Measure and Integral

Jaroslav Lukeš
Jan Malý

matfyzpress
PRAGUE 2005

ISBN 80-86732-68-1
ISBN 80-85863-06-5 (First edition)

Motto: Everybody writes and nobody reads
L. Fejér

Preface

This text is based on lectures in measure and integration theory given by the authors during the past decade at Charles University, and on preliminary lecture notes published in Czech.

It is impossible to thank individually all colleagues and students who assisted in the preparation of this manuscript, but we will just mention Michal Kubeček who helped with the translation and TEX processing.

The authors wish to express their deep gratitude to Professor Stylianos Negrepontis, who was the chief coordinator of TEMPUS project JEP–1980. Without support from him and the Tempus programme the manuscript would never have appeared.

The preparation of this manuscript was partially supported by the grant No. 201/93/2174 of the Czech Grant Agency and by the grant No. 354 of the Charles University.

Prague, 1994 Jaroslav Lukeš and Jan Malý

Preface to the second edition

We have carried out only minor corrections. We wish to thank all who contributed by suggestions and comments.

Prague, 2005 Jaroslav Lukeš and Jan Malý

Contents

List of Basic Notations and Frequently Used Symbols

In this manuscript we use the standard notation.

In all what follows, $\mathbf{N}$, $\mathbb{Z}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{C}$ will denote the sets of all natural, integer, rational, real and complex numbers, respectively.

Extended real number set $\overline{\mathbf{R}}$ consists of $\mathbf{R}$ together with two symbols $-\infty$ and $+\infty$ equipped with the usual algebraic structure and topology. Remind only that $0 \cdot \pm\infty$ and $\pm\infty \cdot 0$ are taken as 0.

If X is a set, $\mathscr{P}(X)$ denotes the collection of all its subsets.

$\mathbf{R}^n$ stands for the Euclidean n-dimensional space under the usual Euclidean norm $|\cdot|$ and the metric $|x - y|$, where $x = [x_1, \dots, x_n]$.

Remember that when multiplied by a matrix (from the left), the vector $x = [x_1, \dots, x_n]$ behaves like a "vertical vector", i.e. like a matrix with one column

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

The horizontal notation is preferred for estetical and typographical reasons.

The *standard* (or *canonical*) basis of the space $\mathbf{R}^n$ is denoted by $\{e_1, \dots, e_n\}$, the vector $e_i = [0, \dots, 0, 1, 0, \dots, 0]$ with 1 at the i^{th} place. The inner product in $\mathbf{R}^n$ is denoted by $x \cdot y$.

By $U(x, r)$ we denote the open ball in a metric space (P, ρ) of radius r round the x. The closed ball is denoted by $B(x, r)$. Thus $U(x, r) = \{y \in P\colon \rho(x, y) < r\}$, $B(x, r) = \{y \in P\colon \rho(x, y) \le r\}$. For the diameter of a set we use the symbol "diam" and for the distance of a pair of sets the symbol "dist".

If nothing else is specified, a *function* on a set X is a mapping of X into $\overline{\mathbf{R}}$. If we want to emphasise that a function does not attain the values $-\infty$ and $+\infty$, we call it a *real* function.

Instead of the notation $\{x \in X\colon f(x) > a\}$ we often use the abbreviated version $\{f > a\}$.

The symbol c_A denotes the *indicator function* of a set $A \subset X$, i.e. the function

$$c_A(x) = \begin{cases} 1 & \text{for } x \in A\,, \\ 0 & \text{for } x \in X \setminus A. \end{cases}$$

We use $f_j \rightrightarrows f$ to denote the uniform convergence of a sequence of functions.

A. Measures and Measurable Functions

1. The Lebesgue Measure

In the history, people were engaged in the problem of measuring lenghts, areas and volumes. In mathematical formulation the task was, for a given set A, to determine its size ("measure") λA. It was required that the volume of a cube or the area of a rectangle or a circle should agree with the well-known formulae. It was also clear by intuition that this measure should be positive and additive, i.e. it should satisfy the equality

$$\lambda \bigcup A_j = \sum \lambda A_j$$

provided $\{A_j\}$ is a finite disjoint collection of sets. For a succesful development of the theory a further condition was imposed: The above equality was claimed to hold even for countable disjoint collections of sets. Moreover, the effort was paid to assign a measure to as many sets as possible.

Now, we are going to show how to proceed on the real line. The same approach will be used later in the Euclidean space $\mathbf{R}^n$ where the proofs will be given.

1.1. Outer Lebesgue Measure. For an arbitrary set $A \subset \mathbf{R}$, define

$$\lambda^* A := \inf\{\sum_{i=1}^{\infty}(b_i - a_i) \colon \bigcup_{i=1}^{\infty}(a_i, b_i) \supset A\}.$$

The value $\lambda^* A$ (which can also be $+\infty$) is called the *outer Lebesgue measure* of a set A.

1.2. Properties of the Outer Lebesgue Measure. One can see immediately that $\lambda^* A \leq \lambda^* B$ if $A \subset B$ and that the measure of a singleton is 0, and without much effort it becomes clear that $\lambda^* I$ is the length of I in case of I interval of any type (see Exercise 1.6). Then it is relatively easy to prove that the outer Lebesgue measure is *translation invariant*: If $A \subset \mathbf{R}$ and $x \in \mathbf{R}$, then $\lambda^* A = \lambda^*(x + A)$. Another important property is the *σ-subadditivity*:

$$\lambda^*(\bigcup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} \lambda^* A_j.$$

In mathematical terminology, the prefix σ usually relates to countable unions and δ to countable intersections.

The question of whether λ^* is an additive set function has a negative answer: There are disjoint sets A, B with

$$\lambda^*(A \cup B) < \lambda^* A + \lambda^* B$$

(cf. 1.8), and we need to find a family of sets (as large as possible) on which the measure λ^* is additive. This task will be solved later in Chapter 4 in a much more general case. Now we just briefly indicate one of its possible solutions in case of the Lebesgue measure.

1.3. Lebesgue Measurable Sets. Let A be a subset of a bounded interval I. Defining the "inner measure" $\lambda_* A = \lambda^* I - \lambda^*(I - A)$, it is natural to investigate the collection of sets for which $\lambda_* A = \lambda^* A$ (cf. Exercise 1.7). This leads to the following definition.

We say that a set $A \subset \mathbf{R}$ is (Lebesgue) *measurable* if $\lambda^* I = \lambda^*(A \cap I) + \lambda^*(I \setminus A)$ for every bounded interval $I \subset \mathbf{R}$. The collection of all measurable sets on $\mathbf{R}$ will be denoted by $\mathfrak{M}$. Not every set is measurable as will be seen in 1.8. The set function $M \mapsto \lambda^* M$, $M \in \mathfrak{M}$ is denoted by λ and called the *Lebesgue measure*. Thus, on measurable sets, the set functions λ^* and λ coincide but for nonmeasurable ones only λ^* is defined.

Another important property of the measure λ is contained in the following theorem which is now presented without proof.

1.4. Theorem. *(a) If $M_1, M_2, \dots$ are elements of $\mathfrak{M}$, then also $M_1 \setminus M_2$, $\bigcap M_n$ and $\bigcup M_n$ are elements of $\mathfrak{M}$. If, in addition, the sets M_n are pairwise disjoint, then*

$$\lambda\Big(\bigcup_n M_n\Big) = \sum_n \lambda M_n .$$

(b) Intervals of any type are in $\mathfrak{M}$.

1.5. Remark. The ingenuity of Lebesgue's approach to the measure consists in considering the *countable* covers of a set A with intervals. If in the definition of $\lambda^* A$ we consider only *finite* covers, we get the notion of so-called *Jordan-Peano content*. In modern analysis this notion is far from being as important as the Lebesgue measure.

1.6. Exercise. If $I \subset \mathbf{R}$ is an interval (of any type), show that $\lambda^* I$ is its length.

Hint. It is sufficient to consider the case $I = [a, b]$. Clearly $\lambda^*[a, b] \le b - a$ (since $[a, b] \subset (a - \varepsilon, b + \varepsilon)$). Suppose $\bigcup\limits_{i=1}^{\infty} (a_i, b_i) \supset [a, b]$. A compactness argument yields the existence of an index n satisfying $\bigcup\limits_{i=1}^{n} (a_i, b_i) \supset [a, b]$. Using induction (with respect to n) it can be shown that $b - a \le \sum\limits_{i=1}^{n} (b_i - a_i)$.

1.7. Exercise. For every bounded set $A \subset \mathbf{R}$, define

$$\lambda_* A := \lambda I - \lambda^*(I \setminus A)$$

where I is a bounded interval containing A. Show that:

(a) the value of $\lambda_* A$ does not depend on the choice of I;

(b) a bounded set $A \subset \mathbf{R}$ is measurable if and only if $\lambda^* A = \lambda_* A$;

(c) a set $M \subset \mathbf{R}$ is measurable if and only if its intersection with each bounded interval is measurable.

In the next part of this chapter we introduce some significant sets on the real line.

1.8. A Nonmeasurable Set. Now we prove the existence of a nonmeasurable subset of $\mathbf{R}$ and consequently prove that the outer Lebesgue measure cannot be additive.

Set $x \sim y$ if $x - y$ is a rational number. It is easy to see that $\sim$ is an equivalence relation on $\mathbf{R}$. Therefore $\mathbf{R}$ splits into an uncountable collection $\mathscr{V}$ of pairwise disjoint classes. A set V belongs to this collection $\mathscr{V}$ if and only if $V = x + \mathbf{Q}$ for some $x \in \mathbf{R}$. By the axiom of choice, there exists a set $E \subset (0,1)$ that shares exactly one point with each set $V \in \mathscr{V}$. We show that E is not in $\mathfrak{M}$.

Let $\{q_n\}$ be a sequence containing all rational numbers from the interval $(-1,+1)$. It is not very difficult to show that the sets $E_n := q_n + E$ are pairwise disjoint and that

$$(0,1) \subset \bigcup_n E_n \subset (-1,2).$$

Assuming that $E \in \mathfrak{M}$, then also $E_n \in \mathfrak{M}$ and Theorem 1.4 gives $\lambda \bigcup_n E_n = \sum_n \lambda E_n$. Distinguishing two cases $\lambda E = 0$ and $\lambda E > 0$ we easily obtain the contradiction.

1.9. Remarks. 1. The proof of the existence of a nonmeasurable set is not a constructive one (it uses the axiom of choice for an uncountable collection of sets). We return to the topic of nonmeasurable sets in Notes 1.22.
2. By a simple argument, an even stronger proposition can be proved: Any measurable set $M \subset \mathbf{R}$ of a positive measure contains a nonmeasurable subset. It is sufficient to realize that $M = \bigcup_{q\in\mathbf{Q}} M \cap (E+q)$ where E is the nonmeasurable set from 1.8 and that any measurable subset of E is of zero (Lebesgue) measure.
3. Van Vleck [1908] "constructed" a set $E \subset [0,1]$ for which $\lambda^* E = 1$ and $\lambda_* E = 0$.

1.10. Exercise. Show that every countable set S is of measure zero.

Hint. Consider covers $\bigcup_{j=1}^{\infty} (r_j - \varepsilon 2^{-j}, r_j + \varepsilon 2^{-j})$ where $\{r_j\}$ is a sequence of all elements of the set S. The assertion also follows from Theorem 1.4 if you realize that singletons have measure zero.

1.11. Examples of Sets of Measure Zero. (a) The set $\mathbf{Q}$ of all rational numbers is countable, thus by Exercise 1.10 it has Lebesgue measure zero.

(b) It can be seen from the hint to the exercise that for every $k \in \mathbf{N}$ there is an open set G_k such that $\mathbf{Q} \subset G_k$ and $\lambda^* G_k \le 1/k$. The set $\bigcap_{k=1}^{\infty} G_k$ has also Lebesgue measure zero, it is dense and uncountable (even residual).

1.12. Cantor Ternary Set. Consider the sequence $\{\mathscr{K}_n\}$ of finite collections of intervals defined in the following way: $\mathscr{K}_0 = \{[0,1]\}$, $\mathscr{K}_1 = \{[0,\frac{1}{3}],[\frac{2}{3},1]\}$. In each step we construct $\mathscr{K}_n$ from $\mathscr{K}_{n-1}$ as the collection of all closed intervals which are the left or right third of an interval from the collection $\mathscr{K}_{n-1}$ (the middle thirds are omitted). Then $\mathscr{K}_n$ is a collection of 2^n disjoint closed intervals, each of them of length 3^{-n}. Let K_n denote the union of the collection $\mathscr{K}_n$. The *Cantor ternary set*[1] C is defined as $\bigcap_n K_n$. It is not difficult to verify that C consists precisely of points of the form $\sum_{i=1}^{\infty} a_i 3^{-i}$ where each a_i is 0 or 2. Roughly speaking, in the Cantor set there are exactly those points of the interval $[0,1]$ whose ternary expansions do not contain the digit 1. The Cantor set has the following properties:

(a) C is a compact set without isolated points;

(b) C is a nowhere dense (and totally disconnected) set;

(c) C is an uncountable set;

(d) the Lebesgue measure of C is zero.

1.13 Discontinua of a Positive Measure. If we construct a set $D \subset [0,1]$ like the Cantor set except that we always omit intervals of length $\varepsilon 3^{-n}$ where $\varepsilon \in (0,1)$ (note that their centres are not the same as those in the construction of the Cantor set), we get a closed nowhere dense set, for which $\lambda D = 1-\varepsilon$. Sets having this property are called the *discontinua of a positive measure*. Another construction: If G is an open subset of the interval $(0,1)$, containing all rational points of this interval and $\lambda G = \varepsilon < 1$ then $[0,1]\setminus G$ is a discontinuum of measure $1-\varepsilon$.

[1] sometimes also called the *Cantor discontinuum*

1.14. Exercise. Prove that there exists a non-Borel subset of the Cantor set and realize that this set is Lebesgue measurable.

Hint. The cardinality argument shows that the set of all Borel subsets of the Cantor set has cardinality of the continuum while the set of all its subsets has greater cardinality.

Instead of this, the following idea can be used. Define

$$\kappa(t) := \inf\{x \in [0,1] : f(x) = t\}$$

where f is the Cantor singular function from 23.1. Show that κ is increasing on the interval $[0,1]$, and therefore it is a Borel function. Suppose E is a nonmeasurable subset of $[0,1]$, $B := \kappa(E)$. Then B (as a subset of the Cantor set) is a measurable set. But since $\kappa^{-1}(B) = E$ (and κ is a Borel function), B cannot be a Borel set.

1.15. Lebesgue Measure on $\mathbf{R}^n$. In the same way as for $\mathbf{R}$, we introduce the Lebesgue measure on $\mathbf{R}^n$. Recall that by an interval in $\mathbf{R}^n$ we understand an arbitrary Cartesian product of n one-dimensional intervals. If $I := (a_1, b_1) \times \cdots \times (a_n, b_n)$ is an open interval, we define its *volume* as

$$\operatorname{vol} I = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n).$$

In the same way we define $\operatorname{vol} I$ for intervals of other types. Given an arbitrary set $A \subset \mathbf{R}^n$, define the *outer Lebesgue measure* of A as the quantity

$$\lambda^* A = \inf\{\sum_{k=1}^{\infty} \operatorname{vol} I_k \colon \bigcup_{k=1}^{\infty} I_k \supset A,\ I_k \text{ is an open interval}\}.$$

We say that a set $A \subset \mathbf{R}^n$ is *measurable* if $\lambda^* T = \lambda^*(A \cap T) + \lambda^*(A \setminus T)$ for every set $T \subset \mathbf{R}^n$. (By analogy with the one-dimensional case we should require this equality to hold just for bounded intervals T. We have chosen the present definition in order to apply the general approach of Chapter 4. Soon we show that there is no difference between these two definitions.) The symbol $\mathfrak{M}$ again denotes the collection of all measurable subsets of $\mathbf{R}^n$. For $M \in \mathfrak{M}$ we denote by $\lambda M := \lambda^* M$ the n-dimensional *Lebesgue measure* of a set M.

1.16. Theorem. *If $\{A_j\}$ is a sequence of (arbitrary) sets of $\mathbf{R}^n$, then*

$$\lambda^*\Big(\bigcup_{j=1}^{\infty} A_j\Big) \le \sum_{j=1}^{\infty} \lambda^* A_j.$$

Proof. The assertion follows from Theorem 4.3. ■

1.17. Theorem. *If M_1, $M_2, \ldots$ are elements of $\mathfrak{M}$, then also $M_1 \setminus M_2$, $\bigcap M_n$ and $\bigcup M_n$ are elements of $\mathfrak{M}$. If, in addition, the sets M_n are pairwise disjoint, then*

$$\lambda\Big(\bigcup_n M_n\Big) = \sum_n \lambda M_n.$$

Proof. The assertion follows from general Theorem 4.5. ■

Compare the following theorem with Exercise 1.6.

1.18. Theorem. *If $I \subset \mathbf{R}^n$ is a bounded interval, $I \subset \bigcup_j Q_j$ where $\{Q_j\}$ is a sequence of open intervals, then*

$$\operatorname{vol} I \le \sum_j \operatorname{vol} Q_j .$$

Thus the n-dimensional Lebesgue measure $\lambda^* I$ is equal to the volume $\operatorname{vol} I$.

Proof. Suppose J is a compact interval contained in I. There exists a p such that the intervals $\{Q_1, \dots, Q_p\}$ cover J. The interval J can be now divided into a finite number of non-overlapping n-dimensional intervals $\{J_i\}$ (distinct elements of $\{J_i\}$ have disjoint interiors) in such a way that the interior of each interval J_i is contained in some of the intervals Q_j. Then

$$\operatorname{vol} J = \sum_i \operatorname{vol} J_i \le \sum_{j=1}^{p} \operatorname{vol} Q_j \le \sum_{j=1}^{\infty} \operatorname{vol} Q_j .$$

Since the difference $\operatorname{vol} I - \operatorname{vol} J$ can be arbitrarily small, the assertion follows. ∎

1.19. Theorem. (a) *Any open subset of $\mathbf{R}^n$ is measurable.*
(b) *If $\lambda^* A = 0$, then A is measurable.*

Proof. The proof of part (b) is obvious; we will prove (a). First we prove that each interval H which is a halfspace (e.g. of the form $(-\infty, c) \times \mathbf{R}^{n-1}$) is measurable. Choose a "test" set T, $\lambda^* T < \infty$, and $\varepsilon > 0$. There exist open intervals $\{Q_j\}$ with

$$\bigcup_j Q_j \supset T \text{ and } \sum_j \operatorname{vol} I_j < \lambda^* T + \varepsilon .$$

Now find open intervals I_j and J_j such that

$$I_j \cup J_j = Q_j, \quad Q_j \cap H \subset I_j,\ Q_j \setminus H \subset J_j \text{ and } \lambda^* I_j + \lambda^* J_j < \lambda^* Q_j + \varepsilon 2^{-j} .$$

Then

$$\lambda^*(T \cap I) + \lambda^*(T \setminus I) \le \sum_j \operatorname{vol} I_j + \sum_j \operatorname{vol} J_j \le \lambda^* T + \varepsilon .$$

We proved the measurability of all intervals H of the form of a halfspace. Now, each open set can be expressed as a countable union of intervals and each interval is a finite intersection of intervals which are halfspaces. ∎

1.20. Theorem. *If $A \subset \mathbf{R}^n$, then*

$$\lambda^* A = \inf \{\lambda G \colon G \text{ open}, \ G \supset A\} .$$

Proof. One inequality follows from the monotonicity of λ^*. Now if $\lambda^* A < \infty$ and $\varepsilon > 0$, then there exist open intervals $I_j \subset \mathbf{R}^n$ such that

$$A \subset \bigcup_j I_j \quad \text{and} \quad \lambda \bigcup_j I_j \le \sum_j \operatorname{vol} I_j < \lambda^* A + \varepsilon .$$

∎

The reader should compare the following theorem and Exercise 15.19.

1.21. Theorem. *Given a set $M \subset \mathbf{R}^n$, the following are equivalent:*

(i) *M is measurable;*
(ii) *for every bounded interval I, $\lambda^* I = \lambda^*(I \cap M) + \lambda^*(I \setminus M)$;*
(iii) *for every $\varepsilon > 0$ there exists an open set $G \supset M$ with $\lambda^*(G \setminus M) < \varepsilon$;*
(iv) *there exists a G_δ-set $D \supset M$ such that $\lambda^*(D \setminus M) = 0$;*
(v) *there exist an F_σ-set B_i and a G_δ-set B_e such that $B_i \subset M \subset B_e$ and $\lambda^*(B_e \setminus B_i) = 0$.*

Proof. The implication (i) $\Longrightarrow$ (ii) is trivial. Assuming (ii), fix $\varepsilon > 0$ and denote $I_k = (-k,k)^n$. By Theorem 1.20 we can find open sets G_k and H_k such that $I_k \cap M \subset G_k$, $I_k \setminus M \subset H_k$, $\lambda G_k \le \lambda^*(I_k \cap M) + 2^{-k}\varepsilon$ and $\lambda H_k \le \lambda^*(I_k \setminus M) + 2^{-k}\varepsilon$. We can assume that G_k and H_k are subsets of I_k. Then we have $G_k \setminus M \subset G_k \cap H_k$. Using (ii) and the measurability of open sets we obtain

$$\lambda I_k + \lambda(G_k \cap H_k) = \lambda G_k + \lambda H_k \le \lambda^*(I_k \cap M) + \lambda^*(I_k \setminus M) + 2^{-k+1}\varepsilon \le \lambda I_k + 2^{-k+1}\varepsilon.$$

Set $G = \bigcup_k G_k$. Then

$$\lambda^*(G \setminus M) \le \sum_{k=1}^{\infty} \lambda(G_k \cap H_k) \le 2\varepsilon$$

so that (iii) holds. That (iii) implies (iv) is evident. It is not very difficult to prove the implication (iv) $\Longrightarrow$ (v). If M satisfies (v), then $M = B_i \cup (M \setminus B_i)$ where the sets B_i and $M \setminus B_i$ are measurable by Theorem 1.19 (each one for a different reason), so that (v) $\Longrightarrow$ (i). ■

1.22. Notes. Originally, H. Lebesgue defined the outer measure on the real line using countable covers formed by intervals, exactly as explained in the text. He defined measurability as in Exercise 1.7.

At the end of the last century, various attempts to define the length or area of geometrical figures appear; in the works of G. Peano [1887] and C. Jordan [1892] even the "measures" of more complicated sets are considered.

The existence of a Lebesgue nonmeasurable set is very closely connected to the axiom of choice (for uncountable collections of sets) and the assertion that such sets exist was first proved by G. Vitali [*1905]. Solovay's result [1970] says that there exist models of the set theory (of course not satisfying the axiom of choice) in which every subset of real numbers is Lebesgue measurable. The existence of a nonmeasurable set can be proved (assuming various set conditions) in other ways as well. Constructions of Bernstein's sets (still assuming the axiom of choice) as examples of nonmeasurable sets are also interesting. Another construction of a nonmeasurable set (the axiom of choice again) based on results of the graph theory comes from R. Thomas [1985]. Using nonstandard methods, it is possible to prove the existence of a nonmeasurable set assuming the existence of ultrafilters (a weaker form of the axiom of choice; cf. M. Davis [*1977]). Recently, M. Foreman and F. Wehrung [1991] proved that the existence of a nonmeasurable set follows from the Hahn-Banach Theorem (which is again a weaker assumption than the axiom of choice).

Let us note that the Lebesgue measure can be extended to a "translation invariant" measure defined on a wider σ-algebra than is the collection of all Lebesgue measurable sets. The construction can be found e.g. in S. Kakutani and J.C. Oxtoby [1950]. However, the Lebesgue measure cannot be extended in a reasonable way to the collection of all subsets of $\mathbf{R}^n$.

It is interesting that in $\mathbf{R}$ or $\mathbf{R}^2$ there exist finitely additive extensions of the Lebesgue measure to the collection of all subsets which can also be invariant with respect to translations and rotations. This was first proved by S. Banach [1923]. However, this result cannot be transferred to spaces of higher dimensions as follows from the famous result of S. Banach and A. Tarski [1924]:

If U and V are arbitrary (!) bounded and open sets in the space $\mathbf{R}^n$, $n \geq 3$, then there exist sets $E_1, \dots, E_k$ and $F_1, \dots, F_k$ such that $E_i \cap E_j = \emptyset = F_i \cap F_j$ for $i \neq j$, $U = \bigcup E_i$, $V = \bigcup F_i$ and E_j are isometric copies of F_j.

In this theorem, which is known as the Banach-Tarski paradox, in general all the sets E_i and F_i cannot be measurable; realize that U, V can be of different measures. More information is contained in S. Wagon [*1985].

2. Abstract Measures

In this chapter we study an abstract notion of measure which stands as a basis for modern integration theory. Also in probability theory, the notion of measure (termed a "probability" there) plays a crucial role. Among many fields of analysis which employ measures in an essential way, let us mention e.g. functional analysis, theory of function spaces and theory of distributions, or mathematical modelling of physical quantities.

So far we have the only nontrivial example of the Lebesgue measure. Further important examples of measures will be introduced later.

Remember that the Lebesgue measure is not defined on the collection of all subsets of $\mathbf{R}$ but on its subcollection which is closed under countable operations. We start with the following definition.

2.1. σ-algebras. A collection $\mathscr{S}$ of subsets of a given set X is called a *σ-algebra* if

(a) $X \in \mathscr{S}$;
(b) if $A \in \mathscr{S}$, then $X \setminus A \in \mathscr{S}$;
(c) if $A_n \in \mathscr{S}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathscr{S}$.

The pair $(X, \mathscr{S})$ is called a *measurable space*.

Clearly every σ-algebra is also closed under countable intersections, under differences and contains the empty set.

Not every collection of sets is a σ-algebra. However, if $\mathscr{T}$ is an arbitrary family of subsets of X, then there exists the *smallest* σ-algebra $\sigma(\mathscr{T})$ which contains $\mathscr{T}$. Such a σ-algebra is simply the intersection of all σ-algebras (in X) which contain $\mathscr{T}$. It surely exists, since there is at least one such a σ-algebra (the σ-algebra $\mathscr{P}(X)$ of *all* subsets of X) and the intersection of any collection of σ-algebras is again a σ-algebra. The collection $\sigma(\mathscr{T})$ is called the *σ-algebra generated by $\mathscr{T}$*.

2.2. Examples. One of the most important examples is the σ-algebra $\mathfrak{M}$ of all Lebesgue measurable sets on the real line. Another important class of examples yields Borel σ-algebras from 2.3. For illustration, we add a few simple examples. Suppose X is an arbitrary set. Then

(a) $\{\emptyset, X\}$ is a σ-algebra;

(b) the collection $\mathscr{P}(X)$ of all subsets of X is a σ-algebra;

(c) $\{A \subset X : A \text{ is countable or } X \setminus A \text{ is countable}\}$ is a σ-algebra.

2.3. Borel Sets. Let P be a topological space. The σ-algebra $\mathscr{B}(P)$ generated by the family of all open subsets of P is called the *Borel σ-algebra of P*; its elements are called *Borel sets*. The Borel σ-algebra $\mathscr{B}(P)$ contains all closed sets, all countable intersections of open sets (these sets are called G_δ *sets*), all countable unions of closed sets (F_σ *sets*), all countable unions of G_δ sets ($G_{\delta\sigma}$ *sets*), all countable intersections of F_σ sets ($F_{\sigma\delta}$) and so on. Let us note that for the complete description of all possible "types" we would have to use (in nontrivial cases) all countable ordinal numbers.

2.4. Measures. Let $\mathscr{S}$ be a collection of subsets of a set X. A nonnegative set function $\mu : \mathscr{S} \to [0, \infty]$ is called a *measure* if

(a) $\mathscr{S}$ is a σ-algebra;
(b) $\mu\emptyset = 0$;
(c) for each sequence $\{A_n\}$ of pairwise disjoint sets from $\mathscr{S}$, $\mu(\bigcup A_n) = \sum \mu A_n$.

The triplet $(X, \mathscr{S}, \mu)$ is termed a *measure space*.

From (b) and (c) it immediately follows that each measure is *monotone* (if $A, B \in \mathscr{S}$, $A \subset B$, then $\mu A \le \mu B$); the property (c) is also called the σ*-additivity* of a measure.

We say that a measure is *finite* if $\mu X < +\infty$, σ*-finite* if there exist sets $M_n \in \mathscr{S}$ such that $\mu M_n < +\infty$ and $X = \bigcup_{n=1}^{\infty} M_n$. If $\mu X = 1$, then we say that μ is a *probability measure*. A measure μ is said to be *complete* if whenewer $B \in \mathscr{S}$ is a null set and $A \subset B$, then also $A \in \mathscr{S}$ (and $\mu A = 0$).

If μ is a measure on X and $E \in \mathscr{S}$, we define $\mu_E\, A = \mu(A \cap E)$ for $A \in \mathscr{S}$. A related notion is the *restriction* $\mu|_A$ of a measure μ to the set $A \in \mathscr{S}$: If we denote by $\mathscr{S}_A$ the σ-algebra $\{M \in \mathscr{S} \colon M \subset A\}$ of subsets of A, we define $\mu|_A(M) = \mu(M)$ for $M \in \mathscr{S}_A$. Finally, if $\mathscr{T} \subset \mathscr{S}$ is a σ-algebra of subsets of X, then the symbol $\mu|_{\mathscr{T}}$ denotes the measure $E \mapsto \mu E,\ E \in \mathscr{T}$.

2.5. Examples. (a) The *Lebesgue measure* on the collection of all Lebesgue measurable sets in $\mathbf{R}^n$. This measure is complete, σ-finite, but not finite.

(b) *Counting measure*. Let X be an arbitrary set, $\mathscr{S} = \mathscr{P}(X)$ the collection of all subsets of X and

$$\mu A = \begin{cases} \text{the number of elements of } A \text{ if } A \text{ is finite,} \\ +\infty \text{ if } A \text{ is infinite.} \end{cases}$$

The counting measure is complete; it is σ-finite if and only if X is countable, and finite just when X is finite.

(c) The *Dirac measure*. Again, let X be an arbitrary set, $x \in X$, $\mathscr{S} = \mathscr{P}(X)$. We define

$$\mu A = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \in X \setminus A. \end{cases}$$

The measure μ is called the Dirac measure at x and it is denoted by ε_x. The Dirac measure is a complete probability measure.

(d) *Trivial measures.* If $\mathscr{S}$ is an arbitrary σ-algebra of subsets of X and $\mu A = 0$ for all $A \in \mathscr{S}$, then μ is an example of a finite measure which is not complete provided $\mathscr{S} \ne \mathscr{P}(X)$.

2.6. Properties of Measures. *Let $(X, \mathscr{S}, \mu)$ be a measure space. Then the the following propositions hold:*

(a) *if $A_1,\ A_2, \dots \in \mathscr{S},\ A_1 \subset A_2 \subset \dots$, then $\mu\left(\bigcup A_n\right) = \lim \mu A_n$;*

(b) *if* $A_1, A_2, \dots \in \mathscr{S}$, $A_1 \supset A_2 \supset \dots$, *and* $\mu A_1 < \infty$, *then* $\mu\left(\bigcap A_n\right) = \lim \mu A_n$;

(c) *if* $A_1, A_2, \dots \in \mathscr{S}$, *then* $\mu\left(\bigcup A_n\right) \le \sum \mu A_n$.

Proof. (a) Since $A_1,\ A_2 \setminus A_1,\ A_3 \setminus A_2, \dots$ are pairwise disjoint, we get

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(A_1 \cup \bigcup_{n=1}^{\infty}(A_{n+1} \setminus A_n)\right) = \mu A_1 + \sum_{n=1}^{\infty} \mu(A_{n+1} \setminus A_n)$$
$$= \lim_k \left(\mu A_1 + \sum_{n=1}^{k-1}(\mu A_{n+1} \setminus A_n)\right) = \lim_{k\to\infty} \mu A_k.$$

(b) By (a), we obtain

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \mu A_1 - \mu\left(A_1 \setminus \bigcap_{n=1}^{\infty} A_n\right) = \mu A_1 - \mu\left(\bigcup_{n=1}^{\infty} A_1 \setminus A_n\right)$$
$$= \mu A_1 - \lim_{n\to\infty} \mu(A_1 \setminus A_n) = \lim_{n\to\infty} \mu A_n.$$

(c) It is sufficient to consider the sequence

$$B_1 = A_1,\ B_2 = A_2 \setminus A_1,\ B_3 = A_3 \setminus (A_1 \cup A_2), \dots,$$

notice that $\bigcup B_n = \bigcup A_n$ and that the sequence $\{B_n\}$ is pairwise disjoint. ■

2.7. Completion of Measures. Now we return to the notion of completeness of a measure. We show that every measure space can be "extended" to a complete measure space.

Let $(X, \mathscr{S}, \mu)$ be a measure space and let $\mathscr{N}$ denote the collection of all sets $A \subset X$ for which there is a $B \in \mathscr{S}$ such that $\mu B = 0$ and $A \subset B$. Further, let $\overline{\mathscr{S}}$ denote the collection of all sets of the form $M \cup N$ where $M \in \mathscr{S}$ and $N \in \mathscr{N}$. We define a set function $\overline{\mu}$ on $\overline{\mathscr{S}}$ by

$$\overline{\mu}(M \cup N) = \mu M, \qquad M \in \mathscr{S},\ N \in \mathscr{N}.$$

Clearly the value of $\overline{\mu}E$ does not depend on the choice of M and N. The set function $\overline{\mu}$ on $\overline{\mathscr{S}}$ is called the *completion* of μ.

2.8. Theorem. *$\overline{\mathscr{S}}$ is a σ-algebra containing $\mathscr{S}$ and $\overline{\mu}$ is a complete measure on $\overline{\mathscr{S}}$ which coincides with μ on $\mathscr{S}$.*

Proof. Obviously, $\mathscr{S} \subset \overline{\mathscr{S}}$. Since both $\mathscr{S}$ and $\mathscr{N}$ are closed under countable unions, the same is true for $\overline{\mathscr{S}}$. If $E \in \overline{\mathscr{S}}$, there are $M \in \mathscr{S}$, $N \in \mathscr{N}$ and $B \in \mathscr{S}$ so that $N \subset B$, $\mu B = 0$ and $E = M \cup N$. Then $X \setminus E = (X \setminus (M \cup B)) \cup (B \setminus N) \in \overline{\mathscr{S}}$. It is easy to verify that $\overline{\mu}$ is a complete measure on $\overline{\mathscr{S}}$ and $\mu = \overline{\mu}$ on $\mathscr{S}$. ■

2.9. Remark. Note that there is a variety of extensions of a measure to a complete measure. The completion described above is uniquely determined and in some sense minimal (cf. Exercise 2.13).

2.10. Exercise. Let f be a nonnegative function on a set X and $\mathscr{S}$ a σ-algebra on X. For $B \in \mathscr{S}$ we define

$$\mu B := \sum_{x \in B} f(x).$$

(Recall that by definition

$$\sum_{x\in B} f(x) = \sup_K \left\{ \sum_{x \in K} f(x) \colon K \subset B,\ K \text{ finite} \right\}.)$$

Show that μ is a measure on $\mathscr{S}$ (so-called *weighted counting measure*).

(a) In a particular case $\mathscr{S} = \mathscr{P}(X)$ and $f = 1$ on X we get the counting measure.

(b) If $\mathscr{S} = \mathscr{P}(X)$, $z \in X$ and $f = c_{\{z\}}$, then μ is the Dirac measure at z.

2.11. Exercise. (a) The trivial measure on $\mathscr{S} = \{\emptyset, \mathbf{R}\}$ is not complete. What is its completion?

(b) The Lebesgue measure on $\mathbf{R}$ considered on the σ-algebra of Borel sets (see Exercise 1.14.a) is not complete.

(c) Suppose $X = \{1, 2, 3, 4\}$, $\mathscr{S} = \{\emptyset, \{1, 2\}, \{3, 4\}, X\}$. Let ω be a measure on $\mathscr{P}(X)$ such that

$$\omega\{1\} = \omega\{2\} = 0,\ \omega\{3\} = \omega\{4\} = 1.$$

If $\mu = \omega|_{\mathscr{S}}$, then μ is not complete. If

$$\mathscr{T} = \mathscr{S} \cup \{\{1\}, \{2\}, \{1, 3, 4\}, \{2, 3, 4\}\}, \quad \nu = \omega|_{\mathscr{T}},$$

then ν is complete. What is the completion of μ?

2.12. Exercise. Prove that the completion of the Lebesgue measure considered on the σ-algebra of all Borel sets in $\mathbf{R}$ is the Lebesgue measure (on the collection of all Lebesgue measurable sets; cf. Theorem 26.1).

2.13. Exercise. Suppose $(X, \mathscr{T}, \nu)$ is the completion of $(X, \mathscr{S}, \mu)$. If μ_1 is a complete measure on a σ-algebra $\mathscr{S}_1$ such that $\mathscr{S} \subset \mathscr{S}_1$ and $\mu = \mu_1|_{\mathscr{S}}$, show that $\mathscr{T} \subset \mathscr{S}_1$ and $\nu = \mu_1|_{\mathscr{T}}$.

2.14. Exercise (Borel-Cantelli lemma). Let $(X, \mathscr{S}, \mu)$ be a measure space and $A_n \in \mathscr{S}$. If $\sum\limits_{n=1}^{\infty} \mu A_n < +\infty$, then $\mu(\limsup A_n) = 0$ (we define $\limsup A_n := \bigcap\limits_{n=1}^{\infty} \bigcup\limits_{k=n}^{\infty} A_k$).

2.15. Exercise (Darboux property). Let $(X, \mathscr{S}, \mu)$ be a measure space. If μ does not have an atom, show that $\{\mu A \colon A \in \mathscr{S}\} = [0, \mu X]$. (A set $A \in \mathscr{S}$ is called an *atom* if $\mu A > 0$ and for each set $B \in \mathscr{S}$, $B \subset A$, either $\mu B = 0$ or $\mu(A \setminus B) = 0$.)

Hint. Fix $0 < \alpha < \mu X$ and set $\mathscr{A} = \{A \in \mathscr{S} \colon \mu A \geq \alpha\}$. There exists a set $C \in \mathscr{A}$ such that $\mu A = \mu C$ for all $A \in \mathscr{A}$, $A \subset C$. In a similar way find $D \in \mathscr{S}$, $D \subset C$, $\mu D \leq \alpha$ with the following property: $\mu B = \mu D$ for each set $B \in \mathscr{S}$ for which $D \subset B \subset C$ and $\mu B \leq \alpha$. Because $C \setminus D$ cannot be an atom, it follows $\mu C = \mu D = \alpha$.

2.16. Notes. In [1895] and [*1898], E. Borel extended the length of intervals to a set function (measure) defined on the collection of all Borel sets. However, H. Lebesgue was the first who created the theory of the integral with the help of this measure. J. Radon in [1913] defined a general notion of (Borel) measures in Euclidean spaces. A.N. Kolmogorov introduced in [*1933] the axiomatic theory of probability measures.

The Darboux property of real-valued measures is a special version of the general *Lyapunov theorem* (A. Lyapunov [1940]) which says that the range of a finite-dimensional nonatomic vector measure is a convex compact set.

3. Measurable Functions

As mentioned in the first chapter, there are serious reasons why the Lebesgue measure is not defined on the collection all subsets of $\mathbf{R}^n$. Likewise, one cannot expect a reasonable theory of integration on the class of all functions; it will be necessary to confine to some reasonable family of functions. This class of course should contain the indicator functions of all measurable sets and should be closed under all common (algebraic as well as limit) operations. As we will see, these requirements are satisfied by the following natural definition.

In what follows, we assume that $\mathscr{S}$ is a σ-algebra of subsets of a given set X.

3.1. Measurable Functions. Suppose $D \in \mathscr{S}$. A function $f : D \to \overline{\mathbf{R}}$ is said to be $\mathscr{S}$-measurable on D if $\{x \in D\colon f(x) > \alpha\} \in \mathscr{S}$ for each $\alpha \in \mathbf{R}$.

A complex function on D is $\mathscr{S}$-measurable if its real and imaginary parts are $\mathscr{S}$-measurable.

3.2. Examples. (a) If $\mathscr{S}$ is a σ-algebra of *all* subsets of X, then every function on X is $\mathscr{S}$-measurable.

(b) Only the constant functions are $\mathscr{S}$-measurable if $\mathscr{S} = \{\emptyset, X\}$.

(c) If $\mathscr{B}$ is the σ-algebra of all Borel sets of a topological space X, then $\mathscr{B}$-measurable functions are called shortly *Borel functions* on X.

3.3. Remark. Since any σ-algebra is closed for complementation, a function f is $\mathscr{S}$-measurable if and only if $\{f \le \alpha\} \in \mathscr{S}$ for each $\alpha \in \mathbf{R}$. Because

$$\{f \ge \alpha\} = \bigcap_{n=1}^{\infty} \left\{f > \alpha - \frac{1}{n}\right\},$$

it is possible to replace the condition $\{f > \alpha\} \in \mathscr{S}$ by $\{f \ge \alpha\} \in \mathscr{S}$) in the definition of $\mathscr{S}$-measurability. Other equivalent conditions are stated in Exercise 3.6.

3.4. Theorem. *Let f, g be $\mathscr{S}$-measurable functions on X. Then*

(a) $\{x \in X\colon f(x) < g(x)\} \in \mathscr{S}$;
(b) $f^{-1}(+\infty)$, $f^{-1}(-\infty) \in \mathscr{S}$.

Proof. Since

$$\{f < g\} = \bigcup_{r \in \mathbf{Q}} (\{f < r\} \cap \{g > r\}),$$

$$\{x \in X : f(x) = +\infty\} = \bigcap_{n=1}^{\infty} \{f > n\},$$

the assertion easily follows. ■

3.5. Properties of $\mathscr{S}$-measurable Functions. *Let f, g, f_n be $\mathscr{S}$-measurable functions (with possibly different domains in $\mathscr{S}$), $\lambda \in \mathbf{R}$ and let φ be a continuous function on an open set $G \subset \mathbf{R}$. Then the following functions are $\mathscr{S}$-measurable where defined (and their definition domains are in $\mathscr{S}$):*

(a) λf, $f + g$, $\max(f, g)$, $\min(f, g)$, $|f|$, fg, f/g;
(b) $\sup f_n$, $\inf f_n$, $\limsup f_n$, $\liminf f_n$ *and* $\lim f_n$;
(c) $\varphi \circ f$.

Proof. We will just indicate the ideas of some of the proofs leaving the others as an exercise for the reader. The assertion (c) is obvious.

(a) Suppose $\alpha \in \mathbf{R}$. Then $\{f + g > \alpha\} = \{f > \alpha - g\}$ (and the function $\alpha - g$ is $\mathscr{S}$-measurable). The functions $|f|$, f^2 and $1/g$ are $\mathscr{S}$-measurable by (c). Then we can use the formulae

$$\max(f, g) = \frac{1}{2}(f + g + |f - g|),$$
$$fg = \frac{1}{2}((f + g)^2 - f^2 - g^2).$$

(b) As a hint let us just note that

$$\{\sup f_n > \alpha\} = \bigcup_n \{f_n > \alpha\}.$$

■

3.6. Exercise. (a) Show that a real-valued function f is $\mathscr{S}$-measurable if and only if $f^{-1}(G) \in \mathscr{S}$ for each open set $G \subset \mathbf{R}$ or, if and only if $f^{-1}(B) \in \mathscr{S}$ for each Borel set $B \subset \mathbf{R}$. *Hint*. Since each open set $G \subset \mathbf{R}$ can be expressed in the form $G = \bigcup(a_n, b_n)$ (where $(a, b) = (-\infty, b) \cap (a, +\infty)$), we can see that a function f is $\mathscr{S}$-measurable if and only if $f^{-1}(G) \in \mathscr{S}$ for each open set $G \subset \mathbf{R}$ (note that $\{f > \alpha\} = f^{-1}((\alpha, +\infty))$!). Then consider the collection

$$\{B \subset \mathbf{R} \colon B \text{ Borel}, \ f^{-1}(B) \in \mathscr{S}\},$$

and show that it is a σ-algebra containing all open sets.

(b) The characterization given in (a) cannot be used for functions having infinite values unless we introduce the notions of open or Borel subsets of $\overline{\mathbf{R}}$.

3.7. Exercise. Let $\{f_n\}$ be a sequence of $\mathscr{S}$-measurable functions. Show that the sets

$$\{x \in X \colon \lim f_n(x) \text{ exists }\} \text{ and } \{x \in X \colon \lim f_n(x) \text{ exists and is finite }\}$$

are in $\mathscr{S}$.

3.8. Simple Functions. By a *simple function* on X we understand a (finite) linear combination of indicator functions of sets from $\mathscr{S}$. In other words, a real- (or complex-)valued function f is simple if f is $\mathscr{S}$-measurable and $f(X)$ is a finite set. Thus every simple function is of the form $\sum_{i=1}^{n} \lambda_i c_{A_i}$ where λ_i are numbers and $A_i \in \mathscr{S}$. Note that this expression is not uniquely determined!

3.9. Theorem. *Let $f \geq 0$ be an $\mathscr{S}$-measurable function on X. Then there exists a sequence $\{f_n\}$ of nonnegative simple functions on X such that $f_n \nearrow f$.*

Proof. For $n \in \mathbf{N}$ and $k = 1, 2, \ldots, n2^n$ set

$$F_{n,k} = \left\{x \in X \colon \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}\right\}$$

and define

$$f_n(x) = \begin{cases} \frac{k-1}{2^n}, & \text{if } x \in F_{n,k}, \\ n & \text{if } x \in X \setminus \bigcup_k F_{n,k}. \end{cases}$$

It is easily seen that f_n are simple and $f_n \nearrow f$. ■

3.10. Exercise. Suppose f, f_n have the same meaning as in the previous theorem. Show that $f_n \rightrightarrows f$ on every set on which f is bounded.

3.11. Exercise. If f is an $\mathscr{S}$-measurable function, then there exists a sequence of simple functions $\{f_n\}$ such that $|f_n| \le |f|$ and $f_n \to f$ on X.

Most frequently we meet the concept of measurability when additionally a measure is in consideration on a given σ-algebra. So suppose in what follows that $(X, \mathscr{S}, \mu)$ is a measure space. The following notion is of great importance in Lebesgue's integration theory, since usually the null sets are negligible.

3.12. Almost Everywhere. We say that a function h is defined *μ-almost everywhere* (briefly *almost everywhere*) on X if its domain $D \in \mathscr{S}$ satisfies $\mu(X \setminus D) = 0$. Suppose f and g are functions defined almost everywhere on X. We say that $f(x) \le g(x)$ *for μ-almost all $x \in X$*, or that $f \le g$ *μ-almost everywhere* if there exists a set N such that $\mu N = 0$ and for all $x \in X \setminus N$ we have $f(x) \le g(x)$. Similarly we understand the expressions "almost everywhere" and "almost all" in other contexts, e.g. when speaking about the equality of functions or about the convergence almost everywhere.

3.13. μ-measurable functions. We say that a function f defined on $D \in \mathscr{S}$ is *μ-measurable* on X if $\mu(X \setminus D) = 0$ and f is $\mathscr{S}_D$-measurable on D.

Let us emphasize that we distinguish strictly between "μ-measurable functions" (defined in general only almost everywhere) and "$\mathscr{S}$-measurable functions on X" (defined everywhere on X).

3.14. Equality almost everywhere. The relation "$f = g$ almost everywhere" is clearly an equivalence relation on the set of all functions (or all μ-measurable functions) on X. The following observations are quite useful.

(a) Let μ be a measure on $(X, \mathscr{S})$ and f be a μ-measurable function defined on $D \in \mathscr{S}$. Then there exists an $\mathscr{S}$-measurable function g on X such that $f = g$ on D; in particular, we can take

$$g = \begin{cases} f & \text{on } D, \\ 0 & \text{on } X \setminus D. \end{cases}$$

(b) If μ is a complete measure on $(X, \mathscr{S})$ and f is a μ-measurable function, then every function g which is equal to f almost everywhere is μ-measurable.

3.15. Exercise. Find an example of a measure space on which there exists an $\mathscr{S}$-measurable function f and an $\mathscr{S}$-nonmeasurable function g such that $f = g$ almost everywhere.

3.16. Exercises. Suppose $(X, \overline{\mathscr{S}}, \overline{\mu})$ is the completion of a measure space $(X, \mathscr{S}, \mu)$.

(a) Show that an everywhere defined function f is $\overline{\mathscr{S}}$-measurable if and only if there exist $\mathscr{S}$-measurable functions g, h such that $g \le f \le h$ on X and $g = h$ μ-almost everywhere in X.

Hint. The proof of one implication is easy. Let f be $\overline{\mathscr{S}}$-measurable. By Exercise 3.11 find a sequence of simple ($\overline{\mathscr{S}}$-measurable) functions $\{f_n\}$ such that $f_n \to f$ on X. Then find $\mathscr{S}$-measurable functions g_n, h_n such that $g_n \le f_n \le h_n$ and $g_n = h_n$ μ-almost everywhere and set

$$g := \limsup g_n, \quad h := \liminf h_n.$$

(b) If f is an $\overline{\mathscr{S}}$-measurable function on X, then there exists an $\mathscr{S}$-measurable function g such that $f = g$ $\overline{\mu}$-almost everywhere.

Hint. The assertion is obvious for indicator functions of sets from $\overline{\mathscr{S}}$, and therefore also for $\overline{\mathscr{S}}$-measurable simple functions. Then use Exercise 3.11.

3.17. Exercise. Let $\{f_n\}$ be a sequence of $\mathscr{S}$-measurable functions on X which converges to a function f μ-almost everywhere. Prove that

(a) if the measure μ is complete, then also f is $\mathscr{S}$-measurable;

(b) if f is defined on X and is $\mathscr{S}$-measurable, then there exists a sequence $\{f_n^*\}$ of $\mathscr{S}$-measurable functions such that $f_n = f_n^*$ μ-almost everywhere and $f_n^* \to f$ everywhere on X;

(c) if the measure μ is complete, then a function f is $\mathscr{S}$-measurable if and only if there exists a sequence of simple functions which converges to f μ-almost everywhere.

3.18. Images and Preimages of Measurable Sets. We know that real-valued $\mathscr{S}$-measurable functions are exactly those for which the preimages of Borel sets are in $\mathscr{S}$. To illustrate the situation, let us present the following example for the Lebesgue measure on $\mathbf{R}$.

(a) Let $g(x) = \frac{1}{2}(x + f(x))$ where f is the Cantor singular function from 23.1. Then g is a continuous and increasing function mapping the interval $[0,1]$ on $[0,1]$. If C is the Cantor set and φ is the inverse function to g on $[0,1]$, then $\varphi^{-1}(C)$ is a (Lebesgue) measurable set of positive measure. As remarked in 1.9.2 there exists a nonmeasurable set $E \subset \varphi^{-1}(C)$. Finally, for $M := \varphi(E)$ we have $M \subset C$, so that M is measurable while $\varphi^{-1}(M) = g(M) = E$ is a nonmeasurable set. Let us note that this M is not even a Borel set (compare also with Exercise 1.14.a).

(b) Without presenting a proof, we mention that the continuous image of a Borel set on $\mathbf{R}$ is always measurable but not necessarily Borel.

4. Construction of Measures from Outer Measures

In this chapter, we will construct measures from the so-called outer measures. This is a very useful method used already before when we introduced the Lebesgue measure in $\mathbf{R}^n$. We also fulfil our promise to prove propositions from the first chapter.

4.1. Outer Measure. By an *outer measure* on a set X we understand a set function γ which assigns to *every* set $A \subset X$ a nonnegative number γA (real or $+\infty$) such that the following conditions are satisfied:

(a) $\gamma\emptyset = 0$;
(b) if $A \subset B$, then $\gamma A \le \gamma B$;
(c) $\gamma\left(\bigcup A_n\right) \le \sum \gamma A_n$.

The property (c) is called the *σ-subadditivity* of an outer measure.

4.2. Examples of Outer Measures. (a) The outer Lebesgue measure in $\mathbf{R}^n$.

(b) The *Hausdorff* outer measure from Chapter 36.

(c) The counting measure.

(d) If $(X, \mathscr{S}, \mu)$ is a measure space, then the set function $\mu^* \colon A \mapsto \inf\{\mu M : M \in \mathscr{S} \colon M \in \mathscr{S},\ M \supset A\}$ is an outer measure. See also Exercise 4.8.

An important example of creating an outer measure is contained in the following theorem.

4.3. Theorem. *Let $\mathscr{G}$ be a collection of subsets of a set X containing $\emptyset$, and $\nu \colon \mathscr{G} \to [0, +\infty]$ be a set function on $\mathscr{G}$ with $\nu\emptyset = 0$. For $A \subset X$ set*

$$\tilde{\nu} A = \inf\{\sum_{n=1}^{\infty} \nu G_n : G_n \in \mathscr{G},\ \bigcup_{n=1}^{\infty} G_n \supset A\}$$

(note that $\inf\emptyset = +\infty$). Then $\tilde{\nu}$ is an outer measure.

Proof. We only have to verify that $\tilde{\nu}$ is σ-subadditive. So suppose $A \subset \bigcup_n A_n$, $\sum \tilde{\nu} A_n < +\infty$. Fix $\varepsilon > 0$ and find $G_n^j \in \mathscr{G}$ so that

$$\bigcup_j G_n^j \supset A_n \text{ and } \sum_j \nu G_n^j < \tilde{\nu} A_n + \frac{\varepsilon}{2^n}.$$

Then

$$\bigcup_{n,j} G_n^j \supset A \text{ and } \sum_n \tilde{\nu} A_n \geq \tilde{\nu} A - \varepsilon.$$

■

4.4. γ-measurable Sets. A set $M \subset X$ is said to be γ-*measurable (in the sense of Carathéodory)* if

$$\gamma T = \gamma(T \cap M) + \gamma(T \setminus M)$$

for each "test" set $T \subset X$ (in other words, if M splits additively each set in X). The collection of all γ-measurable sets will be denoted by $\mathfrak{M}(\gamma)$. To prove that a set M is γ-measurable, it is sufficient to verify the inequality

$$\gamma T \geq \gamma(T \cap M) + \gamma(T \setminus M)$$

for any set T with $\gamma T < +\infty$.

4.5. Theorem. *$\mathfrak{M}(\gamma)$ is a σ-algebra and γ is a complete measure on $\mathfrak{M}(\gamma)$.*

Proof will be divided into a few steps.

(a) It is straightforward to check that $X \in \mathfrak{M}(\gamma)$, that $X \setminus M \in \mathfrak{M}(\gamma)$ provided $M \in \mathfrak{M}(\gamma)$, and that $A \in \mathfrak{M}(\gamma)$ whenever $\gamma A = 0$. Suppose M, $N \in \mathfrak{M}(\gamma)$. We would like to show that also $M \cap N \in \mathfrak{M}(\gamma)$. So choose a test set $T \subset X$. Then

$$\gamma T = \gamma(T \cap M) + \gamma(T \setminus M)$$

and

$$\gamma(T \cap M) = \gamma(T \cap M \cap N) + \gamma((T \cap M) \setminus N).$$

Now we use the test set $T \setminus (M \cap N)$ and γ-measurability of M to get

$$\gamma(T \setminus (M \cap N)) = \gamma((T \cap M) \setminus N) + \gamma(T \setminus M).$$

Thanks to last three equalities, it follows that

$$\gamma T = \gamma(T \cap (M \cap N)) + \gamma(T \setminus (M \cap N)).$$

Since $\mathfrak{M}(\gamma)$ is closed under complements and finite intersections, it is closed also under finite unions.

(b) In order to show that γ is σ-additive on $\mathfrak{M}(\gamma)$, choose $M_n \in \mathfrak{M}(\gamma)$ pairwise disjoint. Setting $T = M_1 \cup M_2$ and using γ-measurability of M_1 we obtain

$$\gamma(M_1 \cup M_2) = \gamma M_1 + \gamma M_2.$$

Thus γ is finitely additive. Further

$$\sum_{n=1}^{\infty} \gamma M_n = \lim_k \sum_{n=1}^{k} \gamma M_n = \lim_k \gamma \left(\bigcup_{n=1}^{k} M_n \right) \le \gamma \left(\bigcup_{n=1}^{\infty} M_n \right),$$

and since the reverse inequality always holds we reach the conclusion.

(c) Let now $M_n \in \mathfrak{M}(\gamma)$ be pairwise disjoint. Our aim is to show that $\bigcup\limits_{n=1}^{\infty} M_n \in \mathfrak{M}(\gamma)$. Choosing a test set $T \subset X$, we have

$$\gamma T = \gamma \left(T \setminus \bigcup_{n=1}^{k} M_n \right) + \gamma \left(T \cap \bigcup_{n=1}^{k} M_n \right) \ge \gamma \left(T \setminus \bigcup_{n=1}^{\infty} M_n \right) + \sum_{n=1}^{k} \gamma(T \cap M_n)$$

for each $k \in \mathbf{N}$. Since γ is σ-subadditive, it readily follows that

$$\gamma T \ge \gamma \left(T \setminus \bigcup_{n=1}^{\infty} M_n \right) + \gamma \left(T \cap \bigcup_{n=1}^{\infty} M_n \right),$$

which is what we wanted. ■

4.6. Exercise. Let γ be a nonnegative function on $\mathscr{P}(X)$, $\gamma\emptyset = 0$. Show that the collection of all "γ-measurable sets" forms a σ-algebra and that γ is additive on it.

Hint. It only needs to examine where monotonicity and σ-subadditivity of the outer measure in the proof of Theorem 4.5 is used.

4.7. Exercise. We say that an outer measure γ is *regular* if for each set $A \subset X$ there exists $M \in \mathfrak{M}(\gamma)$ such that $A \subset M$ and $\gamma A = \gamma M$.

(a) Let γ be a regular outer measure and $A_1 \subset A_2 \subset A_2 \subset \dots$ an increasing sequence of sets. Show that $\gamma\left(\bigcup A_n\right) = \lim \gamma A_n$.

(b) Show that the outer Lebesgue measure is regular.

(c) There exists a decreasing sequence $\{M_n\}$ of subsets of $[0,1]$ such that $\lambda^* M_n = 1$ and $\bigcap\limits_n M_n = \emptyset$ (compare with (a)).

4.8. Exercise. Let $(X, \mathscr{S}, \mu)$ be a measure space and $A \subset X$. Set

$$\mu^* A := \inf\{\mu M \colon M \in \mathscr{S},\ A \subset M\}, \quad \mu_* A := \sup\{\mu M \colon M \in \mathscr{S},\ M \subset A\}.$$

(a) Suppose $\mu^* A < \infty$. Show that $A \in \mathfrak{M}(\mu^*)$ if and only if $\mu_* A = \mu^* A$.

(b) Show that $\mathscr{S} \subset \mathfrak{M}(\mu^*)$ and $\mu = \mu^*$ on $\mathscr{S}$.

4.9. Notes. Carathéodory's characterization of measurable sets appears first in [*1918].

5. Classes of Sets and Set Functions

This chapter will be devoted to a study of various classes of sets and set functions from the point of view of measure theory. We prove theorems on extensions of set functions to larger families of sets as important steps in constructing measures.

5.1. Systems of sets. A family $\mathscr{A}$ of subsets of a set X containing $\emptyset$ is called:

(a) a *semiring* if

(a1) $A \cap B \in \mathscr{A}$ for each $A,\ B \in \mathscr{A}$,

(a2) for $A, B \in \mathscr{A}$, $A \subset B$, there exist pairwise disjoint sets $C_1, \dots, C_n \in \mathscr{A}$ such that $B \setminus A = \bigcup_{j=1}^{n} C_j$;

(b) a *ring* if given $A, B \in \mathscr{A}$, then $A \cup B,\ A \setminus B \in \mathscr{A}$ (so that also $A \cap B \in \mathscr{A}$);

(c) an *algebra* if it is a ring and $X \in \mathscr{A}$;

(d) a *Dynkin class* if

(d1) $X \in \mathscr{A}$,

(d2) $A \setminus B \in \mathscr{A}$ for each $A, B \in \mathscr{A},\ B \subset A$,

(d3) if $A_n \in \mathscr{A}$ are pairwise disjoint, then $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$;

(e) a *π-system* if $A \cap B \in \mathscr{A}$ for each $A, B \in \mathscr{A}$;

(f) a *δ-ring* if $\mathscr{A}$ is a ring closed under countable intersections;

(g) a *σ-ring* if $\mathscr{A}$ is a ring closed under countable unions.

5.2. Premeasure. A nonnegative set function μ is called a *premeasure* on X if μ is defined on a ring $\mathscr{A}$ of subsets of X and satisfies the following conditions:

(a) $\mu\emptyset = 0$,

(b) if $\{A_j\} \subset \mathscr{A}$ is a sequence of pairwise disjoint sets and $\bigcup A_j \in \mathscr{A}$, then $\mu\left(\bigcup A_j\right) = \sum \mu A_j$.

A premeasure is in fact a σ-additive set function defined on a *ring* of sets. We say that a premeasure μ on X is σ-finite if there exists a sequence X_j of sets from $\mathscr{A}$ such that $\mu X_j < \infty$ and $X = \bigcup X_j$.

5.3. Examples. (a) Denote by $\mathscr{I}$ the collection of all intervals on $\mathbf{R}$ including the degenerated ones (i.e. the empty set and the singletons) and by $\mathscr{I}^b$ the collection of all bounded intervals from $\mathscr{I}$. Further, let $\mathscr{I}_l$ be the collection of all intervals of the form $[a, b)$ together with $\emptyset$, $\mathbf{R}$ and the intervals of the form $(-\infty, b)$. Finally, let $\mathscr{I}_l^b = \mathscr{I}_l \cap \mathscr{I}^b$.

(a1) The collections $\mathscr{I}$, $\mathscr{I}_l$, $\mathscr{I}^b$, $\mathscr{I}_l^b$ form semirings which is not true for the collection of all open (or closed) intervals.

(a2) Finite unions of sets from $\mathscr{I}^b$ or $\mathscr{I}_l^b$ form a ring. Finite unions of sets from $\mathscr{I}$ or $\mathscr{I}_l$ form even an algebra $\mathscr{A}$ or $\mathscr{A}_l$. The set function $I \mapsto \operatorname{vol} I$ can be (uniquely) extended to a premeasure on $\mathscr{A}$ or $\mathscr{A}_l$.

(a3) Closing $\mathscr{I}$ or $\mathscr{I}_l$ under countable unions, we do not get even a semiring: Substracting a countable union of open intervals from [0,1] we get the Cantor set which is not a countable union of intervals.

(b) If $\mathscr{S}$ and $\mathscr{T}$ are semirings on X, then the collection of all sets of the form $A \times B$ where $A \in \mathscr{S}$ and $B \in \mathscr{T}$ forms a semiring on $X \times X$.

(c) Let $\mathscr{D}$ be a semiring. Then the set of all finite disjoint unions of elements from $\mathscr{D}$ is a ring (in fact, the smallest ring containing $\mathscr{D}$).

(d) Consider the family $\mathscr{H}$ of all subsets of the set $\{1, 2, \dots, 20\}$ with an even number of elements and show that $\mathscr{H}$ is a Dynkin class but not a semiring.

(e) The collection of all countable subsets of a set X is a σ-ring which for is not a σ-algebra unless X is countable.

(f) The collection of all Lebesgue measurable sets of finite measure forms a δ-ring.

(g) Suppose that $\mathscr{A}$ is the collection of all unions of a finite number of intervals (including the degenerated ones) and $\mathscr{A}' = \{A \cap \mathbf{Q} \colon A \in \mathscr{A}\}$. Then $\mathscr{A}'$ is an algebra of subsets of $\mathbf{Q}$ and $\nu \colon A \cap \mathbf{Q} \mapsto \lambda A$ is a finitely additive set function which is not a premeasure.

(h) Let $\mathscr{A}$ be the algebra of all finite unions of intervals on $(0,1)$. Define a set function ν on $\mathscr{A}$ by the formula

$$\mu(A) = \begin{cases} 1 & \text{if } A \text{ contains some interval of the form } (0,\varepsilon),\ \varepsilon > 0, \\ 0 & \text{in other cases.} \end{cases}$$

Then ν is a finitely additive set function on $\mathscr{A}$ which is not a premeasure.

5.4. Theorem. *Let $\mathscr{G}$ be a ring of subsets of a set X and ν a premeasure on $\mathscr{G}$. If $\tilde{\nu}$ is the outer measure constructed from $\mathscr{G}$ and ν as in Theorem 4.3, then*

(a) $\tilde{\nu} = \nu$ on $\mathscr{G}$;

(b) $\mathscr{G} \subset \mathfrak{M}(\tilde{\nu})$.

Proof. Obviously $\tilde{\nu} \le \nu$. Suppose $G \in \mathscr{G}$, $\tilde{\nu}G < +\infty$ and $\{G_n\} \subset \mathscr{G}$, $\bigcup G_n \supset G$. Then

$$\nu G = \nu\left(\bigcup_n (G_n \cap G)\right) \le \sum_n \nu(G_n \cap G) \le \sum_n \nu G_n,$$

so that $\nu G \le \tilde{\nu} G$.

Now suppose $G \in \mathscr{G}$. Choose a test set $T \subset X$, $\tilde{\nu}T < +\infty$ and $\varepsilon > 0$. There exist $G_n \in \mathscr{G}$ such that

$$T \subset \bigcup_n G_n \quad \text{and} \quad \sum_n \nu G_n \le \tilde{\nu}T + \varepsilon.$$

Then

$$\tilde{\nu}T + \varepsilon \ge \sum_n \nu G_n = \sum_n \left(\nu(G_n \cap G) + \nu(G_n \setminus G)\right) \ge \tilde{\nu}(T \cap G) + \tilde{\nu}(T \setminus G),$$

thus $G \in \mathfrak{M}(\tilde{\nu})$. ■

5.5. Hopf's Extension Theorem. *Let μ be a premeasure on a ring $\mathscr{A}$ of subsets of a set X. Then there exists a measure $\tilde{\mu}$ on $\sigma(\mathscr{A})$ which equals μ on $\mathscr{A}$. This extension of μ is unique provided μ is σ-finite.*

Proof. The existence of a measure $\tilde{\mu}$ is an immediate consequence of Theorems 4.3 and 5.4 (notice that $\sigma(\mathscr{A}) \subset \mathfrak{M}(\tilde{\mu})$). To prove uniqueness, under the σ-finiteness assumptions, let ν be another measure on $\sigma(\mathscr{A})$, $\nu = \mu$ on $\mathscr{A}$. One can easily find out (from the construction of the outer measure $\tilde{\mu}$) that $\nu \le \tilde{\mu}$ on $\sigma(\mathscr{A})$. Then for $A_j \in \mathscr{A}$, $A = \bigcup_j A_j$ we have

$$\mu A = \lim_n \nu\left(\bigcup_{j=1}^n A_j\right) = \lim_n \mu\left(\bigcup_{j=1}^n A_j\right) = \tilde{\mu}A.$$

So if $E \in \sigma(\mathscr{A})$, $\tilde{\mu}E < \infty$, then for a given $\varepsilon > 0$ there exist sets $A_j \in \mathscr{A}$, $A = \bigcup_j A_j$ such that $E \subset A$ and $\tilde{\mu}A < \tilde{\mu}E + \varepsilon$. Hence

$$\tilde{\mu}E \leq \tilde{\mu}A = \nu A = \nu E + \nu(A \setminus E) \leq \nu E + \tilde{\mu}(A \setminus E) \leq \nu E + \varepsilon,$$

thus $\tilde{\mu}E = \nu E$. Finally suppose $X = \bigcup_j X_j$, $\mu X_j < +\infty$; one can assume that X_j are pairwise disjoint. If $E \in \sigma(\mathscr{A})$, then

$$\tilde{\mu}E = \sum_j \tilde{\mu}(E \cap X_j) = \sum_j \nu(E \cap X_j) = \nu E.$$

5.6. Exercise. Let $\mathscr{A}$ be a family of subsets of a set X. Show that there exists the smallest Dynkin class $\mathscr{D}(\mathscr{A})$ containing $\mathscr{A}$.

5.7. Exercise. Show that every σ-algebra is a Dynkin class. A Dynkin class is a σ-algebra if and only if it is a π-system.

5.8. Exercise. Prove that if $\mathscr{A}$ is a π-system, then $\mathscr{D}(\mathscr{A}) = \sigma(\mathscr{A})$.

5.9. Exercise. The assumptions of Hopf's extension theorem can be weakened. Indeed, show that the assertion of Theorem 5.4 is still true if we only assume that μ is finitely additive and σ-subadditive on the semiring $\mathscr{A}$ and that $\mu\emptyset = 0$.

Hint. First note that μ is monotone. Then proceed as in Theorem 5.4 and show that $\sigma(\mathscr{A}) \subset \mathscr{M}(\tilde{\mu})$. In the essential step use the existence of sets $C_n^j \in \mathscr{A}$ with the property $G_n \setminus (G \cap G_n) = \bigcup_j C_n^j$.

5.10. Exercise. Let $\mathscr{A}$ be a π-system on X. Show that if μ_1, μ_2 are probability measures on $\sigma(\mathscr{A})$ which agree on $\mathscr{A}$, then $\mu_1 = \mu_2$ on $\sigma(\mathscr{A})$.

Hint. Let $\mathscr{D} := \{M \in \sigma(\mathscr{A})\colon \mu_1(M) = \mu_2(M)\}$. Show that $\mathscr{D}$ is a Dynkin class and use Exercise 5.7.

5.11. Exercise. Consider

$$\mathscr{A} := \{[a,b) \setminus C\colon a, b \in [0,1]\}, \quad \nu([a,b) \setminus C) := f(b) - f(a),$$

where C is the Cantor set and f the Cantor function from 23.1.

Show that $\mathscr{A}$ is a semiring and that ν is finitely additive but not σ-additive on $\mathscr{A}$.

5.12. Exercise. (a) Let X be an arbitrary set and $n \in \mathbf{N}$. Consider the set function ν which is restriction of the counting measure to the collection

$$\mathscr{G}_n := \emptyset \cup \{A \subset X\colon A \text{ has exactly } n \text{ elements}\}$$

and construct the outer measure $\tilde{\nu}$ as in Theorem 4.3. Investigate the relationship between the collections $\mathscr{G}_n$ and $\mathfrak{M}(\tilde{\nu})$ and compare with Theorem 5.4.

(b) Investigate the same problem in case of $X = (0,1)$, $\mathscr{G} = \mathscr{P}(X)$ and

$$\nu A := \inf\{\sum_{i=1}^{k}(b_i - a_i)\colon \bigcup_{i=1}^{k}(a_i, b_i) \supset A\}$$

(νA is the *Jordan-Peano* content of a set A).

5.13. Exercise. Let μ^* be an outer measure on X constructed from the premeasure μ on an algebra by Theorem 4.3 such that $\mu^* X < +\infty$. For $A \subset X$ set

$$\mu_* A := \mu^* X - \mu^*(X \setminus A).$$

Show that $A \in \mathfrak{M}(\mu^*)$ if and only if $\mu_* A = \mu^* A$ (compare also with Theorem 5.4).

5.14. Capacity on Compact Sets. Let X be a locally compact topological space. A real-valued nonnegative function $\mathscr{C}$ defined on the collection $\mathscr{K}(X)$ of all compact subsets of X is called a *Choquet capacity* if it satisfies the following conditions:

(a) if $K_1 \subset K_2$, then $\mathscr{C}(K_1) \le \mathscr{C}(K_2)$;
(b) if $K_1 \supset K_2 \supset K_3 \dots$, then $\mathscr{C}\left(\bigcap K_n\right) = \lim \mathscr{C}(K_n)$;
(c) $\mathscr{C}(K_1 \cap K_2) + \mathscr{C}(K_1 \cup K_2) \le \mathscr{C}(K_1) + \mathscr{C}(K_2)$ whenever $K_n \in \mathscr{K}(X)$ (so-called *strong subadditivity*).

An important example is the Newtonian capacity in $\mathbf{R}^n$, $n \ge 3$. If we define the *Newtonian potential* of a Radon measure μ on $\mathbf{R}^n$ by

$$\mathscr{N}^\mu(x) := \int_{\mathbf{R}^n} \frac{\mathrm{d}\mu(t)}{|x-t|^{n-2}},$$

we can introduce the *Newtonian capacity* $\operatorname{cap} K$ of a compact set $K \subset \mathbf{R}^n$ as

$$\operatorname{cap} K := \sup\{\mu K \colon \operatorname{supt} \mu \subset K,\ \mathscr{N}^\mu \le 1 \text{ on } \mathbf{R}^n\}.$$

The proof that this capacity satisfies the conditions (a), (b), (c) can be found for instance in [KNV].

5.15. Outer Capacity. Let X be a topological space. A mapping $c : \mathscr{P}(X) \to [0, +\infty]$ is called an *outer capacity* provided it satisfies:

(a) if $A \subset B$, then $cA \le cB$;
(b) if $A_1 \subset A_2 \subset A_3 \subset \dots$, then $c\left(\bigcup A_n\right) = \sup cA_n$;
(c) if $K_1 \supset K_2 \supset K_3 \supset \dots$ and K_n are compact, then $c(\bigcap K_n) = \inf c K_n$.

A set $A \subset X$ is said to be *c-capacitable* if

$$cA = \sup\{cK \colon K \subset A,\ K \text{ compact}\}.$$

1. If $\mathscr{C}$ is a Choquet capacity on a locally compact space (Exercise 5.14) and

$$cA := \inf\{\sup\{\mathscr{C}(K) \colon K \subset G,\ K \text{ compact}\} : G \supset A,\ G \text{ open}\},$$

prove that c is an outer capacity.

2. Suppose c is simultaneously an outer capacity and an outer measure. Investigate the relationship between the notions of c-capacitability and c-measurability in the sense of Carathéodory.

Consider the cases:

(a) X is a two-points set equipped with the discrete topology and $cA = 1$ if $A \ne \emptyset$, $c\emptyset = 0$;

(b) X is $\mathbf{R}$ with the Euclidean topology and c is the outer Lebesgue measure;

(c) X is $\mathbf{R}$ with the discrete topology and c is again the outer Lebesgue measure;

(d) c is the outer capacity derived from the Newtonian capacity. In this case a set $A \subset \mathbf{R}^n$ is c-measurable in the sense of Carathéodory if and only if $cA = 0$ or if $c(\mathbf{R}^n \setminus A) = 0$ (this is not quite easy, see M.M. Rao [*1987]). On the other hand, Choquet's deep result says that each Borel set in $\mathbf{R}^n$ is c-capacitable.

5.16. Notes. "Hopf's extension theorem" is usually attributed to H. Hahn or C. Carathéodory but it is probably originated by M. Fréchet [1915]. The proof using Carathéodory's theorem was given independently by H. Hahn [*1924] and A. N. Kolmogorov [*1933].

The notions of a π-system or of a Dynkin class were investigated by E. B. Dynkin in 1959 as tools for the probability theory. However, families of similar properties were studied already by W. Sierpiński [1928].

An investigation of the theory of capacities in the classical potential theory during the period 1920–1950 is connected with the names of Ch. de la Vallée Poussin, N. Wiener, O. Frostman or M. Brelot among others. These authors studied mainly the Newtonian capacity and examined the role of sets of a "small" capacity. One could say that the capacity theory is a "younger sister" of Lebesgue measure theory. In the 50's the general capacity theory was developed and G. Choquet proved his famous capacitability theorem. An interested reader is referred to a nice article written by Choquet himself [1986] or [1989].

6. Signed and Complex Measures

In this chapter, we will investigate "measures" assuming also negative (or even complex) values.

6.1. Signed Measures. Let $\mathscr{S}$ be a σ-algebra of subsets of X. A set function $\mu\colon \mathscr{S} \to \overline{\mathbf{R}}$ is said to be a *signed measure* on $\mathscr{S}$ if

(a) $\mu\emptyset = 0$;

(b) $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu A_n$ whenever $A_n \in \mathscr{S}$ are pairwise disjoint.

6.2. Remarks. 1. A signed measure can assume at most one of the values $+\infty$ and $-\infty$. Indeed, if $\mu E = +\infty$ and $\mu F = -\infty$ for $E, F \in \mathscr{S}$, then $\mu(E \cap F)$ is finite. Therefore $\mu(E \setminus F) = +\infty, \mu(F \setminus E) = -\infty$ and the equality (b) does not hold for (disjoint) sets $E \setminus F$ and $F \setminus E$.

2. If $\mu(\bigcup A_n)$ is finite, then the series $\sum \mu A_n$ in the condition (b) converges absolutely (indeed, this series converges — to the same value — when its terms are rearranged arbitrarily, which is equivalent to the absolute convergence).

3. Some of basic properties of positive measures hold even for signed measures. Show that the following assertions are true:

(a) if $A_n \in \mathscr{S},\ A_n \nearrow A$, then $\mu A_n \to \mu A$,

(b) if $A_n \in \mathscr{S},\ A_n \searrow A,\ |\mu A_1| < +\infty$, then $\mu A_n \to \mu A$.

4. A signed measure μ is not necessarily monotone: If $A \subset B$, we can no longer assert that $\mu A \le \mu B$.

6.3. Hahn Decomposition. We say that a set $P \in \mathscr{S}$ is *positive* for a signed measure μ if $\mu E \ge 0$ for each set $E \in \mathscr{S}$, $E \subset P$. Analogously we define *negative* sets for μ.

An ordered pair of sets (P, N) is called a *Hahn decomposition* of X for μ if

(a) $P \cup N = X$, $P \cap N = \emptyset$;

(b) P is positive and N is negative (for μ).

6.4. Examples. (a) The empty set is both positive and negative for any signed measure. The pair $(X, \emptyset)$ is a Hahn decomposition for μ if and only if μ is positive.

(b) The pairs $(\mathbf{R}, \emptyset)$, $(\mathbf{R} \setminus \mathbf{Q}, \mathbf{Q})$, $(\mathbf{R} \setminus \{5\}, \{5\})$ are Hahn decompositions of $\mathbf{R}$ for the Lebesgue measure.

(c) If a measure μ_f has a density f with respect to μ (i.e. if $\mu_f A = \int_A f \,\mathrm{d}\mu$ where $f \in \mathscr{L}^*(\mu)$ — cf. 8.19), then the set $P := \{f \ge 0\}$ is positive for μ_f and $(\{f \ge 0\}, \{f < 0\})$ is a Hahn decomposition for μ_f.

6.5. Hahn Decomposition Theorem. *For every signed measure μ there exists a Hahn decomposition. This decomposition is unique in the following sense: If (P_1, N_1) and (P_2, N_2) are two such decompositions, then*

$$\mu(P_1 \cap E) = \mu(P_2 \cap E), \quad \mu(N_1 \cap E) = \mu(N_2 \cap E)$$

for each set $E \in \mathscr{S}$.

Proof. We start with proof of the uniqueness assertion. Suppose that (P_1, N_1) and (P_2, N_2) are Hahn decompositions of X for μ. Then evidently

$$\mu(P_1 \cap E) = \mu(E \cap (P_1 \cap P_2)) = \mu(E \cap P_2)$$

and, in the same way, $\mu(N_1 \cap E) = \mu(N_2 \cap E)$.

To prove the existence of the decomposition, assume that $\mu S < +\infty$ for all $S \in \mathscr{S}$. We proceed via the following steps.

STEP 1: *If $A \in \mathscr{S}$, $\mu A > -\infty$ and $\varepsilon > 0$, then there exists a set $B \subset A$ such that $\mu B \geq \mu A$ and B is positive.* The set B will be constructed as the intersection of a finite or infinite nested sequence $\{A_n\}$ of subsets of A. We start with $A_1 = A$. In the n-th step, if A_n is not positive, we can set

$$k_n = \min\{k \in \mathbf{N}\colon \text{ there exists } E \subset A_n \text{ with } \mu(E) \leq -\frac{1}{k}\}$$

and find $A_{n+1} \subset A_n$ such that $\mu(A_n \setminus A_{n+1}) \leq -\frac{1}{k_n}$. Then we set $B = \bigcap_{n=1}^{\infty} A_n$. The sets $A_n \setminus A_{n+1}$ are pairwise disjoint and their union is $A \setminus B$. Thus

$$+\infty > \mu(B) = \mu(A) - \sum_{n=1}^{\infty} \mu(A_n \setminus A_{n+1}) \geq \mu(A) + \sum_{n=1}^{\infty} \frac{1}{k_n}.$$

It follows that $\mu(B) \geq \mu(A)$ and $k_n \to \infty$. If $E \subset B$ is a measurable subset, then for each $n \in \mathbf{N}$ we have $E \subset A_n$, and thus $\mu(E) > -\frac{1}{k_n}$. Hence B is positive.

STEP 2: To complete the proof, set $s = \sup\{\mu A\colon A \in \mathscr{S}\}$. Since $\emptyset \in \mathscr{S}$, we have $s \geq 0$. There exists a sequence $\{P_n\} \subset \mathscr{S}$ such that $\mu P_n \to s$. In light of the first step we can assume $P_1 \subset P_2 \subset \ldots$ (the union of two positive sets is a positive set!). If $P := \bigcup P_n$, then $P \subset A$ and $\mu P = \lim \mu P_n = s < +\infty$. Moreover, P is positive. Indeed, if $E \subset P$, then $E \cap P_n \nearrow E$ and $\mu(E \cap P_n) \geq 0$. Lastly, to show that the set $N := X \setminus P$ is negative we notice that $\mu(E \cup P) = \mu E + \mu P > s$ for any set $E \subset N$ with $\mu E > 0$. ■

6.6. Variation of a Measure. Let (P, N) be a Hahn decomposition of X for a signed measure μ. Define

$$\mu^+ E = \mu(E \cap P), \quad \mu^- E = -\mu(E \cap N), \quad |\mu|(E) = \mu^+ E + \mu^- E$$

for every set $E \in \mathscr{S}$. By the previous theorem, $\mu^+ E$ and $\mu^- E$ do not depend on the choice of the Hahn decomposition (which is *not* unique!). It is simply checked that the set functions μ^+, μ^- and $|\mu|$ are positive measures on $\mathscr{S}$. They are called the *positive*, *negative* and *total* variations of a signed measure μ. Let us summarize our results in the following theorem.

6.7. Theorem (The Jordan decomposition of a signed measure)**.** *Let μ be a signed measure on $(X,\mathscr{S})$. The functions μ^+, μ^- and $|\mu|$ are positive measures on $\mathscr{S}$ and $\mu = \mu^+ - \mu^-$. If also $\mu = \nu_1 - \nu_2$ where ν_1, ν_2 are positive measures, then $\nu_1 \geq \mu^+$, $\nu_2 \geq \mu^-$.*

Proof. It is enough to prove the last part of the theorem. But for each $E \in \mathscr{S}$, one has

$$\mu^+E = \mu(E \cap P) = \nu_1(E \cap P) - \nu_2(E \cap P) \leq \nu_1(E \cap P) \leq \nu_1 E.$$

An important characterization of the total variation of a signed measure, often used as its definition, is contained in the following theorem.

6.9. Theorem. *Let μ be a signed measure on $(X,\mathscr{S})$ and $E \in \mathscr{S}$. Then*

$$|\mu|(E) = \sup\left\{\sum_{k=1}^{n} |\mu A_k| : A_k \in \mathscr{S} \text{ pairwise disjoint, } \bigcup_{k=1}^{n} A_k = E\right\}.$$

Proof. It is straightforward to check that

$$\sum_k |\mu A_k| = \sum_k \left|\mu^+ A_k - \mu^- A_k\right| \leq \sum_k (\mu^+ A_k + \mu^- A_k) = \sum_k |\mu|(A_k) = |\mu|(E).$$

On the other hand, if (P,N) is a Hahn decomposition of X for μ, set $A_1 = E \cap P$, $A_2 = E \cap N$ and find out that the supremum is even attained.

