox.

9.  
(i) Prove that if \( f \) is elementary and integrable on \( A \), so is \(-f\), and
\[
\int_A (-f) = -\int_A f.
\]
(ii) Show that this holds also if \( f \) is elementary and (extended) real and \( \int_A f \) is orthodox.

§5. Integration of Extended-Real Functions

We shall now define integrals for arbitrary functions \( f : S \to E^* \) in a measure space \((S, \mathcal{M}, m)\). We start with the case \( f \geq 0 \).

**Definition 1.**

Given \( f \geq 0 \) on \( A \in \mathcal{M} \), we define the upper and lower integrals,
\[
\overline{\int}_A f \quad \text{and} \quad \underline{\int}_A f
\]
of \( f \) on \( A \) (with respect to \( m \)) by
\[
(1') \quad \overline{\int}_A f = \int_A f dm = \inf_h \int_A h, \quad \text{over all elementary maps} \ h \geq f \text{ on } A, \text{ and}
\]
\[
(1'') \quad \underline{\int}_A f = \int_A f dm = \sup_g \int_A g, \quad \text{over all elementary and nonnegative maps} \ g \leq f \text{ on } A.
\]

If \( f \) is not nonnegative, we use \( f^+ = f \vee 0 \) and \( f^- = (-f) \wedge 0 \) (§2), and set
\[
(1) \quad \overline{\int}_A f = \int_A f dm = \int_A f^+ - \int_A f^- \quad \text{and}
\]
\[
\underline{\int}_A f = \int_A f dm = \int_A f^+ - \int_A f^-.
\]

By our conventions, these expressions are always defined. The integral 
\( \overline{\int}_A f \) (or \( \underline{\int}_A f \)) is called orthodox iff it does not have the form \( -\infty - \infty \) in (1), e.g., if \( f \geq 0 \) (i.e., \( f^- = 0 \)), or if \( \int_A f < \infty \). An unorthodox integral equals \( +\infty \).

We often write \( \overline{\int} f \) and call it simply the integral (of \( f \)), even if
\[
\overline{\int}_A f \neq \int_A f^2.
\]

“Classical” notation is \( \int_A f(x) \, dm(x) \).

---

1 Those who wish to consider measurable maps only should take Theorem 3 earlier.
2 There is good reason for identifying “integral” with “upper integral.”
Chapter 8. Measurable Functions. Integration

**Definition 2.**

The function $f$ is called **integrable** (or $m$-integrable, or Lebesgue integrable, with respect to $m$) on $A$, iff

$$\int_A f \, dm = \int_A f \, dm \neq \pm \infty.$$  

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 (1') and (1'') to be **simple step functions** only (see §9). It is inferior to Lebesgue theory.

The values of $\int_A f \, dm$ and $\int_A f \, dm$ 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 $\int_A f$ agrees with that of §4. For if $f \geq 0$, $f$ itself is the least of all elementary and nonnegative functions $h \geq f$ and the greatest of all elementary and nonnegative functions $g \leq f$.

Thus by Problem 5 in §4,

$$\int_A f = \min_{h \geq f} \int_A h = \max_{g \leq f} \int_A g,$$

i.e.,

$$\int_A f = \int_A f = \int_A f.$$

If, however, $f \not\geq 0$, this follows by **Definition 2** in §4. This also shows that for **elementary** and (extended) **real** maps,

$$\int_A f = \int_A f \text{ always.}$$

(See also Theorem 3.)

## §5. Integration of Extended-Real Functions

**Note 2.** By **Definition 1,**

$$\int_A f \leq \int_A f \text{ always.}$$

For if $f \geq 0$, then for any elementary and nonnegative maps $g,h$ with $g \leq f \leq h$, we have

$$\int_A g \leq \int_A h$$

by Problem 5 in §4. Thus

$$\int_A f = \sup_g \int_A g$$

is a lower bound of all such $\int_A h$, and so

$$\int_A f \leq \text{glb} \int_A h = \int_A f.$$

In the general formula (1), too,

$$\int_A f \leq \int_A f,$$

since

$$\int_A f^+ \leq \int_A f^+ \text{ and } \int_A f^- \leq \int_A f^-.$$

**Theorem 1.** For any functions $f,g: S \to E^*$ and any set $A \in \mathcal{M}$, we have the following results.

(a) If $f = a$ (constant) on $A$, then

$$\int_A f = \int_A f = a \cdot mA.$$

(b) If $f = 0$ on $A$ or $mA = 0$, then

$$\int_A f = \int_A f = 0.$$

(c) If $f \geq g$ on $A$, then

$$\int_A f \geq \int_A g \text{ and } \int_A f \geq \int_A g.$$

\footnote{Note that integrability is redundant here and in Theorem 2.}
(d) If \( f \geq 0 \) on \( A \), then
\[
\int_A f \geq 0 \quad \text{and} \quad \int_A f = 0.
\]
Similarly if \( f \leq 0 \) on \( A \).

(e) If \( 0 \leq p < \infty \), then
\[
\int_A pf = p \int_A f \quad \text{and} \quad \int_A pf = p \int_A f.
\]

(e') We have
\[
\int_A (-f) = -\int_A f \quad \text{and} \quad \int_A (-f) = -\int_A f
\]
if one of the integrals involved in each case is orthodox. Otherwise,
\[
\int_A (-f) = \infty = \int_A f \quad \text{and} \quad \int_A (-f) = \infty = \int_A f.
\]

(f) If \( f \geq 0 \) on \( A \) and \( A \supseteq B \), \( B \in \mathcal{M} \), then
\[
\int_A f \geq \int_B f \quad \text{and} \quad \int_A f \geq \int_B f.
\]

(g) We have
\[
\left| \int_A f \right| \leq \int_A |f| \quad \text{and} \quad \int_A f \leq \int_A |f|
\]
(but not \( \int_A f \leq \int_A |f| \) in general).

(h) If \( f \geq 0 \) on \( A \) and \( \int_A f = 0 \) (or \( f \leq 0 \) and \( \int_A f = 0 \)), 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 \( f \geq g \geq 0 \) on \( A \).

Take any elementary and nonnegative map \( H \geq f \) on \( A \). Then \( H \geq g \) as well; so by definition,
\[
\int_A g = \inf_{h \geq g} \int_A h \leq \int_A H.
\]
Thus
\[
\int_A f \leq \int_A H
\]
for any such \( H \). Hence also
\[
\int_A g \leq \inf_{H \geq f} \int_A H = \int_A f.
\]

Similarly,
\[
\int_A f \geq \int_A g
\]
if \( f \geq g \geq 0 \).

In the general case, \( f \geq g \) implies \( f^+ \geq g^+ \) and \( f^- \leq g^- \). (Why?)

Thus by what was proved above,
\[
\int_A f^+ \geq \int_A g^+ \quad \text{and} \quad \int_A f^- \leq \int_A g^-.
\]
Hence
\[
\int_A f^+ - \int_A f^- \geq \int_A g^+ - \int_A g^-;
\]
i.e.,
\[
\int_A f \geq \int_A g.
\]

Similarly, one obtains
\[
\int_A f \geq \int_A g.
\]

(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 \( h \geq f \) on \( A \).
By Corollary 1(vii) and Note 3 of §4,
\[ \int_A ph = p \int_A h \]
for any such h. Hence
\[ \int_A pf = \inf_h \int_A ph = \inf_h p \int_A h = p \int_A f. \]
Similarly,
\[ \int_A pf = p \int_A f. \]

The general case reduces to the case \( f \geq 0 \) by formula (1).

(e') Assertion (e') follows from (1) since
\[ (-f)^+ = f^-, \quad (-f)^- = f^+, \]
and \(- (x - y) = y - x \) if \( x - y \) is orthodox. (Why?)

(f) Take any elementary and nonnegative map
\[ h \geq f \geq 0 \] on \( A \).

By Corollary 1(ii) and Note 3 of §4,
\[ \int_B h \geq \int_A h \]
for any such h. Hence
\[ \int_B f = \inf_h \int_B h \leq \inf_h \int_A h = \int_A f. \]

Similarly for \( \int_A f \).

(g) This follows from (c) and (e') since \( f \leq |f| \) implies
\[ \int_A |f| \geq \int_A f \geq - \int_A f. \]
For (h) and later work, we need the following lemmas.
Chapter 8. Measurable Functions. Integration

and set \( v = q - u \).

As the lemma holds for positive functions, we find elementary and nonnegative maps \( h' \) and \( h'' \), with
\[
h' \geq f^+, \quad h'' \leq f^-, \quad \int h' < u < \infty \text{ and } \int h'' > -v.
\]
Let \( h = h' - h'' \). Then
\[
h \geq f^+ - f^- = f,
\]
and by Problem 6 in §4,
\[
\int_A h = \int_A h' - \int_A h'' \quad \text{(for } \int_A h' \text{ is finite!)}
\]
Hence
\[
\int_A h > u + v = q,
\]
and clause (i) is proved in full.

Clause (ii) follows from (i) by Theorem 1(e’)
\[
\int_A f < \infty \quad \text{so } \int_A h = \int_A f.
\]

Proof. If \( f \geq 0, \) choose \( h \geq f \) as in Lemma 2. Let
\[
D = A(h > 0) \quad \text{and } A_n = A(h > \frac{1}{n});
\]
so
\[
D = \bigcup_{n=1}^{\infty} A_n \quad \text{(why?)},
\]
let
\[
q_n = \int_A f + \frac{1}{n}, \quad n = 1, 2, \ldots;
\]
so
\[
q_n \to \int_A f < q_n.
\]

By Lemma 1, for each \( n \) there is an elementary and (extended) real (hence measurable) map \( h_n \geq f \) on \( A \), with
\[
q_n \geq \int_A h_n \geq \int_A f.
\]

Let
\[
h = \inf_n h_n \geq f.
\]

By Lemma 1 in §2, \( h \) is \( \mathcal{M} \)-measurable on \( A \). Also,
\[
(\forall n) \quad q_n > \int_A h_n \geq \int_A h \geq \int_A f
\]
because of Theorem 1(c). Hence
\[
\int_A f = \lim_{n \to \infty} q_n \geq \int_A h \geq \int_A f
\]

as required.

Finally, if \( f = -\infty \), the same proof works with \( q_n = -n \). (Verify!)

Similarly, one finds a measurable map \( g \leq f \), with
\[
\int_A f = \int_A g.
\]

Proof of Theorem 1(h). If \( f \geq 0 \), choose \( h \geq f \) as in Lemma 2. Let
\[
D = A(h > 0) \quad \text{and } A_n = A(h > \frac{1}{n});
\]
so
\[
D = \bigcup_{n=1}^{\infty} A_n \quad \text{(why?)},
\]
and $D, A_n \in \mathcal{M}$ by Theorem 1 of §2. Also,

$$0 = \int_A f = \int_A h \geq \int_{A_n} \left(\frac{1}{n}\right) = \frac{1}{n} m A_n \geq 0.$$  Thus $(\forall n) m A_n = 0$. Hence

$$m D = m \bigcup_{n=1}^{\infty} A_n = m A(h > 0) = 0;$$

so $0 \leq f \leq h \leq 0$ (i.e., $f = 0$) a.e. on $A$.

The case $f \leq 0$ reduces to $(-f) \geq 0$. $\Box$

**Corollary 1.** If

$$\int_A |f| < \infty,$$

then $|f| < \infty$ a.e. on $A$, and $A(f \neq 0)$ is $\sigma$-finite.

**Proof.** By Lemma 1, fix an elementary and nonnegative $h \geq |f|$ with

$$\int_A h < \infty$$

(so $h$ is elementary and integrable).

Now, by Corollary 1(i)–(iii) in §4, our assertions apply to $h$, hence certainly to $f$. $\Box$

**Theorem 2** (additivity). Given $f: S \to E^*$ and an $\mathcal{M}$-partition $\mathcal{P} = \{B_n\}$ of $A \in \mathcal{M}$, we have

$$\int_A f = \sum_n \int_{B_n} f \quad \text{and} \quad \int_A f = \sum_n \int_{B_n} f,$$

provided

$$\int_A f = \left(\int_A f, \text{ respectively}\right)$$

is orthodox, or $\mathcal{P}$ is finite.

Hence if $f$ is integrable on each of finitely many disjoint $\mathcal{M}$-sets $B_n$, it is so on

$$A = \bigcup_n B_n,$$

and formulas (2)(a)(b) apply.

---

4 It suffices that $f$ be integrable on $A$ (apply the same proof to $f^+$ and $f^-$).
If 
\[ \int_A f \]
is orthodox, one of these sums must be finite, and so their difference may be rearranged to yield
\[ \int_A f = \sum_n \left( \int_{B_n} f^+ - \int_{B_n} f^- \right) = \sum_n \int_{B_n} f, \]
proving (a). Similarly for (b).

This rearrangement works also if \( P \) 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 (2\( ^* \)) of Chapter 4, §4. Thus all is proved. \( \Box \)

Corollary 2. If \( mQ = 0 \ (Q \in \mathcal{M}) \), then for \( A \in \mathcal{M} \)
\[ \int_{A-Q} f = \int_A f \text{ and } \int_{A-Q} f = \int_A f. \]

For by Theorem 2,
\[ \int_A f = \int_{A-Q} f + \int_{A\cap Q} f, \]
where
\[ \int_{A\cap Q} f = 0 \]
by Theorem 1(b).

Corollary 3. If \( \int_A f \) (or \( \int_X f \)) is orthodox, so is
\[ \int_X f \left( \int_X f \right) \]
whenever \( A \supseteq X, X \in \mathcal{M} \).

For if \( \int_A f^+, \int_A f^-, \int_A f^+, \text{ or } \int_A f^- \) is finite,
it remains so also when \( A \) is reduced to \( X \) (see Theorem 1(f)). Hence orthodoxy follows by formula (1).

\[ \text{Note 4.} \]
Given \( f: S \to E^* \), we can define two additive (by Theorem 2) set functions \( \overline{\pi} \) and \( \underline{\pi} \) by setting for \( X \in \mathcal{M} \)
\[ \overline{\pi} X = \int_X f \text{ and } \underline{\pi} X = \int_X f. \]
They are called, respectively, the upper and lower indefinite integrals of \( f \), also denoted by
\[ \int f \text{ and } \int f \]
(or \( \overline{\pi} f \) and \( \underline{\pi} f \)).

By Theorem 2 and Corollary 3, if
\[ \int_A f \]
is orthodox, then \( \overline{\pi} \) is \( \sigma \)-additive (and semifinite) when restricted to \( \mathcal{M} \)-sets \( X \subseteq A \). Also,
\[ \overline{\pi} \emptyset = \underline{\pi} \emptyset = 0 \]
by Theorem 1(b).

Such set functions are called signed measures (see Chapter 7, §11). In particular, if \( f \geq 0 \) on \( S \), \( \overline{\pi} \) and \( \underline{\pi} \) are \( \sigma \)-additive and nonnegative on all of \( \mathcal{M} \), hence measures on \( \mathcal{M} \).

Theorem 3. If \( f: S \to E^* \) is \( m \)-measurable (Definition 2 in §3) on \( A \), then
\[ \int_A f = \int_A f. \]

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
\[ A_n = A((1 + \varepsilon)^n \leq f < (1 + \varepsilon)^{n+1}), \ n = 0, \pm 1, \pm 2, \ldots. \]
Clearly, these are disjoint \( \mathcal{M} \)-sets (Theorem 1 of §2), and
\[ A = A_0 \cup A_\infty \cup \bigcup_{n=-\infty}^{\infty} A_n. \]

Thus, setting
\[ g = \begin{cases} 0 & \text{on } A_0, \\ \infty & \text{on } A_\infty, \text{ and} \\ (1 + \varepsilon)^n & \text{on } A_n (n = 0, \pm 1, \pm 2, \ldots) \end{cases} \]
and
\[ h = (1 + \varepsilon)g \text{ on } A, \]
we obtain two elementary and nonnegative maps, with
\[ g \leq f \leq h \text{ on } A. \] (Why?)

By Note 1,
\[ \int_A g = \int_A g. \]

Now, if \( \int_A g = \infty \), then
\[ \int_A f \geq \int_A g \]
yields
\[ \int_A f = \infty. \]

If, however, \( \int_A g < \infty \), then
\[ \int_A h = \int_A (1 + \varepsilon)g = (1 + \varepsilon)\int_A g < \infty; \]
so \( g \) and \( h \) are elementary and integrable on \( A \). Thus by Theorem 2(ii) in §4,
\[ \int_A h - \int_A g = \int_A (h - g) = \int_A ((1 + \varepsilon)g - g) = \varepsilon \int_A g. \]

Moreover, \( g \leq f \leq h \) implies
\[ \int_A g \leq \int_A f \leq \int_A h; \]
so
\[ \left| \int_A f - \int_A g \right| \leq \int_A h - \int_A g \leq \varepsilon \int_A g. \]
As \( \varepsilon \) is arbitrary, all is proved for \( f \geq 0 \).

The case \( f \geq 0 \) now follows by formula (1), since \( f^+ \) and \( f^- \) are \( \mathcal{M} \)-measurable (Theorem 2 in §2).

§5. Integration of Extended-Real Functions

Problems on Integration of Extended-Real Functions

1. Using the formulas in (1) and our conventions, verify that
   (i) \( \overline{\int_A f} = +\infty \) iff \( \overline{\int_A f^+} = \infty; \)
   (ii) \( \overline{\int_A f} = \infty \) iff \( \overline{\int_A f^+} = \infty; \) and
   (iii) \( \overline{\int_A f} = -\infty \) iff \( \overline{\int_A f^-} = \infty \) and \( \overline{\int_A f^+} < \infty. \)
   (iv) Derive a condition similar to (iii) for \( \underline{\int_A f} = -\infty. \)

2. Fill in the missing proof details in Theorems 1 to 3 and Lemmas 1 and 2.

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
   \[ \underline{\int_A f} = \infty. \]
   As Lemmas 1 and 2 surely hold for nonnegative functions, fix a measurable \( F \leq f^+ \) \((F \geq 0), \) with
   \[ \underline{\int_A F} = \underline{\int_A f^+} = \infty. \]
   Arguing as in Theorem 3, find an elementary and nonnegative map \( g \leq F, \) with
   \[ (1 + \varepsilon)\int_A g = \int_A F = \infty; \]
   so \( \underline{\int_A g} = \infty \) and \( 0 \leq g \leq F \leq f^+ \) on \( A. \)
   Let \( A_+ = A(F > 0) \in \mathcal{M} \)
   and \( A_0 = A(F = 0) \in \mathcal{M} \)
   (Theorem 1 in §2). On \( A_+, \)
   \[ g \leq F \leq f^+ = f \] (why?),
while on \( A_0, g = F = 0; \) so
   \[ \underline{\int_{A_0} g} = \int_{A_0} g = \infty \] (why?).
   Now redefine \( g = -\infty \) on \( A_0 \) (only). Show that \( g \) is then the required function.]

4. For any \( f: S \to E^* \), prove the following.
   (a) If \( \overline{\int_A f} < \infty \), then \( f < \infty \) a.e. on \( A. \)
   (b) If \( \overline{\int_A f} \) is orthodox and \( > -\infty \), then \( f > -\infty \) a.e. on \( A. \)
5. Integration of Extended-Real Functions

(i) If \( \int_A |f| < \infty \) and \( \int_A |g| < \infty \), then
\[
\int_A f - \int_A g \leq \int_A |f - g|
\]
and
\[
\int_A f - \int_A g \leq \int_A |f - g|.
\]

[Hint: Use Problems 5 and 6.]

8. Show that any signed measure \( \tau_f \) (Note 4) is the difference of two measures: \( \tau_f = \tau_{f+} - \tau_{f-} \).

§6. Integrable Functions. Convergence Theorems

I. Some important theorems apply to integrable functions.

Theorem 1 (linearity of the integral). If \( f, g: S \to E^* \) are integrable on a set \( A \in \mathcal{M} \) in \( (S, \mathcal{M}, m) \), so is
\[
pf + qg
\]
for any \( p, q \in E^1 \), and
\[
\int_A (pf + qg) = p \int_A f + q \int_A g;
\]
in particular,
\[
\int_A (f \pm g) = \int_A f \pm \int_A g.
\]

Proof. By Problem 5 in §5,
\[
\int_A f + \int_A g \geq \int_A (f + g) \geq \int_A f + \int_A g.
\]
(Here \( \int_A f, \int_A f, \int_A g, \text{ and } \int_A g \)
are finite by integrability; so all is orthodox.)
As
\[
\int_A f = \int_A f \text{ and } \int_A g = \int_A g.
\]
the inequalities turn into equalities, so that
\[ \int_A f + \int_A g = \int_A (f + g) = \int_A (f + g). \]

Using also Theorem 1(iii) from §5, we obtain the desired result for any \( p, q \in E^1 \). \( \square \)

**Theorem 2.** A function \( f : S \to E^* \) is integrable on \( A \) in \( (S, \mathcal{A}, m) \) iff
(i) it is \( m \)-measurable on \( A \), and
(ii) \( \int_A f \) (equivalently \( \int_A |f| \)) is finite.

**Proof.** If these conditions hold, \( f \) is integrable on \( A \) by Theorem 3 of §5.

Conversely, let
\[ \int_A f = \int \mathcal{A} f \not= \pm \infty. \]

Using Lemma 2 in §5, fix measurable maps \( g \) and \( h \) \( (g \leq f \leq h) \) on \( A \), with
\[ \int_A g = \int_A f = \int_A h \not= \pm \infty. \]

By Theorem 3 in §5, \( g \) and \( h \) are integrable on \( A \); so by Theorem 1,
\[ \int_A (h - g) = \int_A h - \int_A g = 0. \]

As
\[ h - g \geq h - f \geq 0, \]

we get
\[ \int_A (h - f) = 0, \]

and so by Theorem 1(h) of §5, \( h - f = 0 \) a.e. on \( A \).

Hence \( f \) is almost measurable on \( A \), and
\[ \int_A f \not= \pm \infty \]

by assumption. From formula (1), we then get
\[ \int_A f^+ \text{ and } \int_A f^- < \infty, \]

and hence
\[ \int_A |f| = \int_A (f^+ + f^-) = \int_A f^+ + \int_A f^+ < \infty \]

by Theorem 1 and by Theorem 2 of §2. Thus all is proved. \( \square \)

Simultaneously, we also obtain the following corollary.

**Corollary 1.** A function \( f : S \to E^* \) is integrable on \( A \) iff \( f^+ \) and \( f^- \) are.

**Corollary 2.** If \( f, g : S \to E^* \) are integrable on \( A \), so also are
\[ f \vee g, f \wedge g, |f|, \text{ and } kf \text{ for } k \in E^1, \]

with
\[ \int_A kf = k \int_A f. \]

Exercise!

For products \( fg \), this holds if \( f \) or \( g \) is bounded. In fact, we have the following theorem.

**Theorem 3** (weighted law of the mean). Let \( f \) be \( m \)-measurable and bounded on \( A \). Set
\[ p = \inf f[A] \text{ and } q = \sup f[A]. \]

Then if \( g \) is \( m \)-integrable on \( A \), so is \( fg \), and
\[ \int_A f|g| = c \int_A |g| \]

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,
\[ (\exists k \in E^1) \ |f| \leq k \]
on \( A \). Hence if \( \int_A |g| = 0 \),
\[ \left| \int_A f|g| \right| \leq \int_A |f| |g| \leq k \int_A |g| = 0; \]

so any \( c \in [p, q] \) yields
\[ \int_A f|g| = c \int_A |g| = 0. \]

If, however, \( \int_A |g| \not= 0 \), the number
\[ c = \left( \int_A f|g| \right) / \int_A |g| \]

is the required constant.

Moreover, as \( f \) and \( g \) are \( m \)-measurable on \( A \), so is \( fg \); and as
\[ \int_A fg \leq |c| \int_A |g| < \infty, \]
$fg$ 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

$$f(x) < c < f(y)$$

for some $x, y \in A$ (why?); hence by the Darboux property, $f(x_0) = c$ for some $x_0 \in A$.

If, however,

$$c \leq \inf f(A) = p,$$

then

$$(f - c)|g| \geq 0$$

and

$$\int_A (f - c)|g| = \int_A f|g| - c \int_A |g| = 0 \ (why?);$$

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 $m(A) = 0$). This also implies $c \in f(A) \in [p, q]$.

Proceed similarly in the case $c \geq q$. □

Corollary 3. If $f$ is integrable on $A \in M$, it is so on any $B \subseteq A (B \in M)$.

Proof. Apply Theorem 1(f) in §5, and Theorem 3 of §5, to $f^+$ and $f^−$. □

II. Convergence Theorems. If $f_n \to f$ on $A$ (pointwise, a.e., or uniformly), does it follow that

$$\int_A f_n \to \int_A f?$$

To give some answers, we need a lemma.

Lemma 1. If $f \geq 0$ on $A \in M$ and if

$$\int_A f > p \in E^*,$$

there is an elementary and nonnegative map $g$ on $A$ such that

$$\int_A g > p,$$

and $g < f$ on $A$ except only at those $x \in A$ (if any) at which

$$f(x) = g(x) = 0.$$

(We then briefly write $g < f$ on $A$.)

§6. Integrable Functions. Convergence Theorems

Proof. By Lemma 1 in §5, there is an elementary and nonnegative map $G \leq f$ on $A$, with

$$\int_A f \geq \int_A G > p.$$

For the rest, proceed as in Problem 7 of §4, replacing $f$ by $G$ there. □

Theorem 4 (monotone convergence). If $0 \leq f_n \not\to f$ (a.e.) on $A \in M$, i.e.,

$$0 \leq f_n \leq f_{n+1} \ (\forall n),$$

and $f_n \to f$ (a.e.) on $A$, then

$$\int_A f_n \not\to \int_A f.$$

Proof for $M$-measurable $f_n$ and $f$ on $A$. By Corollary 2 in §5, we may assume that $f_n \not\to f$ (pointwise) on $A$ (otherwise, drop a null set).

By Theorem 1(c) of §5, $0 \leq f_n \not\to f$ implies

$$0 \leq \int_A f_n \leq \int_A f,$$

and so

$$\lim_{n \to \infty} \int_A f_n \leq \int_A f.$$

The limit, call it $p$, exists in $E^*$, as $\{\int_A f_n\}$. It remains to show that

$$p \geq \int_A f = \int_A f.$$

(We know that

$$\int_A f = \int_A f,$$

by the assumed measurability of $f$; see Theorem 3 in §5.)

Suppose

$$\int_A f > p.$$

Then Lemma 1 yields an elementary and nonnegative map $g \subset f$ on $A$, with

$$p < \int_A g.$$

Let

$$A_n = A(f_n \geq g), \ n = 1, 2, \ldots$$

1 For the general case, see Problem 5.
Then $A_n \in \mathcal{M}$ and

$$A_n \uparrow A = \bigcup_{n=1}^{\infty} A_n.$$  

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

$$\int_A g = s A = \lim_{n \to \infty} s A_n = \lim_{n \to \infty} \int_{A_n} g.$$  

But as $g \leq f_n$ on $A_n$, we have

$$\int_{A_n} g \leq \int_{A_n} f_n \leq \int_A f_n.$$  

Hence

$$\int_A g = \lim_{n \to \infty} \int_{A_n} g \leq \lim_{n \to \infty} \int_A f_n = p,$$
contrary to $p < \int_A g$. This contradiction completes the proof. □

**Lemma 2 (Fatou).** If $f_n \geq 0$ on $A \in \mathcal{M}$ ($n = 1, 2, \ldots$), then

$$\int_A \liminf f_n \leq \liminf \int_A f_n.$$  

**Proof.** Let

$$g_n = \inf_{k \geq n} f_k, \quad n = 1, 2, \ldots;$$

so $f_n \geq g_n \geq 0$ and $\{g_n\} \uparrow$ on $A$. Thus by Theorem 4,

$$\int_A \lim g_n = \lim_{n \to \infty} \int_A g_n = \lim_{n \to \infty} \int_A g_n \leq \lim_{n \to \infty} \int_A f_n.$$  

But

$$\lim_{n \to \infty} g_n = \sup_n g_n = \sup_{k \geq n} f_k = \lim_{n \to \infty} f_n.$$  

Hence

$$\int_A \lim f_n = \int_A \lim g_n \leq \lim_{n \to \infty} \int_A f_n,$$

as claimed. □

**Theorem 5** (dominated convergence). Let $f_n : S \to E$ be $\mathcal{M}$-measurable on $A \in \mathcal{M}$ ($n = 1, 2, \ldots$). Let

$$f_n \to f \text{ (a.e.) on } A.$$  

Then

$$\lim_{n \to \infty} \int_A |f_n - f| = 0,$$

provided that there is a map $g : S \to E^1$ such that

$$\int_A g < \infty$$

and

$$\forall n \quad |f_n| \leq g \text{ a.e. on } A.$$  

**Proof.** Neglecting null sets, we may assume that

$$|f_n| \leq g < \infty$$
on $A$ and $f_n \to f$ (pointwise) on $A$; so $|f| \leq g$ and

$$|f_n - f| \leq |f_n| + |f| \leq 2g$$
on $A$. As $|f| < \infty$, we have

$$|f_n - f| \to 0$$
on $A$. Hence, setting

$$h_n = 2g - |f_n - f| \geq 0,$$

we get

$$2g = \lim_{n \to \infty} h_n = \lim h_n.$$  

We may also assume that $g$ is measurable on $A$. (If not, replace it by a measurable $G \geq g$, with

$$\int_A G = \int_A g < \infty,$$

by Lemma 2 in §5.) Then all

$$h_n = 2g - |f_n - f|$$

are measurable (even integrable) on $A$.  

\[ \text{§6. Integrable Functions. Convergence Theorems} \]
Thus by Lemma 2,
\[
\int_A 2g = \int_A \lim h_n \leq \lim \int_A (2g - |f_n - f|)
\]
\[
= \lim \left( \int_A 2g + \int_A (-|f_n - f|) \right)
\]
\[
= \int_A 2g + \lim \int_A |f_n - f|.
\]
(See Problems 5 and 8 in Chapter 2, §13.)

Canceling \(\int_A 2g\) (finite!), we have
\[
0 \leq -\lim \int_A |f_n - f|.
\]

Hence
\[
0 \geq \lim \int_A |f_n - f| \geq \lim \int_A |f_n - f| \geq 0,
\]
as \(|f_n - f| \geq 0\). This yields
\[
0 = \lim \int_A |f_n - f| = \lim \int_A |f_n - f| = \lim \int_A |f_n - f|,
\]
as required. \(\Box\)

**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
\[
\left| \int_A f_n - \int_A f \right| = \int_A (f_n - f) \leq \int_A |f_n - f| \to 0,
\]
i.e.,
\[
\int_A f_n \to \int_A f.
\]

Moreover, \(f\) is integrable on \(A\), being measurable (why?), with
\[
\int_A |f| \leq \int_A g < \infty.
\]
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”
\[
g \geq |f_n| \text{ with } \int_A g < \infty,
\]
even if \(f\) and the \(f_n\) are integrable.

**Example.**

Let \(m\) be Lebesgue measure in \(A = E^1\), \(f = 0\), and
\[
f_n = \begin{cases} 1 & \text{on } [n, n+1], \\ 0 & \text{elsewhere}. \end{cases}
\]

Then \(f_n \to f\) and \(\int_A f_n = 1\); so
\[
\lim_{n \to \infty} \int_A f_n = 1 \neq 0 = \int_A f.
\]

The trouble is that any
\[
g \geq f_n \quad (n = 1, 2, \ldots)
\]
would have to be \(\geq 1\) on \(B = [1, \infty); \) so
\[
\int_A g \geq \int_B g = 1 \cdot mB = \infty,
\]
instead of \(\int_A g < \infty\).

This example also shows that \(f_n \to f\) alone does not imply
\[
\int_A f_n \to \int_A f.
\]

**Theorem 6** (absolute continuity of the integral). Given \(f: S \to E\) with
\[
\int_X |f| < \infty
\]
and \(\varepsilon > 0\), there is \(\delta > 0\) such that
\[
\int_X |f| < \varepsilon
\]
whenever
\[
mX < \delta \quad (A \supseteq X, \ X \in \mathcal{M}).
\]

**Proof.** By Lemma 2 in §5, fix \(h \geq |f|\), measurable on \(A\), with
\[
\int_A h = \int_A |f| < \infty.
\]
Neglecting a null set, we assume that \(|h| < \infty\) on \(A\) (Corollary 1 of §5). Now, \((\forall n)\) set
\[
g_n(x) = \begin{cases} h(x), & x \in A_n = A(h < n), \\ 0, & x \in -A_n. \end{cases}
\]
Then \(g_n \leq n\) and \(g_n\) is measurable on \(A\). (Why?)
Also, \(g_n \geq 0\) and \(g_n \to h\) (pointwise) on \(A\).
For let \(\varepsilon > 0\), fix \(x \in A\), and find \(k > h(x)\). Then
\[
(\forall n \geq k) \; h(x) \leq n \quad \text{and} \quad g_n(x) = h(x).
\]
So
\[
(\forall n \geq k) \; |g_n(x) - h(x)| = 0 < \varepsilon.
\]
Clearly, \(g_n \leq h\). Hence by Theorem 5
\[
\lim_{n \to \infty} \int_A |h - g_n| = 0.
\]
Thus we can fix \(n\) so large that
\[
\int_A (h - g_n) < \frac{1}{2}\varepsilon.
\]
For that \(n\), let
\[
\delta = \frac{\varepsilon}{2n}
\]
and take any \(X \subseteq A\) \((X \in \mathcal{M})\), with \(mX < \delta\).
As \(g_n \leq n\) (see above), Theorem 1(e) in §5 yields
\[
\int_X g_n \leq \int_X (n) = n \cdot mX < n\delta = \frac{1}{2}\varepsilon.
\]
Hence as \(|f| \leq h\) and
\[
\int_X (h - g_n) \leq \int_A (h - g_n) < \frac{1}{2}\varepsilon
\]
(Theorem 1(f) of §5), we obtain
\[
\int_X |f| \leq \int_X h = \int_X (h - g_n) + \int_X g_n < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon,
\]
as required. \(\Box\)

\section*{Problems on Integrability and Convergence Theorems}

1. Fill in the missing details in the proofs of this section.

2. \(\text{(i) Show that if } f: S \to E^* \text{ is bounded and } m\text{-measurable on } A, \text{ with } mA < \infty, \text{ then } f \text{ is } m\text{-integrable on } A \text{ (Theorem 2) and}
\[
\int_A f = c \cdot mA,
\]
where \(\inf f[A] \leq c \leq \sup f[A]\).

\(\text{(ii) Prove that if } f \text{ also has the Darboux property on } A, \text{ then}
\[
(\exists x_0 \in A) \quad c = f(x_0).
\]
[Hint: Take } g = 1 \text{ in Theorem 3.]