6.10. Complex Measures. A *complex measure* on $(X,\mathscr{S})$ is a σ-additive set function $\mu\colon \mathscr{S} \to \mathbf{C}$ for which $\mu\emptyset = 0$. By σ-additivity we understand that $\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu A_k$ whenever $\{A_k\}$ is a sequence of pairwise disjoint sets from $\mathscr{S}$. Let us note that, as in Remark 6.2.2, the appearing series must converge absolutely or definitely diverge.

Each complex measure μ can be expressed uniquely in the form $\mu = \mu_r + \mathrm{i}\mu_i$ where μ_r and μ_i are (finite!) signed measures on $(X,\mathscr{S})$. In particular, each *finite* signed measure can be understood as a complex measure.

If μ is a complex measure, we define its *total variation* $|\mu|$ by

$$|\mu|(E) = \sup\left\{\sum_{k=1}^{n} |\mu A_k| : A_k \in \mathscr{S} \text{ pairwise disjoint, } \bigcup_{k=1}^{n} A_k = E\right\}.$$

An appeal to Theorem 6.9 reveals that this definition agrees with the previous one for signed measures.

6.11. Theorem. *Let μ be a complex measure on $(X,\mathscr{S})$. Then the total variation $|\mu|$ is a finite positive measure on $(X,\mathscr{S})$.*

Proof. Apparently, $|\mu|(\emptyset) = 0$ and it is simply to verify that $|\mu|$ is finitely additive. If $\mu = \mu_1 - \mu_2 + \mathrm{i}(\mu_3 - \mu_4)$ where $\mu_1 - \mu_2$ is the Jordan decomposition of μ_r and $\mu_3 - \mu_4$ is the Jordan decomposition of μ_i (i.e. μ_j are positive finite measures), then

$$|\mu|(A) \le \mu_1 A + \mu_2 A + \mu_3 A + \mu_4 A,$$

so that $|\mu|(A) < +\infty$. If $A_n \in \mathscr{S}$, $A_n \searrow \emptyset$, then $\lim \mu_j A_n = 0$ for each $j = 1, 2, 3, 4$, and therefore $|\mu|(A_n) \to 0$. A routine argument now shows that μ is σ-additive. ■

6.12. Remark. Note that the range of any complex measure is always a *bounded* subset of the complex plane **C**.

6.13. Exercise. Let μ be a signed measure on $(X, \mathscr{S})$ and $E \in \mathscr{S}$. Show that $\mu^+ E = \sup\{\mu B \colon B \in \mathscr{S},\ B \subset E\}$.

6.14. Exercise. Let μ be a complex measure on $(X, \mathscr{S})$ and $E \in \mathscr{S}$. Prove that $|\mu|(E) = \sup\left\{ \sum_{k=1}^{\infty} |\mu A_k| : A_k \in \mathscr{S} \text{ pairwise disjoint, } \bigcup_{k=1}^{\infty} A_k = E \right\}$.

6.15. Exercise. Let μ be a complex measure on $(X, \mathscr{S})$. Prove that $|\mu|$ is the smallest positive measure ν satisfying $\nu A \ge |\mu| A$ for all $A \in \mathscr{S}$.

6.16 Exercise. Investigate whether $|\mu| = |\mu_r| + |\mu_i|$, or $|\mu| = \sqrt{|\mu_r|^2 + |\mu_i|^2}$.

6.17. Exercise. We denote by $\mathbf{M}(\mathscr{S})$ the set of all finite signed or complex measures on $(X, \mathscr{S})$. If $\mu \in \mathbf{M}(\mathscr{S})$, let $\|\mu\| = |\mu|(X)$. Show that $(\mathbf{M}(\mathscr{S}), \|\cdot\|)$ is a Banach space (i.e. $\mathbf{M}(\mathscr{S})$ is a linear space, $\|\cdot\|$ becomes a norm and $\mathbf{M}(\mathscr{S})$ is complete with respect to it).

Hint. Only the completeness requires a proof. But if $\|\mu_n\|$ is a Cauchy sequence in $\mathbf{M}(\mathscr{S})$, then there exists $\lim \mu_n A$ for each $A \in \mathscr{S}$. Defining $\mu A = \lim \mu_n A$, it is not hard to show that μ is σ-additive and $\|\mu_n - \mu\| \to 0$.

6.18. Exercise. Suppose that $\mu, \mu_n \in \mathbf{M}(\mathscr{S})$, $\|\mu_n - \mu\| \to 0$. Prove that $\mu_n A \to \mu A$ for every set $A \in \mathscr{S}$.

6.19. Exercise. Show that the set $\mathbf{M}(\mathscr{S})$ of all signed measures on $(X, \mathscr{S})$ equipped with the order

$$\mu \le \nu \qquad \text{if } \mu A \le \nu A \text{ for each } A \in \mathscr{S}$$

is a lattice (i.e. for any $\mu, \nu \in \mathbf{M}(\mathscr{S})$ there exist $\sup(\mu, \nu)$ and $\inf(\mu, \nu)$ in $\mathbf{M}(\mathscr{S})$). Notice that $\sup(\mu, \nu)$ does not necessarily mean the set function $A \mapsto \sup(\mu(A), \nu(A))$ — such function, in general, is not a measure!.

Hint. Show that $\frac{1}{2}(\mu + \nu + |\nu - \mu|)$ is $\sup(\mu, \nu)$ and $\frac{1}{2}(\mu + \nu - |\nu - \mu|)$ is $\inf(\mu, \nu)$.

6.20. Finitely Additive Measures — Charges. A real-valued set function ν on an *algebra* of sets $\mathscr{A}$ is called a *charge* if

(a) $\nu\emptyset = 0$;
(b) $\nu(A \cup B) = \nu A + \nu B$ whenever $A, B \in \mathscr{S}$, $A \cap B = \emptyset$.

If ν is a charge, define $\nu^+ A = \sup\{\nu F \colon F \in \mathscr{A},\ F \subset A\}$, $\nu^- A = -\inf\{\nu F \colon F \in \mathscr{A},\ f \subset A\}$, $|\nu|(A) = \nu^+ A + \nu^- A$.

6.21. Exercise. Show that the set functions ν^+, ν^- and $|\nu|$ are positive finitely additive measures on $\mathscr{A}$.

6.22. Exercise. Suppose $X = [0, 1)$. Let $\mathscr{A}$ be the collection of all finite unions of intervals $[a, b) \subset [0, 1)$. Verify that $\mathscr{A}$ is an algebra. Set

$$f(x) = \begin{cases} \frac{1}{x} & \text{for } x \in (0, 1), \\ 0 & \text{for } x = 0. \end{cases}$$

If $A = \bigcup_{i=1}^{n} [a_i, b_i)$, define $\nu A = \sum(f(b_i) - f(a_i))$. Show that neither the "Jordan decomposition theorem", nor the "Hahn decomposition theorem" do hold for ν. Proceed via the following steps:

(a) the definition of νA does not depend on the particular representation of A as $\bigcup[a_i, b_i)$;

(b) ν is a charge on $\mathscr{A}$;

(c) $\nu A \leq 0$ for $A \subset (0, 1)$;

(d) $\nu([0, 1)) = 1$;

(e) $\nu^+([0, 1)) = \nu^-([0, 1)) = +\infty$.

6.23. Exercise. Let ν be a charge on $\mathscr{A}$. Show that $\nu = \nu^+ - \nu^-$ if and only if ν is bounded (i.e. if $\sup\{|\nu A| : A \in \mathscr{A}\} < +\infty$).

6.24. Notes. Signed measures were investigated by H. Lebesgue already in [1910]; he was concerned mostly with measures given by a density (see 8.19).

The existence of a Hahn decomposition of the space and the Jordan decomposition of signed measures were first established in a full generality by Hahn [*1921].

The monograph by K.P.S. Bhaskar Rao and M. Bhaskar Rao [*1983] is devoted to the theory of charges.

B. The Abstract Lebesgue Integral

7. Integration on R

7.1. Newton Integral. Let f be a real-valued function on an interval (a, b). A function F on (a, b) is called an *antiderivative* to f if $F'(x) = f(x)$ for all $x \in (a, b)$. A real number A is called the *Newton integral* of f over (a, b) if there is an antiderivative F of f on (a, b) such that $A = \lim_{x \to b-} F(x) - \lim_{x \to a+} F(x)$. The value of the Newton integral of f on (a, b) does not depend on the choice of F because the difference between any two antiderivatives is a constant. The Newton integral is denoted by $N \int_a^b f$.

7.2. Riemann Integral. Let f be a bounded function on a bounded interval (a, b). If $a = \xi_0 < \xi_1 < \cdots < \xi_m = b$ is a (finite) sequence of points in $[a, b]$ (so called *partition* of the interval $[a, b]$), and $\alpha_1, \ldots, \alpha_m$ a sequence of real numbers, then the value $\sum_{j=1}^{m} \alpha_j (\xi_j - \xi_{j-1})$ is called an *upper Riemann sum* of f if $\alpha_j \geq f$ on every interval (ξ_{j-1}, ξ_j), and a *lower Riemann sum* of f if $\alpha_j \leq f$ on every (ξ_{j-1}, ξ_j). Set

$$\mathscr{R}^* f = \inf\{A : A \text{ is an upper Riemann sum of } f\},$$
$$\mathscr{R}_* f = \sup\{A : A \text{ is a lower Riemann sum of } f\}.$$

A function f is *Riemann integrable* on (a, b) provided $\mathscr{R}^* f = \mathscr{R}_* f$; the common value is then called the *Riemann integral* of f over (a, b) and it is denoted by $R \int_a^b f$.

7.3. Theorem. *Let f be a continuous function on an interval $[a, b]$. Then both the Newton integral and the Riemann integral of f on (a, b) exist and are equal.*

Proof. Since f is uniformly continuous, we easily obtain that $R \int_a^x f$ exists for all $x \in (a, b)$. A short reflection shows that $x \mapsto R \int_a^x f$ is an antiderivative of f and this implies the existence of the Newton integral and the required equality. ∎

7.4. Remark. Each Riemann integrable function f is *absolutely* Riemann integrable, i.e. $|f|$ is Riemann integrable as well. This assertion is no longer true for the Newton integral (consider the integration of the function $x \mapsto \sin x / x$ over $(1, +\infty)$; the change of variables $t = 1/x$ yields an example on a bounded interval).

7.5. Dirichlet Function. The *Dirichlet function* D is defined as the indicator function of the set of all rational numbers. Observe that $D = 0$ almost everywhere. The Dirichlet function is nowhere continuous, it has neither an antiderivative (it does not have the Darboux property) nor the Riemann integral over $(0, 1)$ ($\mathscr{R}^* D = 1$, $\mathscr{R}_* D = 0$).

7.6. Riemann function. The *Riemann function* R is zero on the set of all irrational numbers. If $r = p/q$ is a rational number, p and q are relatively prime and $q > 0$, then $R(r)$ is defined as $1/q$ (zero is supposed to be $0/1$, so $R(0) = 1$). The Riemann function is continuous at x if and only if x is irrational. The Riemann function serves as an example of a "pathological" function which is Riemann but not Newton integrable (it has no antiderivative) on bounded intervals.

7.7. Various Examples. The indicator function of the Cantor ternary set is Riemann integrable. The indicator function of any discontinuum of a positive measure fails to be Riemann integrable. An unbounded Newton integrable function cannot be Riemann integrable. There are examples of bounded Newton integrable functions which are not Riemann integrable (cf. 7.9.f).

7.8. Chebyshev's Inequality for the Riemann Integral. If f is a nonnegative bounded function on (a,b), $\varepsilon > \mathscr{R}^* f$ and $\tau > 0$, then the set $\{f \geq \tau\}$ can be covered by a union of a finite number of intervals, the sum of whose lengths is less than ε/τ (so $\lambda^*\{f \geq \tau\} < \varepsilon/\tau$).

Hint. Find a partition $a = \xi_0 < \xi_1 < \cdots < \xi_m = b$ and numbers c_j such that $c_j \geq f$ on (ξ_{j-1}, ξ_j) and $\sum_{j=1}^{m} c_j(\xi_j - \xi_{j-1}) < \varepsilon$. Then the sum of lengths of those intervals (ξ_{j-1}, ξ_j) for which $c_j \geq \tau$ is less than ε/τ.

7.9. Riemann Integrable Functions. Lebesgue measure theory allows a deeper study of the class of Riemann integrable functions.

(a) For any bounded function f on an interval (a,b) we have

$$\mathscr{R}^* f = \inf\{\int_a^b g : g \text{ piecewise constant, } g \geq f\},$$

$$\mathscr{R}_* f = \sup\{\int_a^b g : g \text{ piecewise constant, } g \leq f\}.$$

(A function f is said to be *piecewise constant* on $[a,b]$ if there exist ξ_j such that $a = \xi_0 < \xi_1 < \cdots < \xi_m = b$ and f is constant on each interval (ξ_{j-1}, ξ_j).)

(b) If f is a bounded function on an interval (a,b), then the function

$$f^* := \inf\{g : g \geq f, \, g \text{ continuous}\}$$

is called the *upper Baire* function of f. The definition of the lower Baire function f_* of f is analogous.

Show that the function f^* is upper semi-continuous and that f is continuous at a point x if and only if $f^*(x) = f_*(x)$.

(c) Show that

$$\mathscr{R}^* f = \inf\{R\int_a^b g : g \geq f, \, g \text{ continuous }\}.$$

(d) Let f be a bounded function on an interval (a,b). Show that the following conditions are equivalent:

(i) f is Riemann integrable;
(ii) for each $\varepsilon > 0$ there exist continuous functions t, s such that $t \leq f \leq s$ and $R\int_a^b (s-t) < \varepsilon$;
(iii) $f_* = f^*$ almost everywhere;
(iv) there exist functions u, v; u upper semi-continuous, v lower semi-continuous such that $v \leq f \leq u$ and $u = v$ almost everywhere;
(v) the set of all points where f is discontinuous is of Lebesgue measure zero.

Hint. Use (b) and (c). For (ii) $\implies$ (iii), according to 7.8, $\lambda^*\{f^* - f_* \geq 1/k\} \leq \lambda^*\{s - t \geq 1/k\} \leq k\varepsilon$ for each $k \in \mathbf{N}$ and $\varepsilon > 0$.

(e) Show that any Riemann integrable function is measurable.

Hint. Any semi-continuous function is measurable. Now use (d) and Theorem 3.14.b.

(f) The above condition (v) permits to construct bounded Newton integrable functions which are not Riemann integrable. An example of Volterra type functions constructed with the aid of closed nowhere dense sets of positive measure (Example 1.13) can be found in A.M. Bruckner [*1978].

7.10. Notes. The origins of the integral calculus are connected with the names of I. Newton and G.W. Leibniz. The modern integration theory has been developed from the beginning of the 19$^{\text{th}}$ century by A.L. Cauchy, L. Dirichlet, B. Riemann, C. Jordan, E. Borel, H. Lebesgue, G. Vitali and others. Some historical treatments on integration are mentioned in 8.25.

8. The Abstract Lebesgue Integral

Elementary expositions of the integration theory in usual textbooks of calculus employ Riemann's or Newton's constructions. However, the examples given in Chapter 7 show that these integrals provide rather small classes of integrable functions. In deeper applications, we need completeness of normed linear spaces of integrable functions, which is achieved with the aid of Lebesgue's integration. A further anvantage of Lebesgue's approach consists in the possibility to integrate over more general domains than intervals.

In the sequel, $(X, \mathscr{S}, \mu)$ will be a measure space.

The Riemann integral of the indicator function of any interval is its length. Accordingly, developing the notion of an integral on X, the requirement that $\int_X c_A \, d\mu = \mu A$ for any $A \in \mathscr{S}$ is quite natural. Further, it is reasonable to impose conditions on additivity and monotonicity of the integral and to ask the family of all integrable functions to be as large as possible.

We introduce the concept of the abstract Lebesgue integral in several steps. First we define $\int_X f \, \mathrm{d}\mu$ in a natural way for nonnegative simple functions and then we extend it to nonnegative μ-measurable functions using their approximation by simple functions. The general case will be completed using the decomposition of a function into its positive and negative parts. Building up this theory, we must be careful in order to legitimate these steps. For instance, we have to show that the definition of $\int_X s \, \mathrm{d}\mu$ in the case of simple functions does not depend on their representation which is not unique. Likewise, we have to deal with problems when infinite values or "indefinite expressions" appear.

The reader may find it instructive to remember the most important example of the n-dimensional Lebesgue measure in $\mathbf{R}^n$. For integration with respect to the Lebesgue measure λ we use a traditional notation

$$\int_E f \, \mathrm{d}x := \int_E f \, \mathrm{d}\lambda \,, \; \int_a^b f \, \mathrm{d}x := \int_{(a,b)} f \, \mathrm{d}\lambda.$$

8.1. Simple Functions. Recall that a simple function is a real $\mathscr{S}$-measurable function s on X having a finite range. Any simple function s can be expressed as

$$s = \sum_{j=1}^{n} \beta_j c_{B_j} \,,$$

where $B_1, \dots, B_n \in \mathscr{S}$ are pairwise disjoint sets and $\beta_1, \dots, \beta_n \in \mathbf{R}$. Of course, this representation of s is not unique.

Remember again that as a matter of convenience we set $0 \cdot \infty = \infty \cdot 0 = 0$.

8.2. Lemma. *Let $A_1, \dots, A_m \in \mathscr{S}$ and $B_1, \dots, B_n \in \mathscr{S}$ be pairwise disjoint sets and let α_i and β_j be nonnegative real numbers. If $\sum_{i=1}^{m} \alpha_i c_{A_i} \le \sum_{j=1}^{n} \beta_j c_{B_j}$, then*

$$\sum_{i=1}^{m} \alpha_i \mu A_i \le \sum_{j=1}^{n} \beta_j \mu B_j .$$

Proof. To simplify the proof define $\alpha_0 = \beta_0 = 0$, $A_0 = X \setminus \bigcup_{i=1}^{m} A_i$ and $B_0 = X \setminus \bigcup_{j=1}^{n} B_j$ (then the collections $\{A_i\}$, $\{B_j\}$ form "partitions" of X). For $i \in \{0,\ldots,m\}$ and $j \in \{0,\ldots,n\}$ either $A_i \cap B_j = \emptyset$ or $\alpha_i \le \beta_j$. Thus

$$\sum_{i=0}^{m} \alpha_i \mu A_i \le \sum_{i=0}^{m}\sum_{j=0}^{n} \alpha_i \mu(A_i \cap B_j) \le \sum_{i=0}^{m}\sum_{j=0}^{n} \beta_j \mu(A_i \cap B_j) \le \sum_{j=0}^{n} \beta_j \mu B_j.$$

■

8.3. Abstract Lebesgue Integral. If $D \in \mathscr{S}$ and s is a nonnegative simple function expressed as $s = \sum_{j=1}^{n} \beta_j c_{B_j}$, where B_j are pairwise disjoint sets and β_j are nonnegative coefficients, define $\int_D s \,\mathrm{d}\mu := \sum_{j=1}^{n} \beta_j \mu(D \cap B_j)$. The previous lemma shows that this value does not depend on the particular representation of s. Next, define

$$\int_D f \,\mathrm{d}\mu := \sup\left\{\int_D s \,\mathrm{d}\mu : \ 0 \le s \le f \text{ on } D,\ s \text{ simple}\right\}$$

if $f \ge 0$ is a μ-measurable function on $D \in \mathscr{S}$. Lemma 8.2 again ensures that the "new" definition and the "old" one agree in case when f is a simple function.

For a $\mathscr{S}_D$-measurable function f on D we define $\int_D f \,\mathrm{d}\mu := \int_D f^+ \,\mathrm{d}\mu - \int_D f^- \,\mathrm{d}\mu$ provided at least one of these integrals is finite. Remember that $f^+ := \max(f, 0)$, $f^- := \max(-f, 0)$.

If f is a function on X and $M \in \mathscr{S}$, then apparently

$$\int_M f \,\mathrm{d}\mu = \int_X f \cdot c_M \,\mathrm{d}\mu$$

and $\int_M f \,\mathrm{d}\mu = \int_M f \,\mathrm{d}\nu$ where $\nu = \mu_{|M}$ is the restriction of μ to the σ-algebra $\mathscr{S}_M := \{A \in \mathscr{S} : \ A \subset M\}$ of subsets of M.

Therefore, it is no loss of generality to restrict our attention to the integration over the whole space X.

It is useful to define the abstract Lebesgue integral even for functions defined only μ-*almost everywhere*. In this case, if f is defined on $D \in \mathscr{S}$ and $\mu(X \setminus D) = 0$, set

$$\int_X f \,\mathrm{d}\mu := \int_D f \,\mathrm{d}\mu$$

if the integral on the right side is defined. It is immediate that this value does not depend on the choice of D.

The symbol $\mathscr{L}^*$ or $\mathscr{L}^*(\mu)$ denotes the family of all μ-measurable functions defined μ-almost everywhere on X for which the abstract Lebesgue integral is defined.

Further denote

$$\mathscr{L}^1 = \mathscr{L}^1(\mu) = \left\{ f \in \mathscr{L}^*(\mu) : \int_X f \,\mathrm{d}\mu \in \mathbf{R} \right\}.$$

If $f \in \mathscr{L}^1(\mu)$, we say that f is *μ-integrable.*

8.4. Lebesgue Integrable Functions on R. (a) Every bounded μ-measurable function on a bounded interval (and so every Riemann integrable function) is Lebesgue integrable. Thus, the functions from Examples 7.5 – 7.6 are Lebesgue integrable as well.

(b) If f is a Riemann integrable function on $[a, b]$, then $\int_a^b f \,\mathrm{d}\lambda \le \mathscr{R}^* f = \mathscr{R}_* f \le \int_a^b f \,\mathrm{d}\lambda$, so that the Lebesgue and the Riemann integrals of f coincide.

(c) If a function has both the Newton and the Lebesgue integrals, they are equal. Proof of this is not easy. We will prove it using the Henstock-Kurzweil integral in Chapter 25.

(d) If f is Newton integrable, then f is μ-measurable since it can be expressed as the limit of a sequence $\{f_n\}$ of continuous functions, in essence, $f_n(x) = n(F(x + \frac{1}{n}) - F(x))$, where F is an antiderivative to f. It can happen that f is not Lebesgue integrable, but this is only the case if f is not absolutely Newton integrable.

(e) Since nonmeasurable functions are rather rare, the question whether or not f is Lebesgue integrable can be reduced to the question whether the integrals $\int_a^b f^+ \,\mathrm{d}x$ and $\int_a^b f^- \,\mathrm{d}x$ are finite or infinite. For instance, the function $f : x \mapsto x^p$ is integrable on $(0, 1)$ if and only if $p > -1$.

The key result is the following monotone convergence theorem (sometimes called Levi's theorem or the Lebesgue monotone convergence theorem) for non-negative functions.

8.5. Theorem. *If $f_n \ge 0$ are μ-measurable functions on X, $f_n \nearrow f$, then $\int_X f_n \,\mathrm{d}\mu \to \int_X f \,\mathrm{d}\mu$.*

Proof. It is clear that the sequence $\left\{\int_X f_n \,\mathrm{d}\mu\right\}$ is nondecreasing, and therefore there exists $\alpha := \lim \int_X f_n \,\mathrm{d}\mu$. Since $f_n \le f$, we have $\alpha \le \int_X f \,\mathrm{d}\mu$ and the assertion is obvious for $\alpha = +\infty$. Assume $\alpha < +\infty$, fix a simple function s, $0 \le s \le f$, and prove that $\int_X s \,\mathrm{d}\mu \le \alpha$. The proof will be given in several steps:

(a) Let $\tau \in (0, 1)$. Define $E_n = \{x \in X : f_n(x) \ge \tau s(x)\}$. Then $E_n \in \mathscr{S}$, $E_n \subseteq E_{n+1}$ and $\bigcup_{n=1}^{\infty} E_n = X$ (if $f(x) = 0$, then $x \in E_1$; if $f(x) > 0$, then $\tau s(x) < \lim f_n(x)$). Thus $\mu(E_n \cap A) \to \mu A$ for every set $A \in \mathscr{S}$.

(b) Let $s = \sum_{j=1}^{k} \beta_j c_{A_j}$, where A_j are pairwise disjoint. Then

$$\int_X f_n \,\mathrm{d}\mu \ge \int_{E_n} f_n \,\mathrm{d}\mu \ge \int_{E_n} \tau s \,\mathrm{d}\mu = \tau \int_{E_n} (\sum_{j=1}^{k} \beta_j c_{A_j}) \,\mathrm{d}\mu = \tau \sum_{j=1}^{k} \beta_j \mu(A_j \cap E_n).$$

(c) Passing to limits in (b) (using (a)), we get

$$\alpha \ge \tau \sum_{j=1}^{k} \beta_j \mu A_j = \tau \int_X s \,\mathrm{d}\mu\,.$$

Since $\tau \in (0, 1)$ is arbitrary, we get $\alpha \ge \int_X s \,\mathrm{d}\mu$, as needed. ∎

8.6. Theorem. *Let $g \in \mathscr{L}^*$ and f be a μ-measurable function, $f = g$ almost everywhere. Then $f \in \mathscr{L}^*$ and $\int_X f \, \mathrm{d}\mu = \int_X g \, \mathrm{d}\mu$.*

Proof. Obvious. ∎

8.7. Remark. For any nonnegative μ-measurable function f on X there exists a sequence of simple functions $s_n \geq 0$ such that $s_n \nearrow f$. From Theorem 8.5 we know that $\int_X f \, \mathrm{d}\mu = \lim \int_X s_n \, \mathrm{d}\mu$. As is often the case, this observation becomes the basis for a definition of the integral of nonnegative μ-measurable functions. Of course, starting with such a definition, it is necessary to prove the independence of $\int_X f \, \mathrm{d}\mu$ on the choice of the sequence $\{s_n\}$.

8.8. Theorem. *If f_1, f_2 are nonnegative μ-measurable functions, then*

$$\int_X (f_1 + f_2) \, \mathrm{d}\mu = \int_X f_1 \, \mathrm{d}\mu + \int_X f_2 \, \mathrm{d}\mu \, .$$

Proof. It might be assumed that both f_1 and f_2 are defined everywhere on X. First suppose that f_1 and f_2 are even simple. Following the same line of proof as in Lemma 8.2, we can find pairwise disjoint $\mathscr{S}$-measurable sets $A_0, \dots, A_m$ and $B_0, \dots, B_n$ and nonnegative real numbers $\alpha_0, \dots, \alpha_m$ and $\beta_0, \dots, \beta_n$ such that $X = \bigcup\limits_{i=0}^{m} A_i = \bigcup\limits_{j=0}^{n} B_j$, $f_1 = \sum\limits_{i=0}^{m} \alpha_i c_{A_i}$ and $f_2 = \sum\limits_{j=0}^{n} \beta_j c_{B_j}$. Then

$$\begin{aligned} \int_X (f_1 + f_2) \, \mathrm{d}\mu &= \sum_{i=0}^{m} \sum_{j=0}^{n} (\alpha_i + \beta_j) \mu(A_i \cap B_j) = \sum_{i=0}^{m} \alpha_i \mu A_i + \sum_{j=0}^{n} \beta_j \mu B_j \\ &= \int_X f_1 \, \mathrm{d}\mu + \int_X f_2 \, \mathrm{d}\mu. \end{aligned}$$

For the general case, find sequences $\{s_n^1\}$ and $\{s_n^2\}$ of simple functions with $s_n^i \nearrow f_i$ and use the first part and Theorem 8.5. ∎

8.9. Theorem (properties of $\mathscr{L}^1(\mu)$). *The following propositions hold:*

(a) *If $f \in \mathscr{L}^1$, then f is finite almost everywhere.*
(b) *If $f, g \in \mathscr{L}^1$, $\alpha, \beta \in \mathbf{R}$, then $\alpha f + \beta g \in \mathscr{L}^1$ and $\int (\alpha f + \beta g) \, \mathrm{d}\mu = \alpha \int_X f \, \mathrm{d}\mu + \beta \int_X g \, \mathrm{d}\mu$ (an appeal to (a) reveals that $\alpha f + \beta g$ is defined almost everywhere).*
(c) *If $f \in \mathscr{L}^1$, then $|f| \in \mathscr{L}^1$ and $\left|\int_X f \, \mathrm{d}\mu\right| \leq \int_X |f| \, \mathrm{d}\mu$ (i.e. the integral is "absolutely convergent").*
(d) *$\mathscr{L}^1$ is a lattice (if $f, g \in \mathscr{L}^1$, then* $\max(f, g), \min(f, g) \in \mathscr{L}^1$*).*
(e) *If f is μ-measurable, $g \in \mathscr{L}^1$, $|f| \leq g$, then $f \in \mathscr{L}^1$.*

Proof. (a) Suppose that $f \in \mathscr{L}^1$, $A := \{x \in X : f(x) = \infty\}$. Then A is measurable and $0 \leq n c_A \leq f^+$ for every $n \in \mathbf{N}$. Thus $0 \leq \int_X n c_A \, \mathrm{d}\mu \leq \int_X f^+ \, \mathrm{d}\mu$ and we get $\mu A \leq \frac{1}{n} \int_X f^+ \, \mathrm{d}\mu < \infty$ for every $n \in \mathbf{N}$. Hence $\mu A = 0$.

(b) It is plain to see that $\int_X \alpha f \, \mathrm{d}\mu = \alpha \int_X f \, \mathrm{d}\mu$ for $f \in \mathscr{L}^1$ and $\alpha \in \mathbf{R}$. Let $f, g \in \mathscr{L}^1$ and write $f = f^+ - f^-$, $g = g^+ - g^-$ and $h = f + g$ (by (a), h is defined almost everywhere). Then

$$h^+ - h^- = f^+ - f^- + g^+ - g^-,$$

and in light of Theorem 8.8,

$$\int_X h^+ \, \mathrm{d}\mu + \int_X f^- \, \mathrm{d}\mu + \int_X g^- \, \mathrm{d}\mu = \int_X f^+ \, \mathrm{d}\mu + \int_X g^+ \, \mathrm{d}\mu + \int_X h^- \, \mathrm{d}\mu \, .$$

The assertion now follows if we observe that the integrals $\int_X h^+ \, \mathrm{d}\mu$ and $\int_X h^- \, \mathrm{d}\mu$ are finite. But this is true as

$$0 \le h^+ = (f+g)^+ \le f^+ + g^+ .$$

(c) If $f \in \mathscr{L}^1$, then $|f| = f^+ + f^- \in \mathscr{L}^1$ by (b), and we have

$$\left| \int_X f \, \mathrm{d}\mu \right| = \left| \int_X f^+ \, \mathrm{d}\mu - \int_X f^- \, \mathrm{d}\mu \right| \le \left| \int_X f^+ \, \mathrm{d}\mu \right| + \left| \int_X f^- \, \mathrm{d}\mu \right| =$$

$$= \int_X f^+ \, \mathrm{d}\mu + \int_X f^- \, \mathrm{d}\mu = \int_X |f| \, \mathrm{d}\mu.$$

(d) The assertion follows immediately from (b) and (c) using

$$\max(f,g) = \frac{1}{2}(f + g + |f+g|).$$

(e) If f is a μ-measurable function, f^+ is μ-measurable as well and

$$0 \le \int_X f^+ \, \mathrm{d}\mu \le \int_X g \, \mathrm{d}\mu < \infty$$

($0 \le f^+ \le |f| \le g$), thus $f^+ \in \mathscr{L}^1$. Analogously $f^- \in \mathscr{L}^1$, and finally $f = f^+ - f^- \in \mathscr{L}^1$. ■

8.10. Remarks. 1. Proposition (b) of the previous theorem shows that $\mathscr{L}^1$ satisfies almost all axioms for a linear space but it is not a true linear space. To get a linear space, the set of all finite everywhere defined functions from $\mathscr{L}^1$ or the space L^1 (which will be introduced in Chapter 10) is to be considered.

2. The sum of an absolutely convergent series $\sum_j a_j$ is the integral of the function $j \mapsto a_j$ over the set $\mathbf{N}$ with respect to the counting measure. On the other hand, the abstract integration theory cannot describe sums of nonabsolutely convergent series. This observation should not be understood as an insufficiency of Lebesgue's theory. The abstract Lebesgue integral described in 8.3 provides the best approach in the general framework of measure spaces. Let us consider a simple experiment: A rearangement of $\mathbf{N}$ can change the sum of a nonabsolutely convergent series while "measures of sets" remain invariant. The sum depends on an additional structure on $\mathbf{N}$, namely, on its ordering. Analogously, taking into account the ordering of the real line (in addition to its measure properties), various kinds of nonabsolutely convergent integrals may be introduced. Let us mention Newton integration on an elementary level and Perron or Henstock-Kurzweil integration on an advanced level (see Section 25). It is instructive to compare the expressions

$$\int_1^\infty \frac{\sin x}{x} \, \mathrm{d}x = \lim_{b \to \infty} \int_1^b \frac{\sin x}{x} \, \mathrm{d}x$$

(where the integral is understood as Newtonian) and

$$\sum_{k=1}^\infty \frac{\sin k}{k} = \lim_{n \to \infty} \sum_{k=1}^n \frac{\sin k}{k}.$$

In the theory of Lebesgue integration, limit theorems play a crucial role. They concern either monotone convergence or dominated convergence.

8.11. Levi's theorem. *Let $\{f_n\}$ be a sequence of μ-measurable functions, $f_n \nearrow f$ almost everywhere and let $\int_X f_1 \,\mathrm{d}\mu > -\infty$. Then $\int_X f \,\mathrm{d}\mu = \lim \int_X f_n \,\mathrm{d}\mu$.*

Proof. We have already proved the theorem when f_n are nonnegative and $f_n \nearrow f$ everywhere. The general case can easily be reduced to Theorem 8.5. First we redefine f and f_n on sets of measure zero in such a way that $f_n \nearrow f$ everywhere. It would be clearly sufficient to assume that $f_n \in \mathscr{L}^1$ for all n. If $g_n = f_n + f_1^-$ (notice that $f_1^- \in \mathscr{L}^1$), then g_n are nonnegative μ-measurable functions and $g_n \nearrow f + f_1^-$. Now, Theorem 8.5 ensures that $\int_X g_n \,\mathrm{d}\mu \to \int_X (f + f_1^-) \,\mathrm{d}\mu$. Since $\int_X f_n \,\mathrm{d}\mu = \int_X g_n \,\mathrm{d}\mu - \int_X f_1^- \,\mathrm{d}\mu$, the assertion easily follows. ■

8.12. Levi's theorem for series. *If f_n are nonnegative μ-measurable functions on X, then $\int_X \sum f_n \,\mathrm{d}\mu = \sum \int_X f_n \,\mathrm{d}\mu$.*

Proof. Use Theorems 8.5 and 8.8. ■

8.13. Lebesgue dominated convergence theorem. *Let $\{f_n\}$ be a sequence of μ-measurable functions, $f_n \to f$ almost everywhere. If there exists a function $h \in \mathscr{L}^1$ such that $|f_n| \le h$ almost everywhere for all n, then $f \in \mathscr{L}^1$ and $\int_X f \,\mathrm{d}\mu = \lim \int_X f_n \,\mathrm{d}\mu$.*

Proof. The proof may be easily reduced to Levi's theorem by considering $f = \limsup f_n = \liminf f_n$. Set $s_n = \sup\{f_n, f_{n+1}, \dots\}$, $t_n = \inf\{f_n, f_{n+1}, \dots\}$. Then $-h \le t_n \le f_n \le s_n \le h$, $s_n \searrow f$, $t_n \nearrow f$ almost everywhere and

$$-\infty < \int_X -h \,\mathrm{d}\mu \le \int_X t_1 \,\mathrm{d}\mu \le \int_X s_1 \,\mathrm{d}\mu \le \int_X h \,\mathrm{d}\mu < +\infty.$$

By Levi's theorem, $\int_X f \,\mathrm{d}\mu = \lim \int_X t_n \,\mathrm{d}\mu = \lim \int_X s_n \,\mathrm{d}\mu$. Since $\int_X t_n \,\mathrm{d}\mu \le \int_X f_n \,\mathrm{d}\mu \le \int_X s_n \,\mathrm{d}\mu$, the assertion follows. ■

8.14. Dominated convergence theorem for series. *Let $\{h_n\}$ be a sequence of μ-measurable functions on X and $g \in \mathscr{L}^1$. Suppose that*

$$\left|\sum_{j=1}^{n} h_j\right| \le g \qquad \text{almost everywhere}$$

for all $n \in \mathbf{N}$ and that the series $\sum\limits_{j=1}^{\infty} h_j$ converges almost everywhere. Then

$$\int_X \sum_{j=1}^{\infty} h_j \,\mathrm{d}\mu = \sum_{j=1}^{\infty} \int_X h_j \,\mathrm{d}\mu.$$

Proof. The theorem follows immediately from the previous one. ■

8.15. Fatou's lemma. *Let $\{f_n\}$ be a sequence of μ-measurable functions and $g \in \mathscr{L}^1$. If $f_n \ge g$ almost everywhere for all $n \in \mathbf{N}$, then*

$$\int_X \liminf f_n \,\mathrm{d}\mu \le \liminf \int_X f_n \,\mathrm{d}\mu.$$

Proof. Set $g_k = \inf\{f_n : k \le n\}$. Then $g_k \nearrow \liminf f_n$ and infer on Levi's theorem again. ■

8.16. Theorem. *Let $f \geq 0$ be a μ-measurable function. If $\int_X f \,\mathrm{d}\mu = 0$, then $f = 0$ almost everywhere.*

Proof. Denote $A_n = \{x \in X : g(x) \geq \frac{1}{n}\}$. Then $\bigcup_{n=1}^{\infty} A_n = \{x \in X : f(x) > 0\}$ and the inequality $0 \leq c_{A_n} \leq nf$ yields $\mu(A_n) = 0$. ■

8.17. Theorem. *Let $f \in \mathscr{L}^1$. If $\int_E f \,\mathrm{d}\mu = 0$ for any measurable set E, then $f = 0$ almost everywhere.*

Proof. If $E := \{f \geq 0\}$, then $\int_E f \,\mathrm{d}\mu = \int_E f^+ \,\mathrm{d}\mu = 0$ and using the previous theorem we get $f^+ = 0$ almost everywhere. Analogously $f^- = 0$ almost everywhere. ■

8.18. Corollary. *Let $f, g \in \mathscr{L}^1$. If*

$$\int_E f \,\mathrm{d}\mu \leq \int_E g \,\mathrm{d}\mu$$

for every measurable set E, then $f \leq g$ almost everywhere.

Proof. Set $h = (f-g)^+$. Then $h \geq 0$ and $0 \leq \int_E h \,\mathrm{d}\mu = \int_{E \cap \{h>0\}} (f-g) \,\mathrm{d}\mu \leq 0$. Thanks to Theorem 8.17, $(f-g)^+ = 0$ almost everywhere. ■

8.19. Indefinite Lebesgue integral. Let $f \in \mathscr{L}^*$. For $E \in \mathscr{S}$ set

$$\mu_f(E) := \int_E f \,\mathrm{d}\mu \,.$$

The set function μ_f is called the *indefinite Lebesgue integral of f*.

8.20. Theorem. *Let $f \in \mathscr{L}^*$. Then μ_f is a signed measure on X. Moreover, $\mu_f(E) = 0$ for every set $E \in \mathscr{S}$ with $\mu E = 0$.*

Proof. It is enough to prove that μ_f is a measure provided f is nonnegative. In this case, notice that $f c_A = \sum_n f c_{A_n}$ if $A = \bigcup_n A_n$ with $A_i \cap A_j = \emptyset$ for $i \neq j$, and cite Theorem 8.12. ■

8.21. Remark. A question arises whether or not for any pair of measures on $\mathscr{S}$, say μ and ν, there is always a nonnegative μ-measurable function f on X such that $\nu = \mu_f$, i.e.

$$\nu E = \int_E f \,\mathrm{d}\mu$$

for all $E \in \mathscr{S}$. Of course, the answer is negative; if such a function exists, then necessarily $\nu E = 0$ whenever $\mu E = 0$. But if μ and ν obey this condition (we say that ν is absolutely continuous with respect to μ), then there is a function f with the ascribed property (at least for σ-finite measures). This will be the subject of the chapter concerning the Radon-Nikodým theorem.

8.22. Exercises. Let $f \in \mathscr{L}^1$. Prove that

(a) μ_f is a finite signed measure on $\mathscr{S}$;

(b) for every $\varepsilon > 0$ there is a $\delta > 0$ such that $|\mu_f(E)| < \varepsilon$ whenever $\mu E < \delta$.

Hint. (a) To verify the equality $\mu_f(A) = \sum \mu_f(A_n)$ ($A_n \in \mathscr{S}$ pairwise disjoint) use Theorem 8.14 for $h_n = f c_{A_n}$.

(b) Suppose there exists an $\varepsilon > 0$ and a sequence $\{E_n\}$ such that $\mu E_n < 2^{-n}$ and $\int_{E_n} |f| \geq \varepsilon$. Set $E = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$. Then $\mu E = 0$ and $\int_E f \,\mathrm{d}\mu \geq \varepsilon$, which is a contradiction.

8.23. Image of a Measure. Let $(X, \mathscr{S}, \mu)$ be a measure space, $(Y, \mathscr{T})$ be a measurable space and $f : X \to Y$ a measurable mapping (i.e. $f^{-1}(E) \in \mathscr{S}$ whenever $E \in \mathscr{T}$). The set function

$$E \mapsto \mu(f^{-1}(E)), \qquad E \in \mathscr{T},$$

is called the *image of the measure* μ under the mapping f and it is denoted by $f(\mu)$.

(a) Show that $f(\mu)$ is a measure on $(Y, \mathscr{T})$.

(b) Let $g : Y \to \overline{\mathbf{R}}$ be a $\mathscr{T}$-measurable function. Show that g is $f(\mu)$-integrable if and only if $g \circ f \in \mathscr{L}^1$. In this case

$$\int_Y g \, \mathrm{d}f(\mu) = \int_X g \circ f \, \mathrm{d}\mu.$$

Hint. First consider simple functions and then pass to limits.

8.24. Exercise. Let f be an increasing differentiable function on $\mathbf{R}$, $\lim_{x \to \pm\infty} f(x) = \pm\infty$. Let μ be a measure on $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$ defined by

$$\mu E = \int_E f' \, \mathrm{d}x.$$

Show that $f(\mu) = \lambda$.

8.25. Notes. The modern integration theory starts with Lebesgue's doctoral thesis [1902] following a short paper [1901]. His definition uses constructed Lebesgue measure on $\mathbf{R}$. The idea of using simple functions (not equal to those introduced here) when defining the integral (still on the real line) comes from F. Riesz ([1912], [1920]).

Further details can be found in historical notes by T. Hawkins [*1970] or by F.A. Medvedev [1975], R. Henstock [1988], T.H. Hildebrandt [1953]. It is clear that H. Lebesgue followed investigations of his predecessors (B. Riemann, C. Jordan, G. Peano, E. Borel and others). Less known is the fact that (roughly) at the same time G. Vitali and W.H. Young published similar results independently.

Theorem 8.11 (concerning the monotone convergence) was proved by Beppo Levi [1906], Lemma 8.15 by P.J.L. Fatou [1906] and Theorem 8.13. by H. Lebesgue [1910].

One more remark: H. Lebesgue developed his theory of integration mainly for the case of "Lebesgue measures" on $\mathbf{R}^n$. Later J. Radon [1913] considered more general measures in $\mathbf{R}^n$ (nowadays called the Radon measures). A general theory of measures on arbitrary σ-algebras was given by M. Fréchet [1915]. Since that time, many results on this subject have been published; the monograph [*1950] by P.R. Halmos is one of the most quoted.

9. Integrals Depending on a Parameter

The Lebesgue dominated convergence theorem has simple but very important consequences on continuity and differentiation of integrals depending on a parameter.

9.1. Theorem. *Let $(X, \mathscr{S}, \mu)$ be a measure space, P a metric space and U a neighbourhood of a point $a \in P$. Suppose that a function $F : U \times X \to \mathbf{R}$ has the following properties:*

(a) *there exists a set $N \subset X$ of measure zero such that for each $x \in X \setminus N$ the function $F(\cdot, x)$ is continuous at a;*
(b) *for each $t \in U$, the function $F(t, \cdot)$ is μ-measurable;*
(c) *there exists a function $g \in \mathscr{L}^1(X)$ such that $|F(t, \cdot)| \le g$ almost everywhere for all $t \in U$.*

Then for each $t \in U$, $F(t,\cdot) \in \mathscr{L}^1(X)$ and the function

$$f : t \mapsto \int_X F(t,\cdot)\,\mathrm{d}\mu$$

is continuous at a.

Proof. The proof that

$$\lim_{t\to a} \int_X F(t,\cdot)\,\mathrm{d}\mu = \int_X F(a,\cdot)\,\mathrm{d}\mu$$

is achieved by showing that

$$\lim_j \int_X F(t_j,\cdot)\,\mathrm{d}\mu = \int_X F(a,\cdot)\,\mathrm{d}\mu$$

for each sequence $t_j \to a$, $t_j \in U$. To prove the last assertion it suffices to use the Lebesgue dominated convergence theorem. ■

9.2. Theorem. *Let $(X, \mathscr{S}, \mu)$ be a measure space, $N \subset X$ a set of measure zero and $I \subset \mathbf{R}$ an open interval. Suppose that a function $F : I \times X \to \mathbf{R}$ has the following properties:*

(a) *For each $x \in X \setminus N$, $F(\cdot, x)$ is differentiable on I;*
(b) *for each $t \in I$, $F(t,\cdot)$ is μ-measurable;*
(c) *there exists a function $g \in \mathscr{L}^1(X)$ such that $\left|\frac{\mathrm{d}}{\mathrm{d}t}F(t,x)\right| \le g(x)$ for each $x \in X \setminus N$ and $t \in I$;*
(d) *there is a $t_0 \in I$ such that $F(t_0,\cdot) \in \mathscr{L}^1(X)$.*

Then $F(t,\cdot) \in \mathscr{L}^1(X)$ for all $t \in I$, the function

$$f : t \mapsto \int_X F(t,\cdot)\,\mathrm{d}\mu$$

is differentiable on I and

$$f'(t) = \int_X \frac{\mathrm{d}}{\mathrm{d}t} F(t,\cdot)\,\mathrm{d}\mu.$$

Proof. Suppose $a, b \in I$, $b \neq a$, $x \in X \setminus N$. By the mean value theorem there exists a ξ between a and b such that

$$\left|\frac{F(b,x) - F(a,x)}{b-a}\right| = \left|\frac{\mathrm{d}}{\mathrm{d}t}F(\xi,x)\right| \le g(x).$$

For $a = t_0$ it follows that the function

$$x \mapsto \frac{F(b,x) - F(a,x)}{b-a}$$

is integrable, thus $F(b,\cdot)$ is integrable. Choose $a \in I$ again. By the Lebesgue dominated convergence theorem,

$$\lim_j \int_X \frac{F(t_j,\cdot) - F(a,\cdot)}{t_j - a}\,\mathrm{d}\mu = \int_X \frac{\mathrm{d}}{\mathrm{d}t} F(t,\cdot)\,\mathrm{d}\mu$$

for each sequence $t_j \to a$ of points of $I \setminus \{a\}$, which yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_X F(a,\cdot)\,\mathrm{d}\mu = \lim_{t\to a}\int_X \frac{F(t,\cdot) - F(a,\cdot)}{t - a}\,\mathrm{d}\mu = \int_X \frac{\mathrm{d}}{\mathrm{d}t} F(a,\cdot)\,\mathrm{d}\mu.$$

■

9.3 Remarks. 1. Notice that in the proofs of the last theorems we made heavy use of the Lebesgue dominated convergence theorem. Of course we could state analogous results based on Levi's theorem.

2. Since the notions of continuity and derivative are local, it suffices to verify the assumption (c) only locally.

3. A number of exercises and clarifying examples can be found in [L-Př].

10. The L^p Spaces

The Lebesgue spaces L^p are important means linking measure theory and functional analysis, and they are essential tools in differential equations theory, probability theory and other branches of modern analysis.

In this chapter, $(X, \mathscr{S}, \mu)$ denotes a fixed measure space and p a number from the interval $[1, \infty]$.

10.1. The Set $\mathscr{L}^p$. (a) Suppose $p < \infty$. We denote by $\mathscr{L}^p = \mathscr{L}^p(X, \mathscr{S}, \mu)$ the set of all μ-measurable functions on X such that

$$\int_X |f|^p \,\mathrm{d}\mu < \infty.$$

The value

$$\|f\|_p := \left(\int |f|^p \,\mathrm{d}\mu\right)^{1/p}$$

is called the L^p*-norm* of a function $f \in \mathscr{L}^p$. We will show soon that $\|\cdot\|_p$ has all important properties of a norm.

(b) We denote by $\mathscr{L}^\infty = \mathscr{L}^\infty(X, \mathscr{S}, \mu)$ the set of all μ-measurable functions f on X such that $|f| \le M$ μ-almost everywhere for some constant M. The least constant M having this property is called the L^∞*-norm* of f and is denoted by $\|f\|_\infty$. Roughly speaking, the difference between $\|f\|_\infty$ and $\sup_X |f|$ is that $\|f\|_\infty$ omits values of f on null sets.

Sometimes it is useful to emphasize the measure or the space X. In this case, abbreviated symbols $\mathscr{L}^p(X)$ or $\mathscr{L}^p(\mu)$ will be used as well.

10.2. Young's inequality. *Let $p, q \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$. If a, b are nonnegative numbers, then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof. We can assume $ab > 0$. Making use of concavity of the function ln,

$$\ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \geq \frac{1}{p}\ln(a^p) + \frac{1}{q}\ln(b^q) = \ln a + \ln b = \ln(ab)$$

holds, and we easily establish the required inequality. ■

10.3. Hölder's inequality. *Suppose that $f \in \mathscr{L}^p$ and $g \in \mathscr{L}^q$, where $p, q \in (1, \infty)$, $\frac{1}{p} + \frac{1}{q} = 1$. Then $fg \in \mathscr{L}^1$ and*

$$\left|\int_X fg \, \mathrm{d}\mu\right| \leq \left(\int_X |f|^p \, \mathrm{d}\mu\right)^{1/p} \left(\int_X |g|^q \, \mathrm{d}\mu\right)^{1/q}.$$

Proof. Denote

$$s = \left(\int_X |f|^p \, \mathrm{d}\mu\right)^{1/p}, \qquad t = \left(\int_X |g|^q \, \mathrm{d}\mu\right)^{1/q}.$$

We may assume that $st > 0$. Thanks to Young's inequality ($a = f(x)/s$, $b = g(x)/t$) we have

$$\frac{f(x)g(x)}{st} \leq \frac{|f(x)|\,|g(x)|}{st} \leq \frac{|f(x)|^p}{ps^p} + \frac{|g(x)|^q}{qt^q}$$

for each $x \in X$. Thus

$$\frac{1}{st}\int_X fg \, \mathrm{d}\mu \leq \frac{\int_X |f|^p \, \mathrm{d}\mu}{ps^p} + \frac{\int_X |g|^q \, \mathrm{d}\mu}{qt^q} \leq \frac{1}{p} + \frac{1}{q} = 1$$

which is what we wanted to prove. ■

10.4. Minkowski's inequality. *Let $p \in [1, \infty]$ and $f, g \in \mathscr{L}^p$. Then $f + g \in \mathscr{L}^p$ and*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p .$$

Proof. It is not hard to verify that $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$.
If $p = \infty$, then $|f| \leq s$ μ-almost everywhere and $|g| \leq t$ μ-almost everywhere, which implies $|f + g| \leq s + t$ μ-almost everywhere. Therefore $\|f + g\|_\infty \leq s + t$, and consequently

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty .$$

With these trivial cases out of the way, there remains the case $1 < p < \infty$. Hölder's inequality yields

$$\begin{aligned}\int_X |f|\,|f+g|^{p-1}\,\mathrm{d}\mu &\le \left(\int_X |f|^p\,\mathrm{d}\mu\right)^{1/p}\left(\int_X |f+g|^{(p-1)q}\,\mathrm{d}\mu\right)^{1/q}\\ &= \left(\int_X |f|^p\,\mathrm{d}\mu\right)^{1/p}\left(\int_X |f+g|^p\,\mathrm{d}\mu\right)^{1-1/p}.\end{aligned}$$

Analogously

$$\int_X |g|\,|f+g|^{p-1}\,\mathrm{d}\mu \le \left(\int_X |g|^p\,\mathrm{d}\mu\right)^{1/p}\left(\int_X |f+g|^p\,\mathrm{d}\mu\right)^{1-1/p}.$$

This entails

$$\begin{aligned}\int_X |f+g|^p\,\mathrm{d}\mu &\le \int_X |f|\,|f+g|^{p-1}\,\mathrm{d}\mu + \int_X |g|\,|f+g|^{p-1}\,\mathrm{d}\mu\\ &\le \left(\left(\int_X |f|^p\right)^{1/p} + \left(\int_X |g|^p\right)^{1/p}\right)\left(\int_X |f+g|^p\,\mathrm{d}\mu\right)^{1-1/p}.\end{aligned}$$

■

10.5. The L^p Spaces. The behavior of the function $f \mapsto \|f\|_p$ on $\mathscr{L}^p$, where $p \in [1, \infty]$, resembles axioms of a norm. However, in general, $\mathscr{L}^p$ is not a linear space and a nonzero function may have zero norm. To apply the theory of normed linear spaces, we identify functions which are equal almost everywhere. Formally, we assign to every function $f \in \mathscr{L}^p$ the class of functions

$$[f] = \{g \in \mathscr{L}^p : g = f \ \mu\text{-almost everywhere on } X\}$$

and define

$$L^p = L^p(X, \mathscr{S}, \mu) = \{[f] : f \in \mathscr{L}^p\}.$$

Then L^p is a linear space equipped with operations

$$[f] + [g] := [f + g], \quad \alpha[f] := [\alpha f] \quad (\alpha \in \mathbf{R}),$$

and with the (true) norm

$$\|\,[f]\,\|_p := \|f\|_p\,.$$

It can easily be seen that these definitions do not depend on the choice of representatives.

It is customary not to distinguish between functions and classes, often even between the spaces $\mathscr{L}^p$ and L^p. For instance, we say that $\{f_j\}$ is a Cauchy sequence in L^p while the meaning is that f_j are functions and $\{[f_j]\}$ is a Cauchy sequence in L^p.

10.6. Completeness of L^p. *Let $\{f_j\}$ be a Cauchy sequence in L^p. Then $\{f_j\}$ is convergent in L^p, i.e. there exists an $f \in \mathscr{L}^p$ such that*

$$\int_X |f - f_j|^p \, \mathrm{d}\mu \to 0.$$

Proof. The easy case $p = \infty$ is left to the reader. Suppose $p < \infty$. First we choose a subsequence $\{g_j\}$ of $\{f_i\}$ such that

$$S := \sum_{j=1}^{\infty} \|g_j - g_{j+1}\|_p < \infty.$$

Denote $h = \sum_{j=1}^{\infty} |g_j - g_{j+1}|$. Then by Levi's Theorem 8.12 and Minkowski's inequality

$$\int_X h^p \, \mathrm{d}\mu = \int_X \left(\sum_{j=1}^{\infty} |g_j - g_{j+1}| \right)^p \mathrm{d}\mu = \lim_{k\to\infty} \int_X \left(\sum_{j=1}^{k} |g_j - g_{j+1}| \right)^p$$

$$\leq \lim_{k\to\infty} \left(\sum_{j=1}^{k} \|g_j - g_{j+1}\|_p \right)^p \leq S^p < \infty.$$

By Theorem 8.9.a there exists a set $M \subset X$ such that $\mu(X \setminus M) = 0$ and $h < \infty$ on M. We can also assume that $|g_1| < \infty$ on M. For each $x \in M$, $\{g_j(x)\}$ is a Cauchy sequence in $\mathbf{R}$, and therefore it is convergent. Thus we can define

$$f(x) = \lim_{j\to\infty} g_j(x)$$

μ-almost everywhere. We prove that $f \in \mathscr{L}^p$ and $\|f - f_j\|_p \to 0$. The sequence $\{|g_j|^p\}$ tends to $\{|f|^p\}$ μ-almost everywhere and

$$|g_j|^p \leq (|g_j - g_1| + |g_1|)^p \leq (h + |g_1|)^p$$

for all $k \in \mathbf{N}$. An appeal to the Lebesgue dominated convergence theorem 8.13 with dominating function $(h + |g_1|)^p$ yields

$$\int_X |f|^p \, \mathrm{d}\mu = \lim_{j\to\infty} \int_X |g_j|^p \, \mathrm{d}x \leq \int_X (h + g_1)^p \, \mathrm{d}\mu < \infty.$$

We see that f is an element of $\mathscr{L}^p$. In a similar way by Lebesgue's theorem (dominating function h^p)

$$\|f - g_j\|_p^p = \int_X |f - g_j|^p \, \mathrm{d}\mu \to 0.$$

Since the sequence $\{f_j\}$ is Cauchy in L^p, we get $\lim_{j\to\infty} \|f_j - g_j\|_p = 0$. Thus $\lim_{j\to\infty} \|f_j - f\|_p = 0$ and f is a limit of the sequence $\{f_j\}$ in L^p. ∎

10.7. Remarks. 1. In the language of functional analysis we have just proved that L^p are *Banach spaces*. The characterization of their duals, which is significant for the theory, is postponed to Exercise 13.17.

2. For a counting measure μ on a set X we write $l^p(X) := L^p(X, \mathscr{P}(X), \mu)$. In particular, for $X = \mathbf{N}$ we get well known spaces of sequences:

- the space ℓ^p, $1 \le p < \infty$ of all sequences $x = \{x_n\}$ such that $\|x\|_p := \left(\sum_n |x_n|^p\right)^{1/p} < \infty$,
- the space ℓ^∞ of all bounded sequences $x = \{x_n\}$ with the norm $\|x\|_\infty := \sup_n |x_n|$.

10.8. Exercise. Suppose $f_n, f \in \mathscr{L}^\infty$. Show that $\|f_n - f\|_\infty \to 0$ if and only if there exists a set $E \in \mathscr{S}$, $\mu E = 0$, such that $f_n \rightrightarrows f$ on $X \setminus E$.

10.9. Exercise. Let $p \in [1, +\infty)$ and $\{f_n\}$ be a Cauchy sequence in $\mathscr{L}^p$. A close inspection of the proof of 10.6 shows that there is a subsequence of $\{f_n\}$ that converges μ-almost everywhere in X. Compare with Theorem 12.4.

10.10. Exercise. (a) Suppose that $p \in [1, +\infty)$. Show that the set

$$\mathscr{J} := \{\sum_j \beta_j c_{B_j} : B_j \in \mathscr{S}, \mu B_j < \infty\}$$

of simple functions is dense in $\mathscr{L}^p$.

Hint. Apparently, $\mathscr{J} \subset \mathscr{L}^p$. Choose $f \in \mathscr{L}^p$. Since f is finite μ-almost everywhere, Exercise 3.11 provides a sequence $\{f_n\}$ of simple functions such that $|f_n| \le |f|$ and $f_n \to f$ almost everywhere. Clearly $f_n \in \mathscr{J}$ and by Lebesgue's theorem with dominating function $2^p |f|^p$, $\|f_n - f\|_p \to 0$.

(b) Show that the set of all simple functions is dense in $\mathscr{L}^\infty$.

10.11. Exercise. Consider $p \in (0, 1)$. Define $\|.\|_p$ and the spaces $\mathscr{L}^p$, L^p in the same way as in the beginning of this chapter. Show that:

(a) L^p is a linear space.
(b) The triangle inequality for $\|.\|_p$ does not hold.

Hint. Employ the indicator functions of two disjoint sets of a positive measure.

(c) The distance function $d_p(f, g) := \int |f - g|^p \,\mathrm{d}\mu$ is a metric on L^p and (L^p, d_p) is a complete metric linear space.

10.12. Notes. The Hölder inequality was first proved by A.L. Cauchy [*1821] in the case of $p = 2$ for finite sums (so, in fact, for the counting measure on a finite set). V.Y. Bunyakowski [1859] proved it (still for $p = 2$) for Riemann integrable functions and the same result was obtained by H.A. Schwarz [1885]. For general p but still for finite sums, the inequality was proved by L.J. Rogers [1888] and by O. Hölder [1889]. The definitive form given in 10.3 comes from F. Riesz [1910], where Minkowski's inequality is proved as well. The original result for finite sums is due to H. Minkowski [1907].

The theory of L^p-spaces was originated by F. Riesz [1906] who defined the L^2-metric and E. Fischer [1907] who proved the completeness of L^2. F. Riesz then defined L^p for other values of p as well and proved their completeness. The notion of a Banach space has its roots in papers by S. Banach [1922] and N. Wiener [1922]. This period culminated when Banach's fundamental monograph [*1932] appeared.

11. Product Measures and the Fubini Theorem

In this chapter, we will consider the following problem: Given two measure spaces $(X, \mathscr{S}, \mu)$ and $(Y, \mathscr{T}, \nu)$, we wish to define a product measure τ on an appropriate σ-algebra $\mathscr{U}$ of subsets of the Cartesian product $X \times Y$. If $M = S \times T$ where $S \in \mathscr{S}$, $T \in \mathscr{T}$ (such sets are called *measurable rectangles*), we require

$$\tau M = \mu S \cdot \nu T\,.$$

An important example is that of Lebesgue measure in $\mathbf{R}^2$ which should arise as the product of one-dimensional Lebesgue measures.

11.1. Product σ-algebra. First we introduce the notion of the product σ-algebra. If $\mathscr{S}$ and $\mathscr{T}$ are some σ-algebras, the *product σ-algebra* $\mathscr{S} \otimes \mathscr{T}$ is the σ-algebra generated by the collection of all measurable rectangles. Thus, $\mathscr{S} \otimes \mathscr{T}$ is the smallest σ-algebra which contains all sets of the form $S \times T$ where $S \in \mathscr{S}$ and $T \in \mathscr{T}$.

For $M \subset X \times Y$, $x \in X$, let

$$M^x := \{y \in Y : [x, y] \in M\} .$$

The set M^x is called the *x-section of M*. Analogously, define the y-section

$$M_y = \{x \in X : [x, y] \in M\}$$

for $y \in Y$.

11.2. Lemma. *If $M \in \mathscr{S} \otimes \mathscr{T}$ and $x \in X$, then $M^x \in \mathscr{T}$ (and, of course, similarly $M_y \in \mathscr{S}$ for $y \in Y$).*

Proof. Denote $\mathscr{A} := \{E \in \mathscr{S} \otimes \mathscr{T} : E^x \in \mathscr{T}\}$. A moment's reflection shows that $\mathscr{A}$ contains all measurable rectangles. A routine argument yields that $(X \times Y \setminus E)^x = Y \setminus E^x$ and $(\bigcup\limits_\alpha E_\alpha)^x = \bigcup\limits_\alpha E^x_\alpha$. We see that $\mathscr{A}$ is a σ-algebra, and therefore $\mathscr{A} = \mathscr{S} \otimes \mathscr{T}$. ■

11.3 Monotone Classes. A family $\mathscr{M}$ of subsets of a set Z is a *monotone class* if it obeys the following conditions:

(a) if $M_1 \subset M_2 \subset M_3 \subset \dots$, $\quad M_n \in \mathscr{M}$, then $\bigcup M_n \in \mathscr{M}$;

(b) if $M_1 \supset M_2 \supset M_3 \supset \dots$, $\quad M_n \in \mathscr{M}$, then $\bigcap M_n \in \mathscr{M}$.

Every monotone class which is simultaneously an algebra forms a σ-algebra.

The following idea, used already in the proof of the previous lemma, will repeat in the proofs of subsequent theorems:

If $\mathscr{A}$ is the family of sets from $\mathscr{S} \otimes \mathscr{T}$ for which a certain proposition holds and if $\mathscr{A}$ contains all measurable rectangles, then $\mathscr{A} = \mathscr{S} \otimes \mathscr{T}$ provided $\mathscr{A}$ is a σ-algebra. It is not always so easy to verify that $\mathscr{A}$ a σ-algebra. Often it is easier to prove that $\mathscr{A}$ is only a monotone class and an algebra (mostly using Levi's theorem). But even in this case $\mathscr{A} = \mathscr{S} \otimes \mathscr{T}$ as follows from the next theorem.

11.4. Theorem. *If $\mathscr{R}$ is an algebra of sets, then the smallest monotone class containing $\mathscr{R}$ is $\sigma(\mathscr{R})$.*

Proof. Let $\mathscr{M}$ denote the smallest monotone class containing $\mathscr{R}$ (it does exist!). It is sufficient to show that $\mathscr{M}$ is an algebra. To this end, it is enough to prove that $\mathscr{M} \subset \mathscr{K}_E := \{B \colon E \setminus B,\ B \setminus E,\ B \cup E \in \mathscr{M}\}$ for each $E \in \mathscr{M}$. We fix a set F and proceed in the following steps:

(a) $\mathscr{K}_F$ is a monotone class;

(b) $\mathscr{R} \subset \mathscr{K}_F$ for $F \in \mathscr{R}$;

(c) $\mathscr{M} \subset \mathscr{K}_F$ for $F \in \mathscr{R}$;

(d) if $E \in \mathscr{M}$, then $\mathscr{R} \subset \mathscr{K}_E$. (Let $F \in \mathscr{R}$; by (c), $E \in \mathscr{K}_F$, and so $F \in \mathscr{K}_E$.) Therefore $\mathscr{R} \subset \mathscr{M} \subset \mathscr{K}_E$ by the definition of $\mathscr{M}$. ■

The basis for a definition of the product measure is in the following proposition.

11.5. Lemma. *Let $(X, \mathscr{S}, \mu)$ and $(Y, \mathscr{T}, \nu)$ be σ-finite measure spaces, $E \in \mathscr{S} \otimes \mathscr{T}$. Then the function $x \mapsto \nu(E^x)$ is $\mathscr{S}$-measurable ($E^x \in \mathscr{T}$ by the previous lemma), the function $y \mapsto \mu(E_y)$ is $\mathscr{T}$-measurable and*

$$\int_X \nu(E^x)\,\mathrm{d}\mu(x) = \int_Y \mu(E_y)\,\mathrm{d}\nu(y).$$

Proof. Let $\mathscr{A}$ be the collection of all sets $E \in \mathscr{S} \otimes \mathscr{T}$ such that all assertions hold for E. If $\mathscr{R}$ denotes the family of all finite disjoint unions of measurable rectangles, it is straightforward to check that $\mathscr{R}$ is an algebra. The assertion will follow from the preceding theorem, provided we can prove that $\mathscr{A}$ is a monotone class containing $\mathscr{R}$. We merely outline the main steps and invite the reader to fill in the details.

(a) $\mathscr{A}$ contains all measurable rectangles.

(b) $\mathscr{R} \subset \mathscr{A}$ (it suffices to show that $\bigcup_{j=1}^{n} E_j \in \mathscr{A}$ whenever $E_j \in \mathscr{A}$ are pairwise disjoint).

(c) If $E_n \in \mathscr{A}$, $E_1 \subset E_2 \subset \dots$, then $\bigcup E_n \in \mathscr{A}$ (use Levi's theorem and properties of sections).

(d) If $E_n \in \mathscr{A}$, $E_1 \supset E_2 \supset \dots$, then $\bigcap E_n \in \mathscr{A}$ (consider first the case when μ and ν are finite measures and proceed as in (c)). ∎

11.6. Product Measure. Let $(X, \mathscr{S}, \mu)$ and $(Y, \mathscr{T}, \nu)$ be σ-finite measure spaces. A measure τ on $\mathscr{S} \otimes \mathscr{T}$ is called a *product measure* of μ and ν (denoted by $\mu \otimes \nu$) if

$$\tau(A \times B) = \mu A \cdot \nu B$$

whenever $A \in \mathscr{S}$, $B \in \mathscr{T}$. Based on the following main theorem, we can claim that the product operation $\otimes$ is properly defined.

11.7. Theorem. *Let $(X, \mathscr{S}, \mu)$ and $(Y, \mathscr{T}, \nu)$ be again σ-finite measure spaces. Then there exists a unique measure τ on $\mathscr{S} \otimes \mathscr{T}$ such that*

$$\tau(A \times B) = \mu A \cdot \nu B$$

whenever $A \in \mathscr{S}$ and $B \in \mathscr{T}$.

Proof. Uniqueness is almost obvious. Indeed, if τ_1 and τ_2 are measures on $\mathscr{S} \otimes \mathscr{T}$ satisfying requirements of the theorem, then the family $\mathscr{D} = \{E \in \mathscr{S} \otimes \mathscr{T} : \tau_1 E = \tau_2 E\}$ contains all measurable rectangles, is closed under finite disjoint unions and is a monotone class. Therefore $\mathscr{D} = \mathscr{S} \otimes \mathscr{T}$ and we see that $\tau_1 = \tau_2$ on $\mathscr{S} \otimes \mathscr{T}$.

To prove the existence of a product measure, define

$$\tau E = \int_X \nu E^x \, d\mu$$

for $E \in \mathscr{S} \otimes \mathscr{T}$. We know that $\tau E = \int_Y \mu E_y \,\mathrm{d}\nu$ according to Lemma 11.5.

A routine argument shows that τ is a measure on $\mathscr{S} \otimes \mathscr{T}$ and $\tau(A \times B) = \mu A \cdot \nu B$ for all measurable rectangles $A \times B$. ∎

11.8. Lemma. *Let $(X, \mathscr{S}, \mu)$, $(Y, \mathscr{T}, \nu)$ be σ-finite measure spaces and f a $\mathscr{S} \otimes \mathscr{T}$-measurable function on $X \times Y$. Then the function $f(x, \cdot)$ is $\mathscr{T}$-measurable on Y for each $x \in X$.*

Proof. Select $\alpha \in \mathbf{R}$, $x \in X$ and observe that

$$\{y \in Y : f(x, y) > \alpha\} = \{(t, y) \in x \times Y : f(t, y) > \alpha\}^x .$$

■

11.9. Fubini's Theorem. *Let $(X, \mathscr{S}, \mu)$, $(Y, \mathscr{T}, \nu)$ be σ-finite measure spaces, and $h \in \mathscr{L}^*(\mu \otimes \nu)$. Then*

$$\int_{X \times Y} h \,\mathrm{d}\mu \otimes \nu = \int_X \left(\int_Y h(x, y) \,\mathrm{d}\nu \right) \mathrm{d}\mu = \int_Y \left(\int_X h(x, y) \,\mathrm{d}\mu \right) \mathrm{d}\nu .$$

Proof. It is no restriction to assume that $h \geq 0$, for we can use the decomposition $h = h^+ - h^-$. The assertion holds whenever $h = c_A$ is the indicator function of a set $A \in \mathscr{S} \otimes \mathscr{T}$ (see the proof of Theorem 11.7), and therefore for nonnegative simple functions as well. If $h \geq 0$ is an arbitrary $\mathscr{S} \otimes \mathscr{T}$-measurable function, there is a sequence $\{h_n\}$ of nonnegative simple functions, $h_n \nearrow h$ (cf. Theorem 3.9). Then $\int_{X \times Y} h_n \,\mathrm{d}\mu \otimes \nu \to \int_{X \times Y} h \,\mathrm{d}\mu \otimes \nu$. On the other hand, $h_n(x, y) \nearrow h(x, y)$ and since the assertion holds for simple functions h_n, by using Levi's theorem (twice) we easily finish the proof. ■

11.10. Remarks. 1. The Lebesgue measure in $\mathbf{R}^{n+k}$ is the completion of the product of Lebesgue measures in $\mathbf{R}^n$ and $\mathbf{R}^k$. We refer to Chapter 26 for a detailed study of this important example.

2. Suppose τ is the product of σ-finite measures μ and ν. Then the measure τ is not necessarily complete, even if μ and ν are complete. (Indeed, the set $\{(x, x) : x \in M\}$ is measurable with respect to the product of Lebesgue measures on $\mathbf{R}$ if and only if $M \subset \mathbf{R}$ is Lebesgue measurable. On the other hand, all subsets of the diagonal $\{(x, x) : x \in \mathbf{R}\}$ are measurable with respect to the completion of $\lambda \otimes \lambda$.) However, for the completed product measure the following variant of Fubini's theorem holds.

11.11. Theorem. *Let $(X, \mathscr{S}, \mu)$, $(Y, \mathscr{T}, \nu)$ be complete σ-finite measure spaces, τ the product of μ and ν and $\overline{\tau}$ the completion of τ. Let $h \in \mathscr{L}^*(\overline{\tau})$. Then the function $h(x, \cdot)$ is ν-measurable for μ-almost all $x \in X$, $h(\cdot, y)$ is μ-measurable for ν-almost all $y \in Y$, and*

$$\int_{X \times Y} h \,\mathrm{d}\overline{\tau} = \int_X \left(\int_Y h(x, y) \,\mathrm{d}\nu \right) \mathrm{d}\mu = \int_Y \left(\int_X h(x, y) \,\mathrm{d}\mu \right) \mathrm{d}\nu .$$

Proof. As usual, it is sufficient to prove the theorem for indicator functions. Let M be a set of the form $A \cup N$, where $A \in \mathscr{S} \otimes \mathscr{T}$ and $N \subset B$ for $B \in \mathscr{S} \otimes \mathscr{T}$, $\tau B = 0$. By Fubini's theorem

$$\int_Y c_B(x, \cdot) \,\mathrm{d}\nu = 0$$

for μ-almost all $x \in X$. Therefore also

$$\int_Y c_N(x,\cdot)\,\mathrm{d}\nu = 0$$

for μ-almost all $x \in X$. Hence we obtain

$$\int_{X\times Y} c_M\,\mathrm{d}\overline{\tau} = \int_X \left(\int_Y c_M(x,y)\,\mathrm{d}\nu\right)\,\mathrm{d}\mu.$$

■

11.12. Exercise. Show that every Dynkin class is a monotone class and find a counterexample that the reverse is not true.

11.13. Notes. The product measures and the reduction of multi-dimensional integration to one-dimensional one were originally studied in case of the Lebesgue measure in $\mathbf{R}^2$ and appear already in H. Lebesgue [*1904]. G. Fubini [1907] claims that the order of integration may be reversed. A complete proof was done by L. Tonelli [1909]. Approximately in the 30's a number of works concerning the abstract product measure theory appeared. The method explained in this chapter is due to H. Hahn [1933].

Let us point out that there are theories concerning products of measures which are not necessarily σ-finite. The theory of infinite products of measures (we understand the infinite number of factors) is also very important.

12. Sequences of Measurable Functions

In this section we will study the relationship among various kinds of convergences of sequences of measurable functions on a measure space $(X, \mathscr{S}, \mu)$. We will consider

(a) the *convergence almost everywhere*;
(b) the *convergence in the norm of L^p spaces*;
(c) the *convergence in measure*;
(d) the *μ-uniform convergence*.

Two last mentioned notions will now be defined.

12.1 Convergence in Measure. Let f, f_n be μ-measurable, almost everywhere finite functions on X. We say that a sequence $\{f_n\}$ *converges in measure* to f if

$$\lim \mu\{x \in X : |f(x) - f_n(x)| \geq \varepsilon\} = 0$$

for each $\varepsilon > 0$.

12.2. μ-uniform Convergence. Let f, f_n be almost everywhere finite functions on X. We say that $\{f_n\}$ converges to f *μ-uniformly* if for every $\varepsilon > 0$ there is $M \in \mathscr{S}$ such that $\mu(X \setminus M) < \varepsilon$ and the convergence $f_n \to f$ is uniform on M.

The limit function of a sequence converging in measure (or μ-uniformly) is unique except for a null set and is finite μ-almost everywhere.

12.3. Theorem. *Let $1 \le p < \infty$ and $\{f_n\}$ be a sequence of μ-measurable, almost everywhere finite functions on X. If $f_n \in \mathscr{L}^p$ and $\|f_n - f\|_p \to 0$, then f_n converge to f in measure.*

Proof. Choose $\varepsilon > 0$ and denote $E_n = \{x \in X : |f(x) - f(x_n)| \ge \varepsilon\}$. Then $\varepsilon^p \mu E_n \le \int_{E_n} |f - f_n|^p \, \mathrm{d}\mu \le \|f - f_n\|_p^p$, which implies the proposition. ∎

12.4. Riesz Theorem. *Let $\{f_n\}$ be a sequence of μ-measurable, almost everywhere finite functions on X. If f_n converge to f in measure, then there exists a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \to f$ almost everywhere.*

Proof. Suppose f_n converge to f in measure. There is a sequence $n_1 < n_2 < n_3 < \dots$ such that

$$\mu\{x \in X : \big|f(x) - f_{n_j}(x)\big| \ge \frac{1}{j}\} < \frac{1}{2^j}.$$

If $A_j := \bigcup_{k=j}^{\infty} \{x \in X : |f(x) - f_{n_k}(x)| \ge \dfrac{1}{k}\}$ and $B := \bigcap_{j=1}^{\infty} A_j$, then $A_1 \supset A_2 \supset A_3 \supset \dots$, $\mu A_1 \le \sum \dfrac{1}{2^k} < +\infty$, and consequently

$$\mu B = \lim \mu A_j \le \lim \sum_{k=j}^{\infty} \frac{1}{2^k} = \lim \frac{1}{2^{j-1}} = 0$$

(notice the connection with Borel-Cantelli's lemma 2.14). Now it is sufficient to prove that $f_{n_k}(x) \to f(x)$ for $x \in X \setminus B$. To this end let $x \in X \setminus B$ and $\varepsilon > 0$ be given. Then there is a $j = j(x)$ such that $x \in X \setminus A_j = \bigcap_{k=j}^{\infty} \{y \in X : |f(y) - f_{n_k}(y)| < \dfrac{1}{k}\}$. If $k_0 > \max(j(x), \frac{1}{\varepsilon})$, then $|f(x) - f_{n_k}(x)| < \frac{1}{k} \le \varepsilon$ for all $k \ge k_0$, as needed. ∎

12.5. Egorov's Theorem. *Let $\mu X < +\infty$ and f, f_n be μ-measurable, almost everywhere finite functions on X. Then $f_n \to f$ almost everywhere if and only if $f_n \to f$ μ-uniformly.*

Proof. Assume that $f_n \to f$ almost everywhere. Denote

$$E := \{x \in X : \, f(x), f_n(x) \text{ are finite and } f_n(x) \to f(x)\}.$$

Obviously $\mu(X \setminus E) = 0$ and it does no harm to assume that $E = X$. Choose $\varepsilon > 0$. Writing $E_{k,m} := \bigcup_{n \ge m} \{x \in X \colon |f_n(x) - f(x)| \ge \frac{1}{k}\}$, we have $E_{k,1} \supset E_{k,2} \supset \dots$ and $\mu E_{k,1} \le \mu X < +\infty$. For fixed $k \in \mathbf{N}$, $\lim_m \mu E_{k,m} = 0$, and we easily find a sequence $\{m_k\}$ such that $\mu\Big(\bigcup_k E_{k,m_k}\Big) < \varepsilon$. Set

$$M = X \setminus \bigcup_k E_{k,m_k}.$$

Then we have $|f_n(x) - f(x)| < \frac{1}{k}$ for each $x \in M$ and $n > m_k$, so that $f_n \to f$ uniformly on M.

The reverse implication is obvious. ■

12.6. Remark. If f and f_n are μ-measurable, almost everywhere finite functions on X and $f_n \to f$ μ-uniformly, then $f_n \to f$ in measure. Indeed, a routine argument shows that $\lim \mu\{|f_n - f| \geq \varepsilon\} = 0$ for any $\varepsilon > 0$. Combining aforementioned results we can state the following *Lebesgue's theorem*:

If $\mu X < +\infty$, f, f_n are μ-measurable, almost everywhere finite functions on X, and if $f_n \to f$ almost everywhere, then $f_n \to f$ in measure.

12.7. Exercise. Suppose $\mu X < \infty$. If f, f_n are μ-measurable, almost everywhere finite functions on X, then $f_n \to f$ in measure if and only if for any subsequence $\{f_{n_k}\}$ there exists its subsequence $\{f_{n_{k_j}}\}$ which converges to f almost everywhere.

Hint. For one of the implications use Theorem 12.4. If $\{f_n\}$ satisfies the "subsequences condition" and does not converge in measure, then there exists an $\varepsilon > 0$ and a subsequence $\{g_n\}$ of $\{f_n\}$ such that $\mu\{|g_n - f| > \varepsilon\} > \varepsilon$ for all n and $g_n \to f$ almost everywhere. By Lebesgue's theorem of 12.6 we obtain a contradiction.

12.8. Exercise. Let $(X, \mathscr{S}, \mu)$ be a measure space, φ an increasing, bounded and continuous function on $[0, \infty]$, $\varphi(0) = 0$, and $\varphi(s+t) \leq \varphi(s) + \varphi(t)$ for all $s, t \geq 0$ (for instance the function $t \mapsto t/(1+t)$). Define

$$\rho(f, g) = \sum_{j=1}^{\infty} 2^{-j} \varphi(\mu\{|f - g| \geq 1/j\}).$$

Show that

(a) ρ is a pseudometric on the space of all the μ-measurable functions on X;

(b) f_n converge to f in measure if and only if $\rho(f, f_n) \to 0$.

12.9. Exercise (Vitali's theorem). Let $1 \leq p < +\infty$. If $f_n \in \mathscr{L}^p$, $f_n \to f$ almost everywhere and $\|f_n\|_p \to \|f\|_p$, then $\|f_n - f\|_p \to 0$.

Hint. Let $\varphi_n = 2^{p-1}(|f_n|^p + |f|^p) - |f_n - f|^p$. Then $\varphi_n \to 2^p |f|^p$ almost everywhere. Since $|f_n - f|^p \leq (|f_n| + |f|)^p \leq 2^{p-1}(|f_n|^p + |f|^p)$, we have $\varphi_n \in \mathscr{L}^1$, $\varphi_n \geq 0$. By Fatou's lemma

$$2^p \int_X |f|^p \leq \liminf \int_X \varphi_n = 2^p \int_X |f|^p - \limsup \int_X |f_n - f|^p .$$

Hence $\limsup \int_X |f_n - f|^p = 0$, finishing the proof.

12.10. Exercise. (a) (Egorov) Let μ be a σ-finite measure on $(X, \mathscr{S})$, f_n μ-measurable functions on X and $f_n \to f$ almost everywhere (f_n, f almost everywhere finite). Show that there exist $E_k \in \mathscr{S}$ of finite measure such that $\mu(X \setminus \bigcup_{k=1}^{\infty} E_k) = 0$ and $f_n \rightrightarrows f$ on every E_k.

(b) Show that the assumption of σ-finiteness of the measure μ in (a) cannot be dropped.

Hint. Let X be the set of all convergent sequences $x = \{x_k\}$ of real numbers and μ the counting measure on X. Define a sequence $\{f_n\}$ of functions on X by $f_n(x) = x_n$. Choose a sequence $\{E_j\}$ of subsets of X on which $\{f_n\}$ converges uniformly and find a convergent sequence $\{y_k\}$ which does not belong to any of the sets E_j (because it converges more slowly than all sequences from E_j).

(c) Show that the μ-uniform convergence implies convergence almost everywhere, or convergence in measure, without the assumption that $\mu X < \infty$.

12.11. Exercise. Find a sequence $\{f_n\}$ of functions on $[0, 1]$ which converges to zero in (Lebesgue) measure whereas the sequence $\{f_n(x)\}$ fails to converge for any $x \in [0, 1]$.

Hint. Consider the indicator functions of intervals $[\frac{k-1}{m}, \frac{k}{m}]$ ordered into a suitable sequence.

12.12. Weak Convergence in L^p. Let $1 \le p < +\infty$ and $q = \frac{p}{p-1}$ ($q = +\infty$ if $p = 1$). Suppose $f,\ f_j \in \mathscr{L}^p$. We say that a sequence $\{f_j\}$ *converges weakly* to f in L^p (denoted $f = w\text{-}\lim f_j$) if $\int_X f_j g \,\mathrm{d}\mu \to \int_X fg \,\mathrm{d}\mu$ for every function $g \in \mathscr{L}^q$.