\(\text{(iii) What results if } A = [a, b] \text{ and } m \text{ = Lebesgue measure?}

3. Prove Theorem 4 assuming that the \(f_n\) are measurable on \(A\) and that
\[
(\exists k) \quad \int_A f_k > -\infty
\]
instead of \(f_n \geq 0\).
[Hint: As } \{f_n\}\rangle, \text{ show that}
\[
(\forall n \geq k) \quad \int_A f_n > -\infty.
\]
If
\[
(\exists n) \quad \int_A f_n = \infty,
\]
then
\[
\int_A f = \lim \int_A f_n = \infty.
\]
Otherwise,
\[
(\forall n \geq k) \quad \left| \int_A f_n \right| < \infty;
\]
so \(f_n\) is integrable. (Why?) By Corollary 1 in §5, assume \(|f_n| < \infty\). (Why?) Apply
Theorem 4 to \(h_n = f_n - f_k\) \((n \geq k)\), considering two cases:
\[
\int_A h < \infty \text{ and } \int_A h = \infty.
\]

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
\[
\int_A F = \int_A f \quad \text{and} \quad \int_A F_n = \int_A f_n.
\]
[Hint: By Lemma 2 of §5, fix measurable maps $h \geq f$ and $h_n \geq f_n$ with the same integrals. Let
\[ F_n = \inf (h \wedge h_k), \quad n = 1, 2, \ldots, \]
and $F = \sup_n F_n \leq h$. (Why?) Proceed.]

5. For $A \in \mathcal{M}$ and any (even nonmeasurable) functions $f, f_n : S \to E^*$, prove the following.

(i) If $f_n \not\rightarrow f$ (a.e.) on $A$, then
\[ \int_A f_n \not\rightarrow \int_A f, \]
provided
\[ (\exists n) \int_A f_n > -\infty. \]
(ii) If $f_n \not\rightarrow f$ (a.e.) on $A$, then
\[ \int_A f_n \not\rightarrow \int_A f, \]
provided
\[ (\exists n) \int_A f_n < \infty. \]

[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?)]

6. Show by examples that

(i) the conditions
\[ \int_A f_n > -\infty \quad \text{and} \quad \int_A f < \infty \]
in Problem 5 are essential; and

(ii) Problem 5(i) fails for lower integrals. What about 5(ii)?

[Hints: (i) Let $A = (0, 1) \subset E^1$, $m = \text{Lebesgue measure}$, $f_n = -\infty$ on $A - \{\frac{1}{n}\}$, $f_n = 1$ elsewhere.

(ii) Let $\mathcal{M} = \{ E^1, \emptyset \}$, $m E^1 = 1$, $m \emptyset = 0$, $f_n = 1$ on $(-n, n)$, $f_n = 0$ elsewhere. If $f = 1$ on $A = E^1$, then $f_n \not\rightarrow f$, but not $\int_A f_n \rightarrow \int_A f$.

Explain!]

7. Given $f_n : S \to E^*$ and $A \in \mathcal{M}$, let
\[ g_n = \inf_{k \geq n} f_k \quad \text{and} \quad h_n = \sup_{k \geq n} f_k \quad (n = 1, 2, \ldots). \]

\[ \int_A f_n \rightarrow \int_A f \]

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
\[ \int_X f_n \rightarrow \int_X f < \infty, \]
then
\[ \int_A f_n \rightarrow \int_A f. \]
(ii) This fails for sign-changing $f_n$.

[Hints: If (i) fails, then
\[ \lim \int_X f_n < \int_X f \text{ or } \lim \int_X f_n > \int_X f. \]
Find a subsequence of
\[ \left\{ \int_X f_n \right\} \text{ or } \left\{ \int_{A-X} f_n \right\} \]
contradicting Lemma 2.

(iii) Let $m = \text{Lebesgue measure}$; $A = (0, 1), X = (0, \frac{1}{2})$,
\[ f_n = \left\{ \begin{array}{ll}
    n & \text{on } (0, \frac{1}{2n}) \\
    -n & \text{on } (1 - \frac{1}{n}, 1). \end{array} \right. \]

$\Rightarrow$9. (i) Show that if $f$ and $g$ are $m$-measurable and nonnegative on $A$, then
\[ \int_A (af + bg) = a \int_A f + b \int_A g. \]

(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.]
10. If
\[ f = \sum_{n=1}^{\infty} f_n, \]
with all \( f_n \) measurable and nonnegative on \( A \), then
\[ \int_A f = \sum_{n=1}^{\infty} \int_A f_n. \]
[Hint: Apply Theorem 4 to the maps \( g_n = \sum_{k=1}^{n} f_k \).]

11. If \( q = \sum_{n=1}^{\infty} \int_A |f_n| < \infty \)
and the \( f_n \) are \( m \)-measurable on \( A \), then
\[ \sum_{n=1}^{\infty} |f_n| < \infty \quad (a.e.) \quad \text{on} \ A \]
and \( f = \sum_{n=1}^{\infty} f_n \) is \( m \)-integrable on \( A \), with
\[ \int_A f = \sum_{n=1}^{\infty} \int_A f_n. \]
[Hint: Let \( g = \sum_{n=1}^{\infty} |f_n| \). By Problem 10,
\[ \int_A g = \sum_{n=1}^{\infty} \int_A |f_n| = q < \infty; \]
so \( g < \infty \quad (a.e.) \quad \text{on} \ A \). (Why?) Apply Theorem 5 and Note 1 to the maps
\[ g_n = \sum_{k=1}^{n} f_k; \]
note that \( |g_n| \leq g \).

12. (Convergence in measure; see Problem 11(ii) of §3).
(i) Prove Riesz' theorem: If \( f_n \to f \) in measure on \( A \subseteq S \), there is a subsequence \( \{f_{n_k}\} \) such that \( f_{n_k} \to f \) (almost uniformly), hence (a.e.), on \( A \).
[Outline: Taking \( \sigma_k = \delta_k = 2^{-k} \),
pick, step by step, naturals \( n_1 < n_2 < \cdots < n_k < \cdots \)
and sets \( D_k \in \mathcal{M} \) such that \( (\forall k) \)
\[ mD_k < 2^{-k} \]
and
\[ \rho'(f_{n_k}, f) < 2^{-k} \]
on \( A - D_k \). (Explain!) Let
\[ E_n = \bigcup_{k=n}^{\infty} D_k, \]
\[ mE_n < 2^{1-n}. \] (Why?) Show that
\[ (\forall n) \quad (\forall k > n) \quad \rho'(f_{n_k}, f) < 2^{1-n} \]
on \( A - E_n \). Use Problem 11 in §3.
(ii) For maps \( f_n : S \to E \) and \( g : S \to E^1 \) deduce that if \( f_n \to f \)
in measure on \( A \) and
\[ (\forall n) \quad |f_n| \leq g \quad (a.e.) \quad \text{on} \ A, \]
then
\[ |f| \leq g \quad (a.e.) \quad \text{on} \ A. \]
[Hint: \( f_{n_k} \to f \) (a.e.) on \( A \).]

13. Continuing Problem 12(ii), let \( f_n \to f \)
in measure on \( A \in \mathcal{M} \) \( (f_n : S \to E) \) and
\[ (\forall n) \quad |f_n| \leq g \quad (a.e.) \quad \text{on} \ A, \]
with
\[ \int_A g < \infty. \]
Prove that
\[ \lim_{n \to \infty} \int_A |f_n - f| = 0. \]
Does
\[ \int_A f_n \to \int_A f? \]
As before, we presuppose an arbitrary (but fixed) measure space $(\mathcal{M}, \mathcal{A}, m)$. Hence

$$\lim_{n \to \infty} \int_{\mathcal{A}} f_n = 0$$

since the series converges.

We then have

$$\sum_{k=1}^{n_0} \int_{D_k} f_n - f_k < 4\varepsilon$$

for $n > n_0$. (Why?) Hence

$$\lim_{n \to \infty} \int_{\mathcal{A}} f_n = 0.$$

See also Problem 7 in §5 and Note 1 of §6 (for measurable functions) as regards

$$\lim_{n \to \infty} \int_{\mathcal{A}} f_n.$$

14. Do Problem 12 in §3 (Lebesgue–Egorov theorems) for $T = E$, assuming

$$(\forall n) \ |f_n| \leq g \ (a.e.)$$

on $A$, with

$$\int_{\mathcal{A}} g < \infty$$

(instead of $mA < \infty$).

[Hint: With $H_n(k)$ as before, it suffices that

$$\lim_{k \to \infty} m(A - H_n(k)) = 0.$$]

(Why?) Verify that

$$(\forall n) \ |f_n| \leq g \ (a.e.)$$

on $A$.

Also, by Theorem 6, fix $\delta > 0$ with

$$\sigma \cdot mH < \varepsilon.$$
In particular, a complex function \( f \) is integrable on \( A \) iff its real and imaginary parts \( (f_{re} \) and \( f_{im} \) are). Then we also say that \( \int_A f \) exists.\(^2\) By (1), we have

\[
\int_A f = \left( \int_A f_{re}, \int_A f_{im} \right) = \int_A f_{re} + i \int_A f_{im}.
\]

If \( f: S \to C^n \), we use (1), with complex components \( f_k \).

With this definition, integration of functions \( f: S \to E^n(C^n) \) reduces to that of \( f_k: S \to E^1(C) \), and one easily obtains the same theorems as in \( \S\S 4-6 \), as far as they make sense for vectors.

**Theorem 1.** A function \( f: S \to E^n \) is integrable on \( A \in \mathcal{M} \) iff it is \( m \)-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 \( \S\S 6 \), for \( k = 1, 2, \ldots, n \), the functions \( f_k^+ \) and \( f_k^- \) are \( m \)-measurable; furthermore,

\[
\int_A f_k^+ \neq \pm \infty \text{ and } \int_A f_k^- \neq \pm \infty.
\]

This implies

\[
\infty > \int_A f_k^+ + \int_A f_k^- = \int_A (f_k^+ + f_k^-) = \int_A |f_k|, \quad k = 1, 2, \ldots, n.
\]

Since \( |f| \) is \( m \)-measurable by Problem 14 in \( \S\S 3 \) (\(| \cdot | \) is a continuous mapping from \( E^n \) to \( E^1 \)), and

\[
|f| = \sum_{k=1}^n |f_k| = \sum_{k=1}^n |f_k|,
\]

we get

\[
\int_A |f| \leq \sum_{k=1}^n \int_A |f_k| = \sum_{k=1}^n \int_A |f_k| < \infty.
\]

Conversely, if \( f \) satisfies

\[
\int_A |f| < \infty
\]

then

\[
(\forall k) \quad \int_A f_k < \infty.
\]

Also, the \( f_k \) are \( m \)-measurable if \( f \) is (see Problem 2 in \( \S\S 3 \)). Hence the \( f_k \) are integrable on \( A \) (by Theorem 2 of \( \S\S 6 \)), and so is \( f \).

The proof for \( C^n \) is analogous. \( \square \)

Similarly for other theorems (see Problems 1 to 4 below). We have already noted that Theorem 5 of \( \S\S 6 \) holds for complex and vector-valued functions. So does Theorem 6 in \( \S\S 6 \). We prove another proposition (Lemma 1) below.

II. Next we consider the general case, \( f: S \to E \) (\( E \) complete). We now adopt Theorem 1 as a definition. (It agrees with Definition 1 of \( \S\S 4 \). 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 1.** If \( f_n \to f \) (uniformly) on \( A \) \((mA < \infty) \), then

\[
\int_A |f_n - f| \to 0.
\]

**Proof.** By assumption,

\[
(\forall \varepsilon > 0) \quad (\exists k) \quad (\forall n > k) \quad |f_n - f| < \varepsilon \text{ on } A;
\]

so

\[
(\forall n > k) \quad \int_A |f_n - f| \leq \int_A \varepsilon = \varepsilon \cdot mA < \infty.
\]

As \( \varepsilon \) is arbitrary, the result follows. \( \square \)

Our goal is to prove results on linearity (Theorem 2) and additivity (Theorem 3) for general \( E \); for a “limited approach,” see Problem 2 for \( E = E^n \).

*Lemma 2.** If

\[
\int_A |f| < \infty \quad (mA < \infty)
\]

and

\[
f = \lim_{n \to \infty} f_n \quad \text{(uniformly) on } A - Q \quad (mQ = 0)
\]

for some elementary maps \( f_n \) on \( A \), then all but finitely many \( f_n \) are elementary and integrable on \( A \), and

\[
\lim_{n \to \infty} \int_A f_n
\]

exists in \( E \); further, the latter limit does not depend on the sequence \( \{f_n\} \).

**Proof.** By Lemma 1,

\[
(\forall \varepsilon > 0) \quad (\exists q) \quad (\forall n, k > q) \quad \int_A |f_n - f| < \varepsilon \text{ and } \int_A |f_n - f_k| < \varepsilon.
\]

\(\square\)
(The latter can be achieved since
\[ \lim_{k \to \infty} \int_A |f_n - f_k| = \int_A |f_n - f| < \varepsilon. \]

Now, as
\[ |f_n| \leq |f_n - f| + |f|, \]

Problem 7 in §5 yields
\[ \forall n > k \quad \int_A |f_n| \leq \int_A |f_n - f| + \int_A |f| < \varepsilon + \int_A |f| < \infty. \]

Thus \( f_n \) is elementary and integrable for \( n > k \), as claimed. Also, by Theorem 2 and Corollary 1(ii), both in §4,
\[ \forall n, k > q \quad \left| \int_A f_n - \int_A f_k \right| = \left| \int_A (f_n - f_k) \right| \leq \int_A |f_n - f_k| < \varepsilon. \]

Thus \( \{ \int_A f_n \} \) is a Cauchy sequence. As \( E \) is complete,
\[ \lim \int_A f_n = \pm \infty \]
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
\[ \left| \lim \int_A g_n - \lim \int_A f_n \right| = \left| \lim \int_A (g_n - f_n) \right| \leq \lim \int_A |g_n - f_n| = 0 \]
by Lemma 1, as \( g_n - f_n \to 0 \) (uniformly) on \( A \). Thus
\[ \lim \int_A g_n = \lim \int_A f_n, \]
and all is proved. \( \Box \)

This leads us to the following definition.

*Definition 2.*

If \( f : S \to E \) is integrable on \( A \in \mathcal{M} \) (\( m A < \infty \)), we set
\[ \int_A f = \int_A f \ dm = \lim_{n \to \infty} \int_A f_n \]
for any elementary and integrable maps \( f_n \) such that \( f_n \to f \) (uniformly) on \( A - Q \), \( mQ = 0 \).

\[ \text{Indeed, } f_n - f_k \to f_n - f \text{ (uniformly) on } A \text{ as } k \to \infty; \text{ so Lemma 1 applies.} \]

§7. Integration of Complex and Vector-Valued Functions

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
\[ \int_A f = \int_A f_n \]
(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
\[ \int_A f = \lim \int_A f_n = \lim \int_B f_n = \int_B f. \]
Thus we now define
\[ \int_A f = \int_B f, \]
even if \( mA = \infty \), provided \( f = 0 \) on \( A - B \), i.e.,
\[ f = f C_B \text{ on } A \]
(\( C_B \) = 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,
\[ \int_A |f| < \infty \]
implies that \( A(f \neq 0) \) is \( \sigma \)-finite. Neglecting \( A(f = 0) \), we may assume that
\[ A = \bigcup B_n, \ mB_n < \infty, \ \text{and } \{B_n\}^\uparrow \]
(if not, replace \( B_n \) by \( \bigcup_{k=1}^n B_k \)); so \( B_n \not\supset A \).

*Lemma 3.* Let \( \phi : S \to E \) be integrable on \( A \). Let \( B_n \not\supset A, \ mB_n < \infty, \)
and set
\[ f_n = \phi C_{B_n}, \quad n = 1, 2, \ldots. \]

Then \( f_n \to \phi \) (pointwise) on \( A \), all \( f_n \) are integrable on \( A \), and
\[ \lim_{n \to \infty} \int_A f_n \]
exists in \( E \). Furthermore, this limit does not depend on the choice of \( \{B_n\} \).

**Proof.** Fix any \( x \in A \). As \( B_n \not\supset A = \bigcup B_n \),
\[ \left( \exists n_0 \right) \forall n > n_0 \quad x \in B_n. \]
By assumption, \( f_n = \phi \) on \( B_n \). Thus
\[
(\forall n > n_0) \quad f_n(x) = \phi(x);
\]
so \( f_n \to \phi \) (pointwise) on \( A \).

Moreover, \( f_n = \phi C_{B_n} \) is \( m \)-measurable on \( A \) (as \( \phi \) and \( C_{B_n} \) are); and
\[
|f_n| = |\phi| C_{B_n}
\]
implies
\[
\int_A |f_n| \leq \int_A |\phi| < \infty.
\]

Thus all \( f_n \) are integrable on \( A \).

As \( f_n = 0 \) on \( A - B_n \) \((mB < \infty)\),
\[
\int_A f_n
\]
is defined. Since \( f_n \to \phi \) (pointwise) and \( |f_n| \leq |\phi| \) on \( A \), Theorem 5 in §6, with \( g = |\phi| \), yields
\[
\int_A |f_n - \phi| \to 0.
\]

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. □

**Definition 3.**
If \( \phi : S \to E \) is integrable on \( A \in \mathcal{M} \), we set
\[
\phi = \lim_{n \to \infty} \int_A \phi \ dm = \lim_{n \to \infty} \int_A f_n,
\]
with the \( f_n \) as in Lemma 3 (even if \( \phi \) has no \( m \)-finite support).

**Theorem 2** (linearity). If \( f, g : S \to E \) are integrable on \( A \in \mathcal{M} \), so is
\[
pf + qg
\]
for any scalars \( p, q \). Moreover,
\[
\int_A (pf + qg) = p \int_A f + q \int_A g.
\]
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
\[
|pf + qg| \leq |p| |f| + |q| |g|
\]
yields
\[
\int_A |pf + qg| \leq |p| \int_A |f| + |q| \int_A |g| < \infty.
\]

Now, as noted above, assume that
\[
f = f C_{B_1} \quad \text{and} \quad g = g C_{B_2}
\]
for some \( B_1, B_2 \subseteq A \) \((mB_1 + mB_2 < \infty)\). Let \( B = B_1 \cup B_2 \); so
\[
f = g = pf + qg = 0 \quad \text{on} \quad A - B;
\]
additionally,
\[
\int_A f = \int_B f, \quad \int_A g = \int_B g, \quad \text{and} \quad \int_A (pf + qg) = \int_B (pf + qg).
\]

Also, \( mB < \infty \); so by Definition 2,
\[
\int_B f = \lim \int_B f_n \quad \text{and} \quad \int_B g = \lim \int_B g_n
\]
for some elementary and integrable maps
\[
f_n \to f \quad \text{uniformly} \quad \text{and} \quad g_n \to g \quad \text{uniformly} \quad \text{on} \quad B - Q, \quad mQ = 0.
\]

Thus
\[
pf_n + qg_n \to pf + qg \quad \text{uniformly} \quad \text{on} \quad B - Q.
\]

But by Theorem 2 and Corollary 1(vii), both of §4 (for elementary and integrable maps),
\[
\int_B (pf_n + qg_n) = p \int_B f_n + q \int_B g_n.
\]

Hence
\[
\int_A (pf + qg) = \int_B (pf + qg) = \lim \int_B (pf_n + qg_n)
\]
\[
= \lim \left( p \int_B f_n + q \int_B g_n \right) = p \int_A f + q \int_A g = p \int_A f + q \int_A g.
\]

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
\[
A = \bigcup_{n=1}^{\infty} B_n, \quad \{B_n \}^\uparrow, \quad mB_n < \infty,
\]
and
\[
f_n = f C_{B_n}, \quad g_n = g C_{B_n}, \quad n = 1, 2, \ldots
\]
Then by definition,
\[ \int_A f = \lim_{n \to \infty} \int_A f_n \quad \text{and} \quad \int_A g = \lim_{n \to \infty} \int_A g_n, \]
and so
\[ p \int_A f + q \int_A g = \lim_{n \to \infty} \left( p \int_A f_n + q \int_A g_n \right). \]
As \( f_n, g_n \) have \( m \)-finite supports, the first part of the proof yields
\[ p \int_A f_n + q \int_A g_n = \int_A (pf_n + qg_n). \]
Thus as claimed,
\[ p \int_A f + q \int_A g = \lim_{n \to \infty} (pf_n + qg_n) = \int_A (pf + qg). \quad \Box \]

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 3** (additivity).
(i) If \( f : S \to E \) is integrable on each of \( n \) disjoint \( \mathcal{M} \)-sets \( A_k \), it is so on their union
\[ A = \bigcup_{k=1}^n A_k, \]
and
\[ \int_A f = \sum_{k=1}^n \int_{A_k} f. \]
(ii) This holds for countable unions, too, if \( f \) is integrable on all of \( A \).

*Proof.* Let \( f \) have \( m \)-finite support: \( f = f_{C_B} \) on \( A, mB < \infty \). Then
\[ \int_A f = \int_B f \quad \text{and} \quad \int_{A_k} f = \int_{B_k} f, \]
where
\[ B_k = A_k \cap B, \quad k = 1, 2, \ldots, n. \]
By Definition 2, fix elementary and integrable maps \( f_k \) (on \( A \)) and a set \( Q \) (\( mQ = 0 \)) such that \( f_k \to f \) (uniformly) on \( B - Q \) (hence also on \( B_k - Q \)), with
\[ \int_A f = \int_B f = \lim_{i \to \infty} \int_B f_i \quad \text{and} \quad \int_{A_k} f = \lim_{i \to \infty} \int_{B_k} f_i, \quad k = 1, 2, \ldots, n. \]
As the \( f_i \) are elementary and integrable, Theorem 1 in §4 yields
\[ \int_A f_i = \int_B f_i = \sum_{k=1}^n \int_{B_k} f_i = \sum_{k=1}^n \int_{A_k} f_i. \]
Hence
\[ \int_A f = \lim_{i \to \infty} \int_B f_i = \lim_{i \to \infty} \sum_{k=1}^n \int_{B_k} f_i = \sum_{k=1}^n \left( \lim_{i \to \infty} \int_{A_k} f_i \right) = \sum_{k=1}^n \int_{A_k} f_i. \]
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
\[ A = \bigcup_{k=1}^\infty A_k \quad \text{(disjoint),} \quad A_k \in \mathcal{M}. \]
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
\[ \int_A g_n = \int_{B_n} g_n = \sum_{k=1}^n \int_{A_k} g_n = \sum_{k=1}^n \int_{A_k} f, \]
since \( g_n = f \) on each \( A_k \subseteq B_n \).

Also, as is easily seen, \(|g_n| \leq |f|\) on \( A \) and \( g_n \to f \) (pointwise) on \( A \) (proof as in Lemma 3). Thus by Theorem 5 in §6,
\[ \int_A |g_n - f| \to 0. \]
As
\[ \left| \int_A g_n - \int_A f \right| = \left| \int_A (g_n - f) \right| \leq \int_A |g_n - f|, \]
we obtain
\[ \int_A f = \lim_{n \to \infty} \int_A g_n, \]
and the result follows by (3). \( \Box \)

**Problems on Integration of Complex and Vector-Valued Functions**

1. Prove Corollary 1(iii)–(vii) in §4 componentwise for integrable maps \( f : S \to E^n (C^n) \).
2. Prove Theorems 2 and 3 componentwise for \( E = E^n (C^n) \).
3. Do it for Corollary 3 in §6.
3. Prove Theorem 1 with
\[ \int_A |f| < \infty \]
replaced by
\[ \int_A |f_k| < \infty, \quad k = 1, \ldots, n. \]
4. Prove that if \( f : S \to E^n (C^n) \) is integrable on \( A \), so is \( |f| \). Disprove the converse.
5. Disprove Lemma 1 for \( mA = \infty \).
*6. Complete the proof of Lemma 3.
*7. Complete the proof of Theorem 3.
*a 8. Do Problem 1 and 2 for \( f : S \to E \).
9. Prove formula (1) from definitions of Part II of this section.
\Rightarrow 10. Show that
\[ \left| \int_A f \right| \leq \int_A |f| \]
for integrable maps \( f : S \to E \). See also Problem 14.
[Hint: If \( mA < \infty \), use Corollary 1(ii) of §4 and Lemma 1. If \( mA = \infty \), “imitate” the proof of Lemma 3.]
11. Do Problem 11 in §6 for \( f_n : S \to E \). Do it componentwise for \( E = E^n (C^n) \).
12. Show that if \( f, g : S \to E^1 (C) \) are integrable on \( A \), then
\[ \left| \int_A fg \right|^2 \leq \int_A |f|^2 \cdot \int_A |g|^2. \]
In what case does equality hold? Deduce Theorem 4(c) in Chapter 3, §31–3, from this result.
[Hint: Argue as in that theorem. Consider the case
\( (\exists t \in E^1) \int_A |f - t g| = 0. \)]
13. Show that if \( f : S \to E^1 (C) \) is integrable on \( A \) and
\[ \left| \int_A f \right| = \int_A |f|, \]
then
\[ (\exists \ c \in C) \quad c f = |f| \quad \text{a.e. on } A. \]
\[ \text{[Hint: Let } a = \int_A f. \text{ The case } a = 0 \text{ is trivial. If } a \neq 0, \text{ let } c = \frac{a}{|f|} \text{; } |c| = 1; c a = |a|.] \]
Let \( r = (\epsilon f)_{\text{re}}. \) Show that \( r \leq |\epsilon f| = |f|, \)
\[ \left| \int_A f \right| = \int_A \epsilon f = \int_A r \leq \int_A |f| = \int_A f, \]
\[ \left| \int_A f \right| = \int_A r = \int_A (\epsilon f)_{\text{re}}, \]
\( (\epsilon f)_{\text{re}} = |\epsilon f| \text{ a.e.}, \text{ and } \epsilon f = |\epsilon f| = |f| \text{ a.e. on } A. \]
15. Show that if \( f : S \to E \) is integrable on \( A \), it is integrable on each \( M \)-set \( B \subset A \). If, in addition,
\[ \int_B f = 0 \]
for all such \( B \), show that \( f = 0 \) a.e. on \( A \). Prove it for \( E = E^n \) first.
[Hint for \( E = E^n : A = A(f > 0) \cup A(f < 0) \). Use Theorems 1(1) and 2 from §5.]
16. In Problem 15, show that
\[ s = \int f \]
is a \( \sigma \)-additive set function on
\[ \mathcal{M}_A = \{ X \in \mathcal{M} \mid X \subset A \} \]
(Note 4 in §5); \( s \) is called the indefinite integral of \( f \) in \( A \).

§7. Integration of Complex and Vector-Valued Functions

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
\[ A \times B, \]
with \( A \in \mathcal{M}, B \in \mathcal{N}, mA < \infty, \) and \( nB < \infty \).

Define a premeasure \( s : \mathcal{C} \to E^1 \) by
\[ s(A \times B) = mA \cdot nB, \quad A \times B \subset \mathcal{C}. \]
Let \( p^* \) be the \( s \)-induced outer measure in \( X \times Y \) and
\[ p : \mathcal{P}^* \to E^* \]
the \( p^* \)-induced measure (“product measure,” \( p = m \times n \)) on the \( \sigma \)-field \( \mathcal{P}^* \) of all \( p^* \)-measurable sets in \( X \times Y \) (Chapter 7, §§5–6).
We consider functions $f: X \times Y \to E^*$ (extended-real).

I. We begin with some definitions.

Definitions.

(1) Given a function $f: X \to Y \to E^*$ (of two variables $x,y$), let $f_x$ or $f(x,\cdot)$ denote the function on $Y$ given by

$$f_x(y) = f(x,y);$$

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 $g: X \to E^*$ by

$$g(x) = \int_Y f(x,\cdot) \, dy,$$

and set

$$\int_X \int_Y f \, dy \, dx = \int_X g \, dx,$$

also written

$$\int_Y f \, dx \, dy = \int_{X \times Y} f \, dx \, dy.$$

This is called the iterated integral of $f$ on $Y$ and $X$, in this order.

Similarly,

$$h(y) = \int_X f^y \, dx$$

and

$$\int_Y \int_X f \, dx \, dy = \int_Y h \, dy.$$

Note that by the rules of $\S 5$, 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

$$\int_X \int_Y f \, dx \, dy = \int_Y \int_X f \, dy \, dx = \int_{X \times Y} f \, dp$$

(the main point).

§8. Product Measures. Iterated Integrals

For monotone sequences

$$f_k: X \times Y \to E^* \quad (k = 1,2,\ldots),$$

we now obtain the following lemma.

Lemma 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

$$g_k(x) = \int_Y f_k(x,\cdot) \, dy$$

and

$$h_k(y) = \int_X f_k(\cdot,y) \, dx.$$

By assumption,

$$0 \leq g_k(x) \nearrow g(x)$$

pointwise on $Y$. Thus by Theorem 4 in §6,

$$\int_Y f_k(x,\cdot) \nearrow \int_Y f(\cdot,y) \, dy,$$

i.e., $g_k \nearrow g$ (pointwise) on $X$, with $g$ as in Definition 2.

Again, by Theorem 4 of §6,

$$\int_X g_k \, dx \nearrow \int_X g \, dx;$$

or by Definition 2,

$$\int_X \int_Y f \, dx \, dy = \lim_{k \to \infty} \int_X \int_Y f_k \, dx \, dy.$$

Similarly for

$$\int_Y \int_X f \, dy \, dx$$

and

$$\int_{X \times Y} f \, dp.$$
so \( mQ = 0 \), and all \( f_k(x, \cdot) \) are \( n \)-measurable on \( Y \), for \( x \in X - Q \). Hence so is

\[
f(x, \cdot) = \lim_{k \to \infty} f_k(x, \cdot).
\]

Similarly for \( f(\cdot, y) \). Thus \( f \) satisfies (b).

Property (a) follows from \( g_h \to g \) and \( h_k \to h \). \( \square \)

Using Problems 9 and 10 from §6, the reader will also easily verify the following lemma.

**Lemma 2.**

(i) If \( f_1 \) and \( f_2 \) are nonnegative, \( p \)-measurable Fubini maps, so is \( a f_1 + b f_2 \) for \( a, b \geq 0 \).

(ii) If, in addition,

\[
\int_{X \times Y} f_1 \, dp < \infty \text{ or } \int_{X \times Y} f_2 \, dp < \infty,
\]

then \( f_1 - f_2 \) is a Fubini map, too.

**Lemma 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

\[
\int_X \int_Y f \, dn \, dm = \sum_{i=1}^{\infty} \int_X \int_Y f_i \, dn \, dm
\]

and

\[
\int_Y \int_X f \, dm \, dn = \sum_{i=1}^{\infty} \int_Y \int_X f_i \, dm \, dn.
\]

**II.** By Theorem 4 of Chapter 7, §3, the family \( \mathcal{C} \) (see above) is a semiring, being the product of two rings,

\[
\{ A \in \mathcal{M} \mid mA < \infty \} \text{ and } \{ B \in \mathcal{N} \mid nB < \infty \}.
\]

(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 1.** The product premeasure \( s \) is \( \sigma \)-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.

§8. Product Measures. Iterated Integrals

**Proof.** Let \( D = A \times B \in \mathcal{C} \); so

\[
C_D(x, y) = C_A(x) \cdot C_B(y).
\]

(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,

\[
g(x) = \int_Y C_D(x, \cdot) \, dn = C_A(x) \int_Y C_B \, dn = C_A(x) \cdot nB;
\]

so \( g = C_A \cdot nB \) is \( \mathcal{M} \)-simple on \( X \), with

\[
\int_X \int_Y C_D \, dn \, dm = \int_X g \, dm = nB \int_X C_A \, dm = nB \cdot mA = sD.
\]

Similarly for \( C_D(\cdot, y) \), and

\[
b(y) = \int_X C_D(\cdot, y) \, dm.
\]

Thus \( C_D \) has Fubini properties (a) and (b), and for every \( D \in \mathcal{C} \)

\[
\int_X \int_Y C_D \, dn \, dm = \int_X \int_Y C_D \, dm \, dn = sD.
\]

To prove \( \sigma \)-additivity, let

\[
D = \bigcup_{i=1}^{\infty} D_i \text{ (disjoint), } D_i \in \mathcal{C};
\]

so

\[
C_D = \sum_{i=1}^{\infty} C_{D_i}.
\]

(Why?) As shown above, each \( C_{D_i} \) has Fubini properties (a) and (b); so by (1) and Lemma 3,

\[
sD = \int_X \int_Y C_D \, dn \, dm = \sum_{i=1}^{\infty} \int_X \int_Y C_{D_i} \, dm \, dn = \sum_{i=1}^{\infty} sD_i;
\]

as required.

Assertion (i) now follows by Theorem 2 in Chapter 7, §6. Hence

\[
sD = pD = \int_{X \times Y} C_D \, dp;
\]

so by formula (1), \( C_D \) also has Fubini property (c), and all is proved. \( \square \)

Next, let \( \mathcal{P} \) be the \( \sigma \)-ring generated by the semiring \( \mathcal{C} \) (so \( \mathcal{C} \subseteq \mathcal{P} \subseteq \mathcal{P}^* \)).
Lemma 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 \( D \subseteq H, \quad H \in \mathcal{C} \).

This is simply Theorem 3 in Chapter 7, §3, with changed notation.

Lemma 5. If \( D \in \mathcal{P} \) (\( \sigma \)-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 \( \mathcal{P} \subseteq \mathcal{R} \).

(i) By Theorem 1, each \( C_D \) (\( D \in \mathcal{C} \)) is a Fubini map; so each \( D \in \mathcal{C} \) is in \( \mathcal{R} \).

(ii) Let

\[
D = \bigcup_{i=1}^{\infty} D_i \text{ (disjoint), } \quad D_i \in \mathcal{R}.
\]

Then

\[
C_D = \sum_{i=1}^{\infty} C_{D_i},
\]

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 \( D \subseteq H, \quad H \in \mathcal{C} \).

Now, \( D \subseteq H \) implies

\[
C_{H - D} = C_H - C_D.
\]

(Why?) Also, by Theorem 1, \( H \in \mathcal{C} \) implies

\[
\int_{X \times Y} C_H \, dp = pH = sH < \infty,
\]

and \( C_H \) is a Fubini map. So is \( C_D \) by assumption. So also is

\[
C_{H - D} = C_H - C_D
\]

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. ☐

We can now establish one of the main theorems, due to Fubini.

Theorem 2 (Fubini I). Suppose \( f : X \times Y \to E^* \) is \( \mathcal{P} \)-measurable on \( X \times Y \) (\( \mathcal{P} \) as above) ron. Then \( f \) is a Fubini map if either

(i) \( f \geq 0 \) on \( X \times Y \), or

(ii) one of

\[
\int_{X \times Y} |f| \, dp, \quad \int_X \int_Y |f| \, dn \, dm, \quad \text{or} \quad \int_Y \int_X |f| \, dm \, dn
\]

is finite.\(^1\)

In both cases,

\[
\int_X \int_Y f \, dn \, dm = \int_Y \int_X f \, dm \, dn = \int_{X \times Y} f \, dp.
\]

Proof. First, let

\[
f = \sum_{i=1}^{\infty} a_i C_{D_i} \quad (a_i \geq 0, \quad D_i \in \mathcal{P}),
\]

i.e., \( f \) is \( \mathcal{P} \)-elementary, hence certainly \( \sigma \)-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,

\[
f = \lim_{k \to \infty} f_k \text{ on } X \times Y
\]

for some sequence \( \{ f_k \} \) 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 \( f^+, \quad f^- \), 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

\[
\int_{X \times Y} f^+ \, dp < \infty
\]

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 \( \sigma \)-finiteness conditions, as shown below.

Lemma 6. Let \( D \in \mathcal{P^*} \) be \( \sigma \)-finite, i.e.,

\[
D = \bigcup_{i=1}^{\infty} D_i \text{ (disjoint)}
\]

for some \( D_i \in \mathcal{P^*} \), with \( pD_i < \infty \) (\( i = 1, 2, \ldots \)).\(^2\)

\(^1\) Note the use of absolute values; without them, Theorem 2 fails (see Problem 5’).

\(^2\) See Note 2 in Chapter 7, §8.
Then there is a $K \in \mathcal{P}$ such that $p(K - D) = 0$ and $D \subseteq K$.

**Proof.** As $\mathcal{P}$ is a $\sigma$-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 7.** If $D \in \mathcal{P}$ is $\sigma$-finite (Lemma 6), then $C_D$ is a Fubini map.

**Proof.** By Lemma 6,

(∃ $K \in \mathcal{P}$) \quad p(K - D) = 0, \quad D \subseteq K.

Let $Q = K - D$, so $pQ = 0$, and $C_Q = C_K - C_D$; that is, $C_D = C_K - C_Q$ and

$$\int_{X \times Y} C_Q \, dp = pQ = 0.$$

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 $\sigma$-finite; so by Lemma 6,

(∃ $Z \in \mathcal{P}$) \quad Q \subseteq Z, \quad pZ = pQ = 0.

Again $C_Z$ is a Fubini map; so

$$\int_X \int_Y C_Z \, dn \, dm = \int_{X \times Y} C_Z \, dp = pZ = 0.$$

As $Q \subseteq Z$, we have $C_Q \leq C_Z$, and so

$$\int_X \int_Y C_Q \, dn \, dm \leq \int_X \left[ \int_Y C_Q(x, \cdot) \, dn \right] dm = \int_{X \times Y} C_Z \, dp = 0.$$

Similarly,

$$\int_Y \int_X C_Q \, dm \, dn = \int_Y \left[ \int_X C_Q(\cdot, y) \, dm \right] dn = 0.$$

Thus setting

$$g(x) = \int_Y C_Q(x, \cdot) \, dn$$
and

$$h(y) = \int_X C_Q(\cdot, y) \, dm,$$
we have

$$\int_X g \, dm = 0 = \int_Y h \, dn.$$

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

$$\int_X \int_Y C_Q \, dn \, dm = \int_Y \int_X C_Q \, dm \, dn = \int_{X \times Y} C_Q \, dp = 0,$$

as required. □

**Theorem 3 (Fubini II).** Suppose $f : X \times Y \to E$ 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 $\sigma$-finite support, i.e., $f$ vanishes outside some $\sigma$-finite set $H \subseteq X \times Y$.

**Proof.** First, let

$$f = \sum_{i=1}^{\infty} a_i C_{D_i}, \quad (a_i > 0, \ D_i \in \mathcal{P}^*),$$

with $f = 0$ on $-H$ (as above).

As $f = a_i \not= 0$ on $A_i$, we must have $D_i \subseteq H$; so all $D_i$ are $\sigma$-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

$$f^+ = f^- = |f| = 0$$
on $-H$.

Thus $f^+$, $f^-$, and $f$ are Fubini maps by part (i) and the argument of Theorem 2. □

**Note 1.** The $\sigma$-finite support is automatic if $f$ is $p$-integrable (Corollary 1 in §5), or if $p$ or both $m$ and $n$ are $\sigma$-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

$$(X_i, \mathcal{M}_i, m_i), \quad i = 1, \ldots, q.$$

One writes

$$p = m_1 \times m_2$$

if $q = 2$ and sets

$$m_1 \times m_2 \times \cdots \times m_{q+1} = (m_1 \times \cdots \times m_q) \times m_{q+1}.$$
Theorems 2 and 3 with similar assumptions then state that the order of integrations is immaterial.

**Note 3.** Lebesgue measure in \( E^n \) 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 \( f : X \times Y \to E^n (C^n) \) if Definitions 2 and 3 are modified as follows (so that they make sense for such maps): In Definition 2, set
\[
g(x) = \int_Y f_x \, dm
\]
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
\[
\int_Y \int_X |f| \, dm \, dn = \int_X \int_Y |f| \, dm \, dn = \int_{X \times Y} |f| \, dp.
\]

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
\[
\int_D f \, dp \quad (D \in \mathcal{P}^*)
\]
reduce to
\[
\int_{X \times Y} f \cdot C_D \, dp.
\]
Thus it suffices to consider integrals over \( X \times Y \).

**Problems on Product Measures and Fubini Theorems**

1. Prove Lemmas 2 and 3.

2. Show that \( \{ A \in \mathcal{M} \mid mA < \infty \} \) is a set ring.

3. Fill in all proof details in Theorems 1 to 3.

4. Do the same for Lemmas 5 to 7.

3. Prove that if \( m \) and \( n \) are \( \sigma \)-finite, so is \( p = m \times n \). Disprove the converse by an example.

4. Prove the following.

   (i) Each \( D \in \mathcal{P} \) (as in the text) is \( \sigma \)-finite.

(ii) All \( \mathcal{P} \)-measurable maps \( f : X \times Y \to E^n \) have \( \sigma \)-finite support.

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?

   (ii) Which \( C \)-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?

5'. Let \( m = n = \text{Lebesgue measure in } [0,1] = X = Y \). Let
\[
f_k = \begin{cases} k(k+1) & \text{on } (\frac{1}{k+1}, \frac{1}{k}) \\ 0 & \text{otherwise.} \end{cases}
\]

Let
\[
f(x, y) = \sum_{k=1}^{\infty} [f_k(x) - f_{k+1}(x)] f_k(y);
\]
the series converges. (Why?) Show that

(i) \( (\forall k) \int_X f_k \, dp = 1 \);

(ii) \( \int_X \int_Y f \, dm \, dn = 1 \neq 0 = \int_Y \int_X f \, dm \, dn \).

What is wrong? Is \( f \) \( \mathcal{P} \)-measurable?

6. Let \( X = Y = [0,1] \), \( m \) as in Example (c) of Chapter 7, §6, \( (S = X) \) and \( n = \text{Lebesgue measure in } Y \).

(i) Show that \( p = m \times n \) is a topological measure under the standard metric in \( E^2 \).

(ii) Prove that \( D = \{(x, y) \in X \times Y \mid x = y\} \in \mathcal{P}^* \).

(iii) Describe \( C \).

[Hints: (i) Any subinterval of \( X \times Y \) is in \( \mathcal{P}^* \); (ii) \( D \) is closed. Verify!]
7. Continuing Problem 6, let \( f = C_D \).
   
   (i) Show that
   \[
   \int_Y \int_X f \, dn \, dm = 0 \neq 1 = \int_Y \int_X f \, dm \, dn.
   \]
   What is wrong?
   
   [Hint: \( D \) is not \( \sigma \)-finite; for if
   \[
   D = \bigcup_{i=1}^{\infty} D_i,
   \]
   at least one \( D_i \) is uncountable and has no finite basic covering values (why?), so \( p^*D_i = \infty \).]
   
   (ii) Compute \( p^*\{(x,0) \mid x \in X\} \) and \( p^*\{(0,y) \mid y \in Y\} \).

8. Show that \( D \in \mathcal{P}^* \) is \( \sigma \)-finite iff
   \[
   D \subseteq \bigcup_{i=1}^{\infty} D_i \text{ (disjoint)}
   \]
   for some sets \( D_i \in \mathcal{C} \).
   
   [Hint: First let \( p^*D < \infty \). Use Corollary 1 from Chapter 7, \S 1.]

9. Given \( D \in \mathcal{P} \), \( a \in X \), and \( b \in Y \), let
   \[
   D_a = \{ y \in Y \mid (a, y) \in D \}
   \]
   and
   \[
   D_b = \{ x \in X \mid (x, b) \in D \}.
   \]
   (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
   \[
   \mathcal{R} = \{ Z \in \mathcal{P} \mid Z_a \in \mathcal{N} \}.
   \]
   Show that \( \mathcal{R} \) is a \( \sigma \)-ring \( \supseteq \mathcal{C} \). Hence \( \mathcal{R} \supseteq \mathcal{P} \), \( D \in \mathcal{R} \); \( D_a \in \mathcal{N} \). Similarly for \( D^b \).
   
   \( \Rightarrow \) 10. Let \( m = n = \) Lebesgue measure in \( E^1 = X = Y \). Let \( f : E^1 \to [0, \infty) \) be \( m \)-measurable on \( X \). Let
   \[
   H = \{ (x,y) \in E^2 \mid 0 \leq y < f(x) \}
   \]
   and
   \[
   G = \{ (x,y) \in E^2 \mid y = f(x,y) \}
   \]
   (the “graph” of \( f \)). Prove that
   
   (i) \( H \in \mathcal{P}^* \) and
   
   \[
   pH = \int_X f \, dm
   \]
   \( (= \) “the area under \( f \)”);
   
   (ii) \( G \in \mathcal{P}^* \) and \( pG = 0 \).
   
   [Hints: (i) First take \( f = C_D \) and elementary and nonnegative maps. Then use Lemma 2 in \S 2 (last clause). Fix elementary and nonnegative maps \( f_k \searrow f \), assuming \( f_k < f \) (if not, replace \( f_k \) by \( (1 - \frac{1}{k}) f \)). Let
   \[
   H_k = \{ (x,y) \mid 0 \leq y < f_k(x) \}.
   \]
   Show that \( H_k \not\rightarrow H \in \mathcal{P}^* \).
   
   (ii) Set
   \[
   \phi(x,y) = y - f(x).
   \]
   Using Corollary 4 of \S 1, show that \( \phi \) 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),
   
   \[
   (\forall x \in E^1) \int_Y C_G(x, \cdot) \, dn = nG_x = 0,
   \]
   as \( G_x = \{ f(x) \}; \) a singleton.]

11. Let
   \[
   f(x,y) = \phi_1(x)\phi_2(y).
   \]
   Prove that if \( \phi_1 \) is \( m \)-integrable on \( X \) and \( \phi_2 \) is \( n \)-integrable on \( Y \), then \( f \) is \( p \)-integrable on \( X \times Y \) and
   
   \[
   \int_{X \times Y} f \, dp = \int_X \phi_1 \cdot \int_Y \phi_2.
   \]

12. Prove Theorem 3(ii) for \( f : X \times Y \to E \) (\( E \) complete).
   
   [Outline: If \( f \) is \( \mathcal{P}^* \)-simple, use Lemma 7 above and Theorem 2 in \S 7.]
   
   If
   \[
   f = \sum_{k=1}^{\infty} a_k C_{D_k}, \quad D_k \in \mathcal{P}^*,
   \]
   let
   \[
   H_k = \bigcup_{i=1}^{k} D_k
   \]
   and \( f_k = f C_{H_k} \), so the \( f_k \) are \( \mathcal{P}^* \)-simple (hence Fubini maps), and \( f_k \to f \) (pointwise) on \( X \times Y \), with \( |f_k| \leq |f| \) and
   \[
   \int_{X \times Y} |f| \, dp < \infty
   \]
   (by assumption). Now use Theorem 5 from \S 6.
Let now \( f \) be \( \mathcal{P}^* \)-measurable; so

\[ f = \lim_{k \to \infty} f_k \] (uniformly)

for some \( \mathcal{P}^* \)-elementary maps \( g_k \) (Theorem 3 in §1). By assumption, \( f = f_C H \) (\( H \) \( \sigma \)-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,

\[ H \subseteq \bigcup_i D_i \text{ (disjoint), } D_i \in \mathcal{C}. \]

Let \( H_i = H \cap D_i \). By the previous step, each \( f_C H_i \) is a Fubini map; so is

\[ f_k = \sum_{i=1}^n f_C H_i, \]

(why?), hence so is \( f = \lim_{k \to \infty} f_k \), by Theorem 5 of §6. (Verify!)

13. Let \( m \) be Lebesgue measure in \( E^1 \), \( p \) is Lebesgue measure in \( E^s \), \( X = (0, \infty) \), and

\[ Y = \{ y \in E^s \mid |y| = 1 \}. \]

Given \( \bar{x} \in E^s \setminus \{0\} \), let

\[ r = |\bar{x}| \text{ and } \bar{u} = \frac{\bar{x}}{r} \in Y. \]

Call \( r \) and \( \bar{u} \) the polar coordinates of \( \bar{x} \neq \bar{0} \).

If \( D \subseteq Y \), set

\[ n^* D = s \cdot p^* \{ r \bar{u} \mid \bar{u} \in D, \ 0 < r \leq 1 \}. \]

Show that \( n^* \) is an outer measure in \( Y \); so it induces a measure \( m \) in \( Y \).

Then prove that

\[ \int_{E^s} f \ dp = \int_X r^{s-1} \ dm(r) \int_Y f(r \bar{u}) \ dn(\bar{u}) \]

if \( f \) is \( p \)-measurable and nonnegative on \( E^s \).

[Hint: Start with \( f = C_A \),

\[ A = \{ r \bar{u} \mid \bar{u} \in H, \ a < r \leq b \}, \]

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.]

§9. Riemann Integration. Stieltjes Integrals

1. 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 \( \mathcal{C}_* \)), such as the volume function \( v \) (Chapter 7, §§1–2).

By a \( \mathcal{C} \)-partition of \( A \in \mathcal{C} \) (or \( A \in \mathcal{C}_* \)), we mean a finite family

\[ \mathcal{P} = \{ A_i \} \subset \mathcal{C} \]

such that

\[ A = \bigcup_i A_i \text{ (disjoint).} \]

As we noted in §5, the Riemann integral,

\[ R \int_A f = R \int_A f \ dm, \]

of \( f : E^n \to E^1 \) can be defined as its Lebesgue counterpart,

\[ \int_A f, \]

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 \( f : E^n \to E^* \) and a \( \mathcal{C} \)-partition

\[ \mathcal{P} = \{ A_1, \ldots, A_q \} \]

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

\[ (1) \quad \underline{S}(f, \mathcal{P}) = \sum_{i=1}^q m A_i \cdot \inf \{ f(A_i) \}, \quad \overline{S}(f, \mathcal{P}) = \sum_{i=1}^q m A_i \cdot \sup \{ f(A_i) \}. \]

(b) The lower and upper Riemann integrals ("\( R \)-integrals") of \( f \) on \( A \) (with respect to \( m \)) are

\[ (2) \quad R \int_A f = R \int_A f \ dm = \sup \mathcal{P} \underline{S}(f, \mathcal{P}) \text{ and } \]

\[ R \int_A f = R \int_A f \ dm = \inf \mathcal{P} \overline{S}(f, \mathcal{P}), \]

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

\[ R \int_A f = R \int_A f. \]

1,2 These expressions exist in \( E^* \) (Chapter 4, §4, (2')).
We then set
\[ R \int_A f = R \int_A f = R \int_A f \ dm = R \int_A f \ dm \]
and call it the Riemann integral (“R-integral”) of \( f \) on \( A \). “Classical” notation:
\[ R \int_A f(\bar{x}) \ dm(\bar{x}). \]

If \( A = [a, b] \subseteq E^1 \), we also write
\[ R \int_a^b f = R \int_a^b f(x) \ dm(x) \]
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
\[ R \int_A f \]
exists (note that this implies the boundedness of \( f \)); note that
\[ R \int_A f \text{ and } R \int_A f \]
are always defined in \( E^1 \).

Below, we always restrict \( f \) to a fixed \( A \subseteq C \) (or \( A \subseteq \mathcal{C}_a \); \( \mathcal{P}', \mathcal{P}'', \mathcal{P}''', \mathcal{P}'''', \mathcal{P}_k \)) denote \( C \)-partitions of \( A \).

We now obtain the following result for any additive \( m : C \rightarrow [0, \infty) \).

**Corollary 1.** If \( \mathcal{P} \) refines \( \mathcal{P}' \) (§3), then
\[ m(A) \leq m(B_k) \]
and
\[ f[B_k] \subseteq f[A_i]; \]
\[ \sup_{i} f[B_k] \leq \sup_{i} f[A_i]; \]
and
\[ \inf_{i} f[B_k] \geq \inf_{i} f[A_i]. \]

So setting
\[ a_i = \inf_{i} f[A_i] \text{ and } b_{ik} = \inf_{i} f[B_{ik}], \]
we get
\[ S(f, \mathcal{P}') \geq \sum_{i} a_i m A_i = \sum_{i} \sum_{k} a_i m B_{ik} \]
\[ \leq \sum_{i,k} b_{ik} m B_{ik} = S(f, \mathcal{P}). \]

Similarly,
\[ S(f, \mathcal{P}'') \leq S(f, \mathcal{P}), \]
and
\[ S(f, \mathcal{P}') \leq S(f, \mathcal{P}'') \]
is obvious from (1). \( \Box \)

**Corollary 2.** For any \( \mathcal{P}' \) and \( \mathcal{P}'' \),
\[ S(f, \mathcal{P}') \leq S(f, \mathcal{P}''). \]

Hence
\[ R \int_A f \leq R \int_A f. \]

**Proof.** Let \( \mathcal{P} = \mathcal{P}' \cap \mathcal{P}'' \) (see §1). As \( \mathcal{P} \) refines both \( \mathcal{P}' \) and \( \mathcal{P}'' \), Corollary 1 yields
\[ S(f, \mathcal{P}'') \leq S(f, \mathcal{P}) \leq S(f, \mathcal{P}') \leq S(f, \mathcal{P}''). \]
Thus, indeed, no lower sum \( S(f, \mathcal{P}') \) exceeds any upper sum \( S(f, \mathcal{P}'') \).

Hence also
\[ \sup_{i} S(f, \mathcal{P}') \leq \inf_{i} S(f, \mathcal{P}''), \]
i.e.,
\[ R \int_A f \leq R \int_A f, \]
as claimed. \( \Box \)
Lemma 1. A map \( f : A \to E^1 \) is R-integrable iff \( f \) is bounded and, moreover,
\[
(3) \quad (\forall \varepsilon > 0) \exists \mathcal{P} \quad S(f, \mathcal{P}) - S(f, \mathcal{P}) < \varepsilon.
\]

Proof. By formulas (1) and (2),
\[
S(f, \mathcal{P}) \leq R \int_A f \leq R \int_A f \leq S(f, \mathcal{P}).
\]
Hence (3) implies
\[
|R \int_A f - R \int_A f| < \varepsilon.
\]
As \( \varepsilon \) is arbitrary, we get
\[
R \int_A f = R \int_A f;
\]
so \( f \) is R-integrable.

Conversely, if so, definitions (b) and (c) imply the existence of \( \mathcal{P'} \) and \( \mathcal{P''} \) such that
\[
S(f, \mathcal{P'}) > R \int_A f - \frac{1}{2} \varepsilon
\]
and
\[
S(f, \mathcal{P''}) < R \int_A f + \frac{1}{2} \varepsilon.
\]
Let \( \mathcal{P} \) refine both \( \mathcal{P'} \) and \( \mathcal{P''} \). Then by Corollary 1,
\[
S(f, \mathcal{P}) - S(f, \mathcal{P}) \leq S(f, \mathcal{P'}) - S(f, \mathcal{P'})
\]
\[
< \left( R \int_A f + \frac{1}{2} \varepsilon \right) - \left( R \int_A f - \frac{1}{2} \varepsilon \right) = \varepsilon,
\]
as required. \( \square \)

Lemma 2. Let \( f \) be C-simple; say, \( f = a_i \) on \( A_i \) for some C-partition \( \mathcal{P} = \{ A_i \} \) of \( A \) (we then write
\[
f = \sum_i a_i C_{A_i}
\]
on \( A \); see Note 4 of §4).

Then
\[
(4) \quad R \int_A f = R \int_A f = \delta(f, \mathcal{P'}) = S(f, \mathcal{P}) = \sum_i a_i mA_i.
\]
Hence any finite C-simple function is R-integrable, with \( R \int_A f \) as in (4).

§9. Riemann Integration. Stieltjes Integrals

Proof. Given any C-partition \( \mathcal{P} = \{ B_k \} \) of \( A \), consider
\[
\mathcal{P}^* \cap \mathcal{P} = \{ A_i \cap B_k \}.
\]
As \( f = a_i \) on \( A_i \cap B_k \) (even on all of \( A_i \)),
\[
a_i = \inf f[A_i \cap B_k] = \sup f[A_i \cap B_k].
\]
Also,
\[
A = \bigcup_{i,k} (A_i \cap B_k) \text{ (disjoint)}
\]
and
\[
(\forall i) \quad A_i = \bigcup_k (A_i \cap B_k);
\]
so
\[
mA_i = \sum_k m(A_i \cap B_k)
\]
and
\[
S(f, \mathcal{P}) = \sum_i \sum_k a_i m(A_i \cap B_k) = \sum_i a_i mA_i = S(f, \mathcal{P}^*)
\]
for any such \( \mathcal{P} \).

Hence also
\[
\sum_i a_i mA_i = \sup_{\mathcal{P}} S(f, \mathcal{P}) = R \int_A f.
\]
Similarly for \( R \int_A f \). 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 \( S \) and \( \bar{S} \) are integrals of C-simple maps, and definition (b) can be restated:
\[
R \int_A f = \sup \{ R \int g : g \text{ is bounded on } A \} \quad \text{and} \quad R \int_A f = \inf \{ R \int h : h \text{ is bounded on } A \}
\]
taking the sup and inf over all C-simple maps \( g, h \) with
\[
g \leq f \leq h \text{ on } A.
\]
(Verify by properties of glb and lub!)

Therefore, we can now develop R-integration as in §§4–5, replacing elementary maps by C-simple maps, with \( S = E^m \). 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 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 1. If \( f_i \to f \) (uniformly) on \( A \) and if the \( f_i \) are \( R \)-integrable on \( A \), so also is \( f \). Moreover,
\[
\lim_{i \to \infty} R \int_A |f - f_i| = 0 \quad \text{and} \quad \lim_{i \to \infty} R \int_A f_i = R \int_A f.
\]

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
\[
(\forall i \geq k) \quad |f - f_i| < \frac{\varepsilon}{mA} \quad \text{on} \quad A.
\]

Verify that
\[
(\forall i \geq k) \quad |\mathfrak{S}(f, \mathcal{P})| < \varepsilon \quad \text{and} \quad |\mathfrak{S}(f - f_i, \mathcal{P})| < \varepsilon;
\]

fix one such \( f_i \) and choose a \( \mathcal{P} \) such that
\[
\mathfrak{S}(f, \mathcal{P}) - \mathfrak{S}(f, \mathcal{P}) < \varepsilon,
\]

which one can do by Lemma 1. Then for this \( \mathcal{P} \),
\[
\mathfrak{S}(f, \mathcal{P}) - \mathfrak{S}(f, \mathcal{P}) < 3\varepsilon.
\]

(Why?) By Lemma 1, then, \( f \) is \( R \)-integrable on \( A \).

Finally,
\[
R \int_A f = R \int_A f_i \leq R \int_A |f - f_i| \leq R \int_A \left( \frac{\varepsilon}{mA} \right) = mA \left( \frac{\varepsilon}{mA} \right) = \varepsilon
\]

for all \( i \geq k \). Hence the second clause of our theorem follows, too. \( \square \)

Corollary 3. If \( f: E^1 \to E^1 \) 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 \( C \)-simple maps.

Now, let \( f \) be regulated (e.g., of the kind specified above).

Then by Lemma 2 of Chapter 5, §10,
\[
f = \lim_{i \to \infty} g_i \quad \text{(uniformly)}
\]

for finite \( C \)-simple \( g_i \).

Thus \( f \) is \( R \)-integrable on \( A \) by Theorem 1. \( \square \)

§9. Riemann Integration. Stieltjes Integrals

II. Henceforth, we assume that \( m \) is a measure on a \( \sigma \)-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 §§4-10 in full.) The measure \( m \) may, but need not, be Lebesgue measure in \( E^n \).

Theorem 2. If \( f: E^n \to E^1 \) is \( R \)-integrable on \( A \in \mathcal{C} \), it is also Lebesgue integrable (with respect to \( m \) as above) on \( A \), and
\[
L \int_A f = R \int_A f.
\]

Proof. Given a \( \mathcal{C} \)-partition \( \mathcal{P} = \{ A_i \} \) of \( A \), define the \( \mathcal{C} \)-simple maps
\[
g = \sum a_i C_{A_i} \quad \text{and} \quad h = \sum b_i C_{A_i},
\]

with
\[
a_i = \inf f[A_i] \quad \text{and} \quad b_i = \sup f[A_i].
\]

Then \( g \leq f \leq h \) on \( A \) with
\[
\mathfrak{S}(f, \mathcal{P}) = \sum a_i m_{A_i} = L \int_A g
\]

and
\[
\mathfrak{S}(f, \mathcal{P}) = \sum b_i m_{A_i} = L \int_A h.
\]

By Theorem 1(c) in §5,
\[
\mathfrak{S}(f, \mathcal{P}) = L \int_A g \leq L \int_A f \leq L \int_A h = \mathfrak{S}(f, \mathcal{P}).
\]

As this holds for any \( \mathcal{P} \), we get
\[
(5) \quad R \int_A f = \sup \mathfrak{S}(f, \mathcal{P}) \leq L \int_A f \leq \mathfrak{S}(f, \mathcal{P}) = R \int_A f.
\]

But by assumption,
\[
R \int_A f = R \int_A f.
\]

Thus these inequalities become equations:
\[
R \int_A f = \int_A f = R \int_A f.
\]
Integrability is with respect to the measure \( m \).

This also shows that an R-integral, when one exists, is always finite.\(^3\)

Thus

\[
\int_A f = \int_A f \neq \pm \infty,
\]

i.e., \( f \) is Lebesgue integrable, and

\[
L \int_A f = R \int_A f,
\]

as claimed. \( \square \)

**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 \( C \)-partitions involve intervals containing both rationals (on which \( f = 1 \)) and irrationals (on which \( f = 0 \)). Thus for any \( P \),

\[
\mathcal{S}(f, P) = 0 \text{ and } \mathcal{S}(f, P) = 1 \cdot mA = 1.
\]

(Why?) So

\[
R \int_A f = \inf \mathcal{S}(f, P) = 1,
\]

while

\[
R \int_A f = 0 \neq R \int_A f.
\]

**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* \( |P| \) of a \( C \)-partition \( P = \{A_1, \ldots, A_k\} \) is the largest of the diagonals \( dA_1 \):

\[
|P| = \max \{dA_1, dA_2, \ldots, dA_k\}.
\]

**Note 4.** For any \( A \in \mathcal{C} \), there is a sequence of \( C \)-partitions \( \mathcal{P}_k \) such that

(i) each \( P_{k+1} \) refines \( P_k \)

(ii) \( \lim_{k \to \infty} |P_k| = 0 \).

\( ^3 \) This also shows that an R-integral, when one exists, is always *finite*.

§9. Riemann Integration. Stieltjes Integrals

To construct such a sequence, bisect the edges of \( A \) so as to obtain \( 2^n \) subintervals of diagonal \( \frac{1}{2^n} dA \) (Chapter 3, §7). Repeat this with each of the subintervals, and so on. Then

\[
|P_k| = \frac{dA}{2^k} \to 0.
\]

**Lemma 3.** Let \( f : A \to E \) be bounded. Let \( \{P_k\} \) satisfy (i) of Note 4. If \( P_k = \{A_1, \ldots, A_{q_k}\} \), put

\[
g_k = \sum_{i=1}^{q_k} C_{A_i} \inf f[A_i]
\]

and

\[
h_k = \sum_{i=1}^{q_k} C_{A_i} \sup f[A_i].
\]

Then the functions

\[
g = \sup_k g_k \text{ and } h = \inf_k h_k
\]

are Lebesgue integrable on \( A \),\(^4\) and

\[
\int_A g = \lim_{k \to \infty} \mathcal{S}(f, P_k) \leq R \int_A f \leq R \int_A f \leq \lim_{k \to \infty} \mathcal{S}(f, P_k) = \int_A h.
\]

**Proof.** As in Theorem 2, we obtain \( g_k \leq f \leq h_k \) on \( A \) with

\[
\int_A g_k = \mathcal{S}(f, P_k)
\]

and

\[
\int_A h_k = \mathcal{S}(f, P_k).
\]

Since \( P_{k+1} \) refines \( P_k \), it also easily follows that

\[
g_k \leq g_{k+1} \leq \sup_k g_k = g \leq f \leq h = \inf_k h_k \leq h_{k+1} \leq h_k.
\]

(Verify!) Thus \( \{g_k\} \uparrow \) and \( \{h_k\} \downarrow \), and so

\[
g = \sup_k g_k = \lim_{k \to \infty} g_k \text{ and } h = \inf_k h_k = \lim_{k \to \infty} h_k.
\]

Also, as \( f \) is bounded,

\[
(\exists K \in E^1) \ \ |f| < K \text{ on } A.
\]

\( ^4 \) Integrability is with respect to the measure \( m \) mentioned above.
The definition of \( g_k \) and \( h_k \) then implies
\[
(\forall k) \quad |g_k| \leq K \text{ and } |h_k| \leq K \quad \text{(why?)},
\]
with
\[
\int_A (K) = K \cdot m_A < \infty.
\]
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
\[
\int_A g = \lim_{k \to \infty} \int_A g_k \text{ and } \int_A h = \lim_{k \to \infty} \int_A h_k.
\]
As
\[
\int_A g_k = S(f, \mathcal{P}_k) \leq R \int_A f
\]
and
\[
\int_A h_k = \overline{S}(f, \mathcal{P}_k) \geq R \int_A f,
\]
passage to the limit in equalities yields (6). Thus the lemma is proved. \( \square \)

**Lemma 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
\[
G_{kp}(\delta_k) \subseteq A_{kp}.
\]
Also, by the definition of \( g_k \) and \( h_k \),
\[
g_k(p) = \inf f[A_{kp}] \text{ and } h_k = \sup f[A_{kp}].
\]
(Why?)

Now fix \( \varepsilon > 0 \). If \( g(p) = h(p) \), then
\[
0 = h(p) - g(p) = \lim_{k \to \infty} |h_k(p) - g_k(p)|;
\]
so
\[
(\exists k) \quad |h_k(p) - g_k(p)| > \inf f[A_{kp}] - \sup f[A_{kp}] < \varepsilon.
\]
As \( G_{kp}(\delta_k) \subseteq A_{kp} \), we get
\[
(\forall x \in G_{kp}(\delta_k)) \quad |f(x) - f(p)| \leq \sup f[A_{kp}] - \inf f[A_{kp}] < \varepsilon,
\]
proving continuity (clause (ii)).

For (i), given \( \varepsilon > 0 \), choose \( \delta > 0 \) so that
\[
(\forall x, y \in A \cap G_{kp}(\delta)) \quad |f(x) - f(y)| < \varepsilon.
\]
Because
\[
(\forall \delta > 0) (\exists k_0) (\forall k > k_0) \quad |\mathcal{P}_k| < \delta
\]
for \( k > k_0 \), \( A_{kp} \subseteq \mathcal{P}_k(\delta) \). Deducate that
\[
(\forall k > k_0) \quad |h_k(p) - g_k(p)| \leq \varepsilon. \quad \square
\]

**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 3.** A map \( f : A \to E^1 \) is R-integrable on \( A \) (with \( m = \text{Lebesgue measure} \) 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
\[
\int_A g = \int_A h
\]
(Corollary 2 in §5).

Hence by formula (6), \( f \) is R-integrable on \( A \).

Conversely, if so, use Lemma 1 with
\[
\varepsilon = 1, \frac{1}{2}, \ldots, \frac{1}{k}, \ldots
\]
to get for each \( k \) some \( \mathcal{P}_k \) such that
\[
\overline{S}(f, \mathcal{P}_k) - \underline{S}(f, \mathcal{P}_k) < \frac{1}{k} \to 0.
\]
By Corollary 1, this will still 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
\[
\overline{S}(f, \mathcal{P}_k) - \underline{S}(f, \mathcal{P}_k) \to 0,
\]
formula (6) shows that
\[
\int_A g = \lim_{k \to \infty} \overline{S}(f, \mathcal{P}_k) = \lim_{k \to \infty} \underline{S}(f, \mathcal{P}_k) = \int_A h,
\]
As \( h \) and \( g \) are integrable on \( A \),

\[
\int_A (h - g) = \int_A h - \int_A g = 0.
\]

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 \( A - Q' - B \),

with \( mB = 0 \) (Note 5). Let \( Q = Q' \cup B \). Then \( mQ = 0 \) and

\[
A - Q = A - Q' - B;
\]

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 \( f: E^n \to E^a \) \((C^a)\).** 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

\[
R \int_A f = \sum_{k=1}^s \epsilon_k R \int_A f_k.
\]

A complex function \( f \) is R-integrable iff \( f_{\text{re}} \) and \( f_{\text{im}} \) are, and then

\[
R \int_A f = R \int_A f_{\text{re}} + i R \int_A f_{\text{im}}.
\]

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 \( s_\alpha \) or the LS measure \( m_\alpha \), where \( \alpha \uparrow \) (Chapter 7, §5, Example (b), and Chapter 7, §9).

Then

\[
R \int_A f \, \text{d}m
\]

is called the Riemann–Stieltjes (RS) integral of \( f \) with respect to \( \alpha \), also written

\[
R \int_A f \, \text{d}\alpha \quad \text{or} \quad R \int_a^b f(x) \, \text{d}\alpha(x)
\]

(\text{the latter if } A = [a,b]; \ f \text{ and } \alpha \text{ are called the integrand and integrator, respectively.})

If \( \alpha(x) = x \), \( m_\alpha \) becomes the Lebesgue measure, and

\[
R \int f(x) \, \text{d}\alpha(x)
\]

turns into

\[
R \int f(x) \, \text{d}x.
\]

Our theory still remains valid; only Theorem 3 now reads as follows.

**Corollary 4.** If \( f \) is bounded and a.e. continuous on \( A = [a,b] \) (under an LS measure \( m_\alpha \)) then

\[
R \int_a^b f \, \text{d}\alpha
\]

exists. The converse holds if \( \alpha \) is continuous on \( A \).