(a) If $f = \text{w-}\lim f_j$, $g = \text{w-}\lim f_j$, then $g = f$ almost everywhere.

(b) If $\|f - f_j\|_p \to 0$, then $f = \text{w-}\lim f_j$.

(c) Suppose $f, f_j \in \mathscr{L}^p$. Let $\mathscr{F}$ be a set of $\mathscr{L}^p$-functions, the linear span of which is dense in L^q. Then the following assertions are equivalent:

(i) $f = \text{w-}\lim f_j$;
(ii) the sequence $\{\|f_j\|_p\}$ is bounded and $\int_X g f_j \,\mathrm{d}\mu \to \int_X gf \,\mathrm{d}\mu$ for every $g \in \mathscr{F}$.

Hint. A typical application of the Banach-Steinhaus theorem from functional analysis.

(d) We can take as $\mathscr{F}$ the set of all indicator functions of sets from $\mathscr{S}$ in case $p = 1$, or the set of all indicator functions of sets from $\mathscr{S}$ of finite measure if $p > 1$; in the special case of the Lebesgue measure, other possibilities are shown in Theorem 31.4.

(e) If $p > 1$, then any norm-bounded sequence of functions from $\mathscr{L}^p$ has a weakly convergent subsequence.

Hint. If the space L^p is separable, proceed in a similar way as in the proof of Theorem 17.10. The nonseparable case is more difficult. The idea is that for $1 < p < \infty$ the space L^p is reflexive and reflexive spaces are characterized by this property, cf. for instance L. Mišík [*1989], Theorem 33.4 or R.B. Holmes [*1975], p.149.

(f) Let $p > 1$. If $f_j \to f$ almost everywhere and if $\{\|f_j\|_p\}$ is a bounded sequence, then $f = \text{w-}\lim f_j$. The proposition holds also if the assumption $f_j \to f$ almost everywhere is replaced by $f_j \to f$ in measure.

(g) The Radon-Riesz theorem. Let $p > 1$, $f = \text{w-}\lim f_j$ and $\|f\|_p = \lim \|f_j\|_p$. Then $\|f - f_j\|_p \to 0$. (What proposition do we obtain by sticking (f) and (g) together?)

Hint. The proof makes substantial use of uniform convexity of L^p spaces for $1 < p < \infty$, see M.M. Rao [*1987], Proposition 5.5.3.

(h) The sequence $f_j : x \mapsto \sin jx$ converges weakly to zero for any p on each bounded interval. None of its subsequences is convergent almost everywhere.

(i) Suppose $f_j = j^{1/p}$ on $(0, 1/j)$ and $f_j(x) = 0$ on $(1/j, 1)$. The sequence $\{f_j\}$ is bounded in the norm of $L^p(0,1)$ and tends to zero almost everywhere. If $p = 1$, then it is not weakly convergent (nor any of its subsequences). If $p > 1$, then it converges weakly but not in norm.

12.13. Weak* Convergence. Assume μ is a σ-finite measure on X and $f, f_j \in L^\infty$. We say that a sequence $\{f_j\}$ *converges weakly** to f (denoted $f = \text{w}^*\text{-}\lim f_j$) if $\int_X f_j g \,\mathrm{d}\mu \to \int_X fg \,\mathrm{d}\mu$ for each function $g \in \mathscr{L}^1$.

(a) If $f = \text{w}^*\text{-}\lim f_j$, $g = \text{w}^*\text{-}\lim f_j$, then $g = f$ almost everywhere.

(b) If $\|f - f_j\|_\infty \to 0$, then $f = \text{w}^*\text{-}\lim f_j$.

(c) Suppose $f, f_j \in \mathscr{L}^\infty$. Let $\mathscr{F}$ be a set of $\mathscr{L}^1$-functions, the linear span of which is dense in L^1. Then the following are equivalent:

(i) $f = \text{w}^*\text{-}\lim f_j$;
(ii) the sequence $\{\|f_j\|_\infty\}$ is bounded and $\int_X g f_j \,\mathrm{d}\mu \to \int_X gf \,\mathrm{d}\mu$ for each function $g \in \mathscr{F}$.

(d) If $f_j \to f$ almost everywhere and if the sequence $\|f_j\|_\infty$ is bounded, then $f = \text{w}^*\text{-}\lim f_j$.

(e) Any norm-bounded sequence $\{f_n\}$ of $\mathscr{L}^\infty$-functions contains a weakly* convergent subsequence.

Hint. Write $X = \bigcup X_k$ where $X_1 \subset X_2 \subset \dots$ and $\mu X_k < \infty$. Suppose $\{f_n\}$ is a norm-bounded sequence of functions from L^∞ and

$$\tilde{f}_{n,k} = \begin{cases} f_n & \text{on } X_k, \\ 0 & \text{on } X \setminus X_k. \end{cases}$$

Then $\{\tilde{f}_{n,k}\}_n$ is bounded even in the L^2-norm and there exists its subsequence $\{\tilde{f}_{n_i,k}\}_n$ which converges weakly in L^2 to a function $\tilde{f}_k$. But this means that $\lim_i \int_X f_{n_i} u \, d\mu = \int_X f_k u \, d\mu$ for each function $u \in L^2$ vanishing on the complement of X_k. The diagonal method (cf. 17.10) provides a subsequence $\{h_n\}$ of $\{f_n\}$ and a limit function f such that $\int_X h_n u \, d\mu \to \int_X f u \, d\mu$ for each function $u \in L^2$ vanishing on the complement of some of the sets X_k. Since the set of all such functions u is dense in L^1, in light of (c) we conclude that $f = \text{w}^*\text{-}\lim f_n$.

(f) The sequence $f_j : x \mapsto \sin jx$ converges weakly* to zero (cite the Riemann-Lebesgue lemma 31.10).

(g) Define $f_j = 0$ on $(0, 1/j)$ and $f_j(x) = 1$ on $(1/j, 1)$. The sequence $\{f_j\}$ is bounded in L^∞-norm and converges both almost everywhere and weakly* to the constant function 1. We have $\|\lim f_j\|_\infty = \lim \|f_j\|_\infty = 1$ but $\lim \|f - f_j\|_\infty \neq 0$.

12.14. Remarks. It may be worth while to take for the interested reader a slightly deeper look from the point of view of functional analysis (see also [Zap]).

(1) Let X is a Banach space with (topological) dual X^*, $x_n, x \in X$, $F_n, F \in X^*$. We say that

(a) $\{x_n\}$ *converges weakly* to x if $F(x_n) \to F(x)$ for any functional $F \in X^*$;

(b) $\{F_n\}$ *converges weakly** to F if $F_n(z) \to F(z)$ for any $z \in X$.

Taking into account 13.17, the previous definitions of the weak and weak* convergence in the L^p spaces now are easier to understand (at least if $1 < p < \infty$ or if μ is σ-finite).

2. We say that a Banach space is *uniformly convex* if for any $\varepsilon > 0$ there is a $\delta > 0$ such that,

$$\text{if } x, y \in X, \|x\| = \|y\| = 1 \text{ and } \left\|\frac{1}{2}(x+y)\right\| > 1 - \delta, \text{ then } \|y - x\| < \varepsilon.$$

Each uniformly convex Banach space is *locally uniformly rotund* which means that if $\|x_n\| = \|x\| = 1$ and $\left\|\frac{1}{2}(x_n + x)\right\| \to 1$, then $\|x_n - x\| \to 0$. The L^p spaces are uniformly convex for $1 < p < \infty$ (it follows from the Clarkson's inequalities, see, for instance, E. Hewitt and K. Stromberg [*1965]).

3. A Banach space X is said to have the *Kadec-Klee property* if $\|x_n - x\| \to 0$, whenever a sequence $\{x_n\} \subset X$ converges weakly to x and $\|x_n\| \to \|x\|$. All locally uniformly rotund Banach spaces (in particular L^p spaces, $1 < p < \infty$) have the Kadec-Klee property. Indeed, if $\{x_n\}$ is a sequence which does not converge to a point x and for which $\|x_n\| = \|x\| = 1$, local uniform rotundity ensures the existence of its subsequence, call it $\{x_n\}$ again, and an $\varepsilon \in (0,1)$ such that $\left\|\frac{1}{2}(x_n + x)\right\| < 1 - \varepsilon$. Using the Hahn-Banach theorem find a $F \in X^*$ with $\|F\| = 1 = F(x)$. Now, it is clear that the sequence $\{F(x_n)\}$ cannot converge to $F(x)$.

4. In interesting cases, the L^1 spaces are not (locally) uniformly convex, nor do they need satisfy a proposition analogous to 12.12.f. Nevertheless, if $f_n \in L^1$, $f_n \to f$ almost everywhere or in measure and $\|f_n\| \to \|f\|$, then $\|f_n - f\| \to 0$ (cf. 12.9).

12.15. Exercise. Let $\mu X < \infty$ and $1 \le p < r \le +\infty$. If $\{\|f_j\|_r\}_j$ is a bounded sequence and $f_j \to f$ almost everywhere (or, in measure), then $\|f - f_j\|_p \to 0$. Show that the assertion does not continue to hold for $r = p$.

12.16. Notes. The notion of convergence in measure was introduced by F. Riesz [1909a] where also Theorem 12.4 was proved. Egorov's theorem was proved for case of the Lebesgue measure by H. Lebesgue [1903] and by D. Egorov [1911].

13. The Radon-Nikodým Theorem and the Lebesgue Decomposition

13.1. Radon-Nikodým Derivative and Absolute Continuity of Measures. Let $(X, \mathscr{S}, \mu)$ be a measure space and $f \in \mathscr{L}^1(\mu)$ a nonnegative function on X. In Chapter 8 we defined the measure μ_f on $(X, \mathscr{S})$ as

$$\mu_f(E) = \int_E f \, d\mu, \qquad E \in \mathscr{S}.$$

Now, if μ and ν are measures on $\mathscr{S}$, one might ask whether there is a nonnegative function $f \in \mathscr{L}^1(\mu)$ such that $\nu = \mu_f$. Immediately we see that such a function in general does not exist. Indeed, every measure of the form μ_f satisfies $\mu_f(E) = 0$, provided $\mu(E) = 0$. However, we show that if $\nu E = 0$ whenever $\mu E = 0$, and μ and ν are σ-finite, then a desired function f, called the *density* or the *Radon-Nikodým derivative* of ν with respect to μ, can be found. Then the density of ν with respect to μ is unique except for a null set and it is sometimes denoted by $\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$.

We say that a measure ν on $\mathscr{S}$ is *absolutely continuous* with respect to μ, and write $\nu \ll \mu$, if $\nu E = 0$ for every $E \in \mathscr{S}$ with $\mu E = 0$.

There are several methods how to prove the existence of Radon-Nikodým derivatives. Our proof is based on the following "variational" approach. Assuming that such a function f exists and is in $\mathscr{L}^2$, consider the functional

$$J_0 : g \mapsto \int_X |g - f|^2 \, \mathrm{d}\mu$$

on the space $L^2(\mu)$. The function f evidently represents the only element of the space $L^2(\mu)$ at which J_0 attains its minimum. For all $g \in L^2(\mu)$

$$J_0(g) = J(g) + \int_X f^2 \, \mathrm{d}\mu$$

where

$$J(g) = \int_X (g^2 - 2fg) \, \mathrm{d}\mu = \int_X g^2 \, \mathrm{d}\mu - 2 \int_X g \, \mathrm{d}\nu.$$

Since the functionals J and J_0 differ only by a constant, J attains its minimum at f. This is also the idea of our proof of the following lemma.

13.2. Lemma. *Let ν, σ be finite measures on $(X, \mathscr{S})$ such that $\nu A \le \sigma A$ for each $A \in \mathscr{S}$. Then there exists a nonnegative function $f \in \mathscr{L}^2(\sigma)$ such that*

$$\nu A = \int_A f \, \mathrm{d}\sigma$$

for any $A \in \mathscr{S}$.

Proof. For $g \in \mathscr{L}^2(\sigma)$ denote

$$J(g) = \int_X g^2 \, \mathrm{d}\sigma - 2 \int_X g \, \mathrm{d}\nu \, .$$

Let

$$c := \inf\{J(g) : g \in L^2(\sigma)\}.$$

Since $\nu A \le \sigma A$ for each $A \in \mathscr{S}$, we have

$$J(g) \ge \int_X (g^2 - 2\,|g|) \, \mathrm{d}\sigma = \int_X (|g| - 1)^2 \, \mathrm{d}\sigma - \sigma(X) \ge -\sigma(X),$$

for $g \in \mathscr{L}^2(\sigma)$, so that $c \in \mathbf{R}$. Hence, there is a sequence $\{f_j\}$ of functions from $\mathscr{L}^2(\sigma)$ such that $J(f_j) \to c$. Since for $g, h \in \mathscr{L}^2(\sigma)$ the "parallelogram law"

$$J(g) + J(h) - 2J\left(\frac{1}{2}(g+h)\right) = \frac{1}{2}\|g-h\|_2^2$$

holds, we have

$$\frac{1}{2}\int_X |f_j - f_k|^2 \,\mathrm{d}\sigma \le J(f_j) + J(f_k) - 2c \to 0$$

as $j, k \to \infty$. Consequently, $\{f_j\}$ is a Cauchy sequence, and therefore there is a limit f of this sequence in $L^2(\sigma)$. Obviously, $J(f) = c$. Choose $A \in \mathscr{S}$. Since

$$J(f) \le J(f + tc_A)$$

for all $t \in \mathbf{R}$, we get

$$0 \le 2t\int_X fc_A \,\mathrm{d}\sigma + t^2\int_X c_A^2 \,\mathrm{d}\sigma - 2t\int_X c_A \,\mathrm{d}\nu = 2t\left(\int_X f\,\mathrm{d}\sigma - \nu A\right) + t^2\sigma A.$$

The last shows that

$$\left|\int_X f\,\mathrm{d}\sigma - \nu A\right| \le \frac{1}{2}|t|\cdot\sigma A$$

(note that t can be positive as well as negative). By letting $t \to 0$ we establish the assertion. ■

13.3. Remark. Proof of the preceding lemma becomes more comprehensible if it is based on the knowledge of the theory of Hilbert spaces. Considering the functional

$$F(g) = \int_X g\,\mathrm{d}\nu, \quad g \in L^2(\sigma).$$

By the Riesz representation theorem on continuous linear functionals on Hilbert spaces there is an element $f \in L^2(\sigma)$ such that

$$F(g) = \int_X fg\,\mathrm{d}\sigma$$

for each $g \in L^2(\sigma)$. Now it is enough to apply the last equality to $g = c_A$, where A runs over $\mathscr{S}$.

13.4. Radon-Nikodým Theorem. *Let μ, ν be finite measures on $(X, \mathscr{S})$, $\nu \ll \mu$. Then there exists a nonnegative function $h \in \mathscr{L}^1(\mu)$ such that*

$$\nu A = \int_A h\,\mathrm{d}\mu$$

for all $A \in \mathscr{S}$. This function h is unique up to μ-almost everywhere equality.

Proof. The uniqueness can be achieved by Theorem 8.17, and we proceed to establish the existence of h. Set $\sigma = \mu + \nu$. By the previous lemma there exists a function $f \in \mathscr{L}^2(\sigma)$ such that

$$\nu A = \int_A f \,\mathrm{d}\sigma$$

for all $A \in \mathscr{S}$. Since

$$0 \le \nu A = \int_A f \,\mathrm{d}\sigma \le \sigma A = \int_A 1 \,\mathrm{d}\sigma\,,$$

an appeal to Corollary 8.18 reveals that $0 \le f \le 1$ σ-almost everywhere. We claim that $f < 1$ σ-almost everywhere. Let $E := \{f = 1\}$. Then

$$\nu E = \int_E f \,\mathrm{d}\sigma = \sigma E,$$

thus $\mu E = 0$. By assumptions we get $\nu E = 0$ and therefore $\sigma E = 0$. Hence, it is no restriction to assume that $0 \le f < 1$ everywhere. For each set $A \in \mathscr{S}$ we have

$$\int_X c_A f \,\mathrm{d}\mu = \int_X c_A (1 - f) \,\mathrm{d}\nu.$$

We see (for simple functions using linearity, then passing to limits) that

$$\int_X g f \,\mathrm{d}\mu = \int_X g(1 - f) \,\mathrm{d}\nu$$

for every nonnegative $\mathscr{S}$-measurable function g on X. For $A \in \mathscr{S}$, put $g = \frac{c_A}{1-f}$. Then

$$\int_X \frac{f}{1-f} \cdot c_A \,\mathrm{d}\mu = \int_X c_A \,\mathrm{d}\nu$$

and the function

$$h := \frac{f}{1-f}$$

has all required properties. Notice that $h \in \mathscr{L}^1(\mu))$ as $\int_X h \,\mathrm{d}\mu = \nu(X) < +\infty$. ∎

13.5. Theorem. *Let μ and ν be finite measures on $(X, \mathscr{S})$. Then the following conditions are equivalent:*

(i) $\nu \ll \mu$;
(ii) *there is a nonnegative function $f \in \mathscr{L}^1(\mu)$ such that*

$$\nu E = \int_E f \,\mathrm{d}\mu$$

for all $E \in \mathscr{S}$;
(iii) *for any $\varepsilon > 0$ there is a $\delta > 0$ such that $\nu E < \varepsilon$ whenever $E \in \mathscr{S}$ and $\mu E < \delta$.*

Proof. The implication (i) $\Longrightarrow$ (ii) was established in Theorem 13.4, (ii) $\Longrightarrow$ (iii) follows by Exercise 8.22.b and (iii) $\Longrightarrow$ (i) is obvious. ■

13.6. Remarks. 1. The Radon-Nikodým theorem can be generalized, its variants hold even for signed or complex measures. Let us state the version for σ-finite measures:

Let μ, ν be σ-finite measures on $(X, \mathscr{S})$, $\nu \ll \mu$. Then there is a nonnegative $\mathscr{S}$-measurable function h (not necessarily from $\mathscr{L}^1(\mu)$) such that

$$\nu A = \int_A h \, \mathrm{d}\mu$$

for all $A \in \mathscr{S}$. This function is unique except on a null set.

Indeed, we find sequences $\{A_n\}, \{B_n\} \subset \mathscr{S}$ of pairwise disjoint sets such that $\mu A_n < +\infty$, $\nu B_n < +\infty$, $\bigcup A_n = \bigcup B_n = X$. Applying Theorem 13.4 for restrictions of μ and ν to $A_n \cap B_k$ we obtain the assertion.

2. We proved the Radon-Nikodým theorem and the Hahn decomposition theorem independently. The reader should find it easy to derive the Radon-Nikodým theorem from the Hahn theorem or vice versa. Both possibilities will be indicated.

Let μ, ν be finite measures on $(X, \mathscr{S})$, $\nu \ll \mu$. We would like to prove the existence of a Radon-Nikodým derivative. For $\alpha \in \mathbf{Q}$ find a Hahn decomposition $X = P_\alpha \cup N_\alpha$ such that $\nu - \alpha\mu$ is a nonnegative measure on P_α and $-(\nu - \mu\alpha)$ is a nonnegative measure on N_α. For $\alpha = 0$ set $P_0 = X$, $N_0 = \emptyset$. Then we find $\mathrm{d}\nu/\mathrm{d}\mu$ in the form

$$f(x) = \sup\{\alpha \in \mathbf{Q} \cap [0, +\infty) \colon x \in P_\alpha\}.$$

On the other hand, if μ is a signed measure, $f = \dfrac{\mathrm{d}\mu^+}{\mathrm{d}\,|\mu|}$ and $P = \{f = 1\}$, $N = X \setminus P$, then (P, N) is a Hahn decomposition of X with respect to μ.

13.7. Integration with Respect to Signed or Complex Measures. Suppose μ is a signed or complex measure on $(X, \mathscr{S})$. Then we define

$$\int_X f \, \mathrm{d}\mu = \int_X af \, \mathrm{d}\,|\mu|$$

where $a = \mathrm{d}\mu/\mathrm{d}\,|\mu|$. (Show that $|a| = 1$ almost everywhere.)

13.8. Singular and Absolutely Continuous Measures. Suppose that μ and ν are positive, signed, or complex measures on the measurable space $(X, \mathscr{S})$.

We say that μ and ν are (mutually) *singular*, which is denoted by $\mu \perp \nu$, if there is a $M \in \mathscr{S}$ with $|\mu|(M) = |\nu|(X \setminus M) = 0$. Obviously, this relation is symmetric.

A measure ν is *absolutely continuous* with respect to μ if $\nu E = 0$, whenever $E \in \mathscr{S}$ and $|\mu|(E) = 0$.

13.9. Examples. (a) Let μ be a signed measure. Using a Hahn decomposition of X for μ, we see immediately that μ^+ and μ^- are mutually singular.

(b) The Dirac measure at 0 and the Lebesgue measure are relatively singular on $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$.

(c) If η is the Lebesgue-Stieltjes measure determined by the Cantor singular function (see Example 23.1 and Exercise 24.8.) and λ the Lebesgue measure, then $\eta \perp \lambda$ on $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$.

13.10. Lebesgue Decomposition Theorem. *Let μ be a (positive) measure on $(X, \mathscr{S})$ and ν a σ-finite or complex measure on $(X, \mathscr{S})$. Then there exists a unique decomposition $\nu = \nu_a + \nu_s$ where $\nu_a \ll \mu$ and $\nu_s \perp \mu$.*

Proof. (See also Exercise 13.12.) We first reduce the general case to the case when ν is a finite positive measure. There are $B_j \in \mathscr{S}$ such that $\mu B_j = 0$ and $\lim \nu B_j = \sup\{\nu B : B \in \mathscr{S},\ \mu B = 0\}$. If $M = \bigcup B_j$, then $\mu M = 0$ and for the measures

$$\nu_a A := \nu(A \setminus M),\ \nu_s A := \nu(A \cap M)$$

the equality $\nu = \nu_a + \nu_s$ holds. Clearly $\nu_s \perp \mu$. If now $\mu B = 0$, then $\nu_a(B) = \nu(B \setminus M) = 0$. This is so because $\nu B' = 0$ for each set $B' \in \mathscr{S}$, $B' \subset X \setminus M$ with $\mu B' = 0$. (Otherwise $\mu(M \cup B') = 0$ and $\nu(B' \cup M) > \nu M$.) If $X = \bigcup X_k$ where $\nu X_k < +\infty$, find sets $M_k \subset X_k$ in the same way as in the first part of the proof and set

$$M = \bigcup M_k, \quad \nu_a A = \nu(A \setminus M), \quad \nu_s A = \nu(A \cap M).$$

Having the proof for positive measures, the general case easily follows.

Uniqueness remains. Suppose that $\nu = \nu'_a + \nu'_s$ is another decomposition with the required properties. Then there is a set $M' \in \mathscr{S}$ for which $\mu M' = 0$ and $\nu'_s(X \setminus M') = 0$. Fix $A \in \mathscr{S}$ and denote $C = A \cap (M \cup M')$, $D = A \setminus (M \cup M')$. We have $\mu C \le \mu M + \mu M' = 0$, so $\nu_a C = \nu'_a C = 0$ and $\nu_s C = \nu'_s C$. Further, $\nu_s D = \nu'_s D = 0$ and therefore $\nu_a D = \nu'_a D$. Since $A = C \cup D$ and $C \cap D = \emptyset$, we get $\nu_s A = \nu'_s A$ and $\nu_a A = \nu'_a A$. ■

13.11. Lebesgue Decomposition. The decomposition $\nu = \nu_a + \nu_s$ is called the *Lebesgue decomposition* of ν relatively to μ, and the measures ν_a and ν_s are called the *absolutely continuous part* and the *singular part* of ν. Notice that there exists a set $M \in \mathscr{S}$ such that $\nu_s = \nu_M$ and $\nu_a = \nu_{X \setminus M}$.

13.12. Exercise. Give the details of an alternative proof of the Lebesgue decomposition theorem: Existence (assuming σ-finitness of both measures) using the Radon-Nikodým theorem and uniqueness from the following exercise (part (c)).

Hint. (a) Existence: Suppose μ, ν are positive and set $\lambda = \mu + \nu$, $f = \mathrm{d}\mu/\mathrm{d}\lambda$. Then the Lebesgue decomposition has the form $\nu_s = \nu_M$, $\nu_a = \nu_{X \setminus M}$, where $M = \{f = 0\}$.

(b) Uniqueness: If $\nu = \nu_a + \nu_s = \nu'_a + \nu'_s$, then $\nu_a - \nu'_a = \nu_s - \nu'_s$ is at the same time absolutely continuous and singular.

13.13. Exercise. Let μ and ν be positive, signed or complex measures on a measurable space $(X, \mathscr{S})$).

(a) Prove that the following are equivalent:

(i) $\mu \perp \nu$;

(ii) $|\mu| \perp |\nu|$;

(iii) $(\mu_r)^+ \perp (\nu_r)^+$, $(\mu_r)^- \perp (\nu_r)^-$; $(\mu_i)^+ \perp (\nu_i)^+$, $(\mu_i)^- \perp (\nu_i)^-$.

(b) If $\nu_1 \perp \mu$, $\nu_2 \perp \mu$, then $\nu_1 + \nu_2 \perp \mu$.

(c) If $\nu \perp \mu$, $\nu \ll \mu$, then $\nu = 0$.

13.14. Exercise. Suppose $S \subset \mathbf{R}$ is countable, $f \ge 0$ and $\mu = \sum_{s \in S} f(s)\varepsilon_s$ (ε_s are the Dirac measures, cf. Exercise 2.10). Describe the Lebesgue decomposition of μ with respect to the Lebesgue measure.

13.15. Exercise. Suppose $\mu \in \mathbf{M}(X)$ (see Exercise 6.17), $V = \{\nu \in \mathbf{M}(X) : \nu \perp \mu\}$. Show that V is a closed linear subspace of the Banach space $\mathbf{M}(X)$.

13.16. Exercise. Let μ, ν be finite (positive) measures on $(X, \mathscr{S})$. Let sup and inf be defined as in Exercise 6.19 and let $\sigma = \mu + \nu$.

(a) Show that the following statements are equivalent:

(i) $\mu \perp \nu$;
(ii) $\inf(\mu, \nu) = 0$;
(iii) $\sup(\mu, \nu) = \mu + \nu$.

(b) Prove that $\dfrac{\mathrm{d}\sup(\mu,\nu)}{\mathrm{d}\sigma} = \max(\dfrac{\mathrm{d}\mu}{\mathrm{d}\sigma}, \dfrac{\mathrm{d}\nu}{\mathrm{d}\sigma})$.

13.17. Duality of L^p Spaces. Suppose $p, q \in [1, +\infty]$, $\frac{1}{p} + \frac{1}{q} = 1$. Consider spaces L^p and L^q determined by a σ-finite measure μ on $(X, \mathscr{S})$.

(a) If $v \in L^q$, then the mapping $u \mapsto \int_X uv \,\mathrm{d}\mu$ is a continuous linear functional on L^p.

(b) Suppose $p < \infty$ and f is a continuous linear functional on L^p. Then there exists a $v \in L^q$ such that $f(u) = \int_X uv \,\mathrm{d}\mu$ for all $u \in L^p$. Moreover $\|v\|_q = \|f\|$ (here, $\|f\| = \sup\{f(u) \colon u \in L^p,\ \|u\|_p \le 1\}$).

Hint. By the Radon-Nikodým theorem there exists a μ-measurable finite function v such that $f(c_E) = \int_E v \,\mathrm{d}\mu$ for each set $E \in \mathscr{S}$. Thus, $f(u) = \int_X uv \,\mathrm{d}\mu$ for each $u \in L^p$. It remains to show that $v \in L^q$ and $\|v\|_q = \|f\|$. There is an increasing sequence $\{E_j\}$ of sets from $\mathscr{S}$ of finite measures such that v is bounded on each E_j and $\bigcup E_j$ is X. Set $u_j = |v|^{q-2}\, v c_{E_j}$. Then $u_j \in L^p$ and $\|u_j\|_p^p = \int_X u_j v \,\mathrm{d}\mu = f(u_j) \le \|f\|\,\|u_j\|_p$ and therefore $\int_{E_j} |v|^q \,\mathrm{d}\mu = \int_X |u_j|^p \,\mathrm{d}\mu \le \left(\|u_j\|_p^{p-1}\right)^q \le \|f\|^q$. An appeal to Levi's theorem reveals that $v \in L^q$ and $\|v\|_q \le \|f\|$. The reverse inequality is as easy consequence of Hölder's inequality.

13.18. Remark. The previous theorem which describes duals of L^p for $1 < p < \infty$ was proved using the Radon-Nikodým theorem, so that we had to confine ourselves to the case of σ-finite measures. However, the theorem holds for arbitrary measures.

In a similar way, it can be proved that any element of L^∞ represents a continuous linear functional on L^1 (by the same formula as in 13.17), but not every element of the dual space to L^1 is of this form. In order to be able to characterize elements of the dual space to L^1 by elements of L^∞, it is necessary to confine ourselves, e.g. to the case of σ-finite measures. One possible description of the dual space of $L^1(\mu)$ in the case of arbitrary measure μ can be found in J. Schwartz [1951].

The assumptions imposed on measures in the Radon-Nikodým theorem can be weakened to the case of so called *localizable* measures. This notion, containing σ-finite case as well as the case of Radon measures, was introduced by I.E. Segal in [1954]. For more information we refer to M.M. Rao [*1987].

Let us remark that the dual spaces of L^∞ can be described using (bounded) finitely additive measures, cf. E. Hewitt and K. Stromberg [*1965].

13.19. Notes. The Radon-Nikodým theorem was first proved by H. Lebesgue [1910] for measures which are absolutely continuous with respect to the Lebesgue measure. Later, it was generalized by M.J. Radon [1913] to Radon measures and by O. Nikodým [1930] to measures on abstract spaces. The Lebesgue decomposition in the case of arbitrary measures is in Saks' monograph [*1937]. There are different proofs of the Radon-Nikodým theorem. One of them is based on the Hahn decomposition and a newer one, using the Riesz theorem of representation of functionals on Hilbert spaces, comes from J. von Neumann. Our proof uses the variational principle and is near to von Neumann's.

In the classical case of the Lebesgue measure, the characterization of dual spaces of L^p spaces is due to F. Riesz [1910] for $p > 1$ and to H. Steinhaus [1919] for $p = 1$.

C. Radon Integral and Measure

14. Radon Integral

14.1. Radon Integral. Throughout this chapter P, will denote a locally compact topological space. The most important examples of locally compact spaces are open and closed subsets of $\mathbf{R}^n$.

The *support* of a function f is the closure of the set $\{x \in P \colon f(x) \neq 0\}$. It is denoted by $\operatorname{supt} f$.

If $K \subset P$ is a compact set, the symbol $\mathscr{C}_K(P)$ stands for the linear space of all continuous functions on P whose support is contained in K. By $\mathscr{C}_c(P)$ we denote the set of all continuous functions on P whose support is compact, i.e. $\mathscr{C}_c(P) = \bigcup\{\mathscr{C}_K(P) \colon K \subset P \text{ is compact }\}$. If the space P is compact, then $\mathscr{C}_c(P)$ and the space $\mathscr{C}(P)$ of all continuous functions on P coincide.

A functional A on $\mathscr{C}_c(P)$ is *positive* if $Af \geq 0$ whenever $f \geq 0$, and *monotone* if $Af \leq Ag$ whenever $f \leq g$. A linear functional is positive if and only if it is monotone (for nonlinear functionals, these two notions are different).

A *Radon integral* on P is any positive linear functional on the space $\mathscr{C}_c(P)$.

14.2. Examples. (a) Fix a point $a \in P$. It is simply to verify that the mapping

$$\varepsilon_a \colon f \mapsto f(a)\,, \quad f \in \mathscr{C}_c(P)$$

is a Radon integral on P. This functional is called the *Dirac integral* at a.

(b) The functional

$$Af := \int_a^b f$$

is a Radon integral on $[a, b]$. Since f is continuous, the integral $\int_a^b f$ always exists in Newton's or Riemann's sense.

(c) The previous example can be generalized. If φ is a nondecreasing function on $\mathbf{R}$ and $[a, b] \subset \mathbf{R}$, we define

$$A_\varphi f = (RS)\int_a^b f\,\mathrm{d}\varphi := \inf\{\sum_{i=1}^n c_i(\varphi(x_i) - \varphi(x_{i-1})) \colon \\ a = x_0 < x_1 < \cdots < x_n = b,\ c_i \geq f \text{ on } [x_{i-1}, x_i]\}$$

for each $f \in \mathscr{C}([a, b])$. The functional A_φ is again a Radon integral and it is called the *Riemann-Stieltjes integral*. If $\varphi(x) = x$, it is the Riemann integral.

(d) The Riemann integral or the Riemann-Stieltjes integral can be defined for functions from $\mathscr{C}_c(\mathbf{R})$. Indeed, observe that $\int_a^b f\,\mathrm{d}\varphi = \int_c^d f\,\mathrm{d}\varphi$ whenever $f \in \mathscr{C}_c(\mathbf{R})$ and $\operatorname{supt} f \subset [a, b] \cap [c, d]$. Thus, if $f \in \mathscr{C}_c(\mathbf{R})$, set $A_\varphi f = \int_a^b f\,\mathrm{d}\varphi$ where $[a, b]$ is an arbitrary interval containing the support of f. The functional A_φ is then a Radon integral on $\mathbf{R}$.

(e) Another example of a Radon integral is the functional

$$f \mapsto \int_G (f \circ \psi)\, g\,,$$

where g is a continuous nonnegative function on an open set $G \subset \mathbf{R}^k$ and $\psi : G \to P$ is a continuous mapping. In this way, the following important examples (f) and (g) can be expressed as well.

(f) Suppose $G \subset \mathbf{R}^k$ is an open set and $\psi : G \to \mathbf{R}^n$ is a diffeomorphism. Then the Radon integral

$$f \mapsto \int_G f \circ \psi \sqrt{\det((\nabla\psi)^*\nabla\psi)}$$

is in fact the (k-dimensional) surface integral of f over $\psi(G)$.

(g) Let $P = \{z \in \mathbf{R}^2 \colon |z| = 1\}$. If f is a continuous function on P and $x = (r\cos\omega,\, r\sin\omega)$, $|x| < 1$, set

$$Af(x) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(1-r^2)f(\cos t, \sin t)}{1 - 2r\cos(t-\omega) + r^2}\,\mathrm{d}t;$$

this integral is called the *Poisson* integral. It is not too hard to see that the mapping

$$f \mapsto Af(x)$$

is a Radon integral on P.

The Poisson integral is used when solving the Laplace (partial differential) equation $\Delta h = 0$. Indeed, given a continuous function f on P, the function $h\colon x \mapsto Af(x)$ is harmonic in the unit disc $U := \{x \in \mathbf{R}^2 \colon |x| < 1\}$ (i.e. h is a solution to the Laplace equation) and

$$\lim_{x\in U,\, x\to z} h(x) = f(z)$$

for each $z \in P$.

An important property of Radon integrals is described in the following theorem.

14.3. Theorem (Daniell's property). *Let A be a Radon integral on P, $f_n \in \mathscr{C}_c(P)$, $f_n \searrow 0$. Then $Af_n \to 0$.*

Proof. There is a limit $b := \lim Af_n \geq 0$. By Dini's theorem $f_n \rightrightarrows 0$ on P (consider the sequence $\{f_n\}$ on the support of f_1), and therefore there exists a sequence $\{n_k\}$ so that $|f_{n_k}| \leq \frac{1}{k^2}$ for each $k \in \mathbf{N}$. Thus the series $\sum f_{n_k}$ converges uniformly on P and if $f := \sum f_{n_k}$, then $f \in \mathscr{C}_c(P)$. Then for each k,

$$bk \leq \sum_{i=1}^{k} Af_{n_i} = A\left(\sum_{i=1}^{k} f_{n_i}\right) \leq Af,$$

and hence $b = 0$. ■

14.4. Semicontinuous Functions. Remember that a function $f : P \to \overline{\mathbf{R}}$ is said to be *lower semicontinuous* if $\{x \in P : f(x) > c\}$ is open for each $c \in \mathbf{R}$. Upper semicontinuous functions are defined in a similar way.

Denote by $C_c^{\uparrow}(P)$ the set of all lower semicontinuous functions on P which are nonnegative outside a compact set and do not attain the value $-\infty$. Analogously we define $C_c^{\downarrow}(P)$.

14.5. Extension of a Radon Integral. Let A be a Radon integral on P. We extend A for larger classes of functions in the following steps.

1. For $f \in \mathscr{C}_c(P)$ let $\int f = Af$.
2. If $f \in C_c^{\uparrow}(P)$, then we define $\int f = \sup\{\int g \colon g \in \mathscr{C}_c(P), g \leq f\}$.

3. Lastly, for an arbitrary function f on P we define *upper* and *lower integrals* as

$$\int^* f = \inf\{\int u \colon u \in C_c^{\uparrow}(P), u \geq f\}, \qquad \int_* f = -\int^* (-f).$$

4. We say that a function f on P is *A-integrable* if $\int_* f = \int^* f$ and this common value, which is denoted by $\int f$, is finite.

14.6. Properties of Extended Radon Integrals. The procedure of extension of Radon integrals requires to prove in each step that the "new integral" agrees with the "old" one for functions which are already integrable with respect to previous definitions. It is also needed to show that

(a) $\int_* f \leq \int^* f$ for every function f on P;

(b) the set of all finite A-integrable functions on P is a linear space and $\int$ is a nonnegative linear functional on it.

14.7. Measures induced by Radon Integrals. The mapping $\mu_A^* : E \mapsto \int^* c_E$ is an outer measure on P. Define $\mathfrak{M}(A)$ as the σ-algebra of all μ_A^*-measurable subsets of P in the sense of Carathéodory (see 4.4.) and μ_A as the restriction of μ_A^* to $\mathfrak{M}(A)$. Then μ_A is a measure on $\mathfrak{M}(A)$. Furthermore, the set of all A-integrable functions on P coincides with the set of all μ_A-integrable functions on P. For every such a function f the equality $\int f = \int_P f \, \mathrm{d}\mu_A$ holds. The measure μ_A is complete and possesses the following properties:

(a) the domain $\mathfrak{M}(A)$ of μ_A contains all Borel subsets of P;
(b) $\mu_A K < \infty$ for every compact $K \subset P$;
(c) $\mu_A G = \sup\{\mu_A K \colon K \subset G, K \text{ compact }\}$ for every open set $G \subset P$;
(d) $\mu_A M = \inf\{\mu_A G \colon G \supset M, G \text{ open }\}$ for every $M \in \mathfrak{M}(A)$.

14.8. Examples. (a) If $z \in P$ and $A \colon f \mapsto f(z)$ is the Dirac integral at z, then μ_A is the Dirac measure at z (note that the Dirac measure is defined on the σ-algebra of all subsets of P).

(b) If A_φ is the Riemann-Stieltjes integral determined on $\mathbf{R}$ by a nondecreasing function φ, then the associated measure $\lambda_\varphi := \mu_{A_\varphi}$ is called the *Lebesgue-Stieltjes measure*. Show that the Lebesgue measure corresponds to $\varphi(x) = x$.

(c) The Radon measure corresponding to the Poisson integral (Example 14.2.g) is called the *harmonic measure* at x.

14.9. Remark. The explanation in 14.5 – 14.7 would be much longer if all proclaimed propositions were proved. In this manuscript, we choose another approach. We will prove the existence of μ_A (the so called Riesz representation theorem) as directly as possible and the extension of a Radon integral to the collection of all A-integrable functions we obtain as $f \mapsto \int_P f \, \mathrm{d}\mu_A$.

Although there are several approaches with different proofs, their most difficult parts are based on similar ideas.

14.10. Signed and Complex Radon Integrals. A linear functional A on the space $\mathscr{C}_c(P)$ is called a *signed Radon integral* on P if for any compact set K there exists a constant a_K such that $|A(f)| \leq a_K \sup_K |f|$ whenever $f \in \mathscr{C}_K(P)$. Analogously we define *complex Radon integrals* on P. (They are functionals on the space $\mathscr{C}_c(P, \mathbf{C})$ of all complex valued continuous functions on P with a compact support.)

Any difference of positive Radon integrals serves as an example of a signed Radon integral. Soon we show that all signed Radon integrals are of this form.

14.11. Variation of Signed and Complex Radon Integrals. Let A be a signed or complex Radon integral on P. The *variation* of A is the (positive) Radon integral $|A|$ defined in the following steps:

1. If $f \in \mathscr{C}_c(P)$ is a nonnegative function, define

$$|A|(f) = \sup\{Ag\colon g \in \mathscr{C}_c(P),\ |g| \le f\}.$$

Apparently, $0 \le |A|(f) < +\infty$. If $f_1, f_2 \in \mathscr{C}_c(P)$ are nonnegative, then $|A|(f_1 + f_2) \ge |A|(f_1) + |A|(f_2)$. To prove the reverse inequality we use the Riesz decomposition lemma: Given $g \in \mathscr{C}_c(P)$, $|g| \le f_1 + f_2$ there exist $g_1, g_2 \in \mathscr{C}_c(P)$ such that $g = g_1 + g_2$, $|g|_j \le f_j$ for $j = 1, 2$. (It is not hard to see that the functions $g_j := f_j g(f_1 + f_2)^{-1}$ on $G := \{f_1 + f_2 > 0\}$ and zero on $P \setminus G$ have all the desired properties.) Thus $|A|(f_1) + |A|(f_2) \ge Ag_1 + Ag_2$, and taking sup over all g, we obtain $|A|(f_1 + f_2) \le |A|(f_1) + |A|(f_2)$.

2. If $f \in \mathscr{C}_c(P)$ is arbitrary, define

$$|A|(f) = |A|(f^+) - |A|(f^-).$$

To prove the additivity of $|A|$, note that $f_1^+ + f_2^+ + (f_1 + f_2)^- = f_1^- + f_2^- + (f_1 + f_2)^+$ and use the previous step. Since $|A|(\gamma f) = \gamma |A|(f)$ for every $f \in \mathscr{C}_c(P)$ and every $\gamma \in \mathbf{R}$, the functional $|A|$ is linear.

14.12. Decomposition of Signed and Complex Radon Integrals. Every complex Radon integral can be decomposed into its real and imaginary part which are signed Radon integrals. If A is a signed Radon integral, then A can be expressed in the form $A = A^+ - A^-$ where $A^+ := \frac{1}{2}(|A| + A)$ (the *positive variation*) and $A^- := \frac{1}{2}(|A| - A)$ (the *negative variation*) are positive Radon integrals. The representation of a signed Radon integral as a difference of positive Radon integrals is not unique; however, the decomposition to the positive and negative parts is "minimal" in a similar sense as the Jordan decomposition of a measure.

14.13. Exercise (product of Radon integrals). Let A_1, A_2 be Radon integrals on locally compact spaces P_1, P_2. Show that there exists exactly one Radon integral A on $P_1 \times P_2$ such that

$$A\,f = A_1 f_1 \cdot A_2 f_2$$

whenever $f_1 \in \mathscr{C}_c(P_1)$, $f_2 \in \mathscr{C}_c(P_2)$ and $f(x_1, x_2) = f_1(x_1) f_2(x_2)$, $x_1 \in P_1$, $x_2 \in P_2$. Prove the similar assertion also for signed and complex Radon integrals.

Hint. According to the Stone-Weierstrass theorem, the set of all linear combinations of functions of the form $f_1(x) \cdot f_2(y)$, where $f_i \in \mathscr{C}_c(P_i)$, is dense in $\mathscr{C}_c(P_1 \times P_2)$.

14.14. Notes. When building up the integration theory, some authors do not start from the original notion of a measure; they consider linear functionals on spaces of functions defined on an arbitrary set and then they extend these functionals to larger collections of functions. More precisely, it is possible to start with a *Riesz lattice* $\mathscr{R}$ of functions on a set X (which is a linear space of functions closed under formations of finite maxima and minima) and with a positive linear functional on $\mathscr{R}$ which satisfies "Daniell's condition". As in 14.5, this functional is extended, a measure is defined in a natural way and its properties are derived. This method was worked out by P.J. Daniell [1918] (for the extension he used sequences of functions) and M.H. Stone in the series of articles [1948] and [1949] (using generalized sequences). Let us note that many authors use different forms of a similar approach (W.H. Young [1904], H.H. Golstine [1941], J. Mařík [1952] and others).

The idea of extending the Radon integral is also due to many authors. However, it seems that Bourbaki's group was the first who worked it out for general locally compact spaces.

Also many authors confine themselves to the integration theory on locally compact spaces. Kakutani's representation theorem (S. Kakutani [1941]) says that any "abstract" functional A on a Riesz lattice $\mathscr{R}$ can be represented as a Radon integral A_K on a locally compact space P_K so that the corresponding spaces of "integrable" functions are isomorphic. However, Kakutani's representation also has some disadvantages — when changing the original functional, the space P changes as well and if the original space is locally compact, Kakutani's representation can lead to a different space. Supposing that the Riesz lattice $\mathscr{R}$ satisfies *Stone's condition* ($\min(1, f) \in \mathscr{R}$ if $f \in \mathscr{R}$), H. Bauer in [1957] constructed a different representation of (P_B, A_B) which removes mentioned disadvantages; the space P_B does not depend on A and if $\mathscr{R}$ is the space of all continuous functions with a compact support on a locally compact space P, P_B can be identified with P, $\mathscr{C}_c(P_B)$ with $\mathscr{R}$ and A_B with A.

Of course, no such representation can save the topology on the original space if it is not locally compact.

Many works are also concerned with integration theories in topological spaces without the assumption of local compactness (for instance in separable metric spaces) but in this case the correspondence between measure and the topological structure is not so fruitful.

15. Radon Measures

15.1. Radon Measures. Properties of the measure μ_A from the previous chapter (14.7) lead to the following definition.

Let P be a locally compact space and $\mathscr{B}(P)$ the σ-algebra of all Borel subsets of P. We say that μ is a *Radon measure* on $(P, \mathscr{S})$ if

(a) $\mathscr{S}$ is a σ-algebra containing $\mathscr{B}(P)$;
(b) $\mu K < \infty$ for every compact set $K \subset P$;
(c) $\mu G = \sup\{\mu K : K \subset G,\ K \text{ compact }\}$ for every open set $G \subset P$;
(d) $\mu A = \inf\{\mu G : G \supset A,\ G \text{ open }\}$ for every $A \in \mathscr{S}$.

15.2. Remarks. 1. There is a one-to-one correspondence between Radon integrals on P and Radon measures on $\mathscr{B}(P)$. To each Radon integral, we can assign a Radon measure μ_A. (See 14.7.; an alternative, more precisely developed approach will be used in 16.4.) Then we can restrict μ_A to $\mathscr{B}(P)$. On the other hand, if μ is a Radon measure on $\mathscr{B}(P)$, then $\mu = \mu_A$ on $\mathscr{B}(P)$, where A is the Radon integral $f \mapsto \int_P f \, d\mu$. Therefore, to give examples of Radon measures it is enough to list examples of Radon integrals.

2. In order to describe the structure of the system of all Radon measures on P, we define an equivalence relation: Radon measures μ_1 on $(P, \mathscr{S}_1)$ and μ_2 on $(P, \mathscr{S}_2)$ are said to be equivalent (at least for the purpose of this remark) if $\mu_1 = \mu_2$ on $\mathscr{B}(P)$. Then, of course, $\mu_1 = \mu_2$ on $\mathscr{S}_1 \cap \mathscr{S}_2$. In each equivalence class we can find two special representatives: the "minimal" one which is defined on $\mathscr{B}(P)$ and the "maximal" one which is complete. However, not every complete Radon measure is "maximal". See also Exercise 15.20. The Borel σ-algebra plays the unifying role, all "minimal" Radon measures are defined on it while domains of "maximal" Radon measures can be different (for instance, there are Lebesgue nonmeasurable sets while every set is measurable with respect to the complete Dirac measure mentioned in Example 14.8.a).

3. It is worthwhile to mention that definitions of Radon measures may vary from one author to another. We have seen that to each Radon integral we can assign an outer Radon measure and the whole scale of equivalent Radon measures. There is no general agreement which of these objects should be called a Radon measure. Some authors even use the term Radon measure for a

measure which is inner regular (i.e. $\mu E = \sup\{\mu K\colon K \subset E,\ K \text{ compact }\}$ for every $A \in \mathscr{S}$). In more general spaces than the σ-compact ones this leads to a different concept than our definition which requires outer regularity. Also the terms *regular Borel measure* or *Borel regular measure* are frequently and ambiguously used for Radon measures or similar objects.

4. Suppose that P is separable, metrizable and that $\mathscr{S} = \mathscr{B}(P)$. Then the property (b) of 15.1 implies both (c) and (d). Indeed, as every open set is a countable union of compact sets, we get (c). If $\mathscr{A}$ is the collection of all Borel sets A satisfying

$$\mu(A \cap U) = \inf\{\mu G\colon G \supset A \cap U, G \text{ open }\}$$

for every open set U, then $\mathscr{A}$ is a σ-algebra containing all open sets. Therefore (d) holds as well.

Throughout, μ is a Radon measure on $(P, \mathscr{S})$.

15.3. Lemma. *If $E \in \mathscr{S}$, $\mu E < \infty$, then*

$$\mu E = \sup\{\mu K\colon K \subset E,\ K \text{ compact }\}.$$

Proof. Given $\varepsilon > 0$, there exist an open set $G \supset E$ and a compact set $K \subset G$ such that $\mu(G \setminus E) < \varepsilon$ and $\mu K > \mu G - \varepsilon$. Let V be an open set containing $G \setminus E$ with $\mu V < \varepsilon$. If $H := K \setminus V$, then H is compact,

$$E \setminus H \subset (E \cap K \setminus H) \cup (E \setminus K) \subset V \cup (G \setminus K)$$

and $\mu H > \mu E - 2\varepsilon$. ■

15.4. Lemma. *Let $f \geq 0$ be a μ-integrable function on P. Then*

(a) $$\int_P f \,\mathrm{d}\mu = \inf\{\int_P g \,\mathrm{d}\mu\colon g \in C_c^{\uparrow}, g \geq f\};$$

(b) $$\int_P f \,\mathrm{d}\mu = \sup\{\int_P h \,\mathrm{d}\mu\colon h \in C_c^{\downarrow}, 0 \leq h \leq f\}.$$

Proof. (a) Choose $\varepsilon > 0$ and find a sequence $\{s_k\}$ of simple functions, $0 \leq s_k \nearrow f$. Since $f = s_1 + \sum\limits_{k=2}^{\infty}(s_k - s_{k-1})$, there exist $A_j \in \mathscr{S}$ and $\alpha_j \geq 0$ so that $f = \sum\limits_{j=1}^{\infty} \alpha_j c_{A_j}$. As f is μ-integrable, we have $\mu A_j < \infty$ for all j and we can find open sets $G_j \supset A_j$ so that $\mu(G_j \setminus A_j) < 2^{-j}\varepsilon$. The function $g := \sum\limits_{j=1}^{\infty} \alpha_j c_{G_j}$ is nonnegative, lower semicontinuous and satisfies

$$\int_P g \,\mathrm{d}\mu \leq \int_P f \,\mathrm{d}\mu + \varepsilon.$$

(b) To prove the asssertion it suffices to show it for simple functions. Let $f = \sum\limits_{j=1}^{n} \alpha_j c_{M_j}$ be a μ-integrable function ($\alpha_1, \dots, \alpha_n \geq 0,\ M_1, \dots, M_n \in \mathscr{S}$ and

$\mu M_j < \infty$). Choose again an $\varepsilon > 0$. By virtue of Lemma 15.3 there are compact sets K_j $(j = 1, \dots, n)$ such that $K_j \subset M_j$ and

$$\mu K_j \geq \mu M_j - \frac{\varepsilon}{n\alpha_j}.$$

Then $h := \sum_{j=1}^{n} \alpha_j c_{K_j}$ is an upper semicontinuous function with a compact support, $0 \leq h \leq f$ and

$$\int_P h \, \mathrm{d}\mu \geq \int_P f \, \mathrm{d}\mu - \varepsilon.$$

■

15.5. Theorem. *Let μ be a complete Radon measure on P. A function f on P is μ-integrable if and only if for every $\varepsilon > 0$ there exist integrable functions $s \in C_c^{\uparrow}$ and $t \in C_c^{\downarrow}$ such that $t \leq f \leq s$ and*

$$\int_P (s - t) \, \mathrm{d}\mu < \varepsilon.$$

Proof. First assume that $f \in \mathscr{L}^1(\mu)$. Let $\varepsilon > 0$ be given. Cite Lemma 15.4 to get nonnegative functions $g \in C_c^{\uparrow}$, $h \in C_c^{\downarrow}$ such that $g \geq f^+$, $h \leq f^-$ and

$$\int_P g \, \mathrm{d}\mu \leq \int_P f^+ \, \mathrm{d}\mu + \varepsilon, \qquad \int_P h \, \mathrm{d}\mu \geq \int_P f^- \, \mathrm{d}\mu - \varepsilon.$$

Then $s := g - h \in C_c^{\uparrow}$, $s \geq f$ and

$$\int_P s \, \mathrm{d}\mu \leq \int f \, \mathrm{d}\mu + 2\varepsilon.$$

In a similar way, we manufacture a $t \in C_c^{\downarrow}$ such that $t \leq f$ and

$$\int_P t \, \mathrm{d}\mu \geq \int f \, \mathrm{d}\mu - 2\varepsilon.$$

This establishes the necessity.

For the converse, suppose that for every $k \in \mathbf{N}$ there exist integrable functions $s_k \in C_c^{\uparrow}$ and $t_k \in C_c^{\downarrow}$ such that $t_k \leq f \leq s_k$ and

$$\int_P (s_k - t_k) \, \mathrm{d}\mu < \frac{1}{k}.$$

No generality is lost with the assumption that the sequence $\{s_k\}$ is nonincreasing, for we may replace it by the sequence $\{s_1, \min(s_1, s_2), \min(s_1, s_2, s_3), \dots\}$, and that the sequence $\{t_k\}$ is nondecreasing. By the Lebesgue dominated convergence theorem

$$\int_P \lim(s_k - t_k) \, \mathrm{d}\mu = 0.$$

Thanks to Theorem 8.16, $\lim(s_k - t_k) = 0$ μ-almost everywhere. Thus $\lim s_k = f$ μ-almost everywhere and by the Lebesgue theorem with the dominating function $|s_1| + |t_1|$ we get that f is μ-integrable. ■

15.6. Corollary. *Let f be a μ-integrable function on P. Then*

$$\int_P f\,\mathrm{d}\mu = \inf\{\int_P s\,\mathrm{d}\mu \colon s \in \mathscr{C}_c^{\uparrow}, s \geq f\}.$$

15.7. Exercise. Show that a finite positive measure μ on $(P, \mathscr{S})$ is Radon if and only if for every $E \in \mathscr{S}$ and for every $\varepsilon > 0$ there exist a compact set K and an open set U such that $K \subset E \subset U$ and $\mu(U \setminus K) < \varepsilon$.

15.8. Signed and Complex Radon Measures. A *signed measure* on $(P, \mathscr{S})$ is said to be *Radon* if its positive and negative variations are Radon measures. A *complex measure* on $\mathscr{S}$ is *Radon* if its real and imaginary parts are signed Radon measures.

Prove the following proposition: For a locally compact space P and a complex measure μ on $(P, \mathscr{S})$, the following assertions are equivalent:

(i) μ is Radon;
(ii) $|\mu|$ is Radon;
(iii) if $E \in \mathscr{S}$ and $\varepsilon > 0$, then there exist a compact set K and an open set U so that $K \subset E \subset U$ and $|\mu A| < \varepsilon$ for each $\mathscr{S}$-measurable set $A \subset U \setminus K$.

Hint. Use the inequality

$$|\mu|\,(U \setminus K) \leq 4 \sup\{\mu A \colon A \in \mathscr{S},\ A \subset U \setminus K\}.$$

To prove this inequality, use the decomposition as in the proof of Theorem 6.11.

15.9. Exercise. Let μ be a Radon measure on P. Prove that the union G of an arbitrary collection $\mathscr{B}$ of μ-null open sets is again a μ-null open set.

Hint. Suppose $K \subset G$ is compact. Then K can be covered by a finite union of sets from $\mathscr{B}$, whence $\mu K = 0$. Now, the definition of a Radon measure (property (c)) shows that $\mu G = 0$.

15.10. Support of a Radon Measure. Let μ be a Radon measure on P. Define the *support* of μ as

$$\operatorname{supt} \mu = P \setminus \bigcup\{G \colon G \text{ open}, \mu G = 0\}.$$

In other words, $\operatorname{supt} \mu$ is the smallest closed set whose complement si of measure zero. According to the previous exercise, such a set always exists. If μ is a signed or complex measure, its support is defined as the support of its total variation $|\mu|$.

(a) Let μ be a positive measure. Show that $z \in \operatorname{supt} \mu$ if and only if $\int_P f\,\mathrm{d}\mu > 0$ for every nonnegative function $f \in \mathscr{C}_c(P)$ with $f(z) > 0$, which happens if and only if $\mu U > 0$ for every open neighbourhood U of z.

(b) If μ_1, μ_2 are Radon measures on P, then $\operatorname{supt}(\mu_1 + \mu_2) \subset \operatorname{supt} \mu_1 \cup \operatorname{supt} \mu_2$ with the equality if μ_1 and μ_2 are positive.

15.11. Exercise. Let μ be a Radon measure and E a Borel set. If E is open or of finite measure, then μ_E (see 2.4) is a Radon measure.

15.12. Exercise. Let μ be a Radon measure and f a nonnegative μ-measurable function on P. Show that μ_f (see 8.19) is a Radon measure in the following cases:

(a) f is a Borel function and μ_f is finite on compact sets;

(b) f is continuous on P.

Hint. Use the part (a).

15.13. Exercise. Suppose that μ is a Radon measure and $f \in L^1(\mu)$. Show that μ_f is a signed Radon measure.

15.14. Exercise. Let μ_0 be a σ-finite Radon measure and $\mu = \mu_a + \mu_s$ the Lebesgue decomposition of a finite (signed or complex) Radon measure μ with respect to μ_0. Show that μ_a and μ_s are Radon measures.

Hint. Use Exercise 15.11.

15.15. Exercise. (a) Show that the assertion of Lemma 15.3 continues to hold, provided E is a countable union of sets of finite measures.

(b) If a Radon measure μ is σ-finite, then

$$\mu E = \sup\{\mu K \colon K \subset E,\ K \text{ compact}\}$$

for every $E \in \mathscr{S}$, in particular for every Borel set $E \subset P$. (Recall that on σ-compact spaces Radon measures are σ-finite. Every locally compact space with a countable base of open sets is metrizable and σ-compact.)

(c) Let P be the Cartesian product $\mathbf{R} \times \mathbf{R}_d$ where $\mathbf{R}_d$ is $\mathbf{R}$ equipped with the discrete topology. Show that P is a locally compact space which is not σ-compact.

For every open set $G \subset P$ set $\mu G = \sum\limits_{y \in \mathbf{R}_d} \lambda G_y$ (see the notation in 11.1) and extend μ to a Radon measure on P.

If $B := \{0\} \times \mathbf{R}_d$, then B is a Borel set, $\mu B = \infty$ and $\mu K = 0$ for every compact set $K \subset P$.

15.16. Exercise. (a) Let μ be a Radon measure on P and $K \subset P$ compact. Show that $\{x \in K \colon \mu\{x\} > 0\}$ is countable. In particular, if μ is a Radon measure on a σ-compact space P, then $\{x \in P \colon \mu\{x\} > 0\}$ is countable.

(b) Let μ be a Radon measure on $\mathbf{R}^n$. Then the set $\{r > 0 \colon \mu\{x \in \mathbf{R}^n \colon |x| = r\} > 0\}$ is countable.

15.17. Exercise. Let μ be a Radon measure on $(P, \mathscr{S})$, $p \in [1, \infty)$. Show that the set $C_c(P)$ is dense in $\mathscr{L}^p(\mu)$.

Hint. With the help of Exercise 10.10.a, it is sufficient to approximate every set $E \in \mathscr{S}$ of finite measure by functions from $\mathscr{C}_K(P)$ in the L^p-norm. If $\varepsilon > 0$, then there exist an open set G and a compact set K (Lemma 15.7) so that $K \subset E \subset G$ and $\mu(G \setminus K) < \varepsilon$. Now, use Urysohn's lemma in order to find a function $\varphi \in \mathscr{C}_c(P)$ with $c_K \le \varphi \le c_G$.

15.18. Exercise. A Radon measure μ on P is called *discrete* (often also *atomic*) if there exists a set $S \subset P$ such that $\mu(P \setminus S) = 0$ and $\mu\{x\} \neq 0$ for all $x \in S$. A measure μ is called *continuous* if $\mu\{x\} = 0$ for each point $x \in P$.

(a) Show that a complex Radon measure μ on P is discrete if and only if there exists a sequence of numbers $\{c_j\}$ and points $x_j \in P$ so that $\sum\limits_j |c_j| < \infty$ and $\mu = \sum\limits_j c_j \varepsilon_{x_j}$. Cf. Exercise 2.10.

(b) Show that every Radon measure μ can be uniquely expressed in the form $\mu = \mu_d + \mu_c$ where μ_d is discrete and μ_c continuous.

Hint. Set $S := \{x \in P \colon \mu\{x\} > 0\}$, $\mu_d A = \mu(A \cap S)$ and $\mu_c A = \mu(A \setminus S)$ (use Exercise 15.16.a).

(c) Show that each Radon measure μ on $\mathbf{R}^n$ can be written uniquely in the form

$$\mu = \mu_d + \mu_a + \mu_s$$

where μ_d is discrete, μ_s is continuous, μ_a is absolutely continuous with respect to the Lebesgue measure λ and μ_s and λ are mutually singular.

15.19. Exercise. Let μ be a σ-finite Radon measure on P and $E \in \mathscr{S}$.

(a) Show that for every $\varepsilon > 0$ there exist an open set G and a closed set F such that $F \subset E \subset G$ and $\mu(G \setminus F) < \varepsilon$.

Hint. Suppose $E = \bigcup\limits_j E_j$, where E_j are of finite measure. Find open sets $G_j \supset E_j$ such that $\mu G_j < \mu E_j + \varepsilon 2^{-j-1}$ and set $G = \bigcup\limits_j G_j$. In a similar way find F (considering the complements).

(b) Show that there exists an F_σ set S and a G_δ set D such that $S \subset E \subset D$ and $\mu(D \setminus S) = 0$ (compare also with Theorem 1.21).

15.20. Exercise. Let μ be a Radon measure on $(P, \mathscr{S})$ and consider the outer measure

$$\mu^* A := \inf\{\mu(G) : \ G \text{ open}, \ G \supset A\}.$$

(a) If $A \in \mathscr{S}$, then A is μ^*-measurable in the sense of 4.4. and $\mu = \mu^*$ on $\mathscr{S}$. Thus the extension of μ to $\mathfrak{M}(\mu^*)$ is the "maximal" representative of μ in the sense of Remark 15.2.2.

(b) Let μ be as in Exercise 15.15.c. Show that the completion of μ is not the "maximal representative".

15.21. Notes. The important step from the Lebesgue measure to the study of more general measures in Euclidean spaces was done by J. Radon [1913]. The notion of a Radon measure is connected with his name, even if this term is not entirely common and some authors use other synonyms instead.

16. Riesz Representation Theorem

In this chapter, let us direct our attention to the relation between Radon measures and Radon integrals on a locally compact space P.

Suppose μ is a Radon measure on P. Then $\mathscr{C}_c(P) \subset \mathscr{L}^1(\mu)$ and the mapping

$$f \mapsto \int_P f \, \mathrm{d}\mu \,, \qquad f \in \mathscr{C}_c(P)$$

is a Radon integral on P.

On the other hand, every Radon integral A on P can be understood as an integral with respect to a Radon measure. In Chapter 14 we indicated a method of proving this proposition, and now we propose to show another method and provide complete proofs.

Symbols μ_A and μ_A^* of this chapter will be used to denote the same objects as previously in Chapter 14, but the definitions are revised.

In the sequel, A denotes a Radon integral on a locally compact space P.

16.1. Outer Radon Measure. If G is an open subset of P, we set

$$\mu_A^*(G) = \sup\{Af\colon f \in \mathscr{C}_c(P), \, 0 \le f \le 1, \, f = 0 \text{ on } P \setminus G\}.$$

Clearly μ_A^* is monotone on open sets. We define

$$\mu_A^*(E) = \inf\{\mu_A^*(G)\colon G \text{ open}, \, G \supset E\}$$

for an arbitrary set $E \subset P$.

The set function μ_A^* will be called the *outer Radon measure* (assigned to the Radon integral A). In the next theorem we will prove that μ_A^* (which will be simply denoted by μ^*) is really an outer measure.

16.2. Properties of Outer Radon Measures. *Let μ^* be an outer Radon measure. Then*

(a) $\mu^* K = \inf\{Ag\colon g \in \mathscr{C}_c(P),\ 0 \le g \le 1,\ g = 1$ *on* $K\}$ *for every compact set* $K \subset P$ *(in particular, μ^* is finite on compact sets);*

(b) $\mu^* G = \sup\{\mu^* K\colon K$ *compact,* $K \subset G\}$ *for every open set* $G \subset P$*;*

(c) μ^* *is an outer measure.*

Proof. (a) Let $K \subset P$ be a compact set and $g \in \mathscr{C}_c(P)$, $0 \le g \le 1$, $g = 1$ on K. Fix $\varepsilon \in (0,1)$ and denote $G = \{x \in P\colon g(x) > 1 - \varepsilon\}$. Obviously $K \subset G$. If $f \in \mathscr{C}_c(P)$, $0 \le f \le 1$, $f = 0$ on $P \setminus G$, then $f \le \frac{1}{1-\varepsilon} g$. Hence $\mu^* G \le \frac{1}{1-\varepsilon} Ag$ and

$$Ag \ge (1-\varepsilon)\mu^* G \ge (1-\varepsilon)\mu^* K\,.$$

We see that $Ag \ge \mu^* K$, and thus

$$\mu^* K \le \inf\{Ag\colon g \in \mathscr{C}_c(P),\ 0 \le g \le 1,\ g = 1 \text{ on } K\}\,.$$

To prove the reverse inequality, select an open set $G \supset K$. Urysohn's lemma provides a function $g \in \mathscr{C}_c(P)$, $0 \le g \le 1$, $g = 0$ on $P \setminus G$ and $g = 1$ on K. Since $Ag \le \mu^* G$, we are done.

(b) Suppose we are given an open set $G \subset P$ and $\varepsilon > 0$. Let $f \in \mathscr{C}_c(P)$ be a function such that $0 \le f \le 1$, $f = 0$ on $P \setminus G$ and choose $\varepsilon > 0$. Since $f_k := \min(f, \frac{1}{k}) \searrow 0$, an appeal to Daniell's property reveals the existence of an $n \in \mathbf{N}$ such that $Af_n < \varepsilon$. If $K := \left\{x \in P\colon f(x) \ge \frac{1}{n}\right\}$, then K is a compact subset of P. According to (a), there exists $g \in \mathscr{C}_c(P)$ such that $0 \le g \le 1$, $g = 1$ on K and $Ag \le \mu^* K + \varepsilon$. Since $f - f_n \le g$, we get

$$Af \le Af_n + Ag \le \mu^* K + 2\varepsilon$$

and we have finished.

(c) Clearly $\mu^* \emptyset = 0$, and $\mu^* S \le \mu^* T$ whenever $S \subset T$. It remains to show that μ^* is σ-subadditive. To this end, let $\varepsilon > 0$ and K_1, $K_2 \subset P$ be given. By Theorem 16.2.a we can find $f_j \in \mathscr{C}_c(P)$, $j = 1,2$ such that $f_j = 1$ on K_j and $Af_j \le \mu^* K_j + \varepsilon$. Then $\mu^*(K_1 \cup K_2) \le A(f_1 + f_2) = Af_1 + Af_2 \le \mu^* K_1 + \mu^* K_2 + 2\varepsilon$ and we see that μ^* is subadditive on compact sets.

Next consider open sets $G_1, G_2 \subset P$ and pick a compact set $K \subset G_1 \cup G_2$. For any point $x \in K$ there is its neighbourhood V_x whose closure is either in G_1 or in G_2. Thanks to the compactness of K, we obtain finite collections of open sets $\{V_i^1\}_i$, $\{V_i^2\}_i$ such that $\overline{V_i^j} \subset G_j$ and $\bigcup_i V_i^1 \cup \bigcup_i V_i^2 \supset K$. Set $K_j = K \cap \bigcup_i \overline{V_i^j}$, $j = 1,2$. Then K_j are compact, $K_j \subset G_j$ and $K = K_1 \cup K_2$. Thus $\mu^* K \le \mu^* K_1 + \mu^* K_2 \le \mu^* G_1 + \mu^* G_2$ and it readily follows that μ^* is subadditive on open sets.

Now, let $\{G_n\}$ be a sequence of open sets. Choose a compact set $K \subset \bigcup_{i=1}^{\infty} G_i$. Then $K \subset \bigcup_{i=1}^{n} G_i$ for some $n \in \mathbf{N}$, whence

$$\mu^* K \le \mu^*\left(\bigcup_{i=1}^{n} G_i\right) \le \sum_{i=1}^{n} \mu^* G_i \le \sum_{i=1}^{\infty} \mu^* G_i\,,$$

and (b) yields that $\mu^*\left(\bigcup_{i=1}^{\infty} G_i\right) \le \sum_{i=1}^{\infty} \mu^* G_i$.

Finally, let $E_n \subset P$ be arbitrary. We wish to show that $\mu^*\left(\bigcup E_n\right) \le \sum \mu^* E_n$. It would be clearly sufficient to assume that $\mu^* E_n < \infty$ for all n. Given $\varepsilon > 0$, we can find open sets G_n such that $E_n \subset G_n$ and $\mu^* G_n < \mu^* E_n + 2^{-n}\varepsilon$. Then

$$\mu^*\left(\bigcup E_n\right) \le \mu^*\left(\bigcup G_n\right) \le \sum \mu^* E_n + \varepsilon\,.$$

As $\varepsilon > 0$ was arbitrary, μ^* is σ-subadditive as needed. ■

16.3. Theorem. *Every Borel subset of P is μ^*-measurable (in the sense of Carathéodory).*

Proof. It is sufficient to prove measurability of open sets. Notice that by virtue of Theorem 16.2.c we have $\mu^*(G_1 \cup G_2) = \mu^* G_1 + \mu^* G_2$ whenever G_1, G_2 are disjoint open sets. Now, given an open set $G \subset P$, a test set $T \subset P$ such that $\mu^* T < \infty$ and $\varepsilon > 0$, there exist open sets V and H such that $V \supset T$, $\mu^* V < \mu^* T + \varepsilon$ and $H \supset V \setminus G$, $\mu^* H < \mu^*(V \setminus G) + \varepsilon$. We can also find a compact set $K \subset V \cap G$ such that $\mu^* K + \varepsilon > \mu^*(V \cap G)$ and an open set W with a compact closure such that $K \subset W \subset \overline{W} \subset V \cap G$. Set $W_0 = V \cap H \setminus \overline{W}$. Then W_0 is an open set, $W \cap W_0 = \emptyset$, $W \cup W_0 \subset V$, $V \setminus G \subset W_0$ and $\mu^* W + \varepsilon > \mu^*(V \cap G)$. Thus

$$\begin{aligned}\mu^*(T \cap G) + \mu^*(T \setminus G) &\le \mu^*(V \cap G) + \mu^*(V \setminus G) < \mu^* W + \varepsilon + \mu^* W_0 \\ &= \mu^*(W \cup W_0) + \varepsilon \le \mu^* V + \varepsilon < \mu^* T + 2\varepsilon.\end{aligned}$$

(In fact, the idea of the proof is quite simple, $T \cap G$ and $T \setminus G$ are approximated by disjoint open sets.) ■

16.4. Measure μ_A. Any Radon integral A on a locally compact space P defines a corresponding outer Radon measure μ_A^*. Carathéodory's construction yields the σ-algebra $\mathfrak{M}_A$ on which μ_A^* is a complete Radon measure. The restriction of μ_A^* to $\mathfrak{M}_A$ will be denoted by μ_A.

16.5. Riesz Representation Theorem. *Let A be a Radon integral on P. Then there exists a complete Radon measure μ on P such that $\mathscr{C}_c(P) \subset \mathscr{L}^1(\mu)$ and $Af = \int_P f \,\mathrm{d}\mu$ for each $f \in \mathscr{C}_c(P)$. The measure μ is unique on $\mathscr{B}(P)$.*

Proof of the existence. If $\mu = \mu_A$, then μ is a complete Radon measure. To complete the proof we only have to show that

$$Af = \int_P f \,\mathrm{d}\mu, \quad f \in \mathscr{C}_c(P).$$

The reader should find it easy to prove that $\mathscr{C}_K(P) \subset \mathscr{L}^1(\mu)$.

Let $f \in \mathscr{C}_c(P)$ be given. It is no restriction to assume that $0 \le f \le 1$. For $n \in \mathbf{N}$ and $k = 0, 1, \ldots, n$ denote

$$f_k := \min\Big(f, \frac{k}{n}\Big), \qquad G_k := \Big\{x \in P \colon f(x) > \frac{k}{n}\Big\}\,.$$

Then using the definition of μ^*G and the properties of integral it is clear that

$$\frac{1}{n}\mu G_k \le A(f_k - f_{k-1}) \le \frac{1}{n}\mu G_{k-1} \quad \text{and} \quad \frac{1}{n}\mu G_k \le \int_P (f_k - f_{k-1})\, d\mu \le \frac{1}{n}\mu G_{k-1}$$

for each $k = 0, 1, \dots, n$. Thus

$$\begin{aligned} \left| Af - \int_P f \,\mathrm{d}\mu \right| &= \left| \sum_{k=1}^{n} \left(A(f_k - f_{k-1}) - \int_P (f_k - f_{k-1})\,\mathrm{d}\mu \right) \right| \\ &\le \sum_{k=1}^{n} \frac{1}{n}(\mu G_{k-1} - \mu G_k) = \frac{1}{n}\mu G_0 = \frac{1}{n}\mu\{x \in P\colon f(x) > 0\}. \end{aligned}$$

Since $\mu\{x \in P\colon f(x) > 0\} < +\infty$, we get $Af = \int_P f \,\mathrm{d}\mu$ and we are done.

Proof of the uniqueness. If ν is a complete Radon measure on P obeying the assumptions of the theorem, we have

$$Af = \int_P f \,\mathrm{d}\nu \le \int_P c_G \,\mathrm{d}\nu = \nu G$$

for any open set $G \subset P$ and $f \in \mathscr{C}_K(P)$ with $0 \le f \le 1$, $f = 0$ on $P \setminus G$. Whence $\mu_A G = \mu_A^* G \le \nu G$. Taking into account Theorem 16.2.a, we have $\mu_A K \ge \nu K$ for each compact set and from regularity of ν and μ_A it immediately follows that ν and μ_A coincide on Borel sets. ■

16.6. Remark. The previous theorem guarantees a one-to-one correspondence between Radon integrals and classes of equivalence of Radon measures (cf. Remark 15.2).

16.7. Other Spaces of Continuous Functions. We denote by $\mathscr{C}_b(P)$ the Banach space of all bounded continuous functions on P equipped with the norm

$$\|f\| = \sup_{x \in P} |f(x)|\,.$$

The space

$$\begin{aligned} \mathscr{C}_0(P) := \{ f \in \mathscr{C}(P)\colon &\text{ for every } \varepsilon > 0 \text{ there exists a compact set } K_\varepsilon \subset P \\ &\text{such that } |f(x)| < \varepsilon \text{ for all } x \text{ outside } K_\varepsilon \} \end{aligned}$$

of all continuous functions on P "vanishing at infinity" is the closure of $\mathscr{C}_c(P)$ in $\mathscr{C}_b(P)$. If μ is a finite Radon measure on P, then

$$f \mapsto \int_P f \,\mathrm{d}\mu$$

is a positive linear functional on $\mathscr{C}_b(P)$ and on $\mathscr{C}_0(P)$ as well. The converse proposition is also true for the space $\mathscr{C}_0(P)$ as an easy consequence of the Riesz Representation Theorem 16.5.

Let A be a positive linear functional on $\mathscr{C}_0(P)$. Then there exists a unique finite complete Radon measure μ on P such that $\mathscr{C}_0(P) \subset \mathscr{L}^1(\mu)$ and $Af = \int_P f \, d\mu$ for every function $f \in \mathscr{C}_0(P)$.

16.8. Exercises. Let P, Q be locally compact spaces, h a continuous mapping from P onto Q and μ a Radon measure on $(P, \mathscr{B}(P))$.

(a) Show that $f \circ h \in \mathscr{L}^1(\mu)$ for every function $f \in \mathscr{C}_c(Q)$ provided $\mu P < +\infty$ or $h^{-1}(F)$ is compact for every compact $F \subset Q$.

(b) Show that the mapping $A\colon f \mapsto \int_P f \circ h \, d\mu$, $f \in \mathscr{C}_c(Q)$, is a Radon integral on Q.

(c) By the Riesz Representation Theorem there exists a unique Radon measure μ' on $(Q, \mathscr{B}(Q))$ which represents A. Prove that $\mu' = h(\mu)$ (notation as in Exercise 8.23).

Hint. Show that $\mu'(K) = f(\mu)(K)$ for every compact set K.

16.9. Product of Radon Measures. Consider locally compact spaces P_1 and P_2. Their Cartesian product $P_1 \times P_2$ is again a locally compact space. If P_1 and P_2 have countable bases, then the product $P_1 \times P_2$ has a countable base as well.

We are now interested in the product of two Radon measures. Two main problems arise:

1. In general, a Radon measure is not necessarily σ-finite and we cannot speak about a product measure in the sense of Chapter 11.

2. Although $\mathscr{B}(P_1) \otimes \mathscr{B}(P_2) \subset \mathscr{B}(P_1 \times P_2)$, the equality in general does not hold. Thus even if the original Radon measures on P_1 and P_2 are σ-finite, their product is not necessarily defined on $\mathscr{B}(P_1 \times P_2)$ and so it cannot be called a Radon measure.

In the sequel, we show a way how to define a product of Radon measures.

(a) Let P_1, P_2 be locally compact spaces with countable bases (and so metrizable). Then $\mathscr{B}(P_1 \times P_2) = \mathscr{B}(P_1) \otimes \mathscr{B}(P_2)$ and if μ_1, μ_2 are Radon mesures on $(P_1), \mathscr{B}(P_1))$, $(P_2, \mathscr{B}(P_2))$, then $\mu_1 \otimes \mu_2$ is a Radon measure on $(P_1 \times P_2, \mathscr{B}(P_1 \times P_2))$.

Hint. It is not hard to verify that $\mathscr{B}(P_1) \otimes \mathscr{B}(P_2) = \mathscr{B}(P_1 \times P_2)$. Hence $\mu_1 \otimes \mu_2$ is defined on $\mathscr{B}(P_1 \times P_2)$. Since $P_1 \times P_2$ has a countable basis, it is sufficient to show that $\mu_1 \otimes \mu_2$ is finite on compact sets. Select a compact set $K \subset P_1 \times P_2$ and denote K_i projections of K onto P_i. Then K_1 and K_2 are compact sets (as continuous images of compact sets) and $K \subset K_1 \times K_2$. Then

$$\mu_1 \otimes \mu_2 \,(K) \le \mu_1 \otimes \mu_2 \,(K_1 \times K_2) = \mu_1 K_1 \cdot \mu_2 K_2 < +\infty.$$

(b) Consider now the general case when P_1 and P_2 have not necessarily countable bases. By Exercise 14.13 there is a Radon integral A on $P_1 \times P_2$ such that

$$Af = A_1 f_1 \cdot A_2 f_2$$

whenever $f_1 \in \mathscr{C}_c(P_1)$, $f_2 \in \mathscr{C}_c(P_2)$ and $f(x_1, x_2) = f_1(x_1) f_2(x_2)$, $x_1 \in P_1$, $x_2 \in P_2$. Show that μ_A is a unique Radon measure μ on $\mathscr{B}(P_1 \times P_2)$ satisfying

$$\mu(E_1 \times E_2) = \mu_1 E_1 \cdot \mu_2 E_2$$

whenever $E_1 \in \mathscr{B}(P_1)$ and $E_2 \in \mathscr{B}(P_2)$. If, in addition, μ_1, μ_2 are σ-finite, then

$$\mu_A(E) = (\mu_1 \otimes \mu_2) E$$

for every set $E \in \mathscr{B}(P_1) \otimes \mathscr{B}(P_2)$.

(c) Using results of 16.10, it is possible to define a product of complex measures as well.

16.10. Representations of Signed and Complex Radon Integrals. (a) Let ν be a signed or complex Radon measure on P and a a bounded Borel function on $\mathscr{M}(P)$. Then the mapping

$$A\colon f \mapsto \int_P af \, d\nu$$

is a complex Radon integral. If, in addition, $\nu P < +\infty$, then A is a continuous linear functional on $\mathscr{C}_0(P)$.

(b) Let ν be a finite signed or complex Radon measure on P. Then the mapping

$$A\colon f \mapsto \int_P f\,\mathrm{d}\nu$$

is a continuous linear functional on the space $\mathscr{C}_0(P)$.

(c) Let A be a complex Radon integral on P (signed Radon integrals are particular cases) and let $|A|$ be as in 14.11. By the Riesz representation theorem there is a Radon measure μ such that $\int f\,d\mu = |A|\,f$ for each $f \in \mathscr{C}_c(P)$. Show that there exists a bounded Borel function a (uniquely determined as an element of $L^\infty(P,\mu)$) such that the equality $Af = \int_P af\,\mathrm{d}\mu$ holds for every function $f \in \mathscr{C}_c(P)$. Moreover, $|a| = 1$ μ-almost everywhere.

Hint. The function a can be defined locally, so the problem can be reduced to the case when P is compact. Our approach will be analogous to that in proving the Radon-Nikodým theorem. The functional A is uniformly continuous on $\mathscr{C}(P)$ with respect to the norm of the Hilbert space $L^2(P,\mu)$, thus it can be continuously extended to the entire space L^2 and the Riesz Representation Theorem on representation of continuous linear functionals on a Hilbert space yields the existence of an element $a \in L^2$ such that $Au = \int au\,\mathrm{d}\mu$ for all $u \in L^2$. Proof of the fact that $|a| = 1$ μ-almost everywhere is more difficult, see e.g. G.K. Pedersen [*1989].

(d) Let A be a continuous linear functional on $\mathscr{C}_0(P)$. Then there exists a unique complete signed (or complex, depends on what field we consider) measure μ on P such that $\mathscr{C}_0(P) \subset \mathscr{L}^1(|\mu|)$ and $Af = \int_P f\,\mathrm{d}\mu$ for each $f \in \mathscr{C}_0(P)$.

16.11. Notes. The famous Riesz Representation Theorem 16.5 was proved for $P = [0,1]$ by F. Riesz [1909b] but continuous linear functionals on $\mathscr{C}([0,1])$ were represented by functions of bounded variation using Riemann-Stieltjes integrals. Other theorems of this type were proved by J. Radon [1913] (for compact subsets of $\mathbf{R}^n$), S. Banach in Appendix to Saks' monograph [*1937], S. Saks [1938] (for compact metric spaces) and S. Kakutani [1941]. The Riesz Representation Theorem for compact spaces and the method of construction of Radon measures directly from Radon integrals is due to J. von Neumann [1934]. Concerning locally compact spaces, A. Weil [1940] was aware about the result. However, the final version for locally compact spaces was established by the Bourbaki group (A. Weil was also its member) in the 40's and appeared in [*1952].

17. Sequences of Measures

Throughout this chapter, P will be a locally compact space. We denote by $\mathscr{M}(P)$ the linear space of all signed Radon integrals on P and by $\mathscr{M}^+(P)$ the set of all (positive) Radon integrals on P. If no confusion can result, we will not distinguish between the Radon integral A and the Radon measure μ_A. Thus, for $\mu \in \mathscr{M}^+(P)$, we will write both $\mu(f)$ $(f \in \mathscr{C}_c(P))$ and μE $(E \subset P)$.