For by Notes 5 and 6, the “only if” in Theorem 3 holds if \( m_\alpha B = 0 \). Here \( B \) consists of countably many endpoints of partition subintervals. But (see Chapter 7, §9) \( m_\alpha \{p\} = 0 \) if \( \alpha \) 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 \( m_\alpha \), 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 \( \sigma_\alpha \) of Problem 9 in Chapter 7, §4, instead of \( s_\alpha \) or \( m_\alpha \). We reserve the name “Stieltjes integrals,” denoted

\[
S \int_a^b f \, \text{d}\alpha
\]

for such integrals, and “RS-integrals” for those based on \( m_\alpha \) or \( s_\alpha \) (this terminology is not standard).

Observe that \( \sigma_\alpha \) need not be \( \geq 0 \). 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).

**Problems on Riemann and Stieltjes Integrals**

1. Replacing \( \mathcal{M} \) by \( \mathcal{C}_1 \), and “elementary and integrable” or “elementary and nonnegative” by “\( \mathcal{C} \)-simple,” prove Corollary 1(ii)(iv)(vii) and Theorems 1(i) and 2(ii), all in §4, and do Problem 5–7 in §4, for R-integrals.

2. Verify Note 1.
2'. Do Problems 5–7 in §5 for R-integrals.

3. Do the following for R-integrals.
   (i) Prove Theorems 1(a)–(g) and 2, both in §5 (C-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 (1'), (1''), and (1) of Definition 1 in §5.
        [Hint: Use Problems 1 and 2.]

4. Fill in all details in the proof of Theorem 1, Lemmas 3 and 4, and
   Corollary 4.

5. For $f, g: E^m \to C^*$, 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.

6. Prove that if $f: A \to C^*$ is bounded and a.e. continuous on $A$, then
   \[
   R \int_A |f| \geq R \int_A f.
   \]
   For $m = \text{Lebesgue measure}$, do it assuming R-integrability only.

7. Prove that if $f, g: A \to E^1$ are R-integrable, then
   (i) so is $f^2$, and
   (ii) so is $fg$.
   [Hints: (i) Use Lemma 1. Let $h = |f| \leq K < \infty$ on $A$. Verify that
           \[
           (\inf h(A_i))^2 = \inf f^2(A_i) \quad \text{and} \quad (\sup h(A_i))^2 = \sup f^2(A_i);
           \]
           so
           \[
           \sup f^2(A_i) - \inf f^2(A_i) = (\sup h(A_i) + \inf h(A_i)) \cdot (\sup h(A_i) - \inf h(A_i)) \leq 2K.
           \]
   (ii) Use
           \[
           f g = \frac{1}{4} (f + g)^2 - (f - g)^2.
           \]
   (iii) For $m = \text{Lebesgue measure}$, do it using Theorem 3.

8. Prove that if $m = \text{the volume function } v$ (or LS function $s_{\alpha}$ for a
   continuous $\alpha$), then in formulas (1) and (2), one may replace $A_i$ by $\overline{A_i}$
   (closure of $A_i$).
   [Hint: Show that here $mA = mA$.
   \[
   R \int_{\overline{A_i}} f = R \int_{A_i} f;
   \]
   and additivity works even if the $A_i$ have some common “faces” (only their interiors
   being disjoint).]

9. (Riemann sums.) Instead of $S$ and $\overline{S}$, Riemann used sums
   \[
   S(f, \mathcal{P}) = \sum_i f(x_i) dm_{A_i},
   \]
   where $m = v$ (see Problem 8) and $x_i$ is arbitrarily chosen from $\overline{A_i}$.
   For a bounded $f$, prove that
   \[
   r = R \int_A f dm
   \]
   exists on $A = [a, b]$ iff for every $\varepsilon > 0$, there is $\mathcal{P}_\varepsilon$ such that
   \[
   |S(f, \mathcal{P}) - r| < \varepsilon
   \]
   for every refinement
   \[
   \mathcal{P} = \{A_i\}
   \]
   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).]

10. Replacing $m$ by the $\sigma_\alpha$ 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,
    \[
    S \int_a^b f \, d\alpha \quad \text{(or } S \int_a^b f \, d\alpha_\alpha).\n    \]
    Here $f, \alpha: E^1 \to E^1$ (monotone or not; even $f, \alpha: E^1 \to C$ will do).
    Prove that if $\alpha: E^1 \to E^1$ is continuous and $\alpha_i$, then
    \[
    S \int_a^b f \, d\alpha = R \int_a^b f \, d\alpha,
    \]
    the $RS$-integral.

11. (Integration by parts.) Continuing Problem 10, prove that
    \[
    S \int_a^b f \, d\alpha
    \]
    exists iff
    \[
    S \int_a^b \alpha \, df
    \]
    does, and then
    \[
    S \int_a^b f \, d\alpha + S \int_a^b \alpha \, df = K,
    \]
    where
    \[
    K = f(b)\alpha(b) - f(a)\alpha(a).
    \]
Chapter 8. Measurable Functions. Integration

[Hints: Take any C-partition \( \mathcal{P} = \{ A_i \} \) of \([a, b]\), with
\[ \overline{A_i} = [y_{i-1}, y_i], \]
say. For any \( x_i \in \overline{A_i} \), verify that
\[ S(f, \mathcal{P}, a) = \sum f(x_i) [\alpha(y_i) - \alpha(y_{i-1})] = \sum f(x_i) \alpha(y_i) - \sum f(x_i) \alpha(y_{i-1}) \]
and
\[ K = \sum f(x_i) \alpha(y_i) - \sum f(x_{i-1}) \alpha(y_{i-1}). \]
Deduce that
\[ K - S(f, \mathcal{P}, a) = S(a, \mathcal{P}', f) = \sum \alpha(x_i) [f(x_i) - f(y_i)] - \sum \alpha(x_{i-1}) [f(y_i) - f(x_{i-1})]; \]
here \( \mathcal{P}' \) results by combining the partition points \( x_i \) and \( y_i \), so it refines \( \mathcal{P} \).
Now, if \( \int_a^b \alpha \, df \) exists, fix \( \mathcal{P}_n \) as in Problem 9 and show that
\[ |K - S(f, \mathcal{P}, a) - \int_a^b \alpha \, df| < \varepsilon \]
whenever \( \mathcal{P} \) refines \( \mathcal{P}_n \).

12. If \( \alpha : E^1 \to E^1 \) is of class CD1 on \([a, b]\) and if
\[ \int_a^b f \, d\alpha \]
exists (see Problem 10), it equals
\[ \int_a^b f(x) \alpha'(x) \, dx. \]
[Hints: Set \( \phi = f \alpha' \), \( \mathcal{P} = \{ A_i \} \), \( \overline{A_i} = [a_{i-1}, a_i] \). Then
\[ S(\phi, \mathcal{P}) = \sum f(x_i) \alpha'(x_i) (a_i - a_{i-1}), \quad x_i \in \overline{A_i}, \]
and (Corollary 3 in Chapter 5, §2)
\[ S(f, \mathcal{P}, a) = \sum f(x_i) [\alpha(a_i) - \alpha(a_{i-1})] = \sum f(x_i) \alpha'(q_i), \quad q_i \in A_i. \]
As \( f \) is bounded and \( \alpha' \) is uniformly continuous on \([a, b]\) (why?), deduce that
\( (\forall \varepsilon > 0) \, (\exists \mathcal{P}_n) \, (\forall \mathcal{P} \) refining \( \mathcal{P}_n \))
\[ |S(\phi, \mathcal{P}) - S(f, \mathcal{P}, a)| < \frac{\varepsilon}{2} \]
and
\[ |S(f, \mathcal{P}, a) - S \int_a^b f \, d\alpha| < \frac{\varepsilon}{2}. \]
Proceed. Use Problem 9.]

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
\[ \int_a^b f \, d\alpha \]
exists, then \( (\exists c \in [p, q]) \) such that
\[ \int_a^b f \, d\alpha = c [\alpha(b) - \alpha(a)]. \]
Similarly, if
\[ R \int_a^b f \, d\alpha \]
exists, then \( (\exists c \in [p, q]) \) such that
\[ R \int_a^b f \, d\alpha = c [\alpha(b) - \alpha(a)]. \]
(ii) If \( f \) also has the Darboux property on \( A \), then \( f = f(x_0) \) for some \( x_0 \in A \).
(iii) If \( \alpha \) is continuous, and \( f \uparrow \) on \( A \), then
\[ \int_a^b f \, d\alpha = [f(b) \alpha(b) - f(a) \alpha(a)] - \int_a^b \alpha \, df \]
equals (Corollary 3 in Chapter 5, §1.)
[Hints: (i) As \( \alpha \uparrow \), we get
\[ p [\alpha(b) - \alpha(a)] \leq \int_a^b f \, d\alpha \leq q [\alpha(b) - \alpha(a)]. \]
(Why?) Now argue as in §6, Theorem 3 and Problem 2.
(ii) Use Problem 11, and apply (i) to \( f \alpha \).
(iii) By Theorem 2 of Chapter 5, §10, \( g \) has a primitive \( \beta \in CD^1 \). Apply Problem 12 to \( S \int_a^b f \, d\beta \).]

§10. Riemann Integration. Stieltjes Integrals

Let \((S, \mathcal{M}, s)\) be a generalized measure space. By Note 1 in §3, a map \( f \) is
\( s \)-measurable iff it is \( \nu_s \)-measurable. This naturally leads us to the following
Definition 1.

A map \( f: S \to E \) 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 \( f(x) \) is \( A \)-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 \( p f(x) \) is \( A \)-defined, i.e., at least two of \( p, f(x), \) and \( sA \) are scalars.

Note 2. If \( s \) is a measure (\( \geq 0 \)), 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 1. If \( m' \) and \( m'' \) are measures, with \( m' \geq m'' \) on \( \mathcal{M} \), then

\[
\int_A |f| \ dm' \geq \int_A |f| \ dm''
\]

for all \( A \in \mathcal{M} \) and any \( f: S \to E \).

Proof. First, take any elementary and nonnegative map \( g \geq |f| \);

\[
g = \sum_i C_A a_i \text{ on } A.
\]

Then (§4)

\[
\int_A g \ dm' = \sum_i a_i \ m' A_i \geq \sum_i a_i \ m'' A_i = \int_A g \ dm''.
\]

Hence by Definition 1 in §5,

\[
\int_A |f| \ dm' = \inf_{g \geq |f|} \int_A g \ dm' \geq \inf_{g \geq |f|} \int_A g \ dm'' = \int_A |f| \ dm'',
\]

as claimed. \( \square \)

Lemma 2.

(i) If \( s: \mathcal{M} \to E^+(C^n) \) with \( s = (s_1, \ldots, s_n) \), 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 \( A \) 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 \( s = (s_1, \ldots, s_n) \), then (Problem 4 of Chapter 7, §11)

\[
v_{s_k} \geq v_{s_k}, \quad k = 1, \ldots, n,
\]

Hence by Definition 1 and Lemma 1, the \( s \)-integrability of \( f \) implies

\[
\infty > \int_A |f| \ dv_{s_k} \geq \int_A |f| \ dv_s.
\]

Also, \( f \) is \( v_{s_k} \)-measurable, i.e., \( \mathcal{M} \)-measurable on \( A - Q \), with

\[
0 = v_s Q \geq v_{s_k} Q \geq 0.
\]

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 (\( \geq 0 \)) and \( v_s = s^+ + s^- \), so that both

\[
v_s \geq s^+ = v_{s_k} \quad \text{and} \quad v_s \geq s^- = v_{s_k}.
\]

Thus the desired result follows exactly as in part (i) of the proof. \( \square \)

We leave Note 3 as an exercise.

Definition 2.

If \( f \) is \( s \)-integrable on \( A \in \mathcal{M} \), we set

(i) in the case \( s: \mathcal{M} \to E^* \),

\[
\int_A f \ ds = \int_A f \ ds^+ - \int_A f \ ds^-,
\]

with \( s^+ \) and \( s^- \) as in formula (3) of Chapter 7, §11;\(^1\)

(ii) in the case \( s: \mathcal{M} \to E^0 (C^n) \),

\[
\int_A f \ ds = \sum_{k=1}^n \tilde{c}_k \int_A f \ ds_k,
\]

with \( \tilde{c}_k \) as in Theorem 2 of Chapter 3, §§1–3;

(iii) if \( s: \mathcal{M} \to C \),

\[
\int_A f \ ds = \int_A f \ ds_{re} + i \cdot \int_A f \ ds_{im}.
\]

(See also Problems 2 and 3.)

Note 4. If \( s \) is a measure, then

\[
s = s^+ = s_{k_e} = s_4
\]

and

\[
0 = s^- = s_{k_m} = s_2;
\]

\(^1\) By choosing \( s^+ \) and \( s^- \) as in formula (3) of Chapter 7, §11, we avoid ambiguity.
so Definition 2 agrees with our previous definitions. Similarly for \( s: \mathcal{M} \rightarrow E^n (C^n) \).

Below, \( s, t, \) and \( u \) are generalized measures on \( \mathcal{M} \) as in Definition 2, while \( f, g: S \rightarrow E \) are functions, with \( E \) a complete normed space, as in Note 1.

**Theorem 1.** The linearity, additivity, and \( \sigma \)-additivity properties (as in §7, Theorems 2 and 3) also apply to integrals

\[
\int_A f \, ds,
\]

with \( s \) as in Definition 2.

**Proof.** (i) **Linearity:** Let \( f, g: S \rightarrow E \) 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,

\[
\int_A (pf + qg) \, ds = \int_A (pf + qg) \, ds^+ - \int_A (pf + qg) \, ds^- = p \int_A f \, ds^+ + q \int_A g \, ds^+ - p \int_A f \, ds^- - q \int_A g \, ds^- = p \int_A f \, ds + q \int_A g \, ds.
\]

Thus linearity holds for signed measures. Via components, it now follows for \( s: \mathcal{M} \rightarrow E^n (C^n) \) as well. Verify!

(ii) **Additivity and \( \sigma \)-additivity** follow in a similar manner. \( \square \)

**Corollary 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

\[
\int_A f \, ds = c \cdot sA.
\]

(ii) If

\[
f = \sum_i a_i C_{Ai}
\]

for an \( \mathcal{M} \)-partition \( \{A_i\} \) of \( A \), then

\[
\int_A f \, ds = \sum_i a_i sA_i \quad \text{and} \quad \int_A |f| \, ds = \sum_i |a_i| sA_i
\]

(both series absolutely convergent).

\( \square \)

§10. Integration in Generalized Measure Spaces

(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 \((v_s) \)-\( \sigma \)-finite (Definition 4 in Chapter 7, §5).

(vi) \( \int_A f \, ds = \int_{A-Q} f \, ds \) if \( v_s Q = 0 \) 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

\[
\int_A f \, ds = \int_A f \, ds^+ - \int_A f \, ds^- = c (s^+ A - s^- A) = c \cdot sA,
\]

as required.

For \( s: \mathcal{M} \rightarrow E^n (C^n) \), the result now follows via components. (Verify!)

(ii) As \( f = a_i \) on \( A_i \), clause (i) yields

\[
\int_{A_i} f \, ds = a_i s A_i, \quad i = 1, 2, \ldots.
\]

Hence by \( \sigma \)-additivity,

\[
\int_A f \, ds = \sum_i \int_{A_i} f \, ds = \sum_i a_i s A_i,
\]

as claimed.

Clauses (iii), (iv), and (v) follow by Corollary 1 in §5 and Theorem 1(b)(ii) there, as \( v_s \) is a measure; (vi) is proved as §5, Corollary 2. We leave (vii) as an exercise. \( \square \)

**Theorem 2** (dominated convergence). If

\[
f = \lim_{i \rightarrow \infty} f_i \ \text{(pointwise)}
\]

on \( A - Q \ (v_s Q = 0) \) and if each \( f_i \) is \( s \)-integrable on \( A \), so is \( f \), and

\[
\int_A f \, ds = \lim_{i \rightarrow \infty} \int_A f_i \, ds,
\]

all provided that

\[
(\forall i) \quad |f_i| \leq g
\]

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.

\(^2\) That is, on \( A - Q, v_s Q = 0 \).
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
\[
\int_A |f_i| \, ds^+ \leq \int_A |f_i| \, dv \leq \int_A g \, dv < \infty;
\]
similarly for
\[
\int_A |f_i| \, ds^-.
\]
As \( s^+ \) and \( s^- \) are measures, Theorem 5 of §6 yields
\[
\int_A f \, ds = \int_A f \, ds^+ - \int_A f \, ds^- = \lim \left( \int_A f_i \, ds^+ - \int_A f_i \, ds^- \right) = \lim \int_A f_i \, ds.
\]
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. \( \square \)

**Theorem 3** (uniform convergence). If \( f_i \to f \) (uniformly) on \( A - Q \) \((v, A < \infty, v, Q = 0)\), and if each \( f_i \) is \( s \)-integrable on \( A \), so is \( f \), and
\[
\int_A f \, ds = \lim_{i \to \infty} \int_A f_i \, ds.
\]

**Proof.** Argue as in Theorem 2, replacing §6, Theorem 5, by §7, Lemma 1. \( \square \)

Our next theorem shows that integrals behave linearly with respect to measures.

**Theorem 4.** Let \( t, u : \mathcal{M} \to E^* (E^n, C^n) \), with \( v \leq \infty \) on \( \mathcal{M} \),\(^{3}\) and let
\[
s = pt + qu
\]
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
\[
\int_A f \, ds = p \int_A f \, dt + q \int_A f \, du.
\]

**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
\[
\int_A f \, dt = \sum_i a_i tA_i \quad \text{and} \quad \int_A f \, du = \sum_i a_i uA_i.
\]

Also, by integrability,
\[
\int_A |f| \, dv = \sum_i |a_i| v_i A_i \quad \text{and} \quad \int_A |f| \, dv = \sum_i |a_i| v_i A_i.
\]

Now, by Problem 4 in Chapter 7, §11,
\[
v_i = v_{t+u} \leq v_t + v_u;
\]
so
\[
\int_A |f| \, dv = \sum_i |a_i| v_i A_i \\
\leq \sum_i |a_i| (v_i A_i + v_u A_i) = \int_A |f| \, dv_t + \int_A |f| \, dv_u < \infty.
\]

As \( f \) is also \( \mathcal{M} \)-measurable (even elementary), it is \( s \)-integrable on \( A \) (by Definition 1), and
\[
\int_A f \, ds = \sum_i a_i sA_i = \sum_i a_i (tA_i + uA_i) = \int_A f \, dt + \int_A f \, du,
\]
as claimed.

Next, suppose \( f \) is \( \mathcal{M} \)-measurable on \( A \) and \( v, A < \infty \). By assumption, \( v_A < \infty \), too; so
\[
v_A \leq v_t A + v_u < \infty.
\]

Now, by Theorem 3 in §1,
\[
f = \lim_{i \to \infty} f_i \quad (\text{uniformly})
\]
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
\[
\int_A f_i \, ds = \int_A f_i \, dt + \int_A f_i \, du.
\]

With \( i \to \infty \), Theorem 3 yields the result.

Finally, let \( v, A = \infty \). By Corollary 1(v), we may assume (as in Lemma 3 of §7) that \( A_i \not\subset A \), with \( v_i A_i < \infty \), and \( v_i A_i < \infty \) (since \( v_i < \infty \), by assumption). Set
\[
f_i = f C_{A_i} \to f \quad \text{(pointwise)}
\]
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 \( v_i A_i < \infty \) and \( v_i A_i < \infty \), the \( f_i \) are also \( s \)-integrable (as

\(^{3}\) Or \(|s| < \infty \); see Theorem 6 in Chapter 7, §11. The restriction is redundant if \( t: \mathcal{M} \to E^n (C^n) \).
shown above), with
\[ \int_A f_i \, ds = \int_{A_i} f_i \, dt = \int_{A_i} f_i \, du = \int_A f_i \, dt + \int_A f_i \, du. \]

With \( i \to \infty \), Theorem 2 now yields the result.

To complete the proof of (b), it suffices to consider, along similar lines, the case \( s = pt \) (or \( s = qt \)). We leave this to the reader.

For (a), see Chapter 7, §11. \( \square \)

Theorem 5. If \( f \) is \( s \)-integrable on \( A \), so is \(|f|\), and
\[ |\int_A f \, ds| \leq \int_A |f| \, dv_s. \]

Proof. By Definition 1, and Theorem 1 of §1, \( f \) and \(|f|\) are \( M \)-measurable on \( A - Q, v_sQ = 0 \), and
\[ \int_A |f| \, dv_s < \infty; \]
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 \( v_sA < \infty \), and then for \( v_sA = \infty \). We omit the details. \( \square \)

Definition 3.

We write
\[ "ds = g dt \text{ in } A" \]
or
\[ "s = \int g dt \text{ in } A" \]
iff \( g \) is \( t \)-integrable on \( A \), and
\[ sX = \int_X g \, dt \]
for \( A \supseteq X, X \in M \).

We then call \( s \) the indefinite integral of \( g \) in \( A \). \( (\int_X g \, dt \) may be interpreted as in Problems 2–4 below.)

Lemma 3. If \( A \in M \) and
\[ ds = g dt \text{ in } A, \]
then
\[ dv_s = |g| \, dv_t \text{ in } A. \]

§10. Integration in Generalized Measure Spaces

Proof. By assumption, \( g \) and \(|g|\) are \( v_t \)-integrable on \( X \), and
\[ sX = \int_X g \, dt \]
for \( A \supseteq X, X \in M \). We must show that
\[ v_sX = \int_X |g| \, dv_t \]
for such \( X \).

This is easy if \( g = c \) (constant) on \( X \). For by definition,
\[ v_sX = \sup_{P} \sum_i |sX_i|, \]
over all \( M \)-partitions \( P = \{X_i\} \) of \( X \). As
\[ sX_i = \int_{X_i} g \, dt = c \cdot tX_i, \]
we have
\[ v_sX = \sup_P \sum_i |c| |tX_i| = |c| \sup_P \sum_i |tX_i| = |c| \, v_tX; \]
so
\[ v_sX = \int_X |g| \, dv_t. \]

Thus all is proved for constant \( g \).

Hence by \( \sigma \)-additivity, the lemma holds for \( M \)-elementary maps \( g \). (Why?)

In the general case, \( g \) is \( t \)-integrable on \( X \), hence \( M \)-measurable and finite on \( X - Q, v_tQ = 0 \). By Corollary 1(iii), we may assume \( g \) finite and measurable on \( X \); so
\[ g = \lim_{k \to \infty} g_k \text{ (uniformly)} \]
on \( X \) for some \( M \)-elementary maps \( g_k, \) all integrable on \( X \), with respect to \( v_t \) (and \( t \)).

Let
\[ s_k = \int g_k \, dt \]
in \( X \). By what we just proved for elementary and integrable maps,
\[ v_{s_k}X = \int_X |g_k| \, dv_t, \quad k = 1, 2, \ldots. \]

Now, if \( v_tX < \infty \), Theorem 3 yields
\[ \int_X |g| \, dv_t = \lim_{k \to \infty} \int_X |g_k| \, dv_t = \lim_{k \to \infty} v_{s_k}X = v_sX. \]
(see Problem 6). Thus all is proved if \( v_1X < \infty \).

If, however, \( v_1X = \infty \), argue as in Theorem 4 (the last step), using the left continuity of \( v_s \) and of

\[
\int |g| \, dv_t.
\]

Verify! \( \square \)

**Theorem 6** (change of measure). If \( f \) is \( s \)-integrable on \( A \in \mathcal{M} \), with

\[
ds = g \, dt \text{ in } A,
\]

then (subject to Note 1) \( fg \) is \( t \)-integrable on \( A \) and

\[
\int_A f \, ds = \int_A f g \, dt.
\]

(\textit{Note the formal substitution of “}g \, dt\text{” for “}ds\text{”})

**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

\[
0 = v_s Q = \int_Q |g| \, dv_t
\]

by Lemma 3. Hence \( 0 = fg \) on \( Q - Z \), \( v_t Z = 0 \). Therefore,

\[
\int_Q fg \, dt = 0 = \int_Q f \, ds
\]

for \( v_s Q = 0 \). Thus we may neglect \( Q \) and assume that \( f \) is finite and \( \mathcal{M} \)-measurable on \( A \).

As \( ds = g \, dt \), Definition 3 and Lemma 3 yield

\[
v_s A = \int_A |g| \, dv_t < \infty.
\]

Also (Theorem 3 in Chapter 8, \( \S \)),

\[
f = \lim_{k \to \infty} f_k \quad (\text{uniformly})
\]

for elementary maps \( f_k \), all \( v_s \)-integrable on \( A \) (Lemma 2 in \( \S \)). As \( g \) is bounded, we get on \( A \)

\[
f g = \lim_{k \to \infty} f_k g \quad (\text{uniformly}).
\]

Moreover, as the theorem holds for \textit{elementary and integrable} maps, \( f_k g \) is \( t \)-integrable on \( A \), and

\[
\int_A f_k \, ds = \int_A f_k g \, dt, \quad k = 1, 2, \ldots.
\]

Since \( v_s A < \infty \) and \( v_t A < \infty \), Theorem 3 shows that \( fg \) is \( t \)-integrable on \( A \), and

\[
\int_A f \, ds = \lim_{k \to \infty} \int_A f_k \, ds = \lim_{k \to \infty} \int_A f_k g \, dt = \int_A f g \, dt.
\]

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 \( A_i \supset A \), with \( v_t A_i < \infty \) (\( \forall i \)).

Now for \( i = 1, 2, \ldots \) set

\[
g_i = \begin{cases} g & \text{on } A_i (|g| \leq i), \\ 0 & \text{elsewhere.} \end{cases}
\]

Then each \( g_i \) is \textit{bounded},

\[
g_i \to g \quad (\text{pointwise}),
\]

and

\[
|g_i| \leq |g|
\]

on \( A \). We also set \( f_i = f C_{A_i} \); so \( f_i \to f \) (pointwise) and \( |f_i| \leq |f| \) on \( A \). Then

\[
\int_A f_i \, ds = \int_{A_i} f_i \, ds = \int_{A_i} f_i g_i \, dt = \int_A f g_i \, dt.
\]

(Why?) Since \( |f_i g_i| \leq |f g| \) and \( f_i g_i \to f g \), the result follows by Theorem 2. \( \square \)

**Problems on Generalized Integration**

Recall that in this section \( E \) is assumed to be a \textit{complete} normed space.

1. Fill in the missing details in the proofs of this section. Prove Note 3.

2. Treat Corollary 1(ii) as a \textit{definition} of

\[
\int_A f \, ds
\]

for \( s: \mathcal{M} \to E \) and \textit{elementary and integrable} \( f \), even if \( E \not= E^n (C^n) \). Hence deduce Corollary 1(i)(vi) for this more general case.

3. Using Lemma 2 in \( \S \), with \( m = v_s \), \( s: \mathcal{M} \to E \), construct

\[
\int_A f \, ds
\]
as in Definition 2 of §7 for the case \( v_A \neq \infty \). Show that this agrees with Problem 2 if \( f \) is elementary and integrable. Then prove linearity for functions with \( v_A \)-finite support as in §7.

4. Define

\[
\int_A f \, ds \quad (s: \mathcal{M} \to E)
\]

also for \( v_A = \infty \).

[Hint: Set \( m = v_A \) in Lemma 3 of §7]

5. Prove Theorems 1 to 3 for the general case, \( s: \mathcal{M} \to E \) (see Problem 4).

[Hint: Argue as in §7]

5'. From Problems 2–4, deduce Definition 2 as a theorem in the case \( E = E^\infty (C^n) \).

6. Let \( s, s_k: \mathcal{M} \to E \ (k = 1, 2, \ldots) \) be any set functions. Let \( A \in \mathcal{M} \) and

\[
\mathcal{M}_A = \{ X \in \mathcal{M} \mid X \subseteq A \}.
\]

Prove that if

\[
(\forall X \in \mathcal{M}_A) \lim_{k \to \infty} s_k X = sX,
\]

then

\[
\lim_{k \to \infty} v_{s_k} A = v_s A,
\]

provided \( \lim v_{s_k} \) exists.

[Hint: Using Problem 2 in Chapter 7, §11, fix a finite disjoint sequence \( \{X_i\} \subseteq \mathcal{M}_A \). Then

\[
\sum_i |s_{X_i}| \leq \lim_{k \to \infty} \sum_i |s_k X_i| \leq \lim_{k \to \infty} v_{s_k} A.
\]

Infer that

\[
v_{s_k} A \leq \lim_{k \to \infty} v_{s_k} A.
\]

Also,

\[
(\forall \varepsilon > 0) (\exists k_0) (\forall k > k_0) \sum_i |s_k X_i| \leq \sum_i |s_{X_i}| + \varepsilon \leq v_{s_k} A + \varepsilon.
\]

Proceed.]

7. Let \( (X, \mathcal{M}, m) \) and \( (Y, \mathcal{N}, n) \) be two generalized measure spaces \( (X \in \mathcal{M}, Y \in \mathcal{N}) \) such that \( mn \) is defined (Note 1). Set

\[
\mathcal{C} = \{ A \times B \mid A \in \mathcal{M}, B \in \mathcal{N}, v_m A \cdot n B < \infty, v_{n_B} B < \infty \}
\]

and \( s(A \times B) = m A \cdot n B \) 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 \( \sigma \)-additivity, use our present Theorem 2 twice. Omit \( P^* \).

8. Continuing Problem 7, let \( \mathcal{P} \) be the \( \sigma \)-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.]

9. Further continuing Problems 7 and 8, define

\[
(\forall D \in \mathcal{P}) \ pD = \int_X \int_Y C_D \, dn \, dm.
\]

Show that \( p \) is a generalized measure, with \( p = s \) on \( \mathcal{C} \), and that

\[
(\forall D \in \mathcal{P}) \int_X f \, dp = \int_{X \times Y} f \, dp
\]

whenever \( H \in \mathcal{P} \), \( f \) is \( p \)-integrable on \( H \), and \( f = 0 \) on \( -H \).

Verify that this is unambiguous, i.e.,

\[
\int_{X \times Y} f \, dp
\]

so defined is independent of the choice of \( H \).

Finally, let \( \mathcal{P}^* \) be the completion of \( p \) (Chapter 7, §6, Problem 15); let \( \mathcal{P}^* \) be its domain.

Develop the rest of the Fubini theory “imitating” Problem 12 in §8.

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 \( \sigma^\alpha \) in an interval \( I \subseteq E^1 \) (or an LS measure \( s_\alpha = \sigma^\alpha \)), there are two monotone functions \( g^\uparrow \) and \( h^\uparrow (\alpha = g - h) \) such that

\[
m_g = s^\uparrow \text{ and } m_h = s^\downarrow.
\]

both satisfying formula (3) of Chapter 7, §11, inside \( I \).

(ii) If \( f \) is \( s_\alpha \)-integrable on \( A \subseteq I \), then

\[
\int_A f \, ds_\alpha = \int_A f \, dm_g - \int_A f \, dm_h
\]

for any \( g^\uparrow \) and \( h^\uparrow \) (finite) such that \( \alpha = g - h \).
*§11. The Radon–Nikodym Theorem. Lebesgue Decomposition*

I. As you know, the indefinite integral

\[ \int f \, dm \]

is a *generalized measure*. We now seek conditions under which a given *generalized measure* \( \mu \) can be represented as

\[ \mu = \int f \, dm \]

for some \( f \) (to be found). We start with two lemmas.

**Lemma 1.** Let \( m, \mu : \mathcal{M} \to [0, \infty) \) be finite measures in \( S \). Suppose \( S \in \mathcal{M} \), \( \mu S > 0 \) (i.e., \( \mu \not= 0 \)) and \( \mu \) is \( m \)-continuous (Chapter 7, §11).

Then there is \( \delta > 0 \) and a set \( P \in \mathcal{M} \) such that \( mP > 0 \) and

\[ (\forall X \in \mathcal{M}) \quad \mu(X) \geq \delta \cdot m(X \cap P). \]

**Proof.** As \( m < \infty \) and \( \mu S > 0 \), there is \( \delta > 0 \) such that

\[ \mu S - \delta \cdot mS > 0. \]

Fix such a \( \delta \) and define a *signed measure* (Lemma 2 of Chapter 7, §11)

\[ \Phi = \mu - \delta m, \]

so that

\[ (\forall Y \in \mathcal{M}) \quad \Phi Y = \mu Y - \delta \cdot mY; \]

hence

\[ \Phi S = \mu S - \delta \cdot mS > 0. \]

By Theorem 3 in Chapter 7, §11 (Hahn decomposition), there is a \( \Phi \)-positive set \( P \in \mathcal{M} \) with a \( \Phi \)-negative complement \( -P = S - P \in \mathcal{M} \).

Clearly, \( mP > 0 \); for if \( mP = 0 \), the \( m \)-continuity of \( \mu \) would imply \( \mu P = 0 \), hence

\[ \Phi P = \mu P - \delta \cdot mP = 0, \]

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),

\[ 0 \leq \mu Y - \delta \cdot mY. \]

Taking \( Y = X \cap P \), we get

\[ \delta \cdot m(X \cap P) \leq \mu(X \cap P) \leq \mu X, \]

as required. \( \square \)

**Lemma 2.** With \( m, \mu, \) and \( S \) as in Lemma 1, let \( \mathcal{H} \) be the set of all maps \( g : S \to E^* \); \( \mathcal{M} \)-measurable and nonnegative on \( S \), such that

\[ \int_X g \, dm \leq \mu X \]

for every set \( X \) from \( \mathcal{M} \).

Then there is \( f \in \mathcal{H} \) with

\[ \int_S f \, dm = \max_{g \in \mathcal{H}} \int_S g \, dm. \]

**Proof.** \( \mathcal{H} \) is not empty; e.g., \( g = 0 \) is in \( \mathcal{H} \). We now show that

\[ (2) \quad (\forall g, h \in \mathcal{H}) \quad g \vee h = \max(g, h) \in \mathcal{H}. \]

Indeed, \( g \vee h \geq 0 \) and \( \mathcal{M} \)-measurable on \( S \), so \( g \) 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

\[ \int_X (g \vee h) = \int_Y (g \vee h) + \int_Z (g \vee h) = \int_Y g + \int_Z h \leq \mu Y + \mu Z = \mu X, \]

proving (2).