This chapter is for information, proofs will be given only to some propositions.

17.1. Strong and Weak Convergence. Let $\mathscr{F}$ be a linear subspace of $\mathscr{C}(P)$ containing $\mathscr{C}_c(P)$. We say that a sequence $\{\mu_n\}$ of Radon integrals on P converges *$\mathscr{F}$-weakly* to a Radon integral μ if

$$\lim_{n\to\infty} \int_P f\,d\mu_n = \int_P f\,d\mu$$

for each $f \in \mathscr{F}$. Notice that the $\mathscr{F}$-weak limit, if exists, is uniquely determined.

The most important case is the $\mathscr{C}_c(P)$-weak convergence which is called the *vague* convergence and denoted by $\mu_n \xrightarrow{\mathrm{v}} \mu$.

From the point of view of functional analysis, the vague convergence is the weak* convergence. In case of spaces of measures, it is common to omit the asterisk. Different spaces of measures are duals to different spaces of continuous functions. If such a space $\mathscr{F}$ is equipped with a norm (or, more generally, with a locally convex topology), then it is possible to introduce a $\mathscr{F}$*-weak* topology on its dual $\mathscr{F}^*$ so that the $\mathscr{F}$-weak convergence of measures is in fact the convergence in this ($\mathscr{F}$-weak) topology. Notice that the space $\mathscr{F}^*$ is not (except a few not very interesting cases) metrizable. Hence the $\mathscr{F}$-weak topology cannot be described by convergence of sequences and, in general, nets have to be used instead.

If P is a compact metric space, then the set $\mathscr{M}^+(P)$ is metrizable in the $\mathscr{C}(P)$-weak topology.

In functional analysis, besides weak convergences we meet also strong convergences. The most important example of a strong type convergence of measures is the convergence $\|\mu_n - \mu\| \to 0$ on the space $\mathscr{M}_b(P)$ of all finite signed (or complex) Radon measures. The norm $\|\mu\|$ defined as $|\mu|(P)$ (like in Exercise 6.17) is the dual norm to the norm of $\mathscr{C}_0(P)$ and $\mathscr{M}_b(P)$ with this norm is a Banach space.

17.2. Comparison of Weak Convergences. Now we will touch the question whether the vague convergence $\mu_n \xrightarrow{v} \mu$ implies the $\mathscr{F}$-weak convergence for a wider space of "test" functions. Proofs of next propositions require the Banach-Steinhaus theorem of functional analysis.

(a) A sequence μ_n converges $\mathscr{C}_0(P)$-weakly to μ if and only if $\mu_n \xrightarrow{v} \mu$ and the sequence $\{\|\mu_n\|\}$ is bounded.

(b) As usually, by a *weak convergence* we understand the $\mathscr{C}_b(P)$-weak convergence where $\mathscr{C}_b(P)$ denotes the set of all bounded continuous functions on P. A sequence μ_n of complex measures from $\mathscr{M}_b(P)$ converges weakly to a measure $\mu \in \mathscr{M}_b(P)$ if and only in $\mu_n \xrightarrow{v} \mu$ and for every $\varepsilon > 0$ there exists a compact set $K \subset P$ so that $|\mu_n|(P \setminus K) < \varepsilon$ for all n.

A sequence μ_n of positive measures from $\mathscr{M}_b^+(P)$ converges weakly to a measure $\mu \in \mathscr{M}_b^+(P)$ if and only if $\mu_n \xrightarrow{v} \mu$ and $\|\mu_n\| \to \|\mu\|$.

(c) Note that the weak convergence implies the $\mathscr{C}_0(P)$-weak convergence and this one implies the vague convergence. If the space P is compact, then $\mathscr{C}_b(P) = \mathscr{C}_c(P) = \mathscr{C}(P)$ and there is no difference between these convergences.

17.3. Examples. (a) If $x_n \to x$, then $\varepsilon_{x_n} \xrightarrow{v} \varepsilon_x$.

(b) Suppose that $\{x_n\}$ is a sequence having no convergent subsequence (for example, take $x_n = n$ for $P = \mathbf{R}$) and $\{\alpha_n\}$ a sequence of real numbers. Then the sequence $\{\alpha_n \varepsilon_{x_n}\}$ converges vaguely to the null measure. This sequence converges (to the null measure) $\mathscr{C}_0(P)$-weakly if and only if α_n is a bounded sequence, and weakly if and only if $\alpha_n \to 0$.

(c) Suppose $x_n \to z$, $y_n \to z$, $x_n \neq y_n$. Then $\varepsilon_{x_n} - \varepsilon_{y_n} \xrightarrow{v} 0$ but $\|\varepsilon_{x_n} - \varepsilon_{y_n}\| \to 2 \neq 0$.

(d) If f is a continuous function on $[0, 1]$, then

$$\int_0^1 f(x)\,\mathrm{d}x = \lim_{k\to+\infty} \frac{1}{k} \sum_{i=1}^{k} f(\frac{i}{k})\,.$$

As if often the case, this equality can become the basis for a definition of the Riemann integral, but it cannot be used to describe the set of Riemann integrable functions. In the language of the weak convergence of measures it can be understood as

$$\frac{1}{k} \sum_{i=1}^{k} \varepsilon_{\frac{i}{k}} \xrightarrow{v} \lambda_{[0,1]}.$$

The above examples show that the weak or vague convergence (unlike the strong convergence, cf. Exercise 6.18) of μ_n to μ does not imply $\mu_n(A) \to \mu(A)$ for all (Borel) sets. Nevertheless, we can state the following theorem for compact spaces. Its modifications hold in locally compact spaces as well.

17.4. Theorem. *Let P be a compact space and $\mu_n, \mu \in \mathscr{M}^+(P)$, $\mu_n \xrightarrow{\mathrm{v}} \mu$.*

(a) *For any lower semicontinuous lower bounded function u on P we have*

$$\int_P u \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int_P u \,\mathrm{d}\mu_n.$$

(b) *If f is a bounded Borel function on P which is continuous μ-almost everywhere, then*

$$\int_P f \,\mathrm{d}\mu = \lim_{n\to\infty} \int_P f \,\mathrm{d}\mu_n.$$

(c) *If $G \subset P$ is an open set, then $\liminf \mu_n(G) \geq \mu(G)$.*

(d) *If $K \subset P$ is a compact set, then $\limsup \mu_n(K) \leq \mu(K)$.*

(e) *If A is a Borel set such that $\mu(\partial A) = 0$, then $\mu_n(A) \to \mu(A)$.*

Proof. (a) is an easy consequence of the definition.

(b) Define f^*, f_* as in 7.9.b. If $f^* = f_*$ μ-almost everywhere, then f is μ-measurable and

$$\int_P f \,\mathrm{d}\mu = \int_P f_* \,\mathrm{d}\mu \leq \liminf_{n\to\infty} \int_P f_* \,\mathrm{d}\mu_n \leq \liminf_{n\to\infty} \int_P f \,\mathrm{d}\mu_n$$
$$\leq \limsup_{n\to\infty} \int_P f \,\mathrm{d}\mu_n \leq \limsup_{n\to\infty} \int_P f^* \,\mathrm{d}\mu_n \leq \int_P f^* \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu.$$

To prove (c), (d) and (e) we can apply (a) and (b) to indicator functions. ■

17.5. Remark. Compare Theorem 17.4 with the well-known result that a bounded function f is Riemann integrable if and only if the set of all points of discontinuity of f is of measure zero (see 7.9.d). It is possible to prove that a bounded Borel function f is Riemann integrable on $[0,1]$ if and only if $\int_{[0,1]} f \,\mathrm{d}\mu_n \to \int_0^1 f \,\mathrm{d}\lambda$ for every sequence μ_n of positive measures converging weakly to λ on $[0,1]$.

17.6. Exercise. Suppose $\mu_n, \mu \in \mathscr{M}^+(P)$. Show that $\mu_n \xrightarrow{\mathrm{v}} \mu$ if and only $\limsup \mu_n(K) \leq \mu(K)$ for each compact set $K \subset P$ and $\liminf \mu_n(G) \geq \mu(G)$ for each open set $G \subset P$.

17.7. Molecular Measures. A Radon measure ν is called *molecular* if $\nu = \sum_{i=1}^{k} \alpha_i \varepsilon_{x_i}$ where $x_1, \dots, x_k \in P$ and $\alpha_1, \dots, \alpha_k$ are positive and $\sum_{i=1}^{k} \alpha_i = 1$. As we have seen in Example 17.3.d, the definition of the Riemann integral is closely related to an approximation of "continuous" Lebesgue measure by "discrete" molecular measures. In a similar way, more general measures can be approximated as the next theorem shows.

17.8. Theorem. *Let μ be a (positive) Radon measure on a compact metric space P. Then there exists a sequence $\{\mu_n\}$ of molecular measures such that $\mu_n \xrightarrow{\mathrm{v}} \mu$.*

Proof. We sketch the main idea of the proof. Thanks to the compactness of P, for every $k \in \mathbf{N}$ there is a finite collection $\{M_k^i\}$ of pairwise disjoint nonempty sets so that $\bigcup_i M_k^i = P$ and $\operatorname{diam} M_k^i \le 2^{-k}$. Choose points $x_k^i \in M_k^i$ and set

$$\mu_k = \sum_i \mu(M_k^i)\varepsilon_{x_k^i}.$$

■

17.9. Remark. The previous theorem can be proved as a nice application of the Krejn-Milman theorem or of the bipolar theorem. However, both of them belong to deeper theorems of functional analysis.

The following theorem has also its interpretation in the language of functional analysis: On dual spaces bounded sets are relatively sequentially compact in the weak* topology.

17.10. Theorem. *Let P be a metric compact space and $\{\mu_n\}$ a sequence of complex Radon measures on P. If $\sup_n \|\mu_n\| < +\infty$, then there exists a weakly convergent subsequence of $\{\mu_n\}$.*

Proof. The method of the proof uses the Cantor diagonal selection process. Since $\mathscr{C}(P)$ is separable, there is a countable dense set $\{f_k\} \subset \mathscr{C}(P)$. Now step by step construct a sequence of sequences of measures $\{\{\mu_n^k\}_n\}_k$ so that $\mu_n^0 = \mu_n$ and every $\{\mu_n^k\}_n$ $(k \ge 1)$ is a subsequence of $\{\mu_n^{k-1}\}_n$ for which the sequence of numbers $\{\mu_n^k(f_k)\}_n$ is convergent. This can be done by the Bolzano-Weierstrass theorem since the sequence $\{\mu_n^{k-1}(f_k)\}_n$ of real numbers is bounded. If $\nu_n := \mu_n^n$, then $\{\nu_n\}$ is a subsequence of $\{\mu_n\}$ and since for every k the sequence $\{\nu_n\}_{n\ge k}$ is a subsequence of $\{\mu_n^k\}_n$, the sequences $\{\nu_n(f_k)\}_n$ are convergent. Choose $f \in \mathscr{C}(P)$ and $\varepsilon > 0$. Find $g \in \{f_k\}$ such that $\|f - g\| < \varepsilon$. There is an n_0 such that $|\nu_i(g) - \nu_j(g)| < \varepsilon$ for all $i, j \ge n_0$. Then by the triangle inequality

$$|\nu_i(f) - \nu_j(f)| \le (1 + 2\sup_n \|\mu_n\|)\varepsilon$$

for all $i, j \ge n_0$. Hence $\{\nu_n(f)\}$ is a Cauchy sequence and therefore it is convergent. Define a functional μ on $\mathscr{C}(P)$ as

$$\mu(f) = \lim_{n\to\infty} \nu_n(f).$$

It remains only to show that μ is a Radon integral on P and μ is a weak limit of $\{\nu_n\}$, which is easy. ■

17.11. Exercise. Let P be a compact metric space and $\mu_n \xrightarrow{\text{v}} \mu$. Show that

$$\|\mu\| \le \liminf \|\mu_n\|$$

(i.e. the norm is a weakly lower semicontinuous function on $\mathscr{M}(P)$).

17.12. Notes. The theory of weak convergence of probability measures has its roots in the probability theory and mathematical statistics (see e.g. P. Billingsley [*1968]) and led, of course, to a study of weak topologies on various subspaces of Radon measures. A similar notion of the vague convergence was probably studied first by Bourbakists (see the second edition of their monograph). Vague convergence is again nothing else than the weak convergence on the space of measures determined by the inductive topology on $\mathscr{C}_c(P)$. Theorem 17.10. is a special case of the Alaoglu-Bourbaki theorem.

18. Luzin's Theorem

If a measure space is endowed with a metric or a topological structure, a natural question arises whether there is a closer relation between measurability and continuity. The following theorems show that for complete Radon measures on locally compact topological spaces this is the case.

18.1. Luzin's Theorem. *For a complete Radon measure μ on a locally compact space P and a μ-almost everywhere finite function f on P, the following conditions are equivalent:*

(i) *f is μ-measurable;*

(ii) *for any $\varepsilon > 0$ and any compact set $K \subset P$ there is an open set G so that $\mu G < \varepsilon$ and $f|_{K \setminus G}$ is continuous;*

(iii) *for any $\varepsilon > 0$ and any compact set $K \subset P$ there exists a continuous function φ on P such that*

$$\mu\{x \in K\colon f(x) \neq \varphi(x)\} < \varepsilon;$$

(iv) *for every compact set $K \subset P$ there exists a sequence $\{\varphi_n\}$ of continuous functions on P such that*

$$\varphi_n \to f \quad \mu\text{-almost everywhere on } K.$$

Proof. (i) $\Longrightarrow$ (ii): Let $\{U_j\}$ be a countable base for the topology on $\mathbf{R}$ (for instance, a sequence of all intervals with rational endpoints). Fix an $\varepsilon > 0$ and a compact set $K \subset P$. There exist open sets $G_j \subset P$ and compact sets $F_j \subset P$ such that

$$F_j \subset K \cap f^{-1}(U_j) \subset G_j \quad \text{and} \quad \mu(G_j \setminus F_j) < 2^{-j}\varepsilon.$$

Set $G = \bigcup(G_j \setminus F_j)$. Clearly G is open and $\mu G < \varepsilon$. Denote $Y = K \setminus G$ and fix an index j. Then $Y \cap G_j = Y \cap F_j$. Thus $Y \cap f^{-1}(U_j) = Y \cap G_j$ is an open subset of Y and we see that $f|_Y$ is continuous.

(ii) $\Longrightarrow$ (iii): Select again $\varepsilon > 0$. Find an open set $G \subset P$ such that $\mu G < \varepsilon$ and $f|_{K \setminus G}$ is continuous. By Tietze's extension theorem, there exists a continuous function φ on P such that $f = \varphi$ on $K \setminus G$.

(iii) $\Longrightarrow$ (iv): Find continuous functions φ_k on P so that $\mu E_k < 2^{-k}$ where $E_k = \{x \in K\colon f(x) \neq \varphi_k(x)\}$. If

$$E := \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k \quad (= \limsup E_k),$$

then $\mu E = 0$ (cf. with the proof of the Borel-Cantelli lemma 2.14). If $x \in K \setminus E$, then there exists n_0 such that $x \notin E_n$ for $n \geq n_0$ and $\varphi_n(x) \to f(x)$.

The implication (iv) $\Longrightarrow$ (i) is obvious. ∎

For locally compact spaces which can be expressed as countable unions of compact sets we get the following corollary.

18.2. Luzin's Theorem. *Let μ be a complete Radon measure on a σ-compact locally compact space P and f a μ-almost everywhere finite function on P. Then the following conditions are equivalent:*

(i) *f is μ-measurable;*

(ii) *for any $\varepsilon > 0$ there exists an open set G so that $\mu G < \varepsilon$ and $f|_{P\setminus G}$ is continuous;*

(iii) *for any $\varepsilon > 0$ there exist a continuous function φ on P and an open set G such that $\mu G < \varepsilon$ and $\varphi = f$ on $P \setminus G$;*

(iv) *there exists a sequence $\{\varphi_n\}$ of continuous functions on P such that*

$$\varphi_n \to f \quad \mu\text{-almost everywhere on } P.$$

Proof. This theorem is an easy consequence of the previous one. ■

18.3. Remarks. 1. Only the equivalence (i) $\Longleftrightarrow$ (ii) is usually called *Luzin's theorem.*

2. Notice that a measurable function can be discontinuous at *all points*(!), consider e.g. the Dirichlet function on $\mathbf{R}$. Luzin's theorem says that the *restricted* function is continuous when omitting a "small" set. A characterization of functions which are continuous at almost all points of the interval $[0, 1]$ was described in 7.9 and 17.5.

3. In general, it is not true that a measurable function is continuous omitting a null set. Consider, for example, the indicator function of a discontinuum of a positive measure (see 1.13).

4. Another interesting characterization of measurable functions for the case of the Lebesgue measure in $\mathbf{R}^n$ is given by Denjoy's theorem 29.9.

5. By Theorem 18.2, every Lebesgue measurable function on $\mathbf{R}$ is a limit of a sequence of continuous functions in the sense of the convergence almost everywhere. The collection of functions which are pointwise limits of continuous functions (so called *functions of the Baire class one*) is not very wide. For example, the Dirichlet function (see 7.5) does not belong to it.

18.4. Exercise. Show that a μ-almost everywhere finite function f on P is μ-measurable if and only if for any compact set $K \subset P$ there exist compact sets $K_n \subset K$ and a μ-null set E such that $K = E \cup \bigcup K_n$ and functions $f|_{K_n}$ are continuous.

18.5. Exercise. Give another proof of the assertion that $\mathscr{C}_c(P)$ is dense in $\mathscr{L}^p$ (cf. Exercise 15.17) with the help of Luzin's theorem.

Hint. First approximate a function from $\mathscr{L}^p$ by a bounded measurable function with compact support and then use Luzin's theorem.

18.6. Notes. Luzin's theorem for the case of the Lebesgue measure was proved by H. Lebesgue [1903] and by N.N. Luzin [1912].

19. Measures on Topological Groups

19.1. Special Case. One of the fundamental properties of the Lebesgue measure λ on the real line is its "translation invariance": If $x \in \mathbf{R}$ and $A \subset \mathbf{R}$ is measurable, then $\lambda(x + A) = \lambda A$. The Lebesgue measure is in fact by this property uniquely determined. Indeed, the following theorem is true (cf. Exercise 26.6):

Let μ be a Radon measure on $\mathbf{R}$. If $\mu([0, 1]) = 1$ and $\mu(x + A) = \mu A$ for every $x \in \mathbf{R}$ and $A \in \mathscr{B}(\mathbf{R})$, then $\mu = \lambda$ on $\mathscr{B}(\mathbf{R})$.

Most of this chapter is devoted to a more general problem. Since the object belongs to elements of harmonic analysis and the reader can consult many textbooks (e.g. G. Bachman [*1964] or K. Ross [*1963]), we merely outline the main ideas and invite the reader to fill in the details.

19.2. Topological Group. Let us start with basic notions. By a *topological group* we understand a group G together with a (Hausdorff) topology such that the group operations $(x, y) \mapsto xy$ and $x \mapsto x^{-1}$ are continuous.

In the sequel, G stands for a topological group whose topology is locally compact. The unit element of G will be denoted by $\mathbf{e}$ and the σ-algebra of all Borel subsets of G by $\mathscr{B}(G)$. Finally, by $\mathscr{S}$ we will denote σ-algebras which appear as domains of measures in consideration.

For $x \in G$ and $A \subset G$, set

$$xA = \{xy\colon y \in A\}, \qquad Ax = \{yx\colon y \in A\}, \qquad A^{-1} = \{x^{-1}\colon x \in A\}.$$

19.3. Haar Measure. A *left Haar measure* on G is every nonzero Radon measure μ on $(G, \mathscr{S})$ which satisfies $\mu(xA) = \mu A$ for each $x \in G$ and $A \in \mathscr{S}$. In a similar way we define a *right Haar measure.* A measure which is simultaneously both a left and a right Haar measure is called briefly a *Haar measure.*

19.4. Examples. (a) The *Lebesgue measure* is a (typical) example of a Haar measure on $\mathbf{R}^n$ (the group operation is the addition).

(b) The *counting measure* is a Haar measure on every group equipped with the discrete topology.

(c) Let $G = (0, +\infty)$ be the multiplicative group of positive real numbers endowed with the Euclidean topology. If

$$\mu A := \int_A \frac{\mathrm{d}x}{x}, \qquad A \text{ Borel},$$

then μ is a Haar measure on G.

(d) Let $G = \mathbf{C} \setminus 0$ be the multiplicative group of all nonzero complex numbers (with the usual topology). If

$$\mu A = \int_A \frac{1}{|z|^2}\,\mathrm{d}\lambda(z)$$

for $A \in \mathscr{B}(G)$, then μ is a Haar measure on G.

19.5. Example. Let G be the multiplicative group of all 2×2-matrices of type

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$$

where $a \in (0, +\infty)$ and $b \in \mathbf{R}$. There exists a one-to-one mapping of G onto $(0, +\infty) \times \mathbf{R}$ given by

$$F\colon \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \to (a, b).$$

Consider on G the (locally compact) topology determined by F from $\mathbf{R}^2$. For $A \in \mathscr{B}(G)$, set

$$\mu A = \int_A \frac{1}{x^2}\,\mathrm{d}x\,\mathrm{d}y, \quad \nu A = \int_A \frac{1}{x}\,\mathrm{d}x\,\mathrm{d}y.$$

(a) Show that μ is a left Haar measure and ν a right Haar measure on G.

(b) Show that $\nu A = \mu(A^{-1})$.

(c) Find a set $A \in \mathscr{B}(G)$ such that $\mu A < \infty$ and $\nu A = \infty$.

(The group G can be viewed as the group of all affine transformations of $\mathbf{R}$ onto $\mathbf{R}$ of the form $t \mapsto at + b$, $a > 0$, $b \in \mathbf{R}$.)

19.6 Remark. If μ is a left Haar measure on G and $\tilde{\mu}$ is defined as $\tilde{\mu}E := \mu(E^{-1})$ (again $E^{-1} \in \mathscr{B}(G)$ if $E \in \mathscr{B}(G)$), then $\tilde{\mu}$ is a right Haar measure on G. In the same way, any right Haar measure determines a left one.

In what follows, we restrict ourselves to a study of left Haar measures. Fundamental properties are contained in the following theorem.

19.7. Theorem (existence and uniqueness of left Haar measure). *(a) On any locally compact group there exists a left Haar measure.*

(b) If μ and ν are complete left Haar measures on G, then there exists $c > 0$ such that $\mu = c\nu$.

Proof of this theorem is quite complicated and takes a lot of effort. There are various existence proofs, some of them as applications of deep theorems of functional analysis. Here we outline a rough idea of an "elementary" existence proof.

Let V be an open neighbourhood of the unit element $\mathbf{e}$ of G. If $E \subset G$ is a compact set, denote by $H_V(E)$ the smallest n such that there exist $x_1, \dots, x_n \in G$ so that $E \subset \bigcup_{j=1}^{n} x_j V$ (existence of a finite cover follows from the compactness of E). Let K be a fixed compact set with a nonempty interior. The required Haar measure is then obtained by some limit procedure for "$V \to \{\mathbf{e}\}$" and by extending set functions

$$E \mapsto \frac{H_V(E)}{H_V(K)}, \qquad E \text{ compact.}$$

A proper and complete description of the entire construction is not easy.

The proof of uniqueness will be given under an additional assumption that μ is even a Haar measure. The general case is similar, but the technical details are more difficult. First, there exists a nonnegative function $h \in \mathscr{C}_c(G)$ so that $h(\mathbf{e}) \neq 0$ and $h(x) = h(x^{-1})$ for $x \in G$ (if $g \in \mathscr{C}_c(G)$, $g(\mathbf{e}) \neq 0$, $g \geq 0$, set $h(x) = g(x) + g(x^{-1})$). Then $\int_G h \, \mathrm{d}\mu > 0$ (see Exercise 19.14.b) and for an arbitrary function $f \in \mathscr{C}_c(G)$ we get

$$\begin{aligned}
\int_G h \, \mathrm{d}\nu \int_G f \, \mathrm{d}\mu &= \int_G h(y) \left(\int_G f(xy) \, \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y) \\
&= \int_{G \times G} h(y) f(xy) \, \mathrm{d}\mu{\otimes}\nu(x, y) \\
&= \int_{G \times G} h(x^{-1}z) f(z) \, \mathrm{d}\mu{\otimes}\nu(x, z) \\
&= \int_G \left(\int_G h(x^{-1}z) f(z) \, \mathrm{d}\nu(z) \right) \mathrm{d}\mu(x) \\
&= \int_G f(z) \left(\int_G h(z^{-1}x) \, \mathrm{d}\mu(x) \right) \mathrm{d}\nu(z) = \int_G h \, \mathrm{d}\mu \int_G f \, \mathrm{d}\nu.
\end{aligned}$$

To give reasons for particular steps when using Fubini's theorem notice that f and h have compact supports and that μ and ν are Radon measures. We can finish the proof by setting $c := \frac{\int_G h \, d\mu}{\int_G h \, d\nu}$. ■

Haar measures have many interesting properties. Let us state some of them.

19.8. Theorem. *Let μ be a left Haar measure on a locally compact group G. Then:*

(a) $\mu U > 0$ *for any nonempty open set* $U \subset G$*;*
(b) $\mu G < +\infty$ *if and only if* G *is compact.*

Proof. (a) Let $U \subset G$ be a nonempty open set. We can assume $\mathbf{e} \in U$. There exists a compact set $K \subset G$ with $\mu K > 0$. Then $K \subset \bigcup_{x \in K} xU$ and the compactness of K yields the existence of $x_1, \dots, x_n \in K$ such that $K \subset \bigcup_{i=1}^{n} x_i U$. Then

$$0 \le \mu K \le \sum_{i=1}^{n} \mu(x_i U) = n\mu(U).$$

(b) Suppose $\mu G < +\infty$. Let $K \subset G$ be a compact set of a positive measure. There exist $x_1, \dots, x_n \in G$ so that the sets $x_i K$, $i = 1, \dots, n$ are pairwise disjoint but all ("remaining") $x \in G$ satisfy $xK \cap (\bigcup_{i=1}^{n} x_i K) \neq \emptyset$. (Indeed – consider finite sequences $x_1, \dots, x_n$ for which $x_1 K, \dots, x_n K$ are pairwise disjoint. Their number is bounded for instance by $\mu G / \mu K$. If $x \in G$, then $xK \cap (\bigcup_{i=1}^{n} x_i K) \neq \emptyset$ and it follows that $G = \left(\bigcup_{i=1}^{n} x_i K \right) K^{-1}$ is compact. ■

19.9. Modular Function. Let μ be a left Haar measure on a locally compact group G. If $x \in G$ and $\mu_x A := \mu(Ax)$ for $A \in \mathscr{B}(G)$, then μ_x is obviously a left Haar measure. By the uniqueness part of Theorem 19.7 there exists $\Delta(x) > 0$ such that $\mu_x = \Delta(x)\mu$. Again, by uniqueness, $\Delta(x)$ does not depend on the choice of a left Haar measure on G. The function $\Delta \colon x \mapsto \Delta(x) : G \to (0, +\infty)$ is called the *modular function* of G.

19.10. Theorem. *Let Δ be the modular function of G. The following conditions are equivalent:*

(i) *every left Haar measure on G is also a right Haar measure;*
(ii) $\Delta = 1$ *on* G.

Proof is easy and is omitted. ■

19.11. Unimodular Group. A locally compact group G is said to be *unimodular* if the modular function $\Delta = 1$ on G. In other words, G is unimodular if the classes of left and right Haar measures on G coincide. Every commutative group is unimodular. However, there exist examples of noncommutative unimodular groups as indicated in the next theorem.

19.12. Theorem. *Any commutative, discrete or compact group is unimodular.*

Proof. If G is discrete, then every left Haar measure is a multiple of the counting measure. Hence, it is also "right invariant". If μ is a left Haar measure on a compact group G and $x \in G$, then $G = Gx$ and $\mu G = \mu(Gx) = \Delta(x)\mu G$. Since $0 < \mu G < \infty$, we obtain $\Delta(x) = 1$. ■

If G is a compact group, then each left Haar measure on G is a Haar measure and we get immediately the following theorem.

19.13. Theorem. *On every compact topological group G there exists a unique complete Haar measure μ satisfying $\mu G = 1$. Moreover, $\mu E = \mu E^{-1}$ for every $E \in \mathscr{B}(G)$.*

19.14. Exercise. Let μ be a left Haar measure. Prove the following assertions.

(a) If $f \in \mathscr{C}_c(G)$ and $y \in G$, then $\int_G f \, d\mu = \int_G f(yx) \, d\mu(x)$. The last equality holds if f is a nonnegative μ-measurable function on G or if $f \in \mathscr{L}^1(\mu)$.

(b) If $f \in \mathscr{C}_c(G)$ is nonnegative and $f(\mathbf{e}) > 0$, then $\int_G f \, d\mu > 0$.

(c) The measure μ is σ-finite if and only if G is σ-compact.

(d) The topology on G is discrete if and only if $\mu\{x\} \neq 0$ for some (and thus for all) $x \in G$.

19.15. Exercise. Calculate the modular function of the group from Example 19.5.

19.16. Exercise . Let Δ be a modular function on a locally compact group G. Show that:

(a) Δ is continuous;
(b) $\Delta(xy) = \Delta(x)\Delta(y)$ for all $x, y \in G$;
(c) $\Delta(\mathbf{e}) = 1$.

19.17. Exercise. Let μ be a left Haar measure on G. Show that the set function $\tilde{\mu}$, where $\tilde{\mu}E := \mu E^{-1}$, is a right Haar measure and $\int_G f(x)\Delta(x^{-1}) \, d\mu = \int_G f \, d\tilde{\mu}$ for every $f \in \mathscr{C}_c(G)$. In other words, the Radon-Nikodým derivative $\dfrac{d\tilde{\mu}}{d\mu}$ equals $\Delta(x^{-1})$.

Hint. Show that the function ν on $\mathscr{B}(G)$ defined by $\nu A := \int_A \Delta(x^{-1}) \, d\mu$ is a right Haar measure. Thus $\nu = K\tilde{\mu}$ for some K. Now use the fact that Δ is a continuous function attaining the value 1 at the unit element $\mathbf{e}$.

19.18. Exercise. Show that the left and right Haar measures on G are mutually absolutely continuous.

Hint. Use Exercise 19.17.

19.19. Convolution of Functions. In the following exercises, let χ be a fixed right Haar measure on $(G, \mathscr{B}(G))$. If $f, g \in \mathscr{L}^1(\chi)$ and $x \in G$, define the *convolution* of f and g at x as

$$f * g\,(x) = \int_G f(xy^{-1})g(y) \, d\chi(y)$$

provided the integral on the right-hand side exists.

19.20. Exercise. If $f, g \in \mathscr{L}^1(\chi)$, then the convolution $f * g$ is defined χ-almost everywhere $f * g \in \mathscr{L}^1(\chi)$ and $\|f * g\|_1 \le \|f\|_1 \, \|g\|_1$. Thus, $L^1(\chi)$ with convolution as multiplication is a Banach algebra.

19.21. Exercise. Show that convolution is commutative if and only if G is commutative.

19.22. Exercise. Show that the Banach algebra $(L^1(\chi), *)$ has a multiplicative unit if and only if G is endowed with the discrete topology. As an example consider the group $\mathbb{Z}$ of all integers with the discrete topology and counting measure.

19.23. Involution. If $f^*(x) := \overline{f(x^{-1})}(\Delta(x))^{-1}$, then the mapping $f \mapsto f^*$ of the space $L^1(\chi)$ onto itself (which is called the *involution*) is an isometry. (Show that $\int_G |f(x^{-1})| (\Delta(x))^{-1} \, \mathrm{d}\chi = \int_G |f(x)| \, \mathrm{d}\chi$ for $f \in \mathscr{C}_c(G)$; hence it follows that $\|f\| = \|f^*\|$ if $f \in \mathscr{L}^1(\chi)$.)

19.24. Convolution of Measures. In the following, G is again a locally compact topological group. By $\mathscr{M}_b(G)$ denote the set of all complex Radon measures on $(G, \mathscr{B}(G))$. According to 17.1, $\mathscr{M}_b(G)$ with the norm defined by $\|\mu\| := |\mu|(G)$ is a Banach space.

We define the *convolution of measures* $\mu, \nu \in \mathscr{M}_b(G)$ as

$$\mu{*}\nu\,(E) := \tau\,\{(x,y) \in G \times G \colon xy \in E\}$$

for any Borel set $E \subset G$, where τ is the product of complex Radon measures μ and ν, cf. 16.9.c. (Verify that $\{(x,y) \in G \times G \colon xy \in E\}$ is a Borel subset of $G \times G$!)

19.25. Exercise. Prove that $\mu * \nu$ is a complex measure on $\mathscr{B}(G)$, $\mu * \nu \in \mathscr{M}_b(G)$ and that the inequality $\|\mu * \nu\| \leq \|\mu\| \cdot \|\nu\|$ holds.

19.26. Exercise. Show that

(a) the operation of convolution of measures is associative;

(b) the convolution of measures on G is commutative if and only if G is commutative.

19.27. Exercise. Prove that if $\mathbf{e}$ is the unit of G, then the Dirac measure $\varepsilon_{\mathbf{e}}$ is the unit of $(\mathscr{M}_b(G), *)$.

19.28. Remark. The reader may find it interested to investigate a relation between convolution of measures and convolution of functions. If $\mu, \nu \in \mathscr{M}_b(G)$, then

$$\mu * \nu\,(E) = \int_{G \times G} c_E(xy) \, \mathrm{d}\mu{\otimes}\nu\,,$$

hence

$$\int_G h \, \mathrm{d}\mu{*}\nu = \int_{G \times G} h(xy) \, \mathrm{d}\mu{\otimes}\nu.$$

for any bounded Borel function h on G. Therefore, for $f, g \in \mathscr{L}^1(\chi)$ we have

$$f * g = \frac{\mathrm{d}(\chi_f * \chi_g)}{\mathrm{d}\chi},$$

where χ_f is the complex Radon measure given by $\chi_f(E) = \int_E f \, \mathrm{d}\chi$. We see that the “new” definition of convolution of functions agrees with the “former” one.

19.29. Exercise. Give a definition of a convolution of a (complex) function $f \in \mathscr{L}^1(\chi)$ and a complex Radon measure $\mu \in \mathscr{M}_b(G)$. Prove that $f * \mu \in \mathscr{L}^1(\chi)$ and that $\|f * \mu\|_1 \leq \|f\|_1 \|\mu\|$. In fact, $L^1(\chi)$ is not only a subalgebra of $\mathscr{M}_b(G)$ but even an ideal.

19.30. Notes. The translation invariant measures (or integrals) on compact Lie groups were studied by F. Peter and H. Weyl [1927]. Remarkable progress was made by proving the existence of a left Haar measure for separable locally compact groups by A. Haar [1933] and by J. von Neumann in [1934] (the existence and uniqueness) for arbitrary compact groups, and in [1936] (uniqueness) for separable locally compact groups and also by A. Weil especially in [*1940]. Their proofs used the axiom of choice. H. Cartan [1940] and G.E. Bredon [1963] then gave proofs without using the axiom of choice. A relatively short proof of the uniqueness result can be found in S. Kakutani [1948]. More detailed historical notes are given by E. Hewitt and K.A. Ross in [*1963]. It is said that Example 19.5. is due to J. von Neumann [1936]. Note that Haar measures are special cases of invariant measures which can be studied in a more general setting, cf. Banach's appendix of Saks' monograph [*1937] or H. Federer [*1969].

D. Integration on R

20. Integral and Differentiation

In the following chapters we will investigate properties of real functions involving the Lebesgue measure and Lebesgue integration on the real line.

In this chapter, our aim is to show an inequality between the measure of $f(E)$ and the integral of f' over E. As a particular case, we obtain Sard's lemma on the real line.

Let K be a positive real number. We say that a real function f is a *K-Lipschitz* function on a set $E \subset \mathbf{R}$ if the inequality

$$|f(x) - f(y)| \leq K\,|x - y|$$

holds for all $x, y \in E$. If f is a K-Lipschitz function on E for some K, we say simply that f is a Lipschitz function on E.

20.1. Lemma. *Let f be a K-Lipschitz function on a set $E \subset \mathbf{R}$. Then $\lambda^* f(E) \leq K\lambda^* E$.*

Proof. The assertion is obvious in the case $\lambda^* E = +\infty$. If $\lambda^* E < +\infty$, select $\varepsilon > 0$ and find a sequence of open intervals (a_j, b_j) with $E \subset \bigcup_j (a_j, b_j)$ and $\sum_j (b_j - a_j) \leq \lambda^* E + \varepsilon$. By hypothesis there are intervals $[\alpha_j, \beta_j]$ such that $f(E \cap (a_j, b_j)) \subset [\alpha_j, \beta_j]$ and $\beta_j - \alpha_j \leq K(b_j - a_j)$. Thus $f(E) \subset \bigcup_j [\alpha_j, \beta_j]$ and

$$\lambda^* f(E) \leq \sum_j (\beta_j - \alpha_j) \leq K \sum_j (b_j - a_j) \leq K(\lambda^* E + \varepsilon).$$

Since the last inequality holds for all $\varepsilon > 0$, we are done. ■

20.2. Lemma. *Let f be a real function on an interval I and $E \subset I$. If $|f'| \leq K$ on E for some $K > 0$, then*

$$\lambda^* f(E) \leq K\lambda^* E.$$

Proof. Choose $K' > K$ and denote

$$E_k = \{x \in E\colon\ |f(x) - f(y)| \leq K' \text{ for all } y \in E \cap (x - \tfrac{1}{k}, x + \tfrac{1}{k})\}.$$

Then $E_1 \subset E_2 \subset \dots$ and $E = \bigcup_k E_k$. Let J be an interval with length less than $\frac{1}{k}$. According to the previous lemma,

$$\lambda^* f(J \cap E_k) \leq K'\,\lambda^*(J \cap E_k).$$

Dividing I into such intervals we obtain $\lambda^* f(E_k) \leq K'\,\lambda^*(E_k)$. An appeal to Exercise 4.7 reveals that $\lambda^* f(E) \leq K'\,\lambda^*(E)$. Since $K' > K$ was arbitrary, we have the required inequality. ■

As a consequence of the preceeding lemma we get one-dimensional version of well-known Sard's lemma. Compare it with the next Theorem 20.4.

20.3. Corollary. *Let f be a real function on an interval $I \subset \mathbf{R}$ and $E \subset I$. If $f' = 0$ on E, then $\lambda f(E) = 0$.*

20.4. Theorem. *Let f be a real-valued function on an interval I and $E \subset I$ a measurable set. Suppose that a finite derivative $f'(x)$ exists at each point $x \in E$. Then f' is measurable on E and*

$$\lambda^* f(E) \leq \int_E |f'| \, .$$

Proof. We first prove the measurability of f' on E. Choose $c \in \mathbf{R}$. For each $k, m \in \mathbf{N}$, the set

$$G_{m,k} := \{x \in I : \text{ there exist } y, z \in I \text{ such that} \\ x - \frac{1}{k} < y < x < z < x + \frac{1}{k} \text{ and } f(z) - f(y) > (c + \frac{1}{m})(z - y)\}$$

is open. Hence the set

$$\{x \in E \colon f'(x) > c\} = E \cap \left(\bigcup_m \bigcap_k G_{m,k} \right)$$

is measurable.

To prove the required inequality, it is no restriction to assume that E is bounded. For an $\varepsilon > 0$ denote

$$E_k = \{x \in E \colon (k-1)\varepsilon \leq |f'(x)| < k\varepsilon\}.$$

Then E_k are pairwise disjoint measurable sets, $E = \bigcup_k E_k$ and an appeal to Lemma 20.2 yields

$$\lambda^* f(E) \leq \sum_k \lambda^* f(E_k) \leq \sum_k k\varepsilon \leq \sum_k \left(\int_{E_k} |f'| + \varepsilon \lambda E_k \right) \\ = \int_E |f'| + \varepsilon \lambda E.$$

■

20.5. Exercise. Under the assumptions of Theorem 20.4 prove that $f(E)$ is a measurable set.
Hint. Use the following Exercise 20.6 and the fact that $E = N \cup \bigcup_n K_n$ where $\lambda N = 0$ and K_n are compact (Theorem 1.21).

20.6. Exercise. Let f be a real-valued function on an interval $I \subset \mathbf{R}$ having a finite derivative at each point of a (Lebesgue) null set $E \subset I$. Show that $\lambda f(E) = 0$.
Hint. Use either Theorem 20.4 or Lemma 20.2 realizing that $E = \bigcup_n \{x \in E : |f'(x)| < n\}$.

20.7. Luzin's (N)-property. We have seen in 3.18 that a continuous image of a (Lebesgue) measurable set does not need to be measurable. On the other hand, if f is differentiable and E measurable, Exercise 20.5 yields that $f(E)$ is measurable. We now consider briefly the question, under what conditions images of measurable sets are again measurable.

We say that a real function f defined on an interval $I \subset \mathbf{R}$ has *Luzin's (N)-property* if the image $f(N)$ of any (Lebesgue) null set $N \subset I$ is again a null set.

(a) (Rademacher) Let f be a continuous function on an interval I. Then the following conditions are equivalent:

(i) f has Luzin's (N)-property;
(ii) $f(M)$ is measurable whenever $M \subset I$ is measurable.

Hint. In light of Theorem 1.21 a measurable set is a countable union of compact sets and a null set. For the converse, use Remark 1.9.2 which says that every set of a positive measure contains a nonmeasurable subset.

(b) Show that the assertion of (a) remains true if f is measurable only (use Luzin's Theorem 18.2). Likewise, the definition of Luzin's (N)-property and other ideas can be generalized to the case when I is a measurable subset of $\mathbf{R}^n$.

(c) Let f be a function having a finite derivative everywhere on an interval I. Then f has Luzin's (N)-property.

Hint. Exercise 20.6.

(d) Every Lipschitz (even locally Lipschitz) function on a measurable set M has Luzin's (N)-property.

20.8. Exercise. Let f be an arbitrary function on an interval $I \subset \mathbf{R}$. If D is the set of points at which f has a finite derivative, then D is a Borel set and the function $x \mapsto f'(x)$ is a Borel function on D.

20.9. Notes. Luzin's (N)-property was introduced by Luzin in [*1915]. Corollary 20.3 which we called the one-dimensional version of Sard's lemma is also often called Luzin's theorem.

21. Functions of Finite Variation and Absolutely Continuous Functions

In this chapter we discuss two classes of functions without using measure theory. Important results concerning these functions will be introduced in following chapters.

21.1. Functions of Finite Variation. Let f be a real-valued function on an interval I. For any interval $[a,b] \subset I$ and any partition $D\colon a = x_0 < x_1 < \cdots < x_m = b$ of $[a,b]$ denote

$$\mathop{\mathrm{V}}_a^b(f,D) = \sum_{j=1}^{m} |f(x_j) - f(x_{j-1})|\,.$$

The extended real number

$$\mathop{\mathrm{V}}_a^b f := \sup\{\mathop{\mathrm{V}}_a^b(f,D)\colon D \text{ is a partition of} [a,b]\}$$

is called the *variation* of f over $[a,b]$. If $\mathop{\mathrm{V}}_a^b f < \infty$ for every interval $[a,b] \subset I$, then f is termed a *function of finite variation*. In this case there exists a function v on I such that

$$v(b) - v(a) = \mathop{\mathrm{V}}_a^b f$$

for any interval $[a, b] \subset I$. Such a function v is determined up to an additive constant and it is called an *(indefinite) variation* of f.

One can easily see that the set of all functions of finite variation on an interval I is a vector space.

If, in addition,

$$\sup\{\overset{b}{\underset{a}{\mathrm{V}}} f \colon [a, b] \subset I\} < +\infty,$$

we say that f is of *bounded variation* on I. For a compact interval $[a, b]$, notions of finite and bounded variation agree and they can be characterized simple by $\overset{b}{\underset{a}{\mathrm{V}}} f < \infty$.

In case of an arbitrary interval I we could say that f is of "locally bounded variation" instead of "finite variation". Indeed, a function f is of finite variation on I if and only if it has bounded variation on each compact subinterval of I. In a similar way, we are going to "localize" notions of absolute continuity and integrability.

21.2 Jordan Decomposition Theorem. *A function f is of finite variation on I if and only if f is a difference of two nondecreasing functions.*

Proof. Monotone functions are of finite variation, and thus differences of monotone functions are also of finite variation. For the converse, if v is an indefinite variation of a function f of finite variation, then v and $v - f$ are nondecreasing and f is their difference. ∎

21.3. Absolutely Continuous Functions. A real-valued function f is said to be *absolutely continuous* on an interval I if given $\varepsilon > 0$, there exists $\delta > 0$ that

$$\sum_{j=1}^{m} |f(b_j) - f(a_j)| < \varepsilon$$

whenever $a_1 < b_1 \le a_2 < b_2 \le \cdots \le a_m < b_m$ are points of I with $\sum_{j=1}^{m} (b_j - a_j) < \delta$. The family of all absolutely continuous functions on an interval I is a vector space which contains all Lipschitz functions.

We say that a function f is *locally absolutely continuous* on an interval I if f is absolutely continuous on every compact subinterval of I.

Any locally absolutely continuous function is continuous and of bounded variation.

21.4. Theorem. *Any absolutely continuous function f on I is the difference of two nondecreasing absolutely continuous functions.*

Proof. It is enough to show that the indefinite variation v of f is absolutely continuous. To this end let an $\varepsilon > 0$ be given. Find $\delta > 0$ such that

$$\sum_{j=1}^{m} |f(b_j) - f(a_j)| < \varepsilon$$

whenever $a_1<b_1 \le a_2<b_2 \le \cdots \le a_m<b_m$ are points of I with $\sum_{j=1}^{m}(b_j - a_j) < \delta$.

Let $A_1<B_2 \le A_2<B_2 \le \cdots \le A_p<B_p$ be points of I with $\sum_{j=1}^{p}(B_j - A_j) < \delta$. Find partitions

$$A_j = a_j^0 < b_j^0 = a_j^1 < \cdots < b_j^{m_j} = B_j$$

of intervals $[A_j, B_j]$ for which

$$v(B_j) - v(A_j) < \sum_{i=1}^{m_j} \left|f(b_j^i) - f(a_j^i)\right| + \frac{1}{p}\varepsilon.$$

Since $\sum_{j,i}(b_j^i - a_j^i) < \delta$, we have

$$\sum_j |v(B_j) - v(A_j)| < \sum_{j,i} \left|f(b_j^i) - f(a_j^i)\right| + \varepsilon < 2\varepsilon\,,$$

as needed. ■

21.5 Exercise. Show that the product of two absolutely continuous functions (or functions of finite variation) on a bounded interval is again an absolutely continuous function (or a function of finite variation).

21.6. Exercise. Prove that every absolutely continuous function has Luzin's (N)-property.

Hint. If N is a null set, find an open set M of a small measure containing N and then use the definition of absolute continuity.

21.7. Notes. C. Jordan discovered functions of finite variation in [1881] and proved the decomposition theorem 21.2.

22. Theorems on Almost Everywhere Differentiation

In this chapter we show that every function of finite variation (in particular, every nondecreasing or Lipschitz function) has a finite derivative almost everywhere. The usual proof of this deep theorem uses Vitali's covering theorem. Here we present a rather elementary proof. Let start with a simple lemma which is due to F. Riesz.

22.1. F. Riesz's Rising Sun Lemma. *Let h be a continuous function on an interval $[a, b]$. Denote*

$$E = \{x \in (a, b) : \text{ there exists } \xi \in (x, b) \text{such that } h(\xi) > h(x)\}.$$

Then E is a union of a sequence of pairwise disjoint open intervals (a_j, b_j) with $h(a_j) \le h(b_j)$.

Proof. It is clear that E is an open set. Hence, E is a union of a sequence of pairwise disjoint maximal open intervals contained in E. Let (α, β) be any such an interval and $x \in (\alpha, \beta)$. Denote

$$M = \{\xi \in (x, \beta) \colon h(\xi) \ge h(x)\}.$$

Since $\beta \notin E$, it follows that $h \le h(\beta)$ in $[\beta, b)$. By hypothesis $M \neq \emptyset$ and $\sup M = \beta$. Thus $h(x) \le h(\beta)$ and letting $x \to \alpha+$, we complete the proof. ■

22.2. Remarks. 1. If (α, β) is a maximal open interval contained in E and $\alpha > a$, then even $h(\alpha) = h(\beta)$.

2. The "mirror" version of the lemma: Let h be a continuous function on an interval $[a, b]$ and denote

$$E = \{x \in (a, b) : \text{ there exists a } \xi \in (a, x) \text{ with } h(\xi) > h(x)\}.$$

Then E is the union of a sequence (a_j, b_j) of pairwise disjoint open intervals with $h(a_j) \ge h(b_j)$.

22.3. Extreme Derivatives. Let f be a function defined on a neighborhood of a point $x \in \mathbf{R}$. Denote

$$D^+ f(x) = \limsup_{t \to 0+} \frac{1}{t}(f(x+t) - f(x)),$$

$$D_+ f(x) = \liminf_{t \to 0+} \frac{1}{t}(f(x+t) - f(x))$$

and analogously $D^- f(x)$ and $D_- f(x)$ for $t \to 0-$. These extended real numbers are called the *Dini derivatives* of f at x. A function f is differentiable at x if all Dini derivatives of f agree at x.

Finally set

$$\overline{D} f(x) = \limsup_{t \to 0} \frac{1}{t}(f(x+t) - f(x))$$

and call the function $\overline{D} f$ the *upper derivative* of f. The lower derivative $\underline{D} f$ is defined in a similar way.

22.4. Lemma. *Every nondecreasing Lipschitz function f on an interval $[a, b]$ has a finite derivative almost everywhere on $[a, b]$.*

Proof. It is no restriction to assume that f is 1-Lipschitz. Lipschitz functions cannot have infinite derivatives. To prove the assertion it suffices to show that $D^+ f \le D_- f$ almost everywhere. Then analogously $D^- f \le D_+ f$ almost everywhere, and

$$0 \le D^+ f \le D_- f \le D^- f \le D_+ f \le D^+ f \le 1$$

almost everywhere as needed. To this end let $0 < p < q < 1$ be given and set

$$M_{p,q} := \{x \in (a, b) \colon D_- f(x) < p < q < D^+ f(x)\}.$$

By Remark 22.2.2 there is a sequence of pairwise disjoint intervals (a_j, b_j) such that

$$\big\{x \in (a, b) \colon D_- f(x) < p\big\} \subset \big\{x \in (a, b) : \text{ there exists } \xi \in (a, x) \text{ with } f(\xi) - p\xi > f(x) - px\big\} = \bigcup_j (a_j, b_j)$$

and

$$f(b_j) - pb_j \le f(a_j) - pa_j.$$

Now apply Lemma 22.1 to $f(x) - qx$ on each interval $[a_k, b_k]$. Again there are pairwise disjoint intervals $(a_{k,j}, b_{k,j}) \subset (a_k, b_k)$ such that

$$\{x \in (a,b) \colon D^+ f(x) > q\} \subset \bigcup_k \{x \in (a_k, b_k) : \text{ there exists } \xi \in (a_k, x) \text{ with } f(\xi) - q\xi < f(x) - qx\} = \bigcup_{k,j} (a_{k,j}, b_{k,j})$$

and

$$f(b_{k,j}) - qb_{k,j} \geq f(a_{k,j}) - qa_{k,j}.$$

Hence

$$\sum_{k,j}(b_{k,j} - a_{k,j}) \leq \frac{1}{q} \sum_{k,j} (f(b_{k,j}) - f(a_{k,j})) \leq \frac{1}{q} \sum_k (f(b_k) - f(a_k)) \leq \frac{p}{q} \sum_k (b_k - a_k) \leq \frac{p}{q}(b-a).$$

Thus, we can define by induction sequences of decreasing collections of intervals; in the $2n^{\text{th}}$ step we obtain a collection of intervals $\{(A_s, B_s)\}$ with

$$\bigcup_s (A_s, B_s) \supset M_{p,q} \quad \text{and} \quad \sum_s (B_s - A_s) \leq \left(\frac{p}{q}\right)^n (b-a).$$

It follows that $M_{p,q}$ is a null set. The assertion now follows from the fact that

$$\{x \in (a,b) \colon D_- f(x) < D^+ f(x)\} \subset \bigcup_{p,q \in (0,1) \cap \mathbf{Q}} M_{p,q}.$$

■

Now we show that also monotone functions are differentiable almost everywhere. The situation is more difficult since monotone functions can be discontinuous. The heart of the proof lies in the observation that, for a nondecreasing function f, the inverse function to $x \mapsto x + f(x)$ (whose domain is not necessarily connected) can be extended to a nondecreasing Lipschitz function on an interval.

22.5. Theorem (Lebesgue). *Every monotone function f on an interval I has a finite derivative at almost all points of I.*

Proof. It would be clearly sufficient to assume that $I = [a, b]$. There exists an interval $[A, B]$ and a function g on $[A, B]$ so that g is Lipschitz and nondecreasing, and $x + f(x) \in [A, B]$, $g(x + f(x)) = x$ for all $x \in [a, b]$. (If f is continuous, then g is simply the inverse function to $x + f(x)$). A moment's thought will convince the reader that

$$\{x \in (a,b) \colon f'(x) \text{ does not exist } \} \subset g(E) \cup g(N)$$

where

$$E = \{y \in (A,B) \colon g'(x) \text{ does not exist } \} \quad \text{and} \quad N = \{y \in (A,B) \colon g'(x) = 0\}.$$

By the previous lemma $\lambda E = 0$. Hence $\lambda g(E) = 0$ by Lemma 20.1. According to Sard's lemma (Corollary 20.3) also $\lambda g(N) = 0$ and the proof is complete. ■

22.6. Corollary. *Every function of bounded variation has a finite derivative almost everywhere.*

We conclude this section by proving an important inequality.

22.7. Theorem. *Let f be a nondecreasing function on an interval $[a,b]$. Then $f' \in \mathscr{L}^1([a,b])$ and*

$$\int_a^b f' \le f(b) - f(a).$$

Proof. For $x > b$ set $f(x) = f(b)$ and define a sequence of functions

$$f_k(x) = k\left(f(x + \frac{1}{k}) - f(x)\right)$$

for $x \in [a,b]$. Then f_k are nonnegative measurable functions on $[a,b]$ (f is measurable!) and $\lim f_k = f'$ almost everywhere. Thus f' is measurable and nonnegative almost everywhere. Fatou's lemma 8.15 and a quick computation establish that

$$\begin{aligned}\int_a^b f' &\le \liminf \int_a^b f_k = \liminf\, k\left(\int_{a+\frac{1}{k}}^{b+\frac{1}{k}} f - \int_a^b f\right)\\ &= \liminf\, k\left(\int_b^{b+\frac{1}{k}} f - \int_a^{a+\frac{1}{k}} f\right) \le \liminf\, k\left(\frac{1}{k}f(b) - \frac{1}{k}f(a)\right)\\ &= f(b) - f(a).\end{aligned}$$

■

22.8. Notes. Theorem 22.5 on differentiability of monotone functions was proved by H. Lebesgue in [*1904] under an additional assumption of continuity of differentiated function. In a full generality, the theorem was proved independently by G. Faber [1918] and by G.C. Young and W.H. Young [1911]. Rising sun lemma 22.1 and proof of 22.4 are due to F. Riesz [1930-32]. To prove Lebesgue's theorem, he uses this lemma and the notion of null sets only. The idea of the proof of 22.5 was used by J. Mignot [1976] and L. Zajíček [1983].

23. Indefinite Lebesgue Integral and Absolute Continuity

In this chapter we examine the formula

$$f(b) - f(a) = \int_a^b f'$$

where the integral is understood as Lebesgue's one and the derivative in the "almost everywhere" sense. First, let us show that formula does not always hold even if f is monotone by considering the following example.

23.1. Cantor Singular Function. We define the Cantor function $f\colon [0,1] \to [0,1]$ in the following way: Let $f(0) = 0$, $f(1) = 1$. For $x \in [\frac{1}{3}, \frac{2}{3}]$ set $f(x) = \frac{1}{2}$. Further set $f(x) = \frac{1}{4}$ for $x \in [\frac{1}{9}, \frac{2}{9}]$ and $f(x) = \frac{3}{4}$ for $x \in [\frac{7}{9}, \frac{8}{9}]$. Each succesive step is essentially the same. If (a,b) is a maximal interval in which f is not yet defined, we subdivide it into thirds and define the value of f in the closed middle third as the arithmetical mean of $f(a)$ and $f(b)$. Then f is defined and uniformly continuous on a dense subset of $[0,1]$. The last step of the definition consists in the continuous extension of f to the whole interval $[0,1]$. There is also an arithmetic definition of the Cantor function. Suppose $x \in [0,1]$ is written in the form

$$x = \sum_{j=1}^{\infty} 3^{-j} x_j$$

where $x_j \in \{0,1,2\}$. Then

$$f(x) = \frac{1}{2} \sum_{j=1}^{\infty} 2^{-j} y_j$$

where

$$y_j = \begin{cases} 1 & \text{if there is } i < j \text{ with } x_i = 1, \\ x_j & \text{otherwise.} \end{cases}$$

We see that $f' = 0$ almost everywhere and

$$f(1) - f(0) = 1 \neq 0 = \int_0^1 f' .$$

This situation cannot occur when f is absolutely continuous as shown by the following theorem.

23.2 Theorem. *Let f be an absolutely continuous function on an interval $[a,b]$. Then $f' \in \mathscr{L}^1([a,b])$ and*

$$f(b) - f(a) = \int_a^b f'.$$

Proof. As every absolutely continuous function is the difference of two monotone absolutely continuous functions, no generality is lost with the assumption that f is nondecreasing. Obviously $f' \geq 0$ almost everywhere and $\int_a^b f' \leq f(b) - f(a)$ thanks to Theorem 22.7. Select $\varepsilon > 0$ and let $\delta > 0$ be as furnished by the definition of absolute continuity of f. There exists an open set G with $\lambda G < \delta$ containing all points of nondifferentiability of f. Let G be expressed as a union of pairwise disjoint intervals (a_j, b_j). Then the definition of absolute continuity of f yields $\sum_{j=1}^k (f(b_j) - f(a_j)) < \varepsilon$ for all $k \in \mathbf{N}$. Thus $\lambda f(G) \leq \varepsilon$. On the other hand, by Theorem 20.4 we have

$$\lambda^* f([a,b] \setminus G) \leq \int_{[a,b]\setminus G} f' \leq \int_a^b f'.$$

Whence

$$f(b) - f(a) = \lambda f([a,b]) \leq \lambda^* f([a,b] \setminus G) + \lambda f(G) \leq \int_a^b f' + \varepsilon.$$

■

23.3 Indefinite Lebesgue Integral. A function φ on an interval $I \subset \mathbf{R}$ is said to be *locally integrable* if it is integrable on every compact subinterval of I, and

f is called an *indefinite Lebesgue integral* of a locally integrable function φ on I provided

$$f(b) - f(a) = \int_a^b \varphi \quad \text{whenever} \quad [a,b] \subset I\,.$$

If $c \in I$ and

$$\Phi(x) = \begin{cases} \displaystyle\int_c^x \varphi, & x \geq c, \\ \displaystyle -\int_x^c \varphi, & x < c\,, \end{cases}$$

then Φ is an indefinite Lebesgue integral of φ and any other indefinite integral of φ differs from Φ by an additive constant. If, in addition, φ is nonnegative, then Φ is nondecreasing. It follows that for a locally integrable function φ, the indefinite integral of φ is a function of finite variation (it is the difference of indefinite integrals of φ^+ and φ^-) and by Corollary 22.6 it has a finite derivative almost everywhere. Notice that according to the Lebesgue dominated convergence theorem any indefinite Lebesgue integral is a continuous function. This assertion can be sharpened considerably as the next theorem shows.

23.4. Theorem. *Let $\varphi \in \mathscr{L}^1(I)$ and f be an indefinite Lebesgue integral of φ on I. Then f is absolutely continuous and $f' = \varphi$ almost everywhere.*

Proof. Since $|f(y) - f(x)| \leq \int_x^y |f|$ for $[x,y] \subset I$, absolute continuity of f follows from Exercise 8.22.b. It remains to show that $f' = \varphi$ almost everywhere. Since f is absolutely continuous, f' exists almost everywhere, $f' \in \mathscr{L}^1(I)$ and for any interval $[a,b] \subset I$

$$\int_a^b f' = f(b) - f(a) = \int_a^b \varphi.$$

It readily follows that

$$\int_E f' = \int_E \varphi$$

for any open set E, and consequently for any measurable set E as well. Then Theorem 8.17 gives the desired conclusion. ■

23.5. Corollary. *For a real-valued function f on an interval $[a,b]$, the following properties are equivalent:*

(i) *f is absolutely continuous on $[a,b]$;*
(ii) *there exists $\varphi \in \mathscr{L}^1([a,b])$ such that $f(x) = f(a) + \int_a^x \varphi$ for all $x \in [a,b]$;*
(iii) *f is differentiable almost everywhere, $f' \in \mathscr{L}^1([a,b])$ and $f(x) = f(a) + \int_a^x f'$ for all $x \in [a,b]$.*

23.6 Corollary. *Suppose f is an absolutely continuous function on $[a,b]$. If $f' = 0$ almost everywhere, then f is constant.*

23.7. Remark. Let φ be a real-valued function on an interval $[a,b]$. Recall that the Newton integral of φ is the difference $f(b) - f(a)$, where f is an antiderivative of φ on the interval $[a,b]$. This definition can be generalized in many ways. We may modify the notion of an "antiderivative" to require the equality $f' = \varphi$ up to a countable set of points. To make this definition reasonable, we should add the assumption that f is continuous in order to guarantee that all "antiderivatives" of f differ up a constant. An analogous problem appears when assuming $f' = \varphi$ only almost everywhere. Now, the example of the Cantor function shows that continuity of f does not guarantee that the increment $f(b) - f(a)$ does not depend on choice of the "generalized antiderivative". However, it is possible to give an alternative definition of the Lebesgue integral of a function φ as the increment of an *absolutely continuous* function f on $[a, b]$ with $f' = \varphi$ almost everywhere. (Notice that this definition does not lead to a true generalization — some antiderivative are not absolutely continuous, see Example 25.1). The definitions of various integrals based on the idea of (in some way) generalized antiderivatives are called the *descriptive* ones. Let us note that these generalizations may consist in omitting "small sets" or in a "generalized differentiation". Furthermore, in Chapter 25 we mention Perron's method which is also included among descriptive approaches.

23.8. Lebesgue Points. Let $I \subset \mathbf{R}$ be an interval and $x \in I$. We say that x is a *Lebesgue point* for a locally integrable function f if

$$\lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} |f(x+t) - f(x)| \, \mathrm{d}t = 0.$$

If F is an indefinite Lebesgue integral of f on an interval I and $x \in I$ is a point where $F'(x) = f(x)$, then

$$\lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} (f(x+t) - f(x)) \, \mathrm{d}t = 0$$

but x does not need to be a Lebesgue point for f. However, it is clear that $F' = f$ at each Lebesgue point for f. The following theorem is thus a strengthening of Theorem 23.4.

23.9. Lebesgue Differentiation Theorem. *Let f be a locally integrable function on an interval I. Then almost every point of I is a Lebesgue point for f.*

Proof. For a fixed $r \in \mathbf{R}$, the function $x \mapsto |f(x) - r|$ is locally integrable on I. By virtue of Theorem 23.4 there exists a set $E_r \subset I$ of Lebesgue measure zero such that

$$\lim_{h \to 0} \frac{1}{2h} \int_{-h}^{h} |f(x+t) - r| \, \mathrm{d}t = |f(x) - r|$$

for every $x \in I \setminus E_r$. If $E := \bigcup_{r \in \mathbf{Q}} E_r$, then $\lambda E = 0$. Now, if $x \in I \setminus E$ and $\varepsilon > 0$, then there exists $r \in \mathbf{Q}$ with $|f(x) - r| < \varepsilon$. Consequently,

$$|f(x+t) - f(x)| \le |f(x+t) - r| + \varepsilon$$

and

$$\limsup_{h \to 0} \frac{1}{2h} \int_{-h}^{h} |f(x+t) - f(x)| \, \mathrm{d}t \le |f(x) - r| + \varepsilon \le 2\varepsilon.$$

■

23.10. Remark. Every continuity point of f is a Lebesgue point for f. An interesting relationship between Lebesgue points and points of approximate continuity will be given in Exercise 29.11.

23.11. Banach-Zarecki Theorem. Let f be a real-valued function on an interval $[a, b]$. The following assertions are equivalent:

(i) f is absolutely continuous on $[a, b]$;
(iv) f is continuous on $[a, b]$, f is of bounded variation and has Luzin's (N)-property.

Hint. We already know that any absolutely continuous function f is continuous and of finite variation. According to Theorem 23.2, $\lambda f(G) \leq \int_G |f'|$ for any open set $G \subset [a, b]$, and a moment's reflection shows that f has Luzin's (N)-property. Conversely, suppose that (iv) holds and let D be the set of all points where f has a finite derivative. By Exercise 20.8, D is measurable. Since $\lambda f([a, b] \setminus D) = 0$, Theorem 20.4 yields

$$|f(s) - f(t)| \leq \int_s^t |f'(\xi)| \, \mathrm{d}\xi$$

for every interval $[s, t] \subset [a, b]$. Now it is sufficient to notice that the function $x \to \int_a^x |f'|$ is absolutely continuous.

23.12. A Characterizations of Lipschitz Functions. Show that for a real-valued function f on an interval $[a, b]$ the following conditions are equivalent:

(i) f is Lipschitz on $[a, b]$;
(ii) for any $\varepsilon > 0$ there exists $\delta > 0$ so that for every finite collection of intervals $[a_j, b_j] \subset [a, b]$ with $\sum_j (b_j - a_j) < \delta$ the inequality $\sum_j |f(b_j) - f(a_j)| < \varepsilon$ holds;
(iii) f is absolutely continuous on $[a, b]$ and f' is bounded on $[a, b] \setminus N$, where $\lambda N = 0$.

(Compare with a similar characterization of absolutely continuous functions.)

23.13. Integration by Parts for Lebesgue Integral. Let f, g be absolutely continuous functions on an interval $[a, b]$. Then

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g.$$

23.14. Notes. Fundamental theorem (23.5) of "the calculus for Lebesgue integrals" was proved by H. Lebesgue [*1904]. The implication (iv) $\Longrightarrow$ (i) of Theorem 23.11 is due to S. Banach [1925].

24. Radon Measures on **R** and Distribution Functions

24.1. Distribution Functions. A *distribution function* of a Radon measure ν on **R** is a nondecreasing and right continuous function F on **R** such that $F(b) - F(a) = \nu(a, b]$ for every interval $(a, b] \subset \mathbf{R}$. It is obvious that any other distribution function of ν can differ from F by only a constant.

If ν is a probability measure, then a distribution function of ν, given as

$$G_\nu(x) := \nu(-\infty, x]\,,$$

is normalized so that $\lim_{x \to -\infty} G_\nu(x) = 0$ and $\lim_{x \to +\infty} G_\nu(x) = 1$.

24.2. Theorem. (a) *Let F be a nondecreasing and right continuous function on* **R**. *Then there exists a unique Radon measure ν_F on $\mathscr{B}(\mathbf{R})$ such that F is the distribution function of ν_F.*

(b) *Let ν be a Radon measure on* **R**. *Then there is a distribution function F of ν.*

Proof. (a) As for the uniqueness, any two such measures agree on all open (or compact) sets in $\mathbf{R}$. The existence can be proved in various ways, depending on what kind of general theorem we wish to use. For instance, we can use Hopf's extension theorem 5.5 or start from the covering collection $\{(a,b]\colon a<b,\ a,b\in\mathbf{R}\}$ and the set function $\nu_F(a,b] := F(b)-F(a)$, create the outer measure and to restrict it to measurable sets. In any case, we have to prove the following property: If $(a,b]\subset\bigcup_{n=1}^{\infty}(a_n,b_n]$, then $F(b)-F(a)\le\sum_{n=1}^{\infty}(F(b_n)-F(a_n))$. To see this, given $\varepsilon>0$, let $\delta_n>0$, $\delta>0$ be so that $F(b_n+\delta_n)<F(b_n)+\varepsilon 2^{-n}$, $F(a+\delta)<F(a)+\varepsilon$. Then consider the cover $\{(a_n,b_n+\delta_n)\}$ of the interval $[a+\delta,b]$.

(b) Setting

$$F_\nu(x) := \begin{cases} \nu(0,x] & \text{for } x>0,\\ 0 & \text{for } x=0,\\ -\nu(x,0] & \text{for } x<0, \end{cases}$$

F_ν is nondecreasing, right continuous, $F_\nu(0)=0$ and $\nu(a,b]=F_\nu(b)-F_\nu(a)$ for every interval $(a,b]\subset\mathbf{R}$. ∎

24.3. Theorem. *The following conditions are equivalent for a measure ν defined on $\mathscr{B}(\mathbf{R})$:*

(i) *ν is a Radon measure;*
(ii) *there exists a distribution function F such that $\nu(a,b]=F(b)-F(a)$ for every bounded interval $(a,b]$;*
(iii) *there exists a nondecreasing function φ on $\mathbf{R}$ such that ν is the Lebesgue-Stieltjes measure λ_φ of Example 14.8.b.*

Proof. Use previous results and the uniqueness part of the Riesz representation theorem 16.5. ∎

24.4. Remarks. 1. According to Remark 15.2 any finite measure on $\mathscr{B}(\mathbf{R})$ is a Radon measure. However, there exist σ-finite measures on $\mathscr{B}(\mathbf{R})$ which are not Radon measures (for example, the sum $\sum_k \delta_{1/k}$ of Dirac measures). Hence, these measures are not Lebesgue-Stieltjes measures and have not distribution functions.

2. The function φ from the last theorem need not be right continuous. However, as a monotone function it has a right limit at each point. If $\tilde\varphi(x) := \lim_{t\to x+}\varphi(t)$, then $\tilde\varphi$ is a distribution function of ν and determines the same Lebesgue-Stieltjes measure as φ. The functions $\tilde\varphi$ and F of Theorem 24.3.ii differ by a constant (cf. Exercise 24.6).

24.5. Exercise. Let F be a distribution function of a Radon measure μ on $\mathbf{R}$ and let μ_F correspond to F as in Theorem 24.2. Show that $\mu=\mu_F$ on $\mathscr{B}(\mathbf{R})$.

24.6. Exercise. Show that any two distribution functions of the same measure differ by a constant.

24.7. Exercise. Let F be a distribution function of a Radon measure μ. Show that:

(a) F is locally absolutely continuous on $\mathbf{R}$ if and only if μ is absolutely continuous with respect to the Lebesgue measure λ. In this case, F' is the Radon-Nikodým derivative $\mathrm{d}\mu/\,\mathrm{d}\lambda$.

(b) Measures μ and λ are mutually singular if and only if $F'=0$ almost everywhere.

(c) $\mu(\{x\})=F(x)-\lim_{t\to x-}F(t)$ for every $x\in\mathbf{R}$. In particular, F is continuous at x if and only if $\mu(\{x\})=0$.

(d) $z \in \operatorname{supt}\mu$ if and only if $F(z-\varepsilon) < F(z+\varepsilon)$ for every $\varepsilon > 0$.

(e) Let $\mu = \mu_d + \mu_a + \mu_s$ be the decomposition of μ into discrete, absolutely continuous and singular part (Exercise 15.18.c). If F_d, F_a, F_s are distribution functions of μ_d, μ_a, μ_s, respectively, then F_a is locally absolutely continuous, F_s is continuous, $F_s' = 0$ almost everywhere and F_d is called the *saltus function* of F. Moreover, $F = F_d + F_a + F_s$.

24.8. Exercise. (a) Let φ equal to the Cantor function (Example 23.1) on $[0,1]$ and $\varphi = c_{(0,\infty)}$ elsewhere on $\mathbf{R}$. Determine $\lambda_\varphi C$, where C is the Cantor set. Show that λ_φ is a continuous measure (see Exercise 15.18) and $\operatorname{supt}\lambda_\varphi = C$ (see Exercise 15.10).

(b) Let $\varepsilon \in [0,1]$. Does there exist a nondecreasing function ψ so that $\psi(0)=0$, $\psi(1)=1$ and $\lambda_\psi C = \varepsilon$?

Hint. Set $\psi(x) = \varepsilon\varphi(x) + (1-\varepsilon)x$, where φ is defined as in (a).

24.9. Exercise. Let F be a distribution function of a Radon measure μ. Assume that F is locally absolutely continuous on $\mathbf{R}$. Show that

$$\int_{\mathbf{R}} g\,\mathrm{d}\mu = \int_{\mathbf{R}} gF'$$

for every $g \in \mathscr{L}^1(\mu)$.

24.10. Fubini's Lemma. Let $\{f_n\}$ be a sequence of nondecreasing functions on an interval $I \subset \mathbf{R}$. Assume that the series $\sum f_n$ converges to a finite limit f on I. Then $f' = \sum f_n'$ almost everywhere on I.

Hint. First prove that $\lambda_f = \sum \lambda_{f_n}$ (notation as in 14.8). To finish the proof, use the Lebesgue decompositions of these measures to absolutely continuous and singular parts.

24.11. Notes. Distribution functions play an important role in probability theory and it is very difficult to trace who introduced them. They were investigated in various forms by Jacob Bernoulli, P.S. de Laplace and others. Lemma 24.10 is due to G. Fubini [1915].

25. Henstock–Kurzweil Integral

In some sense, the Lebesgue integral has the best properties among all integrals which can be defined on any measure space. However, its universality can also be a disadvantage. If we are engaged in integration of functions of one real variable (or several variables, but in what follows we restrict to the real line in order to simplify the explanation), then we can find wider collections of functions on which a reasonable notion of an integral can be introduced. Let us present an illustrative example.

25.1. Example. The function

$$f(x) = \begin{cases} \left(x^2 \cos^2 \frac{1}{x^2}\right)' & \text{for } x \neq 0, \\ 0 & \text{for } x = 0 \end{cases}$$

is not Lebesgue integrable on $[-1,1]$. For $a_j = \dfrac{1}{\sqrt{(j+\frac{1}{2})\pi}}$ and $b_j = \dfrac{1}{\sqrt{j\pi}}$ we have

$$\int_{a_j}^{b_j} |f(x)|\,\mathrm{d}x = \int_{a_j}^{b_j} f(x)\,\mathrm{d}x = \frac{1}{\pi j},$$

and therefore $\int_0^1 |f(x)|\,\mathrm{d}x = +\infty$. On the other hand, f has an antiderivative on $\mathbf{R}$ and the Newton integral $\int_0^1 f(x)\,\mathrm{d}x$ exists.

There are also simple examples of functions which are Lebesgue integrable but not Newton integrable (for example, the function $\operatorname{sign} x$ on $[-1, 1]$). A generalization of both Newton and Lebesgue integrals leads to a nonabsolutely convergent integral which can be introduced in several ways. Here we use an approach due to Henstock and Kurzweil; Perron's one is outlined in 25.10.

25.2. Henstock–Kurzweil integral. For the sake of simplicity we will define the integral for bounded intervals only.

If $[a, b]$ is an interval, denote the set of all (strictly) positive functions on $[a, b]$ by Δ and label functions from Δ as *gauges*. A *partition* is a pair $D = ([a_j, b_j], \xi_j)_{j=1}^m$, where

$$a = a_1 < b_1 = a_2 < b_2 = a_3 < \cdots < b_{m-1} = a_m < b_m = b \text{ and } \xi_j \in [a_j, b_j].$$

If $\delta \in \Delta$ is a gauge, then a partition D is called *δ-fine* (or *subordinated* to $\delta \in \Delta$) whenever $b_j - a_j < \delta(\xi_j)$ for all $j = 1, \dots, m$. The definition of the Henstock–Kurzweil integral is based on the following proposition whose proof is an easy consequence of the compactness of $[a, b]$.

Cousin's Lemma. *Let δ be a gauge from Δ. Then there exists a δ-fine partition. Moreover, if a collection $\mathscr{B}$ of non-overlapping intervals $[a_i, b_i]$ containing points ξ_i with $b_i - a_i < \delta(\xi_i)$ is given, then adding some intervals to $\mathscr{B}$ we get a δ-fine partition.*

Given a real-valued function f on $[a, b]$ and a partition $D = ([a_j, b_j]), \xi_j)_{j=1}^m$ of $[a, b]$, set

$$s(f, D) = \sum_{j=1}^{m} f(\xi_j)(b_j - a_j).$$

We say that f is *Henstock–Kurzweil integrable* on $[a, b]$ if there is a real number K so that for any $\varepsilon > 0$ there is $\delta \in \Delta$ such that

$$|K - s(f, D)| < \varepsilon$$

for each δ-fine partition D of $[a, b]$. The number K is then uniquely determined, it is called the *Henstock–Kurzweil integral* of f and denoted by $K \int_a^b f$. Where no confusion can result, we will drop the prefix "K". This definition is similar to original Riemann's one, the "only" difference being that instead of constants δ in Riemann's definition, a gauge *function* δ appears in Henstock–Kurzweil's definition.

The set of all Henstock–Kurzweil integrable functions is a linear space on which the Henstock–Kurzweil integral is a monotone linear functional.

We are not going to study details of the theory of nonabsolutely convergent integrals; we will concentrate on relations to the Newton and the Lebesgue integral only.

25.3 Theorem. *Let F be a continuous function on $[a, b]$, $F' = f$ on (a, b). Then the Henstock–Kurzweil integral of f on $[a, b]$ exists and equals to $F(b) - F(a)$.*

Proof. Given an $\varepsilon > 0$ and $x \in (a,b)$, there is $\delta(x) > 0$ such that

$$\left|\frac{F(y)-F(x)}{y-x} - f(x)\right| < \varepsilon$$

whenever $y \in [a,b]$, $0 < |y-x| < \delta(x)$. Next, we can find $\delta(a) > 0$ and $\delta(b) > 0$ such that $|f(a)\delta(a)| < \varepsilon$, $|f(b)\delta(b)| < \varepsilon$, $|F(y)-F(a)| < \varepsilon$ for all $y \in [a, a+\delta(a))$ and $|F(y)-F(a)| < \varepsilon$ for all $y \in (b-\delta(b), b]$. Let $D = ([a_j,b_j]),\xi_j)_{j=1}^m$ be a δ-fine partition. Then

$$\begin{gathered}|F(b_j)-F(a_j)-f(\xi_j)(b_j-a_j)| \le \\ |F(b_j)-F(\xi_j)-f(\xi_j)(b_j-\xi_j)| + |F(\xi_j)-F(a_j)-f(\xi_j)(\xi_j-a_j)| \\ < \varepsilon(|b_j-\xi_j|+|\xi_j-a_j|) = \varepsilon(b_j-a_j)\end{gathered}$$

for every $j \in \{1,\dots,m\}$, $\xi_j \in (a,b)$, and

$$|F(b_j)-F(a_j)-f(\xi_j)(b_j-a_j)| \le |F(b_j)-F(a_j)| + |f(\xi_j)(b_j-a_j)| < 2\varepsilon$$

if $\xi_j \in \{a,b\}$. Summing over j, we obtain the estimate

$$|F(b)-F(a)-s(f,D)| < 4\varepsilon + (b-a)\varepsilon.$$

■

25.4. Theorem. *If a function f has a finite Lebesgue integral L, then f also has the Henstock–Kurzweil integral and $K\int_a^b f = L$.*

Proof. If $\varepsilon > 0$, then Theorem 15.5 ensures the existence of a lower semicontinuous function $s > f$ and an upper semicontinuous function $t < f$ such that $s,t \in \mathscr{L}^1([a,b])$ and

$$\int_a^b (s-t) < \varepsilon$$

(we get strict inequalities by adding small constants). For every $x \in [a,b]$ find $\delta(x) > 0$ with $t < f(x) < s$ on $(x-\delta(x), x+\delta(x)) \cap [a,b]$. Let $D = ([a_j,b_j],\xi_j)_{j=1}^m$ be a δ-fine partition of $[a,b]$. Then

$$\int_{a_j}^{b_j} t \le f(\xi_j)(b_j-a_j) \le \int_{a_j}^{b_j} s$$

for each $j \in \{1,\dots,m\}$. Summing we get

$$\int_a^b t \le s(f,D) \le \int_a^b s.$$

Since $\int_a^b t \le L \le \int_a^b s$, we have $|s(f,D)-L| < 2\varepsilon$, and we are done. ■

25.5 Indefinite Henstock–Kurzweil Integral. In the sequel, $[a,b]$ will be a fixed interval and Δ will denote the set of all positive functions on $[a,b]$. If a function f has the Henstock–Kurzweil integral on $[a,b]$ and $[a',b'] \subset [a,b]$, then f has the Henstock–Kurzweil integral on the interval $[a',b']$ as well. Moreover,

$$\int_{c_1}^{c_3} f = \int_{c_1}^{c_2} f + \int_{c_2}^{c_3} f$$

whenever $a \le c_1 < c_2 < c_3 \le b$.

A function F on $[a,b]$ is called an *indefinite Henstock–Kurzweil integral* of f if, for each interval $[a',b'] \subset [a,b]$,

$$F(b') - F(a') = K \int_{a'}^{b'} f\,.$$

If F is an indefinite Henstock–Kurzweil integral of f, then any other indefinite Henstock–Kurzweil integral of f can differ from F by only a constant.

25.6. Saks-Henstock's Lemma. *Let F be an indefinite Henstock–Kurzweil integral of f on $[a,b]$. Then for any $\varepsilon > 0$ there exists a $\delta \in \Delta$ such that*

$$\Big|\sum_{j\in M} F(b_j) - F(a_j) - f(\xi_j)(b_j - a_j)\Big| < \varepsilon$$

whenever $D = ([a_j,b_j],\xi_j)_{j=1}^m$ is a δ-fine partition of $[a,b]$ and $M \subset \{1,\dots,m\}$.

Proof. Choose an $\varepsilon > 0$ and find a gauge $\delta \in \Delta$ such that

$$|F(b) - F(a) - s(f,D)| < \frac{\varepsilon}{2}$$

whenever $D = ([a_j,b_j],\xi_j)_{j=1}^m$ is a δ-fine partition of $[a,b]$. Fix now any such a partition D and $M \subset \{1,\dots,m\}$. For each $j = 1,\dots,m$ there is a partition $D_j = ([a_j^i,b_j^i],\xi_j^i)_{i=1}^{m_j}$ of $[a_j,b_j]$ subordinated to the restriction of δ to $[a_j,b_j]$ such that

$$|F(b_j) - F(a_j) - s(f,D_j)| < \frac{\varepsilon}{2m}.$$

Create a new partition $D^* = ([a_k^*,b_k^*],\xi_k^*)_{k=1}^p$ of $[a,b]$ in such way that every triple (a_k^*,b_k^*,ξ_k^*) is either one of the triplets (a_j^i,b_j^i,ξ_j^i), where $j \notin M$, $i \in \{1,\dots,m_j\}$, or one of the triplets (a_j,b_j,ξ_j), where $j \in M$. This partition is δ-fine and thus

$$|F(b) - F(a) - s(f,D^*)| < \frac{\varepsilon}{2}.$$

Whence combining all inequalities, we get

$$\Big|\sum_{j\in M} F(b_j) - F(a_j) - f(\xi_j)(b_j - a_j)\Big| < \varepsilon.$$

■

25.7. Theorem. *An indefinite Henstock–Kurzweil integral F of f on $[a,b]$ is continuous on $[a,b]$.*

Proof. Let $z \in [a,b]$ and $\varepsilon > 0$ be fixed. By previous Saks–Henstock's lemma find a gauge $\delta \in \Delta$ such that

$$(*) \qquad \Big|\sum_{j\in M} F(b_j) - F(a_j) - f(\xi_j)(b_j - a_j)\Big| < \varepsilon$$

whenever $D = ([a_j, b_j], \xi_j)_{j=1}^m$ is a δ-fine partition of $[a,b]$ and $M \subset \{1,\dots,m\}$. For $x \in [a,b]$, $|x - z| < \delta(z)$, let $D = ([a_j, b_j], \xi_j)_{j=1}^m$ be a partition such that $\xi_k = z$ for some $k \in \{1,\dots,m\}$ and such that the endpoints of the interval $[a_k, b_k]$ are x and z. Applying (*) with $M = \{k\}$, we obtain

$$|F(x) - F(z)| < \varepsilon + |f(z)|\,|x - z|\,.$$

■

Based on Henstock–Kurzweil's definition, the following theorem (even in higher dimensions) was proved by J. Král [1985].