Let

\[ k = \sup_{g \in \mathcal{H}} \int_S g \, dm \in E^*. \]

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

\[ \lim_{n \to \infty} \int_S g_n \, dm = k. \]

(Verify!) Set

\[ f = \lim_{n \to \infty} g_n. \]

It exists since \( \{g_n\} \uparrow \). By Theorem 4 in §6,

\[ k = \lim_{n \to \infty} \int_S g_n = \int_S f. \]
Also, \( f \) is \( \mathcal{M} \)-measurable and \( \geq 0 \) on \( S \), as all \( g_n \) are; and if \( X \in \mathcal{M} \), then
\[
(\forall n) \quad \int_X g_n \leq \mu X;
\]
hence
\[
\int_X f = \lim_{n \to \infty} \int_X g_n \leq \mu X.
\]
Thus \( f \in \mathcal{H} \) and
\[
\int_S f = k = \sup_{g \in \mathcal{H}} \int_S g,
\]
i.e.,
\[
\int_S f = \max_{g \in \mathcal{H}} \int_S g \leq \mu S < \infty.
\]
This completes the proof. \( \Box \)

**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 1 (Radon-Nikodym).** If \((S, \mathcal{M}, m)\) is a \( \sigma \)-finite measure space, if \( S \in \mathcal{M} \), and if \( \mu : \mathcal{M} \to E^n(C^n) \) is a generalized \( m \)-measurable continuous, then
\[
\mu = \int f \, dm \text{ on } \mathcal{M}
\]
for at least one map \( f : S \to E^n(C^n) \), \( \mathcal{M} \)-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 \( \mu \), with respect to \( m \).

**Proof.** Via components (Theorem 5 in Chapter 7, §11), all reduces to the case \( \mu : \mathcal{M} \to E^1 \).

Then Theorem 4 (Jordan decomposition) in Chapter 7, §11, yields
\[
\mu = \mu^+ - \mu^-,
\]
where \( \mu^+ \) and \( \mu^- \) are finite measures \( \geq 0 \), 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 \( S \),
\[
\int f \leq \mu < \infty,
\]
and
\[
\int_S f \, dm = k = \sup_{g \in \mathcal{H}} \int_S g \, dm.
\]
We claim that \( f \) is the required map.

Indeed, let
\[
\nu = \mu - \int f \, dm;
\]
so \( \nu \) is a finite \( m \)-continuous measure \( \geq 0 \) 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
\[
(\forall X \in \mathcal{M}) \quad \nu X \geq \delta \cdot m(X \cap P).
\]
Now let
\[
g = f + \delta \cdot C_P;
\]
so \( g \) is \( \mathcal{M} \)-measurable and \( \geq 0 \). Also,
\[
(\forall X \in \mathcal{M}) \quad \int_X g = \int_X f + \delta = \int_X C_P = \int_X f + \delta \cdot m(X \cap P) \\
\leq \int_X f + \nu(X \cap P) \\
\leq \int_X f + \nu X = \mu X
\]
by our choice of \( \delta \) and \( \nu \). Thus \( g \in \mathcal{H} \). On the other hand,
\[
\int_S g = \int_S f + \delta = \int_S C_P = \int_S f + \delta \cdot mP > k,
\]
contrary to
\[
k = \sup_{g \in \mathcal{H}} \int_S g.
\]
This proves that \( \int f = \mu \), indeed.

Now suppose there is another map \( h \in \mathcal{H} \) with
\[
\mu = \int h \, dm = \int f \, dm \neq \infty;
\]
so
\[ \int (f - h) \, dm = 0. \]
(Why?) Let
\[ Y = S(f \geq h) \quad \text{and} \quad Z = S(f < h); \]
so \( Y, Z \in \mathcal{M} \) (Theorem 3 of §2) and \( f - h \) is sign-constant on \( Y \) and \( Z \). Also, by construction,
\[ \int_Y (f - h) \, dm = 0 = \int_Z (f - h) \, dm. \]
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,
\[ mS(f \neq h) = 0. \]
Thus all is proved for the case \( mS < \infty \).
Next, let \( m \) be \( \sigma \)-finite:
\[ S = \bigcup_{k=1}^{\infty} S_k \quad \text{(disjoint)} \]
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
\[ \int_X f_k \, dm = \mu X \]
for all \( \mathcal{M} \)-sets \( X \subseteq S_k \). Fixing such an \( f_k \) for each \( k \), define \( f : S \to E^1 \) by
\[ f = f_k \quad \text{on} \quad S_k, \quad k = 1, 2, \ldots. \]
Then (Corollary 3 in §1) \( f \) is \( \mathcal{M} \)-measurable and \( \geq 0 \) on \( S \).
Taking any \( X \in \mathcal{M} \), set \( X_k = X \cap S_k \). Then
\[ X = \bigcup_{k=1}^{\infty} X_k \quad \text{(disjoint)} \]
and \( X_k \in \mathcal{M} \). Also,
\[ (\forall k) \quad \int_{X_k} f \, dm = \int_{X_k} f_k \, dm = \mu X_k. \]
Thus by \( \sigma \)-additivity (Theorem 2 in §5),
\[ \int_X f \, dm = \sum_{k=1}^{\infty} \int_{X_k} f \, dm = \sum_k \mu X_k = \mu X < \infty \quad (\mu \text{ is finite!}). \]
Thus \( f \) is as required, and its “uniqueness” follows as before. \( \square \)

**Note 2.** By Definition 3 in §10, we may write
\[ "d\mu = f \, dm" \]
for
\[ "\int f \, dm = \mu." \]

**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
\[ \mu = \int f \, dm \quad \text{and} \quad mS(f \neq h) = 0 \]
then are replaced by
\[ \mu = \int f \, ds \quad \text{and} \quad v_s S(f \neq h) = 0. \]

**II.** Theorem 1 requires \( \mu \) 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 \( s, t : \mathcal{M} \to E \) \((\mathcal{M} \subseteq 2^E)\), 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
\[ (\forall X \in \mathcal{M} \mid X \subseteq -P) \quad sX = 0. \]
(We then briefly say “\( s \) resides in \( P \).”)
For generalized measures, this means that
\[ (\forall X \in \mathcal{M}) \quad sX = s(X \cap P). \]
Why?

**Corollary 1.** If the generalized measures \( s, u : \mathcal{M} \to E \) are \( t \)-singular, so is \( k s \) for any scalar \( k \) \((\text{if} \ s \text{ is scalar valued,} \ k \text{ may be a vector})\).
So also are \( s \pm u \), provided \( t \) is additive.

(Exercise! See Problem 3 below.)

**Corollary 2.** If a generalized measure \( s : \mathcal{M} \to E \) 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 \( X \in \mathcal{M} \),
\[ s(X - P) = 0 \quad (\text{for} \ X - P \subseteq -P) \]
and
\[ v_t (X \cap P) = 0 \quad (\text{for} \ X \cap P \subseteq P). \]
As \( s \ll t \), we also have \( s(X \cap P) = 0 \) by Definition 3(i) in Chapter 7, §11. Thus by additivity,

\[
  sX = s(X \cap P) + s(X - P) = 0,
\]
as claimed. \( \square \)

**Theorem 2** (Lebesgue decomposition). Let \( s, t : \mathcal{M} \to E \) be generalized measures.

If \( v_s \) is \( t \)-finite (Definition 3(iii) in Chapter 7, §11), there are generalized measures \( s', s'' : \mathcal{M} \to E \) such that

\[
  s' \ll t \text{ and } s'' \perp t
\]

and

\[
  s = s' + s''.
\]

**Proof.** Let \( v_0 \) be the restriction of \( v_s \) to

\[
  \mathcal{M}_0 = \{ X \in \mathcal{M} \mid v_0X = 0 \}.
\]

As \( v_s \) is a measure (Theorem 1 of Chapter 7, §11), so is \( v_0 \) (for \( \mathcal{M}_0 \) is a \( \sigma \)-ring; verify!).

Thus by Problem 13 in Chapter 7, §6, we fix \( P \in \mathcal{M}_0 \), with

\[
  v_sP = v_0P = \max \{ v_sX \mid X \in \mathcal{M}_0 \}.
\]

As \( P \in \mathcal{M}_0 \), we have \( v_sP = 0 \); hence

\[
  |sP| \leq v_sP < \infty
\]

(for \( v_s \) is \( t \)-finite).

Now define \( s', s'', v' \), and \( v'' \) by setting, for each \( X \in \mathcal{M} \),

\[
  (4) \quad s'X = s(X - P);

  (5) \quad s''X = s(X \cap P);

  (6) \quad v'X = v_s(X - P);

  (7) \quad v''X = v_s(X \cap P).
\]

As \( s \) and \( v_s \) are \( \sigma \)-additive, so are \( s', s'', v' \), and \( v'' \). (Verify!) Thus \( s', s'' : \mathcal{M} \to E \) are generalized measures, while \( v' \) and \( v'' \) are measures \( \geq 0 \).

We have

\[
  (\forall X \in \mathcal{M}) \quad sX = s(X - P) + s(X \cap P) = s'X + s''X;
\]
i.e.,

\[
  s = s' + s''.
\]

Similarly one obtains \( v_s = v' + v'' \).

Also, by (5), since \( X \cap P = \emptyset \),

\[
  -P \subseteq X \text{ and } X \in \mathcal{M} \implies s''X = 0,
\]
while \( v_tP = 0 \) (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

\[
  (3Y \in \mathcal{M}) \quad v_tY = 0
\]

(i.e., \( Y \in \mathcal{M}_0 \)), but

\[
  0 < v'Y = v_s(Y - P);
\]

So by additivity,

\[
  v_s(Y \cup P) = v_sP + v_s(Y - P) > v_sP;
\]

with \( Y \cup P \in \mathcal{M}_0 \), contrary to

\[
  v_sP = \max \{ v_sX \mid X \in \mathcal{M}_0 \}.
\]

This contradiction completes the proof. \( \square \)

**Note 4.** The set function \( x'' \) in Theorem 2 is bounded on \( \mathcal{M} \). Indeed, \( s'' \perp t \) yields a set \( P \in \mathcal{M} \) such that

\[
  (\forall X \in \mathcal{M}) \quad s''(X - P) = 0;
\]

and \( v_tP = 0 \) implies \( v_sP < \infty \). (Why?) Hence

\[
  s''X = s''(X \cap P) + s''(X - P) = s''(X \cap P);
\]

As \( s = s' + s'' \), we have

\[
  |s''| \leq |s| + |s'| \leq v_s + v'_s;
\]

so

\[
  |s''X| = |s''(X \cap P)| \leq v_sP + v'_sP.
\]

But \( v'_sP = 0 \) by \( t \)-continuity (Theorem 2 of Chapter 7, §11). Thus \( |s''| \leq v_sP < \infty \) on \( \mathcal{M} \).

**Note 5.** The Lebesgue decomposition \( s = s' + s'' \) in Theorem 2 is unique. For if also

\[
  u' \ll t \text{ and } u'' \perp t
\]

and

\[
  u' + u'' = s = s' + s'';
\]

then with \( P \) as in Problem 3, \((\forall X \in \mathcal{M})\)

\[
  s''(X \cap P) + s''(X \cap P) = u'(X \cap P) + u''(X \cap P)
\]

\[\text{§11. The Radon–Nikodym Theorem. Lebesgue Decomposition 343}\]
and \( v_t(X \cap P) = 0 \). But
\[
s'(X \cap P) = 0 = u'(X \cap P)
\]
by \( t \)-continuity; so (8) reduces to
\[
s''(X \cap P) = u''(X \cap P),
\]
or \( s''X = u''X \) (for \( s'' \) and \( u'' \) reside in \( P \)). Thus \( s'' = u'' \) on \( M \).

By Note 4, we may cancel \( s'' \) and \( u'' \) in
\[
s' + s'' = u' + u''
\]
to obtain \( s' = u' \) also.

**Note 6.** If \( E = E^n(C^n) \), the \( t \)-finiteness of \( v_s \) in Theorem 2 is redundant, for \( v_s \) is even bounded (Theorem 6 in Chapter 7).

We now obtain the desired generalization of Theorem 1.

**Corollary 3.** If \( (S, M, m) \) is a \( \sigma \)-finite measure space \( (S \in M) \), then for any generalized measure
\[
\mu: M \to E^n(C^n),
\]
there is a unique \( m \)-singular generalized measure
\[
s'' : M \to E^n(C^n)
\]
and a (“essentially” unique) map
\[
f : S \to E^n(C^n),
\]
\( M \)-measurable and \( m \)-integrable on \( S \), with
\[
\mu = \int f \, dm + s''.
\]

(Note 3 applies here.)

**Proof.** By Theorem 2 and Note 5, \( \mu = s' + s'' \) for some (unique) generalized measures \( s', s'': M \to E^n(C^n) \), 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. \( \square \)

**Problems on Radon-Nikodym Derivatives and Lebesgue Decomposition**

1. Fill in all proof details in Lemma 2 and Theorem 1.  

2. Verify the statement following formula (3). Also prove the following:  
   (i) If \( P \in M \) along with \( -P \in M \), then \( s \perp t \) implies \( t \perp s \);  
   (ii) \( s \perp t \) iff \( v_s \perp v_t \).

*§11. The Radon-Nikodym Theorem. Lebesgue Decomposition*

3. Prove Corollary 1.  
   [Hints: Here \( M \) is a \( \sigma \)-ring. Suppose \( s \) and \( u \) reside in \( P' \) and \( P'' \), respectively, and \( v_1P = 0 = v_2P'' \). Let \( P = P' \cup P'' \in M \). Verify that \( v_1P = 0 \) (use Problem 8 in Chapter 7, §11). Then show that both \( s \) and \( u \) reside in \( P \).]

4. Show that if \( s : M \to E^r \) is a signed measure in \( S \in M \), then \( s^+ \perp s^- \) and \( s^- \perp s^+ \).

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 \((\geq 0)\), so are \( s' \) and \( s'' \).


*§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,
\[
m : M' \to E^r
\]
is Lebesgue measure in \( E^n \) (presupposed in such terms as “a.e.”, etc.); \( s \) is an arbitrary set function. For convenience, we set
\[
s'(\bar{p}) = 0
\]
and
\[
\int f \, dm = 0,
\]
unless defined otherwise; thus \( s' \) and \( \int f \) exist always.

We start with several lemmas that go back to Lebesgue.

**Lemma 1.** With the notation of Definition 3 of Chapter 7, §10, the functions
\[
\mathcal{D}s, \overline{D}s, \text{ and } s'
\]
are Lebesgue measurable on \( E^n \) for any set function
\[
s : M' \to E^r \quad (M' \supseteq \mathcal{K}).
\]

**Proof.** By definition,
\[
\mathcal{D}s(\bar{p}) = \inf \{ h_t(\bar{p}) : t \in T, \}
\]
where
\[ h_r(\bar{p}) = \sup \left\{ \frac{sI}{mI} \bigg| I \in \mathcal{K}_p^r \right\} \]
and
\[ \mathcal{K}_p^r = \left\{ I \in \mathcal{K} \bigg| \bar{p} \in I, \ dI < \frac{1}{r} \right\}, \quad r = 1, 2, \ldots. \]
As is easily seen (verify!),
\[ E^n(h_r > a) = \bigcup \left\{ I \in \mathcal{K} \bigg| a < \frac{sI}{mI}, \ dI < \frac{1}{r} \right\}, \quad a \in E^*. \]
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
\[ \mathcal{D}s = \inf_r h_r, \]
by Lemma 1 of §2 and Definition 3 in Chapter 7, §10. Similarly for \( \mathcal{D}s \).
Hence by Corollary 1 in §2, the set
\[ A = E^n(\mathcal{D}s = \mathcal{D}s) \]
is measurable. As \( s' = \mathcal{D}s \) on \( A \), \( s' \) is measurable on \( A \) and also on \( -A \) (by convention, \( s' = 0 \) on \( -A \)), hence on all of \( E^n \). \( \square \)

**Lemma 2.** With the same notation, let \( s : \mathcal{M}' \to E^* (\mathcal{M}' \supseteq \mathcal{K}) \) 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
\[ \mathcal{D}s > a \quad \text{on} \quad A, \]
then
\[ a \cdot mA \leq sB. \]
**Proof.** Fix \( \varepsilon > 0 \). By regularity (Definition 4 in Chapter 7, §7), there is an open set \( G \supseteq B \), with
\[ sB + \varepsilon \geq sG. \]
Now let
\[ \mathcal{K}^\varepsilon = \left\{ I \in \mathcal{K} \bigg| I \subseteq G, \ sI \leq (a - \varepsilon) mI \right\}. \]
As \( \mathcal{D}s > a \), the definition of \( \mathcal{D}s \) 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
\[ m\left( A - \bigcup_k I_k \right) = 0 \]
and
\[ mA \leq m(A - \bigcup I_k) + m \bigcup I_k = 0 + m \bigcup I_k = \sum_k mI_k. \]
As
\[ \bigcup I_k \subseteq G \quad \text{and} \quad sB + \varepsilon \geq sG \]
(by our choice of \( \mathcal{K}^\varepsilon \) and \( G \)), we obtain
\[ sB + \varepsilon \geq s \bigcup I_k = \sum_k sI_k \geq (a - \varepsilon) \sum_k mI_k \geq (a - \varepsilon) mA. \]
Thus
\[ (a - \varepsilon) mA \leq sB + \varepsilon. \]
Making \( \varepsilon \to 0 \), we obtain the result. \( \square \)

**Lemma 3.** If
\[ t = s \pm u, \]
with \( s, t, u : \mathcal{M}' \to E^* \) and \( \mathcal{M}' \supseteq \mathcal{K} \), and if \( u \) is differentiable at a point \( \bar{p} \in E^n \), then
\[ \mathcal{D}t = \mathcal{D}s \pm u' \quad \text{and} \quad \mathcal{D}t = \mathcal{D}s \pm u' \quad \text{at} \quad \bar{p}. \]

The proof, from definitions, is left to the reader (Chapter 7, §12, Problem 7).

**Lemma 4.** Any \( m \)-continuous measure \( s : \mathcal{M}' \to E^1 \) 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
\[ (\forall X \in \mathcal{M}^* \ | \ mX < \delta) \quad sX < \varepsilon. \]
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
\[ m(A - F) < \delta \quad \text{and} \quad m(G - A) < \delta. \]
Thus by our choice of \( \delta \),
\[ s(A - F) < \varepsilon \quad \text{and} \quad s(G - A) < \varepsilon, \]
as required. \( \square \)
Corollary 2

Differentiability follows by §12, but we obtain it anyway.

§ of Chapter 7, Theorem 4.

A.

= \mathcal{M}^*. \quad \text{If the } s_k \text{ are a.e. differentiable, then}

\overline{D}s = Ds = \lim_{k \to \infty} s'_k \text{ a.e.}

Proof. Let first \( s_k \not\nearrow s \). Set

\( t_k = s - s_k \).

By Corollary 2 in Chapter 7, §11, all \( t_k \) are \( m \)-continuous, hence strongly regular (Lemma 4). Also, \( t_k \searrow 0 \) (since \( s_k \not\nearrow s \)). Hence

\[ t_k |I| \geq t_{k+1} |I| \geq 0 \]

for each cube \( I \); and the definition of \( \overline{D}t_k \) implies that

\( \overline{D}t_k \geq \overline{D}t_{k+1} \geq \overline{D}t_{k+1} \geq 0 \).

As \( \{ \overline{D}t_k \} \downarrow \), \( \lim_{k \to \infty} \overline{D}t_k \) exists (pointwise). Now set

\[ A_r = E^n \left( \lim_{k \to \infty} \overline{D}t_k \geq \frac{1}{r} \right), \quad r = 1, 2, \ldots. \]

By Lemma 1 (and Lemma 1 in §2), \( A_r \in \mathcal{M}^* \). Since

\[ \overline{D}t_k \geq \lim_{k \to \infty} \overline{D}t_k \geq \frac{1}{r} \]

on \( A_r \), Lemma 2 yields

\[ \frac{1}{r} m A_r \leq t_k |I| \leq t_k A_r. \]

As \( t_k \searrow 0 \), we have

\[ \frac{1}{r} m A_r \leq \lim_{k \to \infty} t_k A_r = 0. \]

Thus

\[ m A_r = 0, \quad r = 1, 2, \ldots. \]

Also, as is easily seen,

\[ E^n \left( \lim_{k \to \infty} \overline{D}t_k > 0 \right) = \bigcup_{r=1}^{\infty} E^n \left( \lim_{k \to \infty} \overline{D}t_k \geq \frac{1}{r} \right) = \bigcup_{r=1}^{\infty} A_r \]

and

\[ m \bigcup_{r=1}^{\infty} A_r = 0. \]

Hence

\[ \lim_{k \to \infty} \overline{D}t_k \leq 0 \text{ a.e.} \]

As

\[ \overline{D}t_k \geq Ds \geq 0 \]

(see above), we get

\[ \lim_{k \to \infty} \overline{D}t_k = 0 = \lim_{k \to \infty} Dt_k \text{ a.e. on } E^n. \]

Now, as \( t_k = s - s_k \) and as the \( s_k \) are differentiable, Lemma 3 yields

\[ \overline{D}t_k = \overline{D}s - s'_k \text{ and } Dt_k = Ds - s'_k \text{ a.e.} \]

Thus

\[ \lim_{k \to \infty} (\overline{D}s - s'_k) = 0 = \lim(Ds - s'_k), \]

i.e.,

\[ \overline{D}s = \lim_{k \to \infty} s'_k = Ds \text{ a.e.} \]

This settles the case \( s_k \not\nearrow s \).

In the case \( s_k \nearrow s \), one only has to set \( t_k = s_k - s \) and proceed as before. (Verify!) \( \square \)

Lemma 6. Given \( A \in \mathcal{M}^*, mA < \infty \), let

\[ s = \int C_A \, dm \]

on \( \mathcal{M}^* \). Then \( s \) is a.e. differentiable, and

\[ s' = C_A \text{ a.e. on } E^n. \]

\((C_A = \text{characteristic function of } A)\)

Proof. First, let \( A \) be open and let \( \bar{p} \in A \).

Then \( A \) contains some \( G_\delta(\bar{p}) \) and hence also all cubes \( I \in \mathcal{K} \) with \( dI < \delta \) and \( \bar{p} \in I \).

Thus for such \( I \in \mathcal{K} \),

\[ sI = \int_I C_A \, dm = \int_I (1) \, dm = mI; \]

i.e.,

\[ \frac{sI}{mI} = 1 = C_A(\bar{p}), \quad \bar{p} \in A. \]

Hence by Definition 1 of Chapter 7, §12,

\[ s'(\bar{p}) = 1 = C_A(\bar{p}) \]

if \( \bar{p} \in A \); i.e., \( s' = C_A \) on \( A \).

\(^1\) Differentiability follows by Theorem 4 of Chapter 7, §12, but we obtain it anyway.
We claim that \( \mathcal{T}_s = s' = 0 \) a.e. on \(-A\).

To prove it, note that
\[
s = \int C_A \, dm
\]
is a finite (why?) \( m \)-continuous measure on \( \mathcal{M}^* \). By Lemma 4, \( s \) is strongly regular. Also, as \( sI \geq 0 \) for any \( I \in \mathcal{K} \), we certainly have
\[
\mathcal{T}_s \geq D_s \geq 0.
\]
(Why?) Now let
\[
B = E^n(\mathcal{T}_s > 0) = \bigcup_{r=1}^{\infty} B_r,
\]
where
\[
B_r = E^n\left( \mathcal{T}_s \geq \frac{1}{r} \right), \quad r = 1, 2, \ldots,
\]
We have to show that \( m(B - A) = 0 \).

Suppose
\[
m(B - A) > 0.
\]
Then by (2), we must have \( m(B_r - A) > 0 \) for at least one \( B_r \); we fix this \( B_r \). Also, by (3),
\[
\mathcal{T}_s \geq \frac{1}{r} \text{ on } B_r - A
\]
(even on all of \( B_r \)). Thus by Lemma 2,
\[
0 < \frac{1}{r} m(B_r - A) \leq s(B_r - A) = \int_{B_r - A} C_A \, dm.
\]

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),
\[
m(E^n(\mathcal{T}_s > 0) - A) = 0;
\]
i.e.,
\[
\mathcal{T}_s = D_s = 0 \text{ a.e. on } -A.
\]
We see that
\[
s' = 0 = C_A \text{ a.e. on } -A,
\]
and
\[
s' = 1 = C_A \text{ on } A,
\]
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 \( k \in \mathbb{N} \) an open set \( G_k \supseteq A \), with
\[
m(G_k - A) < \frac{1}{k} \text{ and } G_k \supseteq G_{k+1}.
\]
Let
\[
s_k = \int C_{G_k} \, dm.
\]
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,
\[
\mathcal{T}_s = D_s = \lim_{k \to \infty} C_{G_k} = C_A \text{ (a.e.)}.
\]
The lemma is proved. \( \square \)

**Theorem 1.** Let \( f : E^n \to E^r(E^r, C^r) \) be \( m \)-integrable, at least on each cube in \( E^n \). Then the set function
\[
s = \int f \, dm
\]
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
\[
\mathcal{M}_0 = \{ X \in \mathcal{M}^* | X \subseteq J \}.
\]
This does not affect \( s' \) on \( J \); for as \( J \) is open, any sequence of cubes
\[
I_k \to \tilde{p} \in J
\]
terminates inside \( J \) anyway.

When so restricted,
\[
s = \int f
\]
is a generalized measure in \( J \); for \( \mathcal{M}_0 \) is a \( \sigma \)-ring (verify!), and \( f \) is integrable on \( J \). Also, \( m \) is strongly regular, and \( s \) is \( m \)-continuous.

‡ Recall that \( \int f \) is always defined by our convention.
First, suppose $f$ is $\mathcal{M}_0$-simple on $J$, say,

$$f = \sum_{i=1}^{q} a_i C_{A_i},$$

say, with $0 < a_i < \infty$, $A_i \in \mathcal{M}_0$, and

$$J = \bigcup_{i=1}^{q} A_i \text{ (disjoint)}.$$

Then

$$s = \int f = \sum_{i=1}^{q} a_i \int C_{A_i}.$$

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

$$s' = \sum_{i=1}^{q} a_i \left( \int C_{A_i} \right)' = \sum_{i=1}^{q} a_i C_{A_i} = f \text{ (a.e.)},$$

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 $f_k \not\to f$ for some simple maps $f_k \geq 0$. Let

$$s_k = \int f_k \text{ on } M_0, \quad k = 1, 2, \ldots.$$ 

Then all $s_k$ and $s = \int f$ are finite measures and $s_k \not\to s$, by Theorem 4 in §6. Also, by what was shown above, each $s_k$ is differentiable a.e. on $J$, with $s_k' = f_k$ (a.e.). Thus as in Lemma 5,

$$D_s = Ds = s' = \lim_{k \to \infty} s_k' = \lim_{k \to \infty} f_k = f \text{ (a.e.) on } J,$$

with $s' = f \not\equiv \pm \infty$ (a.e.), as $f$ is integrable on $J$. Thus all is proved. \hfill \Box

II. So far we have considered Lebesgue ($\mathcal{K}$) differentiation. However, our results easily extend to $\Omega$-differentiation (Definition 2 in Chapter 7, §12).

The proof is even simpler. Thus in Lemma 1, the union in formula (1) is countable (as $\mathcal{K}$ is replaced by the countable set family $\Omega$); hence it is $\mu$-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 $\mu$ on $\mathcal{M}$. Once the lemmas are established (reread the proofs!), we obtain the following.

Theorem 2. Let $S$, $\rho$, $\Omega$, and $\mu: \mathcal{M} \to E^*$ be as in Definition 2 of Chapter 7, §12. Let $f: S \to E^*(E', C')$ be $\mu$-integrable on each $A \in \mathcal{M}$ with $\mu A < \infty$.

Then the set function

$$s = \int f \, d\mu$$

is $\Omega$-differentiable, with $s' = f$, (a.e.) on $S$.

Proof. Recall that $S$ is a countable union of sets $U_n \in \Omega$ with $0 < \mu U_n < \infty$.

As $\mu^*$ is $\mathcal{G}$-regular, each $U_n$ lies in an open set $J_n^* \in \mathcal{M}$ with

$$\mu J_n^* < \mu U_n + \varepsilon_n < \infty.$$

Also, $f$ is $\mu$-measurable (even integrable) on $J_n^*$. Dropping a null set, assume that $f$ is $\mathcal{M}$-measurable on $J = J_n^*$.

From here, proceed exactly as in Theorem 1, replacing $m$ by $\mu$. \hfill \Box

Both theorems combined yield the following result.

Corollary 1. If $s: \mathcal{M}^* \to E^*(E', C')$ is an $m$-continuous and $m$-finite generalized measure in $E^*$, then $s$ is $\mathcal{K}$-differentiable a.e. on $E^*$, and $ds = s' \, dm$ (see Definition 3 in §10) in any $A \in \mathcal{M}^*$ ($m A < \infty$).\footnote{The restriction $mA < \infty$ is redundant if $s$ is finite.}

Similarly for $\Omega$-differentiation.

Proof. Given $A \in \mathcal{M}^*$ ($m A < \infty$), there is an open set $J \supseteq A$ such that

$$m J < m A + \varepsilon < \infty.$$ 

As before, restrict $s$ and $m$ to

$$\mathcal{M}_0 = \{ X \in \mathcal{M}^* \mid X \subseteq J \}.$$ 

Then by assumption, $s$ is finite and $m$-continuous on $\mathcal{M}_0$ (a $\sigma$-ring); so by Theorem 1 in §11,

$$s = \int f \, dm$$

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

$$s = \int f = \int s'$$

on $\mathcal{M}_0$.

This implies $ds = s' \, dm$ in $A$.

For $\Omega$-differentiation, use Theorem 2. \hfill \Box
Corollary 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
\[ \int_A f \, ds = \int_A fs' \, dm. \]
Similarly for $\Omega$-derivatives, with m replaced by $\mu$.

Proof. By Corollary 1, $ds = s' \, dm$ in A. Thus Theorem 6 of §10 yields the result. $\square$

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 $s_\alpha$ is surely finite on $s_\alpha$-measurable subsets of A; so Corollaries 1 and 2 show that
\[ \int_A f \, ds_\alpha = \int_A fs'_\alpha \, dm = \int_A f \alpha' \, dm, \]
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 $\alpha$ is absolutely continuous in the stronger sense (Problem 2 in Chapter 4, §8).

Indeed, assuming the latter, fix $\varepsilon > 0$ and choose $\delta$ as in Definition 3 of Chapter 7, §11. Then if $mX < \delta$, we have $X \subseteq \bigcup I_k$ (disjoint)
for some intervals $I_k = (a_k, b_k)$, with
\[ \delta > \sum mI_k = \sum (b_k - a_k). \]
Hence
\[ |sX| \leq \sum |sI_k| < \varepsilon. \]
(Why?) Similarly for the converse.$^5$

Problems on Differentiation and Related Topics

1. Fill in all proof details in this section. Verify footnote 4 and Note 2.

2. Given a measure $s : \mathcal{M}' \rightarrow E^\ast$ ($\mathcal{M}' \supseteq \mathcal{K}$), prove that
   (i) $s$ is topological;
   (ii) its Borel restriction $\sigma$ is strongly regular; and
   (iii) $Ds$, $\mathcal{L}s$, 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 $\mathcal{K}$-differentiation proved here.

$^4$ The restriction $mA < \infty$ is redundant if $s$ is finite.

$^5$ Note that $s[a] = 0$ if $s$ is m-continuous.

3. What analogues to 2(i)–(iii) apply to $\Omega$-differentiation in $E^n$? In $(S, \rho)$?

4. (i) Show that any m-singular measure $s$ in $E^n$, finite on $\mathcal{K}$, has a zero derivative (a.e.).
   (ii) For $\Omega$-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 $\sigma$). Suppose $mE^n(\mathcal{K}s > 0) > a > 0$
and find a contradiction to Lemma 2.]

5. Give another proof for Theorem 4 in Chapter 7, §12.
   [Hint: Fix an open cube $J \in \mathcal{K}$. By Problem 2(ii), restrict s and m to $\mathcal{M}_0 = \{X \in \mathcal{B} \mid X \subseteq J\}$ 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 $\Omega$-differentiation, assume $s$ regular; argue as in Corollary 1, using Corollary 2 of §11.]

6. Prove that if
\[ F(x) = L \int_a^x f \, dm \quad (a \leq x \leq b), \]
with $f : E^1 \rightarrow E^\ast$ ($E^n$, $C^n$) 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 $f \geq 0$, $F'$ a.e. on $A$. Let $s = \int f \, dm$ on $\mathcal{M}'$. Let $t = ms$ be the $F$-induced LS measure. Show that $s = t$ on intervals in $A$; so $s' = t' = F'$ a.e. on $A$ (Problem 9 in Chapter 7, §11). Use Theorem 1.]

$^6$ Here $L \int_a^x f \, dm = \int_{[a,x]} f \, dm$; $m =$Lebesgue measure.
Chapter 9
Calculus Using Lebesgue Theory

§1. L-Integrals and Antiderivatives

I. Lebesgue theory makes it possible to strengthen many calculus theorems. We shall start with functions on $E^1$, $f: E^1 \to E$. (A reader who has omitted the “starred” part of Chapter 8, §7, will have to set $E = E^* (E^n, C^m)$ throughout.)

By L-integrals of such functions, we mean integrals with respect to Lebesgue measure $m$ in $E^1$. Notation:

$$L\int_a^b f = L\int_a^b f(x)\,dx = L\int_{(a,b]} f$$

and

$$L\int_b^a f = -L\int_a^b f.$$  

For Riemann integrals, we replace “$L$” by “$R$.” We compare such integrals with antiderivatives (Chapter 5, §5), denoted

$$\int_a^b f,$$

without the “$L$” or “$R$.” Note that

$$L\int_{[a,b]} f = L\int_{(a,b)} f,$$

e tc., since $m\{a\} = m\{b\} = 0$ here.

**Theorem 1.** Let $f: E^1 \to E$ be L-integrable on $A = [a,b]$. Set

$$H(x) = L\int_a^x f, \quad x \in A.$$
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 $V_H[A] < \infty$.\(^3\)

**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

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

for $x \in (c, p)$. Then

$$\left(\forall x \in (c, p) \right) \left| L \int_{c}^{p} (f - q) \right| \leq L \int_{c}^{p} |f - q| \leq L \int_{c}^{p} (\varepsilon) = \varepsilon \cdot m|, p| = \varepsilon (p - x).$$

But

$$L \int_{c}^{p} (f - q) = L \int_{c}^{p} f - L \int_{c}^{p} q,$$

and

$$L \int_{c}^{p} q = q (p - x).$$

Thus

$$|H(p) - H(x) - q(p - x)| \leq \varepsilon (p - x);$$

i.e.,

$$\left| \frac{H(p) - H(x)}{p - x} - q \right| \leq \varepsilon \quad (c < x < p).$$

Hence

$$f(p) = q = \lim_{x \to p^-} \frac{\Delta H}{\Delta x} = H'_V (p).$$

If $f$ is right continuous at $p \in [a, b)$, a similar formula results for $H'_V (p)$. This proves clause (i).

---

\(^1\) This is true even in the stronger sense, as in Problem 2 of Chapter 5, §8, or in §2 (next to this).

\(^2\) Recall that $V_H[A]$ is the total variation of $H$ on $A$ (Chapter 5, §§7-8).

\(^3\) Part (ii) is true even if $f$ is not L-integrable, only $L \int_{a}^{b} |f| < \infty$ is needed.

\(|1. L\text{-Integrals and Antiderivatives} 359|

(ii) Let $\varepsilon > 0$ be given. Then Theorem 6 in Chapter 8, §6, yields a $\delta > 0$ such that

$$\left| L \int_{A}^{x} f \right| \leq L \int_{A}^{x} |f| \leq \varepsilon$$

whenever

$$mX < \delta$$

and $A \supseteq X$, $X \in M$.

Here we may set

$$X = \bigcup_{i=1}^{r} A_i \quad (\text{disjoint})$$

for some intervals

$$A_i = (a_i, b_i) \subseteq A$$

so that

$$mX = \sum_{i} mA_i = \sum_{i} (b_i - a_i) < \delta.$$
We can now prove a generalized version of the so-called fundamental theorem of calculus, widely used for computing integrals via antiderivatives.

**Theorem 2.** If \( f : E^1 \to E \) 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

\[
L \int_a^b f = F(x) - F(a) \quad \text{for all } x \in A.
\]

**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 \(-A: \)

\[
F(x) = \begin{cases} 
F(a) & \text{if } x < a, \\
F(b) & \text{if } x > b,
\end{cases}
\]

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 \( F = f \) on \( E^1. \)

Also, for \( n = 1, 2, \ldots \) and \( t \in E^1, \) set

\[
f_n(t) = n \left( F \left( t + \frac{1}{n} \right) - F(t) \right) = \frac{F(t + 1/n) - F(t)}{1/n}.
\]

Then

\[
f_n \to F' = f \quad \text{on } -Q;
\]

i.e., \( f_n \to f \) (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 \( F \) is 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

\[
|f| \leq K < \infty \quad \text{on } A - P;
\]

as \( mP = 0, \)

\[
\int_A |f| \leq \int_A (K) = K \cdot mA < \infty,
\]

proving integrability. Now, as

\[
F = \int f \quad \text{on any interval } [t, t + 1/n],
\]

we also have

\[
\begin{align*}
\forall t \in E^1 \quad |F(t + 1/n) - F(t)| & \leq \sup_{t \in Q} |F'(t)| \frac{1}{n} \leq K/n. \\
Hence \quad |f_n(t)| & = n |F(t + 1/n) - F(t)| \leq K;
\end{align*}
\]

i.e., \( |f_n| \leq K \) for all \( n. \)

Thus \( f \) and \( f_n \) satisfy Theorem 5 of Chapter 8, §6, with \( g = K. \) By Note 1 there,

\[
\lim_{n \to \infty} \int_a^b f_n = L \int_a^b f.
\]

In the next lemma, we show that also

\[
\lim_{n \to \infty} L \int_a^b f_n = F(x) - F(a),
\]

which will complete the proof. \( \square \)

**Lemma 1.** Given a finite continuous \( F : E^1 \to E \) and given \( f_n \) as in (3), we have

\[
\lim_{n \to \infty} L \int_a^b f_n = F(x) - F(a) \quad \text{for all } x \in E^1.
\]

**Proof.** As before, \( F \) and \( f_n \) are bounded, continuous, and \( L \)-integrable on any \( [a, x] \) or \( [x, a]. \) Fixing \( a, \) let

\[
H(x) = L \int_a^x F, \quad x \in E^1.
\]

By Theorem 1 and Note 2, \( H = \int F \) also in the sense of Chapter 5, §5, with \( F = H' \) (derivative of \( H \)) on \( E^1. \)

Hence by Definition 2 the same section,

\[
\int_a^x F = H(x) - H(a) = H(x) - 0 = L \int_a^x F;
\]

i.e.,

\[
L \int_a^x F = \int_a^x F,
\]

and so

\[
L \int_a^x f_n(t) \, dt = n \int_a^x F \left( t + \frac{1}{n} \right) \, dt - n \int_a^x F(t) \, dt
\]

\[
= n \int_{a+1/n}^{x+1/n} F(t) \, dt - n \int_a^x F(t) \, dt.
\]

---

\[^5\text{See Definition 1 from Chapter 5, §5.}\]
We computed
\[ \int F(t + 1/n) \, dt \]
by Theorem 2 in Chapter 5, §5, with \( g(t) = t + 1/n \). Thus by additivity,
\[ L \int_a^b f_n = n \int_{a+1/n}^{x+1/n} F - n \int_a^x F = n \int_a^x F - n \int_{a+1/n}^{x+1/n} F. \]
But
\[ n \int_a^x F = \frac{H(x + \frac{1}{n}) - H(x)}{n} \to H'(x) = F(x). \]
Similarly,
\[ \lim_{n \to \infty} n \int_{a+1/n}^{x+1/n} F = F(a). \]
This combined with (5) proves (4), and hence Theorem 2, too. \( \square \)

We also have the following corollary.

**Corollary 1.** If \( f: E^1 \to E^n (E^m, C^n) \) is R-integrable on \( A = [a, b] \), then
\[ (\forall x \in A) \quad R \int_a^x f = L \int_a^b f = F(x) - F(a), \]
provided \( F \) is primitive to \( f \) on \( A \).

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 \( F' = f \) 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
\[ F \neq \int f. \]
For examples, see Problems 2–5.

**Corollary 2.** If \( f \) is relatively continuous and finite on \( A = [a, b] \) and has a bounded derivative on \( A - Q \) (\( Q \) countable), then \( f' \) is L-integrable on \( A \) and
\[ L \int_a^x f' = f(x) - f(a) \quad \text{for} \ x \in A. \]
This is simply Theorem 2 with \( F, f, P \) replaced by \( f, f', Q \), respectively.

---

\( ^6 \) We assumed that \( E = E^* (E^n, C^n) \) since R-integrals were defined for that case only.

---

**Corollary 3.** If in Theorem 2 the primitive
\[ F = \int f \]
is exact on some \( B \subseteq A \), then
\[ f(x) = \frac{d}{dx} L \int_a^x f, \quad x \in B. \]
(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,
\[ L \int_a^x f = F(x) - F(a) = \int_a^x f. \]
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 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 \( f: E^1 \to E^1 \) (real or not) has a primitive on \( g[A - Q] \), exact on \( g[A - Q] \), and that \( f \) is bounded on \( g[A - Q] \).

Then \( f \) is L-integrable on \( g[A] \), the function
\[ (f \circ g) \dot{g}' \]
is L-integrable on \( A \), and
\[ L \int_a^b f(g(x)) \dot{g}'(x) \, dx = L \int_p^q f(y) \, dy, \]
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 2 (Bonnet).** Suppose \( f: E^1 \to E^1 \) is \( \geq 0 \) 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
\[ L \int_a^b fg = f(a) \cdot L \int_a^c g \quad \text{for some} \ c \in A. \]

**Proof.** The L-integrability of \( fg \) follows by Theorem 3 in Chapter 8, §8, 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 \( C \)-partition 
\[ P_n = \{ A_{ni} \} \quad (i = 1, 2, \ldots, q_n) \]
of \( A \) so that
\[ (\forall n) \quad \frac{1}{n} > \frac{1}{n} (f, P_n) - \frac{1}{n} (f, P_n) = \sum_{i=1}^{q_n} w_{ni} m A_{ni}, \]
where we have set
\[ w_{ni} = \sup f[A_{ni}] - \inf f[A_{ni}]. \]

Consider any such \( P = \{ A_i \}, i = 1, \ldots, q \) (we drop the “\( n \)” for brevity). If
\[ A_i = [a_{i-1}, a_i], \]
then since \( f \downarrow \),
\[ w_i = f(a_{i-1}) - f(a_i) \geq |f(x) - f(a_{i-1})|, \quad x \in A_i. \]
Under Lebesgue measure (Problem 8 of Chapter 8, §9), we may set
\[ A_i = [a_{i-1}, a_i] \quad (\forall i) \]
and still get
\[ L \int_A f g = \sum_{i=1}^{q} \int_{A_i} f(a_{i-1}) L \int_{A_i} g(x) \, dx \]
\[ + \sum_{i=1}^{q} L \int_{A_i} [f(x) - f(a_{i-1})] g(x) \, dx. \]
(Verify!) Here \( a_0 = a \) and \( a_q = b \).

Now, set \( G(x) = L \int_a^x g \)
and rewrite the first sum (call it \( r \) or \( r_n \)) as
\[ r = \sum_{i=1}^{q} f(a_{i-1}) [G(a_i) - G(a_{i-1})] \]
\[ = \sum_{i=1}^{q} G(a_i) [f(a_{i-1}) - f(a_i)] + G(b) f(a_{q-1}), \]
or
\[ r = \sum_{i=1}^{q} G(a_i) w_i + G(b) f(a_{q-1}), \]

because \( f(a_{i-1}) - f(a_i) = w_i \) 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 \( A \).

§1. L-Integrals and Antiderivatives

As \( f_\downarrow \) and \( f \geq 0 \) on \( A \), we have
\[ w_i \geq 0 \text{ and } f(a_{q-1}) \geq 0. \]
Thus, replacing \( G(b) \) and \( G(a_i) \) by \( K \) (or \( k \)) in (13) and noting that
\[ \sum_{i=1}^{q} w_i = f(a) - f(a_{q-1}), \]
we obtain
\[ k f(a) \leq r \leq K f(a); \]
more fully, with \( k = \min G[A] \) and \( K = \max G[A] \),
\[ (\forall n) \quad k f(a) \leq r_n \leq K f(a). \]

Next, let \( s \) (or rather \( s_n \)) be the second sum in (12). Noting that
\[ w_i \geq |f(x) - f(a_{i-1})|, \]
suppose first that \( |g| \leq B \) (bounded) on \( A \).

Then for all \( n \),
\[ |s_n| \leq \sum_{i=1}^{q} L \int_{A_{ni}} (w_{ni} B) = B \sum_{i=1}^{q} w_{ni} m A_{ni} < \frac{B}{n} \to 0 \quad (\text{by (11)}). \]
But by (12),
\[ L \int_A f g = r_n + s_n \quad (\forall n). \]
As \( s_n \to 0 \),
\[ L \int_A f g = \lim_{n \to \infty} r_n, \]
and so by (14),
\[ k f(a) \leq L \int_A f g \leq K f(a). \]
By continuity, \( f(a) G(x) \) takes on the intermediate value \( L \int_A f g \) at some \( c \in A \); so
\[ L \int_A f g = f(a) G(c) = f(a) L \int_a^c f, \]

since
\[ G(x) = L \int_a^x f. \]
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!) \( \square \)
Corollary 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

\[
L \int_a^b fg = f(a) L \int_a^c g + f(b) L \int_c^b g \quad \text{for some } c \in A.
\]

Proof. If, say, \( f \uparrow \) on \( A \), set

\[
h(x) = f(x) - f(b).
\]

Then \( h \geq 0 \) and \( h \uparrow \) on \( A \); so by Lemma 2,

\[
\int_a^b gh = h(a) L \int_a^c g \quad \text{for some } c \in A.
\]

As

\[
h(a) = f(a) - f(b),
\]

this easily implies (15).

If \( f \downarrow \), apply this result to \( -f \) to obtain (15) again. \( \square \)

Note 3. We may restate (15) as

\[
\exists c \in A \quad L \int_a^b fg = p \int_a^c g + q \int_c^b g,
\]

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

\[
f(a) = p \quad \text{and } f(b) = q.
\]

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 \( F : E^1 \to E \) is called an \( L \)-primitive or an indefinite \( L \)-integral of \( f : E^1 \to E \), on \( A = [a, b] \) iff \( f \) is \( L \)-integrable on \( A \) and

\[
F(x) = c + L \int_a^x f
\]

for all \( x \in A \) and some fixed finite \( c \in E \).

§1. \( L \)-Integrals and Antiderivatives

Notation:

\[
F = L \int f \quad \text{(not } F = \int f\text{)}
\]

or

\[
F(x) = L \int f(x) \, dx \quad \text{on } A.
\]

By (16), all \( L \)-primitives of \( f \) on \( A \) differ by finite constants only.

If \( E = E^*(E^n, C^n) \), 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 3.** Let

\[
F = L \int f \quad \text{on } A = [a, b]
\]

for some \( f : E^1 \to E^*(E^n, C^n) \).

Then \( F \) is differentiable, with

\[
F' = f \quad \text{a.e. on } A.
\]

In classical notation,

\[
f(x) = \frac{d}{dx} \int_a^x f(t) \, dt \quad \text{for almost all } x \in A.
\]

A proof was sketched in Problem 6 of Chapter 8, §12. (It is brief but requires more “starred” material than used in §2.)

**Theorem 4.** Let \( F : E^1 \to E^n (C^n) \) be differentiable on \( A = [a, b] \) (at \( a \) and \( b \) differentiability may be one sided). Let \( F' = f \) be \( L \)-integrable on \( A \).

Then

\[
L \int_a^x f = F(x) - F(a) \quad \text{for all } x \in A.
\]

Problems on \( L \)-Integrals and Antiderivatives

1. Fill in proof details in Theorems 1 and 2, Lemma 1, and Corollaries 1–3.

2. Verify Note 2.

2'. Let \( F \) be Cantor’s function (Problem 6 in Chapter 4, §5). Let

\[
G = \bigcup_{k,i} G_{ki}
\]

\((G_{ki} \text{ as in that problem})\). So \([0, 1] - G = P \) (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,

$$R \int_0^1 F' = L \int_0^1 F' = L \int_G F' = 0$$

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.

3. Let

$$F = \begin{cases} 0 & \text{on } [0, \frac{1}{2}), \\
1 & \text{on } [\frac{1}{2}, 1]. \end{cases}$$

Show that

$$R \int_0^1 F' = 0$$

exists, yet

$$F(1) - F(0) = 1 - 0 = 1.$$ 

What is wrong?

[Hint: A genuine primitive of $F'$ (call it $\phi$) has to be relatively continuous on $[0, 1]$; find $\phi$ and show that $\phi(1) - \phi(0) = 0$.]

4. What is wrong with the following computations?

(i) $L \int_{-1}^1 \frac{dx}{x^2} = -\frac{1}{x} \bigg|_{-1}^1 = -1$.

(ii) $L \int_{-1}^1 \frac{dx}{x} = \ln |x| \bigg|_{-1}^1 = 0$. 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)?

5. Let

$$F(x) = x^2 \cos \frac{\pi}{x^2}, \quad F(0) = 1.$$ 

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 $A$.

§1. L-Integrals and Antiderivatives

(iii) Let

$$a_n = \sqrt{\frac{2}{4n+1}} \quad \text{and} \quad b_n = \frac{1}{\sqrt{2n}} \quad \text{for } n = 1, 2, \ldots.$$ 

Show that

$$A \supseteq \bigcup_{n=1}^{\infty} [a_n, b_n] \quad \text{(disjoint)}$$

and

$$L \int_{a_n}^{b_n} f = \frac{1}{2n}$$

so

$$L \int_a^b f \geq L \int \bigcup_{n=1}^{\infty} [a_n, b_n] f \geq \sum_{n=1}^{\infty} \frac{1}{2n} = \infty,$$

and $f = F'$ is not L-integrable on $A$.

What is wrong? Is there a contradiction to Theorem 2?

6. Consider both

(a) $f(x) = \frac{\sin x}{x}, \quad f(0) = 1$, and

(b) $f(x) = \frac{1 - e^{-x}}{x}, \quad f(0) = 1$.

In each case, show that $f$ is continuous on $A = [0, 1]$ and

$$R \int_A f \leq 1$$

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

$$\int f$$

by Theorem 3 of Chapter 5, §9.

Find

$$R \int_A f$$

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.”]
7. Take \( A, G_n = (a_n, b_n), \) and \( P \) (\( mP > 0 \)) as in Problem 17(ii) of Chapter 7, §8.

Define \( F = 0 \) on \( P \) and

\[
F(x) = (x - a_n)^2 (b - x)^2 \sin \frac{1}{(b - a_n)(x - a_n)(x - b_n)} \quad \text{if } x \notin P.
\]

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 \( x \to p \) 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

\[
\frac{\Delta F}{\Delta x} \leq (x - a_n) (b - a)^2 \leq |\Delta x| (b - a)^2;
\]

so \( F'_n(p) = 0 \). (What if \( x \to p^+ \) over \( P' \)?) Similarly, show that \( F'_n = 0 \) on \( P \).

Prove however that \( F'(x) \) oscillates from 1 to \(-1\) as \( x \to a_n^+ \) or \( x \to b_n^- \), hence also as \( x \to p \in P \) (why?): so \( F' \) is discontinuous on all of \( P \), with \( mP > 0 \). Now use Theorem 3 in Chapter 8, §9.

\[\Rightarrow 8.\] If

\[ Q \subseteq A = [a, b] \]

and \( mQ = 0 \), find a continuous map \( g : A \to E^1, g \geq 0, g \uparrow \), with

\[ g' = +\infty \quad \text{on } Q. \]

[Hints: By Theorem 2 of Chapter 7, §8, fix \( \forall n \) an open \( G_n \supseteq Q \), with \( mG_n < 2^{-n} \). Set

\[
g_n(x) = m(G_n \cap [a, x])
\]

and

\[ g = \sum_{n=1}^{\infty} g_n \]

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

\[ g_n(x) = g_n(p) + (x - p), \]

so

\[
\frac{\Delta g_n}{\Delta x} = 1
\]

and

\[
\frac{\Delta g}{\Delta x} = \sum_{n=1}^{\infty} \frac{\Delta g_n}{\Delta x} \to \infty.\]

§1. L-Integrals and Antiderivatives


(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

\[ g[A] - g(Q) \subseteq g[A - Q]; \]

\( f \) satisfies Theorem 2 on \( g[A] \), with \( P = g[Q] \), while \( (f \circ g)g' \) satisfies it on \( A \).

\[ \Rightarrow 10. \] Show that if \( h : E^1 \to E^* \) is \( L \)-integrable on \( A = [a, b] \), and

\[
(\forall x \in A) \quad L \int_a^x h = 0,
\]

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,

\[ H \subseteq \bigcup_{k \in \mathbb{N}} B_k \quad \text{(disjoint)} \]

for some intervals \( B_k \subseteq A \), with

\[ \sum_{k} mB_k < mH + \varepsilon = mH + mK = mA. \]

(Why?) Set \( B = \bigcup_{k} B_k \); so

\[
\int_B h = \sum_{k} \int_{B_k} h = 0
\]

(for \( L \int h = 0 \) on intervals \( B_k \)). Thus

\[
\int_{A - B} h = \int_A h - \int_B h = 0.
\]

But \( B \supseteq H \); so

\[ A - B \subseteq A - H = K, \]

where \( h > 0 \), even though \( m(A - B) > 0 \). (Why?)

Hence find a contradiction to Theorem 1(b) of Chapter 8, §5. Similarly, disprove that \( mA(h < 0) = \varepsilon > 0 \).

\[ \Rightarrow 11. \] Let \( F^\uparrow \) on \( A = [a, b] \), \( |F| < \infty \), with derived function \( F' = f \). Taking Theorem 3 from Chapter 7, §10, for granted, prove that

\[
L \int_a^x f \leq F(x) - F(a), \quad x \in A.
\]

[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 \( f_n \to f \) (a.e.) on \( A \). Argue as in Lemma 1 using Fatou's lemma (Chapter 8, §6, Lemma 2).]
12. (“Truncation.”) Prove that if \( g: S \to E \) is \( M \)-integrable on \( A \in M \) in a measure space \((S,M,m)\), then for any \( \varepsilon > 0 \), there is a bounded, \( M \)-measurable and integrable on \( A \) map \( g_0: S \to E \) such that

\[
\int_A |g - g_0| \, dm < \varepsilon.
\]

[Outline: Redefine \( g = 0 \) on a null set, to make \( g \) \( M \)-measurable on \( A \). Then for \( n = 1, 2, \ldots \) set

\[
g_n = \begin{cases} g & \text{on } A(|g| < n), \\ 0 & \text{elsewhere.}
\end{cases}
\]

(The function \( g_n \) is called the \( n \)th truncate of \( g \).
Each \( g_n \) is bounded and \( M \)-measurable on \( A \) (why?), and

\[
\int_A |g| \, dm < \infty
\]

by integrability. Also, \( |g_n| \leq |g| \) and \( g_n \to g \) (pointwise) on \( A \). (Why?)]

Now use Theorem 5 from Chapter 8, §6, to show that one of the \( g_n \)s may serve as the desired \( g_0 \).

13. Fill in all proof details in Lemma 2. Prove it for unbounded \( g \).

[Hints: By Problem 12, fix a bounded \( g_0 \), \(|g_0| \leq B \), with

\[
E \int_A |g - g_0| < \frac{1}{2} \frac{\varepsilon}{f(a) - f(b)}
\]

Verify that

\[
|s_n| \leq \sum_{k=1}^n \int_{A_k} w_{n_k} |g| \leq \sum_{k=1}^n \int_{A_k} w_{n_k} |g| + \sum_{k=1}^n \int_{A_k} w_{n_k} |g - g_0|
\]

\[
\leq B \sum_{k=1}^n w_{n_k} |M_{n_k}| + \sum_{k=1}^n \int_{A_k} [f(a) - f(b)] |g - g_0|
\]

\[
< \frac{n}{2} \sum_{k=1}^n \left| \int_{A_k} [f(a) - f(b)] |g - g_0| \right| < \frac{1}{n} \left( \frac{1}{2} \varepsilon + \frac{1}{2} \varepsilon \right).
\]

For all \( n > 2/\varepsilon \), we get \( |s_n| < \frac{1}{2} \varepsilon + \frac{1}{2} \varepsilon = \varepsilon \). Hence \( s_n \to 0 \). Now finish as in the text.]

14. Show that Theorem 4 fails if \( F \) is not differentiable at some \( p \in A \).

[Hint: See Problems 2 and 3.]

§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 1.

Given \( F: E^1 \to E, p \in E^1 \), and \( q \in E \), we write

\[ q \sim DF(p) \]

and call \( q \) an \( F \)-derivate at \( p \) iff

\[ q = \lim_{k \to \infty} \frac{F(x_k) - F(p)}{x_k - p} \]

for at least one sequence \( x_k \to p \) (\( x_k \neq p \)).

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

\[ x_k = p + \frac{1}{k} \to p \]

let

\[ y_k = \frac{F(x_k) - F(p)}{x_k - p}, \quad k = 1, 2, \ldots \]

By the compactness of \( E^* \) (Chapter 4, §6, example (d)), \( \{y_k\} \) must have a subsequence \( \{y_{k_n}\} \) with a limit \( q \in E^* \) (e.g., take \( q = \lim y_{k_n} \)), and so \( q \sim DF(p) \).

We also obtain the following lemma.

Lemma 1. If \( F: E^1 \to E^* \) has no negative derivatives on \( A - Q \), where \( A = [a, b] \) and \( mQ = 0 \), and if no derivative of \( F \) on \( A \) equals \(-\infty \), then \( F \) is \( M \)-integrable on \( A \).

Proof. First, suppose \( F \) has no negative derivatives on \( A \) at all. Fix \( \varepsilon > 0 \) and set

\[ G(x) = F(x) + \varepsilon x. \]

Seeking a contradiction, suppose \( a \leq p < q \leq b \), yet \( G(q) < G(p) \). Then if

\[ r = \frac{1}{2}(p + q), \]

one of the intervals \([p, r]\) and \([r, q]\) (call it \([p_1, q_1]\)) satisfies \( G(q_1) < G(p_1) \).

Let

\[ r_1 = \frac{1}{2}(p_1 + q_1). \]

Again, one of \([p_1, r_1]\) and \([r_1, q_1]\) (call it \([p_2, q_2]\)) satisfies \( G(q_2) < G(p_2) \). Let

\[ r_2 = \frac{1}{2}(p_2 + q_2), \]

and so on.

\[ ^1 \text{"DF(p)" stands for "an F-derivate at p."} \]
Thus obtain contracting intervals \( [p_n, q_n] \), with
\[
G(q_n) < G(p_n), \quad n = 1, 2, \ldots.
\]
Now, by Theorem 5 of Chapter 4, §6, let
\[
p_o \in \bigcap_{n=1}^{\infty} [p_n, q_n].
\]
Then set \( x_n = q_n \) if \( G(q_n) < G(p_o) \), and \( x_n = p_n \) otherwise. Then
\[
\frac{G(x_n) - G(p_o)}{x_n - p_o} < 0
\]
and \( x_n \to p_o \). By the compactness of \( E^* \), fix a subsequence
\[
\frac{G(x_{n_k}) - G(p_o)}{x_{n_k} - p_o} \to c \in E^*,
\]
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 derivatives of \( G \) are \( > 0 \). (Why?)

This contradiction shows that \( a \leq p < q \leq b \) implies \( G(p) \leq G(q) \), i.e.,
\[
F(p) + \varepsilon p \leq F(q) + \varepsilon q.
\]
Making \( \varepsilon \to 0 \), 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 \( g' = \infty \) on \( Q \). Given \( \varepsilon > 0 \), set
\[
G = F + \varepsilon g.
\]
As \( g^\uparrow \), we have
\[
(\forall x, p \in A) \quad \frac{G(x) - G(p)}{x - p} \geq \frac{F(x) - F(p)}{x - p}.
\]
Hence \( DG(p) \geq 0 \) if \( p \notin Q \).

If, however, \( p \in Q \), then \( g'(p) = \infty \) implies \( DG(p) \geq 0 \). (Why?) Thus all \( DG(p) \geq 0 \); so by what was proved above, \( G^\uparrow \) on \( A \). It follows, as before, that \( F^\uparrow \) on \( A \), also. The lemma is proved. \( \square \)

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:
\[
|f| \leq K \quad \text{on } A.
\]

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
\[
L \int_a^x f \leq L \int_y^y f,
\]
i.e., \( F(x) \leq F(y) \); so \( F^\uparrow \) and \( F^\uparrow \) \( \geq 0 \) 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
\[
f_n(t) = \frac{F(t + \frac{1}{n}) - F(t)}{\frac{1}{n}, \text{ so by what was proved above,}} \quad \Rightarrow \quad G(t) \text{ a.e.}
\]

Since all \( f_n \) are \( m \)-measurable on \( A \) (why?), so is \( F' \). Moreover, as \( |f| \leq K \), we obtain (as in Lemma 1 of §1)
\[
|f_n(x)| = \int_a^x |f(t + \frac{1}{n}) - f(t)| dt \leq \int_a^x \frac{K}{n} = K.
\]

Thus by Theorem 5 from Chapter 8, §6 (with \( g = K \)),
\[
L \int_a^x F' = \lim_{n \to \infty} L \int_a^x f_n = L \int_a^x f
\]
(Lemma 1 of §1). Hence
\[
L \int_a^x (F' - f) = 0, \quad x \in A,
\]
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 \( F^\uparrow \) \( \geq 0 \) on \( A \).

If so, we use “truncation”: For \( n = 1, 2, \ldots \), set
\[
g_n = \left\{ \begin{array}{ll}
f & \text{on } A(f \leq n), \\
0 & \text{elsewhere.}
\end{array} \right.
\]
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
\[
0 \leq g_n \leq f
\]
on \( A \). By the first part of the proof, then,
\[
\frac{d}{dx} L \int_a^x g_n = g_n \quad \text{a.e. on } A, \quad n = 1, 2, \ldots.
\]
Also, set (\forall n)\n\[ F_n(x) = L \int_a^x (f - g_n) \geq 0; \]
so \( F_n \) is monotone (\|) on \( A \). (Why?)

Thus by Theorem 3 in Chapter 7, §10, each \( F_n \) has a derivative at almost every \( x \in A \),
\[ F'_n(x) = \frac{d}{dx} \left( L \int_a^x f - L \int_a^x g_n \right) = F'(x) - g_n(x) \geq 0 \quad \text{a.e. on} \ A. \]

Making \( n \to \infty \) and recalling that \( g_n \to f \) on \( A \), we obtain
\[ F'(x) - f(x) \geq 0 \quad \text{a.e. on} \ A. \]

Thus
\[ L \int_a^x (F' - f) \geq 0. \]

But as \( F' \) (see above), Problem 11 of §1 yields
\[ L \int_a^x F' \leq F(x) - F(a) = L \int_a^x f; \]
so
\[ L \int_a^x (F' - f) = L \int_a^x F' - L \int_a^x f \leq 0. \]

Combining, we get
\[ (\forall x \in A) \quad L \int_a^x (F' - f) = 0; \]
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 \( f \).\(^2\)

Let (\forall n)
\[ g_n = \begin{cases} f & \text{on } A(f \leq n), \\ 0 & \text{on } A(f > n); \end{cases} \]
so \( g_n \to f \) (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,
\[ \lim_{n \to \infty} L \int_a^x g_n = L \int_a^x f, \quad x \in A. \]

Now, set
\[ F_n(x) = F(x) - L \int_a^x g_n. \]

\(^2\) Not \( f \geq 0 \), though, since Corollary 1 in Chapter 8, §6, does not apply to differentiation.

Then by Theorem 3 of §1 (already proved),
\[ F'_n(x) = F'(x) - \frac{d}{dx} L \int_a^x g_n = f(x) - g_n(x) \geq 0 \quad \text{a.e. on} \ A \]
(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
\[ \frac{1}{x - p} L \int_a^x g_n \leq n, \]
even if \( x < p \). (Why?) Hence
\[ \frac{\Delta F_n}{\Delta x} \geq \frac{\Delta F}{\Delta x} - n, \]
as
\[ F_n(x) = F(x) - L \int_a^x g_n, \]

Thus none of the \( F_n \)-derivates on \( A \) can be \(-\infty\).

By Lemma 1, then, \( F_n \) is monotone (\|) on \( A \); so \( F_n(x) \geq F_n(a) \), i.e.,
\[ F(x) - L \int_a^x g_n \geq F(a) - L \int_a^a g_n = F(a), \]
or
\[ F(x) - F(a) \geq L \int_a^x g_n, \quad x \in A, n = 1, 2, \ldots. \]

Hence by (1),
\[ F(x) - F(a) \geq L \int_a^x f, \quad x \in A. \]

For the reverse inequality, apply the same formula to \(-f\). Thus we obtain the desired result:
\[ (2) \quad F(x) = F(a) + L \int_a^x f \quad \text{for } x \in A. \quad \square \]

**Note 1.** Formula (2) is equivalent to \( F = L \int f \) on \( A \) (see the last part of §1). For if (2) holds, then
\[ F(x) = c + L \int_a^x f, \]
with \( c = F(a) \); so \( F = L \int f \) by definition.

Conversely, if
\[ F(x) = c + L \int_a^x f, \]
set \( x = a \) to find \( c = F(a) \).
II. Absolute continuity redefined.

Definition 2.
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
\[
\sum_{i=1}^{k} (b_i - a_i) < \delta \quad \text{implies} \quad \sum_{i=1}^{k} |f(b_i) - f(a_i)| < \varepsilon
\]
for any disjoint intervals \((a_i, b_i)\), 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 1. If \( f, g, h: E^1 \rightarrow E^a(C) \) are absolutely continuous on \( A = [a, b] \),
so are
\[
f \pm g, \, hf, \, \text{and} \, |f|.
\]

So also is \( f/h \) if
\[
(\exists \varepsilon > 0) \quad |h| \geq \varepsilon \quad \text{on} \, A.
\]

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
\[
f \vee g, \, f \wedge g, \, f^+ \quad \text{and} \quad f^-
\]
are absolutely continuous along with \( f \) and \( g \).

Corollary 1. A function \( F: E^1 \rightarrow E^a(C^a) \) is absolutely continuous on \( A = [a, b] \) if 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, \( F_r \) and \( F_i \), are.

Corollary 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 2. If \( F: E^1 \rightarrow E^a(C^a) \) 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
\[
F = g - h
\]
for some nondecreasing \( g \) and \( h \) (Theorem 3 in Chapter 5, §7).

§2. More on \( L \)-integrals and absolute continuity

Now, by Theorem 3 from Chapter 7, §10, \( g \) and \( h \) are a.e. differentiable on \( A \). Hence so is
\[
g - h = F.
\]

Moreover, \( g' \geq 0 \) and \( h' \geq 0 \) since \( g \uparrow \) and \( h \uparrow \).

Thus for the \( L \)-integrability of \( F' \), proceed as in Problem 11 in §1, i.e., show that \( F' \) is measurable on \( A \) and that
\[
L \int_a^b F' = L \int_a^b g' - L \int_a^b h'
\]
is finite. This yields the result. \( \square \)

Theorem 2 (Lebesgue). If \( F: E^1 \rightarrow E^a(C^a) \) is absolutely continuous on \( A = [a, b] \), then the following are true:

(i*) \( F \) is a.e. differentiable, and \( F' \) is \( L \)-integrable, on \( A \).
(ii*) If, in addition, \( F' = 0 \) 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 \( F' = 0 \) a.e. on \( A \). Fix any
\[
B = [a, c] \subseteq A
\]
and let \( Z \) consist of all \( p \in B \) at which the derivative \( F' = 0 \).

Given \( \varepsilon > 0 \), let \( K \) be the set of all closed intervals \([p, x], \, p < x \), such that
\[
\left| \frac{\Delta F}{\Delta x} \right| = \left| \frac{F(x) - F(p)}{x - p} \right| < \varepsilon.
\]

By assumption,
\[
\lim_{x \to p} \frac{\Delta F}{\Delta x} = 0 \quad (p \in Z),
\]
and \( m(B - Z) = 0 \); \( B = [a, c] \in M^* \). If \( p \in Z \), and \( x - p \) is small enough, then
\[
\left| \frac{\Delta F}{\Delta x} \right| < \varepsilon,
\]
i.e., \( [p, x] \in K \).