25.8. Theorem. *Let F be an indefinite integral of a Henstock–Kurzweil integrable function f on $[a,b]$. Then f is a measurable function, the derivative F' exists at almost all points of (a,b) and $F' = f$ almost everywhere.*

Proof. The continuity of F (Theorem 25.7) implies that $\overline{D}F$ is measurable (even Borel) since

$$\overline{D}F(x) = \lim_{n\to\infty}\lim_{k\to\infty} f_{n,k}\,,$$

where functions $f_{n,k}$ defined as

$$f_{n,k}(x) = \sup\left\{\frac{F(y) - F(x)}{y - x} : y \in [a,b],\ \frac{1}{n+k} \le |y - x| \le \frac{1}{n}\right\}$$

are continuous. Now we prove that $\underline{D}F = \overline{D}F$ almost everywhere which establishes the existence of the derivative almost everywhere and its measurability.

We will be done once we show that, for every $n \in \mathbf{N}$, the level sets

$$L_n := \{x \in [a,b] : \underline{D}F(x) < f(x) - \frac{1}{n}\} \quad \text{and}$$
$$U_n := \{x \in [a,b] : \overline{D}F(x) > f(x) + \frac{1}{n}\}$$

have measure zero. For this purpose we will fix $n \in \mathbf{N}$ and estimate the (outer) measure of the set $L := L_n$. We are now in a position to invoke Saks–Henstock's lemma 25.6: Given $\varepsilon > 0$, there is a gauge function $\delta \in \Delta$ such that

$$\Big|\sum_{j\in M} F(b_j) - F(a_j) - f(\xi_j)(b_j - a_j)\Big| < \varepsilon$$

whenever $D = ([a_j, b_j], \xi_j)_{j=1}^m$ is a δ-fine partition of $[a, b]$ and $M \subset \{1, \ldots, m\}$. Let $\mathscr{V}$ denote the family of all intervals $[a', b']$ for which there exists an $x \in [a', b'] \cap L$ such that $b' - a' < \delta(x)$ and

$$\frac{F(b') - F(a')}{b' - a'} < f(x) - \frac{1}{n}.$$

Then $\mathscr{V}$ is a Vitali cover of L, and so, by Corollary 27.3, there exists a finite, pairwise disjoint subfamily $\{[a_i, b_i]\}_i$ of $\mathscr{V}$ with points $\xi_i \in [a_i, b_i]$ such that

$$\lambda^*(L \setminus \bigcup_i [a_i, b_i]) < \varepsilon.$$

Now, use Cousin's lemma in 25.2 to extend the collection $([a_i, b_i], \xi_i)_i$ to a family $([a_j^*, b_j^*], \xi_j^*)$ subordinated to δ. If $M := \{j\colon$ there exists an i with $[a_j^*, b_j^*] = [a_i, b_i]\}$, the property of gauge δ yields that

$$\sum_i (f(\xi_i)(b_i - a_i) - (F(b_i) - F(a_i))) < \varepsilon\,,$$

and the definition of $\mathscr{V}$ that

$$\sum_i (f(\xi_i)(b_i - a_i) - (F(b_i) - F(a_i))) > \frac{1}{n} \sum_i (b_i - a_i).$$

Thus $\sum_i (b_i - a_i) < n\varepsilon$, which in turn implies that

$$\lambda^*(L) \leq n\varepsilon + \lambda^*(L \setminus \bigcup_i [a_i, b_i]) < (n+1)\varepsilon.$$

■

The following proposition illustrates the relation between Henstock-Kurzweil and Lebesgue integrals — their equivalence for *nonnegative* functions.

25.9. Theorem. *If a nonnegative function f has the Henstock–Kurzweil integral on $[a, b]$, then f has also the Lebesgue integral on $[a, b]$ and $L \int_a^b f = K \int_a^b f$.*

Proof. According to Theorem 25.6, f is measurable. It remains to show that $L \int_a^b f < \infty$ and use Theorem 25.4. The proof may now be finished in a stroke: Since $f \geq 0$, the indefinite Henstock–Kurzweil integral F of f is nondecreasing on $[a, b]$ and thus by Theorems 22.7 and 25.8 $L \int_a^b f = L \int_a^b F' \leq F(b) - F(a) < +\infty$. ■

25.10. Perron Integral. There are also other ways how to define an integral which is equivalent to the Henstock–Kurzweil integral (i.e. it provides the same class of "integrable" functions on which it equals to the Henstock–Kurzweil integral). In the following, we outline Perron's approach to generalizing both the Newton and the Lebesgue integrals on intervals. Let us start with two notes:

(a) A descriptive definition of the Newton integral of a function f requires the existence of an antiderivative (a function F with $F' = f$). This is a very strong condition as it reduces the class of Newton integrable functions. Among others, this condition implies that f shares the Darboux property and is of the Baire class one. A descriptive definition of the Lebesgue integral requires the existence of an absolutely continuous function F with $F' = f$ only almost everywhere.

(b) The following particular case of Theorem 15.5 is known as the Vitali-Carathéodory theorem: A function f is Lebesgue integrable on an interval I if for any $\varepsilon > 0$ there exists an upper semicontinuous function g and a lower semicontinuous function h with $g \le f \le h$ on I and $\int_I (h - g) < \varepsilon$.

Now we are in the position to introduce Perron's integral. Let f be a function on an interval $[a, b]$. A function M is termed to be a *majorant* of f if $f(x) \le \underline{D}M(x) \ne -\infty$ for every $x \in [a, b]$. If $f \ge \overline{D}m \ne +\infty$, we say that m is a *minorant* of f.

A function f is said to be *Perron integrable* on an interval $[a, b]$ if for any $\varepsilon > 0$ there exists a majorant M and a minorant m so that

$$M(b) - M(a) - (m(b) - m(a)) < \varepsilon.$$

Since $\underline{D}(M - m) \ge \underline{D}M - \overline{D}m \ge 0$ (and since every function with a nonnegative lower derivative is nonnegative) we may define the *Perron integral* of f as

$$\begin{aligned} P\int_a^b f &= \inf\{M(b) - M(a) \colon M \text{ is a majorant of } f\} \\ &= \sup\{m(b) - m(a) \colon m \text{ is a minorant of } f\}. \end{aligned}$$

We may also define upper and lower Perron integrals of f, and f will be Perron integrable when they are equal and finite.

25.11 Remark. Lebesgue integrable functions may be characterized in the following way:

A function f on an interval $[a, b]$ is *Lebesgue* integrable if and only if for any $\varepsilon > 0$ there is a majorant M and a minorant m, both of them *absolutely continuous*, so that

$$M(b) - M(a) - (m(b) - m(a)) < \varepsilon.$$

To prove this proposition, one can use the Vitali-Carathéodory characterization (Theorem 15.5) again (the indefinite integrals of semicontinuous functions serve as majorants and minorants).

It is quite interesting to notice that in the definition of the Perron integral we can restrict to continuous majorants:

A function f on an interval $[a, b]$ is *Perron* integrable if and only if for any $\varepsilon > 0$ there is a majorant M and a minorant m, both of them *continuous*, so that

$$M(b) - M(a) - (m(b) - m(a)) < \varepsilon.$$

Finally, the Riemann integrable functions may be characterized in a similar way:

A function f on an interval $[a, b]$ is *Riemann* integrable if and only if for any $\varepsilon > 0$ there is a majorant M and a minorant m, both of them *with continuous derivative*, so that

$$M(b) - M(a) - (m(b) - m(a)) < \varepsilon.$$

25.12. Exercise. Find the gauge function δ from the definition of the Henstock–Kurzweil integral if f is:

(a) Newton integrable;
(b) Lebesgue integrable;
(c) an indicator function of an open set;
(d) an indicator function of a null set.

25.13 Exercise. Give an example to show that the indefinite Henstock–Kurzweil integral need not be an absolutely continuous function.

25.14. Historical Notes. H. Lebesgue already knew that his integral on the real line does not integrate all derivatives. The task was, to find a wider class of "integrable functions" containing both the Lebesgue and Newton integrable functions. The problem was solved by A. Denjoy [1912] who used a constructive method based on a transfinite approach, and by N. N. Luzin [1912] using a descriptive definition based on further generalization of the notion of absolute continuity. The method of envelopes was developed by O. Perron [1914]; it was proved that his and Denjoy's integrals are equivalent. Nowadays, this integral is usually called the *Denjoy-Perron integral* or the *restricted Denjoy integral*. Another approach was developed independently by R. Henstock [1961] and J. Kurzweil [1957] who returned to a Riemann type definition. The integral is often called the *complete Riemann* or the *Henstock-Kurzweil integral*. Simple definitions of Henstock's and Kurzweil's approach and simple proofs allowed a deeper study of the Henstock–Kurzweil integral and surprising applications. Among recent publications let us mention R. Henstock [*1988], Peng Yee Lee [*1989], Š. Schwabik [*1992] and Š. Schwabik and Guoju Ye [*2005].

E. Integration on $\mathbf{R}^n$

26. Lebesgue Measure and Integral on $\mathbf{R}^n$

Recall that the outer Lebesgue measure λ_n^* is defined as

$$\lambda_n^* A := \inf\left\{\sum \operatorname{vol} I_k : \bigcup I_k \supset A, I_k \text{ are open intervals}\right\}$$

and the Lebesgue measure λ_n is defined as the restriction of λ_n^* to the σ-algebra $\mathfrak{M}_n = \mathfrak{M}_n(\lambda_n^*)$ of all Lebesgue measurable sets. Where no confusion can result, the subscript "n" will be omitted.

26.1. Theorem. *The Lebesgue measure λ_n is a complete Radon measure.*

Proof. Use Theorem 1.21 and Remark 15.2.4 to prove that λ_n is a Radon measure. Completeness is a consequence of Caratheodory's construction, see Theorem 4.5. ■

26.2. Theorem. *The Lebesgue measure λ_{n+k} on $\mathfrak{M}_{n+k}$ is the completion of the product measure $\lambda_n \otimes \lambda_k$ on $\mathfrak{M}_n \otimes \mathfrak{M}_k$.*

Proof. Recall that $A \in \mathfrak{M}_n$ if and only if A is a union of a Borel set and a set of measure zero. Indeed, such a property is possessed by any complete Radon measure and then we may refer to Theorem 26.1. The proof is now divided into three steps.

1. First we show that $\mathfrak{M}_n \otimes \mathfrak{M}_k \subset \mathfrak{M}_{n+k}$. For this purpose it is enough to show that any measurable rectangle $A \times B$ for $A \in \mathfrak{M}_n$, $B \in \mathfrak{M}_k$ is in $\mathfrak{M}_{n+k}$. We write $A = D_A \cup N_A$ and $B = D_B \cup N_B$, where D_A, D_B are Borel sets and N_A, N_B have measure zero. Then $D_A \times D_B$ is a Borel set and $(A \times B) \setminus (D_A \times D_B) = (A \times N_B) \cup (N_A \times B)$ has measure zero.

2. Conversely, $\mathscr{B}_{n+k} = \mathscr{B}_n \otimes \mathscr{B}_k \subset \mathfrak{M}_n \otimes \mathfrak{M}_k$. Since $(\mathfrak{M}_{n+k}, \lambda_{n+k})$ is the completion of $(\mathscr{B}_{n+k}, \lambda_{n+k})$, $\mathfrak{M}_{n+k} \subset \mathfrak{M}_n \otimes \mathfrak{M}_k$.

3. Since $\lambda_n \otimes \lambda_k = \lambda_{n+k}$ on open intervals of $\mathbf{R}^{n+k}$, using Hopf's extension theorem 5.5 and uniqueness of the completion we conclude that $\lambda_n \otimes \lambda_k = \lambda_{n+k}$ on $\mathfrak{M}_{n+k}$.

26.3. Remark. It is not very hard to prove that $\mathscr{B}_n \otimes \mathscr{B}_k = \mathscr{B}_{n+k}$ in Euclidean spaces. For measurable sets, the situation is much more complicated. Indeed,

$$\mathscr{B}_{n+k} \subset \mathfrak{M}_n \otimes \mathfrak{M}_k \subset \mathfrak{M}_{n+k} \quad \text{and} \quad \mathscr{B}_{n+k} \neq \mathfrak{M}_n \otimes \mathfrak{M}_k \neq \mathfrak{M}_{n+k}.$$

For the first counterexample consider the Cartesian product $H \times H$, where H is a measurable non-Borel subset of $\mathbf{R}$ (Exercise 1.14.b) and use Lemma 11.2. The second one can be find in Remark 11.10.

26.4. Exercise. Show that the Lebesgue measure is the only complete Radon measure on $\mathbf{R}^n$ that assigns to each interval its volume.

26.5. Exercise. The Lebesgue measure is translation-invariant: If $A \in \mathfrak{M}_n$ and $x \in \mathbf{R}^n$, then $A + x \in \mathfrak{M}_n$ and $\lambda_n(A + x) = \lambda_n A$.

26.6. Exercise. Let ν be a nontrivial translation-invariant complete Radon measure on $\mathbf{R}^n$. Show that there is $c > 0$ such that $\nu = c\lambda_n$.

Hint. Let $c = \nu(0,1)^n$. Evaluate the ν-measure of arbitrary intervals and use Exercise 26.4 to compare λ_n and $\frac{1}{c}\nu$.

26.7. Lebesgue Measurable Functions. Since $\mathfrak{M}_n$ contains the family of all Borel sets, each Borel (in particular, each continuous) function (on $\mathbf{R}^n$) is measurable. A deeper and complete characterization of measurable functions (e.g. on open or closed subsets of $\mathbf{R}^n$) is contained in Luzin's Theorem 18.2.

Basic tools of the integral calculus of functions of several variables are substitution theorem and Fubini's theorem. The last one allow to reduce the multidimensional integration to a succession of one-dimensional integrals.

26.8. Introduction to Fubini's Theorem. If $M \subset \mathbf{R}^{n+k}$ is a measurable set, let $\mathscr{L}^*(M)$ denote the set of all functions f which have (finite or infinite) integral $\int_M f \, \mathrm{d}\lambda_{n+k}$ (which will be traditionally denoted by $\int_M f(x,y) \, \mathrm{d}x \, \mathrm{d}y$). Given $x \in \mathbf{R}^n$, recall that M^x stands for the set $\{y \in \mathbf{R}^k : [x,y] \in M\}$.

26.9. Fubini's Theorem. *Let $M \subset \mathbf{R}^{n+k}$ be a measurable set and $f \in \mathscr{L}^*(M)$. Then the integral*

$$g(x) := \int_{M_x} f(x, \cdot) \, \mathrm{d}\lambda_k$$

exists for almost all $x \in \mathbf{R}^n$, and

$$\int_M f \, \mathrm{d}\lambda_{n+k} = \int_{\mathbf{R}^n} g \, \mathrm{d}\lambda_n \, .$$

In other words,

$$\int_M f(x,y) \, \mathrm{d}x \, \mathrm{d}y = \int_{\mathbf{R}^n} \left(\int_{M^x} f(x,y) \, \mathrm{d}y \right) \mathrm{d}x.$$

Proof. To prove the assertion it suffices to use Theorems 11.11. and 26.2. ∎

26.10. Remarks. 1. The assumption $f \in \mathscr{L}^*(M)$ is satisfied if, for instance, $f \in \mathscr{L}^1(M)$ (Tonelli's theorem), or if f is a non-negative measurable function (Fubini's theorem in a narrow sense). When applying Fubini's theorem, we usually start with Fubini's theorem in a narrow sense in order to show that $|f|$ is integrable, and then we use Tonelli's theorem.

2. Let πM denote the projection of M into $\mathbf{R}^n$ (i.e. πM is the set of all points x for which M^x is nonempty). Then

$$\int_M f(x,y) \, \mathrm{d}x \, \mathrm{d}y = \int_{\pi M} \left(\int_{M^x} f(x,y) \, \mathrm{d}y \right) \mathrm{d}x$$

provided πM is measurable. In general, the measurability of M does not imply the measurability of πM. Observe that πM is measurable whenever M is open (then πM is also open), or compact (then πM is compact), or a countable union of compact sets (e.g. a closed set). Projections of Borel sets are also measurable but the proof is difficult. The study of projections of Borel sets led to the notion of *analytic sets* (see D.L. Cohn [*1980]).

26.11. Example. We evaluate the integral $\int_B f(z) \, \mathrm{d}z$, where f is an integrable function on the ball $B := \{z \in \mathbf{R}^3 \colon |z| < 1\}$. Let us write z in the form $[x,y]$, where $x = [z_1, z_2] \in \mathbf{R}^2$ and $y = z_3 \in \mathbf{R}$. Then $B^x = \{y \in \mathbf{R} \colon |y| < \sqrt{1 - |x|^2}\}$ and $\pi B = \{x \in \mathbf{R}^2 \colon |x| < 1\}$. Hence

$$\begin{aligned} \int_B f(z) \, \mathrm{d}z &= \int_{\pi B} \left(\int_{-\sqrt{1-|x|^2}}^{\sqrt{1-|x|^2}} f(x,y) \, \mathrm{d}y \right) \mathrm{d}x \\ &= \int_{-1}^{1} \left(\int_{-\sqrt{1-z_1^2}}^{\sqrt{1-z_1^2}} \left(\int_{-\sqrt{1-z_1^2-z_2^2}}^{\sqrt{1-z_1^2-z_2^2}} f(z_1, z_2, z_3) \, \mathrm{d}z_3 \right) \mathrm{d}z_2 \right) \mathrm{d}z_1 . \end{aligned}$$

26.12. Jacobian and Derivative. If $G \subset \mathbf{R}^k$ is an open set and $\varphi = [\varphi_1, \dots, \varphi_n]\colon G \to \mathbf{R}^n$ has all partial derivatives at a point $z \in G$, then the *Jacobi matrix* $(\frac{\partial \varphi_i}{\partial x_j}(z))_{\substack{i=1,\dots,n \\ j=1,\dots,k}}$ of the mapping φ at z will be denoted by $\nabla\varphi(z)$.

We say that φ is *(Fréchet) differentiable* at z if there exists a linear mapping $L\colon \mathbf{R}^k \to \mathbf{R}^n$ such that

$$\lim_{h\to 0} \frac{\varphi(z+h) - \varphi(z) - L(h)}{|h|} = 0\,.$$

In this case, L is represented by the matrix $\nabla\varphi(z)$, it is called the *derivative* of φ at z and it is denoted by $\varphi'(z)$. If all partial derivatives of φ exist and are continuous at z, then φ is differentiable at z.

A mapping which has continuous all partial derivatives of order $\le k$ is called a $\mathscr{C}^k$*-mapping*, or a mapping of *class* $\mathscr{C}^k$. A $\mathscr{C}^\infty$*-mappings*, or mappings of *class* $\mathscr{C}^\infty$ are those possessing continuous partial derivatives of all orders.

In case $k = n$, the Jacobi matrix $\nabla\varphi(z)$ is a square matrix and its determinant $J_\varphi(z)$ is called the *Jacobian* of φ at z.

We state the following theorem without proof, later we provide a proof of a more general Theorem 34.18.

26.13. Change of Variable Formula. *Let $G \subset \mathbf{R}^n$ be an open set and $\varphi\colon G \to \mathbf{R}^n$ a one-to-one $\mathscr{C}^1$-mapping such that $J_\varphi(z) \neq 0$ for all $z \in G$. Let f be a function on $\varphi(G)$ and $E \subset \varphi(G)$ a measurable set. Then*

$$\int_E f(x)\,\mathrm{d}x = \int_{\varphi^{-1}(E)} f(\varphi(t))\,|J_\varphi(t)|\,\mathrm{d}t$$

provided either of these integrals exists.

26.14. Polar Coordinates. Consider the mapping $\varphi\colon [r,t] \mapsto [r\cos t, r\sin t]$. Then φ is of class $\mathscr{C}^\infty$ in $\mathbf{R}^2$,

$$\varphi'(r,t) = \begin{pmatrix} \cos t, & -r\sin t \\ \sin t, & r\cos t \end{pmatrix}$$

and $J_\varphi(r,t) = r$. The mapping φ is not one-to-one. For calculation of integrals, the substitution $x = \varphi(r,t)$ is often useful (usually $G = (0,\infty)\times(0,2\pi)$). In this situation, the hypothesis of the change of variable formula in 26.13 are satisfied (notice that $\varphi(G) \neq \mathbf{R}^2$ but $\lambda(\mathbf{R}^2 \setminus \varphi(G)) = 0$).

26.15. Lemniscate. Calculate the area of the set M surrounded by the lemniscate

$$M := \{[x,y] \in \mathbf{R}^2\colon (x^2+y^2)^2 < 2a^2(x^2-y^2)\}\,.$$

Hint. If φ denotes a mapping of 26.14 (polar coordinates) and if

$$L := \{[r,t] \in \mathbf{R}^2\colon r \in (0, \sqrt{2a^2\cos 2t}),\ t \in (0, \frac{\pi}{4})\}\,,$$

then $\varphi(L) = M \cap \{[x,y] : x > 0, y > 0\}$. By the change of variable formula 26.13 we get

$$\lambda_2(M) = 4\int_L r\,\mathrm{d}r\,\mathrm{d}t = 4a^2\int_0^{\frac{\pi}{4}} \cos 2t\,\mathrm{d}t = 2a^2.$$

26.16. Spherical Coordinates. Consider the mapping $\varphi = [x, y, z]$, where

$$x(r,t,\theta) = r\cos\theta\,\cos t, \quad y(r,t,\theta) = r\cos\theta\,\sin t, \quad z(r,t,\theta) = r\sin\theta$$

and $[r, t, \theta] \in G := (0, \infty) \times (-\pi, \pi) \times (-\frac{1}{2}\pi,\ \frac{1}{2}\pi)$. Then

$$\varphi'(r,t,\theta) = \begin{pmatrix} \cos\theta\cos t, & -r\sin t\cos\theta, & -r\cos t\sin\theta \\ \cos\theta\sin t, & r\cos t\cos\theta, & -r\sin t\sin\theta \\ \sin\theta, & 0, & r\cos\theta \end{pmatrix}.$$

Thus $J_\varphi(r,t,\theta) = r^2\cos\theta$. Again $\varphi(G) \neq \mathbf{R}^3$. However, the set $\mathbf{R}^3 \setminus \varphi(G) = \{[x,y,z] \in \mathbf{R}^3 : y = 0 \text{ and } x \le 0\}$ is of measure zero.

26.17. Exercise. Let κ_n denote the volume of the n-dimensional unit ball.

(a) Let φ be a nonnegative measurable function on $(0, R)$. Prove that

$$\int_{B(0,R)} \varphi(|x|)\,\mathrm{d}x = n\kappa_n \int_0^R r^{n-1}\varphi(r)\,\mathrm{d}r.$$

Hint. Consider first the case when φ is piecewise constant and then approximate φ by piecewise constant functions.

(b) Show that $\kappa_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$.

Hint. Write $I_n = \int_{\mathbf{R}^n} e^{-|x|^2}\,dx$. Using Fubini's theorem we obtain $I_n = I_1^n$. By the previous part (a)

$$I_n = n\kappa_n \int_0^\infty r^{n-1}\,e^{-r^2}\,dr = \kappa_n\,\frac{n}{2}\,\Gamma(\frac{n}{2}) = \kappa_n\Gamma(\frac{n}{2}+1)\,.$$

It is easy to compute $I_2 = \kappa_2\Gamma(2) = \pi$. Hence $I_n = \pi^{n/2}$ and we are done.

26.18. Convolution of Functions and Measures. Let f, g be measurable functions on $\mathbf{R}^n$. The *convolution* of f and g is the function $f * g$ defined as

$$f{*}g\,(x) = \int_{\mathbf{R}^n} f(x-y)g(y)\,\mathrm{d}y$$

at those points x for which the integral exists.

Let f be a function on $\mathbf{R}^n$ and μ a (signed) Radon measure on $\mathbf{R}^n$. If the expression $f * \mu(x) := \int_{\mathbf{R}^n} f(x-y)\,\mathrm{d}\mu(y)$ makes sense for almost all $x \in \mathbf{R}^n$, then the function $f * \mu$ is called the *convolution* of f and μ.

26.19. Theorem. *If $f, g \in \mathscr{L}^1(\mathbf{R}^n)$, then the function $y \mapsto f(x-y)g(y)$ is in $\mathscr{L}^1(\mathbf{R}^n)$ for almost all $x \in \mathbf{R}^n$, $f * g \in \mathscr{L}^1$ and $\|f * g\|_1 \le \|f\|_1\,\|g\|_1$.*

Proof. First we have to show that the function $[x,y] \mapsto f(x-y)g(y)$ is measurable on $\mathbf{R}^{2n}$. Obviously, the function $[x,y] \mapsto g(y)$ is measurable on $\mathbf{R}^{2n}$ and we need to prove the measurability of $[x,y] \mapsto f(x-y)$. To this end it is enough to show that the set $M := \{[x,y] \in \mathbf{R}^n \times \mathbf{R}^n : x - y \in E\}$ is measurable for every measurable set $E \subset \mathbf{R}^n$. But this is apparent since M is of the form $E \times \mathbf{R}^n$ in the coordinate system $x' = x - y$, $y' = x + y$. Next, we use Fubini's Theorem 26.10 (in a narrow sense) to obtain

$$\begin{aligned}\int_{\mathbf{R}^{2n}} |f(x-y)g(y)|\,\mathrm{d}x\,\mathrm{d}y &= \int_{\mathbf{R}^n} \left| g(y) \left(\int_{\mathbf{R}^n} |f(x-y)|\,\mathrm{d}x \right) \right| \mathrm{d}y \\ &= \|f\|_1\,\|g\|_1 < \infty.\end{aligned}$$

Using Fubini's theorem again (now Tonelli's one) the desired conclusion follows. ■

26.20. Young's Convolution Theorem. *Let* $1 \le p, q, r \le \infty$, $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$. *If* $f \in \mathscr{L}^p$, $g \in \mathscr{L}^q$, *then* $f * g = g * f$ *is defined almost everywhere, it is an element of* $\mathscr{L}^r$ *and* $\|f * g\|_r \le \|f\|_p \|g\|_q$.

Proof. With trivial cases out of the way (cf. also the proof of Theorem 26.21) we assume $p, q, r < \infty$. Denoting $\tilde{f} = |f|^p$, $\tilde{g} = |g|^q$, $\tilde{f}$ and $\tilde{g}$ are in $\mathscr{L}^1$. By the previous theorem the convolution $\tilde{f} * \tilde{g}$ is defined almost everywhere and it is in $\mathscr{L}^1$. For a fixed x, the Hölder inequality yields

$$\begin{aligned} \left| \int_{\mathbf{R}^n} f(x-y) g(y) \, \mathrm{d}y \right| &\le \int_{\mathbf{R}^n} |f(x-y)|^{\frac{r-p}{r}} |g(y)|^{\frac{r-q}{r}} \left(|f(x-y)|^{\frac{p}{r}} |g(y)|^{\frac{q}{r}} \right) \\ &\le \left(\int_{\mathbf{R}^n} |f(x-y)|^p \right)^{\frac{r-p}{pr}} \left(\int_{\mathbf{R}^n} |g(y)|^q \right)^{\frac{r-q}{qr}} \left(\int_{\mathbf{R}^n} |f(x-y)|^p |g(y)|^q \right)^{\frac{1}{r}} \\ &= \|f\|_p^{\frac{r-p}{r}} \|g\|_q^{\frac{r-q}{r}} ((\tilde{f} * \tilde{g})(x))^{\frac{1}{r}}. \end{aligned}$$

By Fubini's theorem we have the equality

$$\begin{aligned} \int_{\mathbf{R}^n} \tilde{f} * g &= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} \tilde{f}(x-y) \tilde{g}(y) \, \mathrm{d}y \right) \mathrm{d}x = \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} \tilde{f}(x-y) \tilde{g}(y) \, \mathrm{d}x \right) \mathrm{d}y \\ &= \|f\|_p^p \|g\|_q^q , \end{aligned}$$

and consequently

$$\begin{aligned} \|f * g\|_r^r &= \int_{\mathbf{R}^n} \left| \int_{\mathbf{R}^n} f(x-y) g(y) \, \mathrm{d}y \right|^r \mathrm{d}x \le \|f\|_p^{r-p} \|g\|_q^{r-q} \int_{\mathbf{R}^n} (\tilde{f} * \tilde{g})(x) \, \mathrm{d}x \\ &\le \|f\|_p^{r-p} \|g\|_q^{r-q} \|f\|_p^p \|g\|_q^q = \|f\|_p^r \|g\|_q^r . \end{aligned}$$

In fact, a precise proof should contain also a verification of measurability of the function

$$x \mapsto \int_{\mathbf{R}^n} f(x-y) g(y) \, \mathrm{d}y.$$

This is clear if f and g are bounded and of bounded support, as $f(x-y)g(y)$ is then integrable on $\mathbf{R}^n \times \mathbf{R}^n$, cf. 26.18. The general case can be obtained by approximation $f_k \to f$, $g_k \to g$, where

$$f_k(x) = \begin{cases} f(x) & \text{if } |x| \le k \text{ and } |f(x)| \le k, \\ 0 & \text{otherwise} \end{cases}$$

and g_k's are defined analogously. ∎

26.21. Theorem. *If* $f \in \mathscr{L}^p$ $(1 \le p \le \infty)$ *and if* μ *is a finite signed Radon measure on* $\mathbf{R}^n$, *then* $f * \mu \in \mathscr{L}^p$ *and*

$$\|f * \mu\|_p \le \|f\|_p \; |\mu| (\mathbf{R}^n).$$

Proof. (a) Let $1 < p < +\infty$. The Hölder inequality gives (for a fixed x)

$$\left|\int_{\mathbf{R}^n} f(x-y)\,\mathrm{d}\mu(y)\right| \le \left(\int_{\mathbf{R}^n} |f(x-y)|^p\,\mathrm{d}\,|\mu|\,(y)\right)^{\frac{1}{p}} (|\mu|\,(\mathbf{R}^n))^{\frac{p-1}{p}},$$

whence (by Fubini's theorem)

$$\begin{aligned}\int_{\mathbf{R}^n} |f*\mu|^p\,\mathrm{d}x &\le (|\mu|\,(\mathbf{R}^n))^{p-1}\int_{\mathbf{R}^n}\left(\int_{\mathbf{R}^n}|f(x-y)|^p\,\mathrm{d}x\right)\mathrm{d}\,|\mu|\,(y)\\ &= |\mu|\,(\mathbf{R}^n)^p\,\|f\|_p^p.\end{aligned}$$

(b) If $p = 1$, then Fubini's theorem yields immediately

$$\int_{\mathbf{R}^n} |f*\mu|\,\mathrm{d}x \le \int_{\mathbf{R}^n}\left(\int_{\mathbf{R}^n}|f(x-y)|\,\mathrm{d}x\right)\mathrm{d}\,|\mu|\,(y) = |\mu|\,(\mathbf{R}^n)\,\|f\|_1.$$

(c) If $p = +\infty$, then

$$|f*\mu(x)| \le \int_{\mathbf{R}^n} |f(x-y)|\,\mathrm{d}\,|\mu|\,(y) \le \|f\|_\infty\,|\mu|\,(\mathbf{R}^n)$$

for almost all $x \in \mathbf{R}^n$. ∎

26.22. Remark. Suppose $1 \le p \le \infty$, $\frac{1}{p}+\frac{1}{q}=1$, $f \in \mathscr{L}^p$, $g \in \mathscr{L}^q$. Then $f*g \in \mathscr{L}^\infty$ by Young's convolution theorem 26.20 and $\|f*g\|_\infty \le \|f\|_p\,\|g\|_q$. Even stronger assertion holds: The convolution $f*g$ is uniformly continuous on $\mathbf{R}^n$, and if $1 < p < \infty$, then $f*g \in \mathscr{C}_0(\mathbf{R}^n)$. These properties can be proved using methods of Chapter 31 (see Exercise 31.9).

26.23. Exercise. Suppose $f, g, h \in \mathscr{L}^1$. Show that $(f*g)*h = f*(g*h)$ almost everywhere.

26.24. Exercise. Now we indicate another way how to define the *Jordan-Peano volume* on Euclidean spaces (compare with Exercise 5.12). Let m be a positive integer and $i = [i_1,\dots,i_n] \in \mathbb{Z}^n$ be a "multiindex". Denote by $Q_{m,i}$ the closed cube $[(i_1-1)2^{-m}, i_1 2^{-m}] \times \cdots \times [(i_n-1)2^{-m}, i_n 2^{-m}]$. If m is fixed and i runs over $\mathbb{Z}^n$, the cubes $Q_{m,i}$ are non-overlapping (i.e. their interiors are pairwise disjoint) and cover $\mathbf{R}^n$. If $E \subset \mathbf{R}^n$, let $\mathscr{J}^*(E,m)$ denote 2^{-mn} times the number of elements of the set $\{i\colon Q_{m,i}\cap E \ne \emptyset\}$ and $\mathscr{J}_*(E,m)$ denote 2^{-mn} times the number of elements of the set $\{i\colon Q_{m,i} \subset E\}$,

$$\mathscr{J}^*E = \inf_m \mathscr{J}^*(E,m) \quad\text{and}\quad \mathscr{J}_*E = \sup_m \mathscr{J}_*(E,m).$$

Prove that:

(a) $\mathscr{J}_*E \le \lambda_*E \le \lambda^*E \le \mathscr{J}^*E$ for any set $E \subset \mathbf{R}^n$;

(b) $\mathscr{J}_*G = \lambda G$ for any open set G;

(c) $\mathscr{J}^*K = \lambda K$ for every compact set K;

(d) for a bounded set E, $\mathscr{J}^*E = \mathscr{J}_*E$ if and only if $\lambda(\partial E) = 0$.

26.25. Exercise. Let $G \subset \mathbf{R}^n$ be an open set. Show that there exists a disjoint collection of balls in G whose union covers G except on a null Lebesgue set.

Hint. The assertion is an easy consequence of Vitali's covering theorem 27.2. Here we indicate an elementary proof. We can assume that G is of a finite measure. Set $G_0 = G$. By virtue of Exercise 26.24 there are closed cubes $Q_1,\dots,Q_p$ whose interiors are pairwise disjoint and whose union has a measure greater than $\frac{1}{2}\lambda G$. Set $c = 2^{-n}\lambda B(0,1/2)$. Then in the interior of every cube having a measure α there is a closed ball of measure $c\alpha$. Hence we get closed pairwise disjoint balls $B_j \subset Q_j$ whose union F_1 has a measure greater than $\frac{c}{2}\lambda G$. Set $G_1 = G_0 \setminus F_1$. Inductively, there are $F_{i+1} \subset G_i := G_{i-1}\setminus F_i$ which are unions of pairwise disjoint closed balls with $\lambda G_i \le (1-\frac{c}{2})\lambda G_{i-1} \le \cdots \le (1-\frac{c}{2})^i\lambda G$.

26.26. Exercise. Let $E \subset \mathbf{R}^n$ be an arbitrary set. Prove that

$$\lambda^* E = \inf\{\sum_j \lambda U(x_j, r_j) \colon \bigcup_j U(x_j, r_j) \supset E\}.$$

Hint. It may be supposed that $\lambda^* E < +\infty$. Choose an $\varepsilon > 0$ and find an open set $G \supset E$ with $\lambda G < \lambda^* E + \varepsilon$. Thanks to Exercise 26.25 there exists a countable disjoint collection $\{V_j\}$ of open balls so that $\{V_j\}$ covers G except on a null set N and

$$\sum_j \lambda V_j = \lambda \bigcup_j V_j \le \lambda G + \varepsilon\,.$$

Find a countable family $\{I_j\}$ of open intervals such that $\bigcup_j I_j \supset N$ and $\sum_j \lambda I_j \le \varepsilon$. According to Exercise 26.24 there exist closed cubes Q_j^m with $\bigcup_m Q_j^m \supset I_j$ and $\sum_m \lambda Q_j^m < 2\lambda I_j$. For each cube Q_j^m find an open ball W_j^m with the same centre and with radius equal to the diameter of this cube (so that W_j^m contains Q_j^m). It is not hard to check that there exists a constant c (depending only on the dimension n) such that $\lambda W_j^m \le c\lambda Q_j^m$. Whence

$$\sum_{j,m} \lambda W_j^m \le c \sum_{j,m} \lambda Q_j^m \le 2c \sum_j \lambda I_j \le 2c\varepsilon.$$

To finish the proof, it is enough to consider the union of all balls from collections $\{V_j\}$ and $\{W_j^m\}$.

27. Covering Theorems

Covering theorems provide an important tool for deriving deeper results in measure and integration theory. We will present two Vitali's type theorems. The task is to find, for a given cover, a countable or finite disjoint subcover in such way that the difference is a set of small measure.

27.1. Vitali Cover. We say that a collection $\mathscr{V}$ of closed balls is a *Vitali cover* of a set $A \subset \mathbf{R}^n$ (or that $\mathscr{V}$ has the *Vitali property*) if for each $x \in A$ and each $r > 0$ there is $B \in \mathscr{V}$ so that $x \in B \subset B(x, r)$.

First we state the classical theorem of Vitali.

27.2. Vitali's Covering Theorem. *Let $\mathscr{V}$ be a Vitali cover of a set $A \subset \mathbf{R}^n$. Then there exists a pairwise disjoint countable subcollection $\mathscr{A} \subset \mathscr{V}$ such that*

$$\lambda(A \setminus \bigcup \mathscr{A}) = 0.$$

Proof. It is no restriction to assume that $A \subset H \subset \mathbf{R}^n$, where H is an open set having finite measure. Indeed, if we prove the theorem in this case, we can apply it successively to $H = U(0,1)$, $H = U(0,2) \setminus B(0,1)$, $H = U(0,3) \setminus B(0,2), \dots$ and to find a countable pairwise disjoint collection of the given family which covers $A \setminus \{x \in \mathbf{R}^n \colon |s| \in \mathbf{N}\}$ (notice that $\lambda\{x \in \mathbf{R}^n \colon |x| \in \mathbf{N}\} = 0$). It may be also supposed that no finite union of pairwise disjoint balls of $\mathscr{V}$ covers A. Denote $\mathfrak{W} = \{(x, r) \colon B(x, r) \in \mathscr{V}\}$. Now we are going to construct inductively a sequence $\{B_j\}$ of pairwise disjoint balls of $\mathscr{V}$. In the 0^{th} step we have the empty collection of balls. Assume that in the k^{th} step ($k = 1, 2, \dots$) the closed balls $B_1, \dots, B_{k-1} \in \mathscr{V}$ are selected and denote

$$s_k = \sup\Big\{r > 0\colon B(x,r) \in \mathscr{V},\ B(x,r) \cap \big(\bigcup_{j=1}^{k-1} B_j\big) = \emptyset,\ B(x,r) \subset H\Big\}.$$

The Vitali property and the additional assumptions ensure that $s_k > 0$. Further $+\infty > s_1 \geq s_2 \geq \dots$ since H is of finite measure. Now choose a ball $B_k = B(x_k, r_k) \in \mathscr{V}$ with $r_k > s_k/2$ and $B_k \cap B_j = \emptyset$ for $j < k$.

We get a sequence $\mathscr{A} := \{B_k,\ k \in \mathbf{N}\}$ of pairwise disjoint sets from $\mathscr{V}$. We show that $\mathscr{A}$ covers A except on a set of Lebesgue measure zero. To this end let $p \in \mathbf{N}$ and $x \in A \setminus \bigcup \mathscr{A}$ be given. We now invoke the Vitali property and find a ball $B = B(y, s) \in \mathscr{V}$ with $x \in B$ and $B \cap B_j = \emptyset$ for $j = 1, \dots, p-1$. Clearly $s \leq s_p$. Since $\lambda H < \infty$, we have $\sum\limits_k r_k^n < \infty$ and $\lim s_k = \lim r_k = 0$. Therefore we can find $q \geq p$ with $s_{q+1} < s \leq s_q$. By definition of s_{q+1} there exists $i \leq q$ with $B_i \cap B \neq \emptyset$. If z denotes a point of $B_i \cap B$, then

$$|x - x_i| \leq |x - y| + |y - z| + |z - x_i| \leq s + s + r_i \leq 2s_q + r_i \leq 2s_i + r_i \leq 5r_i\,,$$

whence $x \in B(x_i, 5r_i)$. Since $i \geq p$, we get

$$A \setminus \bigcup \mathscr{A} \subset \bigcup_{i=p}^{\infty} B(x_i, 5r_i)\,,$$

and thus

$$\lambda(A \setminus \bigcup \mathscr{A}) \leq \sum_{i=p}^{\infty} 5^n \lambda B(x_i, r_i)\,.$$

Taking the limit as $p \to \infty$ it follows that $\lambda(A \setminus \bigcup \mathscr{A}) = 0$, and we have finished. ■

27.3. Corollary. *Let $A \subset H \subset \mathbf{R}^n$, where H is an open set of finite measure. If a collection $\mathscr{V}$ of closed balls in H is a Vitali cover of A and $\varepsilon > 0$, then there exists a finite collection of pairwise disjoint balls $\mathscr{A}_\varepsilon \subset \mathscr{V}$ such that*

$$\lambda^*(A \setminus \bigcup \mathscr{A}_\varepsilon) < \varepsilon.$$

Proof. Set $\mathscr{A}_\varepsilon = \{B_1, \dots, B_p\}$, where B_j are as in the proof of the previous theorem and p is sufficiently large. ■

Although Theorem 27.2 is sufficiently strong for most of the applications, we state also more powerful covering theorems which will be useful in Chapter 28.

27.4. Besicovitch Theorem. *There exists a positive integer N (depending only on the dimension n of the space $\mathbf{R}^n$) with the following property:*

Whenever $A \subset \mathbf{R}^n$ is a set and Δ is a bounded positive function on A, then there exist a finite family of countable sets $D_1, \dots, D_N \subset A$ such that for each $k \in \{1, 2, \dots, N\}$, the collection $\{B(x, \Delta(x)) \colon x \in D_k\}$ is pairwise disjoint, and

$$A \subset \bigcup \{U(x, \Delta(x)) \colon x \in \bigcup_{k=1}^{N} D_k\}.$$

Proof. Thanks to the compactness of $S =: \{x \in \mathbf{R}^n \colon |x| = 1\}$, there is a finite set $T \subset S$ such that

$$S \subset \bigcup_{t \in T} U(t, \tfrac{1}{40}).$$

Let ξ denote number of elements of T and $N := 2\xi + 1$. First let us assume that A is bounded. A disjoint sequence $\{x_j\}$ of points of A will be chosen inductively as follows: Denote

$$s_1 = \sup\{\Delta(x) \colon x \in A\}$$

and find $x_1 \in A$ with $\Delta(x_1) > \frac{7}{8} s_1$. In the j^{th} step $(j > 1)$ denote

$$M_j = A \setminus \bigcup_{i<j} U(x_i, \Delta(x_i)).$$

If $M_j = \emptyset$, then we stop and set $D = \{x_i \colon i = 1, \dots, j-1\}$. Otherwise, let $s_j = \sup\{\Delta(x) \colon x \in M_j\}$, find $x_j \in M_j$ with $\Delta(x_j) > \frac{7}{8} s_j$ and set $D = \{x_i \colon i \in \mathbf{N}\}$.

We claim that $A \subset \bigcup_j U_j$, where $U_j := U(x_j, \Delta(x_j))$. This is clear if D is finite. In the general case, since A is bounded, $\{x_j\}$ has a Cauchy subsequence $\{x_{j_k}\}$. Observe that $\Delta(x_i) \le |x_i - x_j|$ (indeed, for $i < j$, x_j lies outside the ball $U(x_i, \Delta(x_i))$), and therefore

$$\inf\{\Delta(x_i) \colon i \in \mathbf{N}\} \le \inf\{|x_i - x_j| : i, j \in \mathbf{N},\, i \ne j\} = 0.$$

Given now $x \in A$, there is j with $\Delta(x_j) < \frac{7}{8}\Delta(x)$. It is then apparent that $\Delta(x) > s_j$, so that $x \notin M_j$. Whence $x \in \bigcup_{i<j} U_i$.

Next we define a function p. For every $x_j \in D$, set

$$P_j = \{i \in \{1, \dots, j-1\} \colon B_i \cap B_j \ne \emptyset\},$$

where $B_j := B(x_j, \Delta(x_j)$, and define inductively $p(x_1) = 1$,

$$p(x_j) = \min\{k \in \mathbf{N} \colon k \ne p(x_i) \text{ for all } i \in P_j\}.$$

Set $D_k = \{x_j \colon p(x_j) = k\}$. Plainly the collection of balls $\{B_j \colon x_j \in D_k\}$ is pairwise disjoint for all $k \le N$, and to complete the proof we only have to show that $p(x_j) \le N$ for each $x_j \in D$. We will be done once we prove that, for each $x_j \in D$, number of balls B_i, $i < j$, intersecting B_j is less than N.

We shall suppose that the last assertion does not hold, and derive a contradiction. Select $z := x_j \in D$. If number of balls B_i, $i < j$ intersecting B_j is $\geq N$, then there exist $z_1, z_2, z_3 \in D$ and $t \in S$ so that

$$z_k = x_{j_k} \quad \text{for } j_1 < j_2 < j_3 < i$$

and

$$\left| \frac{z - z_k}{|z - z_k|} - t \right| < \frac{1}{40} \quad \text{for} \quad k = 1, 2, 3\,.$$

For $k = 1, 2, 3$ denote

$$r = \Delta(z),\ r_k = \Delta(z_k),\ R_k = |z - z_k|\,.$$

We will show that r, r_k and R_k satisfy the following system of 16 inequalities

$$\begin{array}{ll}
(\mathrm{I}) & r > 0, \\
(\mathrm{II}_k) & R_k \leq r_k + r, \\
(\mathrm{III}_k) & r < \frac{8}{7} r_k, \\
(\mathrm{IV}_1) & r_3 < \frac{8}{7} r_2, \\
(\mathrm{IV}_2) & r_3 < \frac{8}{7} r_1, \\
(\mathrm{IV}_3) & r_2 < \frac{8}{7} r_1, \\
(\mathrm{V}_k) & r_k \leq R_k, \\
(\mathrm{VI}_1) & r_2 \leq |R_2 - R_3| + \frac{1}{8} r_3, \\
(\mathrm{VI}_2) & r_1 \leq |R_1 - R_3| + \frac{1}{8} r_3, \\
(\mathrm{VI}_3) & r_1 \leq |R_1 - R_2| + \frac{1}{8} r_2,
\end{array}$$

which does not have any solution. Therewith we get the required contradiction.

Now, we will derive the system of inequalities (I) – (VI_3). The inequality (II_k) says that the ball $B(z_k, r_k)$ intersects $B(z, r)$. The inequalities (III), (IV) follow from the following inequalities

$$\Delta(x_j) \leq s_i < \frac{8}{7} \Delta(x_i) \quad \text{for} \quad i < j\,.$$

Since $\Delta(x_i) < |x_j - x_i|$ for $i < j$, we get (V) and (VI). For instance, to obtain (VI_1) we use the definition of N, (II_3) and (III_3)) in order to show that

$$\begin{aligned}
r_2 \leq |z_2 - z_3| &\leq \left| z_2 - z - \frac{|z_3 - z|}{|z_2 - z|} (z_2 - z) \right| + \left| z_3 - z - \frac{|z_3 - z|}{|z_2 - z|} \right| \\
&\leq |R_2 - R_3| + R_3 \left| \frac{z_3 - z}{|z_3 - z|} - \frac{z_2 - z}{|z_2 - z|} \right| \leq |R_2 - R_3| + 2R_3 \frac{1}{40} \\
&\leq |R_2 - R_3| + \frac{1}{20}(r_3 + r) \leq |R_2 - R_3| + \frac{15}{20 \cdot 7} r_3 \leq |R_2 - R_3| + \frac{1}{8} r_3.
\end{aligned}$$

Now we show that the system of inequalities (I)–(VI) does not possess a solution. First, note that r_k and R_k are positive by (I), (III) and (V). By (II_1), (II_2), (V_1), (V_2), (IV_3) and (III_2) we have

$$|R_1 - R_2| \le r + |r_1 - r_2| = r + r_1 + r_2 - 2\min(r_1, r_2) \le r + r_1 + r_2 - \frac{7}{4} r_2 < r_1 + \frac{3}{7} r_2 .$$

On the other hand, (VI_1), (VI_2) and (IV_1) yield

$$|R_1 - R_3| + |R_2 - R_3| \ge r_1 - \frac{1}{8} r_3 + r_2 - \frac{1}{8} r_3 \ge r_1 + \frac{5}{7},$$

whence

$$|R_1 - R_2| < |R_1 - R_3| + |R_3 - R_2| .$$

Similarly we obtain

$$|R_1 - R_3| < |R_1 - R_2| + |R_2 - R_3| ,$$

and

$$|R_2 - R_3| < |R_1 - R_3| + |R_2 - R_1| .$$

We see that the system (I)–(VI_3) has no solution and this contradiction yields the required conclusion.

For the general case when A is unbounded, let $K > 2 \sup_{x\in A} \Delta(x)$, and denote

$$A_i = \{x \in A \colon (i-1)K \le |x| < iK\}.$$

By the above argument we find for each $i \in \mathbf{N}$ corresponding sets $D_k^i \subset A_i$, $k = 1, \dots, N$. Setting

$$D_k = \begin{cases} \bigcup\{D_k^i \colon i \text{ even }\} & \text{for } k \le N, \\ \bigcup\{D_{k-N}^i \colon i \text{ odd }\} & \text{for } N < k \le 2N. \end{cases}$$

we see that number $2N$ obeys required properties. ■

27.5. Lemma. *Let μ be a Radon measure on $\mathbf{R}^n$ and N as in Theorem 27.4. Suppose further that $A \subset H \subset \mathbf{R}^n$, where H is an open set with $\mu H < +\infty$. Given a positive bounded function Δ on A such that $B(x, \Delta(x)) \subset H$ for all $x \in A$, then there exist an open set $U \subset H$ and a finite set $F \subset A$ such that $A \subset U$, the collection of balls $B(x, \Delta(x))_{x\in F}$ is pairwise disjoint and*

$$\mu\left(U \setminus \bigcup_{x\in F} B(x, \Delta(x))\right) < \left(1 - \frac{1}{2N}\right)\mu U.$$

Proof. By the previous theorem we can find countable sets $D_1, \dots, D_N \subset A$ such that for $D := \bigcup_{k=1}^{N} D_k$ and $k \in \{1, 2, \dots, N\}$, the collection of balls

$$\{B(x, \Delta(x)) \colon x \in D_k\}$$

is pairwise disjoint and

$$A \subset \bigcup_{x \in D} B(x, \Delta(x)).$$

Denote

$$E_k = \bigcup_{x \in D_k} B(x, \Delta(x)), \qquad U = \bigcup_{x \in D} U(x, \Delta(x)).$$

We may find and index $q \in \{1, \dots, \}$ with $\mu E_q \geq \frac{1}{N}\mu U$ and a finite set $F \subset D_q$ such that

$$\mu \bigcup_{x \in F} B(x, \Delta(x)) > \frac{1}{2N}\mu U.$$

Now, it is not hard to see that

$$\mu\left(U \setminus \bigcup_{x \in F} B(x, \Delta(x))\right) < \left(1 - \frac{1}{2N}\right)\mu U.$$

■

27.6. Vitali's Covering Theorem for Radon Measures. *Let μ be a Radon measure on $\mathbf{R}^n$ and $A \subset H \subset \mathbf{R}^n$, where H is an open of finite μ-measure. Let further $\mathscr{V}$ be a family of closed balls of $\mathbf{R}^n$ having the following Vitali's type property: Given $x \in A$, there exist $r_j \searrow 0$ such that $B(x, r_j) \in \mathscr{V}$. Then for any $\varepsilon > 0$ there exists an open set $G \subset \mathbf{R}^n$ with $\mu G < \varepsilon$ and a finite family $\mathscr{A}$ of pairwise disjoints balls of $\mathscr{V}$ such that*

$$A \subset G \cup \bigcup \mathscr{A}.$$

Proof. We will construct recursively sequences of sets $\{G_k\}$ and $\{A_k\}$ according to the following rule: Set $G_0 = H$, $A_0 = A$. Suppose that in the k^{th} step the open sets $G_0 \supset \cdots \supset G_{k-1}$ and sets $A_0 \supset \cdots \supset A_{k-1}$ have been chosen. By Lemma 27.5 there is an open set U_k, $A_{k-1} \subset U_k \subset G_{k-1}$ and a closed set $M_k \subset G_{k-1}$ so that M_k is the union of a finite pairwise disjoint collection of balls of $\mathscr{V}$, and

$$\mu(U_k \setminus M_k) < \left(1 - \frac{1}{2N}\right)\mu U_k.$$

Set

$$A_k = A_{k-1} \setminus M_k, \qquad G_k = U_k \setminus M_k.$$

Our construction guarantees that $A \setminus G_k$ is covered by the selected balls (which constitute a finite pairwise disjoint subcollection of $\mathscr{V}$ whose union is $M_1 \cup \cdots \cup M_k$) and

$$\mu G_k < \left(1 - \frac{1}{2N}\right)\mu G_{k-1} \leq \cdots \leq \left(1 - \frac{1}{2N}\right)^k \mu G_0.$$

Taking sufficiently large k, we get the desired conclusion. ■

27.7. Corollary. *Let μ be a complete Radon measure on $\mathbf{R}^n$, $A \subset \mathbf{R}^n$ and $\mathscr{V}$ be a collection of closed balls in $\mathbf{R}^n$. If for any $x \in A$ there exist a sequence $r_j \searrow 0$ with $B(x, r_j) \in \mathscr{V}$, then there exists a countable pairwise disjoint collection $\mathscr{A} \subset \mathscr{V}$ such that*

$$\mu\left(A \setminus \bigcup \mathscr{A}\right) = 0.$$

Proof. The proposition is a straightforward consequence of Theorem 27.6 if A is bounded. For the general case, we use a similar idea as in the proof of Theorem 27.2. Since it can happen that $\mu\{x\colon |x| = r\} \neq 0$, we use Exercise 15.16.b in order to find a sequence $\{r_1, r_2, \dots\}$, $r_j \nearrow +\infty$ for which $\mu\{x\colon |x| = r_j\} = 0$ for every $j \in \mathbf{N}$.

27.8. Remark. Stating Theorem 27.6 and Corollary 27.7 we assumed that centers of given balls lie in the covered set. Hence, Theorem 27.2 and its Corollary 27.3 are not direct consequences of Theorem 27.6 and Corollary 27.7.

27.9. Exercise. Let (P, ρ) be a separable metric space, $A \subset P$ and $\tau > 1$. Suppose there is $\mathfrak{W} \subset P \times (0, 1]$ such that the collection of balls $\{B(x, r)\colon (x, r) \in \mathfrak{W}\}$ covers A. Show that there exists a countable collection $\mathfrak{S} \subset \mathfrak{W}$ such that the balls $B(x, r)$, $(x, r) \in \mathfrak{S}$ are pairwise disjoint and

$$A \subset \bigcup_{(x,r) \in \mathfrak{S}} U(x, (1 + 2\tau)r).$$

Hint. Step by step, find for any $k \in \mathbf{N}$ a countable colection $\mathfrak{S}_k \subset \mathfrak{W}_k$, where

$$\mathfrak{W}_k := \left\{(x, r) \in \mathfrak{W}\colon \tau^{-k} < r \leq \tau^{-k+1}, B(x, r) \cap \bigcup \{B(y, s)\colon (y, s) \in \mathfrak{S}_1 \cup \dots \cup \mathfrak{S}_{k-1}\} = \emptyset\right\}$$

in such a way that the balls $B(x, r)$, $(x, r) \in \mathfrak{S}_k$ are pairwise disjoint and

$$B(x, r) \cap \bigcup \{B(y, x)\colon (y, s) \in \mathfrak{S}_k\} \neq \emptyset$$

for all $(x, r) \in \mathfrak{W}_k$.

(This can be done, for instance, in the following way: Let $\{q_j\}$ be a dense sequence of points of P. We start from $\mathfrak{S}_k^0 = \emptyset$ and go on recursively. Let $j \geq 1$ be a positive integer. If there exists $(x, r) \in \mathfrak{W}_k$ so that $q_j \in B(x, r)$ and $B(x, r) \cap B(y, s) = \emptyset$ for all $(y, s) \in \mathfrak{S}_k^{j-1}$, then set $\mathfrak{S}_k^j = \mathfrak{S}_k^{j-1} \cup \{(x, r)\}$. In the opposite case leave $\mathfrak{S}_k^j = \mathfrak{S}_k^{j-1}$.)

Now set $\mathfrak{S} = \bigcup_{k=1}^{\infty} \mathfrak{S}_k$. The balls $B(x, r)$, $(x, r) \in \mathfrak{S}$ are plainly pairwise disjoint. Take $x \in A$. We are going to prove that

$$x \in \bigcup_{(u,r) \in \mathfrak{S}} U(u, (1 + 2\tau)r).$$

Since

$$A \subset \bigcup_{(u,r) \in \mathfrak{W}} B(u, r),$$

there is $(y, s) \in \mathfrak{W}$ with $x \in B(y, s)$. Further find $k \in \mathbf{N}$ satisfying $\tau^{-k} < s \leq \tau^{-k+1}$. Then either $(y, s) \in \mathfrak{S}_k$ (and there is nothing to prove), or $B(y, s)$ intersects some of the "previous" balls. More precisely, there exists $z \in B(u, r) \cap B(y, s)$, where $(u, r) \in \mathfrak{S}_1 \cup \dots \cup \mathfrak{S}_k$. Obviously $s < \tau r$, and therefore

$$\rho(x, u) \leq \rho(x, y) + \rho(y, z) + \rho(z, u) \leq 2s + r < (2\tau + 1)r.$$

Hence $x \in U(u, (2\tau + 1)r)$.

27.10. Exercise. Let $\mathscr{U}$ be a collection of open balls in $\mathbf{R}^n$, $M = \bigcup \mathscr{U}$, $c < \lambda M$. Prove that there exist disjoint balls $U_1, \ldots, U_k \in \mathscr{U}$ such that

$$\sum_{j=1}^{k} \lambda U_j > \frac{c}{3^n}.$$

Hint. Find a compact set $K \subset M$ and a finite subcollection $\mathscr{U}' \subset \mathscr{U}$ so that $\lambda K > c$ and $\bigcup \mathscr{U}' \supset K$. Select the ball $U_1 \in \mathscr{U}'$ with the greatest radius. Again, select among the balls of $\mathscr{U}'$ which do not intersect U_1 the ball U_2 with the greatest radius. Suppose the balls $U_1, \ldots, U_{j-1}$ have been chosen and find among the balls of $\mathscr{U}'$ which do not intersect $U_1, \ldots, U_{j-1}$ the ball U_j with the greatest radius. After a finite number of steps, we get a sequence $\{U_1, \ldots, U_k\}$ such that we cannot add another ball. If $x \in K$, then $x \in U$ for some $U \in \mathscr{U}'$ and the construction implies that $U \cap U_i \neq \emptyset$ for some i. Let i be the smallest of such indices. Then the radius R of the ball U is not greater than the radius of U_i, so that x lies in the ball with the same center as U_i and radius $3R$. Hence

$$c < \lambda K \leq 3^n \sum_{j=1}^{k} \lambda U_j.$$

27.11. Exercise. Prove Vitali's covering theorem 27.2 using Exercise 27.10 in a similar way as Theorem 27.6 was derived from Lemma 27.5.

27.12. Notes. The famous covering theorem is due to G. Vitali [1908] who proved it for closed intervals and the Lebesgue measure. Later on, Vitali's covering theorem was generalized in various directions by many authors (H. Lebesgue [1910], S. Banach [1924], C. Carathéodory (second edition of the book [*1918] in 1927)). Another step forward was made by A.S. Besicovitch [1945], [1946] and by A.P. Morse [1947]. The origin of the simple covering theorem which appeared in Exercise 27.10 is in Wiener's article [1939].

28. Differentiation of Measures

28.1. Derivative of a Measure. In the following, μ will stand for a Radon measure on $\mathbf{R}^n$ and λ will denote again the Lebesgue measure.

For any $x \in \mathbf{R}^n$, define

$$\overline{D}_\mu(x) := \limsup_{r \to 0+} \frac{\mu B(x,r)}{\lambda B(x,r)} \quad \text{and} \quad \underline{D}_\mu(x) := \liminf_{r \to 0+} \frac{\mu B(x,r)}{\lambda B(x,r)}.$$

If $\overline{D}_\mu(x) = \underline{D}_\mu(x)$, we call this common value the (symmetric) *derivative* of μ (with respect to λ) at x and denoted it by $D_\mu(x)$.

28.2. Remarks. 1. Recall that $\lambda B(x,r) = \dfrac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)} r^n$, see Exercise 26.17.b

2. The functions $x \mapsto \underline{D}_\mu(x)$ and $x \mapsto \overline{D}_\mu(x)$ are Borel-measurable. This assertion follows from the fact that the function $x \mapsto \mu B(x,r)$ is upper semicontinuous, $x \mapsto \lambda B(x,r)$ is continuous and that when defining $\overline{D}_\mu$, we can restrict to $r \in \mathbf{Q}$.

Compare with the following rather surprising theorem (O. Hájek [1957]): *If F is an arbitrary function on an interval $I \subset \mathbf{R}$, then the function $\overline{D}F$ (see 22.3) is Borel on I.* The proof for continuous F is indicated in the proof of Theorem 25.8. An analogous theorem fails for Dini derivatives, there are examples of functions for which all four Dini derivatives are nonmeasurable.

3. Let μ be a Radon measure on $\mathbf{R}$ and F its distribution function (see 24.1). Then $F'(x) = D_\mu(x)$ provided $F'(x)$ exists. Moreover, the following is true: *If $F'(x) \leq a$ for $x \in A$, then $\mu A \leq a\lambda A$. Similarly, $\mu B \geq a\lambda B$ provided $F' \geq a$ on B.* Compare this result with the next lemma.

28.3. Lemma. *Let $A \subset \mathbf{R}^n$ be a Borel set and $a > 0$.*

(a) *If $\underline{D}_\mu(x) \le a$ for all $x \in A$, then $\mu A \le a\lambda A$.*

(b) *If $\overline{D}_\mu(x) \ge a$ for all $x \in A$, then $\mu A \ge a\lambda A$.*

Proof. With the trivial case $\lambda A = \infty$ out of the way, assume $\lambda A < \infty$ and fix an $\varepsilon > 0$. There is an open set $G \supset A$ with $\lambda G \le \lambda A + \varepsilon$. By Vitali's covering theorem 27.6 for Radon measures we can find disjoint closed balls $B(x_i, r_i) \subset G$ such that

$$\mu B(x_i, r_i) \le (a+\varepsilon)\lambda B(x_i, r_i) \quad \text{and} \quad \mu\big(A \setminus \bigcup_i B(x_i, r_i)\big) = 0.$$

(If μ is absolutely continuous with respect to λ, we can also use Vitali's covering theorem 27.2.) Then

$$\mu A \le \sum_i \mu B(x_i, r_i) \le (a+\varepsilon) \sum_i \lambda B(x_i, r_i) \le (a+\varepsilon)\lambda G \le (a+\varepsilon)(\lambda A + \varepsilon)\,,$$

whence (a) follows.

The inequality in part (b) can be proved similarly using Vitali's covering theorem 27.2 for the Lebesgue measure. ∎

28.4. Theorem. *Any Radon measure μ on $\mathbf{R}^n$ has a finite derivative D_μ λ-almost everywhere on $\mathbf{R}^n$.*

Proof. Without loss of generality we can restrict to a compact interval $I \subset \mathbf{R}^n$. For $k \in \mathbf{N}$ and $r, s \in \mathbf{Q}^+$, $s < r$, set

$$A_k = \{x \in I \colon \overline{D}_\mu(x) \ge k\}, \quad A(r,s) = \{x \in I \colon \underline{D}_\mu(x) \le s < r \le \overline{D}_\mu(x)\}.$$

By the previous lemma,

$$k\lambda A_k \le \mu A_k \le \mu I < +\infty \quad \text{and} \quad r\lambda A(r,s) \le \mu A(r,s) \le s\lambda A(r,s) \le s\lambda I < +\infty.$$

Hence $\lambda\big(\bigcap_k A_k\big) = \lim_k \lambda A_k = 0$ and $\lambda A(r,s) = 0$ (realize that $0 \le s < r$). The desired conclusion readily follows observing that

$$0 \le \underline{D}_\mu(x) \le \overline{D}_\mu \le +\infty, \quad \{x \in I \colon \overline{D}_\mu(x) = +\infty\} = \bigcap_k A_k$$

and

$$\{x \in I \colon \underline{D}_\mu(x) < \overline{D}_\mu(x)\} = \bigcup_{r,s \in \mathbf{Q}^+} A(r,s).$$

∎

28.5. Remark. Consider now the simple case when F is the distribution function of a Radon measure μ on $\mathbf{R}$. We know that F has a (finite) derivative F' almost everywhere (Lebesgue's theorem 22.5), that $\int_a^b F' \,\mathrm{d}\lambda \le F(b) - F(a)$ (Theorem 22.7), and that $\int_a^b F' \,\mathrm{d}\lambda = F(b) - F(a)$ for any interval $[a,b]$ if and only if $\mu \ll \lambda$ (Corollary 23.5 and Exercise 24.7). In the last case, F' agrees almost everywhere with the Radon-Nikodým derivative $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}$. An analogous assertion holds for differentiation of measures and it is contained in the next Theorem 28.6.

Proof of the following theorem is based on Lemma 28.3. A proof based on the covering theorem of Exercise 27.10 can be found in W. Rudin [*1974].

28.6. Theorem. *Let μ be a Radon measure on $\mathbf{R}^n$ and $B \subset \mathbf{R}^n$ a Borel set. Then:*

(a) $\int_B D_\mu \, \mathrm{d}\lambda \le \mu B$ *(in particular, D_μ is locally λ-integrable);*

(b) $\int_B D_\mu \, \mathrm{d}\lambda = \mu B$*, provided $\mu \ll \lambda$.*

Proof. Suppose $\beta > 1$, $k \in \mathbb{Z}$ and set

$$B_k(\beta) = \{x \in B \colon \beta^k \le D_\mu(x) < \beta^{k+1}\}, \quad M = \bigcup_k B_k(\beta).$$

According to Theorem 28.4, $\lambda(B \setminus M) = 0$. If $\mu \ll \lambda$, then also $\mu(B \setminus M) = 0$. Using Lemma 28.3.b, we get

$$\begin{aligned}\int_B D_\mu \, \mathrm{d}\lambda = \int_M D_\mu \, \mathrm{d}\lambda = \sum_{k=-\infty}^{+\infty} \int_{B_k(\beta)} D_\mu \, \mathrm{d}\lambda \le \sum_{k=-\infty}^{+\infty} \beta^{k+1} \lambda B_k(\beta) \\ \le \beta \sum_{k=-\infty}^{+\infty} \mu B_k(\beta) \le \beta \mu B \,,\end{aligned}$$

and by part (a) of Lemma 28.3,

$$\int_B D_\mu \, \mathrm{d}\lambda \ge \sum_{k=-\infty}^{+\infty} \int_{B_k(\beta)} D_\mu \, \mathrm{d}\lambda \ge \sum_{k=-\infty}^{+\infty} \mu B_k(\beta) = \beta^{-1} \mu M = \beta^{-1} \mu B.$$

When taking the limit as $\beta \to 1+$, we get both (a) and (b). ■

28.7. Theorem. *The following statements about a Radon measure μ on $\mathbf{R}^n$ are equivalent:*

(i) $\mu \ll \lambda$*;*

(ii) $\mu B = \int_B D_\mu \, \mathrm{d}\lambda$ *for every Borel set $B \subset \mathbf{R}^n$;*

(iii) D_μ *is the Radon-Nikodým derivative of μ with respect to λ;*

(iv) $\underline{D}_\mu < +\infty$ *μ-almost everywhere on $\mathbf{R}^n$.*

Proof. The previous theorem says that (i) $\Longrightarrow$ (ii), thus the equivalence (i) $\Longleftrightarrow$ (ii) $\Longleftrightarrow$ (iii) is obvious. To show that (i) implies (iv), use Theorem 28.6. Assume now (iv) and let $A \subset \mathbf{R}^n$, $\lambda A = 0$. Using Lemma 28.3.a, we get

$$\mu\{x \in A \colon \underline{D}_\mu(x) \le k\} \le k \, \lambda A = 0$$

for every $k \in \mathbf{N}$, whence plainly $\mu A = 0$. ■

28.8. Remarks. 1. It is not difficult to state similar results for signed Radon measures.

2. If μ is a Radon measure on $\mathbf{R}^n$ and $\mu = \mu_s + \mu_a$ is its Lebesgue decomposition into the singular and absolutely continuous part, then D_μ is the Radon-Nikodým derivative of μ_a with respect to λ.

28.9. Exercise. Let $G \subset \mathbf{R}^n$ be an open set, $f \colon G \to \mathbf{R}^n$ be a diffeomorphism. If B is a Borel subset of G, set $\mu B := \lambda f(B)$. Show that μ is a Radon measure on G and

(a) $D_\mu(x) = |J_f(x)|$ for every $x \in G$;

(b) $|J_f|$ is the Radon-Nikodým derivative $\frac{\mathrm{d}\mu}{\mathrm{d}\lambda}$.

28.10. Notes. We could define the (symmetric) derivative of a measure μ at x using cubes with the center at x whose lenght of edges converges to zero.

Furthermore, we could consider more general sequences of sets "shrinking" in a suitable way to x. We could consider, for instance, all sequences of intervals containing x whose diameters converge to zero. Let us note that for these general "derivatives of measures" analogous theorems to those of this paragraph fail to hold, and even analogies of Theorem 29.2 (which is their direct consequence) do not hold. An example can be found in M. de Guzmán [*1975], Chapter V., §2.

On the topic of differentiation of measures the reader is invite to consult for instance, W. Rudin [*1974], M. de Guzmán [*1975] or J. Lukeš, J. Malý and L. Zajíček [*1986].

29. Lebesgue Density Theorem and Approximately Continuous Functions

29.1. Density points. Let λ be again the Lebesgue measure on $\mathbf{R}^n$, $M \subset \mathbf{R}^n$ a measurable set and $x \in \mathbf{R}^n$. We say that x is a *point of density*, or a *density point* of M if

$$\lim_{r \to 0+} \frac{\lambda(M \cap U(x,r))}{\lambda U(x,r)} = 1 .$$

29.2. Lebesgue Density Theorem. *Almost every point of a measurable set M is its density point.*

Proof. If $\mu A := \lambda(A \cap M)$ for every (Lebesgue) measurable set $A \subset \mathbf{R}^n$, then μ is a Radon measure on $\mathfrak{M}_n$ absolutely continuous with respect to λ. Since $\mu A = \int_A c_M \, \mathrm{d}\lambda$ for $A \in \mathfrak{M}_n$, c_M is the Radon-Nikodým derivative of μ with respect to λ. On the other hand, Theorem 28.7 tell us that D_μ is also the Radon-Nikodým derivative of μ with respect to λ. Hence $D_\mu = c_M$ almost everywhere according to the Radon-Nikodým Theorem (Remark 13.6.1) . If we realize that $D_\mu(x) = \lim\limits_{r \to 0+} \dfrac{\mu U(x,r)}{\lambda U(x,r)}$ and $c_M(x) = 1$ for $x \in M$, we obtain the assertion. ■

29.3. Remark. When $n = 1$, the Lebesgue density theorem is an easy consequence of Theorem 23.4 according to which the derivative of an indefinite Lebesgue integral of a (locally) integrable function c_M equals c_M almost everywhere.

29.4. Density Topology. A measurable set $M \subset \mathbf{R}^n$ is called *d-open* if each point of M is its point of density. For instance, the set of all irrational numbers, or any open subset of $\mathbf{R}^n$ are d-open. We are going to show that the collection d of all d-open sets on $\mathbf{R}^n$ forms a topology which will be labelled as the *density topology*. Since d contains all open subsets of $\mathbf{R}^n$, it is finer than the Euclidean topology of $\mathbf{R}^n$.

29.5. Theorem. *The collection of all d-open sets forms a topology.*

Proof. Plainly $\emptyset$ and $\mathbf{R}^n$ are d-open, and d is closed under the formation of finite intersections. If $\mathscr{A}$ is a collection of d-open sets, we have to prove that $T := \bigcup \mathscr{A} \in d$. If we show that T is measurable, then every point of T is apparently a density point of T. We can assume that $T \subset I$, where $I \subset \mathbf{R}^n$ is a compact interval. Denoting $\mathscr{S} = \{S\colon$ there exists a countable collection $\mathscr{A}_0 \subset \mathscr{A}$ such that $S = \bigcup \mathscr{A}_0\}$, there exists $S \in \mathscr{S}$ with $\lambda S = \sup\{\lambda M\colon M \in \mathscr{S}\}$. Given $x \in T$, there exists $A \in \mathscr{A}$ so that x is a point of density of A. Since $\lambda(A \cup S) = \lambda S$, we get

$\lambda(A \setminus S) = 0$. Hence x is a point of density of $A \cap S$ and also a point of density of S. Of course, x cannot be a point of density of $I \setminus S$. Since by the Lebesgue density theorem almost every point of $I \setminus S$ is a point of density of $I \setminus S$, it follows that $\lambda(T \setminus S) = 0$. Hence $T = S \cup (T \setminus S)$ is measurable. ■

29.6. Remark. The density topology in $\mathbf{R}^n$ shares a lot of interesting properties. Let us note that the density topology is not metrizable, it is not normal (it is completely regular), the only d-compact sets are the finite ones, and that the Baire category theorem holds for d.

29.7. Approximately Continuous Functions. A function f defined on a neighbourhood of a point $z \in \mathbf{R}^n$ is said to be *approximately continuous* at z if there exists a measurable set $M \subset \mathbf{R}^n$ such that z is a density point of M and $\lim_{x \to z,\, x \in M} f(x) = f(z)$. If f is approximately continuous at each point of a given set, we say that f is approximately continuous. It is clear that each continuous function is approximately continuous.

29.8. Theorem. *A function f is approximately continuous at z if and only if f is d-continuous at z.*

Proof. Suppose f is d-continuous at z and defined on a neighbourhood $U = U(z, r_0)$ of z. Then z is a point of density of each of the sets $M_j := \{x \in U\colon |f(x) - f(z)| < \frac{1}{j}\}$. Find a decreasing sequence of radii $r_k > 0$ satisfying

$$\frac{\lambda(U(z,r) \setminus M_j)}{\lambda U(z,r)} < 2^{-j-k}$$

for all $j = 1, \dots, k$ and $r \in (0, r_k)$. Set $A_j = U(z, r_j) \setminus M_j$ and $M = U \setminus \bigcup_{j=1}^{\infty} A_j$. We have

$$\frac{\lambda(U(z,r) \cap A_j)}{\lambda U(z,r)} \le 2^{-j}$$

for all $r > 0$. Choose $k \in \mathbf{N}$ and $r \in (0, r_k)$. Then

$$\frac{\lambda(U(z,r) \cap A_j)}{\lambda U(z,r)} \le \begin{cases} 2^{-j-k} & \text{for } j \le k\,, \\ 2^{-j} & \text{for } j > k\,. \end{cases}$$

Whence

$$\frac{\lambda(U(z,r) \setminus M)}{\lambda U(z,r)} \le 2^{-k+1}\,,$$

and z is a density point of M. If $x \in M \cap U(z, r_k)$, then $x \in M_k$ and consequently $|f(x) - f(z)| < \frac{1}{k}$. Therefore $\lim_{x \to z, x \in M} f(x) = f(z)$. The reverse implication is obvious. ■

29.9. Denjoy's Theorem. *A function $f\colon \mathbf{R}^n \to \mathbf{R}$ is Lebesgue measurable if and only if f is approximately continuous at almost all points of $\mathbf{R}^n$.*

Proof. Let f be approximately continuous at almost all points. Denote by N the set of all points of approximate discontinuity of f. Take any $c \in \mathbf{R}$. We will show that the set $M := \{x \in \mathbf{R}\colon f(x) > c\}$ is measurable. Since M is a

d-neighbourhood of any point $x \in M \setminus N$, it follows that $M \setminus N$ is also a d-neighbourhood of x. Hence $M \setminus N$ is d-open, and thus measurable. Since N is a null set, M is measurable as well.

Now assume that f is measurable and select an $\varepsilon > 0$. Luzin's theorem 18.2 provides us with a continuous function g on $\mathbf{R}^n$ and an open set G so that $\mu G < \varepsilon$ and $f = g$ in $\mathbf{R}^n \setminus G$. By the Lebesgue density theorem almost every point of the set $\mathbf{R}^n \setminus G$ is its point of density. Thus f is approximately continuous at almost all points of $\mathbf{R}^n \setminus G$ and we can easily conclude that f is approximately continuous at almost all points of $\mathbf{R}^n$. ∎

29.10. Remark. Compare the last equivalence with the Lebesgue theorem 7.9 according to which a bounded function f is Riemann integrable if and only if f is continuous almost everywhere.

29.11. Exercise. Let f be a bounded function defined on a neighbourhood of a point $z \in \mathbf{R}^n$. Then f is approximately continuous at z if and only if z is a Lebesgue point for f (the definition of Lebesgue points in $\mathbf{R}^n$ is analogous to the one-dimensional case in 23.8). Show that the assumption of boundedness is essential for one of the implications.

29.12. Notes. The notion of approximate continuity was introduced by A. Denjoy [1915]; he proved also that every Lebesgue measurable function is approximately continuous at almost all points. The converse assertion was proved by V. Stepanov in [1924]. The density theorem 29.2 is due to H. Lebesgue [*1904]. This theorem also holds for nonmeasurable sets if the outer Lebesgue measure is used in the definition of density points. The notion of the density topology was studied much later in the 1950's. Many of its properties and generalizations and recent applications can be found in the monograph by J. Lukeš, J. Malý and L. Zajíček [*1986].

30. Lipschitz Functions

30.1. Lipschitz Mappings. Let (P_1, ρ_1), (P_2, ρ_2) be metric spaces and $\beta > 0$. Recall that a mapping $f\colon P_1 \to P_2$ is said to be *β-Lipschitz* if

$$\rho_2(f(x), f(y)) \leq \beta \rho_1(x, y)$$

for every $x, y \in P_1$. We say that f is *Lipschitz* on P_1 if there exists $\beta > 0$ for which f is β-Lipschitz. A mapping $f\colon P_1 \to P_2$ is called *locally Lipschitz* if for any $z \in P_1$ there exists its neighbourhood U such that f is Lipschitz on U (in this case, the constant β can vary from point to point).

30.2. Lemma. *Let f be a continuous function on an open set $G \subset \mathbf{R}^n$ and $i \in \{1, \ldots, n\}$. Then the set of those points at which $\frac{\partial f}{\partial x_i}$ fails to exist is a Borel set.*

Proof. To simplify the proof, we can assume that $G = \mathbf{R}^n$. Denote

$$u_{m,k}(x) = \sup \left\{ \frac{f(x + te_i) - f(x)}{t} : t \in (-\frac{1}{m}, \frac{1}{m}), |t| \geq \frac{1}{k} \right\},$$

$$v_{m,k}(x) = \inf \left\{ \frac{f(x + te_i) - f(x)}{t} : t \in (-\frac{1}{m}, \frac{1}{m}), |t| \geq \frac{1}{k} \right\},$$

$$u(x) = \inf_m \sup_{k>m} u_{m,k}(x) \quad , \quad v(x) = \sup_m \inf_{k>m} v_{m,k}(x).$$

Then, for $k > m$, $u_{m,k}$ and $v_{m,k}$ are continuous. Since the set of points where $\frac{\partial f}{\partial x_i}(x)$ exists equals $\{x \in \mathbf{R}^n : -\infty < u(x) = v(x) < +\infty\}$, the assertion follows. ∎

30.3. Rademacher's Theorem. *Let f be a Lipschitz function on an open set $G \subset \mathbf{R}^n$. Then f is differentiable almost everywhere in G.*

Proof. Assume that f is a β-Lipschitz function. Let E be the set of all points where f fails to have some of the partial derivatives. Using Fubini's theorem, the one-dimensional theorem on differentiability of Lipschitz functions (Lemma 22.4) easily implies that E is of measure zero (notice that E is measurable by Lemma 30.2).

Now, for $p, q \in \mathbf{Q}^n$ and $m \in \mathbf{N}$ denote

$$S_{p,q,m} = \Big\{ x \in G \setminus E \colon p_i < \frac{f(x+te_i) - f(x)}{t} < q_i \text{ for all } i = 1, \dots, n \\ \text{and for } t \in \Big(-\frac{1}{m}, \frac{1}{m} \Big) \setminus \{0\} \Big\}.$$

Let $\tilde{S}_{p,q,m}$ be the set of all density points of $S_{p,q,m}$. By the Lebesgue density theorem 29.2 $\lambda(S_{p,q,m} \setminus \tilde{S}_{p,q,m}) = 0$. If

$$N := \bigcup_{p,q,m} (S_{p,q,m} \setminus \tilde{S}_{p,q,m}),$$

then N is of measure zero. We claim that f is differentiable at each point of $G \setminus (N \cup E)$.