It easily follows that \( K \) covers \( Z \) in the Vitali sense (verify); so for any \( \delta > 0 \), Theorem 2 of Chapter 7, §10 yields disjoint intervals
\[
I_k = [p_k, x_k] \in K, \quad I_k \subseteq B,
\]
with
\[
m^*(Z - \bigcup_{k=1}^q I_k) < \delta.
\]
hence also
\[ m \left( B - \bigcup_{k=1}^{q} I_k \right) < \delta \]
(for \( m(B - Z) = 0 \)). But
\[ B - \bigcup_{k=1}^{q} I_k = [a, c] - \bigcup_{k=1}^{q-1} [p_k, x_k] \]
\[ = [a, p_1) \cup \bigcup_{k=1}^{q-1} [x_k, p_{k+1}) \cup [x_q, c] \] (if \( x_k < p_k < x_{k+1} \));
so
\[ m \left( B - \bigcup_{k=1}^{q} I_k \right) = (p_1 - a) + \sum_{k=1}^{q-1} (p_{k+1} - x_k) + (c - x_q) < \delta. \]

Now, as \( F \) is absolutely continuous, we can choose \( \delta > 0 \) so that (3) implies
\[ |F(p_1) - F(a)| + \sum_{k=1}^{q-1} |F(p_{k+1}) - F(x_k)| + |F(c) - F(x_q)| < \varepsilon. \]
But \( I_k \in K \) also implies
\[ |F(x_k) - F(p_k)| < \varepsilon(x_k - p_k) = \varepsilon \cdot mI_k. \]
Hence
\[ \left| \sum_{k=1}^{q} [F(x_k) - F(p_k)] \right| < \varepsilon \sum_{k=1}^{q} mI_k \leq \varepsilon \cdot mB = \varepsilon (c - p). \]

Combining with (4), we get
\[ |F(c) - F(a)| \leq \varepsilon(1 + c - a) \to 0 \text{ as } \varepsilon \to 0; \]
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 3.** A map \( F : E^1 \to E^1 (E^o, C^o) \) is absolutely continuous on \( A = [a, b] \) iff
\[ F = L \int f \text{ on } A \]
for some function \( f \)' and then
\[ F(x) = F(a) + L \int_{a}^{x} f, \quad x \in A. \]

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,
\[ F(x) = F(a) + L \int_{a}^{x} f, \quad x \in A. \]

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
\[ H(x) = L \int_{a}^{x} f, \quad x \in A. \]
Then \( H \), too, is absolutely continuous and so \( F - H \). Also, by Theorem 3 of §1,
\[ H' = f = F', \]
and so
\[ (F - H)' = 0 \text{ a.e. on } A. \]

By Theorem 2, \( F - H = c \); i.e.,
\[ F(x) = c + H(x) = c + L \int_{a}^{x} f, \]
and so \( F = L \int f \) on \( A \), as claimed. \( \square \)

**Corollary 3.** If \( f, F : E^1 \to E^1 (E^o, C^o) \), we have
\[ F = L \int f \]
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
\[ F = \int f \]
with \( L \)-primitives
\[ F = L \int f, \]
such as \( F' \), the derived function of \( F \).
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 3.** We call $p \in E^1$ a Lebesgue point (“L-point”) of $f : E^1 \to E$ if

(i) $f$ is $L$-integrable on some $G_p(\delta)$;
(ii) $q = f(p)$ is finite; and
(iii) $\lim_{x \to p} \frac{1}{|x-p|} L \int_p^x |f - q| = 0$.

The Lebesgue set of $f$ consists of all such $p$.

**Corollary 4.** Let

$$F = L \int f \text{ on } A = [a, b].$$

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,

$$F(x) = c + L \int_p^x f, \quad x \in G_p(\delta),$$

and

$$\frac{1}{|\Delta x|} \left| L \int_p^{x+\Delta x} (f - q) \right| \leq \frac{1}{|\Delta x|} L \int_p^x |f - q| \to 0$$

as $x \to p$. (Here $q = f(p)$ and $\Delta x = x - p$.)

Thus with $x \to p$, we get

$$\frac{|F(x) - F(p)|}{x-p} = \frac{1}{|x-p|} \left| L \int_p^{x+\Delta x} f - (x-p) q \right| = \frac{1}{|x-p|} \left| L \int_p^x f - L \int_p^{x+\Delta x} q \right| \to 0,$$

as required. □

**Corollary 5.** Let $f : E^1 \to E^n (C^n)$. 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 $f$.

(Exercise!)
As

\[ |f(p) - r| = |q - r| < \varepsilon, \]

we get

\[ \frac{1}{|x-p|} L \int_p^x |f - r| \leq \left( \frac{1}{|x-p|} L \int_p^x |f - q| \right) + |q - r| \]

\[ \leq \frac{\varepsilon}{3} + \varepsilon = \frac{2\varepsilon}{3} \]

Hence

\[ L \int_p^x |f - r| < \frac{2\varepsilon}{3} |x-p|. \]

Combining with (6), we have

\[ \frac{1}{|x-p|} L \int_p^x |f - q| < \varepsilon + \frac{2\varepsilon}{3} = \varepsilon \]

whenever |x-p| < \delta. Thus

\[ \lim_{x \to p} \frac{1}{|x-p|} L \int_p^x |f - q| = 0, \]

as required. \[ \square \]

**Problems on L-Integrals and Absolute Continuity**

1. Fill in all details in the proof of Lemma 1 and of Theorems 3 and 4 from §1.

2. Prove Theorem 1 and Corollaries 1, 2, and 5.

3. Disprove the converse to Corollary 4. (Give an example!)

4. Show that if \( F: E^1 \to E \) is L-integrable on \( A = [a, b] \) and continuous at \( p \in A \), then \( p \) is an L-point of \( F \).

5. Let \( F = 1 \) on \( R \) (rationals) and \( F = 0 \) on \( E^1 - R \) (Dirichlet function).

Show that \( F \) has exactly three derivatives (0, +\( \infty \), and -\( \infty \)) at every \( p \in E^1 \).

6. We say that \( F \) is a Lipschitz map, or has the uniform Lipschitz property on \( A \), iff

\[ (\exists K \in E^1) \ (\forall x, y \in A) \ |F(x) - F(y)| \leq K|x - y|. \]

Prove the following:

(i) Any such \( F \) is absolutely continuous on \( A = [a, b] \).

(ii) If all derivatives of \( f \) satisfy

\[ |Df(x)| \leq k < \infty, \ x \in A = [a, b], \]

then \( f \) is a Lipschitz map on \( A \).

7. Let \( g: E_1 \to E_1 \) and \( f: E^1 \to E \) (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 \( A \).

8. Prove that if \( F: E^1 \to E^1 \) 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 \( \delta \) as in Definition 2. As \( m \) is regular, there is an open \( G \),

\[ Q \subseteq G \subseteq (a, b), \]

with \( mG < \delta \). By Lemma 2 of Chapter 7, §2,

\[ G = \bigcup_{k=1}^{\infty} I_k \ (disjoint) \]

for some \( I_k = (a_k, b_k) \).

Let \( u_k = \inf F[I_k], v_k = \sup F[I_k] \); so

\[ F[I_k] \subseteq [u_k, v_k] \]

and

\[ m^* F[I_k] \leq v_k - u_k. \]

Also,

\[ \sum (b_k - a_k) = \sum mI_k = mG < \delta. \]

From Definition 2, show that

\[ \sum_{k=1}^{\infty} (v_k - u_k) \leq \varepsilon \]

(first consider partial sums). As

\[ F[Q] \subseteq F[G] \subseteq \bigcup_k F[I_k], \]

get

\[ m^* F[Q] \leq \sum_k m^* F[I_k] = \sum_k (v_k - u_k) \leq \varepsilon \to 0. \]
9. Show that if \( F \) is as in Problem 8 and if
\[
A = [a, b] \supseteq B, \quad B \in \mathcal{M}^* 
\]
(L-measurable sets), then
\[
F[B] \in \mathcal{M}^*. 
\]
("\( F \) 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 \( B \in \mathcal{F} \), then
\[
B = \bigcup_i B_i, \quad B_i \in \mathcal{F};
\]
so by (i),
\[
F[B] = \bigcup_i F[B_i] \in \mathcal{F} \subseteq \mathcal{M}^*. 
\]
(iii) If \( B \in \mathcal{M}^* \), then by Theorem 2 of Chapter 7, §8,
\[
(3) \text{ if } F \in \mathcal{F} \quad \exists K \in \mathcal{F} \quad K \subseteq B, m(B - K) = 0. 
\]
Now use Problem 8, with \( Q = B - K. \) ]

\( \Rightarrow 10. \) (Change of variable.) Suppose \( g : E^1 \to E^1 \) is absolutely continuous and one-to-one on \( A = [a, b] \), while \( f : E^1 \to E^1 \) (\( E^1, C^n \)) is \( L \)-integrable on \( g[A] \).

Prove that \( (f \circ g)' \) is \( L \)-integrable on \( A \) and
\[
L \int_a^b (f \circ g)' \, \mu = L \int_p^q f, 
\]
where \( p = g(a) \) and \( q = g(b) \).

[Hints: Let \( F = L f \) and \( H = F \circ g \).] By Theorem 3, \( \exists K \in \mathcal{M} \) (end), \( F \) and \( H \) are absolutely continuous on \( g[A] \) and \( A \), respectively, and \( H' \) is \( L \)-integrable on \( A \). So by Theorem 3,
\[
H = L \int H' = L \int (f \circ g)g', 
\]
as \( H' = (f \circ g)' \) a.e. on \( A. \]

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) \( x^2. \)

§2. More on \( L \)-Integrals and Absolute Continuity

(f) \( x \sin(1/x); \)
(g) \( x^2 \sin x^{-2} \) (Problem 5 in §1);
(h) \( \sqrt{x^3} \sin(1/x) \) (verify that \( |f'(x)| \leq \frac{1}{x} + x^{-\frac{3}{2}} \)).

[Hint: Use Problems 6 and 7.]

§3. Improper (Cauchy) Integrals

Cauchy extended \( R \)-integration to unbounded sets and functions as follows.

Given \( f : E^1 \to E \) and assuming that the right-hand side \( R \)-integrals and limits exist, define (first for unbounded sets, then for unbounded functions)

(i) \( \int_a^\infty f \, \mu = \lim_{x \to \infty} R \int_a^x f; \)

(ii) \( \int_{-\infty}^a f \, \mu = \lim_{x \to -\infty} R \int_x^a f. \)

If both
\[
\int_0^\infty f \, \mu \quad \text{and} \quad \int_{-\infty}^0 f 
\]
exist, define
\[
\int_{-\infty}^\infty f = \int_{(-\infty,0)} f + \int_{[0,\infty)} f. 
\]

Now, suppose \( f \) is unbounded near some \( p \in A = [a, b], \) i.e., unbounded on \( A \cap G_p \)
for every deleted globe \( G_p \) about \( p \) (such points are called singularities). Then (again assuming existence of the \( R \)-integrals and limits), we define

(1) in case of a singularity \( p = a, \)
\[
\int_{a+}^b f = \int_{[a,b]} f + \lim_{x \to a+} R \int_x^b f; 
\]
(2) if \( p = b, \)
\[
\int_b^{a-} f = \int_{[a,b]} f + \lim_{x \to b-} R \int_a^x f; 
\]
(3) if \( a < p < b \) and if
\[
\int_a^p f \quad \text{and} \quad \int_p^b f. 
\]
exist, then
\[ \int_a^b f = \int_a^{p-} f + \int_p^b f + \int_b^{p+} f. \]

The term
\[ \int_p^b f = \int_{(p,p]} f \]
is necessary if RS- or LS-integrals are used.\(^1\)

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.\(^2\)

We call all such integrals improper or Cauchy (C) integrals. A C-integral is said to converge if 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
\[ C \int_a^b f \, dm, \quad C \int_{[a,b)} f, \quad \text{etc.} \]

“Classical” notation:

\[ C \int f(x) \, dm(x) \text{ or } C \int f(x) \, dx \]

(the latter if \(m\) is Lebesgue measure). We omit the “C” if confusion with proper integrals \(\int_a^b f\) 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 \(\int_a^b f\) splits accordingly. (Additivity applies to C-integrals; see Problem 9, below.)

**Examples.**

(A) The integral
\[ L \int_{-1}^{1/2} \frac{dx}{x^2} \]
has a singularity at 0. By Theorem 1 below,\(^3\) we get
\[
L \int_{-1}^{1/2} \frac{dx}{x^2} = \int_{-1}^0 \frac{dx}{x^2} + \int_0^{1/2} \frac{dx}{x^2} = \lim_{x \to 0^-} \left( -\frac{1}{x} - 1 \right) + \lim_{x \to 0^+} \left( -2 + \frac{1}{x} \right) = \infty + \infty = \infty.
\]

(B) We have
\[ C \int_{1/2}^\infty \frac{dx}{x^2} = \lim_{t \to \infty} \left( -\frac{1}{x^2} \right) = 2. \]

Hence
\[ C \int_{-1}^\infty \frac{dx}{x^2} = C \int_{-1}^{1/2} \frac{dx}{x^2} + C \int_{1/2}^\infty \frac{dx}{x^2} = \infty + 2 = \infty. \]

(C) The integral
\[ L \int_{-1}^1 \frac{|x|}{x} \, dx \]
has no singularities (consider deleted globes about 0). The primitive \(F(x) = |x|\) exists (example (b) in Chapter 5, §5); so
\[ L \int_{-1}^1 \frac{|x|}{x} \, dx = |x| \bigg|_{-1}^1 = 0. \]

In the rest of this section, we state our theorems mainly for
\[ C \int_a^\infty f, \]
but they apply, with similar proofs, to
\[ C \int_{-\infty}^b f, \quad C \int_a^{b-} f, \quad \text{etc.} \]
The measure \(m\) is as explained above.

**Theorem 1.** Let \(A = [a, \infty), f : E^1 \to E \) (\(E\) complete).

(i) If \(f \geq 0\) on \(A\), then
\[ C \int_a^\infty f \, dm \]
exists (\(\leq \infty\)) and equals
\[ \int_A f \, dm. \]

\(^{1}\) For RS- and LS-integrals, we may well have \(\int_a^b f \neq 0, \int_{[a,b]} f \neq \int_{(a,b)} f\), etc.

\(^{2}\) This also applies if an infinite interval has an inside singularity.

\(^{3}\) It applies to finite intervals \(A\), too.

\(^{4}\) That is, the proper integral.
(ii) The map \( f \) is \( m \)-integrable on \( A \) iff
\[
C \int_a^\infty |f| < \infty
\]
and \( f \) is \( m \)-measurable on \( A \); then again,
\[
C \int_a^\infty f \, dm = \int_A f \, dm.
\]

**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
\[
F(x) = \int_a^x f \, dm, \quad x \geq a.
\]
Then by Theorem 1(f) in Chapter 8, §5, \( F \uparrow \) on \( A \); for \( a \leq x \leq y \) implies
\[
F(x) = \int_a^x f \leq \int_a^y f = F(y).
\]
Now, by the properties of monotone limits,
\[
\lim_{x \to \infty} F(x) = \lim_{x \to \infty} \int_a^x f = C \int_a^\infty f
\]
exists in \( E^*_u \); so by Theorem 1 of Chapter 4, §2, it can be found by making \( x \) run over some sequence \( x_k \to \infty \), say, \( x_k = k \).

Thus set
\[
A_k = [a, k], \quad k = 1, 2, \ldots.
\]
Then \( \{A_k\} \uparrow \) and
\[
\bigcup A_k = A = [a, \infty),
\]
i.e., \( A_k \not\nearrow A \).

Moreover, by Note 4 in Chapter 8, §5, the set function \( s = \int f \) is \( \sigma \)-additive and semifinite (\( \geq 0 \)). Thus by Theorem 2 of Chapter 7, §4 (left continuity)
\[
(1) \quad \int_A f \, dm = \lim_{k \to \infty} \int_{A_k} f = \lim_{k \to \infty} \int_a^k f = C \int_a^\infty f,
\]
proving (i).

(ii) By clause (i),
\[
C \int_a^\infty |f| = \int_A |f| \, dm
\]
exists, as \( |f| \geq 0 \). Hence
\[
C \int_a^\infty |f| < \infty
\]
plus measurability amounts to integrability (Theorem 2 of Chapter 8, §6).

Moreover,
\[
C \int_a^\infty |f| < \infty
\]
implies the convergence of \( C \int_a^\infty f \) (see Corollary 1 below). Thus as
\[
\lim_{x \to \infty} \int_a^x f
\]
exists, we proceed exactly as before (here \( s = \int f \) is finite), proving (ii) also. \( \Box \)

**Note 2.** If \( E \subseteq E^*_u \), formula (1) results even if \( f \) is not \( m \)-measurable.\(^5\)

**Note 3.** While \( f \) cannot be integrable unless \( |f| \) is (Corollary 2 of Chapter 8, §6), it can happen that
\[
C \int_A f
\]
converges even if
\[
C \int |f| = \infty
\]
(this is called conditional convergence). A case in point is
\[
C \int_a^\infty \sin \frac{1}{x} \, dx;
\]
see Problem 8.

Thus \( C \)-integrals may be finite where proper integrals are \( \infty \) 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, §2. (Review it!) As we shall see, \( C \)-integrals resemble infinite series.

**Theorem 2** (Cauchy criterion). Let \( A = [a, \infty), f : E^1 \to E \), \( E \) complete. Suppose
\[
\int_a^x f \, dm
\]
exists for each \( x \in A \). (This is automatic if \( E \subseteq E^*_u \); see Chapter 8, §5.) Then
\[
C \int_a^\infty f
\]
converges iff for every \( \varepsilon > 0 \), there is \( b \in A \) such that
\[
(2) \quad \left| \int_b^x f \, dm \right| < \varepsilon \quad \text{whenever } b \leq v \leq x < \infty,
\]
\(^6\)This is true provided \( \int f \, dm \) is finite or orthodox, so that \( s = \int f \) is semifinite.
and

\( (2') \quad \left| \int_a^b f \, dm \right| < \infty. \)

**Proof.** By additivity (Chapter 8, §5, Theorem 2; Chapter 8, §7, Theorem 3),

\[ \int_a^x f = \int_a^b f + \int_b^x f \]

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

\[ C \int_a^\infty f \]

converges, let

\[ r = \lim_{x \to \infty} \int_a^x f \, dm \neq \pm \infty. \]

Then for any \( \varepsilon > 0 \), there is some

\[ b \in [a, \infty) = A \]

such that

\[ \left| \int_a^x f \, dm - r \right| < \frac{1}{2} \varepsilon \quad \text{for } x \geq b. \]

(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

\[ \left| \int_a^x f \, dm - r \right| < \frac{1}{2} \varepsilon \]

and

\[ \left| r - \int_a^x f \, dm \right| < \frac{1}{2} \varepsilon. \]

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 \( F : A \to E \) by

\[ F(x) = \int_a^x f \, dm, \]

so

\[ C \int_a^\infty f = \lim_{x \to \infty} F(x) \]

---

6 Here and later, for LS integrals, replace \( \int_a^x f \) by \( \int_{(a,x]} \) and \( \int_a^b f \) by \( \int_{[a,b)} \).

§3. Improper (Cauchy) Integrals

if this limit exists. By (2),

\[ |F(x)| = \left| \int_a^x f \, dm \right| \leq \int_a^b f \, dm + \int_b^x f \, dm < \int_a^b f \, dm + \varepsilon \]

if \( x \geq b \). Thus \( F \) is finite on \( [b, \infty) \), and so we may again use the standard metric

\[ \rho(F(x), F(v)) = |F(x) - F(v)| = \left| \int_a^x f \, dm - \int_a^v f \, dm \right| \leq \left| \int_v^x f \, dm \right| < \varepsilon \]

if \( x, v \geq b \). The existence of

\[ C \int_a^\infty f \, dm = \lim_{x \to \infty} F(x) \neq \pm \infty \]

now follows by Theorem 2 of Chapter 4, §2. (We shall henceforth presuppose this "starred" theorem.)

Thus all is proved. \( \square \)

**Corollary 1.** Under the same assumptions as in Theorem 2, the convergence of

\[ C \int_a^\infty |f| \, dm \]

implies that of

\[ C \int_a^\infty f \, dm. \]

Indeed,

\[ \left| \int_v^x f \right| \leq \int_v^x |f| \]

(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 2** (comparison test). If \( |f| \leq |g| \) a.e. on \( A = [a, \infty) \) for some \( f, g : E^1 \to E \), then

\[ C \int_a^\infty |f| \leq C \int_a^\infty |g|; \]

so the convergence of

\[ C \int_a^\infty |g| \]

implies that of

\[ C \int_a^\infty |f|. \]

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

$$E^1 \times E^1 = E^2,$$

and let $p = m \times n$, where $m$ and $n$ are Lebesgue measure or LS measures in $E^1$. Let

$$A = [a, b], \quad B = [c, d], \quad \text{and} \quad D = A \times B.$$

Then the integral

$$\int_B \int_A f \, dm \, dn = \int_Y \int_X f \, C_D \, dm \, dn$$

is also written

$$\int_c^d \int_a^b f \, dm \, dn$$
or

$$\int_c^d \int_a^b f(x, y) \, dm(x) \, dn(y).$$

As usual, we write “$dx$” for “$dm(x)$” if $m$ is Lebesgue measure in $E^1$; similarly for $n$.

We now define

$$C \int_a^\infty \int_c^\infty f \, dn \, dm = \lim_{b \to \infty} \int_a^b \left( \lim_{d \to \infty} \int_c^d f(x, y) \, dn(y) \right) \, dm(x)$$

(3)

$$= C \int_c^\infty \int_a^\infty f(x, y) \, dn(y) \, dm(x),$$

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

$$C \int_c^\infty \int_a^\infty f \, dm \, dn, \quad C \int_{-\infty}^\infty \int_{-\infty}^\infty f \, dm \, dn, \quad \text{etc.}$$

Theorem 3. Let $f : E^2 \to E^*$ be $p$-measurable on $E^2$ ($p, m, n$ as above). Then we have the following.

(i) The Cauchy integrals

$$C \int_{-\infty}^\infty \int_{-\infty}^\infty |f| \, dm \, dn \quad \text{and} \quad C \int_{-\infty}^\infty \int_{-\infty}^\infty |f| \, dm \, dn$$

exist ($\leq \infty$), and both equal

$$\int_{E^2} |f| \, dp.$$

(ii) If one of these integrals is finite, then

$$C \int_{-\infty}^\infty \int_{-\infty}^\infty f \, dm \, dn \quad \text{and} \quad C \int_{-\infty}^\infty \int_{-\infty}^\infty f \, dm \, dn$$

converge, and both equal

$$\int_{E^2} f \, dp.$$

(Similarly for $C \int_a^b \int_{-\infty}^\infty f \, dm \, dn$, etc.)

Proof. As $m$ and $n$ are $\sigma$-finite (finite on intervals!), $f$ surely has $\sigma$-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 4 (passage to polars). Let $p = \text{Lebesgue measure in } E^2$. Suppose $f : E^2 \to E^*$ is $p$-measurable on $E^2$. Set

$$F(r, \theta) = f(r \cos \theta, r \sin \theta), \quad r > 0.$$

Then

(a) $C \int_{-\infty}^\infty \int_{-\infty}^\infty f \, dx \, dy = C \int_0^\infty r \, dr \int_0^{2\pi} F \, d\theta, \quad \text{and}$

(b) $C \int_{-\infty}^\infty \int_{-\infty}^\infty f \, dx \, dy = C \int_0^\infty r \, dr \int_0^{\pi/2} F \, d\theta,$

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”

$$\{(r, \theta) \mid r_1 < r \leq r_2, \theta_1 \leq \theta \leq \theta_2 \}$$

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

$$pD = \frac{1}{2}(r_2^2 - r_1^2)(\theta_2 - \theta_1)$$

(the difference between two circular sectors).

Hence the integrals in (a) and (b) can also be treated as proper integrals.
For \( f = C_D \), formulas (a) and (b) easily follow from

\[
pD = L \int_{g^2} C_D \, dp.
\]

(Verify!) Now, curved rectangles \textit{behave like half-open intervals}

\[(t_1, r_2) \times (\theta_1, \theta_2)\]

in \( E^2 \), since \textbf{Theorem 1} in Chapter 7, \S 1, and \textbf{Lemma 2} of Chapter 7, \S 2, apply with the same proof. Thus they form a \textit{semiring generating the Borel field in} \( E^2 \).

Hence show (as in Chapter 8, \S 8) that \textbf{Theorem 4} holds for \( f = C_D \) (\( D \in B \)). Then take \( D \in \mathcal{M}^* \). Next let \( f \) be elementary and nonnegative, and so on, as in \textbf{Theorems 2} and \textbf{3} in Chapter 8, \S 8. \( \Box \)

\textbf{Examples} (continued).

(D) Let

\[
J = L \int_0^\infty e^{-x^2} \, dx;
\]

so

\[
J^2 = \left( C \int_0^\infty e^{-x^2} \, dx \right) \left( C \int_0^\infty e^{-y^2} \, dy \right)
= C \int_0^\infty \int_0^\infty e^{-(x^2 + y^2)} \, dx \, dy. \quad \text{(Why?)}
\]

Set

\[
f(x, y) = e^{-(x^2 + y^2)}
\]

in \textbf{Theorem 4(b)}. Then \( F(r, \theta) = e^{-r^2} \); hence

\[
J^2 = C \int_0^\infty r \, dr \left( \int_0^{\pi/2} e^{-r^2} \, d\theta \right)
= C \int_0^\infty re^{-r^2} \, dr \cdot \frac{\pi}{2} = \frac{\pi}{4} \pi e^{-1} \big|_0^\infty = \frac{\pi}{4} \pi.
\]

(Here we computed

\[
\int re^{-r^2} \, dr
\]

by substituting \( r^2 = t \).) Thus

\[
C \int_0^\infty e^{-x^2} \, dx = L \int_0^\infty e^{-x^2} \, dx = \sqrt{\frac{\pi}{4}} \pi = \frac{1}{2} \sqrt{\pi}.
\]

\[\boxed{\text{(4)}}\]

\[\text{§3. Improper (Cauchy) Integrals} \quad \text{397}\]

\textbf{Problems on Cauchy Integrals}

1. Fill in all proof details in \textbf{Theorems 1}–\textbf{3}. Verify also at least some of the cases other than \( \int_0^\infty f \). Check the validity for \textit{LS-integrals} (footnote 6).

2. Prove \textbf{Theorem 4} in detail.

2’. Verify Notes 2 and 3 and examples (A)–(D).

3. Assuming \( a > 0 \), verify the following:

\[
\begin{align*}
(i) \int_1^\infty \frac{1}{t} e^{-t} \, dt & \leq \int_1^\infty e^{-t} \, dt = \frac{1}{e}. \\
& \quad \text{[Hint: Use \textit{Corollary 2}]} \\
(ii) \int_1^\infty e^{-at} \, dt & = \frac{e^{-a}}{a}. \\
(iii) \int_0^\infty e^{-at} \, dt & = \frac{1}{a}. \\
(iv) \int_0^\infty e^{-at} \sin bt \, dt & = \frac{b}{a^2 + b^2}.
\end{align*}
\]

4. Verify the following:

\[
\begin{align*}
(i) \int_1^\infty \int_1^\infty e^{-xy} \, dy \, dx & = \int_1^\infty \frac{1}{x} e^{-x} \, dx \leq \frac{1}{e} \quad \text{(converges, by 3(i))}. \\
(ii) \int_0^\infty \int_0^\infty e^{-xy} \, dy \, dx & = \int_1^\infty \int_0^\infty e^{-xy} \, dy \, dx = \int_1^\infty \frac{1}{x} (1 - e^{-x}) \, dx \geq \int_1^\infty \frac{1}{x} - e^{-x} \, dx = \infty.
\end{align*}
\]

Does this contradict formula (4) in the text, or \textbf{Problem 5}, which follows?

5. Let \( f(x, y) = e^{-xy} \) and

\[
g(x) = L \int_0^1 e^{-xy} \, dy;
\]

so \( g(0) = 1 \). (Why?)

(i) Is \( g \) \textit{R-integrable} on \( A = [0, 1] \)? Is \( f \) so on \( A \times A \)?

(ii) Find \( g(x) \) using \textit{Corollary 1} in \S 1.

(iii) Find the value of

\[
R \int_0^1 \int_0^1 e^{-xy} \, dy \, dx = R \int_0^1 g
\]

to within 1/10.

[Hint: Reduce it to \textbf{Problem 6(b)} in \S 1.]
6. Let \( f, g : E^1 \to E^* \) be \( m \)-measurable on \( A = [a, b), b \leq \infty \). Prove the following:

(i) If
\[
C \int_a^b f^+ < \infty \quad \text{or} \quad C \int_a^b f^- < \infty,
\]
then \( C \int_a^b f \) exists and equals
\[
C \int_a^b f^+ - C \int_a^b f^- = \int_A f \, dm \quad \text{(proper)}.
\]

(ii) If \( \int_a^b f \) converges conditionally only, then
\[
\int_a^b f^+ - \int_a^b f^- = \infty.
\]

(iii) In case \( C \int_a^b |f| < \infty \), we have
\[
C \int_a^b (|f| + g) = \infty
\]
if \( C \int_a^b |g| = \infty \); also,
\[
C \int_a^b (f + g) = C \int_a^b f + C \int_a^b g
\]
if \( C \int_a^b g \) exists (finite or not).

7. Suppose \( f : E^1 \to E^* \) is \( m \)-integrable and sign-constant on each
\[
A_n = [a_n, a_{n+1}), \quad n = 1, 2, \ldots,
\]
but changes sign from \( A_n \) to \( A_{n+1} \), with
\[
\bigcup_{n=1}^\infty A_n = [a, \infty)
\]
and \( \{a_n \} \uparrow \) fixed.

Prove that if
\[
\left| \int_{A_n} f \, dm \right| < 0
\]
as \( n \to \infty \), then
\[
C \int_a^\infty f
\]
converges.

[Hint: Use Problem 10 in Chapter 4, §13.]

8. Let
\[
f(x) = \frac{\sin x}{x}, \quad f(0) = 1.
\]

Prove that
\[
C \int_0^\infty f(x) \, dx
\]
converges conditionally only.

[Hints: Use Problem 7. Show that
\[
C \int_0^\infty |f| = L \int_{(0, \infty)} |f| = L \int_0^\infty f^+ = L \int_0^\infty f^- = \infty.
\]

9. (Additivity.) Given \( f : E^1 \to E \) (E complete) and \( a < b < c \leq \infty \), suppose that
\[
\int_a^c f \, dm \neq \pm \infty
\]
(proper) exists for each \( x \in [a, c) \). Prove the following:

(a) \( C \int_a^b f \) and \( C \int_a^c f \) converge.

(b) If \( C \int_b^c f \) converges, so does
\[
C \int_a^c f = C \int_a^b f + C \int_b^c f.
\]

(c) Countable additivity does not necessarily hold for C-integrals.

[Hint: Use Problem 8 suitably splitting \([0, \infty)\).]

10. (Refined comparison test.) Given \( f, g : E^1 \to E \) (E complete) and \( b \leq \infty \), prove the following:

(i) If for some \( a < b \) and \( k \in E^1 \),
\[
|f| \leq |kg| \quad \text{on} \quad [a, b)
\]
then
\[
\int_a^b |g| < \infty \implies \int_a^b |f| < \infty.
\]

(ii) Such \( a, k \in E^1 \) do exist if
\[
\lim_{t \to k^-} |f(t)| < \infty
\]
exists.
(iii) If this limit is not zero, then
\[ \int_a^b |g| < \infty \text{ iff } \int_a^b |f| < \infty. \]

(Similarly in the case of \( \int_a^b \) with \( a \geq -\infty \).)

11. Prove that
(i) \( \int_1^\infty t^p \, dt < \infty \text{ iff } p < -1; \)
(ii) \( \int_0^1 t^p \, dt < \infty \text{ iff } p > -1; \)
(iii) \( \int_0^\infty t^p \, dt = \infty. \)

12. Use Problems 10 and 11 to test for convergence of the following:
(a) \( \int_0^\infty \frac{t^{3/2}}{1 + t^2} \, dt; \)
(b) \( \int_1^\infty \frac{dt}{t\sqrt{1 + t^2}}; \)
(c) \( \int_0^\infty \frac{P(t)}{Q(t)} \, dt; \)
(Q, P polynomials of degree \( s \) and \( r \), \( s > r \); \( Q \neq 0 \) for \( t \geq a \));
(d) \( \int_0^1 \frac{dt}{\sqrt{1 - t^2}}; \)
(e) \( \int_0^1 t^p \ln t \, dt; \)
(f) \( \int_0^1 \frac{dt}{\ln t}; \)
(g) \( \int_0^\pi \tan^p t \, dt. \)

13. (The Abel–Dirichlet test.) Given \( f, g : E^1 \to E^1 \), suppose that
(a) \( f \uparrow \), with \( \lim_{t \to \infty} f(t) = 0; \)
(b) \( g \) is L-measurable on \( A = [a, \infty) \); and
(c) \( \exists K \in E^1 \) (\( \forall x \in A \) \( |L\int_x^\infty g| < K. \)

Then \( \int_a^\infty f(x)g(x) \, dx \) converges.

[Outline: Set \( G(x) = \int_a^x g; \)
so \( |G| < K \) on \( A \). By Lemma 2 of \( \S1 \), \( fg \) is L-integrable on each \( [u, v] \subset A \), and
(\( \exists c \in [u, v] \)) such that
\[ |L\int_u^v f| = |f(u)\int_u^v g| = |f(u)(G(c) - G(u))| < 2Kf(u). \]
Now, by (a),
\[ (\forall \varepsilon > 0) (\exists k \in A) (\forall u \geq k) |f(u)| < \frac{\varepsilon}{2K}; \]
so
\[ (\forall v \geq u \geq k) |L\int_u^v f| < \varepsilon. \]
Now use Theorem 2.

Now extend this to \( g : E^1 \to E^\infty (C^\infty). \]

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 \( q = \lim_{t \to \infty} f(t) \) and \( F = q - f; \) so
\( f g = q g - F g. \)
Apply Problem 13 to
\( C \int_a^\infty F(x)g(x) \, dx. \)]

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) \( \int_0^\infty \frac{\cos t}{\sqrt{t}} \, dt. \)

[Hint: Integrate \( \int_u^\infty \frac{\cos t}{\sqrt{t}} \, dt \) by parts; then let \( u \to 0 \) and \( v \to \infty \).]
(c) \( \int_1^\infty \frac{\cos t}{t^p} \, dt. \)
(d) \( \int_0^\infty \sin t^2 \, dt. \)

[Hint: Substitute \( t^2 = u; \) then use (a).]
16. The Cauchy principal value (CPV) of \( C \int_{-\infty}^{\infty} f(t) \, dt \) is defined by

\[
(CPV) \int_{-\infty}^{\infty} f = \lim_{\varepsilon \to 0} \int_{-\infty}^{\varepsilon} f(t) \, dt
\]

(if it exists). Prove the following:

(i) If \( C \int f(t) \, dt \) exists, so does \( (CPV) \int f \), and the two are equal. Disprove the converse. [Hint: Take \( f(t) = \text{sign}(t)/\sqrt{|t|} \).

(ii) Do the same for

\[
(CPV) \int_{a}^{b} f = \lim_{\delta \to 0+} \left( \int_{a}^{a+\delta} f + \int_{b-\delta}^{b} f \right),
\]

\( p \) being the only singularity in \((a, b)\).

§4. Convergence of Parametrized Integrals and Functions

1. We now consider C-integrals of the form

\[
C \int f(t, u) \, dm(t),
\]

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 \( f : E^{2} \to E \) (\( E \) complete) even if not stated explicitly. As before, we give our definitions and theorems for the case

\[
C \int_{a}^{\infty}.
\]

The other cases \( C \int_{-\infty}^{a}, \, C \int_{b}^{-}, \text{etc.} \) are analogous; they are treated in Problems 2 and 3. We assume

\[
a, b, c, x, t, u, v \in E^{1}
\]

throughout, and write “\( dt \)” for “\( dm(t) \)” iff \( m \) is Lebesgue measure.