To this end, take $x \in G \setminus (N \cup E)$ and $\varepsilon \in (0,1)$. Pick $p, q \in \mathbf{Q}^n$ so that

$$q_i - \varepsilon < p_i < \frac{\partial f}{\partial x_i}(x) < q_i, \qquad i = 1, \dots, n.$$

Then there exists $m \in \mathbf{N}$ with $x \in S := S_{p,q,m}$. Since x is outside of N, x is even a density point of S. Find $\delta \in (0, \frac{1}{m})$ such that $\lambda(U(x,r) \setminus S) \le \big(\frac{\varepsilon}{2}\big)^n \lambda(U(x,r))$ for all $r \in (0, 2\delta)$. In particular, notice that $U(x, (1+\varepsilon)\tau) \setminus S$ does not contain any ball of radius $\varepsilon\tau$ if $\tau \in (0, \delta)$. Choose $y \in U(x, \delta)$ and denote $y^i = (y_1, \dots, y_i, x_{i+1}, \dots, x_n)$. For any $i \in \{1, \dots, n\}$, let U_i be the ball with center y^i and radius $\varepsilon |y - x|$. The choice $\tau = |y - x|$ yields that $S \cap U_i \ne \emptyset$ for all i. If $z^i \in S \cap U_i$ and $w^i := z^{i-1} + (y_i - x_i)e_i$, then

$$\big|w^i - y^i\big| = \big|z^{i-1} - y^{i-1}\big| \le \varepsilon |y - x|,$$
$$p_i < \frac{f(w^i) - f(z^{i-1})}{y_i - x_i} < q_i \quad \text{and} \quad p_i < \frac{\partial f}{\partial x_i}(x) < q_i.$$

Whence

$$\left| f(w^i) - f(z^{i-1}) - \frac{\partial f}{\partial x_i}(x)(y_i - x_i) \right| \le (q_i - p_i)|y_i - x_i| \le \varepsilon |y - x|.$$

Summarizing, we get

$$\begin{aligned}
&\left|f(y) - f(x) - \sum_i \frac{\partial f}{\partial x_i}(x)(y_i - x_i)\right| \\
&\quad \le \left|\sum_i \left(f(w^i) - f(z^{i-1}) - \frac{\partial f}{\partial x_i}(x)(y_i - x_i)\right)\right| \\
&\qquad + \sum_i \left(\left|f(w^i) - f(y^i)\right| + \left|f(z^{i-1}) - f(y^{i-1})\right|\right) \\
&\quad \le \varepsilon(n + 2\beta n)\,|y - x|\,.
\end{aligned}$$

Thus $f'(x)$ does exist. ■

30.4. Lemma. *Let $(f_\alpha)_\alpha$ be a family of β-Lipschitz functions on $\mathbf{R}^n$. Then the function $\sup_\alpha f_\alpha$ is β-Lipschitz provided it is finite at least at one point.*

Proof. It is quite obvious. ■

30.5. McShane's Extension Theorem. *Let f be a β-Lipschitz function on a set $E \subset \mathbf{R}^n$. Then there exists a β-Lipschitz extension f^* of f to all of $\mathbf{R}^n$. If, in addition, E is bounded, then we can find a β-Lipschitz extension f^{**} of f with compact support.*

Proof. Set

$$f^*(x) = \sup_{y \in E}\{f(y) - \beta\,|y - x|\}.$$

By the previous lemma, f^* is a β-Lipschitz function and $f^* = f$ on E. If E is bounded, then there exists an $\alpha > 0$ such that $|f(x)| < \alpha - \beta\,|x|$ for all $x \in E$ and we set

$$f^{**}(x) = \begin{cases} \min(f^*(x), \alpha - \beta\,|x|) & \text{if } |x| < \alpha/\beta \text{ and } f^*(x) \ge 0; \\ \max(f^*(x), -\alpha + \beta\,|x|) & \text{if } |x| < \alpha/\beta \text{ and } f^*(x) < 0; \\ 0 & \text{if } |x| \ge \alpha/\beta. \end{cases}$$

■

30.6. Remark. Let f be a β-Lipschitz mapping from $E \subset \mathbf{R}^n$ into $\mathbf{R}^k$. Then the coordinates of f can be extended separately by the previous theorem and we get a $k\beta$-Lipschitz extension $f^*\colon \mathbf{R}^n \to \mathbf{R}^k$ of f. However, this extension fails to be β-Lipschitz. There exists a stronger extension theorem for mappings due to Kirszbraun which guarantees the existence of a β-Lipschitz extension. The proof of this assertion is more difficult.

30.7. Exercise. Suppose $E \subset \mathbf{R}^n$ and $f\colon E \to \mathbf{R}^n$ is a β-Lipschitz mapping. Prove that $\lambda^* f(E) \le \beta^n \lambda^* E$.

Hint. We can assume that $\lambda^* E < +\infty$. Fix an $\varepsilon > 0$. By Lemma 26.26 find a sequence $\{U(x_j, r_j)\}$ of open balls so that $E \subset \bigcup U(x_j, r_j)$ and $\sum_j \lambda U(x_j, r_j) < \lambda^* E + \varepsilon$. Since $f(U(x_j, r_j)) \subset U(f(x_j), \beta r_j)$ for every j,

$$\lambda^*(f(E)) \le \sum_j \lambda U(f(x_j), \beta r_j) = \beta^k \sum_j \lambda U(x_j, r_j) \le \beta^k(\lambda^* E + \varepsilon).$$

30.8. Notes. Theorem 30.3 is due to H. Rademacher [1919]. An elementary proof of it (different from ours) was given by A. Nekvinda and L. Zajíček [1984].

Extensions of Lipschitz functions in metric space were closely related to original proofs of the Hahn-Banach theorem. Theorem 30.5 of extension of (nonlinear) Lipschitz functions was proved independently by M.D. Kirszbraun [1934] and by E.J. McShane [1934]. Another material can be found in G.J. Minty [1970].

31. Approximation Theorems

31.1. Space $\mathscr{D}(\Omega)$. Let $\Omega \subset \mathbf{R}^n$ be an open set. We denote by $\mathscr{D}(\Omega)$ the linear space of all infinitely differentiable functions on Ω with compact support in Ω. Our aim is to prove that $\mathscr{D}(\Omega)$ is often dense in other function spaces. The first task is to construct a nontrivial infinitely smooth function on $\mathbf{R}^n$ with compact support.

Set

$$\chi_1(x) = \begin{cases} \alpha e^{1/(|x|^2-1)} & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1, \end{cases}$$

where the constant α is chosen in such a way that $\int_{\mathbf{R}^n} \chi_1(x)\,\mathrm{d}x = 1$. Denote

$$\chi_k(x) = k^n \chi_1(kx).$$

Then $\int_{\mathbf{R}^n} \chi_k(x \pm y)\,\mathrm{d}y = 1$ for any $k \in \mathbf{N}$ and $x \in \mathbf{R}^n$. A simple calculation shows that the functions χ_k are infinitely differentiable (first reduce the proof to a problem of differentiability of functions of one variable).

Now, if f is a locally integrable function, then Theorem 9.2 gives immediately that $\chi_k * f$ are infinitely differentiable functions as well. Likewise, the convolutions $\chi_k * \mu$ are infinitely differentiable if μ is a Radon measure.

31.2. Lemma. *Suppose $p \in [1, \infty)$, $f \in \mathscr{L}^p(\mathbf{R}^n)$ and $k \in \mathbf{N}$. Then $\chi_k * f \in \mathscr{L}^p(\mathbf{R}^n)$ and $\|\chi_k * f\|_p \leq \|f\|_p$.*

Proof. According to generalized Young's inequality 26.20,

$$\|\chi_k * f\|_p \leq \|f\|_p \cdot \|\chi_k\|_1 \leq \|f\|_p$$

(notice that $\|\chi_k\|_1 = 1$). ■

31.3. Theorem. *Suppose $p \in [1, \infty)$. For any $f \in \mathscr{L}^p(\mathbf{R}^n)$, we have $\|f - \chi_k * f\|_p \to 0$.*

Proof. Denote $E_j = \{x \in \mathbf{R}^n\colon |x| \leq j \text{ and } |f(x)| \leq j\}$ and $F_j = f c_{E_j}$. Since $\bigcap_j (\mathbf{R}^n \setminus E_j)$ has measure zero, $f_j - f \to 0$ almost everywhere and by the Lebesgue theorem (dominating function $|f|^p$), $\|f - f_j\|_p \to 0$. Fix an $\varepsilon > 0$ and find $j \in \mathbf{N}$ with $\|f - f_j\|_p < \varepsilon$. Whence by Lemma 31.2

$$\|\chi_k * f_j - \chi_k * f\|_p = \|\chi_k * (f_j - f)\|_p < \varepsilon$$

for all $k \in \mathbf{N}$. Now $\chi_k * f_j \to f_j$ (as $k \to \infty$) almost everywhere. Indeed, it is routine to verify that $\chi_k * f_j \to f_j$ at all points of approximate continuity of

f_j and this set, according to Denjoy's theorem 29.9, is of full measure. By the Lebesgue convergence theorem with dominating function $(2j)^p c_{B(0,j+1)}$ we have

$$\lim_{k\to 0} \|f_j - \chi_k * f_j\|_p = 0.$$

Hence,

$$\|f - \chi_k * f\|_p \le \|f - f_j\|_p + \|f_j - \chi_k * f_j\|_p + \|\chi_k * f_j - \chi_k * f\|_p < 3\varepsilon\,,$$

provided k is large enough. ■

31.4. Theorem. *Let $\Omega \subset \mathbf{R}^n$ be an open set and $p \in [1,\infty)$. Then $\mathscr{D}(\Omega)$ is a dense subset of $\mathscr{L}^p(\Omega)$.*

Proof. Set $\Omega_j = \{x \in \Omega\colon |x| < j \text{ and } \operatorname{dist}(x,\partial\Omega) > \frac{1}{j}\}$. Take a function $f \in \mathscr{L}^p(\Omega)$ and denote $g_j = f c_{\Omega_j}$. As in the previous proof, it is clear that $\|f - g_j\|_p \to 0$. For all $k > j$ we get $\chi_k * g_j \in \mathscr{D}(\Omega)$. Now the estimate

$$\|f - \chi_k * g_j\|_p \le \|f - g_j\|_p + \|g_j - \chi_k * g_j\|_p$$

yields the assertion. ■

31.5. Theorem. *Let f be a continuous function with compact support contained in an open set $\Omega \subset \mathbf{R}^n$. Then $\chi_k * f \rightrightarrows f$ on Ω.*

Proof. We can assume that $\Omega = \mathbf{R}^n$. Since f is uniformly continuous, given $\varepsilon > 0$ there is $k_0 \in \mathbf{N}$ such that $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in \mathbf{R}^n$, $|x - y| < \frac{1}{k_0}$. If $k > k_0$ and $x \in \mathbf{R}^n$, then

$$\begin{aligned} |f(x) - \chi_k * f(x)| &= \left| \int_{\mathbf{R}^n} (f(x) - f(y))\chi_k(x-y)\,\mathrm{d}y \right| \\ &= \left| \int_{B(x,1/k)} (f(x) - f(y))\chi_k(x-y)\,\mathrm{d}y \right| \\ &\le \varepsilon \int_{\mathbf{R}^n} \chi_k(x-y)\,\mathrm{d}y = \varepsilon\,. \end{aligned}$$

■

31.6. Theorem. *Let μ be a Radon measure on Ω. Then*

(a) *there exist sequences $\{\mu_j\}$ of measures and $\{f_j\}$ of functions from $\mathscr{D}(\mathbf{R}^n)$ such that $f_j = \mathrm{d}\mu_j/\,\mathrm{d}\lambda$ and $\mu_j \xrightarrow{\text{w}} \mu$;*

(b) *if, in addition, μ has compact support, then we can take measures μ_j with densities $\chi_j * \mu$.*

Proof. Let us start with the case (b). Assume again that $\Omega = \mathbf{R}^n$ and take $f \in \mathscr{C}_c(\mathbf{R}^n)$. Then by the previous theorem $f * \chi_j \rightrightarrows f$, and it readilly follows that

$$\int_{\mathbf{R}^n} f(x)(\chi_j * \mu)(x)\,\mathrm{d}x = \int_{\mathbf{R}^n} f * \chi_j\,\mathrm{d}\mu \to \int_{\mathbf{R}^n} f\,\mathrm{d}\mu.$$

For the proof of (a), we first approximate μ by μ_K (see 2.4) where K is a suitable compact set, and then use the convolution as above. ■

31.7. Theorem. *Let f be a β-Lipschitz function on $\mathbf{R}^n$. Then $\chi_k * f$ are (infinitely differentiable) β-Lipschitz functions, $|\chi_k * f(x) - f(x)| \le \beta/k$ for all $x \in \mathbf{R}^n$ and $(\chi_k * f)' \to f'$ almost everywhere in $\mathbf{R}^n$.*

Proof. Since

$$(\chi_k * f)' = \chi_k * f',$$

$\chi_k * f$ are β-Lipschitz functions. For any $x \in \mathbf{R}^n$ we have

$$\begin{aligned}|\chi_k * f(x) - f(x)| &\le \int_{B(x,\frac{1}{k})} |f(x) - f(y)|\, \chi_k(x-y)\, \mathrm{d}y \\ &\le \frac{\beta}{k} \int_{B(x,\frac{1}{k})} \chi_k(x-y)\, \mathrm{d}y = \frac{\beta}{k}.\end{aligned}$$

We can see that $(\chi_k * f')(x) \to f'(x)$ at all points x, where the derivative $f'(x)$ exists and is approximately continuous. But the Rademacher theorem 29.9 and Denjoy's theorem 30.3 tell us that this happens almost everywhere. ∎

31.8. Exercise. Give an alternative proof of the fact that $\mathscr{D}(\Omega)$ is dense in $\mathscr{L}^p(\Omega)$ for $1 \le p < \infty$.

Hint. If $f \in \mathscr{C}_c(\Omega)$, then by Theorem 31.5 the functions $\chi_k * f$ converge uniformly to f and their supports are contained in a compact set (not depending on k). This yields that $\chi_k * f$ converge to f in the $\mathscr{L}^p$-norm as well. Now it is sufficient to use the density of $\mathscr{C}_c(\Omega)$ in $\mathscr{L}^p(\Omega)$ (Exercise 15.17).

31.9. Exercise. Use Theorem 31.4 to prove that the convolution $f * g$ is uniformly continuous provided $f \in \mathscr{L}^p$ and $g \in \mathscr{L}^q$, $\frac{1}{p} + \frac{1}{q} = 1$ and $p, q > 1$. (This will prove the proposition stated in Remark 26.22.)

31.10. Riemann-Lebesgue Lemma. Suppose $f \in \mathscr{L}^1(\mathbf{R}^n)$. Then

$$\lim_{|x|\to\infty} \int_{\mathbf{R}^n} e^{ix\cdot t} f(t)\, \mathrm{d}t = 0.$$

Hint. If $g \in \mathscr{D}(\mathbf{R}^n)$, then integration by parts gives

$$-\int_{\mathbf{R}^n} |x|^2\, e^{ix\cdot t} g(t)\, \mathrm{d}t = \int_{\mathbf{R}^n} e^{ix\cdot t} \sum_{j=1}^{n} \frac{\partial^2 g}{\partial t_j^2}(t)\, \mathrm{d}t$$

whence

$$\left| \int_{\mathbf{R}^n} e^{ix\cdot t} g(t)\, \mathrm{d}t \right| \le |x|^{-2}\, \|\Delta g\|_1 .$$

Now we invoke the fact that $\mathscr{D}(\mathbf{R}^n)$ is dense in $\mathscr{L}^1(\mathbf{R}^n)$: Given $f \in \mathscr{L}^1(\mathbf{R}^n)$ and $\varepsilon > 0$, by Theorem 31.4 there is $g \in \mathscr{D}(\mathbf{R}^n)$ (g depends on f and ε) such that $\|f - g\|_1 < \varepsilon$. Then

$$\left| \int_{\mathbf{R}^n} e^{ix\cdot t} (f(t) - g(t))\, \mathrm{d}t \right| \le \varepsilon,$$

whence

$$\left| \int_{\mathbf{R}^n} e^{ix\cdot t} f(t)\, \mathrm{d}t \right| \le |x|^{-2}\, \|\Delta g\|_1 + \|f - g\|_1 .$$

32. Distributions

Since twenties, physicists started to work with "generalized functions". These "functions" were determined by their "average densities" in a neighbourhood of each point. P.A.M. Dirac introduced the "δ-function" having the following properties:

$$\delta(x) = 0 \text{ for } x \neq 0, \quad \delta(0) = \infty \quad \text{and} \quad \int_U \delta(x) = 1$$

for every open ball U containing the origin. Thus the "average density of the δ-function" in the ball $U(x,r)$ is

$$f_r(x) := \begin{cases} \frac{1}{\lambda U(x,r)} & \text{if } x \in U(0,r), \\ 0 & \text{otherwise.} \end{cases}$$

(Notice that $x \in U(0,r)$ if and only if $0 \in U(x,r)$.) We can see that

$$\delta(x) = \lim_{r \to 0+} f_r(x).$$

If now φ is a continuous function on $\mathbf{R}^n$, denote

$$Z_r(\varphi) = \int_{\mathbf{R}^n} f_r \varphi, \qquad \delta(\varphi) = \varphi(0).$$

Using the definition of continuity, one can show that

$$\lim_{r \to 0+} Z_r(\varphi) = \delta(\varphi)$$

for any continuous function φ on $\mathbf{R}^n$. In other words, the δ-function is a "weak" limit of the sequence $\{Z_r\}$ of functionals while "values" of the δ-function are pointwise limits of the sequence $\{f_r\}$. The great number of quotation-marks in the last sentences indicates that we have to give exact definitions and thus also interpretations of the notion used. Thus while the physicists succesfully used the δ-function and many other "generalized functions", the mathematical theory of distributions (or, generalized functions) arose much later in the end of the 30's and it is connected with names of S.L. Sobolev, and mainly with L. Schwartz.

And one more note. We know that there exist continuous functions which have a derivative at no point. One of the great advantages of the theory of distributions is the fact that each of them does have a derivative. However, we should be aware of the difference between classical derivatives and derivatives in the sense of distributions. Among others, the derivative of a function in sense of distributions is not necessarily a function (cf. Examples 32.5). In order that the distributional derivative of a continuous function f to be again a function (more exactly, a regular distribution, cf. 32.3.1), f has to be locally absolutely continuous. In particular, f must have the classical derivative almost everywhere.

In what follows, we consider an open set $\Omega \subset \mathbf{R}^n$. If $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ is a n-tuple of nonnegative integers (so-called multiindex), set $|\alpha| = \alpha_1 + \dots + \alpha_n$ and denote the differential operator

$$\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

by the symbol D^α. Finally, recall that $\mathscr{D}(\Omega)$ denotes the linear space of all infinitely differentiable functions with compact support contained in Ω (see 31.1).

32.1. Notion of a Distribution. If $\{f_k\}$ is a sequence of functions from $\mathscr{D}(\Omega)$, we say that $f_k \overset{\mathscr{D}}{\to} 0$ if there exists a compact set $K \subset \Omega$ such that $\operatorname{supt} f_k \subset K$ for every k and $D^\alpha f_k \rightrightarrows 0$ on K for every multiindex α. Every linear functional T on $\mathscr{D}(\Omega)$ satisfying

$$T(f_k) \to 0 \quad \text{whenever} \quad f_k \overset{\mathscr{D}}{\to} 0$$

is called a *distribution* (on Ω). Thus every distribution is determined by its values on $\mathscr{D}(\Omega)$.

Let f be a locally integrable function on Ω (i.e. integrable on each compact subset of Ω, in particular, f can be a continuous function). Set

$$T_f(\varphi) = \int_\Omega f\varphi \,\mathrm{d}\lambda$$

for $\varphi \in \mathscr{D}(\Omega)$ (notice that the integral exists). It is not very difficult to prove that T_f is a distribution: Suppose $\varphi_k \in \mathscr{D}(\Omega)$ and $\varphi_k \overset{\mathscr{D}}{\to} 0$. If K is a compact set on whose complement all φ_k's vanish, then

$$|T_f(\varphi_k)| \le \max_{t \in K} |\varphi_k(t)| \int_K |f| \,\mathrm{d}\lambda$$

and $\varphi_k \rightrightarrows 0$ on K. It follows that $T(\varphi_k) \to 0$.

Thus every locally integrable function defines a distribution. We say that T is a *regular distribution* if there is a function $f \in \mathscr{L}^1_{\text{loc}}(\Omega)$ with $T = T_f$. In this sense, we can identify regular distributions and locally integrable functions.

32.2. Remarks. 1. The space $\mathscr{D}(\Omega)$ is usually endowed with a locally convex topology in such a way that $f_k \to 0$ in this topology if and only if $f_k \overset{\mathscr{D}}{\to} 0$. So distributions are continuous linear functionals on $\mathscr{D}(\Omega)$.

2. If a linear functional Z on $\mathscr{D}(\Omega)$ is a "pointwise" limit of a sequence of distributions $\{Z_j\}$ (i.e. if $Z(\varphi) = \lim Z_j(\varphi)$ for each $\varphi \in \mathscr{D}(\Omega)$), then Z is also a distribution. This is a consequence of the Banach-Steinhaus theorem of functional analysis, which is valid for $\mathscr{D}(\Omega)$.

3. If f, g are locally integrable functions and $f = g$ almost everywhere, then apparently $T_f = T_g$. The converse is also true: If the regular distributions T_f, T_g are equal, then $f = g$ almost everywhere.

For this purpose, it is sufficient to prove that $f\eta = g\eta$ for every $\eta \in \mathscr{D}(\Omega)$. Suppose that $h = (f - g)\eta$ on Ω and $h = 0$ outside Ω. Then $h \in L^1(\mathbf{R}^n)$ and $T_h = 0$ in $\mathbf{R}^n$. Thus $\chi_k * h = 0$ for every k and Theorem 31.3 (the notation is the same) yields $h = 0$ almost everywhere.

This assertion shows that we can really identify regular distributions and locally integrable functions.

32.3. Examples. 1. In 32.1 we "embedded" locally integrable functions into the space of distributions. Another class of objects contained in the space of distribution are Radon measures. Let μ be a (positive) Radon measure on Ω. The functional T_μ defined as

$$T_\mu(\varphi) = \mu(\varphi) \quad (= \int_\Omega \varphi \, \mathrm{d}\mu)$$

is again a distribution. In this sense, we can understand every Radon measure as a distribution. For example, the Dirac measure (the "δ-function") defined by $T_\delta(\varphi) = \varphi(0)$ is a distribution. It is not regular. Indeed, if there were a locally integrable function f with $T_f = T_\delta$, we could easily deduce that $f = 0$ almost everywhere. But T_δ is not the zero distribution.

2. It is not true that every distribution is regular or is defined by some Radon measure. As the example consider the functional $\varphi \mapsto \varphi'(0)$. However, there is a simple criterion which enables to decide whether a distribution T is equal to T_μ. Namely, if μ is a Radon measure, then the distribution T_μ is a positive functional. Conversely, if T is a distribution with $T(\varphi) \geq 0$ whenever $\varphi \in \mathscr{D}(\Omega)$ is nonnegative, then there exists a Radon measure μ such that $T = T_\mu$. The last proposition is an easy consequence of the Riesz representation theorem 16.5.

3. For $\varepsilon > 0$, denote $O(\varepsilon) = (-\infty, -\varepsilon) \cup (\varepsilon, +\infty)$ and

$$R_\varepsilon(\varphi) = \int\limits_{O(\varepsilon)} \frac{\varphi(x)}{x} \, \mathrm{d}x \quad \text{for } \varphi \in \mathscr{D}(\mathbf{R})$$

(notice that the integral exists). It is not difficult to see that R_ε is a distribution on $\mathbf{R}$. Further, for every $\varphi \in \mathscr{D}(\mathbf{R})$ there is a finite limit $\lim\limits_{\varepsilon \to 0+} R_\varepsilon(\varphi)$. If we denote it by $T_{\frac{1}{x}}(\varphi)$, then $T_{\frac{1}{x}}$ is a distribution. The reader may find it instructive to prove this assertion directly, or using Remark 32.2.2. Notice also that the function $\frac{1}{x}$ is not locally integrable on $\mathbf{R}$. Thus, the distribution $T_{\frac{1}{x}}$ which is often identified with the function $\frac{1}{x}$ is not regular and also cannot be determined by a Radon measure.

The set of all distributions is a linear space which is the topological dual to $\mathscr{D}(\Omega)$ (cf. Remark 32.2.1).

As mentioned above, every distribution is differentiable. To motivate the exact definition, consider a function f with continuous derivative f' on $\mathbf{R}$ (notice that both f and f' are locally integrable). In this case it is natural to require that $T_{f'}$ should be the derivative of T_f. Since,

$$T_{f'}(\varphi) = \int_{\mathbf{R}} f'\varphi = -T_f(\varphi')$$

for any $\varphi \in \mathscr{D}(\mathbf{R})$ we are led to the following general definition.

32.4. Differentiation of Distributions. Let T be a distribution on Ω. If α is a multiindex, we define the α^{th} (partial) derivative D^α of the distribution T as

$$D^\alpha T(\varphi) = (-1)^{|\alpha|} T(D^\alpha \varphi), \quad \varphi \in \mathscr{D}(\Omega).$$

In particular,

$$\frac{\partial T}{\partial x_j}(\varphi) = -T\left(\frac{\partial \varphi}{\partial x_j}\right).$$

In the sense of identification of regular distributions and locally integrable functions, we often speak about distributional derivatives of functions provided the result of differentiation is a function.

If f is a $\mathscr{C}^1$- function, then its distributional derivative is the classical one: $(T_f)' = T_{f'}$. For the one-dimensional case see the next Remarks. The multi-dimensional case involving partial derivatives is then an easy consequence of Fubini's theorem.

32.5. Remarks and Examples. 1. It remains to show that $D^\alpha T$ is again a distribution. But the linearity of $D^\alpha T$ is obvious, and supposing $\varphi_k \xrightarrow{\mathscr{D}} 0$, then also $D^\alpha \varphi_k \xrightarrow{\mathscr{D}} 0$ and we get $D^\alpha T(\varphi_k) \to 0$ (we used the fact that the mapping $\varphi \mapsto D^\alpha \varphi$ is continuous on $\mathscr{D}(\Omega)$).

2. Notice that any distribution has derivatives of all degrees. However, the derivative of a "function" fails to be a "function". As the next examples show, it can be, for example, a "measure" . For instance, any continuous function which does not possess a derivative at any point is infinitely differentiable — but only in the sense of distributions!

3. Assume now that u and f are locally integrable functions on an interval (a, b) and that the distributional derivative of u is f. Then there exists a locally absolutely continuous function v and a constant c such that $u = v + c$ almost everywhere and $v' = f$ almost everywhere. Indeed, consider a function $\eta \in \mathscr{D}((a, b))$ with $\int_a^b \eta = 1$ and denote by v the indefinite Lebesgue integral of f. We know that v is a locally absolutely continuous function. The integration by parts 23.13 shows that the distributional derivative of $w := u - v$ is a zero function. If $c = \int_a^b w\eta$, $\varphi \in \mathscr{D}((a, b))$ is an arbitrary function and $\psi(x) = \int_a^x \left(\varphi(t) - \eta(t) \int_a^b \varphi(s)\, \mathrm{d}s\right) \mathrm{d}t$, then $\psi \in \mathscr{D}((a, b))$ and

$$\int_a^b (w - c)\varphi = \int_a^b w\varphi - \int_a^b w\eta \int_a^b \varphi = \int_a^b w\psi' = 0.$$

Thus $T_{w-c} = 0$, whence $w = c$ almost everywhere by 32.2.3.

4. Let f be a continuous function having the derivative f' almost everywhere on an interval I. Then f' is a distributional derivative of f if and only if f is locally absolutely continuous on I.

5. Consider the function $f(x) = \max(0, x)$ and the so-called *Heaviside function* Y on $\mathbf{R}$ which is defined as the indicator function of the interval $(0, +\infty)$. Then f is not "classically differentiable" at zero. However, its distributional derivative is the Heaviside function Y. Integration by parts shows easily that

$$(T_f)''(\varphi) = (T_Y)'(\varphi) = -\int_0^{+\infty} \varphi' = \varphi(0) = T_\delta(\varphi)$$

whenever $\varphi \in \mathscr{D}(\Omega)$. We see that the derivative of the distribution T_Y is the Dirac "δ-function". In a similar way we find that

$$(T_Y)''(\varphi) = -\varphi'(0) \quad \text{for} \quad \varphi \in \mathscr{D}(\Omega).$$

Show that $(T_Y)''$ is neither a "function" nor a "measure".

6. Let f be a real function whose classical derivative is not locally Lebesgue integrable (cf. Example 25.1). Then $(T_f)'$ is not a regular distribution.

7. The distributional derivative of the Cantor function of Example 23.1 is not the zero function but a measure "concentrated" on the Cantor set (see Exercise 24.8.a).

8. The function $\ln(|x|)$ is locally Lebesgue integrable. Its derivative in the sense of distributions is the distribution $T_{\frac{1}{x}}$ (This gives one way how to prove that $T_{\frac{1}{x}}$ is a distribution.)

9. For $\varphi \in \mathscr{D}(\mathbf{R})$, set

$$U(\varphi) = \sum_{k=1}^{\infty} \varphi^{(k)}(k)\,.$$

Show that U is a distribution on $\mathbf{R}$ which is neither regular nor determined by a Radon measure.

10. Let $\Delta := \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ be the Laplace operator and $u(x) := |x|^{2-n}$. Our aim is to find Δu in the distributional sense. First, we have that $\frac{\partial u}{\partial x_i} = (2-n)\frac{x_i}{|x|^n}$ is a locally integrable function. Now the distributional derivative $\frac{\partial^2 u}{\partial x_i^2}$ is a complicated distribution containing an "integral average" similar to the distribution of Example 32.3.3. But we have an easier task, namely to find out the sum $\sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2}$ of these distributions. If κ_n denotes the volume of the unit ball in $\mathbf{R}^n$ (cf. Exercise 26.17), then we prove that

$$\int_{\mathbf{R}^n} \sum_{i=1}^{n} \frac{x_i}{|x|^n} \frac{\partial \varphi}{\partial x_i}(x) = -n\kappa_n \varphi(0)$$

for every test function $\varphi \in \mathscr{D}(\mathbf{R}^n)$. Hence

$$\Delta u = n(2-n)\kappa_n \delta_0 .$$

To this end let γ be a nonincreasing infinitely smooth function on $(0,\infty)$, $\gamma(t) = 1$ for $t < 1$, $\gamma(t) = 0$ for $t > 2$, $|\gamma'(t)| \le 2$ and $\eta_r(x) = \gamma(x/r)$. We use the decomposition

$$\varphi = \varphi(0)\eta_r + (\varphi - \varphi(0))\eta_r + \varphi(1-\eta_r).$$

Then

$$\int_{\mathbf{R}^n} \sum_{i=1}^{n} \frac{x_i}{|x|^n} \frac{\partial \varphi}{\partial x_i}(x) = A_r + B_r + C_r + D_r ,$$

where

$$\begin{aligned} A_r &= \varphi(0) \int_{\mathbf{R}^n} \sum_{i=1}^{n} \frac{x_i}{|x|^n} \frac{\partial \eta_r}{\partial x_i}(x)\,\mathrm{d}x \\ &= \varphi(0) \int_{\mathbf{R}^n} |x|^{1-n} r^{-1} \gamma'(|x|/r)\,\mathrm{d}x = \varphi(0) n\kappa_n \int_0^{2r} r^{-1}\gamma'(t/r)\,\mathrm{d}t = -n\kappa_n\varphi(0) \end{aligned}$$

(we used Exercise 26.17),

$$\begin{aligned} B_r &= \int_{\mathbf{R}^n} \sum_{i=1}^{n} \frac{x_i}{|x|^n} (\varphi(x) - \varphi(0)) \frac{\partial \eta_r}{\partial x_i}(x)\,\mathrm{d}x, \\ C_r &= \int_{\mathbf{R}^n} \sum_{i=1}^{n} \frac{x_i}{|x|^n} \eta_r(x) \frac{\partial \varphi}{\partial x_i}(x)\,\mathrm{d}x \end{aligned}$$

and

$$D_r = \int_{\mathbf{R}^n} \sum_{i=1}^{n} \frac{x_i}{|x|^n} \frac{\partial\,(\varphi(1-\eta_r))}{\partial x_i}(x)\,\mathrm{d}x.$$

Set

$$\psi_i(x) := \begin{cases} \frac{x_i}{|x|^n}\varphi(x)(1-\eta_r(x)) & \text{if } x \ne 0; \\ 0 & \text{if } x = 0. \end{cases}$$

An easy calculation shows that $\psi_i \in \mathscr{D}(\mathbf{R}^n)$ and

$$\sum_{i=1}^{n} \frac{\partial \psi_i}{\partial x_i}(x) = \sum_{i=1}^{n} \frac{x_i}{|x|^n} \frac{\partial\,(\varphi(1-\eta_r))}{\partial x_i}(x).$$

Thus

$$D_r = \int_{\mathbf{R}^n} \sum_{i=1}^{n} \frac{\partial \psi_i}{\partial x_i} \, \mathrm{d}x = 0.$$

The proof will be finished by estimating the integrals B_r and C_r and taking the limit as $r \to 0$. Since $\varphi \in \mathscr{D}(\mathbf{R}^n)$, there exists a constant K such that $|\nabla \varphi(x)| \le K$ and $|\varphi(x) - \varphi(0)| \le K\,|x|$ for all $x \in U(0,1)$. Obviously, $|\nabla \eta_r| \le \frac{2}{r}$. We obtain

$$|B_r| \le \int_{B(0,2r)} \sum_{i=1}^{n} \left| \frac{x_i}{|x|^n} \right| |(\varphi(x) - \varphi(0)| \left| \frac{\partial \eta_r}{\partial x_i}(x) \right| \mathrm{d}x \le \frac{4nKr}{r} \int_{B(0,2r)} |x|^{1-n} \, \mathrm{d}x \to 0$$

and

$$|C_r| \le \int_{B(0,2r)} \sum_{i=1}^{n} \eta_r(x) \left| \frac{x_i}{|x|^n} \right| \left| \frac{\partial \varphi}{\partial x_i}(x) \right| \mathrm{d}x \le nK \int_{B(0,2r)} |x|^{1-n} \, \mathrm{d}x \to 0.$$

Another operation with distributions is the mutiplication. Since we prove the Schwartz impossibility theorem which says that on the space of all distibutions the multiplication cannot be defined in a reasonable way, we restrict to the multiplication by smooth functions. However, if h is a smooth function on $\mathbf{R}$ and f is continuous, then

$$T_{hf}(\varphi) = T_f(h\varphi) \quad \text{for} \quad \varphi \in \mathscr{D}(\Omega)$$

and this equality leads to the following definition.

32.6. Multiplication of Distributions by Smooth Functions. Let T be a distribution on Ω and $h \in \mathscr{C}^\infty(\Omega)$. We define the distribution $hT = Th$ by

$$hT(\varphi) = T(h\varphi)\,, \quad \varphi \in \mathscr{D}(\Omega).$$

32.7. Remarks and Examples. 1. It is easy to show that hT and Th are distributions.
2. We have

$$xT_\delta(\varphi) = T_\delta(x\varphi(x)) = 0,$$
$$xT_{\frac{1}{x}}(\varphi) = T_{\frac{1}{x}}(x\varphi(x)) = \int_{\mathbf{R}} \varphi = T_1(\varphi)$$

and we see that the multiplication in the space of distributions is almost the same as the "pointwise" multiplication.

32.8. Schwartz Impossibility Theorem. *On the space of distributions on* $\mathbf{R}$, *a multiplication cannot be defined in such a way that it would be commutative and associative and that* $xT_\delta = 0$ *and* $xT_{\frac{1}{x}} = T_1$.

Proof. The following shows that it is impossible to define a multiplication:

$$0 = 0 \cdot T_{\frac{1}{x}} = (xT_\delta)T_{\frac{1}{x}} = T_\delta(xT_{\frac{1}{x}}) = T_\delta \cdot 1 = T_\delta.$$

■

32.9. Convergence of Sequences of Distributions. Let $\{T_k\}$ be a sequence of distributions on Ω. We say that T_k converge to a distribution T on Ω, denoted by $T_k \to T$, if $T_k(\varphi) \to T(\varphi)$ for every test function $\varphi \in \mathscr{D}(\Omega)$. Thus the convergence in the sense of distributions is nothing else than the pointwise convergence, or the weak*-convergence on $\mathscr{D}'(\Omega)$.

32.10. Examples. (a) If $T_k := T_{\sin kx}$ on $\mathbf{R}$, then $\{T_k\}$ converges to zero (i.e. to the distribution T_0).

(b) Let T_k be the regular distribution on $\mathbf{R}$ determined by the function $\frac{\sin kx}{\pi x}$. Show that $T_k \to T_\delta$ (when keeping the convention of 32.1, the sequence of functions $\{\frac{\sin kx}{\pi x}\}$ converges to the Dirac measure δ in the sense of distributions).

Hint. Choose $\varphi \in \mathscr{D}(\mathbf{R})$. If $\varphi = 0$ outside the interval $[-A, A]$, then

$$\pi T_k(\varphi) = \int_{-A}^{A} \frac{(\varphi(x) - \varphi(0))}{x} \sin kx \, \mathrm{d}x + \varphi(0) \int_{-A}^{A} \frac{\sin kx}{x} \, \mathrm{d}x.$$

By the Riemann-Lebesgue Lemma 31.10, the limit of the first term is zero. Further it is known that

$$\lim_{k\to\infty} \int_{-kA}^{kA} \frac{\sin t}{t} \, \mathrm{d}t = \pi \,,$$

whence the assertion follows.

(c) Show that the sequence of functions $\{\chi_k\}$ (cf. 31.1) converges on $\mathbf{R}^n$ to the Dirac measure δ in the sense of distributions.

Hint. The proposition is an easy consequence of Theorem 31.5.

(d) Suppose $f_k, f \in \mathscr{L}^1(\Omega)$. If $\lim \int_K |f_k - f| \to 0$ for every compact set $K \subset \Omega$, then $T_{f_k} \to T_f$.

32.11. Theorem. *Let α be a multiindex, T, T_k distributions. If $T_k \to T$ on Ω, then $D^\alpha T_k \to D^\alpha T$.*

Proof. The proposition is an immediate consequence of definitions. ■

32.12. Example. We define that $\sum T_n = T$ provided $\sum T_n(\varphi) = T(\varphi)$ for any test function $\varphi \in \mathscr{D}(\Omega)$. The previous theorem tell us that every convergent series of distributions can be differentiated term by term. Here there is an illustrating example:

Let f be a 2π-periodic function on $\mathbf{R}$,

$$f(x) = \begin{cases} \frac{1}{2}(\pi - x) & \text{for } x \in (0, \pi], \\ -\frac{1}{2}(\pi + x) & \text{for } x \in [-\pi, 0), \\ 0 & \text{for } x = 0. \end{cases}$$

It follows from the theory of Fourier series that

$$f(x) = \sum_{k=1}^{\infty} \frac{\sin kx}{k}$$

for every $x \in \mathbf{R}$. Show that:

(a) This equality holds in the sense of distributions as well.
(b) The derivative of the distribution T_f is a 2π-periodic measure μ whose restriction to the interval $[-\pi, \pi]$ equals $-\frac{1}{2} + \pi\delta$.

On the other hand, Theorem 32.11 implies that $(T_f)' = \sum_{k=1}^{\infty} T_{\cos kx}$.

Therefore, in the sense of distributions we have the equality

$$\sum_{k=1}^{\infty} \cos kx = \mu.$$

We see that we can assign to the divergent series $\sum_{k=1}^{\infty} \cos kx$ a "sum" but this one is a distribution and not a function.

32.13. Notes. The theory of distributions in the framework of duality between topological vector spaces was created by L. Schwartz in the late 1930's and was fully published in [1947] (because of the scientific isolation during the second world war he published his theory afterwards). He also introduced the fundamental space $\mathscr{S}(\mathbf{R}^n)$ suitable for the Fourier transform of distributions.

Some related considerations were implicitly contained in earlier works of other authors. In fact, L. S. Sobolev [1936] also introduced basic notions of the theory of distributions. Since he was led by a special question concerning the solution of the Cauchy problem he did not realize the power of his own ideas.

Concerning the history of the theory of distributions, we suggest also Lützen's monograph [*1982]. Other material can be found in J. Horváth [1970] or J. Horváth [*1966].

33. Fourier Transform

Throughout this chapter, $u \cdot v$ denotes the inner product of n-dimensional vectors u and v. By a function we understand a *complex-valued* function and in this way we modify the definitions of function spaces like L^p.

The Fourier transform is an operator whose calculus provides a number of formulae. Roughly speaking, the Fourier transform transforms differentiation with respect to j^{th} variable into multiplication by x_j and convolution into product. Due to these properties it has wide applications in function spaces theory, theory of partial differential equations, operator calculus and other fields.

When reading this chapter, we suggest to notice the parallel between the theory of Fourier series and the Fourier transform (cf. Remark 33.14).

33.1. Fourier Transform of L^1-Functions. Formally, we can write the Fourier transform in the form

$$\widehat{f}(x) := (2\pi)^{-n/2} \int_{\mathbf{R}^n} e^{-ix\cdot y} f(y)\,\mathrm{d}y.$$

This formula can be understood as a pointwise definition provided $f \in \mathscr{L}^1(\mathbf{R}^n)$. Notice that the Lebesgue integral always exists.

We start with simple formulae.

33.2. Theorem. *If $f, g \in \mathscr{L}^1(\mathbf{R}^n)$, then*

(a) $\widehat{f * g} = (2\pi)^{n/2} \widehat{f}\,\widehat{g}$;

(b) $\int_{\mathbf{R}^n} \widehat{f} g = \int_{\mathbf{R}^n} f \widehat{g}$ *(the so-called multiplication formula).*

Proof. The proof will follow using Fubini's theorem easily. It only needs to be observed that $\widehat{f} g \in \mathscr{L}^1$ whenever $f, g \in \mathscr{L}^1$. ■

The following theorem plays a key role in applications of the Fourier transform to the theory of partial differential equations.

33.3. Theorem. *Suppose $f \in \mathscr{L}^1(\mathbf{R}^n)$ and $1 \le j \le n$.*

(a) *If $x_j f \in \mathscr{L}^1(\mathbf{R}^n)$, then*

$$\frac{\partial \widehat{f}}{\partial x_j} = (\widehat{-ix_j f}).$$

(b) *Suppose that* f, $\frac{\partial f}{\partial x_j}$ *are continuous on* $\mathbf{R}^n$, $\frac{\partial f}{\partial x_j} \in \mathscr{L}^1(\mathbf{R}^n)$ *and*

$$\lim_{|x|\to\infty} f(x) = 0.$$

Then

$$\widehat{\partial f/\partial x_j}(x) = i x_j \widehat{f}(x).$$

Proof. (a) The differentiation under the integral sign of

$$\widehat{f}(x) = (2\pi)^{-n/2} \int_{\mathbf{R}^n} e^{-ix\cdot t} f(t)\,\mathrm{d}t$$

leads to

$$\frac{\partial \widehat{f}}{\partial x_j}(x) = -(2\pi)^{-n/2} \int_{\mathbf{R}^n} e^{-ix\cdot t} i t_j f(t)\,\mathrm{d}t.$$

(b) Using Fubini's theorem and integration by parts we obtain

$$\begin{aligned}(2\pi)^{-n/2} \int_{\mathbf{R}^n} e^{-ix\cdot t} \frac{\partial f}{\partial t_j}(t)\,\mathrm{d}t &= -(2\pi)^{-n/2} \int_{\mathbf{R}^n} f(t) \frac{\partial}{\partial t_j} e^{-ix\cdot t}\,\mathrm{d}t \\ &= i x_j (2\pi)^{-n/2} \int_{\mathbf{R}^n} f(t) e^{-ix\cdot t}\,\mathrm{d}t\,.\end{aligned}$$

■

33.4. Lemma. *If* $h(x) = e^{-|x|^2/2}$, *then* $\widehat{h} = h$.

Proof. Assume first that $n = 1$. By the previous lemma both h and $\widehat{h}$ are solutions to the differential equation

$$u' + xu = 0\,.$$

Since $h(0) = \widehat{h}(0)$ (see Exercise 26.17.b), we get $\widehat{h} = h$. Assuming we had already proved the theorem for $n = 1$, we get

$$\begin{aligned}\int_{\mathbf{R}^n} e^{-ix\cdot t} e^{-|t|^2/2}\,\mathrm{d}t &= \int_{\mathbf{R}^n} e^{-ix_1 t_1 - t_1^2/2} \ldots e^{-ix_n t_n - t_n^2/2}\,\mathrm{d}t_1 \ldots \mathrm{d}t_n \\ &\int_{\mathbf{R}} e^{-ix_1 t_1 - t_1^2/2}\,\mathrm{d}t_1 \cdots \int_{\mathbf{R}} e^{-ix_n t_n - t_n^2/2} (2\pi)^{n/2} e^{-|x|^2/2}.\end{aligned}$$

■

33.5. Fourier Transform of Distributions. We start with a motivation. It would be quite natural to define the Fourier transform of a distribution in such way that the Fourier transform of the regular distribution T_f would be the distribution $T_{\widehat{f}}$. But if $f \in \mathscr{L}^1(\mathbf{R}^n)$ and $\varphi \in \mathscr{D}(\mathbf{R}^n)$, the multiplication formula 33.2.b gives

$$\int_{\mathbf{R}^n} \widehat{f}\varphi = \int_{\mathbf{R}^n} f\widehat{\varphi}.$$

and since the Fourier transform of a function from $\mathscr{D}(\mathbf{R}^n)$ fails to be in $\mathscr{D}(\mathbf{R}^n)$ we cannot proceed in this way. (Notice that the zero function $\varphi = 0$ is the only function of $\mathscr{D}(\mathbf{R}^n)$ with $\widehat{\varphi} \in \mathscr{D}(\mathbf{R}^n)$.)

Observe also that the Fourier transform of a function from $\mathscr{L}^1$ needs not to be an element of $\mathscr{L}^1$. If f is the indicator function of the interval $[-1, 1]$, then

$$\widehat{f}(x) = \sqrt{\frac{2}{\pi}}\,\frac{\sin x}{x}.$$

So we search for a subset of $\mathscr{L}^1(\mathbf{R}^n)$ which would be closed with respect to the Fourier transform and to differentiation (and thus also with respect to multiplication by polynomials, cf. Theorem 33.3), and in this way we are naturally led to the notion of the Schwartz space $\mathscr{S}(\mathbf{R}^n)$.

We say that a $\mathscr{C}^\infty$–function f on $\mathbf{R}^n$ belongs to the *Schwartz space* $\mathscr{S}(\mathbf{R}^n)$ if $p\,D^\alpha f$ is a bounded function for any multiindex α and for any polynomial p on $\mathbf{R}^n$. Equivalently, it is easy to see that $f \in \mathscr{S}(\mathbf{R}^n)$ if and only if $p\,D^\alpha f \in \mathscr{L}^1(\mathbf{R}^n)$ whenever p is a polynomial and α a multiindex.

The space of test functions $\mathscr{D}(\mathbf{R}^n)$ is a subset of $\mathscr{S}(\mathbf{R}^n)$. The function $e^{-|x|^2}$ provides an example of a function from $\mathscr{S}(\mathbf{R}^n)$ which does not have compact support.

With the help of Theorem 33.3 and Lemma 33.6 it can be easily shown that $\mathscr{S}(\mathbf{R}^n)$ is closed with respect to the Fourier transform, and that the mapping $\varphi \mapsto \widehat{\varphi}$ is a linear isomorphism of $\mathscr{S}(\mathbf{R}^n)$ onto itself.

A linear functional S on $\mathscr{S}(\mathbf{R}^n)$ is called a *tempered distribution* if it is continuous in the following sense: If $\{\varphi_j\}$ is a sequence of functions from $\mathscr{S}(\mathbf{R}^n)$ and $pD^\alpha\varphi_j \to 0$ for any multiindex α and any polynomial p, then $T(\varphi_j) \to 0$. We can easily verify that every tempered distribution (restricted to $\mathscr{D}(\mathbf{R}^n)$) is a distribution.

The *Fourier transform* $\widehat{T}$ of a tempered distribution T is defined as

$$\widehat{T}(\varphi) = T(\hat{\varphi})\,, \quad \varphi \in \mathscr{S}(\mathbf{R}^n).$$

Apparently $\widehat{T}$ is again a tempered distribution.

So far, the Fourier transform is defined for functions from $\mathscr{L}^1(\mathbf{R}^n)$ and for tempered distributions. Any locally integrable function f determines a (regular) distribution, and many of them even tempered distributions. If $f \in \mathscr{L}^p(\mathbf{R}^n)$ $(1 \le p \le \infty)$, then

$$S_f : \varphi \mapsto \int_{\mathbf{R}^n} f\varphi\,, \quad \varphi \in \mathscr{S}(\mathbf{R}^n)$$

is a tempered distribution. In the same way as in the previous chapter, we identify the function $f \in \mathscr{L}^p(\mathbf{R}^n)$ and the tempered distribution S_f. The multiplication formula yields that $\widehat{S}_f = S_{\hat{f}}$ and we see that the new definition of the Fourier transform of tempered distributions is an extension of the original one for functions from $\mathscr{L}^1(\mathbf{R}^n)$.

Since the Fourier transform is now defined also for functions outside of $\mathscr{L}^1(\mathbf{R}^n)$, the question arises whether, given function f (say $f \in \mathscr{L}^p$), a function g can be found so that $\widehat{S}_f = S_g$ (then we could say that the Fourier transform of f is g).

We know that in the case of $\mathscr{L}^1$ the answer is positive as we can take $g = \widehat{f}$. On the other hand, Exercise 33.9 shows that the Fourier transform of a constant function (which is a function of $\mathscr{L}^\infty(\mathbf{R}^n)$) is not a function.

Let us note that even in the case when the Fourier transform of a function f is a function g (in the sense mentioned above), the function g is not determined pointwise (as in the original definition of the Fourier transform for $f \in \mathscr{L}^1(\mathbf{R}^n)$) but only except a null set.

In what follows, we restrict our attention to the case $p = 2$ and we state the main theorem which guarantees a good theory for the Hilbert space $L^2(\mathbf{R}^n)$. First we define the *inverse Fourier transform* by the formula

$$\tilde{f}(y) = (2\pi)^{-n/2} \int_{\mathbf{R}^n} e^{ix\cdot y} f(x)\,\mathrm{d}x\,, \quad f \in \mathscr{L}^1(\mathbf{R}^n);$$
$$\tilde{T}(\varphi) = T(\tilde{\varphi})\,,\ \varphi \in \mathscr{S}(\mathbf{R}^n), \quad T \text{ is a tempered distribution}$$

If $f \in L^2(\mathbf{R}^n)$, then we define the *Fourier transform* $\widehat{f}$ as such $g \in L^2(\mathbf{R}^n)$ for which $\widehat{S}_f = S_g$. The *inverse Fourier transform* $\tilde{f}$ of a function $f \in L^2(\mathbf{R}^n)$ is defined in a similar way. Before we verify that these definitions make sense, let state a useful lemma.

33.6. Lemma. *The Schwartz space $\mathscr{S}(\mathbf{R}^n)$ is a dense subset of $L^2(\mathbf{R}^n)$ and*

$$\tilde{\hat{f}} = f\,, \quad \hat{\tilde{f}} = f\,, \quad \|\widehat{f}\|_2 = \|f\|_2$$

for $f \in \mathscr{S}(\mathbf{R}^n)$.

Proof. In Theorem 31.4 we proved even that $\mathscr{D}(\mathbf{R}^n)$ is a dense subset of $L^2(\mathbf{R}^n)$.

Now we are going to prove the formula $\tilde{\hat{f}} = f$. Denote $h(x) = e^{-|x|^2/2}$. Remember, according to Lemma 33.4, that $\widehat{h} = h$. Using the multiplication formula 33.2.b we obtain

$$\int_{\mathbf{R}^n} \widehat{f}(t) h(\frac{t}{k})\,\mathrm{d}t = \int_{\mathbf{R}^n} f(\frac{t}{k}) \widehat{h}(t)\,\mathrm{d}t.$$

When taking the limit as $k \to \infty$ it follows that

$$\int_{\mathbf{R}^n} \widehat{f}(t)\,\mathrm{d}t = \int_{\mathbf{R}^n} f(0)\widehat{h}(t)\,\mathrm{d}t = (2\pi)^{n/2} f(0).$$

We make the substitution $z = x + y$. The integrals then become

$$\begin{aligned} f(y) &= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} (2\pi)^{-n} e^{-ix\cdot t} f(t+y)\,\mathrm{d}t \right) \mathrm{d}x \\ &= \int_{\mathbf{R}^n} \left(\int_{\mathbf{R}^n} (2\pi)^{-n} e^{iy\cdot t} e^{-iz\cdot t} f(z)\,\mathrm{d}z \right) \mathrm{d}t \\ &= \int_{\mathbf{R}^n} (2\pi)^{-n/2} e^{iy\cdot t} \widehat{f}(t)\,\mathrm{d}t = \tilde{\hat{f}}(y). \end{aligned}$$

The formula $f = \hat{\hat{f}}$ can be proved in a similar way. Notice that for the Fourier transform of the complex conjugated function we have $\hat{\bar{f}} = \bar{\hat{f}}$ and thus the multiplication formula yields

$$\int_{\mathbf{R}^n} |f|^2 \, \mathrm{d}x = \int_{\mathbf{R}^n} f\bar{f} \, \mathrm{d}x = \int_{\mathbf{R}^n} f\bar{\hat{\hat{f}}} \, \mathrm{d}x = \int_{\mathbf{R}^n} f\hat{\hat{\bar{f}}} \, \mathrm{d}x = \int_{\mathbf{R}^n} \hat{f}\bar{\hat{f}} \, \mathrm{d}x = \int_{\mathbf{R}^n} \left|\hat{f}\right|^2 \mathrm{d}x.$$

■

33.7. Plancherel Theorem. *For any $f \in L^2(\mathbf{R}^n)$ there exists $g \in L^2(\mathbf{R}^n)$ such that $\widehat{S}_f = S_g$, $\tilde{S}_g = S_f$ and $\|f\|_2 = \|g\|_2$.*

Proof. By virtue of Lemma 33.6 there are $f_j \in \mathscr{S}(\mathbf{R}^n)$ such that $f_j \to f$ in $L^2(\mathbf{R}^n)$. By the same lemma, $\{\widehat{f_j}\}$ is a Cauchy sequence in $L^2(\mathbf{R}^n)$ and thus there exists $g \in \mathrm{L}^2(\mathbf{R}^n)$ with

$$\left\|g - \widehat{f_j}\right\|_2 \to 0.$$

By passing to the limit we get

$$\widehat{S}_f = S_g, \qquad \tilde{S}_g = S_f \quad \text{and} \quad \|f\|_2 = \|g\|_2 .$$

■

33.8. Remark. The Plancherel theorem says that the (original) Fourier transform can be uniquely extended from $\mathscr{S}(\mathbf{R}^n)$ to $L^2(\mathbf{R}^n)$. The result is a continuous linear isometry $\mathscr{F}$ of the space $L^2(\mathbf{R}^n)$ onto itself. Thus $\mathscr{F}$ is a unitary mapping and one can easily find out that also the inner product in $L^2(\mathbf{R}^n)$ is invariant with respect to it.

33.9. Exercise. Let δ be the Dirac δ-function. Prove that $\widehat{\delta}$ is the constant $(2\pi)^{-n/2}$ and that the Fourier transform of this constant is δ.

33.10. Exercise. Find a function u whose Fourier transform $\widehat{u}$ is a solution to the equation

$$-\triangle\widehat{u} + \widehat{u} = \delta$$

in the sense of distribution (here $\triangle$ is the Laplace operator, cf. 32.5.10).

33.11. Fourier Transform of a Function in $L^1(\mathbf{R}^n)$. (a) Suppose $f \in \mathscr{L}^1(\mathbf{R}^n)$. Then $\widehat{f}$ is a uniformly continuous function and $\lim_{|x|\to\infty} \widehat{f}(x) = 0$ (in other words, $\widehat{f} \in \mathscr{C}_0(\mathbf{R}^n)$).

Hint. The assertion concerning the limit at infinity is a consequence of the Riemann-Lebesgue lemma 31.10. As for the proof of uniform continuity it is no restriction to assume that f is from the space $\mathscr{D}(\mathbf{R}^n)$ since we know that $\mathscr{D}(\mathbf{R}^n)$ is dense in $L^1(\mathbf{R}^n)$ (Theorem 31.4). Suppose that the support of f is contained in $U(0,R)$. We use the inequality

$$\left|e^{iy\cdot t} - e^{ix\cdot t}\right| \le |t| \; |y-x|$$

to get

$$\left|\widehat{f}(y) - \widehat{f}(x)\right| = \left|\int_{\mathbf{R}^n} (e^{iy\cdot t} - e^{ix\cdot t}) f(t) \, \mathrm{d}t\right| \le R \; |y-x| \; \|f\|_1$$

from which the uniform continuity of $\widehat{f}$ follows.

(b) Remark that any function of $\mathscr{C}_0(\mathbf{R}^n)$ does not need to be the Fourier transform of a function from $L^1(\mathbf{R}^n)$. Consider the example of the $\mathscr{C}_0(\mathbf{R})$–function $g : x \mapsto \dfrac{x}{\sqrt{1+x^2}\log(2+x^2)}$. If $f \in \mathscr{L}^1(\mathbf{R})$, then Fubini's theorem together with boundedness of the function $a \mapsto \int_1^a \frac{1}{z} e^{-iz} \, \mathrm{d}z$ on $(1,+\infty)$ imply that $\int_1^x \frac{\widehat{f}(t)}{t} \, \mathrm{d}t$ is a bounded function on $(1,+\infty)$. Hence apparently $\widehat{f} \ne g$.

Note that, in contrast to the equality $\{\widehat{f} \colon f \in L^2(\mathbf{R}^n)\} = L^2(\mathbf{R}^n)$, a satisfactory characterization of the range $\{\widehat{f} \colon f \in \mathscr{L}^1(\mathbf{R})\}$ of the Fourier transform on $L^1(\mathbf{R})$ is not known.

33.12. Exercise. For $f \in L^2(\mathbf{R}^n)$ define

$$\mathscr{G}f(x) = \int_{\mathbf{R}} f(t) \frac{e^{-itx} - 1}{-it}\,\mathrm{d}t.$$

Show that $\mathscr{G}f$ is an indefinite Lebesgue integral of the (locally integrable) function $\widehat{f}$.

33.13. Remark. Define the "Fourier transform" F on $\mathscr{L}^1(\mathbf{R}^n)$ by the formula

$$Ff = \frac{1}{a}\int_{\mathbf{R}^n} e^{-ibx\cdot t} f(t)\,\mathrm{d}t$$

and denote $c := (2\pi)^n\,|b|^{-n}\,a^{-2}$. Then (under certain assumptions)

$$\begin{aligned} F(f * g) &= a\,Ff\,Fg; \\ F(Ff(x)) &= c\,f(-x); \\ \frac{\partial(Ff)}{\partial x_j}(x) &= -ibx_j Ff(x); \\ \|Ff\|_2 &= c\,\|f\|_2\,. \end{aligned}$$

Browsing through different manuscripts, various choices of a and b appear (e.g. $a = 1$, $a = (2\pi)^{n/2}$, $b = \pm 1$, $b = \pm 2\pi$).

33.14. Fourier Transform and Fourier Series. Let T be the interval $[-\pi, \pi]$. We assign to any function $u \in L^1(T)$ the sequence $\{\widehat{u}(k)\}_{k\in\mathbb{Z}}$ of its *Fourier coefficients*

$$\widehat{u}(k) := \frac{1}{2\pi}\int_T e^{-ikt} u(t)\,\mathrm{d}t.$$

Conversely, to a sequence $\{c_k\} \in l^1(\mathbb{Z})$ we assign the function

$$\tilde{c}(x) := \sum_{k\in\mathbb{Z}} c_k e^{ikx}$$

(the sum of the convergent *trigonometrical series* with coefficients c_k). The (formal) series

$$\sum_{k\in\mathbb{Z}} \widehat{u}(k) e^{ikx}$$

is called the *Fourier series* of a function u.

We leave the theory of Fourier series out of this manuscript (the books by V. Jarník [JII] or W. Rudin [*1974] are suggested). However, we state here without proofs some results indicating an analogy between the theories of Fourier series and Fourier transform.

The sequence of Fourier coefficients $\{\widehat{u}(k)\}$ is an analogy of the Fourier transform $\widehat{f}(x)$ and the sum $\tilde{c}(x)$ is an analogy of the inverse Fourier transform $\tilde{f}(x)$. In some sense, the theory of Fourier series is less complicated because of $L^2(T) \subset L^1(T)$ and $l^1(\mathbb{Z}) \subset l^2(\mathbb{Z})$. On the other hand, we lose the symmetry between the Fourier transform and the inverse Fourier transform.

(a) If $u \in Ł^1(T)$, then $\lim\limits_{|k|\to\infty} \widehat{u}(k) = 0$. The proposition is a consequence of the Riemann-Lebesgue lemma 31.10; the proof is similar to that for the case of Fourier transform (cf. 33.11.a). Also in this case, not every sequence in $c_0(\mathbb{Z})$ is a sequence of Fourier coefficients of a function in $L^1(T)$.

(b) If $\{c_k\} \in l^1(\mathbb{Z})$, then $\tilde{c}$ is a continuous function on T with $\tilde{c}(-\pi) = \tilde{c}(\pi)$ (compare with 33.11.a). Let us note again that not every continuous function can be expressed in this way.

(c) If $u \in L^2(T)$, then $\{\widehat{u}(k)\} \in l^2(\mathbb{Z})$ (this follows from the Bessel inequality). Conversely, for any sequence $\{c_k\} \in l^2(\mathbb{Z})$ there exists a unique function $u \in L^2(T)$ with $c_k = \widehat{u}(k)$ (the Riesz-Fischer theorem). This function can be obtained as the limit of the sequence $\{s_k\} \subset L^2(T)$, where

$$s_k(x) = \sum_{j=-k}^{k} c_k e^{ikx}, \qquad x \in T.$$

Moreover, $\|u\|_2 = \sqrt{2\pi}\ \|\{\widehat{u}(k)\}\|_2$ (the Parseval identity). Compare with the Plancherel theorem 33.7.

(d) The mapping $u \mapsto \{\widehat{u}(k)\}$ and $\{c_k\} \mapsto \tilde{c}$ can be also generalized to wider classes of functions or sequences. We are not going to mention the details. Just notice that the mapping $\{c_k\} \mapsto u$, $\{c_k\} \in l^2(\mathbb{Z})$ given by the Riesz-Fischer theorem is in fact an extension of the mapping $\{c_k\} \mapsto \tilde{c}$, $\{c_k\} \in l^1(\mathbb{Z})$. However, in the case of l^2, the function u is determined only as an element of $L^2(T)$, i.e. except of a null set.

(e) If $u \in \mathscr{C}^1(\mathbf{R})$ is a 2π-periodic function and $v = u'$, then

$$\widehat{v}(k) = ik\widehat{u}(k)\,, \quad k \in \mathbb{Z}.$$

Compare again with Theorem 33.3.

33.15. Notes. The theory of Fourier series and the Fourier transform has a very long history. They were D. Bernoulli and L. Euler who first used what is now called Fourier series. J. B. J. Fourier used trigonometric series in connection with problems on heat and wrote his famous book [*1822]. He also studied the Fourier transform in the form

$$F(x) = \int_0^\infty f(x)\cos tx\,dx$$

and formally derived the inversion formula.

Common features of the Fourier transform and the Fourier series (see the list of parallels in 33.14) indicated that there should exist a unifying theory covering both. This was established by A. Weil [1940] who presented Fourier transform on locally compact Abelian groups. This approach led further to harmonic analysis on groups, see E. Hewitt and K.A. Ross [*1970] or G.B. Folland [*1975].

Note that M. Plancherel extended the definition of the Fourier transform for functions in $L^2(\mathbf{R})$ so that he obtained the unitary operator on this Hilbert space (cf. Remark 33.8).

Besides the Fourier transform, various integral transforms like the Laplace or Mellin transform are also examined.

Roughly speaking, an integral transform is a one-to-one mapping between two function spaces defined by a formula of the form

$$\widehat{u}(x) = \int_Y K(x,y)\,u(y)\,\mathrm{d}\nu(y).$$

Often an inverse formula can be found as

$$u(y) = \int_X K'(x,y)\,\widehat{u}(x)\,\mathrm{d}\mu(x).$$

The theory of Fourier series can be including as a particular case since series could be considered as integrals with respect to the counting measure.

F. Change of Variable and k-dimensional Measures

34. Change of Variable Theorem

Throughout this chapter $\mathbf{R}^n$ will be the Euclidean space and $k \le n$ a non-negative integer. Our aim is to prove the general change of variable theorem for k-dimensional measures.

We want to "measure" various subsets of $\mathbf{R}^n$ whose dimension is $k \le n$. For instance, we are interested in the problem how to calculate the length of curves or the area of surfaces in $\mathbf{R}^3$. In miscellaneous fields of analysis we encounter various methods leading to this goal in different degrees of generality. If a set $A \subset \mathbf{R}^n$ is an isometric copy of a set $E \subset \mathbf{R}^k$, it is reasonable to introduce its k-dimensional measure as the Lebesgue measure of its preimage E. Consequently, it is not difficult to introduce the k-dimensional measure of "flat" sets. Now, if we consider a set A on a curved surface, we divide it into "almost flat" parts A_j and sum evaluating measures of flat sets E_j which are "close to A_j". In the limiting case, we obtain the intuitive meaning of measure of A. Our task is therefore to introduce a concept of k-dimensional measures which would be general enough, easily describing and transparent. We will propose two independent solutions of this problem in the following chapters.

The integral with respect to the k-dimensional measure is the so-called curve ($k = 1$), or surface ($k > 1$) integral. We prove a change of variable formula for this integral which will include Theorem 26.13 as a special case.

We start by recalling useful notions from linear algebra.

The *inner product* of vectors x and y will be denoted again by $x \cdot y$ and the norm of an element x of $\mathbf{R}^n$ simply by $|x|$. By e_j we denote the basic vector $[0, \dots, 1, \dots, 0]$ of $\mathbf{R}^n$ having jth term 1 and all other terms zero.

Further, $\mathbf{M}_{n,k}$ will be the set of all matrices with n rows and k columns. Each such a matrix represents a linear mapping of $\mathbf{R}^k$ into $\mathbf{R}^n$.

If $L : \mathbf{R}^k \to \mathbf{R}^n$ is a linear mapping, then there is a unique mapping $L^* : \mathbf{R}^n \to \mathbf{R}^k$ such that $Lx \cdot y = x \cdot L^*y$ for all $x \in \mathbf{R}^k$ and $y \in \mathbf{R}^n$. The mapping L^* is called the *adjoint* to L. If L is represented by a matrix A, then L^* is represented by the transposed matrix to A and this one will be denoted by A^T. The *norm* of the mapping L is defined by the formula

$$\|L\| = \sup\{|Lx| : x \in \mathbf{R}^k, |x| \le 1\}.$$

Further we denote $\left\|L^{-1}\right\| = \sup\{|x| : x \in \mathbf{R}^k, |Lx| \le 1\}$.

We suppose that the reader is familiar with the notion of a determinant. We will relate it not only to square matrices but also to objects which are representable by means of square matrices: to k-tuples of vectors from $\mathbf{R}^k$, k-tuples of linear forms over $\mathbf{R}^k$, or to linear mappings of $\mathbf{R}^k$ into itself.

34.1. Properties of Determinants. *Let $A, B \in \mathbf{M}_{k,k}$. Then*

(a) $\det A = \det A^T$,

(b) $\det AB = \det A \det B$.

34.2. Isometric Mappings. Let $\mathbf{V}, \mathbf{W}$ be k-dimensional linear spaces with inner products. Remember that a linear mapping $L : \mathbf{W} \to \mathbf{V}$ is said to be an *isometry* provided it preserves distances of points. Then, of course, it preserves also the norm and the inner product (indeed, the inner product can be expressed in terms of the norm according to the formula $4x \cdot y = \|x+y\|^2 - \|x-y\|^2$). The matrices of isometric mappings of $\mathbf{R}^k$ into $\mathbf{R}^k$ are termed *unitary* or *orthogonal*. A matrix A is orthogonal if and only if its rows (columns) form an orthonormal basis of $\mathbf{R}^k$. The product of unitary matrices is a unitary matrix.

34.3. Lemma. *Let $\mathbf{W}$ be a k-dimensional linear space with an inner product. Then there exists a linear isometry $Q\colon \mathbf{R}^k \to \mathbf{W}$.*

Proof. Let $u_1, \dots, u_k$ be an orthonormal basis of $\mathbf{W}$. The linear transformation which sends $e_i \mapsto u_i$, $i = 1, \dots, k$, is an isometry. ■

34.4. Example.
Let $\mathbf{W} \subset \mathbf{R}^3$ be a vector space generated by the vectors $[1,1,1]$ and $[0,1,2]$. We are looking for a linear isometry of $\mathbf{R}^2$ into $\mathbf{W}$. We arrive at this finding an orthonormal basis u_1, u_2 of the space $\mathbf{W}$. Solving the equation $[1,1,1] \cdot \big([0,1,2] + \alpha[1,1,1]\big) = 0$ we obtain $\alpha = -1$, so that the vector $[-1,0,1] = [0,1,2] - [1,1,1]$ is orthogonal to $[1,1,1]$. If we set $u_1 = 3^{-1/2}[1,1,1]$, $u_2 = 2^{-1/2}[-1,0,1]$, then the vectors u_1, u_2 are unit vectors. The matrix of the required linear mapping has as columns u_1, u_2.

34.5. Lemma. *Let $Q \in \mathbf{M}_{k,k}$ be a unitary matrix. Then $|\det Q| = 1$.*

34.6. Lemma. *Let $A \in \mathbf{M}_{k,k}$. Then there exist unitary matrices Q_1, $Q_2 \in \mathbf{M}_{k,k}$ and a diagonal matrix $D \in \mathbf{M}_{k,k}$ such that $A = Q_1 D Q_2$.*

Proof. An elementary theorem of linear algebra tell us that there is a unitary matrix U and a symmetric matrix R such that $A = UR$. Further, there exist a unitary matrix Q and a diagonal matrix D such that $R = Q^T D Q$. Now it remains to put $Q_1 = UQ^T$, $Q_2 = Q$. ■

Now, go back to measure theory. We denote, as usual, by λ the Lebesgue measure.

34.7. Lemma. *Let $L\colon \mathbf{R}^k \to \mathbf{R}^k$ be a linear mapping. Then, for any measurable set $E \subset \mathbf{R}^k$, the set $L(E)$ is measurable and $\lambda(L(E)) = |\det L|\, \lambda E$.*

Proof. The measurability of $L(E)$ is obvious. By Lemma 34.6 there are isometric linear mappings Q_1, Q_2 and a diagonal mapping $D = (d_{i,j})_{i,j=1}^k$ of the space $\mathbf{R}^k$ into itself such that $L = Q_1 D Q_2$. Let $K = [a_1, b_1] \times \cdots \times [a_k, b_k]$. Then (to simplify the notation we will assume that $d_{i,i} \geq 0$, but the same result is obtained also in the remaining cases) $D(K) = [d_{1,1}a_1,\, d_{1,1}b_1] \times \cdots \times [d_{k,k}a_k,\, d_{k,k}b_k]$, so that

$$\lambda(D(K)) = |d_{1,1}(b_1 - a_1) \dots d_{k,k}(b_k - a_k)| = |\det D|\, \lambda K.$$

Since isometric linear mappings are measure preserving and their composition with another linear mapping M do not change the absolute value of the determinant of M, it follows that for any measurable set $E \subset \mathbf{R}^n$ we have

$$\lambda(L(E)) = \lambda(DQ_2(E)) = |\det D|\ \lambda(Q_2(E)) = |\det D|\ \lambda E = |\det L|\ \lambda E.$$

■

34.8. k-dimensional Measures. Let σ be a measure on a σ-algebra Σ containing $\mathscr{B}(\mathbf{R}^n)$. Denote

$$\sigma^* E = \{\inf \sigma S \colon S \supset E,\ S \in \Sigma\}.$$

We say that σ is a *k-dimensional measure on* $\mathbf{R}^n$ if:

(a) $\sigma^* I(E) = \lambda^* E$ whenever $I : \mathbf{R}^k \to \mathbf{R}^n$ is an isometric mapping and $E \subset \mathbf{R}^k$;

(b) $\sigma^* \varphi(E) \le \beta^k \sigma^* E$ for each β-Lipschitz mapping $\varphi \colon E \to \mathbf{R}^n$ and $E \subset \mathbf{R}^n$.

A k-dimensional measure is not uniquely determined by (a) and (b) (cf. Remark 34.31). Nevertheless, on reasonable sets all k-dimensional measures coincide. Notice also that for any k-dimensional measure σ, the property (a) implies that $\sigma([0,1]^k \times \{0\}^{n-k}) = 1$.

By Exercise 30.7, the Lebesgue measure satisfies the property (b) (and thus also (a)) in the case $k = n$. Using (a), we see immediately that the Lebesgue measure is the unique n-dimensional measure on $\mathbf{R}^n$.

34.9. Existence of k-dimensional Measures. *There exists a k-dimensional measure on* $\mathbf{R}^n$*.*

Proof. For $E \subset \mathbf{R}^n$, set

$$\mu^* E = \inf\{\sum_j \lambda_k G_j\}\,,$$

where the infimum is taken over all sequences of open sets $\{G_j\}$, $G_j \subset \mathbf{R}^k$, which admit 1-Lipschitz functions φ_j on G_j such that $E \subset \bigcup_j \varphi_j(G_j)$. If such a sequence $\{G_j\}$ does not exist, we set $\mu^* E = \infty$.

Obviously μ^* is an outer measure on $\mathbf{R}^n$.

Let Σ be the σ-algebra of all μ^*-measurable sets (see 4.4). Similarly as in Theorem 1.19 we prove that every Borel set belongs to Σ: It is sufficient to observe that for any open halfspace H and any continuous mapping φ of an open set $G \subset \mathbf{R}^k$ into $\mathbf{R}^n$, the set G is a disjoint union of $G \cap \varphi^{-1}(H)$ (and this is clearly open) and $G \setminus \varphi^{-1}(H)$ (this is a union of a countable family of closed sets; therefore a measurable set which can be approximated by an open superset with arbitrarily close measure). Obviously

$$\mu^* E = \inf\{\mu F \colon F \in \Sigma,\ F \supset E\}$$

as the sets of the form $\bigcup_j \varphi_j(G_j)$ are measurable (G_j is a union of a sequence of compact sets and each continuous image of a compact set is compact). According to Exercise 30.7 we obtain that μ is a k-dimensional measure. ■

34.10. Volume. Let $\mathbf{W}$ be a k-dimensional linear space with an inner product and $L \colon \mathbf{R}^k \to \mathbf{W}$ a linear mapping. Let $Q \colon \mathbf{R}^k \to \mathbf{W}$ be an isometry and $a := \left|\det(Q^{-1}L)\right|$. Then $\lambda(Q^{-1}L(E)) = a\,\lambda E$ for any Borel set $E \subset \mathbf{R}^k$. The constant a is independent of the choice of the isometric mapping Q and has a similar meaning to that of the multiplier $|\det L|$ of Lemma 34.7 (the case $n = k$). The expression $\left|\det(Q^{-1}L)\right|$ may be simplified: Using 34.1 we have

$$\begin{aligned}(\det(Q^{-1}L))^2 &= \det((Q^{-1}L)^T Q^{-1}T)) = \det\big(Q^{-1}Le_i \cdot Q^{-1}Le_j\big)_{i,j=1}^k \\ &= \det\big(Le_i \cdot Le_j\big)_{i,j=1}^k.\end{aligned}$$

The expression

$$\sqrt{\det(u_i \cdot u_j)_{i,j=1}^k}$$

is said to be the *volume of a k-tuple of vectors* $(u_1, \dots, u_k)$ and will be denoted by $\operatorname{vol}(u_1, \dots, u_k)$. It has a geometric interpretation as the "k-dimensional measure" of the set

$$\{c_1u_1 + \dots + c_ku_k : (c_1, \dots, c_k) \in [0,1]^k\}.$$

We associate the *volume* also to the linear mapping L by

$$\operatorname{vol} L := \operatorname{vol}(Le_1, \dots, Le_k).$$

Let us notice that the estimate $\operatorname{vol} L \le \|L\|^k$ is valid. We may summarize the consequences for the k-dimensional measure in the following theorem.

34.11. Theorem. *Let* $\mathbf{W}$ *be a k-dimensional linear subspace of* $\mathbf{R}^n$ *and* σ *be a k-dimensional measure on* $\mathbf{R}^n$*. Let* $L\colon \mathbf{R}^k \to \mathbf{W}$ *be a linear mapping. Then*

$$\sigma(L(E)) = \operatorname{vol} L\, \lambda E$$

for any measurable set $E \subset \mathbf{R}^k$*.*

34.12. Theorem. *Let* $\mathbf{W}$ *be a k-dimensional linear space and* $L\colon \mathbf{R}^k \to \mathbf{W}$, $M\colon \mathbf{R}^k \to \mathbf{R}^k$ *be linear mappings. Then* $\operatorname{vol}(LM) = |\det M| \operatorname{vol} L$*.*

Proof. Let $Q\colon \mathbf{R}^k \to \mathbf{W}$ be a linear isometry. By 34.1

$$\begin{aligned}(\operatorname{vol} LM)^2 &= \det((Q^{-1}LM)^T Q^{-1}LM) = \det M^T \det((Q^{-1}L)^T Q^{-1}L) \det M \\ &= (\det M)^2 (\operatorname{vol} L)^2.\end{aligned}$$

■

34.13. Cauchy-Binet formula. Now we derive a formula which makes the calculation of the volume of a mapping $L\colon \mathbf{R}^k \to \mathbf{R}^n$ easier, at least for some combinations of n and k. The elements of the set $\{1, \dots, n\}^k$ (i.e. the ordered k-tuples of indices) will be called *multiindices*. We will denote $I = I(k,n)$ the set of all multiindices $\alpha \in \{1, \dots, n\}^k$ which are *increasing*, i.e. $\alpha_1 < \dots < \alpha_k$. Consider a matrices $A = (a_{m,j})_{\substack{m=1,\dots,n\\ j=1,\dots,k}}$ and $B = (b_{m,j})_{\substack{m=1,\dots,n\\ j=1,\dots,k}}$. For each multiindex $\alpha \in I$ write $A_\alpha = (a_{\alpha_i,j})_{\substack{i=1,\dots,k\\ j=1,\dots,k}}$, $B = (b_{\alpha_i,j})_{\substack{i=1,\dots,k\\ j=1,\dots,k}}$. Now we show that

$$\det A^T B = \sum_{\alpha \in I} \det A_\alpha^T B_\alpha.$$

A routine calculation shows that

$$\det A^T B = \det(\sum_{m=1}^{n} a_{m,i} b_{m,j})_{i,j=1}^{k} = \sum_{\beta\in\{1,\dots,m\}^k} \det(a_{\beta_i,i}, b_{\beta_i,j})_{i,j=1}^{k}.$$

Since

$$\det(a_{\beta_i,i}, b_{\beta_i,j})_{i,j=1}^{k} = 0$$

whenever any index in β is repeated, we obtain

$$\begin{aligned}\det A^T B &= \sum_{\beta\in\{1,\dots,m\}^k} \det(a_{\beta_i,i}, b_{\beta_i,j})_{i,j=1}^{k} \\ &= \sum_{\alpha\in I} \sum_{\beta\in\{\alpha_1,\dots,\alpha_k\}^k} \det(a_{\beta_i,i}, b_{\beta_i,j})_{i,j=1}^{k} = \sum_{\alpha\in I} \det(A_\alpha^T B_\alpha).\end{aligned}$$

Now the promised application to the volume follows: We set both A and B to be the matrices of the mapping L. From the above computation we obtain

$$(\operatorname{vol} L)^2 = \det A^T A = \sum_{\alpha\in I} \det A_\alpha^T A_\alpha = \sum_{\alpha\in I} (\det A_\alpha)^2.$$

34.14. Example. Let

$$u_1 = [0,2,1], \quad u_2 = [1,0,1],$$

so that

$$A = \begin{pmatrix} 0 & 1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix}.$$

Then we evaluate

$$(\operatorname{vol}(u_1,u_2))^2 = \det(A^T A) = \det \begin{pmatrix} 5 & 1 \\ 1 & 2 \end{pmatrix} = 9\,,$$

or

$$(\operatorname{vol}(u_1,u_2))^2 = \big(\det \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}\big)^2 + \big(\det \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\big)^2 + \big(\det \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}\big)^2 = (-2)^2 + (-1)^2 + 2^2 = 9.$$

34.15. Lemma. *Let σ be a k-dimensional measure in $\mathbf{R}^n$. Let $L : \mathbf{R}^k \to \mathbf{R}^n$ be an injective linear mapping, $G \subset \mathbf{R}^k$ an open set, $F \subset G$ a measurable set and $\varepsilon > 0$. Further, let $\varphi\colon G \to \mathbf{R}^n$ be an almost everywhere differentiable mapping G for which*

$$|\varphi(s) - \varphi(t) - L(s-t)| \le \varepsilon\,|s-t|\,, \quad t, s \in F.$$

Denote $\beta = 1 + \varepsilon\,\|L^{-1}\|$, $\alpha = (1 - \varepsilon\,\|L^{-1}\|)^{-1}$.

Then the following assertions hold:

(a) *the mapping $\varphi \circ L^{-1}$ is β-Lipschitz on $L(F)$;*
(b) *if $\varepsilon\,\|L^{-1}\| < 1$, then the mapping φ is one-to-one and the mapping $L\circ\varphi^{-1}$ is α-Lipschitz on $\varphi(F)$;*
(c) *for almost every $t \in F$ we have $\|\varphi'(t) - L\| \le \varepsilon$ and $\operatorname{vol}\varphi'(t) \le \beta^k \operatorname{vol} L$;*
(d) *if $\varepsilon\,\|L^{-1}\| < 1$, then $\operatorname{vol}\varphi'(t) \ge \alpha^{-k} \operatorname{vol} L$ for almost all $t \in F$.*

(e) $\sigma^*\varphi(E) \le \beta^k \operatorname{vol} L\, \lambda^* E$ *for any set* $E \subset F$;
(f) *if* $\varepsilon \left\|L^{-1}\right\| < 1$, *then*

$$\beta^{-k}\alpha^{-k} \int_E \operatorname{vol}\varphi'(t)\,\mathrm{d}t \le \sigma^*\varphi(E) \le \beta^k\alpha^k \int_E \operatorname{vol}\varphi'(t)\,\mathrm{d}t$$

for any set $E \subset F$.

Proof. (a) and (b): Let $t, s \in F$. Then

$$\begin{aligned}|\varphi(s) - \varphi(t)| &\le |\varphi(s) - \varphi(t) - L(s-t)| + |Ls - Lt| \le \varepsilon\,|s-t| + |Ls - Lt| \\ &\le (\varepsilon \left\|L^{-1}\right\| + 1)\,|Ls - Lt|\end{aligned}$$

(hence (a) follows) and

$$\begin{aligned}|Ls - Lt| &\le |\varphi(s) - \varphi(t) - L(s-t)| + |\varphi(s) - \varphi(t)| \\ &\le \varepsilon \left\|L^{-1}\right\|\, |Ls - Lt| + |\varphi(s) - \varphi(t)|\,.\end{aligned}$$

The last inequality may be also simplified

$$(1 - \varepsilon \left\|L^{-1}\right\|)\,|L(s) - L(t)| \le |\varphi(s) - \varphi(t)|\ ,$$

which yields (b).

(c) and (d): By the Lebesgue Density Theorem 29.2, almost every point of F is its point of density. If, in addition, the derivative $A := \varphi'(t)$ exists at such a point t, then obviously $\|A - L\| \le \varepsilon$. Since the mapping AL^{-1} is (as in (a)) β-Lipschitz, we have $\left\|AL^{-1}\right\| \le \beta$ and $\operatorname{vol} AL^{-1} \le \beta^k$, or $\operatorname{vol} A \le \beta^k \operatorname{vol} L$. Similarly we obtain $\operatorname{vol} A \ge \alpha^{-k} \operatorname{vol} L$ if $\varepsilon \left\|L^{-1}\right\| < 1$.

(e) and (f): Using the definition of the k-dimensional measure and the preceding considerations, it follows that

$$\sigma^*\varphi(E) \le \beta^k \sigma^* L(E) = \beta^k \operatorname{vol} L\, \lambda^* E.$$

Assuming $\varepsilon \left\|L^{-1}\right\| < 1$ we obtain

$$\sigma^*\varphi(E) \le \beta^k \operatorname{vol} L\, \lambda^* E \le \beta^k\alpha^k \int_E \operatorname{vol}\varphi'(t)\,\mathrm{d}t$$

and

$$\sigma^*\varphi(E) \ge \alpha^{-k} \operatorname{vol} L\, \lambda^* E \ge \beta^{-k}\alpha^{-k} \int_E \operatorname{vol}\varphi'(t)\,\mathrm{d}t.$$

■

Now we will deal with locally Lipschitz mappings. We will frequently and tacitly use the fact that (in view of the Rademacher Theorem 30.3) any Lipschitz mapping φ has its derivative $\varphi'(t)$ almost everywhere, and that this is measurable (as a function of t).

34.16. Lemma. *Let $G \subset \mathbf{R}^k$ be an open set, $E \subset G$ a measurable set and $\varphi\colon G \to \mathbf{R}^n$ a locally Lipschitz mapping. Let $\mathscr{F}$ be a closed or open subset of $\mathbf{M}_{n,k}$ and ε a positive function on $\mathscr{F}$. Then there exists a countable disjoint partition $E \cap \{\varphi' \in \mathscr{F}\} = \bigcup D_j$, where D_j are measurable and for each j there is $L \in \mathscr{F}$ such that*

$$|\varphi(t) - \varphi(s) - L(s-t)| < \varepsilon(L)\,|s-t|$$

for each $s, t \in D_j$.

Proof. First suppose that $\mathscr{F}$ is compact. The system of sets

$$\{M \in \mathscr{F}\colon\ \|M - L\| < \varepsilon(L)\},$$

where L runs over $\mathscr{F}$, is an open cover of $\mathscr{F}$. We can thus find its countable subcover and divide $E \cap \{\varphi' \in \mathscr{F}\}$ into a disjoint union of measurable sets E^i, where for each i there is $L \in \mathscr{F}$ such that for all $t \in E^i$ we have $\|\varphi'(t) - L\| < \varepsilon(L)$. Let i be fixed. We have $E^i = \bigcup\limits_p E^i_p$, where

$$E^i_p := \big\{t \in E^i\colon \text{for each } s \in U(t, \tfrac{2}{p}) \text{ we have} \\ |\varphi(s) - \varphi(t) - L(s-t)| < \varepsilon(L)\,|s-t|\big\}.$$

If we divide E^i_p into measurable pairwise disjoint parts $E^i_{p,q}$ with diameter less than $1/p$, then $E^i_{p,q}$ have all the required properties. If $\mathscr{F}$ is closed or open in $\mathbf{M}_{n,k}$, it is a countable union of compact sets and we use the preceding part. ∎

34.17. Sard Theorem. *Let σ be a complete k-dimensional measure on $\mathbf{R}^n$. Let $G \subset \mathbf{R}^k$ be an open set and $\varphi\colon G \to \mathbf{R}^k$ a locally Lipschitz mapping. Let $Z := \{t\colon\ \operatorname{vol}\varphi'(t) = 0\}$. Then $\sigma(\varphi(Z)) = 0$.*

Proof. It is enough to show that $\sigma(\varphi(E)) = 0$ for any set $E := Z \cap \{\|\varphi'\| \le m\}$. If $\mathscr{K} := \{M \in \mathbf{M}_{n,k}\colon\ \operatorname{vol} M = 0 \text{ a } \|M\| \le m\}$, then $\mathscr{K}$ is a closed subset of $\mathbf{M}_{n,k}$. Choose $\delta > 0$. With each $L \in \mathscr{K}$ we associate $\varepsilon(L) > 0$ such that $2^k \|L\|^{k-1} \varepsilon(L) < \delta$ and $|Lt| \ge \varepsilon(L)\,|t|$ for every t perpendicular to $\operatorname{Ker} L := \{s \in \mathbf{R}^k\colon Ls = 0\}$. According to Lemma 34.16 there is a disjoint partition E into a countable union of measurable sets D_j such that for each j there is $L \in \mathscr{K}$ satisfying

$$|\varphi(t) - \varphi(s) - L(s-t)| < \varepsilon(L)\,|s-t|$$

for each $s, t \in D_j$. Let j be fixed. The dimension d of the space $\operatorname{Ker} L$ is less than k. Let P be the orthogonal projection of $\mathbf{R}^k$ onto $\operatorname{Ker} L$ and $\mathbf{W}$ be a (k-d)-dimensional subspace of $\mathbf{R}^n$ orthogonal to $L(\mathbf{R}^k)$. We find an isometric mapping Q of the space $\operatorname{Ker} L$ to $\mathbf{W}$ and set $Mt = Lt + \varepsilon(L)QPt$. It is easily verified that for $\varepsilon = \varepsilon(L)$ we have $\|M - L\| \le \varepsilon$, $\left\|M^{-1}\right\| \le 1/\varepsilon$ and $\operatorname{vol} M = \varepsilon^{k-d} \operatorname{vol} L \le \varepsilon^{k-d} m^d \le \varepsilon m^{k-1}$. By Lemma 34.15 we get

$$\sigma^* \varphi(D_j) \le (2\varepsilon \left\|M^{-1}\right\|)^k \operatorname{vol} M\, \lambda D \le 2^k m^{k-1} \varepsilon\, \lambda D_j \le \delta\, \lambda D_j.$$

Taking the union and letting $\delta \to 0$ we obtain $\sigma^* \varphi(E) = 0$. ∎

34.18. Change of Variable Theorem. *Let σ be a complete k-dimensional measure on $\mathbf{R}^n$. Let $G \subset \mathbf{R}^k$ be an open set and $\varphi\colon G \to \mathbf{R}^n$ a locally Lipschitz mapping. If u is a measurable function on G and $w(x) := \sum\{u(t)\colon t \in \varphi^{-1}(x)\}$, then*

$$\int_{\varphi(G)} w(x)\,\mathrm{d}\sigma(x) = \int_G u(t)\,\mathrm{vol}\,\varphi'(t)\,\mathrm{d}t\,,$$

provided the integral on the right-hand side converges.

Proof. Apparently it is enough to prove the theorem when u is the indicator function of a measurable set $E \subset \mathbf{R}^k$. If E is of measure zero, it is mapped by a locally Lipschitz mapping again to a set of measure zero. If $E \subset \{\mathrm{vol}\,\varphi' = 0\}$, then according to the Sard Theorem 34.17 we have $\sigma(\varphi(E)) = 0$ and the integral on the right-hand side is also zero. Hence we may assume that $E \subset \{\mathrm{vol}\,f' > 0\}$. If $\mathscr{F} := \{M \in \mathbf{M}_{n,k}\colon \mathrm{vol}\,M > 0\}$, then $\mathscr{F}$ is an open subset of $\mathbf{M}_{n,k}$. Choose $\tau > 1$. To each matrix $L \in \mathbf{M}_{n,k}$ find $\varepsilon(L) > 0$ such that $\varepsilon(L)\,\|L^{-1}\| < 1 - \tau^{-(1/2k)}$ (then the constants α, β of Lemma 34.15 satisfy the estimate $\beta^k \le \alpha^k < \sqrt{\tau}$). By Lemma 34.16 there is a disjoint partition $E = \bigcup\limits_j D_j$, where D_j are measurable sets and for every j there is $L \in \mathscr{K}$ such that

$$|\varphi(t) - \varphi(s) - L(s-t)| < \varepsilon(L)\,|s-t|$$

for all $s, t \in D_j$. Then by virtue of Lemma 34.15, the mapping φ is one-to-one on D_j. We first reduce the general case to the case when φ is one-to-one on E. Thanks to Lemma 34.15 we have

$$\tau^{-1}\sigma^*\varphi(D_j) \le \int_{D_j} \mathrm{vol}\,\varphi'(t)\,\mathrm{d}t \le \tau\sigma^*\varphi(D_j).$$

For each compact set $K \subset D_j$ the set $\varphi(K)$ is compact and thus measurable. There is a sequence $\{K_i\}$ of compact subset D_j such that $\lambda K_i \to \lambda D_j$. Then the set $F_j := \bigcup\limits_i \varphi(K_i)$ is a σ-measurable subset $\varphi(D_j)$ and

$$\int_{D_j} \mathrm{vol}\,\varphi'(t)\,\mathrm{d}t \le \tau\sigma F_j.$$

Denote $F = \bigcup\limits_j F_j$. If we sum over j, we obtain

$$\tau^{-1}\sigma^*\varphi(E) \le \int_E \mathrm{vol}\,\varphi'(t)\,\mathrm{d}t \le \tau\sigma\varphi(F)\,.$$

Whence, passing to the limit as $\tau \to 1$ we get

$$\sigma^*\varphi(E) \le \int_E \mathrm{vol}\,\varphi'(t)\,\mathrm{d}t \le \sigma\varphi(F).$$

Since $F \subset E$, it follows

$$\sigma\varphi(E) = \int_E \operatorname{vol}\varphi'(t)\,\mathrm{d}t\,,$$

which is the desired formula in the case of the indicator function of the set E.

Now, if φ is not one-to-one on E, the same procedure as above yields a disjoint partition $E = \bigcup_j E_j$, where E_j are measurable sets and φ is one-to-one on E_j. According to the preceding part of the proof, it follows that

$$\int_{\varphi(G)} c_{E_j}\,\mathrm{d}\sigma = \sigma\varphi(E_j) = \int_{E_j} \operatorname{vol}\varphi'(t)\,\mathrm{d}t.$$

Summing over j we complete the assertion. ■

34.19. Corollary. *Let σ be a complete k-dimensional measure on $\mathbf{R}^n$. Let $G \subset \mathbf{R}^k$ be an open set and $\varphi\colon G \to \mathbf{R}^n$ a locally Lipschitz mapping. Let f be a σ-measurable function on $\varphi(G)$ and $E \subset G$ a measurable set. If we define the Banach indicatrix $N(x,\varphi,E)$ as the cardinality of the set $E \cap \varphi^{-1}(x)$, then*

$$\int_{\varphi(G)} N(x,\varphi,E)\,f(x)\,\mathrm{d}\sigma(x) = \int_E f(\varphi(t))\operatorname{vol}\varphi'(t)\,\mathrm{d}t\,,$$

provided either integral exists.

In particular, if the mapping φ is, in addition, one-to-one, then

$$\int_{\varphi(G)} f(x)\,\mathrm{d}\sigma(x) = \int_G f(\varphi(t))\operatorname{vol}\varphi'(t)\,\mathrm{d}t.$$

Proof. The assertion is a direct consequence of Theorem 34.18. It remains only to verify that $f \circ \varphi$ is measurable on $\{\operatorname{vol}\varphi' > 0\}$ and $N(x,\varphi,E)$ is σ-measurable.

If $E \subset \{\operatorname{vol}\varphi' = 0\}$, then an appeal to the Sard Theorem 34.17 reveals that $\sigma\varphi(E) = 0$, so that $N(x,\varphi,E) = 0$ almost everywhere. Therefore, we may suppose that $E \subset \{\operatorname{vol}\varphi' > 0\}$ and similarly as in the proof of the preceding theorem we may restrict to the case when φ is one-to-one on E. Further, it is no restriction to assume that f is the indicator function of a σ-measurable set $H \subset \varphi(G)$. In this case there are Borel sets $B, N \subset \varphi(G)$ such that $B \subset H$, $H \setminus B \subset N$ and $\sigma N = 0$. By Theorem 34.18 we have $\lambda(E \cap \varphi^{-1}(N)) = 0$, and therefore $\lambda(E \cap \varphi^{-1}(H \setminus B)) = 0$. We see that the set $E \cap \varphi^{-1}(H) = E \cap \bigl(\varphi^{-1}(B) \cup \varphi^{-1}(H \setminus B)\bigr)$ is measurable. ■

34.20. Bilipschitz Mappings. Let (P_1,ρ_1), (P_2,ρ_2) be metric spaces, $G \subset P_1$ and $\beta \in [1,\infty)$ a constant. We say that a mapping $\varphi\colon G \to P_2$ is *β-bilipschitz* on G if

$$\beta^{-1}\rho_1(x,y) \le \rho_2(\varphi(x),\varphi(y)) \le \beta\rho_1(x,y)$$

for all $x, y \in G$. In other words, whenever both φ and φ^{-1} are β-Lipschitz mappings. A mapping is termed *bilipschitz* if it is β-bilipschitz for some β.

A mapping $\psi\colon G \to P_2$ is said to be *locally bilipschitz* if each point $x \in G$ has a neighborhood U_x and β_x such that the restriction of ψ to U_x is β_x-bilipschitz.

34.21. Remarks. 1. Each bilipschitz mapping is one-to-one, it is even homeomorphic. However, a locally bilipschitz mapping fails to be one-to-one; and even if it is one-to-one, it is not necessarily homeomorphic.

2. Notice that 1-bilipschitz mappings are nothing else than isometries.

3. Let ψ be a bilipschitz mapping of an open subset of $\mathbf{R}^k$ to $\mathbf{R}^n$. Then ψ is differentiable almost everywhere (by the Rademacher Theorem 31.2). Moreover, if $\psi'(t)$ exists, then the rank of $\psi'(t)$ is k.

34.22. Regular Mappings and Diffeomorphisms. Every *regular mapping* (i.e. a $\mathscr{C}^1$-mapping of an open set $G \subset \mathbf{R}^k$ to $\mathbf{R}^k$ whose Jacobian is nowhere on G vanishing) serves as an example of a locally bilipschitz mapping. Also any *regular mapping* of an open set $G \subset \mathbf{R}^k$ into $\mathbf{R}^n$ ($k \leq n$), (which is defined as a $\mathscr{C}^1$-mapping whose Jacobi matrix has rank k at each point G) is a locally bilipschitz mapping.

Any one-to-one regular mapping φ of an open set $G \subset \mathbf{R}^k$ into $\mathbf{R}^k$ is a *diffeomorphism*, i.e. φ is a homeomorphism for which both φ and φ^{-1} are $\mathscr{C}^1$-mappings.

34.23. Examples. The mappings given by *polar* or *spherical* coordinates of Chapter 26 are typical examples of regular mappings. If $a \in U(0,1)$ is a constant vector, then $f : x \mapsto |x|\,(\frac{x}{|x|} - a)$, $x \in U(0,1)$ may serve as as example of a mapping which is $(1-|a|)^{-1}$-bilipschitz but not of class $\mathscr{C}^1$ (as it is not differentiable at the origin).

34.24. Integration on k-dimensional Surfaces. In applications, k-dimensional measures are mostly used on the so-called k-dimensional surfaces. These objects admit local parametrizations by means of Lipschitz (or even bilipschitz) mappings, and consequently the k-dimensional measure is uniquely determined on subsets of k-dimensional surfaces by the Change of Variable Theorem.

Given a set $\Omega \subset \mathbf{R}^n$, we always consider it as a metric space whose topology is inherit from those of $\mathbf{R}^n$. A set $\Omega \subset \mathbf{R}^n$ is called a *k-dimensional surface* whenever for each point $x \in \Omega$ there exists a locally bilipschitz homeomorphic mapping φ of an open set $G \subset \mathbf{R}^k$ into Ω such that $x \in \varphi(G)$, and $\varphi(G)$ is a relatively open subset of Ω. The mapping φ is then called a *parametrization* of $\varphi(G)$.

If each point of Ω has a neighborhood parametrizable by mappings of class $\mathscr{C}^\ell$ ($\ell \geq 1$), we say that the surface Ω itself is of class $\mathscr{C}^\ell$.

Roughly speaking, k-dimensional surfaces are sets which locally look like a bilipschitz deformation of an open part of $\mathbf{R}^k$. Equivalently, we may describe this property in terms of *charts*, which are defined as inverse mappings to parametrizations. A k-dimensional *chart* is thus defined as a locally bilipschitz homeomorphic mapping of an open subset U of Ω onto an open subset of $\mathbf{R}^k$. A set Ω is a k-dimensional surface if and only if for each point $x \in \Omega$ there is a k-dimensional chart defined on a neighborhood of x (relative to Ω).

Because of the Lindelöf property of $\mathbf{R}^n$ each k-dimensional surface is a countable union of surfaces of the form $\varphi(G)$, where φ is a parametrization.

Another method of characterizing k-dimensional surfaces of class $\mathscr{C}^\ell$ is to use an implicit description: A set $\Omega \subset \mathbf{R}^n$ is a k-dimensional surface of class $\mathscr{C}^\ell$ if and only if each $x \in \Omega$ admits a neighborhood W_x in $\mathbf{R}^n$ and a mapping

$g\colon W_x \to \mathbf{R}^{n-k}$ of class $\mathscr{C}^\ell$ such that the matrix $g'(x)$ has rank $n-k$ and $W_x \cap \Omega = W_x \cap \{g = 0\}$.

(It is an easy exercise to use the Implicit Function Theorem in order to get a parametrization. Conversely, if $\varphi\colon G \to \Omega$ is a parametrization of class $\mathscr{C}^\ell$ and $x = \varphi(t)$, then there exists the projection Π of the space $\mathbf{R}^n$ onto $\mathbf{R}^k$ such that $(\Pi \circ \varphi)'(t) \neq 0$. Then by the Inverse Mapping Theorem there exists (after an eventual restriction of the domain of φ) the inverse mapping $(\Pi \circ \varphi)^{-1}$. The equation $g(x) = 0$, where $g(x) = x - \varphi\big((\Pi \circ \varphi)^{-1}(\Pi(x))\big)$, is the desired implicit description of Ω on the neighborhood of x.)

Let Ω be a k-dimensional surface and σ a k-dimensional measure on $\mathbf{R}^n$. The integral $\int_\Omega f \, \mathrm{d}\sigma$ is sometimes labelled as the *curve or surface integral* of f.

34.25. Example (helix). Let $X = \{[x, y, z] \in \mathbf{R}^3 : x = \cos z,\ y = \sin z\}$. We may parametrize X by the mapping

$$\varphi(t) = [\cos t,\ \sin t,\ t]\,, \qquad t \in \mathbf{R}.$$

Then

$$\varphi'(t) = [-\sin t,\ \cos t,\ 1]$$

and

$$\operatorname{vol} \varphi'(t) = \sqrt{(-\sin t)^2 + \cos^2 t + 1} = \sqrt{2}.$$

34.26. Example (sphere). Let $S = \{[x, y, z] \in \mathbf{R}^3 : x^2 + y^2 + z^2 = 1\}$. We introduce three possibilities of a parametrization.

(a) Remember that the spherical coordinates are given by formulas $\varphi_s = [x, y, z]$, where

$$\begin{aligned} x(t, a) &= \cos a\, \cos t, \\ y(t, a) &= \cos a\, \sin t, \\ z(t, a) &= \sin a\,, \end{aligned}$$

$[t, a] \in G := (0, 2\pi) \times (-\frac{1}{2}\pi,\ \frac{1}{2}\pi)$. Then $\varphi(G) = X := S \setminus N$, where N is the "meridian" $\{[x, y, z] \in S\colon y = 0, x \in [0, 1]\}$. Obviously, two-dimensional measure of N is zero, so that the difference between S and X may be neglected. We have

$$\varphi_s'(t, a) = \begin{pmatrix} -\sin t \cos a, & -\cos t \sin a \\ \cos t \cos a, & -\sin t \sin a \\ 0, & \cos a \end{pmatrix}$$

and

$$\operatorname{vol} \varphi_s'(t, a) = \sqrt{\cos^2 a \sin^2 a + \sin^2 t \cos^4 a + \cos^2 t \cos^4 a} = \cos a.$$

(b) Let S, X, N be as in (a). For the parametrization of X we may also use the "cylindrical coordinates" $\varphi_c = [x, y, z]$, where

$$\begin{aligned} x(t, h) &= r \cos t, \\ y(t, h) &= r \sin t, \qquad (r \text{ denotes } \sqrt{1 - h^2}), \\ z(t, h) &= h \end{aligned}$$

on $G := \{[t, h] \in \mathbf{R}^2 : t \in (0, 2\pi),\ h \in (-1, 1)\}$. Then

$$\varphi_c'(t, h) = \begin{pmatrix} -r \sin t, & -\frac{h}{r} \cos t \\ r \cos t, & -\frac{h}{r} \sin t \\ 0, & 1 \end{pmatrix}$$

and

$$\operatorname{vol} \varphi_c'(t, h) = \sqrt{h^2 + (1 - h^2)(\sin^2 t + \cos^2 t)} = 1.$$

(c) Let S_+ be the hemisphere $\{[x,y,z] \in S : z > 0\}$. Then S_+ can be parametrized by its projection into the plane determined by the axes x and y. Consider the parametrization $\varphi_p = [x,y,z]$, where

$$\begin{aligned} x(s,t) &= s, \\ y(s,t) &= t, \\ z(s,t) &= \sqrt{1-s^2-t^2}, \end{aligned}$$

on $\{(s,t) \in \mathbf{R}^2 \colon s^2+t^2 < 1\}$. Then

$$\varphi_p'(s,t) = \begin{pmatrix} 1, & 0 \\ 0, & 1 \\ \frac{-s}{\sqrt{1-s^2-t^2}}, & \frac{-t}{\sqrt{1-s^2-t^2}} \end{pmatrix}$$

and

$$\operatorname{vol} \varphi_p'(t) = \frac{1}{\sqrt{1-s^2-t^2}}.$$

34.27. Example. By means of the parametrization of Example 34.26.b we will compute two-dimensional surface measure of the unit sphere. Let S, X, G have the same meaning as in the quoted example. Then

$$\sigma(S) = \sigma(X) = \int_G \operatorname{vol} \varphi' \,\mathrm{d}t \,\mathrm{d}h = \int_{-1}^1 (\int_0^{2\pi} \mathrm{d}t) dh = 4\pi.$$

34.28. Example. Let σ be a two-dimensional measure on $S_- := \{(x,y,z) \in \mathbf{R}^3 \colon z = -\sqrt{1-x^2-y^2}\}$. The integral

$$\int_{S_-} z^2 \,\mathrm{d}\sigma$$

posseses (up to constants) the physical meaning of the force by which (under the unit gravitation acceleration) a ball with the unit density half immersed in a liquid is lifted (computed by the integration of the gravitation forces).

Use spherical coordinates on $G = \{(t,a) \in \mathbf{R}^2 : t \in (0,2\pi),\ a \in (-\pi/2, 0)\}$. Then

$$\int_{S_-} z^2 \,\mathrm{d}\sigma = \int_G \sin^2 a \cos a \,\mathrm{d}t \,\mathrm{d}a = 2\pi \int_{-\pi/2}^0 \sin^2 a \cos a \,\mathrm{d}a = \frac{2}{3}\pi\,,$$

which is the half of the volume of the unit sphere. Therefore, we have verified Archimedes principle in our particular case.

34.29. Example (the length of the graph of a function). Let Γ be the graph of a Lipschitz function $f : [0,1] \to \mathbf{R}$ and σ a 1-dimensional measure on $\mathbf{R}^2$. Let $G = (0,1)$ and $\varphi(t) = (t, f(t))$, $t \in G$. Then the assumptions of Theorem 34.18 are satisfied and

$$\sigma(\Gamma) = \int_0^1 \operatorname{vol} \varphi'(t) \,\mathrm{d}t = \int_0^1 \sqrt{1+(f'(t))^2} \,\mathrm{d}t.$$

34.30. Exercise. Derive the formula

$$\sigma(\Omega) = 2\pi \int_a^b f(t)\sqrt{1+(f'(t))^2} \,\mathrm{d}t$$

in the case of two-dimensional measure of the surface Ω drawn in $\mathbf{R}^3$ by the rotation of the graph of a function f around the z-axis. Suppose that the function $f \colon (a,b) \to (0,+\infty)$ is Lipschitz and set

$$\Omega = \{[x,y,z] \in \mathbf{R}^3 \colon z \in (a,b),\ \sqrt{x^2+y^2} = f(z)\}.$$

34.31. Rectifiable Sets. A set $H \subset \mathbf{R}^n$ is said to be *k-rectifiable* if there is a Lipschitz mapping of a bounded measurable set $E \subset \mathbf{R}^k$ to H. It is a consequence of the Change of Variable Formula 34.19 that on the σ-algebra generated by k-rectifiable sets all k-dimensional measures coincide. On the other hand, there exist closed sets on which k-dimensional measures may differ (for example, the k-dimensional measure of 34.9 may be different from the Hausdorff measure of Chapter 36). The constructions of such sets are rather complicated.

Notice that n-rectificable sets in $\mathbf{R}^n$ are just bounded Lebesgue measurable sets.

34.32. Notes. The concept of k-dimensional measures comes from A. Kolmogorov [1932]. Our exposition follows the monograph by H. Federer [*1969] (see also L. C. Evans and R. E. Gariepy [*1992]) and ideas due to D. Preiss. The classical Sard theorem (or, Morse-Sard theorem) for $\mathscr{C}^1$–mappings was proved by A.P. Morse [1939] and A. Sard [1942].

35. The Degree of a Mapping

Let $G \subset \mathbf{R}^k$ be an open set and $\varphi\colon G \to \mathbf{R}^k$ a locally Lipschitz mapping. Then

$$\int_G |J_\varphi(t)| \,\mathrm{d}t = \int_{\varphi(G)} N(x, \varphi, G) \,\mathrm{d}x,$$

where $N(x, \varphi, G)$ is the number of points of the set $\varphi^{-1}(x)$ (the *Banach indicatrix*). It is natural to ask what is the meaning of the integral

$$\int_G J_\varphi(t) \,\mathrm{d}t.$$

This problem leads to the notion of the degree of a mapping which turns out to have many applications. Let us mention, for instance, its importance for the topology of Euclidean spaces and solvability problems for nonlinear "algebraic" equations.

First we need to prepare several auxiliary computational results.

35.1. Lemma. *Let $G \subset \mathbf{R}^k$ be an open set and $\psi_1, \dots, \psi_k$ twice continuously differentiable functions on G. Then we have*
(a)

$$\sum_{q=1}^k (-1)^{q+1} \frac{\partial}{\partial t_q} \det\Big(\frac{\partial \psi_i}{\partial t_j}\Big)_{\substack{i=2,\dots,k \\ j=1,\dots,q-1,q+1,\dots,k}} = 0.$$

(b)

$$\sum_{q=1}^k (-1)^{q+1} \frac{\partial}{\partial t_q} \Big(\psi_1 \det\Big(\frac{\partial \psi_i}{\partial t_j}\Big)_{\substack{i=2,\dots,k \\ j=1,\dots,q-1,q+1,\dots,k}}\Big) = \det(\nabla\psi_1, \dots, \nabla\psi_k).$$

Proof. (a) The left-hand side is equal to the sum

$$\sum_{q=1}^k \sum_{p=1}^k (-1)^{q+1} \det(a_{i,j,p,q})_{\substack{i=2,\dots,k \\ j=1,\dots,q-1,q+1,\dots,k}},$$

where

$$a_{i,j,p,q} = \begin{cases} \frac{\partial \psi_i}{\partial t_j} & \text{if } j \neq p \neq q\,, \\ \frac{\partial^2 \psi_i}{\partial t_q \partial t_j} & \text{if } j = p \neq q\,, \\ 0 & \text{if } p = q. \end{cases}$$

Now it remains to realize that using the interchange of the order of differentiation we get

$$(-1)^{q+1}\det(a_{i,j,p,q})_{\substack{i=2,\dots,k\\ j=1,\dots,q-1,q+1,\dots,k}} = (-1)^{p}\det(a_{i,j,q,p})_{\substack{i=2,\dots,k\\ j=1,\dots,p-1,p+1,\dots,k}}$$

for all p, q.

(b) If we apply the chain rule to the left-hand side of the equality, we obtain

$$\begin{aligned}
&\sum_{q=1}^{k}(-1)^{q+1}\frac{\partial}{\partial t_q}\Big(\psi_1 \det\big(\frac{\partial\psi_i}{\partial t_j}\big)_{\substack{i=2,\dots,k\\ j=1,\dots,q-1,q+1,\dots,k}}\Big)\\
&=\sum_{q=1}^{k}(-1)^{q+1}\frac{\partial\psi_1}{\partial t_q}\det\big(\frac{\partial\psi_i}{\partial t_j}\big)_{\substack{i=2,\dots,k\\ j=1,\dots,q-1,q+1,\dots,k}}\\
&+\psi_1\sum_{q=1}^{k}(-1)^{q+1}\frac{\partial}{\partial t_q}\det\big(\frac{\partial\psi_i}{\partial t_j}\big)_{\substack{i=2,\dots,k\\ j=1,\dots,q-1,q+1,\dots,k}}.
\end{aligned}$$

The first term on the right-hand side is just the expansion of the determinant $\det(\nabla\psi_1,\dots,\nabla\psi_k)$ by the first row, the second one vanishes by the part (a). ■

35.2. Lemma. *Let $\psi_1,\dots,\psi_k$ be Lipschitz functions on $\mathbf{R}^k$ vanishing outside a compact subset of $\mathbf{R}^k$. Then*

$$\int_{\mathbf{R}^k}\det(\nabla\psi_1,\dots,\nabla\psi_k)\,\mathrm{d}t = 0.$$

Proof. We proceed in two steps. First, we assume that the functions $\psi_1\dots\psi_k$ are of class $\mathscr{C}^2$. By Lemma 35.1.b there are functions $\eta_1,\dots,\eta_k$ vanishing outside compact subsets of $\mathbf{R}^k$ such that

$$\det(\nabla\psi_1,\dots,\nabla\psi_k) = \sum_{j=1}^{k}\frac{\partial\eta_j}{\partial t_j}.$$

For each $j=1,\dots,k$ and $[t_1,\dots,t_{j-1},t_{j+1},\dots,t_k]\in\mathbf{R}^{k-1}$ we have

$$\int_{\mathbf{R}}\frac{\partial\eta_j}{\partial t_j}(t_1,\dots,t_{j-1},\xi,t_{j+1},\dots,t_k)\,\mathrm{d}\xi = 0.$$

Fubini's theorem yields

$$\int_{\mathbf{R}^k}\frac{\partial\eta_j}{\partial t_j}(t)\,\mathrm{d}t = 0.$$

Summing up these equalities over $j=1,\dots,k$ we get the required formula.

In the second step we show that the formula is valid without smoothness assumptions. Suppose that all functions $\psi_1,\dots,\psi_k$ are β-Lipschitz. Then

$$\int_{\mathbf{R}^k}\det(\nabla\chi_q*\psi_1,\dots,\nabla\chi_q*\psi_k)\,\mathrm{d}t = 0$$

(notation as in 31.1). Letting $q\to\infty$, by the Lebesgue dominated convergence theorem with constant dominating function we get the desired equality. ■

35.3. Corollary. *Let $G \subset \mathbf{R}^k$ be a bounded open set and $\varphi_1, \dots, \varphi_k$, $\psi_1, \dots, \psi_k$ Lipschitz functions on $\overline{G}$. If $\varphi_i = \psi_i$ on ∂G, $i = 1, \dots, k$, then*

$$\int_G \det(\nabla\varphi_1, \dots, \nabla\varphi_k)\,\mathrm{d}t = \int_G \det(\nabla\psi_1, \dots, \nabla\psi_k)\,\mathrm{d}t.$$

Proof. By McShane's theorem 30.5 the functions φ_i, ψ_i can be extended to the whole of $\mathbf{R}^k$ to be Lipschitz functions of compact support. Further, we redefine ψ_i on $\mathbf{R}^k \setminus G$ to coincide there with φ_i. Thanks to the preceding lemma it follows that

$$\begin{aligned}\int_G \det(\nabla\varphi_1, \dots, \nabla\varphi_k)\,\mathrm{d}t &= -\int_{\mathbf{R}^k\setminus G} \det(\nabla\varphi_1, \dots, \nabla\varphi_k)\,\mathrm{d}t \\ &= -\int_{\mathbf{R}^k\setminus G} \det(\nabla\psi_1, \dots, \nabla\psi_k)\,\mathrm{d}t \\ &= \int_G \det(\nabla\psi_1, \dots, \nabla\psi_k)\,\mathrm{d}t.\end{aligned}$$

■

35.4. Lemma. *Let $G \subset \mathbf{R}^k$ be a bounded open set, f an integrable function on $\mathbf{R}^k$ and φ, ψ Lipschitz mappings of $\overline{G}$ into $\mathbf{R}^k$. If $\varphi = \psi$ on ∂G, then*

$$\int_G f(\varphi(t))\,J_\varphi(t)\,\mathrm{d}t = \int_G f(\psi(t))\,J_\psi(t)\,\mathrm{d}t.$$

Proof. Let $\varphi = [\varphi_1, \dots, \varphi_k]$, $\psi = [\psi_1, \dots, \psi_k]$. No generality is lost with the assumption that $f \in \mathscr{D}(\mathbf{R}^k)$, for $\mathscr{D}(\mathbf{R}^k)$ is dense in $L^1(\mathbf{R}^k)$. Find a $\mathscr{C}^1$-function g on $\mathbf{R}^k$ so that $\frac{\partial g}{\partial x_1} = f$. Then by Corollary 35.3

$$\int_G \det(\nabla(g\circ\varphi), \nabla\varphi_2, \dots, \nabla\varphi_k)\,\mathrm{d}t = \int_G \det(\nabla(g\circ\psi), \nabla\psi_2, \dots, \nabla\psi_k)\,\mathrm{d}t.$$

Further,

$$\int_G \det(\nabla(g\circ\varphi), \nabla\varphi_2, \dots, \nabla\varphi_k)\,\mathrm{d}t = \sum_{i=1}^k \int_G \Big(\frac{\partial g}{\partial x_i}\circ\varphi\Big)\,\det(\nabla\varphi_i, \nabla\varphi_2, \dots, \nabla\varphi_k)\,\mathrm{d}t.$$

Among the terms of the sum on the right, only the first one may differ from zero and it equals $\int_G (f\circ\varphi)\,J_\varphi\,\mathrm{d}t$. If we combine this result with an analogous computation for ψ_i, we obtain the desired equality. ■

35.5. Lemma. *Let $G \subset \mathbf{R}^k$ be a bounded open set and $\varphi, \psi \colon \overline{G} \to \mathbf{R}^k$ Lipschitz mappings. Let $B(y, r)$ be a closed ball of $\mathbf{R}^k$ which does not intersect any of the segments joining $\varphi(t)$ and $\psi(t)$, $t \in \partial G$. Let f be an integrable function on $\mathbf{R}^k$ with support in $U(y, r)$. Then*

$$\int_G f(\varphi(t))\,J_\varphi(t)\,\mathrm{d}t = \int_G f(\psi(t))\,J_\psi(t)\,\mathrm{d}t.$$

Proof. Denote

$$K = \{t \in \overline{G} : \text{the segment joining } \varphi(t) \text{ and } \psi(t) \text{ intersects } B(y,r)\}.$$

Then K is a compact set contained in G. The function η which is one on K and zero outside G is Lipschitz on $K \cup (\mathbf{R}^k \setminus G)$. According to McShane's extension theorem 30.5, η can be extended (under the same notation) as a Lipschitz function on $\mathbf{R}^k$. Set $\zeta = \varphi + \eta(\psi - \varphi)$. Then $\zeta = \varphi$ on ∂G and $f(\zeta(t)) = f(\psi(t))$ for all $t \in G$. and by Lemma 35.4 we obtain

$$\int_G f(\varphi(t))\, J_\varphi(t)\, \mathrm{d}t = \int_G f(\zeta(t))\, J_\zeta(t)\, \mathrm{d}t = \int_G f(\psi(t))\, J_\psi(t)\, \mathrm{d}t.$$

■

35.6. Lemma. *Let $G \subset \mathbf{R}^k$ be a bounded open set and $\varphi : \overline{G} \to \mathbf{R}^k$ a Lipschitz mapping. If $U \subset \mathbf{R}^k$ is a connected open set disjoint with $\varphi(\partial G)$, then there exists $a \in \mathbf{R}$ such that*

$$\int_G f(\varphi(t))\, J_\varphi(t)\, \mathrm{d}t = a \int_{\mathbf{R}^k} f(x)\, \mathrm{d}x$$

for every function $f \in \mathscr{L}^1(\mathbf{R}^k)$ with support in U.

Proof. To prove the assertion it suffices to show it when U is an open cube, and it is enough to verify the equality for indicator functions of cubes $K(z,\rho) := [-\rho,\rho]^k + z$ contained in U, where $z \in U$ and ρ is a rational number. Select ρ and set $U_\rho := \{z \in U : K(z,\rho) \subset U\}$. Then for $z, y \in U_\rho$, we have

$$\int_G c_{K(y,\rho)}(\varphi(t))\, J_\varphi(t)\, \mathrm{d}t = \int_G c_{K(z,\rho)}(\psi(t))\, J_\psi(t)\, \mathrm{d}t\,,$$

where $\psi(t) = \varphi(t) + y - z$. For y close enough to z the functions φ, ψ satisfy the hypotheses of Lemma 35.5 and, in view of connectedness of U_ρ, it follows that the function

$$y \mapsto \int_G c_{K(y,\rho)}(\varphi(t))\, J_\varphi(t)\, \mathrm{d}t$$

is constant on U_ρ. Hence

$$\int_G c_{K(y,\rho)}(\varphi(t))\, J_\varphi(t)\, \mathrm{d}t = a(\rho)\lambda K(y,\rho).$$

Since the Jacobian of a Lipschitz mapping is a bounded function and G is a bounded set, the constant $a(\rho)$ is finite. If ρ_1 is a rational number and ρ_2 is an integer multiple of ρ_1, then the cube with edges $2\rho_2$ may be filled by cubes of edge $2\rho_1$, so that $a(\rho_2) = a(\rho_1)$. Therefore we can conclude that $a(\rho)$ does not depend on ρ. ■

35.7. Degree of a Mapping. Let $G \subset \mathbf{R}^k$ be a bounded open set and $\varphi : \overline{G} \to \mathbf{R}^k$ a continuous mapping. By Tietze's extension theorem, φ is continuously extendable as a continuous function of compact support to the whole of $\mathbf{R}^k$ (with the same notation). By Theorem 31.5 there exist (even $\mathscr{C}^\infty$) functions φ_j on $\mathbf{R}^k$ such that $\varphi_j \rightrightarrows \varphi$. Let $y \in \mathbf{R}^k \setminus \varphi(\partial G)$. We find a neighborhood U of the point y whose closure does not intersect $\varphi(\partial G)$. We may assume that $\overline{U} \cap \varphi_j(\partial G) = \emptyset$ for all j. Using the preceding lemmas, there are finite constants a_j such that

$$\int_G f(\varphi_j(t))\, J_{\varphi_j}(t)\, \mathrm{d}t = a_j \int_{\mathbf{R}^k} f(x)\, \mathrm{d}x$$

for each integrable function f on $\mathbf{R}^n$ with support in U. Further, an appeal to Lemma 35.5 makes it clear that the sequence $\{a_j\}$ is constant for $j \geq j_0$, and that its limit a is independent of the choice of the sequence $\{\varphi_j\}$. We will soon be able to show that a is an integer and we call a to be the *degree of the mapping* φ at y on G. We denote it by $\deg(y, \varphi, G)$.

35.8. Change of Variable Formula. *Let $G \subset \mathbf{R}^k$ be a bounded open set and $\varphi \colon \overline{G} \to \mathbf{R}^k$ a Lipschitz mapping. Let f be a measurable function on $\mathbf{R}^k$, $f = 0$ almost everywhere on $\varphi(\partial G)$. Then*

$$\int_G f(\varphi(t))\, J_\varphi(t)\, \mathrm{d}t = \int_{\mathbf{R}^k} \deg(x, \varphi, G)\, f(x)\, \mathrm{d}x\,,$$

provided the integral on the left-hand side converges.

Proof. Making similar measurability considerations to those of the proof of Corollary 34.19, we conclude the theorem using the preceding lemmas and the definition of a degree. ■

35.9. Properties of a Degree. *Let $G \subset \mathbf{R}^k$ be a bounded open set and $\varphi \colon \overline{G} \to \mathbf{R}^k$ a continuous mapping.*

(a) *The function $y \mapsto \deg(y, f, G)$ is constant on each component of $\mathbf{R}^k \setminus f(\partial G)$.*
(b) *Let $\psi \colon \overline{G} \to \mathbf{R}^k$ be a continuous mapping and let the segment joining $\varphi(t)$ and $\psi(t)$, $t \in \partial G$, do not contain y. Then $\deg(y, \varphi, G) = \deg(y, \psi, G)$.*
(c) *If φ is one-to-one and $J_\varphi > 0$ almost everywhere in G, then $\deg(y, \varphi, G) = 1$ for all $y \in \varphi(G)$.*
(d) *If $\deg(y, \varphi, G) \neq 0$, then the equation $\varphi(t) = y$ has a solution in G.*
(e) *If $G_1, \dots, G_m$ are disjoint open subsets of G and $\varphi(t) \neq y$ for $t \in G \setminus (G_1 \cup \dots \cup G_m)$, then*

$$\deg(y, \varphi, G) = \deg(y, \varphi, G_1) + \dots + \deg(y, \varphi, G_m).$$

(f) *If φ is $\mathscr{C}^1$ and $J_\varphi(t) \neq 0$ for all $t \in \varphi^{-1}(y)$, then*

$$\deg(y, \varphi, G) = \sum \{\operatorname{sign} J_\varphi(t) \colon t \in \varphi^{-1}(y)\}\,.$$

(g) *The degree $\deg(y, \varphi, G)$ is an integer number.*

Proof. The assertions (a) and (b) are obvious. Combining Change of Variable Formula in 35.8 and 34.19, we get (c).

(d) Assume that the equation $\varphi(t) = y$ has no solution. Since $\varphi(\overline{G})$ is a compact set, there is an open neighborhood U of y such that $\overline{U} \cap \varphi(\overline{G}) = \emptyset$. Find a Lipschitz function ψ close enough to φ such that $\overline{U} \cap \psi(\overline{G}) = \emptyset$ and $\deg(y, \psi, G) \neq 0$. Then $G \cap \psi^{-1}(U) = \emptyset$ and by 35.8

$$0 = \int_{G \cap \psi^{-1}(U)} J_\psi = \deg(y, \psi, G)\, \lambda U \neq 0\,,$$

which is a contradiction.

(e) We may assume that φ is a Lipschitz mapping. The set $K := \varphi(\overline{G} \setminus (G_1 \cup \cdots \cup G_m))$ is compact and does not contain y. We find an open neighborhood U of y whose closure does not intersect K. Then by 35.8

$$\begin{aligned} \deg(y, \varphi, G) &= \int_{G \cap \varphi^{-1}(U)} J_\varphi = \int_{G_1 \cap \varphi^{-1}(U)} J_\varphi + \cdots + \int_{G_m \cap \varphi^{-1}(U)} J_\varphi \\ &= \deg(y, \varphi, G_1) + \cdots + \deg(y, \varphi, G_m). \end{aligned}$$

(f) Let $s \in \varphi^{-1}(y)$. Since $J_\varphi(s) \neq 0$, there is a neighborhood V of s such that φ is one-to-one on V, $\varphi(V)$ is an open set (the Inverse Mapping Theorem) and J_φ has a constant sign on V. Then

$$\lambda\varphi(V)\, \deg(y, \varphi, V) = \int_V J_\varphi \,\mathrm{d}t = \operatorname{sign} J_\varphi(s) \int_V |J_\varphi| \,\mathrm{d}t = \operatorname{sign} J_\varphi(s)\, \lambda\varphi(V)\,.$$

It is also clear that the set $\varphi^{-1}(y)$ is isolated, and therefore $\varphi^{-1}(y)$ is a finite subset of $\overline{G}$. We complete the proof using (e).

(g) We can assume that φ is of class $\mathscr{C}^1$. Let $E = \{t \in G \colon J_\varphi(t) = 0\}$. By the Sard Theorem 34.17, $\lambda(\varphi(E)) = 0$. Choose $y \in \mathbf{R}^k \setminus \varphi(\partial G)$. Let U be a neighborhood of y which does not intersect $\varphi(\partial G)$. Then there is $x \in U \setminus \varphi(E)$. Since $\deg(y, \varphi, G) = \deg(x, \varphi, G)$, according to (f) $\deg(y, \varphi, G)$ is an integer. ■

35.10. Open Mapping Theorem. *Let $\varphi \colon G_0 \to \mathbf{R}^k$ be a bilipschitz mapping on an open set $G_0 \subset \mathbf{R}^k$. Then $\varphi(G_0)$ is an open set.*

Proof. Let $t \in G_0$ and $y = \varphi(t)$. Let G be a bounded open set, $t \in G \subset \overline{G} \subset G_0$, and U be an open neighborhood of y which does not intersect $\varphi(\partial G)$. If we prove that $\deg(y, \varphi, G) \neq 0$, then by 35.9.d the equation $\varphi(s) = x$ has a solution for any $x \in U$. Now, there is an open neighborhood V of t such that $\varphi(V) \subset U$. Denote $V^+ = \{s \in V \colon J_{\varphi(s) > 0}\}$ and $V^- = \{s \in V \colon J_{\varphi(s) < 0}\}$. Notice that by the Rademacher Theorem 30.3, φ' exists almost everywhere. From the definition of a bilipschitz mapping it is clear that $J_\varphi(t) = 0$ is impossible, and at least one of the sets V^+, V^- contains a compact set K of positive measure. We obtain $0 \neq \int_K J_\varphi = \deg(y, \varphi, G)\, \lambda\varphi(K)$. Hence $\deg(y, \varphi, G) \neq 0$. ■

35.11. Theorem on Orientation. *Let φ be a locally bilipschitz mapping of a connected open set $G_0 \subset \mathbf{R}^k$ into $\mathbf{R}^k$. Then $J_\varphi > 0$ almost everywhere in G_0, or $J_\varphi < 0$ almost everywhere in G_0.*

Proof. Let $G \subset \overline{G} \subset G_0$ be a bounded open set and $t \in G$. There is a neighborhood U of $\varphi(t)$ which does not intersect $\varphi(\partial G)$, and a neighborhood V of t such that $\varphi(V) \subset U$. As in the proof of the Open Mapping Theorem we deduce that $\deg(\varphi(t), \varphi, G) \neq 0$. By 35.9.g, the degree is an integer, so that $|\deg(\varphi(t), \varphi, G)| \geq 1$. We have

$$\Big|\int_V J_\varphi(t)\,\mathrm{d}t\Big| = |\deg(\varphi(t), \varphi, G)|\ \lambda\varphi(V) \geq \lambda\varphi(V) = \int_V |J_\varphi(t)|\ \mathrm{d}t.$$

Thus, J_φ has a constant sign almost everywhere in V. From the connectedness of G_0 we obtain the assertion. ∎

35.12. Notes. Although the origins of the idea of degree go back to K.F. Gauss and A.L. Cauchy, for smooth mappings and smooth sets they were considered at the turn of the century by H. Kronecker, H. Poincaré, E. Picard, P. Bohl or J. Hadamard. The essential step for the developing the theory of degree of mappings in finite dimensional spaces and its application is due to L.E.J. Brouwer [1912]. Later on, J. Leray and J. Schauder introduced the topological degree also in infinite-dimensional spaces. Today, the significance of a degree, mainly in nonlinear functional analysis, is undeniable. There is a rich bibliography and many sources for study the degree theory. We refer the reader e.g. to S. Fučík and J. Milota [FM], J. Stará and O. John [SJ], J.T. Schwartz [*1969], S. Fučík, J. Nečas, J. Souček and V. Souček [*1973], K. Deimling [*1985], I. Fonseca and W. Gangbo [*1995] and P. Drábek and J. Milota [*2004].

36. Hausdorff Measures

In Chapter 34 we proved the existence of a k-dimensional measures on $\mathbf{R}^n$ using a relatively simple method. This approach to k-dimensional measures is appropriate for purposes of applications to the curve and surface integrals.

In theoretical parts of modern analysis we encounter Hausdorff measures occuring more frequently in various connections (even for noninteger values of k).

According to 34.31, on rectifiable sets, and in particular on k-dimensional surfaces, the normalized Hausdorff measure and the measure constructed in 34.9 coincide.

Without loss of clarity we take a slightly deeper look in a more general setting supposing that p (the "dimension") is a nonnegative real number and (P, ρ) is a metric space on which we are going to construct the p-dimensional Hausdorff measure.

36.1. Outer Hausdorff Measure. Let $A \subset P$. Denote

$$\mathscr{H}_p(A, \delta) = \inf\Big\{\sum_{j=1}^{\infty}(\operatorname{diam} A_j)^p : \bigcup_{j=1}^{\infty} A_j \supset A, \operatorname{diam} A_j \leq \delta\Big\} \qquad \text{for} \qquad \delta > 0,$$

$$\mathscr{H}_p(A) = \sup_{\delta>0} \mathscr{H}_p(A, \delta) \quad (= \lim_{\delta\to 0+} \mathscr{H}_p(A, \delta)\ ).$$

The set function $A \mapsto \mathscr{H}_p(A)$ is called the *p-dimensional (outer) Hausdorff measure*.

As we show later, if $P = \mathbf{R}^n$ and $k \leq n$ is a nonnegative integer, there is a constant κ_k such that $\mathscr{H}_k/\kappa_k$ is a k-dimensional measure on $\mathbf{R}^n$ in the sense of definition 34.8. The measure $\mathscr{H}_k/\kappa_k$ is called the *normalized Hausdorff measure*.

36.2. Metric Outer Measure. An outer measure γ on P is called a *metric outer measure* if $\gamma(A \cup B) = \gamma A + \gamma B$ whenever $A, B \subset P$ and $\inf\{\rho(x,y) : x \in A,\ y \in B\} > 0$.

36.3. Remarks. 1. In the definition of the Hausdorff measure $\mathscr{H}_p(A)$ we considered arbitrary coverings of A but we may confine to coverings consisting of closed or open sets.

2. Notice that the n-dimensional Hausdorff measure of a set K which is a ball or a cube in $\mathbf{R}^n$ equals $c(\operatorname{diam} K)^n$ for a suitable constant c (cf. Lemma 36.12).

3. The set functions $\mathscr{H}_p(\cdot, \delta)$ are not metric outer measures and cannot be used to describe length or area. Notice that the "one-dimensional" measure $\mathscr{H}_1(K, \delta)$ of the unit square K in $\mathbf{R}^2$ is a finite number although K contains infinitely many segments of length one. This is why we cannot replace the definition of Hausdorff measure by a more simple formula

$$A \mapsto \inf\{\sum_{j=1}^{\infty} (\operatorname{diam} A_j)^p : \bigcup_{j=1}^{\infty} A_j \supset A\}$$

(which is in fact $\mathscr{H}_p(A, \infty)$).

4. Further examples of k-dimensional measures in $\mathbf{R}^n$ (which, in general, do not coincide with the normalized Hausdorff measure) can be obtained using other covering families. For example, the *spherical* measure is defined using coverings formed by open balls, see H. Federer [*1969].

The next series of theorems shows that $\mathscr{H}_p$ is a metric outer measure. Therefore, we can apply Carathodory's method and to derive that each Borel set is $\mathscr{H}_p$-measurable, and that the restriction of $\mathscr{H}_p$ to the Borel σ-algebra is a measure.

36.4. Theorem. *Let γ be an metric outer measure on P. Then each Borel subset of P is γ-measurable.*

Proof. It would be clearly sufficient to prove that closed sets are γ-measurable. To this end let a closed set $F \subset P$ be given. Choose a test set $T \subset P$, $\gamma T < +\infty$ and denote

$$P_j = \{x \in T : \frac{1}{j+1} \le \operatorname{dist}(x, F) < \frac{1}{j}\}, \quad j = 1, 2, \dots,$$
$$P_0 = \{x \in T : \operatorname{dist}(x, F) \ge 1\}.$$

Then the sets P_0, P_2, $P_4, \dots$ have positive distances, so that

$$\sum_{j=0}^{m} \gamma P_{2j} = \gamma\Big(\bigcup_{j=0}^{m} P_{2j}\Big) \le \gamma T$$

for all $m \in \mathbf{N}$. Similarly $\sum_{j=0}^{m} \gamma P_{2j+1} \le \gamma T$, and we see that the series $\sum_{j=0}^{\infty} \gamma P_j$ is convergent. Since for each $m \in \mathbf{N}$, the distance between $\bigcup_{j=0}^{m} P_j$ and $T \cap F$ is positive we have

$$\gamma(T \setminus F) \le \gamma\Big(\bigcup_{j=0}^{m} P_j\Big) + \gamma\Big(\bigcup_{j=m+1}^{\infty} P_j\Big) \le \gamma T - \gamma(T \cap F) + \sum_{j=m+1}^{\infty} \gamma P_j.$$

Letting $m \to \infty$ we obtain

$$\gamma(T \setminus F) \le \gamma(T) - \gamma(T \cap F).$$

■

36.5. Remark. If γ is an outer measure on P for which every Borel set is γ-measurable, then γ is already a metric outer measure.

36.6. Theorem. *$\mathscr{H}_p$ is a metric outer measure on P.*

Proof. Theorem 4.3 tell us that $A \mapsto \mathscr{H}_p(A, \delta)$ is an outer measure for every $\delta > 0$. Letting $\delta \to 0$ we see that $\mathscr{H}_p$ is an outer measure. Let A, B be sets of positive distance and $\delta_0 < \inf\{\rho(x, y) : x \in A, y \in B\}$. Let $M \subset A \cup B$, $\operatorname{diam} M \le \delta_0$. Then either $M \subset A$ or $M \subset B$. Hence

$$\mathscr{H}_p(A \cup B, \delta) = \mathscr{H}_p(A, \delta) + \mathscr{H}_p(B, \delta)$$

for all $\delta \in (0, \delta_0)$. Thus

$$\mathscr{H}_p(A \cup B) = \mathscr{H}_p(A) + \mathscr{H}_p(B).$$

■

36.7. Corollary. *Any Borel subset of P is $\mathscr{H}_p$-measurable.*

36.8. Exercise. (a) Let $0 \le p < q$. If $\mathscr{H}_p(A) < +\infty$, then $\mathscr{H}_q(A) = 0$.

(b) The number $\inf\{p \ge 0 : \mathscr{H}_p(A) = 0\}$ is called the *Hausdorff dimension* of a set A. Compute the Hausdorff dimension of the Cantor set in $[0, 1]$.

(c) For any $p \in (0, 1)$, a "generalized Cantor set" $B \subset [0, 1]$ may be constructed with $\mathscr{H}_p(B) = 1$.

36.9. Remark. The definition of the p-dimensional Hausdorff measure admits also noninteger values of p. The Hausdorff measures with noninteger dimensions are not directly linked with the topics of the following chapters nevertheless they have a great importance e.g. in the theory of singular sets (to describe the "size" of "negligible" sets, for example the set of discontinuities for a solution of system of partial differential equations, or the set of points of divergence of a Fourier series; the applications to "removable singularities" are also frequent). The concept of Hausdorff measures with noninteger dimension is also a starting point for the famous fractal theory (see e.g G.A. Edgar [*1990] and K.J. Falconer [*1985]).

36.10. Exercise. Show that each 0-dimensional Hausdorff measure is the counting measure.

36.11. Theorem. *Let P and P' be metric spaces, $\mathscr{H}_p$ the p-dimensional Hausdorff measure on P and $\mathscr{H}'_p$ the p-dimensional Hausdorff measure on P'. Let E be a subset of P and $f : E \to P'$ a β-Lipschitz mapping. Then*

$$\mathscr{H}'_p(f(E)) \le \beta^p \mathscr{H}_p(E).$$

Proof. For each $\delta > 0$ and each sequence $\{A_j\}$ of subsets of P with $\bigcup_{j=1}^{\infty} A_j \supset E$, $\operatorname{diam} A_j < \delta$ we have

$$\mathscr{H}_p(f(E), \beta\delta) \le \sum_{j=1}^{\infty} (\operatorname{diam} f(A_j))^p \le \beta^p \sum_{j=1}^{\infty} (\operatorname{diam} A_j)^p.$$

Hence the desired inequality easily follows. ■

36.12. Lemma. *Let $\mathscr{H}_k$ be the k-dimensional Hausdorff measure on $\mathbf{R}^n$, $K = [0,1]^k \times \{0\}^{n-k}$. Then $0 < \mathscr{H}_k(K) < +\infty$.*

Proof. Given $\delta > 0$ there is $m \in \mathbf{N}$ such that $\dfrac{\sqrt{k}}{m} < \delta$. We divide K to m^k cubes K_j with edges $\dfrac{1}{m}$ and diameters $\dfrac{\sqrt{k}}{m} < \delta$. Then

$$\mathscr{H}_k(K, \delta) \le \sum_{j=1}^{m^k} (\operatorname{diam} K_j)^k = m^k \Big(\frac{\sqrt{k}}{m}\Big)^k = k^{k/2}.$$

Hence $\mathscr{H}_k(K) \le k^{k/2}$.

Conversely, let λ be the "Lebesgue measure" on K. Let $A \subset K$ and $x \in A$. Then

$$\begin{aligned} A &\subset B(x, \operatorname{diam} A) \\ &\subset \big[x_1 - \operatorname{diam} A, x_1 + \operatorname{diam} A\big] \times \cdots \times \big[x_k - \operatorname{diam} A, x_k + \operatorname{diam} A\big], \end{aligned}$$

thus

$$\lambda A \le 2^k (\operatorname{diam} A)^k.$$

If $\{A_j\}$ is a sequence of subsets of K, $\bigcup_{j=1}^{\infty} A_j = K$, then

$$\sum_{j=1}^{\infty} (\operatorname{diam} A_j)^k \ge 2^{-k} \sum_{j=1}^{\infty} \lambda A_j \ge 2^{-k} \lambda K = 2^{-k}.$$

Whence taking the infimum we obtain the desired lower estimate for $\mathscr{H}_k(K)$. ■

36.13. Normalized Hausdorff Measures on $\mathbf{R}^n$. Let k be a nonnegative integer, $k \le n$ and $\kappa_k := \mathscr{H}_k([0,1]^k \times 0^{n-k})$. Then $\mathscr{H}_k/\kappa_k$ is a k-dimensional measure on $\mathbf{R}^n$ which will be labelled as the normalized Hausdorff measure.

The constant κ_k is equal to the number $(4/\pi)^{k/2}\,\Gamma(1 + \frac{k}{2})$. The computation is not easy, see C.A. Rogers [*1970].

36.14. Remarks. 1. The hints for computation k-dimensional measures of concrete rectifiable sets are given in Chapter 34. We will not present the (difficult) examples of nonrectifiable sets.

2. Without essential changes of proofs a similar theory to that of this chapter can be built also for the case of spherical measures (cf. Remark 36.3.4). The computation of the corresponding constant analogous to κ_k is much easier, in fact it follows immediately from Exercise 26.26.

3. In terms of the Hausdorff measure the following version of the Sard theorem can be proved:

Let $G \subset \mathbf{R}^k$ be an open set and $f : G \to \mathbf{R}^n$ an arbitrary mapping. Let E be the set of all points $t \in G$ at which the derivative $f'(t)$ exists and $\operatorname{vol} f'(t) = 0$ *(which means that the rank of the matrix $f'(t)$ is less than k). Then $\mathscr{H}_k(f(E)) = 0$.*

36.15. Notes. C. Carathéodory developed in [1914] the theory of the one-dimensional ("linear") measure in n-dimensional Euclidean spaces. In the same paper he mentions the possibility of introducing k-dimensional measures in an n-dimensional space (for an integer k). The k-dimensional measure for arbitrary positive $k > 0$ on $\mathbf{R}^n$ was introduced by F. Hausdorff [1919]. The interesting Theorem 36.4 is due to C. Carathodory [*1918]. The theory of Hausdorff measures was developed very intensively, particularly A.S. Besikovitch published a great amount of papers devoted to this topic. From monographs on Hausdorff measures we recommend C.A. Rogers [*1970], K. J. Falconer [*1986] and P. Mattila [*1995].

G. Surface and Curve Integrals

37. Integral Calculus in Vector Analysis

In this chapter we continue a study of curve and surface integrals and state a change of variable formula. Moreover, we derive formulae concerning relations between the integration over the "interior" (of a set or a surface) and the integration over the "boundary". Later on we will see that these formulae are particular cases of general Stokes' Theorem of the next chapter. They are, in fact, multidimensional generalizations of the famous Leibniz formula

$$\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a)\,.$$

The reader perhaps appreciates that we include the explanation of three-dimensional (and therefore most important from the point of view of the classical physics) situations without a deeper excursion to multilinear algebra.

Recall, that the integral calculus on k-dimensional surfaces in $\mathbf{R}^n$ is based on the notion of a k-dimensional measure (34.8) which exists (Theorem 34.9) and on k-dimensional surfaces is uniquely determined (Remark 34.31). Underline that for a basic understanding of the topic it is not essential to know how k-dimensional measures were constructed.

Since all theorems of this chapter are only special cases of more general results of Chapter 38, we will not disturb the presentation by their proofs.

For the one-dimensional measure on $\mathbf{R}^n$ we will use the notation s, while S will be reserved for the $(n-1)$-dimensional measure on $\mathbf{R}^n$.

37.1. Vector Field, Gradient, Divergence, Curl. By a *vector field* on a set X we understand a mapping f of a set X into $\mathbf{R}^n$.

Let g be a function of class $\mathscr{C}^1$ on an open set $U \subset \mathbf{R}^n$. Then its *gradient* $\operatorname{grad} g$ on U is defined as the vector field $x \mapsto [\frac{\partial g}{\partial x_1}(x), \dots, \frac{\partial g}{\partial x_n}(x)]$. (The difference between the gradient and the derivative for functions of class $\mathscr{C}^1$ consists only in the convention that the gradient is a vector while the derivative is a linear form.)

Let $f = [f_1, \dots, f_n]$ be a continuously differentiable (i.e. of class $\mathscr{C}^1$) vector field on an open set $U \subset \mathbf{R}^n$. Then the *divergence* of the field f is the function $\operatorname{div} f$ on U defined as

$$\operatorname{div} f = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}.$$

If f is a continuously differentiable vector field on an open set $U \subset \mathbf{R}^2$, its *curl* is the function $\operatorname{curl} f$ defined by the formula

$$\operatorname{curl} f = \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}.$$

Finally, if f is a continuously differentiable vector field on an open set $U \subset \mathbf{R}^3$, we introduce its *curl* $\operatorname{curl} f$ as the vector field on U defined by

$$\operatorname{curl} f = \Big[\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}, \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}, \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\Big].$$

In higher-dimensional spaces the curl corresponds to the bilinear form $(u, v) \mapsto f'(x)u \cdot v - f'(x)v \cdot u$.

Let $\mathbf{V}$ be an n-dimensional vector space with an inner product. A choice of an orthonormal basis $(u_1, \dots, u_n)$ in $\mathbf{V}$ transfers the calculus in $\mathbf{V}$ into a calculus in $\mathbf{R}^n$. The coordinate mapping $L : \mathbf{V} \to \mathbf{R}^n$ assigns with each vector $x \in \mathbf{V}$ a vector $Lx \in \mathbf{R}^n$ of its coordinates with respect to $(u_1, \dots, u_n)$ in such a way that if $x = t_1 u_1 + \cdots + t_n u_n$, then $Lx = [t_1, \dots, t_n]$. Since the basis $(u_1, \dots, u_n)$ is orthonormal, we have $t_i = u_i \cdot x$. Let $U \subset \mathbf{V}$ be an open set, g a real function on U and $f : U \to \mathbf{V}$ a vector field. Denote $G = L(U)$, $\tilde{g} = g \circ L^{-1}$ and $\tilde{f} = L \circ f \circ L^{-1}$. Then $\tilde{f}$ is a real function on G and $\tilde{g}$ is a mapping of G into $\mathbf{R}^n$. Of course, the coordinates and the matrices of derivatives $\tilde{f}$ and $\tilde{g}$ depend on the choice of a basis $(u_1, \dots, u_n)$. We can introduce $\operatorname{grad} g(x) := L^{-1}(\operatorname{grad} \tilde{g}(Lx))$, $\operatorname{div} f(x) := \operatorname{div} \tilde{f}(Lx)$, and $\operatorname{curl} f(x) := \operatorname{curl} \tilde{f}(Lx)$ (if $n = 2$) or $\operatorname{curl} f(x) := L^{-1} \operatorname{curl} f(Lx)$ (if $n = 3$). Then the operators grad, div and curl do not depend on the choice of an orthonormal basis in $\mathbf{V}$.

37.2. Example (in $\mathbf{R}^2$). Let $f(x) = [x_1^2 + x_2^2, x_2 - x_1]$. Then $\operatorname{div} f = \frac{\partial(x_1^2+x_2^2)}{\partial x_1} + \frac{\partial(x_2-x_1)}{\partial x_2} = 2x_1 + 1$ and $\operatorname{curl} f = \frac{\partial(x_2-x_1)}{\partial x_1} - \frac{\partial(x_1^2+x_2^2)}{\partial x_2} = -1 - 2x_2$.

37.3. Example (in $\mathbf{R}^3$). Let $f(x) = [x_1, x_2^2, x_3 x_1]$. Then $\operatorname{div} f = \frac{\partial x_1}{\partial x_1} + \frac{\partial x_2^2}{\partial x_2} + \frac{\partial x_3 x_1}{\partial x_3} = 1 + 2x_2 + x_1$, $\operatorname{curl} f = [\frac{\partial x_3 x_1}{\partial x_2} - \frac{\partial x_2^2}{\partial x_3}, \frac{\partial x_1}{\partial x_3} - \frac{\partial x_3 x_1}{\partial x_1}, \frac{\partial x_2^2}{\partial x_1} - \frac{\partial x_1}{\partial x_2}] = [0, -x_3, 0]$.

37.4. Exercise. Let g be a function of class $\mathscr{C}^1$ defined on a neighborhood of x and $u = \operatorname{grad} g(x) \neq 0$. Show that the function which associates with an unit vector v the derivative of g at x in the direction v attains its maximum at $u/|u|$.

37.5. Vector Product. A vector $w \in \mathbf{R}^n$ is said to be the *vector product* (also called the *cross product*) of vectors $u_1, \dots, u_{n-1}$ and denoted by $u_1 \times \cdots \times u_{n-1}$, if for each vector $v \in \mathbf{R}^n$ we have $w \cdot v = \det(v, u_1, \dots, u_{n-1})$. Thus in particular the i-th coordinate of the vector w is

$$e_i \cdot w = \det(e_i, u_1, \dots, u_{n-1}) = (-1)^{i+1} \det(e_j \cdot u_m)_{j=1,\dots,i-1,i+1,\dots n}^{m=1,\dots,n-1}.$$

Notice that the vector product is a binary operation if and only if $n = 3$. The vector product is always perpendicular to its factors and its norm is

$$|u_1 \times \cdots \times u_{n-1}| = \operatorname{vol}(u_1, \dots, u_{n-1}).$$

37.6. Example. In $\mathbf{R}^3$ we have

$$[1,1,1] \times [1,2,3] = \left[\det \begin{pmatrix} 1, & 1 \\ 2, & 3 \end{pmatrix}, -\det \begin{pmatrix} 1, & 1 \\ 1, & 3 \end{pmatrix}, \det \begin{pmatrix} 1, & 1 \\ 1, & 2 \end{pmatrix}\right] = [1, -2, 1].$$

37.7. Example. In $\mathbf{R}^2$, the vector "product" of the vector $[1, 2]$ is the vector $[2, -1]$.

37.8. Exercise. Compute in $\mathbf{R}^4$ the vector product $[1, 0, 1, 0] \times [0, 0, 0, 1] \times [0, -2, 0, 0]$.

37.9. Orientation. Let Ω be a k-dimensional surface and $x \in \Omega$. By $\mathscr{P}_x = \mathscr{P}_x(\Omega)$ denote the set of all parametrizations of neighborhoods of x in Ω.

By a *(local) orientation* of Ω at x we understand a mapping, which associates with every parametrization $\varphi \in \mathscr{P}_x$ at x a *positive* or a *negative* sign so that $\det(\varphi_2^{-1} \circ \varphi_1)' > 0$ almost everywhere on a neighborhood of $\varphi_1^{-1}(x)$ whenever $\varphi_1 \in \mathscr{P}_x$ and $\varphi_2 \in \mathscr{P}_x$ are both positive or both negative parametrizations, and $\det(\varphi_2^{-1} \circ \varphi_1)' < 0$ almost everywhere on a neighborhood of $\varphi_1^{-1}(x)$ whenever $\varphi_1 \in \mathscr{P}_x$ is a positive parametrization and $\varphi_2 \in \mathscr{P}_x$ is a negative parametrization. It can be proved that a local orientation is always available.

By an *orientation* of a surface Ω we understand its local orientation at each point satisfying the following additional property: If φ is a positive parametrization at x, then it is positive also at all points of some neighborhood of x.

Although each point has exactly two local orientations, some problems can occur when orienting a whole surface. There exist surfaces without a possibility of a (global) orientation (the well known Möbius strip, see Example 39.7). When orientations exist, their number is even. An orientable surface with n connected components possess 2^n orientations.

An n-dimensional surface in $\mathbf{R}^n$ is nothing else than an open subset of $\mathbf{R}^n$. It possesses its *natural orientation* in which the identical parametrizations are positive.

0-dimensional surfaces are countable sets of isolated points, compact 0-dimensional surfaces have only a finite number of points. Any oriented 0-dimensional surface F is decomposed into sets F^+ and F^-. At points of F^+ all parametrizations are positive and at all points of F^- are negative.

37.10. Curves. One-dimensional surfaces are called *curves*. Let $\gamma \subset \mathbf{R}^n$ be an oriented curve. If $\varphi\colon G \to \gamma$ is a positive parametrization, then for almost every $t \in G$ we define the *unit tangent vector* $\mathbf{t}(x)$ at the point $x = \varphi(t)$ by the formula $\mathbf{t}(x) = u/|u|$, where $u = \frac{d\varphi}{dt}(t)$. Then the vector $\mathbf{t}(x)$ is defined up to an s-null set and unit tangent vectors defined by means of different positive parametrizations coincide except on an s-null set (for more details see 38.14).

The field $\mathbf{t}(x)$ obtained in this way determines the orientation. Namely, a parametrization $\varphi\colon G \to \gamma$ is positive whenever

$$\frac{d\varphi}{dt}(t) = \mathbf{t}(\varphi(t)) \left|\frac{d\varphi}{dt}(t)\right|$$

for almost all $t \in G$ and negative if

$$\frac{d\varphi}{dt}(t) = -\mathbf{t}(\varphi(t)) \left|\frac{d\varphi}{dt}(t)\right|$$

for almost all $t \in G$.

The field of unit tangent vectors will be sometimes called shortly the *tangent field*.

The integral $\int_\gamma f \cdot \mathbf{t}\, \mathrm{d}s$ is called a *curve integral* of the vector field f. It satisfies the following Change of Variable Formula.

37.11. Theorem. *Let γ be an oriented curve. Let $G \subset \mathbf{R}$ be an open set and $\varphi : G \to \gamma$ be a positive parametrization. Then for any vector field $f = [f_1, \ldots, f_n]$, $f_j \in \mathscr{L}^1(\gamma, s)$, we have*

$$\int_{\varphi(G)} f \cdot \mathbf{t}\, \mathrm{d}s = \int_G \big(f_1(\varphi(t))\varphi_1'(t) + \cdots + f_n(\varphi(t))\varphi_n'(t)\big)\, \mathrm{d}t.$$

37.12. Example. Let $\gamma = \{[x, y, z] \in \mathbf{R}^3 : x = \cos z,\ y = \sin z,\ 0 < z < 2\pi\}$ be a helix (see Example 34.25) and $f(x, y, z) = [0, x, 3z^2]$. We orient γ in such a way that the unit tangent vector would have a positive z-coordinate. Parametrizing $\varphi(t) = [x(t), y(t), z(t)] = [\cos t,\ \sin t,\ t]$ we compute

$$\mathbf{t} = \frac{1}{\sqrt{2}}[-\sin z, \cos z, 1] = \frac{1}{\sqrt{2}}[-y, x, 1]\,,$$

and therefore

$$\int_\gamma f \cdot \mathbf{t}\, \mathrm{d}s = \int_0^{2\pi} (x\, y' + 3z^2\, z')\, \mathrm{d}t = \int_0^{2\pi} (\cos t (\sin t)' + 3t^2)\, \mathrm{d}t = \pi + 8\pi^3.$$

37.13. Normal Field on $(n-1)$-dimensional Surfaces. Let $\Gamma \subset \mathbf{R}^n$ be an oriented $(n-1)$-dimensional surface. If $\varphi \colon G \to \Gamma$ is a positive parametrization, then for almost all $t \in G$ we define the *unit normal vector* $\mathbf{n}(x)$ at the point $x = \varphi(t)$ by the formula

$$\mathbf{n}(x) = \frac{w_1 \times \cdots \times w_{n-1}}{|w_1 \times \cdots \times w_{n-1}|},$$

where $w_j = \frac{\partial \varphi}{\partial t_j}(t)$. Again, the vector $\mathbf{n}(x)$ is defined except on a S-null set and using different parametrizations yields the same result outside a S-null set (for more details see 38.14).

The field $\mathbf{n}(x)$ obtained in this way determines the orientation. A parametrization $\varphi \colon G \to \Gamma$ is positive provided for almost all $t \in G$ we have

$$\frac{\partial \varphi}{\partial t_1}(t) \times \cdots \times \frac{\partial \varphi}{\partial t_{n-1}}(t) = \mathbf{n}(\varphi(t)) \left| \frac{\partial \varphi}{\partial t_1}(t) \times \cdots \times \frac{\partial \varphi}{\partial t_{n-1}}(t) \right|,$$

and negative if the above equality holds having the opposite sign on the right-hand side.

The field of unit normal vectors will be briefly called the *normal field*.

The integral $\int_\Gamma f \cdot \mathbf{n}\, \mathrm{d}S$ is called the *surface integral* of the vector field f. It satisfies the following Change of Variable Formula.

37.14. Theorem. *Let Γ be an $(n-1)$-dimensional oriented surface. Let $G \subset \mathbf{R}^{n-1}$ be an open set and $\varphi : G \to \Gamma$ a positive parametrization. Then for any vector field $f = [f_1, \ldots, f_n]$, where $f_j \in \mathscr{L}^1(\Gamma, S)$, we have*

$$\int_{\varphi(G)} f \cdot \mathbf{n}\, \mathrm{d}S = \int_G f(\varphi(t)) \cdot \Big(\frac{\partial \varphi}{\partial t_1} \times \cdots \times \frac{\partial \varphi}{\partial t_{n-1}}\Big)\, \mathrm{d}t.$$

37.15. Example. We evaluate the integral

$$\int_\Gamma [x, y, z^2] \cdot \mathbf{n}\, \mathrm{d}S,$$

where

$$\Gamma = \{[x, y, z] \in \mathbf{R}^3 \colon z^2 = x^2 + y^2,\ 0 < z < 1\}$$

and $\mathbf{n}$ is supposed to be oriented "outwards from the cone"

$$\Omega := \{[x, y, z] \in \mathbf{R}^3 \colon x^2 + y^2 < z^2, 0 < z < 1\}$$

(see 37.21). We use the parametrization

$$\varphi(r, t) = [r\cos t,\ r\sin t, r]$$

on $G := \{[r, t] \colon r \in (0, 1) \text{ and } t \in (0, 2\pi)\}$. Obviously $\varphi(G)$ differs from Γ only in a set of measure zero, hence there is no difference between integration over Γ and over $\varphi(G)$. We have

$$\varphi'(r, t) = \begin{pmatrix} \cos t, & -r\sin t \\ \sin t, & r\cos t \\ 1, & 0 \end{pmatrix},$$

so that

$$\frac{\partial \varphi}{\partial r} \times \frac{\partial \varphi}{\partial t} = [-r\cos t,\ -r\sin t,\ r].$$

In case of a positive parametrization this vector would be a positive multiple of the unit normal vector and thus it would be directed outwards from Ω. Since it is directed inwards, the parametrization φ is negative (like in 37.21 it is possible to precise "outwards" and "inwards"). The formula will differ only in the sign. We have

$$\begin{aligned} \int_\Gamma [x, y, z^2] \cdot \mathbf{n}\, \mathrm{d}S &= \int_G [r\cos t,\ r\sin t,\ r^2] \cdot [r\cos t,\ r\sin t,\ -r]\, \mathrm{d}r\, \mathrm{d}t \\ &= \int_G (r^2 - r^3)\, \mathrm{d}r\, \mathrm{d}t = \frac{\pi}{6}. \end{aligned}$$

37.16. Example. Let $\Gamma := \{[x, y, z] \in \mathbf{R}^3 : x^2 + y^2 = z < 1\}$. Let the unit normal vector

$$\mathbf{n}(x, y, z) = \pm \frac{1}{\sqrt{4z + 1}} [2x, 2y, -1])$$

be oriented by the choice of its sign $+$. Then $[t, r] \mapsto [x(t, r), y(t, r), z(t, r)] := [r\cos t, r\sin t, r^2]$, $t \in (0, 2\pi)$, $r \in (0, 1)$, is a positive parametrization and its range differs from Γ only in a set of measure zero. Let $f(x, y, z) = [2x, 0, 0]$. Then

$$\begin{aligned} \int_\Gamma f \cdot \mathbf{n}\, \mathrm{d}S &= \int_{(0,2\pi)\times(0,1)} 2x\ \det(\nabla y, \nabla z)\, \mathrm{d}t\, \mathrm{d}r \\ &= \int_{(0,2\pi)\times(0,1)} 2r\cos t\ \det(\nabla(r\sin t), \nabla(r^2))\, dt\, \mathrm{d}r \\ &= \int_{(0,2\pi)\times(0,1)} 4r^3\ cos^2 t\, \mathrm{d}t\, dr = \pi. \end{aligned}$$

37.17. Surfaces with Lipschitz Boundaries. One of the most important formula of the integral calculus on surfaces is the general Stokes' Theorem and its special cases. For the purpose of a formulation of these results we need to introduce the notion of a surface with a Lipschitz boundary.

Denote by $\mathbf{H}^k_+$, $\mathbf{H}^k_-$ the halfspaces $(0,\infty)\times\mathbf{R}^{k-1}$ and $(-\infty,0)\times\mathbf{R}^{k-1}$, respectively. Further denote by $\mathbf{i}$ the mapping of $\mathbf{R}^{k-1}$ onto $\partial\mathbf{H}^k_-$ defined as

$$\mathbf{i}([s_1,\dots,s_{k-1}]) = [0,s_1,\dots,s_{k-1}].$$

Let $\Omega\subset\mathbf{R}^n$ be a bounded oriented k-dimensional surface. The k-boundary of Ω is defined as $\overline{\Omega}\setminus\Omega$. It is the topological boundary of Ω if $k=n$. Suppose that the k-boundary Γ of Ω is an oriented $k-1$-dimensional surface. We say that Γ is a *Lipschitz k-boundary* of Ω if for every point $z\in\Gamma$ there exist an open set $G\subset\mathbf{R}^k$, a neighborhood U of x and a homeomorphic locally bilipschitz mapping $\varphi : G\cap\overline{\mathbf{H}^k_-}\to\Omega\cup\Gamma$ such that $z\in\varphi(\partial\mathbf{H}^k_-)$, $\varphi(G\cap\mathbf{H}^k_-)=\Omega\cap U$, $\varphi(G\cap\partial\mathbf{H}^k_-)=\Gamma\cap U$ and one of the following situations occurs:

(a) $\varphi|_{G\cap\mathbf{H}^k_-}$ is a positive parametrization of $\Omega\cap U$ and $\varphi_{|G\cap\partial\mathbf{H}^k_-}\circ\mathbf{i}$ is a positive parametrization of $\Gamma\cap U$;

(b) $\varphi|_{G\cap\mathbf{H}^k_-}$ is a negative parametrization of $\Omega\cap U$ and $\varphi|_{G\cap\partial\mathbf{H}^k_-}\circ\mathbf{i}$ is a negative parametrization of $\Gamma\cap U$.

If $k>1$ and a parametrization φ satisfies (b), the by a "mirror-like" modification we get a parametrization satisfying (a).

37.18. Introduction to Curve Integral Theorem. Let $\gamma\subset\mathbf{R}^n$ be an oriented curve with a Lipschitz 1-boundary F. As we already know, the orientation on γ is formed by a field of unit tangent vectors $\mathbf{t}$ and F is a finite set consisting from a "positive part" F^+ and a "negative part" F^-.

The relation between orientations γ and F is given in Definition 37.17. Less precisely but transparently: The tangent field is directed from points of the set F^- towards points of the set F^+. If $\mathbf{t}(a)$ is a continuous extension of $\mathbf{t}$ to the point $a\in F^-$ (warning: its existence is not guaranteed by our assumptions), then

$$\mathbf{t}(a) = \lim_{\substack{x\to a\\ x\in\gamma}}\frac{x-a}{|x-a|}.$$

The result in $b\in F^+$ is similar, but with an opposite sign.

Under these assumptions the following result is valid.

37.19. Curve Integral Theorem. *Let g be a continuously diferentiable function on a neighborhood of $\overline{\gamma}$. Then*

$$\sum_{b\in F^+} g(b) - \sum_{a\in F^-} g(a) = \int_\gamma \operatorname{grad} g\cdot\mathbf{t}\,\mathrm{d}s.$$

37.20. Example. Let h be a Lipschitz function on $[-1,1]$, $h(-1)=h(1)=0$. Let $\gamma=\{[x,y] : y=h(x), |x|<1\}$. If we choose an orientation of the unit tangent vector $\mathbf{t}$ to γ so that its x-coordinate is positive, then

$$\mathbf{t}(x,y) = \frac{1}{\sqrt{1+(h'(x))^2}}[1,h'(x)]\,.$$

We have to evaluate the integral $\int_\gamma f\cdot\mathbf{t}\,\mathrm{d}s$, where $f(x,y)=[3x^2\,\cos y,\ -x^3\,\sin y]$. A direct computation does not seem to be very hopeful, particularly if the function h is rather complicated. Nevertheless, since $f=\operatorname{grad} g$ for $g=x^3\,\cos y$, by the Curve Integral Theorem we get

$$\int_\gamma f\cdot\mathbf{t}\,\mathrm{d}s = g([1,0])-g([-1,0])=2.$$

37.21. Introduction to Gauss Theorem. Next we introduce the Gauss Theorem which is also called the Gauss-Ostrogradski Theorem, or the Divergence Theorem.

Let $\Omega\subset\mathbf{R}^n$ be a bounded open set with a Lipschitz boundary Γ (i.e. suppose that the conditions of 37.17 are satisfied) and consider a natural orientation on Ω. Thus the orientation of Γ is uniquely determined and it is represented by the field $\mathbf{n}$ of unit normal vectors. Roughly speaking, we can say that $\mathbf{n}(x)$ is the unit vector which is perpendicular to the boundary of Ω at x and is oriented out of Ω. Therefore $\mathbf{n}(x)$ is also labelled as the vector of the *outer normal*. By the formulation "out" we understand that for a nonvanishing vector $u\in\mathbf{R}^n$ and a small positive t we have $x+tu\in\Omega$ provided $u\cdot\mathbf{n}(x)<0$, and $x+tu\notin\Omega$ provided $u\cdot\mathbf{n}(x)>0$. Of course, such a situation occurs only at the points of smoothness of Γ.

Under the above described assumptions the following result holds.