If

\[
C \int_{a}^{\infty} f(t, u) \, dm(t)
\]

converges for each \( u \) in a set \( B \subseteq E^{1} \), we can define a map \( F : B \to E \) by

\[
F(u) = C \int_{a}^{\infty} f(t, u) \, dm(t) = \lim_{\varepsilon \to \infty} \int_{a}^{\varepsilon} f(t, u) \, dm(t).
\]

This means that

(i) \( (\forall \, u \in B) \, (\forall \varepsilon > 0) \, (\exists \, b > a) \, (\forall \, x \geq b) \, \left| \int_{a}^{b} f(t, u) \, dm(t) - F(u) \right| < \varepsilon \),

so \( |F| < \infty \) on \( B \).

Here \( b \) depends on both \( \varepsilon \) 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 \( \varepsilon \) alone. We then say that

\[
C \int_{a}^{\infty} f(t, u) \, dm(t)
\]

converges uniformly on \( B \) (i.e., for \( u \in B \)), and write

\[
F(u) = C \int_{a}^{\infty} f(t, u) \, dm(t) \text{ (uniformly) on } B.
\]

Explicitly, this means that

(ii) \( (\forall \varepsilon > 0) \, (\exists \, b > a) \, (\forall \, u \in B) \, (\forall \, x \geq b) \, \left| \int_{a}^{x} f(t, u) \, dm(t) - F(u) \right| < \varepsilon \).

Clearly, this implies (i), but not conversely. We now obtain the following.

Theorem 1 (Cauchy criterion). Suppose

\[
\int_{a}^{\varepsilon} f(t, u) \, dm(t)
\]

exists for \( x \geq a \) and \( u \in B \subseteq E^{1} \). (This is automatic if \( E \subseteq E^{*} \); see Chapter 8, §5.)

Then

\[
C \int_{a}^{\infty} f(t, u) \, dm(t)
\]

converges uniformly on \( B \) iff for every \( \varepsilon > 0 \), there is \( b > a \) such that

(iii) \( (\forall \, v, x \in [b, \infty)) \, (\forall \, u \in B) \, \left| \int_{a}^{x} f(t, u) \, dm(t) \right| < \varepsilon \),

and

\[
\left| \int_{a}^{b} f(t, u) \, dm(t) \right| < \infty.
\]

1 This statement shall imply that \( C \int_{a}^{\infty} f(t, u) \, dm(t) \neq \pm \infty \) exists for \( x \geq a, \, u \in B \).

2 For LS-integrals, replace \( \int_{a}^{\infty} \) by \( \int_{[x, v]} \) here and in the proof below.
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) $u \in B$, 

$$C \int_0^\infty f(t, u) \, dm(t)$$

satisfies Theorem 2 of §3. Hence

$$F(u) = \lim_{x \to \infty} \int_a^x f(t, u) \, dm(t) \neq \pm \infty$$

exists for every $u \in B$ (pointwise). Now, from (3), writing briefly $\int f$ for $\int f(t, u) \, dm(t)$, we obtain

$$\left| \int_v^\infty f \right| = \left| \int_v^\infty f - \int_a^v f \right| < \varepsilon$$

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

$$F(u) - \int_a^v f \leq \varepsilon$$

whenever $v \geq b$.

But by our assumption, $b$ depends on $\varepsilon$ alone (not on $u$). Thus unfixing $u$, we see that (5) establishes the uniform convergence of

$$\int_0^\infty f,$$

as required.\(^3\) □

**Corollary 1.** Under the assumptions of Theorem 1,

$$C \int_0^\infty f(t, u) \, dm(t)$$

converges uniformly on $B$ if

$$C \int_0^\infty |f(t, u)| \, dm(t)$$

does.

Indeed,

$$\left| \int_v^\infty f \right| \leq \int_v^\infty |f| < \varepsilon.$$

\(^3\) Note that Theorem 1 essentially depends on the assumed completeness of $E$. 

**Corollary 2** (comparison test). Let $f : E^2 \to E$ and $M : E^2 \to E^*$ satisfy

$$|f(t, u)| \leq M(t, u)$$

for $u \in B \subseteq E^1$ and $t \geq a$.

Then

$$C \int_a^\infty M(t, u) \, dm(t)$$

converges uniformly on $B$ if

$$C \int_a^\infty |f(t, u)| \, dm(t)$$

does.

Indeed, Theorem 1 applies, with

$$\left| \int_v^\infty f \right| \leq \int_v^\infty M < \varepsilon.$$

Hence we have the following corollary.

**Corollary 3** ("$M$-test"). Let $f : E^2 \to E$ and $M : E^1 \to E^*$ satisfy

$$|f(t, u)| \leq M(t)$$

for $u \in B \subseteq E^1$ and $t \geq a$. Suppose

$$C \int_a^\infty M(t, u) \, dm(t)$$

converges. Then

$$C \int_a^\infty |f(t, u)| \, dm(t)$$

converges (uniformly) on $B$. So does

$$C \int_a^\infty f(t, u) \, dm(t)$$

by Corollary 1.

**Proof.** Set

$$h(t, u) = M(t) \geq |f(t, u)|.$$

Then Corollary 2 applies (with $M$ replaced by $h$ there). Indeed, the convergence of

$$C \int h = C \int M$$

is trivially "uniform" for $u \in B$, since $M$ does not depend on $u$ at all. □
406 Chapter 9. Calculus Using Lebesgue Theory

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 4. Suppose

$$C \int_a^\infty f(t, u) \, dm(t)$$

converges (pointwise) on $B \subseteq E^1$. Then this convergence is uniform iff

$$\lim_{u \to \infty} C \int_u^\infty f(t, u) \, dm(t) = 0 \text{ (uniformly) on } B,$$

i.e., iff

$$(\forall \varepsilon > 0) \, (\exists b > a) \, (\forall u \in B) \, (\forall v \geq b) \, \left| C \int_v^\infty f(t, u) \, dm(t) \right| < \varepsilon.$$

The proof (based on Theorem 1) is left to the reader, along with that of the following corollary.

Corollary 5. Suppose

$$\int_a^b f(t, u) \, dm(t) \neq \pm \infty$$

exists for each $u \in B \subseteq E^1$.

Then

$$C \int_a^\infty f(t, u) \, dm(t)$$

converges (uniformly) on $B$ iff

$$C \int_b^\infty f(t, u) \, dm(t)$$

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

$$H : E^2 \to E \text{ (} E \text{ complete)},$$

we say that $H(x, y)$ converges to $F(y)$, uniformly on $B$, as $x \to q \,(q \in E^*)$, and write

$$\lim_{x \to q} H(x, y) = F(y) \text{ (uniformly) on } B$$

iff we have

$$(6) \quad (\forall \varepsilon > 0) \, (\exists G_{-\varepsilon}) \, (\forall y \in B) \, (\forall x \in G_{-\varepsilon}) \, \left| H(x, y) - F(y) \right| < \varepsilon,$$

hence $|F| < \infty$ on $B$.

If here $q = \infty$, the deleted globe $G_{-\varepsilon}$ has the form $(b, \infty)$. Thus if

$$H(x, u) = \int_u^\infty f(t, u) \, dt,$$

(6) turns into (2) as a special case. If (6) holds with “$(\exists G_{-\varepsilon})$” and “$(\forall y \in B)$” interchanged, as in (1), convergence is pointwise only.

As in Chapter 8, §8, we denote by $f(x, y)$, or $f^y$, the function of $x$ alone (on $E^1$) given by

$$f^y(x) = f(x, y).$$

Similarly,

$$f_x(y) = f(x, y).$$

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 2. Assume $f, g : E^2 \to E^1$ 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 $(\downarrow \text{ or } \uparrow)$ on $A$;
(iv) $|f| < K \in E^1$ (bounded) on $A \times B$.

Then

$$C \int_a^\infty f(t, u) g(t, u) \, dt$$

converges uniformly on $B$.

Proof. Given $\varepsilon > 0$, use assumption (i) and Theorem 1 to choose $b > a$ so that

$$|L \int_v^x g(t, u) \, dt| < \frac{\varepsilon}{2K},$$

written briefly as

$$\left| L \int_v^x g^u \right| < \frac{\varepsilon}{2K},$$

for all $u \in B$ and $x > v \geq b$, with $K$ as in (iv).

4 Briefly: “$f(t, u)$ is monotone in $t$, and $g(t, u)$ is measurable in $t \,(t \in A)$.” It should be well noted that all $f^u$ and $g^u$ are functions of $t$ on $E^1$. 

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

\[
L \int_v^x f^u g^u = f^u(v) L \int_v^x g^u + f^u(x) L \int_v^x g^u
\]

for some \( c \in [v, x] \).

Combining with (7) and using (iv), we easily obtain

\[
\left| L \int_v^x f(t, u) g(t, u) \, dt \right| < \varepsilon
\]

whenever \( u \in B \) and \( x > v \geq b\). (Verify!)

Our assertion now follows by Theorem 1. \( \square \)

**Theorem 3** (Abel–Dirichlet test). Let \( f, g: E^2 \to E^\ast \) satisfy

(a) \( \lim_{t \to \infty} f(t, u) = 0 \) (uniformly) for \( u \in B \);

(b) each \( f^u(u \in B) \) is nonincreasing (↓) on \( A = [0, \infty) \);

(c) each \( g^u(u \in B) \) is \( L\)-measurable on \( A \); and

(d) (\( \exists K \in E^1 \)) (\( \forall x \in A \)) (\( \forall u \in B \)) \( |L \int_v^x g(t, u) \, dt| < K \).

Then

\[
C \int_0^\infty f(t, u) g(t, u) \, dt
\]

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

\[
\left| L \int_v^x f^u g^u \right| = \left| f^u(v) L \int_v^x g^u \right| \leq K f(v, u)
\]

for \( u \in B \) and \( x > v \geq a \).

Then use assumption (a) to fix \( k \) so that

\[
|f(t, u)| < \frac{\varepsilon}{2K}
\]

for \( t > k \) and \( u \in B\). \( \square \)

**Note 2.** Via components, Theorems 2 and 3 extend to the case \( g: E^2 \to E^n \) (Continuity).

**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.

§4. Convergence of Parametrized Integrals and Functions

**Examples.**

(A) The integral

\[
\int_0^\infty \frac{\sin tu}{t} \, dt
\]

converges uniformly on \( B_3 = [\delta, \infty) \) if \( \delta > 0 \), and pointwise on \( B = [0, \infty) \).

Indeed, we can use Theorem 3, with

\[
g(t, u) = \sin tu
\]

and

\[
f(t, u) = \frac{1}{t}, \quad f(0, u) = 1,
\]

say. Then the limit

\[
\lim_{t \to \infty} \frac{1}{t} = 0
\]

is trivially uniform for \( u \in B_3 \), as \( f \) is independent of \( u \). Thus assumption (a) is satisfied. So is (d) because

\[
\left| \int_0^x \sin tu \, dt \right| = \left| \frac{1}{u} \int_0^x \sin \theta \, d\theta \right| \leq \frac{1}{u} \cdot 2.
\]

(Explain!) The rest is easy.

Note that Theorem 2 fails here since assumption (i) is not satisfied.

(B) The integral

\[
\int_0^\infty \frac{1}{t} e^{-tu} \sin at \, dt
\]

converges uniformly on \( B = [0, \infty) \). It does so absolutely on \( B_3 = [\delta, \infty) \), if \( \delta > 0 \).

Here we shall use Theorem 2 (though Theorem 3 works, too). Set

\[
f(t, u) = e^{-tu}
\]

and

\[
g(t, u) = \frac{\sin at}{t}, \quad g(0, u) = a.
\]

Then

\[
\int_0^\infty g(t, u) \, dt
\]

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 \( B_3 \), use Corollary 3 with

\[
M(t) = e^{-\delta t},
\]
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 \( G_{\text{depl}} \) is fixed.

**Theorem 4.** Suppose

\[
\lim_{x \to y} H(x, y) = F(y) \quad \text{(uniformly)}^{5}
\]

for \( y \in B \subseteq E^1 \). Then we have the following:

(i) If all \( H_k \) \((x \in G_{\text{depl}})\) are continuous\(^6\) or \( m \)-measurable on \( B \), so also is \( F \).

(ii) The same applies to \( m \)-integrability on \( B \), provided \( mB < \infty \); and then

\[
\lim_{x \to y} \int_B |H_k - F| = 0;
\]

hence

\[
\lim_{x \to y} \int_B H_k = \int_B F = \int_B \left( \lim_{x \to y} H_k \right). \tag{8'}
\]

Formula (8') is known as the rule of passage to the limit under the integral sign.

**Proof.** (i) Fix a sequence \( x_k \to q \) \((x_k \text{ in the deleted globe } G_{\text{depl}})\), and set

\[
H_k = H_{x_k} \quad (k = 1, 2, \ldots).
\]

The uniform convergence

\[
H(x, y) \to F(y)
\]

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 \( H_k \) be \( m \)-integrable on \( B \); let

\[
mB < \infty.
\]

Then the \( H_k \) are \( m \)-measurable on \( B \), and so is \( F \), by (i). Also, by (6),

\[(\forall \varepsilon > 0) \ (\exists \varepsilon \in G_{\text{depl}}) \ (\forall x \in G_{\text{depl}}) \int_B |H_k - F| \leq \int_B (\varepsilon) = \varepsilon mB < \infty,
\]

proving (8). Moreover, as

\[
\int_B |H_k - F| < \infty,
\]

\( H_k - F \) is \( m \)-integrable on \( B \), and so is

\[
F = H_k - (H_k - F).
\]

Hence

\[
\left| \int_B H_k - \int_B F \right| = \int_B (H_k - F) \leq \int_B |H_k - F| \to 0,
\]

as \( x \to q \), by (8). Thus (8') is proved, too. \( \Box \)

Quite similarly (keeping \( E \) complete and using sequences), we obtain the following result.

**Theorem 5.** Suppose that

(i) all \( H_k \) \((x \in G_{\text{depl}})\) are continuous and finite on a finite interval \( B \subseteq E^1 \),

and differentiable on \( B - Q \), for a fixed countable set \( Q \);

(ii) \( \lim_{x \to y} H(x, y) \neq \pm \infty \) exists for some \( y_0 \in B \); and

(iii) \( \lim_{x \to y} D_x H(x, y) = f(y) \) (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,

\[
F(y) = \lim_{x \to y} H(x, y) \quad \text{(uniformly) for } y \in B.
\]

**Outline of proof.** Note that

\[
D_x H(x, y) = \frac{d}{dy} H_k(y).
\]

Use Theorem 1 of Chapter 5, §9, with \( F_n = H_{x_n} \), \( x_n \to q \). \( \Box \)

**Note 4.** If \( x \to q \) over a path \( P \) (clustering at \( q \)), one must replace \( G_{\text{depl}} \) and \( G_{\text{depl}} \) by \( P \cap G_{\text{depl}} \) and \( P \cap G_{\text{depl}} \) in (6) and in Theorems 4 and 5.
Problems on Uniform Convergence of Functions and C-Integrals

1. Fill in all proof details in Theorems 1–5, Corollaries 4 and 5, and examples (A) and (B).

2. Using (6), prove that
\[ \lim_{x \to q} H(x, y) \text{ (uniformly)} \]
exists on \(B \subseteq E^3\) iff
\( (\forall \varepsilon > 0) (\exists G_{\varepsilon}) (\forall y \in B) (\forall x, x' \in G_{\varepsilon}) \quad |H(x, y) - H(x', y)| < \varepsilon. \)

Assume \(E\) complete and \(|H| < \infty\) on \(G_{\varepsilon} \times B.

[Hint: "Imitate" the proof of Theorem 1, using Theorem 2 of Chapter 4, §2.]

3. State formulas analogous to (1) and (2) for \(\int_{-\infty}^{a}, \int_{a}^{b}, \text{ and } \int_{a}^{\infty}. \)

4. Do at least some of the cases involved.

[Hint: Use Theorem 1 of §3 and Problem 1', if already solved.]

5. Prove that
\[ \lim_{x \to q} H(x, y) = F(y) \text{ (uniformly)} \]
on \(B\) iff
\[ \lim_{n \to \infty} H(x_n, \cdot) = F \text{ (uniformly)} \]
on \(B\) for all sequences \(x_n \to q \quad (x_n \neq q)\).

[Hint: "Imitate" Theorem 1 in Chapter 4, §2. Use Definition 1 of Chapter 4, §12.]

6. Prove that if
\[ \lim_{x \to q} H(x, y) = F(y) \text{ (uniformly)} \]
on \(A\) and on \(B\), then this convergence holds on \(A \cup B\). Hence deduce similar propositions on \(C\)-integrals.

7. Show that the integrals listed below violate Corollary 4 and hence do not converge uniformly on \(P = (0, \delta)^2\) though proper L-integrals exist for each \(u \in P\). Thus show that Theorem 1(ii) does not apply to uniform convergence.

(a) \[\int_{0}^{1} \frac{udt}{(t^2 - u^2)};\]
(b) \[\int_{0}^{1} \frac{u^2 - t^2}{(t^2 + u^2)^2} dt;\]
(c) \[\int_{0}^{1} \frac{tu(t^2 - u^2)}{(t^2 + u^2)^2} dt.\]

[Hint for (b)]: To disprove uniform convergence, fix any \(\varepsilon, v > 0\). Then
\[\int_{0}^{v} \frac{u^2 - t^2}{(t^2 + u^2)^2} dt = \frac{v}{v^2 + u^2} \to \frac{1}{v}\]
as \(u \to 0\). Thus if \(v < \frac{1}{2}\),
\[\exists u \in P \quad \int_{0}^{v} \frac{u^2 - t^2}{(t^2 + u^2)^2} dt > \frac{1}{2v} > \varepsilon.\]

§4. Convergence of Parametrized Integrals and Functions

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) \[\int_{0}^{1} e^{-ut^2} dt; \quad U = [\delta, \infty); \quad P = (0, \delta).\]

[Hint: Set \(M(t) = e^{-at}\) for \(t \geq 1\) (Corollaries 3 and 5).]

(b) \[\int_{0}^{1} e^{-ut^2} \cos t dt \quad (u \geq 0); \quad U = [\delta, \infty).\]

(c) \[\int_{0}^{1} t^{n-1} dt; \quad U = [\delta, \infty).\]

(c') \[\int_{0}^{1} t^{n-1}(\ln t)^n dt; \quad U = [\delta, \infty).\]

(d) \[\int_{0}^{1} t^{-\alpha} \sin t dt; \quad U = [0, \delta], \quad 0 < \delta < 2; \quad P = [\delta, 2); \quad M(t) = t^{1-\delta}.\]

[Hint: Fix \(v\) so small that
\[(\forall t \in (0, v)) \quad \sin t \geq \frac{1}{2}\]
Then, if \(u \to 2\),
\[\int_{0}^{v} t^{-\alpha} \sin t dt \geq \frac{1}{2} \int_{0}^{v} dt \int_{0}^{t^{-1}} \to \infty.\]

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) \[\int_{0}^{1} \frac{u dt}{(t^2 - u^2)};\]
(b) \[\int_{0}^{1} \frac{u^2 - t^2}{(t^2 + u^2)^2} dt;\]
(c) \[\int_{0}^{1} \frac{tu(t^2 - u^2)}{(t^2 + u^2)^2} dt.\]

[Hint for (b)]: To disprove uniform convergence, fix any \(\varepsilon, v > 0\). Then
\[\int_{0}^{v} \frac{u^2 - t^2}{(t^2 + u^2)^2} dt = \frac{v}{v^2 + u^2} \to \frac{1}{v}\]
as \(u \to 0\). Thus if \(v < \frac{1}{2}\),
\[\exists u \in P \quad \int_{0}^{v} \frac{u^2 - t^2}{(t^2 + u^2)^2} dt > \frac{1}{2v} > \varepsilon.\]
8. In example (A), disprove uniform convergence on $P = (0, \infty)$.
[Hint: Proceed as in Problem 6.]

9. Do example (B) using Theorem 3 and Corollary 5. Disprove uniform convergence on $B$.

10. Show that
$$\int_{0^+}^{\infty} \frac{\sin tu}{t} \cos t \ dt$$
converges uniformly on any closed interval $U$, with $\pm 1 \notin U$.
[Hint: Transform into $\frac{1}{2} \int_{0}^{\infty} \left( \sin((u+1)t) + \sin((u-1)t) \right) dt$.]

11. Show that
$$\int_{0}^{\infty} t \sin t^3 \sin tu \ dt$$
converges (uniformly) on any finite interval $U$.
[Hint: Integrate $\int_{0}^{y} t \sin t^3 \sin tu \ dt$ by parts twice. Then let $y \to \infty$ and $x \to 0$.]

12. Show that
$$\int_{0^+}^{\infty} e^{-tu} \cos \frac{t}{u^n} \ dt \quad (0 < a < 1)$$
converges (uniformly) for $u \geq 0$.
[Hint: For $t \to 0^+$, use $M(t) = t^{-a}$. For $t \to \infty$, use example (B) and Theorem 2.]

13. Prove that
$$\int_{0^+}^{\infty} \frac{\cos tu}{t^n} \ dt \quad (0 < a < 1)$$
converges (uniformly) for $u \geq \delta > 0$, but (pointwise) for $u > 0$.
[Hint: Use Theorem 3 with $g(t, u) = \cos tu$ and
$$\left| \int_{0}^{t^n} \frac{\sin tu}{u} \ dt \right| \leq \frac{1}{\delta}$$
For $u > 0$,
$$\int_{0}^{t^n} \frac{\cos tu}{t^n} \ dt = u^{n-1} \int_{0}^{t^n} \cos \frac{z}{u} \ dz \to \infty$$
if $v = 1/u$ and $u \to 0$. Use Corollary 4.]

14. Given $A, B \subseteq E^1$ ($mA < \infty$) and $f : E^2 \to E$, 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 $A$.

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 $[a, x] \subseteq E^1$, $x \geq a$.

Then show that $F$ is relatively continuous, hence integrable, on $B$ and that
$$\int_{B} F = \lim_{x \to \infty} \int_{a}^{x} H_x,$$
where
$$H(x, y) = \int_{a}^{x} f(t, y) \ dm(t).$$
(Passage to the limit under the $f$-sign.)
[Hint: Use Problem 14 and Theorem 4; note that
$$C \int_{0}^{\infty} f(t, y) \ dm(t) = \lim_{x \to \infty} H(x, y)$$ (uniformly).]
Index

Abel–Dirichlet test
  for convergence of improper integrals, 400
  for uniform convergence of parametrized C-integrals, 408
Absolute
  extrema, 82
  maxima, 82
  minima, 82
Absolute continuity of the integral, 275
Absolute convergence of improper integrals, 393
Absolutely continuous functions on $E^1$, 378
  and L-integrals, 380
Absolutely continuous with respect to a set function $t$, 197
Additive extensions of set functions, 129
Additive set functions, 126, 126
Additivity of the integral, 260, 290
Additivity of volume
  countable, 104
  of intervals, 101
  $\sigma$-additivity, 104
Almost everywhere (a.e.), 231
  convergence of functions, 231
Almost measurable functions, 231, 231
Almost uniform convergence of functions, 239
Egorov’s theorem, 240, 283
Antiderivatives, 357
  and L-integrals, 357
  and R-integrals, 362
  change of variable in, 363
  primitives, 359
Baire categories (of sets), 70
  sets of Category I, 71
  sets of Category II, 71
Baire’s theorem, 71
Banach spaces, 76
integration of functions with values in, 285–291, 305
  open map principle, 75
  uniform boundedness principle, 75
Banach-Steinhaus uniform boundedness principle, 75
Basic covering of a set, 138
Basic covering value of a set, 138
Basis of a vector space, 16
Bicontinuous maps, 70
Bijective functions, 52
  linear maps, 53
Borel fields, 162
  measurable functions, 222
  measures, 162
  restrictions of measures, 162
  sets, 162
Boundedness, linear, 9
  $C^*_s$-sets, 104
  volume of, 107
  $C$-simple sets, 99
  $C^*_s$, family of C-simple sets, 99
C-integrals, see Improper integrals
  parametrized, 402; see also Parametrized C-integrals
Cantor’s set, 76
Carathéodory property (CP), 145, 146, 157
Cauchy criterion
  for convergence of improper integrals, 391
  for uniform convergence of parametrized C-integrals, 403
Cauchy integrals (C-integrals), see Improper integrals
  parametrized, 402; see also Parametrized C-integrals
Cauchy principal value (CPV), 402
Index

418

419

Basic, 138
M-coverings of a set, 137
Ω-coverings of a set, 213
Vitali, 180; see also Vitali coverings
CP, the Carathéodory property, 145
Critical points, 82
Darboux sums (upper and lower), 307
Decompositions
Lebesgue, 342
of generalized measures, 344
Derivatives
of point functions, 373
of set functions, 187
Derivatives
directional, see Directional derivatives of set functions, 210
Radon-Nikodym, 338, 351
partial, see Partial derivatives
Determinants
functional, 49
of matrices, 47, 96
Differentiable functions, 17
and directional derivatives, 19
chain rule for, 7
continuously, 38, 57
differentials of, 17
and partial derivatives, 19, 22
in a normed space, 17
m times differentiable, 38
Differentiable set functions, 210
Differentials, 17
chain rule for, 7
of functions in a normed space, 17
of order m, 39
Differentiation of set functions, 210–216
Differential, 211
Lebesgue differentiation, 211, 351
Ω-differentiation, 211, 353
Directional derivatives, 1
Differentiable functions and, 19
Finite Increment Law for, 7
higher order, 35
of linear maps, 15
Discriminant of a quadratic polynomial, 80
Disjoint set families, 99
Domination convergence theorem, 273, 327
Dot products, linear functionals on $E^n$ and $C^m$, 10
Double series, 110, 115
$E^n$ intervals in, 97
volume of open sets in, 108
Elementary functions, 218
integrable, 241
integrals of, 241
integration of, 241–250
Euler’s theorem for homogeneous functions, 34
Extended-real functions
integration of, 251–267; see also Integration of extended-real functions integrable, 252
lower integrals of, 251
upper integrals of, 251
Extremum, extrema
absolute, 82
conditional, 88
local, 79, 89
Fatou’s lemma, 272
Fields of sets, 116
generated by a set family, 117
Finite Increments Law for directional derivatives 7
Finite set functions, 125
Finite with respect to a set function $t$ (t-
finite), 197
Finely additive set functions, 126, 126
Fréchet’s theorem, 237
Fubini map, 294
theorem, 298, 301, 305, 334
Functional determinants, 49
Functionals, linear, see Linear functionals
Functions. See also Maps
bijective, 52
continuous, 161
differentiable, see Differentiable functions
homeomorphisms, 70
homogeneous, 34
implicit function theorem, 64
inverse function theorem, 61
partially derived, 2
Fundamental theorem of calculus, 360
Generalized integration, 323ff.
change of measure, 332
dominated convergence theorem, 327
Fubini property in, 334
indefinite integrals in, 330
Generalized measure spaces, 194
integration in, 323ff.
Generalized measures, 194
completion of, 205
decomposition of, 344
signed measures, 194, 199
Gradient of a function, 20
Hadamard’s theorem, 96
Hahn decomposition theorem, 201
Hereditary set families, 123
Homeomorphisms, 70
Homogeneous functions, 34
Euler’s theorem for, 34
Implicits
differentiation, 66, 87
function theorem, 64
Improper integrals, 388
absolute convergence of, 393
Cauchy criterion for, 391
Cauchy principal value (CPV) of, 402
comparison test for, 393, 399
conditional convergence of, 391
Abel–Dirichlet test for convergence of, 400
iterated, 394
convergence of, 388
singularities of, 387
Indefinite integrals, 263, 293, 330
indefinite $L$-integrals, 366
independence, linear, 16
Inner products representing linear functionals on $E^n$ and $C^m$, 10
Integrable functions
elementary, 241
extended-real, 252
with values in complete normed spaces, 285
Riemann, 307, 317
Integrals
Cauchy (C-integrals), 388; see also Improper integrals
in generalized measure spaces, 323ff.
indefinite, 263, 293, 330
improper, 388; see also Improper integrals iterated, 294
Lebesgue, 357
Lebesgue integrals and Riemann integrals, 313
lower, 251

Index
of elementary functions, 241
orthodox, 247
parametrized C-integrals, 402; see also
Parametrized C-integrals
Riemann (R-integrals), 308ff.; see also
Riemann integrals
Riemann–Stieltjes, 318
Stieltjes, 319, 321ff.
unorthodox, 247
upper, 251
with respect to Lebesgue measure (L-
integrals), 357
Integration
absolute continuity of the integral, 275
additivity of the integral, 260, 290
by parts, 321
dominated convergence theorem, 273, 327
Fatou’s lemma, 272
in generalized measure spaces, 323ff.
of elementary functions, 241–250
of extended-real functions, 251–267
of functions with values in Banach
spaces, 285–291, 305
linearity of the integral, 267, 288
monotone convergence theorem, 271
weighted law of the mean, 269
Intervals in $E^n$, 97
additivity of volume of, 101
simple step functions on, 218
step functions on, 218
Inverse function theorem, 61
Iterated integrals, 294
iterated improper integrals, 394
Fubini map, 294
Fubini theorem, 298, 301, 305, 334
Jacobian matrix, 18
Jacobians, 49
Jordan components, 203
Jordan decompositions, 202
Jordan components, 203
Jordan outer content, 140
$\mathcal{F}$ (the set of all cubes in $E^n$), 186, 210
$L$-measurable, see Lebesgue-measurable.
$L$-integrable, see Lebesgue-integrable.
$L$-integrals, 357
and absolutely continuous functions, 380
indefinite, 366
$L$-primitive, 366
Lagrange form of the remainder in Taylor’s
Theorem, 42
Lagrange multipliers, 89
Lebesgue
decompositions, 342
extensions, 154, 168
Lebesgue-integrable functions, 241
Lebesgue-measurable functions, 222
Lebesgue-measurable sets, 168
measure, 168–175
outer measure, 138
nonmeasurable sets under Lebesgue
measure, 173
points of functions, 382
premeasure, 126, 138, 168
premeasure space, 138
sets of functions, 382
Lebesgue–Stieltjes
measurable functions, 222
measures, 176
measures in $E^n$, 179
outer measures, 146, 176
premeasures, 176
set functions, 127, 135, 176
signed Lebesgue–Stieltjes measures,
206, 335
Left-continuous set functions, 131
Linear boundedness, 9
Linear functionals, 7
on $E^n$ and $C^n$ as dot products, 10
Linear independence, 16
Linear maps, 7
as a normed linear space, 13
bijective, 53
bounded, 9
continuous, 9, 13
directional derivatives of, 15
matrix representation of composite, 12
matrix representation of, 11
norm of, 13
uniformly continuous on $E^n$ or $C^n$, 10
Linear operator, 7
Linear subspaces of a vector space, 16
Linearity of the integral
of extended-real functions, 267
of functions with values in Banach
spaces, 288
Lipschitz condition, 7, 384
uniform, 7
Local
extremum, extrema, 79, 89
maximum, maxima, 79
minimum, minimina, 79
Lower
Darboux sums, 307
integrals, 251
Riemann integrals, 308
LS, see Lebesgue–Stieltjes
Luzin’s theorem, 234
M-test for uniform convergence of param-
eterized C-integrals, 405
Maps. See also Functions
bicontinuous, 70
clopen, 61
closed, 59
linear, see Linear maps
open, 59
open map principle, 75
Matrix, matrices
as elements of a vector space, 15
determinants of, 47, 96
Jacobian, 18
$n \times n$ matrices as a noncommutative ring
with identity, 15
of composite linear maps, 12
representation of a linear map, 11
Maximum, maxima
absolute, 82
conditional, 88
local, 79
Meagre sets, 71
Measurable covers of sets, 156
Measurable functions
almost, 231
Borel, 222
Fréchet’s theorem, 237
Lebesgue (L), 222
Lebesgue–Stieltjes (LS), 222
Luzin’s theorem, 234
$M$-measurable functions, 218
$m$-measurable functions, 231
Tietze’s theorem, 236
Measurable sets, 147
nonmeasurable sets under Lebesgue
measure, 173
outer, 149
Measurable spaces, 217
Measure spaces, 147
almost measurable functions on sets in,
231
probability spaces as, 148
topological, 162
Measures, 147, 194. See also Set functions
Borel restrictions of, 162
generalized measures, 194
Lebesgue premeasure, 126
Lebesgue–Stieltjes, 127, 135, 176
left-continuous, 131
monotone, 136, 147
outer measures, 138; see also Outer measures
premeasures, 137
regular, 140, 155
right-continuous, 131
rotation-invariant, 192
σ-additive, 126
σ-finite, 140
σ-subadditive, 137
semifinite, 126
signed measures, 194, 199
signed Lebesgue–Stieltjes measures, 206, 335
singular with respect to a set function t
(t-singular), 341
total variation of, 194
totally σ-finite, 140, 169
translation-invariant, 171
volume of sets, see Volume
Set rings, 101, 115
generated by a set family, 117
Signed Lebesgue–Stieltjes measures, 206
induced by a function of bounded variation, 206
integration in, 335
Signed measure spaces, 194, 199
Hahn decomposition theorem, 201
Jordan components, 203
Jordan decompositions, 202
negative sets in, 199
positive sets in, 199
Simple functions, 218
simple step functions, 218
 Singular with respect to a set function t
(t-singular), 341
 Singularities of improper integrals, 387
 Span of vectors in a vector space, 16
 Step functions, 218
 simple, 218
 Stieltjes integrals, 319, 321ff.
 integration by parts, 321
 laws of the mean, 322
 Strongly regular measures, 162, 234, 237, 347
 Sylvester’s theorem, 80
 Symmetric difference of sets, 122
 Symmetric quadratic forms, 80
 Sylvester’s theorem, 80
 Taylor polynomial, 43
 Taylor’s Theorem, 40
 generalized, 45
 Lagrange form of remainder, 42
 Taylor polynomial, 43
 Tietze’s theorem, 236
 Topological measure spaces, 162
 Topological spaces, 161
 compact regular (CR) set functions on,
 continuous functions between, 161
 metric spaces as, 161
 pseudometric spaces as, 165
 Topologies, 161
 closed sets in, 161
 open sets in, 161
 Total variation of set functions, 194
 Totally σ-finite set functions, 140, 169
 Translation-invariant set functions, 171
 Uniform boundedness principle of Banach
 and Steinhaus, 75
 Uniformly normal Vitali coverings, 192
 Universal Vitali coverings, 192
 Unorthodox integrals, 247
 Upper
 Darboux sums, 307
 integrals, 251
 Riemann integrals, 307
 V-coverings, see Vitali coverings
 Vectors
 span of a set of, 16
 Vector spaces
 basis of, 16
 dimension of, 16
 linear subspaces of, 16
 matrices as elements of, 15
 span of vectors in, 16
 Vitali coverings, 180
 normal, 192
 uniformly normal, 192
 universal, 192
 Volume
 additivity of volume of intervals, 101
 monotonicity of, 109
 of Δ-sets in $E^n$, 107
 of open sets in $E^n$, 108
 Weighted law of the mean, 269