37.22. Gauss Theorem. *Let f be a vector field of class $\mathscr{C}^1$ on a neighborhood of $\overline{\Omega}$. Then*

$$\int_\Gamma f\cdot\mathbf{n}\,\mathrm{d}S=\int_\Omega \operatorname{div} f\,d\lambda.$$

37.23. Example (the ball, spherical coordinates). Let $\Omega=\{[x,y,z]\in\mathbf{R}^3\colon x^2+y^2+z^2<1\}$, $\Gamma=\{[x,y,z]\in\mathbf{R}^3\colon x^2+y^2+z^2=1\}$. Consider the mapping given by the spherical coordinates: $\psi=[x,y,z]$, where

$$\begin{aligned} x(r,t,a)&=r\cos a\,\cos t,\\ y(r,t,a)&=r\cos a\,\sin t,\\ z(r,t,a)&=r\sin a,\end{aligned}$$

and $[r,t,a]\in H:=(0,\infty)\times(-\pi,\pi)\times(-\frac12\pi,\ \frac12\pi)$. We have

$$\psi'(r,t,a)=\begin{pmatrix}\cos a\cos t, & -r\sin t\cos a, & -r\cos t\sin a\\ \cos a\sin t, & r\cos t\cos a, & -r\sin t\sin a\\ \sin a, & 0, & r\cos a\end{pmatrix}.$$

Then the mapping $\varphi(s,t,a)=\psi(s+1,t,a)$, $[s,t,a]\in(-1,0]\times(-\pi,\pi)\times(-\frac12\pi,\ \frac12\pi)$ satisfies requirements of 37.17 for $[x,y,z]\in\psi(H)$. (If $[x_0,y_0,z_0]$ is not in $\psi(H)$, we can use mappings $[s,t,a]\mapsto\psi(s+1,t-t_0,a-a_0)$ for suitable t_0 and a_0.) After verifying that $\det\psi'=r^2\cos a>0$ and computing the outer normals we get

$$\begin{aligned} w_2&:=\frac{\partial[\cos a\,\cos t,\ \cos a\,\sin t,\ \sin a]}{\partial t}=[-\cos a\,\sin t,\ \cos a\,\cos t,\ 0],\\ w_3&:=\frac{\partial[\cos a\,\cos t,\ \cos a\,\sin t,\ \sin a]}{\partial a}=[-\sin a\,\cos t,\ -\sin a\,\sin t,\ \cos a],\\ w_2\times w_3&=[\cos^2 a\,\cos t,\ \cos^2 a\,\sin t,\ \cos a\,\sin a],\\ |w_2\times w_3|&=\cos a\,.\end{aligned}$$

Finally

$$\mathbf{n}(x,y,z) = \frac{w_2 \times w_3}{|w_2 \times w_3|} = [\cos t\,\cos a,\,\sin t\,\cos a,\,\sin a] = [x,y,z].$$

37.24. Example. (a) By means of the Gauss Theorem we will evaluate (not for the first time) the surface measure of the sphere $\Gamma := \{(x,y,z) \in \mathbf{R}^3 \colon x^2+y^2+z^2 = 1\}$. We utilize the fact that Γ is a Lipschitz boundary of the ball $\Omega = \{(x,y,z) \in \mathbf{R}^3 \colon x^2+y^2+z^2 < 1\}$. We compute the two-dimensional measure of the sphere by integrating the constant 1 which we express in the form $1 = f \cdot \mathbf{n}$, where we choose e.g. $f(x,y,z) = [x,y,z]$. Actually, $[x,y,z] \cdot \mathbf{n} = [x,y,z] \cdot [x,y,z] = x^2+y^2+z^2 = 1$ (for a different set it would be necessary to find a different f). We obtain

$$\int_\Gamma 1\,\mathrm{d}S = \int_\Gamma [x,y,z] \cdot \mathbf{n}\,\mathrm{d}S = \int_\Omega \operatorname{div}[x,y,z]\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\Omega)} 3\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 4\pi.$$

(b) As an exercise evaluate

$$\int_\Gamma x^2\,\mathrm{d}S \qquad (\text{result: } \frac{4}{3}\,\pi).$$

37.25. Example (the cube). Let Ω be a cube $(0,1)^3$ and Γ its boundary oriented by the outer normal (e.g. on $\{0\} \times (0,1)^2$ the unit normal vector is $[-1,0,0]$). The condition 37.17 is clearly satisfied for z lying on this side, and less obvious (but still satisfied) if z lies on an edge or even at a vertex. Let, for example, z be the vertex $[0,0,0]$. Then a mapping φ with the desired properties can be found in the form $\varphi(t) = [-t_1+\max(0,-t_2,-t_3),\ -t_1+\max(t_2,0,t_2-t_3),\ -t_1+\max(t_3,t_3-t_2,0)]$, $t \in (-\frac12, 0] \times (-\frac12, \frac12)^2 \cap \{|t_2-t_3| < \frac12\}$.

37.26. Introduction to Green Theorem. Consider the situation described in the Gauss theorem for the two-dimensional case. Then Γ is a curve, the orientation of which can be expressed not only by means of the normal field, but also by means of the tangent field. The vector $\mathbf{n}(x)$ is s-almost everywhere the "vector product of the vector" $\mathbf{t}(x)$, i.e. $(\mathbf{n}(x), \mathbf{t}(x))$ is a positive orthonormal basis of $\mathbf{R}^2$. (Roughly speaking, $\mathbf{t}(x)$ circulates around Ω anti-clockwise.)

37.27. Green Theorem. *Let $g = [g_1, g_2]$ be a continuously differentiable vector field on a neighborhood of $\overline{\Omega}$. Then*

$$\int_\Gamma g \cdot \mathbf{t}\,\mathrm{d}s = \int_\Omega \operatorname{curl} g\,\mathrm{d}x.$$

37.28. Exercise. Compute the unit tangent vector and the unit normal vector to the unit circle oriented as the (Lipschitz) boundary to the unit disc.

37.29. Example. Using the Green Theorem we compute the contents of the figure $\Omega := \{(x,y) \in \mathbf{R}^2 \colon \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} < 1\}$. The boundary (ellipse) Γ can be (up to a set of one-dimensional measure zero) parametrized by the mapping $\varphi(t) = [x(t), y(t)] := (a\cos t, b\sin t)$, $t \in (0, 2\pi)$. Let f be a vector field on $\mathbf{R}^n$ whose curl is 1, e.g. $f = [0,x]$, $f = [-y,0]$, or $f = [-\frac12 y, \frac12 x]$. The the Green theorem yields

$$\int_\Omega 1\,\mathrm{d}x\,\mathrm{d}y = \int_\Gamma [0,x] \cdot \mathbf{t}(x,y)\,\mathrm{d}s = \int_0^{2\pi} xy'\,\mathrm{d}t = \int_0^{2\pi} ab\cos^2 t\,\mathrm{d}t = \pi ab.$$

37.30. Exercise. Using the Green theorem evaluate the content of the figure surrounded by the asteroide $\{(a\cos^3 t,\, b\sin^3 t) \colon t \in [0, 2\pi)\}$.

37.31. Introduction to Stokes' Theorem. Let $\Gamma \subset \mathbf{R}^3$ be an oriented two-dimensional surface with a Lipschitz k-boundary γ. Suppose that a normal field $\mathbf{n}$ on Γ and a tangent field $\mathbf{t}$ on γ are given. Less precisely, but more transparently we can say that the tangent field $\mathbf{t}(x)$ again circulates anti-clockwise provided we observe the situation against the direction of the normal vector $\mathbf{n}(x)$. (Warning: in contrast to the Green Theorem, here $\mathbf{n}$ means the normal field to the surface.)

Under the above stated assumptions the following theorem holds.

37.32. (Special) Stokes' Theorem. *Let f be a vector field of class $\mathscr{C}^1$ on a neighborhood of $\overline{\Gamma}$. Then*

$$\int_\gamma f \cdot \mathbf{t} \, \mathrm{d}s = \int_\Gamma \mathbf{n} \cdot \operatorname{curl} f \, dS.$$

37.33. Example. Let h be a positive even function of class $\mathscr{C}^1$ on $[-1,1]$, $h(1)=0$. Let $\Gamma = \{[x,y,z] : z = h(r)\}$, $r = \sqrt{x^2+y^2}$. Let the unit normal vector have the orientation for which its z-coordinate is positive, i.e.

$$\mathbf{n}(x,y,z) = \frac{1}{\sqrt{1+(h'(r))^2}} \, [-h'(r)\frac{x}{r}, \, -h'(r)\frac{y}{r}, \, 1].$$

Let us evaluate $\int_\Gamma g \cdot \mathbf{n} \, \mathrm{d}S$ for $g(x,y,z) = [-xe^z, -ye^z, 2e^z]$. Since $g = \operatorname{curl} f$ for $f(x,y,z) = [-ye^z, xe^z, 0]$ and $f = [-y,x,0]$ on $\gamma := \{[x,y,z] \in \mathbf{R}^3 : r=1, z=0\}$ is a Lipschitz 2-boundary of Γ, the Stokes formula gives

$$\int_\Gamma g \cdot \mathbf{n} \, \mathrm{d}S = \int_\gamma f \cdot \mathbf{t} \, ds = \int_\gamma [-y,x] \cdot \mathbf{t}(x,y) \, \mathrm{d}s = \int_0^{2\pi} (\sin^2 t + \cos^2 t) \, \mathrm{d}t = 2\pi.$$

For the evaluation of the integral over γ we used polar coordinates: $x = \cos t$, $y = \sin t$, $t \in (0, 2\pi)$.

37.44. Notes. Main formulas of the calculus of curve and surface integral were established in 19$^{\text{th}}$ century. The divergence theorem was discovered by K. F. Gauss, G. Green and M. V. Ostrogradski. The Stokes theorem is due to G. G. Stokes. See also 39.24.

38. Integration of Differential Forms

In this chapter we present the theorem which contains as special cases the Curve integral theorem, the Gauss theorem and Stokes' theorem of Chapter 37. For this purpose we need more advanced tools of exterior algebra.

38.1. k-covectors and k-vectors. Let $\mathbf{V}$ be a n-dimensional vector space and $\mathbf{V}^*$ its dual space. The value of a linear form $V \in \mathbf{V}^*$ at a vector $u \in \mathbf{V}$ will be denoted by $\langle V, u \rangle$.

A mapping $W : \mathbf{V}^k \to \mathbf{R}$ is called a *k-linear form* (on $\mathbf{V}$) if it is linear in each variable separately. As a special case we get *linear forms* for $k=1$, or *bilinear forms* for $k=2$.

A k-linear form W on $\mathbf{V}$ is said to be a *k-covector* if W is *antisymmetric* in the following sense: If π is a transposition (a permutation, which transposes exactly one pair of elements) of the set $\{1, \dots, k\}$, then

$$\left\langle W, \, (u_{\pi(1)}, \dots, u_{\pi(k)}) \right\rangle \; = \; - \left\langle W, \, (u_1, \dots, u_k) \right\rangle.$$

Given a k-covector W, we have

$$\left\langle W,\ (\sum_{j=1}^{k} a_{1,j}u_j, \ldots, \sum_{j=1}^{k} a_{k,j}u_j)\right\rangle \;=\; \det A\ \langle W,\ (u_1,\ldots,u_k)\rangle$$

for any matrix $A = (a_{i,j})_{i,j=1}^{k}$ and any ordered k-tuple $(u_1,\ldots,u_k)$ of vectors of $\mathbf{V}$.

Dually, a *k-vector* on $\mathbf{V}$ will be defined as a k-covector on $\mathbf{V}^*$. Thus a k-vector assigns a real number to a k-tuple of linear forms on $\mathbf{V}$. In particular, a 1-covector is the same as a linear form; a 1-vector u on $\mathbf{V}$ will be identified with a vector $v \in \mathbf{V}$, if u assigns to each linear form $V \in \mathbf{V}^*$ the value $\langle V, v\rangle$. A real number c will be identified with a 0-form, or a 0-vector, which assigns the number c to each 0-tuple of vectors or forms, respectively. The vector space of all k-covectors on $\mathbf{V}$ will be denoted by $\Lambda^k(\mathbf{V})$ and the vector space of all k-vectors on $\mathbf{V}$ by $\Lambda_k(\mathbf{V})$. So $\Lambda^k(\mathbf{V}) = \Lambda_k(\mathbf{V}^*)$, $\Lambda_k(\mathbf{V}) = \Lambda^k(\mathbf{V}^*)$, $\Lambda_1(\mathbf{V}) = \mathbf{V}$, $\Lambda^1(\mathbf{V}) = \mathbf{V}^*$, and $\Lambda_0(\mathbf{V}) = \Lambda^0(\mathbf{V}) = \mathbf{R}$.

38.2. Example. A bilinear form is representable by a matrix. For the sake of simplicity assume $\mathbf{V} = \mathbf{R}^n$, then $(a_{i,j})_{i,j}$ is a matrix of a bilinear form V if

$$\langle V, (u,v)\rangle \;=\; \sum_{i,j} a_{i,j}u_iv_j$$

for each $u, v \in \mathbf{R}^n$. Among bilinear forms we select 2-covectors as bilinear forms with an antisymmetric matrix, i.e. $a_{i,j} = -a_{j,i}$ (in particular $a_{i,i} = 0$). An example of a matrix of a 2-covector in $\mathbf{R}^3$ is

$$\begin{pmatrix} 0, & 2, & -3 \\ -2, & 0, & 5 \\ 3, & -5, & 0 \end{pmatrix}.$$

The (canonical) inner product represented by the unit matrix is also a bilinear form but it fails to be a 2-covector.

38.3. Exterior Product. By an *exterior product* of an ordered k-tuple $(V_1,\ldots V_k)$ of linear forms on $\mathbf{V}$ we mean the k-covector $V_1 \wedge \cdots \wedge V_k$ defined by the formula

$$\langle V_1 \wedge \cdots \wedge V_k,\ (u_1,\ldots,u_k)\rangle \;=\; \det(\langle V_i,\ u_j\rangle)_{i,j=1}^{k}, \quad (u_1,\ldots,u_k) \in \mathbf{V}^k.$$

Dually, the *exterior product* of an ordered k-tuple $(u_1,\ldots u_k)$ of vectors from $\mathbf{V}$ will be the k-vector $u_1 \wedge \cdots \wedge u_k$ defined by the formula

$$\langle (V_1,\ldots,V_k),\ u_1 \wedge \cdots \wedge u_k\rangle \;=\; \det(\langle V_i,\ u_j\rangle)_{i,j=1}^{k}, \quad (V_1,\ldots,V_k) \in (\mathbf{V}^*)^k.$$

Notice that the exterior product is invariant with respect to even permutations of factors. An odd permutation changes (only) the sign. The exterior product vanishes if the factors are linearly dependent (in particular, if two factors coincide).

Now we will investigate coordinates of k-vectors and k-covectors in $\mathbf{R}^n$. Let X_i denote the i-th coordinate form $[u_1,\ldots,u_n] \mapsto u_i$. Recall that in the context of these chapters by a *multiindex* we understand an ordered k-tuple $\alpha = [\alpha_1,\ldots,\alpha_k]$

of indices from $\{1,\dots,n\}$, and that the set of all such multiindices is denoted by $\{1,\dots,n\}^k$. A multiindex α is called *increasing* if $\alpha_1 < \cdots < \alpha_k$. The set of all increasing multiindices from $\{1,\dots,n\}^k$ is denoted by $I(k,n)$. Each k-covector W posseses (uniquely determined) real coordinates W_α, $\alpha \in \{1,\dots,n\}^k$, namely

$$W_\alpha = \langle W,\ (e_{\alpha_1},\dots,e_{\alpha_k})\rangle\,.$$

Dually, each k-vector w has (uniquely determined) real coordinates w_α, $\alpha \in \{1,\dots,n\}^k$, namely

$$w_\alpha = \langle (X_{\alpha_1},\dots,X_{\alpha_k}),\ w\rangle\,.$$

Let W be a k-covector (take into account analogous considerations for k-vectors). From the antisymmetry it follows

$$W_{\pi(\alpha)} = -W_\alpha$$

for each transposition π of indices.

In particular $W_\alpha = 0$ whenever an index occurs in the multiindex at least twice. For a complete description of a k-covector we need only coordinates corresponding to increasing multiindices.

We have

$$W = \sum_{\alpha\in I(k,n)} W_\alpha\, X_{\alpha_1}\wedge\cdots\wedge X_{\alpha_k}\,,$$

so that the basis of $\Lambda^k(\mathbf{R}^n)$ is $\{X_{\alpha_1}\wedge\cdots\wedge X_{\alpha_k} : \alpha\in I(k,n)\}$. Dually, we have

$$w = \sum_{\alpha\in I(k,n)} w_\alpha\, e_{\alpha_1}\wedge\cdots\wedge e_{\alpha_k}\,,$$

so that the basis of $\Lambda_k(\mathbf{R}^n)$ is $\{e_{\alpha_1}\wedge\ldots\wedge e_{\alpha_k} : \alpha\in I(k,n)\}$.

A similar consideration leads to a description of the basis of $\Lambda^k(\mathbf{V})$ and $\Lambda_k(\mathbf{V})$ for a general n-dimensional vector space $\mathbf{V}$. (It remains only to replace $(e_1,\dots,e_n)$ by a fixed basis $\mathbf{V}$ and $(X_1,\dots,X_n)$ by its dual basis.) It follows that the dimension of the spaces $\Lambda^k(\mathbf{V})$ and $\Lambda_k(\mathbf{V})$ is equal to the number of the multiindices of $I(k,n)$ which is $\binom{n}{k}$.

By means of coordinates a duality pairing between k-vectors and k-covectors can be defined: If $v\in\Lambda_k(\mathbf{V})$ and $W\in\Lambda^k(\mathbf{V})$, then we introduce

$$\langle W, v\rangle :\ = \sum_{\alpha\in I(k,n)} W_\alpha v_\alpha.$$

In particular, given $v_1,\dots,v_k\in\mathbf{V}$ and $W_1,\dots,W_k\in\mathbf{V}^*$, we have

$$\begin{aligned}\langle W,\ v_1\wedge\cdots\wedge v_k\rangle &= \langle W,(v_1,\dots,v_k)\rangle\,,\\ \langle W_1\wedge\cdots\wedge W_k,\ v\rangle &= \langle (W_1,\dots,W_k),v\rangle\,.\end{aligned}$$

Notice that not every k-covector is an exterior product of linear forms and not every k-vector is an exterior product of vectors, see Example 38.7.

38.4. Example (in $\mathbf{R}^3$).

$$\begin{aligned}(e_1+e_2-e_3)\wedge(2e_2+e_3) &= 2e_1\wedge e_2+2e_2\wedge e_2-2e_3\wedge e_2+e_1\wedge e_3+e_2\wedge e_3-e_3\wedge e_3\\ &= 2e_1\wedge e_2+e_1\wedge e_3+(2+1)e_2\wedge e_3.\end{aligned}$$

38.5. Example (in $\mathbf{R}_3$, notice how the sign is changed under the transposition).

$$\begin{aligned}X_2\wedge(X_1+X_3)\wedge(X_1-X_3) &= X_2\wedge X_1\wedge X_1-X_2\wedge X_1\wedge X_3\\ &\quad +X_2\wedge X_3\wedge X_1-X_2\wedge X_3\wedge X_3\\ &= X_1\wedge X_2\wedge X_3-X_3\wedge X_2\wedge X_1=2X_1\wedge X_2\wedge X_3.\end{aligned}$$

38.6. Example. In $\mathbf{R}^4$ we have

$$\langle X_1\wedge X_2+X_2\wedge X_3+X_3\wedge X_4,\ 2e_1\wedge e_2-e_1\wedge e_4+e_2\wedge e_4\rangle\ =2.$$

38.7. Example. The 2-covector $W=X_1\wedge X_2+X_2\wedge X_3+X_3\wedge X_4$ in $\mathbf{R}^4$ cannot be written as an exterior product $V_1\wedge V_2$ of linear forms. We prove this by a contradiction: Assume that $W=V_1\wedge V_2$. Let $A\colon\mathbf{R}^4\to\mathbf{R}^2$ be the linear mapping $x\mapsto[V_1x,V_2x]$. Since $W_{(1,2)}\neq 0$, we have $Ae_1\neq 0$. Further, $W_{(1,3)}=W_{(1,4)}=0$, so that Ae_3 and Ae_4 are multiples of Ae_1. This is a contradiction, as $W_{(3,4)}\neq 0$.

38.8. Example. There is no chance to construct a counterexample similar to 38.7 in $\mathbf{R}^3$. Indeed, let W be a 2-covector on $\mathbf{R}^3$. If $W_{1,3}=0$, then

$$W=(W_{(1,2)}X_1-W_{(2,3)}X_3)\wedge X_2.$$

If $W_{1,3}\neq 0$, then

$$W=(\frac{W_{(2,3)}}{W_{(1,3)}}\,X_2+X_1)\wedge(W_{(1,3)}X_3+W_{(1,2)}X_2).$$

38.9. Differential Forms. A mapping $\omega: E\to\Lambda^k(\mathbf{R}^n)$, where $E\subset\mathbf{R}^n$, is called a *differential k-form* (or, simply a differential form) on E. We identify differential 0-forms with functions. Any differential k-form ω on $E\subset\mathbf{R}^n$ is representable in coordinates

$$\omega=\sum_{\alpha\in I(k,n)}\omega_\alpha \mathrm{d}x_{\alpha_1}\wedge\cdots\wedge\mathrm{d}x_{\alpha_k},$$

where $\mathrm{d}x_i$ denotes the constant differential 1-form $x\mapsto X_i$ and ω_α are functions on E. We say that a differential form ω is of class $\mathscr{C}^\ell$ if all its coordinates ω_α are of class $\mathscr{C}^\ell$. Similarly we define other properties of differential forms (e.g. measurability) using coordinates .

38.10. Example. We illustrate the calculation with differential forms by the following example in $\mathbf{R}^4$:

$$\begin{aligned}&(\mathrm{d}x_2+\mathrm{d}x_4)\wedge(x_4\,\mathrm{d}x_1-x_1\,\mathrm{d}x_4)\wedge(\mathrm{d}x_2+\mathrm{d}x_3)=(x_4\,\mathrm{d}x_2\wedge\mathrm{d}x_1+x_4\,\mathrm{d}x_4\wedge\mathrm{d}x_1-x_1\,\mathrm{d}x_2\wedge\mathrm{d}x_4\\ &\qquad -x_1\,\mathrm{d}x_4\wedge\mathrm{d}x_4)\wedge(\mathrm{d}x_2+\mathrm{d}x_3)\\ &\quad=(-x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_2-x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_4-x_1\,\mathrm{d}x_2\wedge\mathrm{d}x_4)\wedge(\mathrm{d}x_2+\mathrm{d}x_3)\\ &\quad=-x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_2-x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_4\wedge\mathrm{d}x_2-x_1\,\mathrm{d}x_2\wedge\mathrm{d}x_4\wedge\mathrm{d}x_2\\ &\qquad -x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_3-x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_4\wedge\mathrm{d}x_3-x_1\,\mathrm{d}x_2\wedge\mathrm{d}x_4\wedge\mathrm{d}x_3\\ &\quad=x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_4-x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_2\wedge\mathrm{d}x_3-x_4\,\mathrm{d}x_1\wedge\mathrm{d}x_4\wedge\mathrm{d}x_3+x_1\,\mathrm{d}x_2\wedge\mathrm{d}x_3\wedge\mathrm{d}x_4.\end{aligned}$$

38.11. Differential. Let

$$\omega = \sum_{\alpha \in I(k-1,n)} \omega_\alpha \mathrm{d}x_{\alpha_1} \wedge \cdots \wedge \mathrm{d}x_{\alpha_{k-1}}$$

be a differential $(k-1)$-form of class $\mathscr{C}^1$ on an open set $U \subset \mathbf{R}^n$. Then its *differential* $\mathrm{d}\omega$ is defined as the differential k-form on U given by the formula

$$\mathrm{d}\omega(x) = \sum_{\alpha \in I(k-1,n)} \sum_{i=1}^{n} \frac{\partial \omega_\alpha}{\partial x_i}(x)\, \mathrm{d}x_i \wedge \mathrm{d}x_{\alpha_1} \wedge \cdots \wedge \mathrm{d}x_{\alpha_{k-1}}.$$

The *differential* of a $\mathscr{C}^1$-function f on U is the differential 1-form

$$df(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, \mathrm{d}x_i\,.$$

In particular, the differential of the coordinate function $x \mapsto x_i$ is $\mathrm{d}x_i$ which corresponds to the notation introduced above.

38.12. Example (in $\mathbf{R}^2$).

$$\begin{aligned} d(x_2\,\mathrm{d}x_1 - x_1 \sin x_2\,\mathrm{d}x_2) &= \mathrm{d}x_2 \wedge \mathrm{d}x_1 - \sin x_2\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 - x_1 \cos x_2\,\mathrm{d}x_2 \wedge \mathrm{d}x_2 \\ &= -\mathrm{d}x_1 \wedge \mathrm{d}x_2 - \sin x_2\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 = (-1 - \sin x_2)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2. \end{aligned}$$

38.13. Example (in $\mathbf{R}^3$).

(a)
$$\begin{aligned} d(x_1\,\mathrm{d}x_1 + x_1x_3\,\mathrm{d}x_2) &= \mathrm{d}x_1 \wedge \mathrm{d}x_1 + x_3\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 + x_1\,\mathrm{d}x_3 \wedge \mathrm{d}x_2 \\ &= x_3\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 - x_1\,\mathrm{d}x_2 \wedge \mathrm{d}x_3. \end{aligned}$$

(b)
$$\begin{aligned} d(x_1x_2\,\mathrm{d}x_1 \wedge \mathrm{d}x_3) &= x_2\,\mathrm{d}x_1 \wedge \mathrm{d}x_1 \wedge \mathrm{d}x_3 + x_1\,\mathrm{d}x_2 \wedge \mathrm{d}x_1 \wedge \mathrm{d}x_3 \\ &= -x_1\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3. \end{aligned}$$

38.14. Orientation of k-dimensional Surfaces. In the preceding chapters we have seen that the orientation of a curve forms a vector field (of the unit tangent vectors) on it while the orientation of an $n-1$-dimensional surface also forms a vector field (now that of unit normal vectors) but in an entirely different way. We will see that both tangent and normal fields are particular cases of a general object. An orientation of a k-dimensional surface determines a k-vector tangent field and the $(n-k)$-covector normal field on it. (In fact the normal should also be a covector, but it is customary to understand it as a vector – the inner product structure makes this identification possible.)

Let σ be a k-dimensional measure on $\mathbf{R}^n$ and Ω be a k-dimensional oriented surface in $\mathbf{R}^n$. If $\varphi\colon G \to \Omega$ is a positive parametrization, then for almost all $t \in G$ we define the *unit tangent k-vector* $\xi(x) \in \Lambda_k(\mathbf{R}^n)$ at the point $x = \psi(t)$ by the formula

$$\xi(x) = \frac{w_1 \wedge \cdots \wedge w_k}{\mathrm{vol}(w_1, \ldots, w_k)},$$

where $w_j = \frac{\partial \varphi}{\partial t_j}(t)$. Then the k-vector $\xi(x)$ is defined σ-almost everywhere and does not depend on the particular choice of the parametrization φ.

We will not use the normal $n-k$-covector in full generality, nevertheless for the sake of completeness we will outline the definition: It associates with a $n-k$-tuple $u_1, \dots, u_{n-k}$ of vectors in $\mathbf{R}^n$ the number

$$\operatorname{vol}^{-1}(w_1, \dots, w_k)\ \det(u_1, \dots, u_{n-k}, w_1, \dots, w_k).$$

The vector space T_x generated by the vectors $w_1, \dots, w_k$ (it will be termed the tangent space in the next chapter) is σ-almost everywhere independent of the choice of a parametrization.

Notice that the dimension of $\Lambda_k(T_x)$ is 1. Consider two orientations of a surface Ω. Suppose that $\varphi_1 : G_1 \to \Omega$, $\varphi_2 : G_2 \to \Omega$ are parametrizations such that φ_1 is positive under the first orientation and φ_2 is positive under the second orientation. Let $U \subset \varphi_1(G_1) \cap \varphi_2(G_2)$ be a connected open (relatively in Ω) set. Let ξ_j be a field of unit tangent k-vectors on U determined by the j-th orientation (by the parametrization φ_j). The above conducted dimension considerations show that $\xi_2(x) \in \{\xi_1(x), -\xi_1(x)\}$ σ-almost everywhere on U. From Theorem 35.11 it follows that either $\xi_2 = \xi_1$ σ-almost everywhere, or $\xi_2 = -\xi_1$ σ-almost everywhere.

We can deduce that the orientation of Ω can be reconstructed from the knowledge of the unit tangent k-vector field. A parametrization $\varphi : G \to \Omega$ is *positive* if for almost all $t \in G$

$$\frac{\partial \varphi}{\partial t_1}(t) \wedge \cdots \wedge \frac{\partial \varphi}{\partial t_k}(t) = \xi(\varphi(t)) \operatorname{vol}\Big(\frac{\partial \varphi}{\partial t_1}(t), \dots, \frac{\partial \varphi}{\partial t_k}(t)\Big),$$

and *negative* if the above stated equality holds with the opposite sign.

38.15. Example. Let $\Omega := \{x \in \mathbf{R}^4 : x_1 > 0,\ x_3 = x_1 \cos x_2,\ x_4 = x_1 \sin x_2\}$. We have to find the unit tangent 2-vector field on Ω knowing that the parametrization $\varphi : (0,\infty) \times (-\pi, \pi) \to \Omega$,

$$\varphi(t_1, t_2) = [t_1, t_2,\ t_1 \cos t_2,\ t_1 \sin t_2]$$

is positive. We have

$$w_1 := \frac{\partial \varphi}{\partial t_1}(t) = [1, 0,\ \cos t_2,\ \sin t_2],$$
$$w_2 := \frac{\partial \varphi}{\partial t_2}(t) = [0, 1,\ -t_1 \sin t_2,\ t_1 \cos t_2],$$

and thus

$$\operatorname{vol}^2(w_1, w_2) = 1 + t_1^2 \cos^2 t_2 + t_1^2 \sin^2 t_2 + \cos^2 t_2 + \sin^2 t_2 + t_1^2(\cos^2 t_2 + \sin^2 t_2)^2 = 2 + 2t_1^2$$

and

$$\xi(x) = \frac{[1, 0, x_3/x_1, x_4/x_1] \wedge [0, 1, -x_4, x_3]}{\sqrt{2 + 2\,x_1^2}}.$$

The 2-vector $\xi(x)$ has six coordinates, namely $\xi_{(1,2)}(x)$, $\xi_{(1,3)}(x)$, $\xi_{(1,4)}(x)$, $\xi_{(2,3)}(x)$, $\xi_{(2,4)}(x)$ and $\xi_{(3,4)}(x)$. For example,

$$\xi_{(1,3)}(x) = \frac{\det \begin{pmatrix} 1, & 0 \\ x_3/x_1, & -x_4 \end{pmatrix}}{\sqrt{2 + 2\,x_1^2}} = \frac{-x_4}{\sqrt{2 + 2\,x_1^2}}.$$

38.16. Integration of Differential Forms. Let Ω be an oriented k-dimensional surface, ξ a unit tangent k-vector field and σ a k-dimensional measure on Ω. Then the integral of a differential k-form ω is defined as

$$\int_\Omega \omega := \int_\Omega \langle \omega, \xi \rangle \, \mathrm{d}\sigma$$

provided the right-hand side makes sense. Here, the symbol Ω represents all the structure $(\Omega, \text{orientation})$: Without the knowledge of ξ it would be impossible to determine the sign of the integral.

38.17. Particular Cases. Notions of the preceding chapter can be now revisited from the point of view of the general approach using differential forms and k-vector fields. Let us start with trivial cases.

If $k = 0$, then Ω is a finite set, $\xi(x) = \pm 1$ on Ω and the integral is only a sum of values. A differential 0-form ω is a function and

$$\int_\Omega \omega = \sum_{x \in \Omega} \xi(x)\omega(x).$$

If $k = n$, then Ω is an open subset of $\mathbf{R}^n$, $\xi = e_1 \wedge \cdots \wedge e_n$ (but theoretically, the "unnatural" case $\xi = -e_1 \wedge \cdots \wedge e_n$ should not be excluded), a differential n-form ω is described by one coordinate, and we have

$$\int_\Omega f(x) \, \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n = \int_\Omega f \, \mathrm{d}\lambda.$$

A bit more interesting is the one-dimensional case of a curve. Then $\xi(x)$ is a 1-vector $\mathbf{t}(x)$, a differential 1-form is expressed by a vector field g and

$$\int_\Omega g_1 \, \mathrm{d}x_1 + \cdots + g_n \, \mathrm{d}x_n = \int_\Omega g \cdot \mathbf{t} \, \mathrm{d}s.$$

Finally the case of a normal field is also included: Let $k = n - 1$. Then the vector field $\mathbf{n}(x)$ and the $(n-1)$-vector field $\xi(x)$ are linked by the relation

$$\xi = \mathbf{n}_1 \, e_2 \wedge \cdots \wedge e_n - \mathbf{n}_2 \, e_1 \wedge e_3 \wedge \cdots \wedge e_n + (-1)^n \mathbf{n}_n \, e_1 \wedge \cdots \wedge e_{n-1}.$$

A differential $(n-1)$-form is (again, but in a different way) represented by a vector field g and

$$\int_\Omega g_1 \, \mathrm{d}x_2 \wedge \cdots \wedge \mathrm{d}x_n - g_2 \, \mathrm{d}x_1 \wedge \mathrm{d}x_3 \wedge \cdots \wedge \mathrm{d}x_n$$
$$+ \cdots + (-1)^{n-1} g_n \, \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{n-1} = \int_\Omega g \cdot \mathbf{n} \, d\sigma.$$

38.18. Change of Variable Formula. *Let Ω be an oriented k-dimensional surface, ξ a unit tangent k-vector field on Ω and σ a k-dimensional measure on Ω. Let $\varphi : G \to \Omega$ be a positive parametrization. Then for any differential form*

$$\omega = \sum_{\alpha \in I(k,n)} \omega_\alpha \mathrm{d}x_{\alpha_1} \wedge \cdots \wedge \mathrm{d}x_{\alpha_k} \,,$$

whose coordinates are in $\mathscr{L}^1(\Omega, \sigma)$ we have

$$\begin{aligned} \int_{\varphi(G)} \omega &= \int_G \sum_{\alpha \in I(k,n)} \omega_\alpha \circ \varphi \, d\varphi_{\alpha_1} \wedge \cdots \wedge d\varphi_{\alpha_k} \\ &= \int_G \sum_{\alpha \in I(k,n)} \omega_\alpha \circ \varphi \, \det(\nabla \varphi_{\alpha_1}, \ldots, \nabla \varphi_{\alpha_k}). \end{aligned}$$

Proof. The assertion is an obvious consequence of 34.19 and definitions. ∎

38.19. Example. We evaluate the integral of the differential form $z \, \mathrm{d}x \wedge dy$ over the sphere $\Omega = \{[x, y, z] \in \mathbf{R}^3 \colon x^2 + y^2 + z^2 = 1\}$. We will use the parametrization $\varphi(t, a) = [x(t, a), y(t, a), z(t, a)]$, where

$$\begin{aligned} x(t, a) &= \cos a \, \cos t, \\ y(t, a) &= \cos a \, \sin t, \\ z(t, a) &= \sin a, \end{aligned}$$

$[t, a] \in G := (0, 2\pi) \times (-\frac{1}{2}\pi, \, \frac{1}{2}\pi)$ and the orientation is supposed to make φ positive. Recall, that the uncovered part of the sphere has the $(n-1)$-dimensional measure zero. We have

$$\begin{aligned} dx &= -\sin t \cos a \, \mathrm{d}t - \cos t \sin a \, \mathrm{d}a, \\ dy &= \cos t \cos a \, \mathrm{d}t - \sin t \sin a \, \mathrm{d}a, \\ dz &= \cos a \, \mathrm{d}a. \end{aligned}$$

Thus

$$\begin{aligned} \int_\Omega z \, \mathrm{d}x \wedge dy &= \int_G \sin a(-\sin t \cos a \, \mathrm{d}t - \cos t \sin a \, \mathrm{d}a) \wedge (\cos t \cos a \, \mathrm{d}t - \sin t \sin a \, \mathrm{d}a) \\ &= \int_G \sin^2 a \cos a \sin^2 t \, \mathrm{d}t \wedge da - \sin^2 a \cos a \cos^2 t da \wedge dt \\ &= \int_G \sin^2 a \cos a (\sin^2 t + \cos^2 t) \, \mathrm{d}t \, \mathrm{d}a = 2\pi \int_{-\pi/2}^{\pi/2} \sin^2 a \cos a \, \mathrm{d}a \\ &= 2\pi \int_{-1}^{1} u^2 \, du = \frac{4}{3}\pi. \end{aligned}$$

38.20. Example. Let $\Omega := \{x \in \mathbf{R}^4 \colon x_1^2 + x_2^2 = x_3^2 + x_4^2 = 1, \; x_1 > 0, \, x_3 > 0\}$. Consider the parametrization $\varphi(t) = [\cos t_1, \, \sin t_1, \, \cos t_2, \, \sin t_2]$, $t \in G := (-\frac{\pi}{2}, \frac{\pi}{2})^2$. Suppose that the orientation makes φ positive. Then

$$\nabla \varphi(t) = \begin{pmatrix} -\sin t_1, & 0 \\ \cos t_1, & 0 \\ 0, & -\sin t_2 \\ 0, & \cos t_2 \end{pmatrix}$$

so that

$$\frac{\partial\varphi}{\partial t_1}\wedge\frac{\partial\varphi}{\partial t_2}(t) = \sin t_1\, \sin t_2\, e_1\wedge e_3 - \sin t_1\, \cos t_2\, e_1\wedge e_4 \\ - \cos t_1\, \sin t_2\, e_2\wedge e_3 + \cos t_1\, \cos t_2\, e_2\wedge e_4$$

and

$$\operatorname{vol}\Big(\frac{\partial\varphi}{\partial t_1},\frac{\partial\varphi}{\partial t_2}\Big)(t) = \sqrt{\sin^2 t_1\, \sin^2 t_2 + \sin^2 t_1\, \cos^2 t_2 + \cos^2 t_1\, \sin^2 t_2 + \cos^2 t_1\, \cos^2 t_2} = 1.$$

(The expression under the root is the sum of squares of coordinates of the 2-vector $\frac{\partial\varphi}{\partial t_1}\wedge\frac{\partial\varphi}{\partial t_2}$.) It follows that

$$\xi(x) = x_2x_4\, e_1\wedge e_3 - x_2x_3\, e_1\wedge e_4 - x_1x_4\, e_2\wedge e_3 + x_1x_3\, e_2\wedge e_4$$

is a tangent 2-vector field. We can compute (taking into account the sign of φ only) for example

$$\int_\Omega x_1\,\mathrm{d}x_2\,\mathrm{d}x_4 = \int_G \cos t_1 \det\begin{pmatrix}\cos t_1, & 0\\ 0, & \cos t_2\end{pmatrix} = \int_G \cos^2 t_1\, \cos t_2\,\mathrm{d}t_1\,\mathrm{d}t_2 = \pi.$$

38.21. General Stokes Theorem. *Let $\Omega\subset\mathbf{R}^n$ be a bounded oriented k-dimensional surface with a Lipschitz k-boundary Γ. Let ω be a $\mathscr{C}^1$ differential $k{-}1$-form on a neighborhood of $\overline{\Omega}$. Then*

$$\int_\Gamma \omega = \int_\Omega d\omega.$$

Proof. We present a proof for a differential form ω of the form

$$\omega = \omega_\alpha \mathrm{d}x_{\alpha_1}\wedge\cdots\wedge\mathrm{d}x_{\alpha_{k-1}},$$

where $\alpha\in I(k{-}1,n)$. Using compactness of $\overline{\Omega}$ and the definition of a surface with a Lipschitz k-boundary we get open balls $U(x_1,r_1),\dots U(x_m,r_m)$ covering $\overline{\Omega}$, open sets $G_q\subset\mathbf{R}^k$ and bilipschitz mappings $\varphi_q\colon G\cap\overline{\mathbf{H}^k_-}\to\overline{\Omega}$ such that, for each $q=1,\dots,m$, we have $U(x_q,r_q)\cap\Omega=\varphi_q(G_q\cap\mathbf{H}^k_-)$, $U(x_q,r_q)\cap\Gamma=\varphi_q(G_q\cap\partial\mathbf{H}^k_-)$, $\varphi_{q|G_q\cap\mathbf{H}^k_-}$ is a positive parametrization of $\Omega\cap U(x_q,r_q)$ and $\varphi_{q|G_q\cap\partial\mathbf{H}^k_-}\circ\mathbf{i}$ is a positive parametrization of $\Gamma\cap U(x_q,r_q)$. (Without loss of generality we assume the case (a) of 37.17.) Let χ_q be Lipschitz functions positive on $U(x_q,r_q)$, vanishing outside $U(x_q,r_q)$ and satisfying $\sum\limits_{q=1}^m\chi_q=1$ on $\overline{\Omega}$. (This is in fact a partition of the unity, cf. 39.11. We can choose e.g. $\tilde\chi_q(x)=\max(0,r_q-|x-x_q|)$ and $\chi_q(x)=\tilde\chi_q(x)/\sum\limits_{i=1}^m\tilde\chi_i(x)$.) Let q be fixed, $U=U_q$ and $\varphi=\varphi_q$. We define a function η on $\mathbf{R}^k$ by the formula

$$\eta(t)=\begin{cases}1 & \text{if } t_1\le 0,\\ 1-t_1 & \text{if } 0<t_1<1,\\ 0 & \text{if } t_1\ge 1.\end{cases}$$

Let $\psi_1,\dots,\psi_k$ be Lipschitz function with a compact support on $\mathbf{R}^k$ satisfying

$$\psi_1=(\chi_q\omega_\alpha)\circ\varphi,\quad \psi_2=\varphi_{\alpha_1},\quad \dots,\ \psi_k=\varphi_{\alpha_{k-1}}\qquad \text{on } G\cap\overline{\mathbf{H}^k_-},$$
$$\psi_1=0\qquad \text{on}\overline{\mathbf{H}^k_-}\setminus G,$$
$$\psi_i(t)=\psi_i(0,t_2,\dots,t_k),\qquad \text{if } 0<t_1<1.$$

The existence of these functions follows from McShane's theorem 30.5. Applying Corollary 35.3 to $\eta\psi_1,\psi_2,\dots,\psi_k$ we obtain

$$\begin{aligned}0&=\int_{\mathbf{R}^k}\det(\nabla(\eta\,\psi_1),\nabla\psi_2,\dots,\nabla\psi_k)\,\mathrm{d}t\\&=\int_{\mathbf{R}^k}\psi_1\det(\nabla\eta,\nabla\psi_2,\dots,\nabla\psi_k)\,\mathrm{d}t+\int_{\mathbf{R}^k}\eta\,\det(\nabla\psi_1,\nabla\psi_2,\dots,\nabla\psi_k)\,\mathrm{d}t.\end{aligned}$$

The first integrand can differ from zero only on the strip $\{0\le t_1\le 1\}$, otherwise $\nabla\eta=0$. Fubini's theorem yields

$$\begin{aligned}-\int_{\mathbf{R}^k}\psi_1\det(\nabla\eta,\nabla\psi_2,\dots,\nabla\psi_k)\,\mathrm{d}t&=\int_{\mathbf{R}^k}\Big(\int_0^1\psi_1dt_1\Big)\det\Big(\frac{\partial\psi_i}{\partial t_j}\Big)_{i,j=2}^k\mathrm{d}t_2\dots dt_k\\&=\int_{G\cap\partial\mathbf{H}^k_-}\psi_1\det\Big(\frac{\partial\psi_i}{\partial t_j}\Big)_{i,j=2}^k\mathrm{d}t_2\dots dt_k\\&=\int_\Gamma\chi_q\omega.\end{aligned}$$

The second integrand vanishes on $\mathbf{H}^k_-\setminus G$ (where $\psi_1=0$), on the strip $\{0<t_1<1\}$ (where the partial derivatives $\partial\psi_i/\partial t_1$ are zero) and for $t_1\ge 1$ (where $\nabla\eta=0$). Thus

$$\begin{aligned}\int_{\mathbf{R}^k}\eta\,\det(\nabla\psi_1,\nabla\psi_2,\dots,\nabla\psi_k)\,\mathrm{d}t&=\int_{G\cap\mathbf{H}^k_-}\det(\nabla\psi_1,\dots,\nabla\psi_k)\,\mathrm{d}t\\&=\int_\Omega d(\chi_q\omega).\end{aligned}$$

Summing over $q=1,\dots,m$ we get the desired equality. ■

38.22. Example. Let $\Omega:=\{x\in\mathbf{R}^4\colon 1<x_1^2+x_2^2=x_3^2+x_4^2<4\}$, $\Gamma:=\{x\in\mathbf{R}^4\colon x_1^2+x_2^2=x_3^2+x_4^2=1\}\cup\{x\in\mathbf{R}^4\colon x_1^2+x_2^2=x_3^2+x_4^2=4\}$. We want to verify that for a suitable orientation, Γ is a Lipschitz 3-boundary of Ω. Using the definition, we show this for the point $[1,0,1,0]$. Set

$$\varphi(t)=[(1-t_1)\cos t_2,(1-t_1)\sin t_2,(1-t_1)\cos t_3,(1-t_1)\sin t_3],$$

$t\in G:=(-1,1)\times(-\pi,\pi)^2$. We have

$$\nabla\varphi(t)=\begin{pmatrix}-\cos t_2, & -(1-t_1)\sin t_2, & 0\\ -\sin t_2, & (1-t_1)\cos t_2, & 0\\ -\cos t_3, & 0, & -(1-t_1)\sin t_3\\ -\sin t_3, & 0, & (1-t_1)\cos t_3\end{pmatrix}.$$

Hence we easily compute that $\mathrm{vol}(\frac{\partial\varphi}{\partial t_2}(t), \frac{\partial\varphi}{\partial t_3}(t)) = 1$ (similarly as in Example 38.20) and

$$\mathrm{vol}\big(\frac{\partial\varphi}{\partial t_1}(t),\ \frac{\partial\varphi}{\partial t_2}(t),\ \frac{\partial\varphi}{\partial t_3})(t)\big) = \sqrt{2}(1-t_1)^2.$$

Let the orientations of Ω and Γ be chosen in such a way that the parametrization φ is positive. We easily verify that the unit tangent 3-vector field ξ on Ω has the form

$$\xi(x) = \frac{1}{\sqrt{x_1^2+x_2^2+x_3^2+x_4^2}}\big(x_1\, e_2\wedge e_3\wedge e_4 - x_2\, e_1\wedge e_3\wedge e_4 - x_3\, e_1\wedge e_2\wedge e_4 + x_4\, e_1\wedge e_2\wedge e_3\big)$$

(which corresponds to the normal field

$$\mathbf{n}(x) = \frac{1}{\sqrt{x_1^2+x_2^2+x_3^2+x_4^2}}[x_1, x_2, -x_3, -x_4],$$

see 38.19), and the unit tangent 2-vector field η on Γ is

$$\eta(x) = a\big(x_2x_4\, e_1\wedge e_3 - x_2x_3\, e_1\wedge e_4 - x_1x_4\, e_2\wedge e_3 + x_1x_3\, e_2\wedge e_4\big),$$

where a equals $-1/4$ for $x_1^2+x_2^2=4$ and 1 for $x_1^2+x_2^2=1$.

38.23. Example. Let $\Omega := \{[x,y,z] \in \mathbf{R}^3 \colon z = \sqrt{x^2+y^2} < \frac{3}{5}x+1\}$, $\Gamma := \{[x,y,z] \in \mathbf{R}^3 \colon z = \sqrt{x^2+y^2} = \frac{3}{5}x+1\}$. (A part of the conical surface surrounded by an ellipse .) Choose the orientation Ω in such a way that the z-coordinate of the normal is positive. The evaluation of the integral $\int_\Omega dy\wedge dz$ is simplified when applying Stokes' formula. We use the mapping

$$\varphi(r,t) = [\frac{15}{16} + \frac{25}{16}r\,\cos t,\ \frac{5}{4}r\,\sin t,\ \sqrt{(\frac{15}{16}+\frac{25}{16}r\,\cos t)^2 + (\frac{5}{4}r\,\sin t)^2}$$

which is a positive parametrization; $\varphi((0,1)\times(0,2\pi))$ covers Ω up to a set of two-dimensional measure zero. The mapping $\psi : t \mapsto \varphi(1,t)$, $t\in(0,2\pi)$ is then a positive parametrization of $\Gamma\setminus\{[\frac{5}{2},0,\frac{5}{2}]\}$. Since $z = \frac{3}{5}x+1$ on Γ, we have $\psi_3 = \frac{3}{5}\psi_1+1$ and $d\psi_3 = \frac{3}{5}\,\mathrm{d}\psi_1 = -\frac{3}{5}\,\frac{25}{16}\sin t\,\mathrm{d}t$. Thus

$$\int_\Omega dy\wedge dz = \int_\Gamma y\,\mathrm{d}z = -\int_0^{2\pi}\frac{5\cdot 3\cdot 25}{4\cdot 5\cdot 16}\,\sin^2 t\,\mathrm{d}t = -\frac{75}{64}\pi.$$

38.24. Notes. The exterior multiplication was invented by H. Grassmann in the 19th century. See also 39.24.

39. Integration on Manifolds

A natural generalization of k-dimensional surfaces in $\mathbf{R}^n$ are manifolds. The main difference consist in the fact that a surface is considered as a subset of $\mathbf{R}^n$ while in the case of manifolds we abstract from the embedding into $\mathbf{R}^n$. In this chapter we also introduce notions which we omitted in the previous chapters like the notion of the tangent space.

However, this chapter should not be understood as an introduction to the analysis on manifolds. It contains only a direct way of presentation integration theory. We do not give proofs which, in principle, are mostly the same as proofs of analogous results of the preceding chapters.

Now, the whole theory could be built in the spirit of Lipschitz mappings like in the last chapters. We hope that the reader welcome at least once the simplicity of a $\mathscr{C}^1$-presentation.

39.1. Manifolds. Let Ω be a metrizable topological space. A homeomorphic mapping μ of an open set $U \subset \Omega$ into $\mathbf{R}^k$ is called a (k-dimensional) *chart* (on Ω) provided $\mu(U)$ is an open subset of $\mathbf{R}^k$. The domain U of the chart μ will be denoted by U_μ. A system $\mathscr{A}$ of charts on Ω is called a $\mathscr{C}^\ell$- *atlas* (on Ω) if $\{U_\mu : \mu \in \mathscr{A}\}$ is a covering of Ω and all superpositions $\nu \circ \mu^{-1}$, where $\mu, \nu \in \mathscr{A}$, are of class $\mathscr{C}^\ell$. In this case $\Omega = (\Omega, \mathscr{A})$ is called a (k-dimensional) *manifold* of class $\mathscr{C}^\ell$. If the order of differentiability of a manifold is not specified, then $\mathscr{C}^1$ is tacitly understood.

Let us emphasize that our manifolds are supposed to be metrizable which is not always the case of other authors. Omitting this assumption we get some peculiar examples which are not important from the point of view of applications.

If the set Ω will be equipped by two different atlases $\mathscr{A}_1$, $\mathscr{A}_2$, we understand the manifolds $(\Omega, \mathscr{A}_1)$, $(\Omega, \mathscr{A}_2)$ to be different. (Sometimes it is preferred to identify manifolds whose atlases are in a sense equivalent.)

A set $G \subset \mathbf{R}^k$ is identified with the manifold $(G, \{\mathbf{id}\})$, where $\mathbf{id}$ is the identity mapping on G.

Let $(\Omega, \mathscr{A})$, $(\Omega', \mathscr{A}')$ be two manifolds (not necessarily of the same dimension) and $U \subset \Omega$ be an open set. We say that a mapping $f : U \to \Omega'$ is *of class* $\mathscr{C}^\ell$ *(measurable, differentiable) at a point* $x \in \Omega$ if all superpositions $\nu \circ f \circ \mu^{-1}$, $\mu \in \mathscr{A}$, $\nu \in \mathscr{A}'$, are of class $\mathscr{C}^\ell$ (measurable, differentiable) at $\mu(x)$. A set $E \subset \Omega$ is called *measurable* if $\mu(E \cap U_\mu)$ is λ-measurable for each $\mu \in \mathscr{A}$, and a *null* set if $\lambda(\mu(E \cap U_\mu)) = 0$ for each $\mu \in \mathscr{A}$. A homeomorphic mapping f is called a *diffeomorphism* provided both f and f^{-1} are $\mathscr{C}^1$.

Notice the essential difference with the previous concept of a surface in $\mathbf{R}^n$. If Ω is a subset of $\mathbf{R}^n$, then a metric and a linear structure induced on Ω gives an easy way to a differentiation. Having a manifold, the only information about the structure (except the topological one) is given by the atlas. Without the knowledge of the atlas we are not able to decide whether a mapping of an open subset of $\mathbf{R}^k$ to the manifold is differentiable.

39.2. Embedding. Let $(\Omega, \mathscr{A})$ be a k-dimensional manifold of class $\mathscr{C}^\ell$. Let $f \colon \Omega \mapsto \mathbf{R}^n$ be a diffeomorphism and $f\mathscr{A} = \{\mu \circ f^{-1} : \mu \in \mathscr{A}\}$. The mapping f is called an *embedding* of class $\mathscr{C}^\ell$ into $\mathbf{R}^n$ if $(f(\Omega), f\mathscr{A})$ is again a manifold of class $\mathscr{C}^\ell$.

Most frequently we meet the identical embedding of a manifold being itself a topological subspace of $\mathbf{R}^n$. A possibility to introduce the structure of an embedded manifold (an atlas) on a set $\Omega \subset \mathbf{R}^n$ is to use coordinate charts $x \mapsto [x_{\alpha_1}, \dots, x_{\alpha_k}]$, where α is a multiindex of $\{1, \dots, n\}^k$.

39.3. Orientation. Let $k \geq 1$. A diffeomorphism ψ of an open set $G \subset \mathbf{R}^k$ into $\mathbf{R}^k$ is called *positive* if $J_\psi > 0$ on G, and *negative* if $J_\psi < 0$ on G. We say that $(\Omega, \mathscr{A})$ is an *oriented manifold*, or that $\mathscr{A}$ is an *oriented atlas on* Ω, if all superpositions $\nu \circ \mu^{-1}$, where $\mu, \nu \in \mathscr{A}$, are positive.

Notice that to any oriented manifold $(\Omega, \mathscr{A})$ of dimension $k \geq 1$ there exists a manifold $(\Omega, \overline{\mathscr{A}})$ with an "opposite" orientation, where

$$\overline{\mathscr{A}} = \{[-\mu_1, \mu_2, \dots, \mu_k] \colon [\mu_1, \dots, \mu_k] \in \mathscr{A}\}.$$

We say, that a connected manifold $(\Omega, \mathscr{A})$ of a dimension $k \geq 1$ is *orientable* if there is an oriented atlas $\mathscr{A}' \subset \mathscr{A} \cup \overline{\mathscr{A}}$. A disconnected manifold is orientable if all its connected components are orientable.

0-dimensional manifolds consist of isolated points. The orientation of such a manifold is nothing else than the assignment of a sign plus or minus with every of its points. Let $(\Omega, \mathscr{A})$ be an oriented 0-dimensional manifold. If $\mu \in \mathscr{A}$ and $z \in U_\mu$, then $U_\mu = \{z\}$ and $\mu(z) = 0$. If z is a positive point, then the chart μ is positive and conversely.

39.4. Tangent Spaces and Derivative of a Mapping. Let $(\Omega, \mathscr{A})$ be a k-dimensional manifold and $x \in \Omega$. Let $\mu \in \mathscr{A}$ be a chart whose domain contains x, and $\varphi = \mu^{-1}$. Then the *tangent space* T_x to Ω at the point x is generated by vectors $\frac{\partial \varphi}{\partial t_1}(\mu(x)), \ldots, \frac{\partial \varphi}{\partial t_k}(\mu(x))$. If the manifold Ω is embedded into $\mathbf{R}^n$, it is not necessary to define these vectors since partial derivatives of φ are elements of $\mathbf{R}^n$. There is a lack of such an interpretation for abstract manifolds. To associate a meaning with symbols $\frac{\partial \varphi}{\partial t_j}(\mu(x))$, we can imagine the following construction: The vector $\frac{\partial \varphi}{\partial t_j}(\mu(x))$ is represented as the linear form $f \mapsto \frac{\partial (f \circ \varphi)}{\partial t_j}$ on the vector space of all functions on Ω which are differentiable in x. A tedious computation shows that such a definition of a tangent space is independent of the choice of μ.

Let $(\mathscr{X}, \mathscr{M})$ and $(\mathscr{Y}, \mathscr{N})$ be two manifolds (not necessarily of the same dimension) and $f : \mathscr{X} \to \mathscr{Y}$. If $x \in \mathscr{X}$, $y \in \mathscr{Y}$, $f(x) = y$ and f is differentiable at x, then the *derivative* of f at x is defined as the mapping which assigns with each vector $u \in T_x(\mathscr{X})$ the vector $f'(x)u \in T_y(\mathscr{Y})$. Namely, if $u = \frac{\partial \varphi}{\partial t_j}(\mu(x))$, where $\mu \in \mathscr{M}$ and $\varphi = \mu^{-1}$, we define $f'(x)u = \frac{\partial (f \circ \varphi)'}{\partial t_j}(\mu(x))$.

Let the k-dimensional manifold $(\Omega, \mathscr{A})$ be oriented $\mu \in \mathscr{A}$ and $x \in U_\mu$. The basis $(u_1, \ldots, u_k)$ of the tangent space $T_x(\Omega)$ is called *positive* provided $\det(\mu'(x)u_1, \ldots, \mu'(x)u_k) > 0$, and *negative* if $\det(\mu'(x)u_1, \ldots, \mu'(x)u_k) < 0$. Notice that neither of these notions depend on the particular choice of a chart.

39.5. Example. Let Ω be the sphere $\{x \in \mathbf{R}^n : |x|^2 = 1\}$.

(a) The structure of an oriented manifold of class $\mathscr{C}^\infty$ can be formed on Ω by the atlas of the coordinate charts: $\mathscr{A} = \{\mu_q : q \in \{-n, \ldots, -1, 1, \ldots, n\}\}$, where

$$\begin{aligned}
\mu_1(x) &= [x_2, \ldots, x_n], \quad x_1 > 0,\\
\mu_{-1}(x) &= [-x_2, \ldots, x_n], \quad x_1 < 0,\\
\mu_2(x) &= [x_1, x_3, \ldots, x_n], \quad x_2 > 0,\\
\mu_{-2}(x) &= [-x_1, x_3, \ldots, x_n], \quad x_2 < 0,\\
&\ldots\\
\mu_n(x) &= [x_1, \ldots, x_{n-1}], \quad x_n > 0,\\
\mu_{-n}(x) &= [-x_1, \ldots, x_{n-1}], \quad x_n < 0.
\end{aligned}$$

(b) There is no atlas on Ω composed from a sole chart μ. Indeed, Ω is compact, μ is continuous and thus $\mu(\Omega)$ should be compact as well. However, there are no compact open sets in $\mathbf{R}^{n-1}$.

(c) We find the tangent space at a point $x \in \Omega$, for instance by means of μ_n for x satisfying $x_n > 0$, using $\varphi = \mu_n^{-1}$. Thus, if $G = \{t \in \mathbf{R}^{n-1} : |t| < 1\}$, then $\varphi(t) = [t_1, \ldots, t_{n-1}, \sqrt{1 - |t|^2}]$, $t \in G$, and the tangent space at the point $x = \varphi(t)$ is generated by the vectors

$$[1, 0, \ldots, 0, \frac{-t_1}{\sqrt{1-|t|^2}}] = [1, 0, \ldots, 0, -\frac{x_1}{x_n}], \ldots, [0, \ldots, 0, 1, \frac{-t_{n-1}}{\sqrt{1-|t|^2}}] = [0, \ldots, 0, 1, -\frac{x_{n-1}}{x_n}].$$

(d) Consider the mapping $g\colon \mathbf{R}^2 \to \mathbf{R}^3$ defined as $g(x) = [x_1^2, x_2^2, \sqrt{2}\, x_1 x_2]$ and denote its restriction to the unit circle Ω_2 in $\mathbf{R}^2$ by f. Then f maps Ω_2 to the unit sphere Ω_3 in $\mathbf{R}^3$. The derivative $f'(x)$ maps a vector $u \in T_x(\Omega_2) \subset \mathbf{R}^2$ to the vector $f'(x)(u) \in T_{f(x)}(\Omega_3)$, and $f'(x)(u) = g'(x)u = u_1 \frac{\partial g}{\partial x_1}(x) + u_2 \frac{\partial g}{\partial x_2}(x) = [2u_1x_1,\, 2u_2x_2,\, \sqrt{2}u_1x_2 + \sqrt{2}u_2x_1]$.

39.6. Example. Let $0 < r < R$, and

$$\{[x,y,z] \in \mathbf{R}^3 : (\sqrt{x^2+y^2} - R)^2 + z^2 = r^2\}$$

be an anuloid. We will use the parametrization $\varphi_q(s,t) = [x,y,z]$, $[s,t] \in G_q$, where

$$\begin{aligned} x &= (R + r\cos s)\,\cos t,\\ y &= (R + r\cos s)\,\sin t,\\ z &= r\,\sin s,\\ G_1 &= (0,2\pi) \times (0,2\pi),\\ G_2 &= (0,2\pi) \times (-\pi,\pi),\\ G_3 &= (-\pi,\pi) \times (0,2\pi),\\ G_4 &= (-\pi,\pi) \times (-\pi,\pi). \end{aligned}$$

Then the atlas $\mathscr{A} = \{\varphi_1^{-1}, \ldots, \varphi_4^{-1}\}$ forms the structure of an oriented manifold of class $\mathscr{C}^\infty$ on Ω.

39.7. Example. Let $\Omega = \varphi(G)$, where $G = (-1/2, 1/2) \times (-2\,\pi,\, 2\,\pi)$ and

$$\varphi(s,t) = [(1 + s\cos\frac{t}{2})\cos t,\ (1 + s\cos\frac{t}{2})\sin t,\ s\sin\frac{t}{2}].$$

Then Ω has the shape of the *Möbius strip*. Let $\mathscr{A}$ be the atlas of all $\mathscr{C}^1$ charts on Ω. We will prove that the manifold $(\Omega, \mathscr{A})$ is not orientable. We have

$$\varphi'(s,t) = \begin{pmatrix} \cos t\cos\frac{t}{2}, & -\sin t - s\sin t\cos\frac{t}{2} - \frac{s}{2}\cos t\sin\frac{t}{2} \\ \sin t\cos\frac{t}{2}, & \cos t + s\cos t\cos\frac{t}{2} - \frac{s}{2}\sin t\sin\frac{t}{2} \\ \sin\frac{t}{2}, & \frac{s}{2}\cos\frac{t}{2} \end{pmatrix}.$$

Assume that there exists an oriented atlas $\mathscr{A}' \subset \mathscr{A}$. Let $H = \{[s,t] \in G\colon\ \det\nabla(\mu\circ\varphi)(s,t) > 0$ whenever $\mu \in \mathscr{A}'$ and $\varphi(s,t) \in U_\mu\}$. Then using the connectedness of G it follows that $H = G$ or $H = \emptyset$. Let $\mu \in \mathscr{A}'$, $[-1,0,0] \in U_\mu$. Then $\det(\mu\circ\varphi)$ should have the same sign at the point $[0,-\pi]$ as at the point $[0,\pi]$, which leads to a contradiction.

39.8. Differential Forms. A *differential k-form* on an n-dimensional manifold Ω is defined as a mapping $\omega : x \in \Omega \mapsto \omega(x) \in \Lambda^k(T_x(\Omega))$. The calculation with differential forms on manifolds is transferred into a calculation with differential forms in $\mathbf{R}^k$ by means of a pullback. Let $G \subset \mathbf{R}^m$ be an open set and $\varphi : G \to \Omega$ a $\mathscr{C}^1$ mapping. Let φ be a differential k-form on Ω. Then the differential k-form $\varphi^\sharp\omega$ which is called the *pullback* of φ on G is defined as

$$\left\langle (\varphi^\sharp\omega)(t),\ (u_1,\ldots,u_k)\right\rangle \ = \ \langle \omega(\varphi(t)),\ (\varphi'(t)u_1,\ldots,\varphi'(t)u_k)\rangle\,, \quad u_1\ldots,u_k \in \mathbf{R}^m.$$

The pullback of a differential form on Ω

$$\omega(x) = \sum_{\alpha\in I(k,n)} \omega_\alpha(x) d\mu_{\alpha_1} \wedge \cdots \wedge d\mu_{\alpha_k}\,(x)$$

is a differential form on G

$$\varphi^\sharp\omega(t) = \sum_{\alpha\in I(k,n)} \omega_\alpha(\varphi(t)) d(\mu_{\alpha_1}\circ\varphi) \wedge \cdots \wedge d(\mu_{\alpha_k}\circ\varphi)\,(t),$$

which can be, of course, expressed in coordinates in $\mathbf{R}^m$ (see Example 39.10).

We say that a differential form ω on Ω is of *class* $\mathscr{C}^\ell$ or *measurable* if the same holds for its pullbacks $(\mu^{-1})^\sharp\omega$, $\mu \in \mathscr{A}$.

The *differential* of a differential $k-1$-form ω is defined as a differential k-form $d\omega$ such that $(\mu^{-1})^\sharp d\omega = d((\mu^{-1})^\sharp\omega)$ for each chart $\mu \in \mathscr{A}$. Every $\mathscr{C}^1$ differential form has a differential (this is not entirely easy), and if Ω is an open subset of $\mathbf{R}^n$, it coincides with the differential introduced in the preceding chapter.

If a manifold Ω is embedded into $\mathbf{R}^n$, then any differential form

$$\omega = \sum_{\alpha \in I(k,n)} \omega_\alpha \mathrm{d}x_{\alpha_1} \wedge \cdots \wedge \mathrm{d}x_{\alpha_k}$$

on a neighborhood of Ω induces a differential form $\tilde\omega$ on Ω, namely

$$\tilde\omega = \sum_{\alpha \in I(k,n)} \omega_\alpha d\tilde x_{\alpha_1} \wedge \cdots \wedge d\tilde x_{\alpha_k}\,,$$

where $\tilde x_i$ are the coordinate functions $x \mapsto x_i$ on Ω. Notice that $\omega(x) \in \Lambda^k(\mathbf{R}^n)$ while $\tilde\omega(x) \in \Lambda^k(T_x(\Omega))$. The space $T_x(\Omega)$ can be understood as a subspace of $\mathbf{R}^n$. The difference is immaterial from the point of view of integration. Nevetheless, a certain carefulness is recommended. Indeed, two different elements of $\Lambda^k(\mathbf{R}^n)$ can coincide on $T_x(\Omega)$ so that a differential form on Ω can have more distinct descriptions in coordinates (related to $\mathbf{R}^n$). This phenomenon is demonstrated by the next example.

39.9. Example. Let Ω be the unit circle $\{[x,y] \in \mathbf{R}^2 : x^2+y^2 = 1\}$. Then $x\,\mathrm{d}x + y\,\mathrm{d}y$ induces the zero differential form on Ω.

39.10. Example. Let Ω be the conical surface $\{x \in \mathbf{R}^3 : x_3 = \sqrt{x_1^2+x_2^2}\}$. Let μ be the chart which is the inverse to the mapping $\varphi : G \to \Omega$, $G = (0,1)\times(0,2\pi)$, $\varphi(t) = [\varphi_1(t), \varphi_2(t), \varphi_3(t)] = [t_1\ \cos t_2,\ t_1\ \sin t_2,\ t_1]$, and let $\omega(x) = x_1\,\mathrm{d}x_2 \wedge \mathrm{d}x_3$ be the differential form on Ω. Then

$$\varphi^\sharp\omega = \varphi_1\,\mathrm{d}\varphi_2 \wedge d\varphi_3 = t_1 \cos t_2\,\mathrm{d}(t_1\ \sin t_2) \wedge dt_1 = (-t_1^2\ \cos^2 t_2)dt_1 \wedge dt_2.$$

39.11. Partition of Unity. Let $(\Omega, \mathscr{A})$ be manifold. A system $\{\chi_\tau\}_{\tau\in\mathscr{T}}$ of nonnegative functions of class $\mathscr{C}^1$ on Ω is called a *partition of unity on* Ω (subordinated to a covering $\{U_\mu\}_{\mu\in\mathscr{A}}$) if for every $\tau \in \mathscr{T}$ there is $\mu \in \mathscr{A}$ such that $\overline{\{\chi_\tau > 0\}} \subset U_\mu$ and, in addition, each point $x \in \Omega$ has a neighborhood V with a finite set $\mathscr{T}_V \subset \mathscr{T}$ such that $\chi_\tau = 0$ on V, provided $\tau \notin \mathscr{T}_V$, and $\sum_{\tau\in\mathscr{T}_V} \chi_\tau = 1$ on V. So, $\sum_{\tau\in\mathscr{T}} \chi_\tau = 1$ on Ω and this sum is "locally finite".

The partition of unity exists, cf. 39.22.

39.12. Riemannian Metric. If we want to introduce a k-dimensional measure on a manifold, the idea of copying the definition from $\mathbf{R}^n$ is not the best one. An analogy with the Change of Variable Formula of 34.19 is more straightforward. In this case we need to define a volume of a k-tuple of tangent vectors. The definition of a volume (if we omit the possibility of its axiomatic introduction) is based on an inner product. If the given manifold is embedded into $\mathbf{R}^n$, on each tangent space we have to our disposal an inner product from $\mathbf{R}^n$. In a general case we need to consider an inner product on tangent spaces as an additional structure.

Let $(\mathscr{X}, \mathscr{M})$ be an n-dimensional manifold of class $\mathscr{C}^1$ and $\mathbf{g}$ a mapping associating with every $x \in \mathscr{X}$ a positive definite bilinear form $\mathbf{g}^x$ on $T_x(\mathscr{X})$. Then we can express $\mathbf{g}^x$ in coordinates with respect to a chart $\mu \in \mathscr{M}$ in such a way that for any vectors u, $v \in T_x(\mathscr{X})$ we have

$$\mathbf{g}^x(u,v) = \sum_{i,j=1}^{n} \mathbf{g}^x_{i,j} \hat{u}_i \hat{v}_j ,$$

where $\hat{u}_i = \mu_i'(x)u$ and $\hat{v}_j = \mu_j'(x)v$. If all coordinate functions $x \mapsto \mathbf{g}^x_{i,j}$, $i, j = 1, \ldots, n$, are continuous, we call $\mathbf{g}$ a *Riemannian metric* on $\mathscr{X}$. The structure $(\mathscr{X}, \mathscr{M}, \mathbf{g})$ is called a *Riemannian manifold*. Any manifold of class $\mathscr{C}^1$ admits a Riemannian structure; it follows easily using the partition of unity.

39.13. Example. If $\mathscr{X} := \{[x,y,z] \in \mathbf{R}^3 : x = r \cos z, y = r \sin z, \text{ where } r = \sqrt{x^2+y^2}\}$ is a helix, then $\mathscr{X}$ is a two-dimensional manifold. If we introduce a metric on $T_{[x,y,z]}$ by the formula

$$\mathbf{g}^{[x,y,z]}(u,v) = Pu \cdot Pv,$$

where P is the projection $[x,y,z] \mapsto [x,y]$, then we get a Riemannian manifold which does not have an isometric embedding to $\mathbf{R}^n$. (The mapping P is of course locally an isometric embedding into $\mathbf{R}^2$ but globaly it is not one-to-one.) Such manifolds are useful in complex analysis, the example demonstrates that non-imbedded manifolds are not only "useless abstractions".

39.14. Example. Let Ω be an open unit circle in $\mathbf{R}^2$. Set

$$\mathbf{g}^x(u,v) = u \cdot in + \frac{(u \cdot x)(v \cdot x)}{|x|^2}, \quad x \in \Omega,\ u, v \in \mathbf{R}^2.$$

Then $\mathbf{g}$ is a Riemannian metric which gives the shape of a hemisphere to the manifold Ω. Indeed, the mapping $f : x \to [x_1, x_2, \sqrt{x_1^2 + x_2^2}]$ which maps $\mathscr{X}$ to the "actual hemisphere" $f(\mathscr{X})$ (endowed with the Euclidean inner product) preserves the inner product. Namely, for all $u, v \in \mathbf{R}^2$ and $x \in \Omega$ we have

$$(f'(x)u) \cdot (f'(x)v) = \mathbf{g}^x(u,v),$$

hence f is an "isometric mapping". On the other hand, Ω is not isometric to any open subset of $\mathbf{R}^2$ (it is not possible to "make the hemisphere flat"). From the geometrical point of view, the shape of the manifold expressed by the Riemannian metric is more important than the original "underlying space".

39.15. k-dimensional Measures on Riemannian Manifolds. Suppose that $(\mathscr{X}, \mathscr{A}, \mathbf{g})$ is a Riemannian manifold, $x \in \mathscr{X}$ and $(u_1, \ldots, u_k) \in (T_x(\mathscr{X}))^k$. We define a volume of this k-tuple of vectors similarly as in 34.10:

$$\operatorname{vol}(u_1, \ldots, u_k) = \det(\mathbf{g}^x(u_i, u_j))_{i,j=1}^k .$$

This immediately introduces also the volume of a linear mapping $L : \mathbf{R}^k \to T_x(\mathscr{X})$. Let σ be a measure on the σ-algebra of all measurable subsets of $\mathscr{X}$. We say that σ is a *k-dimensional measure on X* if for each $\mu \in \mathscr{A}$ and each measurable set $E \subset U_\mu$ we have

$$\sigma E = \int_{\varphi^{-1}(E)} \operatorname{vol} \varphi'(t)\, \mathrm{d}t\,, \quad \text{where} \quad \varphi = \mu^{-1}\,.$$

Since the integral does not depend on a particular choice of μ, the partition of unity leads to the existence of a "k-dimensional measure" on a k-dimensional Riemannian manifold.

39.16. Integration of Differential Forms on Riemannian Manifolds. Let $(\Omega, \mathscr{A}, \mathbf{g})$ be a k-dimensional Riemannian manifold. We define the *unit tangent k-vector* $\xi(x)$ to Ω at a point $x \in \Omega$ by the formula

$$\xi(x) = \frac{u_1 \wedge \cdots \wedge u_k}{\operatorname{vol}(u_1, \ldots, u_k)},$$

where $(u_1, \ldots, u_k)$ is a positive basis of $T_x(\Omega)$ (the definition does not depend on the choice of the base). Now, we can introduce the integration of differential forms by means of integration by k-dimensional measure σ on Ω similarly as we have proceeded in 38.16: Namely, if ω is an integrable differential form on Ω, then

$$\int_\Omega \omega = \int_\Omega \langle \omega, \xi \rangle \, \mathrm{d}\sigma.$$

39.17. Example. We evaluate the integral

$$\int_\Omega \frac{r x_2}{\sqrt{4r^4 + 5r^2 + 1}} \, \mathrm{d}\sigma,$$

where $r = \sqrt{x_1^2 + x_2^2}$, $\Omega := \{x \in \mathbf{R}^4 : x_1 = r\cos x_4,\ x_2 = r \sin x_4,\ x_3 = r^2,\ x_4 \in (0, \pi),\ r \in (0,1)\}$ and σ is a two-dimensional measure on Ω.

Let (e_s, e_t) be the canonical basis of $\mathbf{R}^2$ and $(e_1, \ldots, e_4)$ the canonical basis of $\mathbf{R}^4$. The manifold Ω will be parametrized by the mapping $\varphi(s,t) = [s\cos t,\ s \sin t,\ s^2,\ t]$, $[s,t] \in G := (0,1) \times (0, \pi)$. We introduce the structure of an oriented manifold on Ω by the atlas $\{\varphi^{-1}\}$. Denote $L = \varphi'(s,t)$, $w = Le_s \wedge Le_t$. Then

$$\begin{aligned} w &= \frac{\partial \varphi}{\partial s} \wedge \frac{\partial \varphi}{\partial t}(s,t) = [\cos t,\ \sin t,\ 2s,\ 0] \wedge [-s\sin t,\ s\cos t,\ 0, 1] \\ &= [\frac{x_1}{r},\ \frac{x_2}{r},\ 2r,\ 0] \wedge [-x_2, x_1, 0, 1] \\ &= r\, e_1 \wedge e_2 + 2r x_2\, e_1 \wedge e_3 + \frac{x_1}{r}\, e_1 \wedge e_4 - 2r x_1\, e_2 \wedge e_3 + \frac{x_2}{r}\, e_2 \wedge e_4 + 2r\, e_3 \wedge e_4, \end{aligned}$$

and

$$|w|^2 = r^2 + 4r^2 x_2^2 + \frac{x_1^2}{r^2} + 4r^2 x_1^2 + \frac{x_2^2}{r^2} + 4r^2 = 4r^4 + 5r^2 + 1.$$

For the unit tangent k-vector we have

$$\xi = \frac{w}{|w|},$$

the integrand is expressed as

$$\frac{r x_2}{\sqrt{4r^4 + 5r^2 + 1}} = \frac{1}{2} \langle \mathrm{d}x_1 \wedge \mathrm{d}x_3,\ \xi \rangle,$$

and thus

$$\begin{aligned} \int_\Omega \frac{r x_2}{\sqrt{4r^4 + 5r^2 + 1}} \, \mathrm{d}\sigma &= \frac{1}{2} \int_\Omega \mathrm{d}x_1 \wedge \mathrm{d}x_3 \\ &= \frac{1}{2} \int_G (\cos t \, \mathrm{d}s - s \sin t \, \mathrm{d}t) \wedge 2s\, \mathrm{d}s = \int_G s^2 \sin t \, \mathrm{d}s \, \mathrm{d}t = \frac{2}{3}. \end{aligned}$$

39.18. Integration on General Manifolds. The definition 39.16 is a logical conclusion of the approach of the preceding chapter, where the presence of the inner product was quite obvious. Nevertheless, for the purpose of integration of differential forms on manifolds neither Riemannian structure nor a k-dimensional measure are needed. Indeed, we can realize that the integral expressions given by the Change of Variable Formula do not depend on these structures. In the general case we can proceed as follows: Let $(\Omega, \mathscr{A})$ be a k-dimensional oriented manifold. Let ω be a measurable differential k-form on $\mathscr{A}$ and E a measurable subset of Ω. The integral $\int_E \omega$ is defined in two steps: First, assume $E \subset U_\mu$ for some $\mu \in \mathscr{M}$. Then we define

$$\int_E \omega = \int_{\mu(E)} (\mu^{-1})^\sharp \omega.$$

From the Change of Variable Formula it follows that this integral (if it makes sense) does not depend on the choice μ.

The second step is based on the partition of unity. Let $\{\chi_\tau\}_{\tau \in \mathscr{T}}$ be a partition of unity on $(\Omega, \mathscr{A})$.

For any measurable set $E \subset \Omega$ denote

$$I(E) = \sum_{\tau \in \mathscr{T}} \int_{E \cap \{\chi_\tau > 0\}} \chi_\tau \omega$$

(the expression $I(E)$ does not necessarily makes sense). We say, that the *integral* $\int_E \omega$ *converges*, or that the differential form ω is *integrable* on E if for any measurable set $E' \subset \Omega$, the expression $I(E')$ makes sense and it is a finite number, and for any sequence $\{E_q\}$ of parwise disjoint measurable sets $E_q \subset E$, the series

$$\sum_{q=1}^{\infty} I(E_q)$$

converges (absolutely). If the integral $\int_E \omega$ converges, we set

$$\int_E \omega = I(E).$$

Since we were careful enough, the value of such a defined integral depends neither on the "partition" of E, nor on the partition of unity. On the other hand, it depends on the orientation of Ω: The reverse orientation forces the converse of the sign of the integral.

Let $G \subset \mathbf{R}^k$ be an open set and $\varphi : G \to \Omega$ a diffeomorphism. We say, that φ is a *positive parametrization* if all superpositions $\mu \circ \varphi$, $\mu \in \mathscr{A}$, have a positive Jacobian. For positive parametrizations the following change of variable formula is valid:

$$\int_{\varphi(G)} \omega = \int_G \varphi^\sharp \omega$$

provided either of these integrals exists.

39.19. Example. Let Ω be the set of all decreasing solution of the differential equation $x''(s) - \frac{(x'(s))^2}{x(s)} = 0$ satisfying the condition $0 < x(0) < 1$. Let $\mathscr{A} = \{\mu : \Omega \to \mathbf{R}^2 :$ there are a, $b \in \mathbf{R}$, such that $a < b$ and $\mu(x) = [x(a), x(b)]$ for all $x \in \Omega\}$. Then $(\Omega, \mathscr{A})$ is a two-dimensional oriented manifold and $\varphi : t \mapsto e^{t_2 s + t_1}$, $t \in (-\infty, 0)^2$ is a positive parametrization of Ω. The tangent space $T_x(\Omega)$, $x = \varphi(t)$, is representable as a two-dimensional vector space of functions generated by the functions $\frac{\partial \varphi}{\partial t_1}(t) : s \mapsto e^{t_2 s + t_1}$ and $\frac{\partial \varphi}{\partial t_2}(t) : s \mapsto s\, e^{t_2 s + t_1}$. For each $\tau \in \mathbf{R}$, let ε_τ be the function on Ω defined as $\varepsilon_\tau(x) = x(\tau)$. Then $\omega := d\varepsilon_0 \wedge d\varepsilon_1$ is a differential form which associates with u_1, $u_2 \in T_x(\Omega)$ the number

$$\det \begin{pmatrix} u_1(0), & u_2(0) \\ u_1(1), & u_2(1) \end{pmatrix}.$$

We have

$$\begin{aligned} \varphi^\sharp \omega &= d(e^{t_1}) \wedge d(e^{t_1 + t_2}) = e^{t_1}\, \mathrm{d}t_1 \wedge (e^{t_1+t_2}\, \mathrm{d}t_1 + e^{t_1+t_2}\, \mathrm{d}t_2) \\ &= e^{2t_1 + t_2}\, \mathrm{d}t_1 \wedge dt_2, \end{aligned}$$

so that

$$\int_\Omega \omega = \int_{(-\infty,0)^2} e^{2t_1+t_2}\, \mathrm{d}t_1\, \mathrm{d}t_2 = \frac{1}{2}.$$

39.20. Introduction to General Stokes' Theorem on Manifolds. Let $(\mathscr{X}, \mathscr{A})$ be a k-dimensional oriented manifold and $\Omega \cup \Gamma$ be a compact subset of $\mathscr{X}$. Let $(\Gamma, \mathscr{B})$ be a $(k-1)$-dimensional oriented manifold. We suppose that for each point $z \in \Omega \cup \Gamma$ there exist an open set $G \subset \mathbf{R}^k$, a neighborhood U of the point z and a homeomorphic mapping $\varphi : G \cap \overline{\mathbf{H}^k_-} \to \Omega \cup \Gamma$, such that $z \in \varphi(\partial \mathbf{H}^k_-)$, $\varphi(G \cap \mathbf{H}^k_-) = \Omega \cap U$, $\varphi(G \cap \partial \mathbf{H}^k_-) = \Gamma \cap U$ and one of the following cases occurs:

(a) $\varphi_{|G \cap \mathbf{H}^k_-}$ is a positive parametrization of $\Omega \cap U$ and $\varphi_{|G \cap \partial \mathbf{H}^k_-} \circ \mathbf{i}$ is a positive parametrization of $\Gamma \cap U$.

(b) $\varphi_{|G \cap \mathbf{H}^k_-}$ is a negative parametrization of $\Omega \cap U$ and $\varphi_{|G \cap \partial \mathbf{H}^k_-} \circ \mathbf{i}$ is a negative parametrization of $\Gamma \cap U$.

(Recall that, by the conventions of this chapter, any parametrization is a diffeomorphism.)

39.21. General Stokes' Theorem on Manifolds. *Let ω be a $\mathscr{C}^1$ differential $(k-1)$-form on $\mathscr{X}$. Then*

$$\int_\Gamma \omega = \int_\Omega d\omega.$$

39.22. Existence of the Partition of Unity. Let $(\Omega, \mathscr{A})$ be a k-dimensional (topological) manifold. We introduce a temporary term of an *admissible family of functions* for a system $\{f_\tau\}_{\tau \in \mathscr{T}}$ of nonnegative functions on Ω which satisfies the following conditions: For every τ from the index set $\mathscr{T}$ there exists $\mu \in \mathscr{A}$ such that $\overline{\{f_\tau > 0\}} \subset U_\mu$, and $f_\tau \circ \mu^{-1}$ is an infinitely differentiable function. Further, each point $x \in \Omega$ has a neighborhood V with a finite set $\mathscr{T}_V \subset \mathscr{T}$ such that $f_\tau = 0$ on V provided $\tau \notin \mathscr{T}_V$. The only requirement on partition of unity being not satisfied by an admissible family of function is that the sum is 1. However, if $\{f_\tau\}_{\tau \in \mathscr{T}}$ is an admissible family of function whose sum S is positive, then $\{f_\tau / S\}_{\tau \in \mathscr{T}}$ is a partition of unity. Its quality depends on the quality of the atlas. If the atlas $\mathscr{A}$ is $\mathscr{C}^\ell$, or locally Lipschitz, then also the partition of unity will have the same property.

In the next step let $K \subset W$ be subsets of Ω, K compact, W open. We show that there is an admissible family of function (even a finite one) whose sum is positive on K and vanishing outside W. For each point $a \in K$ we find $\mu_a \in \mathscr{A}$ and a radius $r_a > 0$ so that $a \in U_{\mu_a}$ and $B(\mu_a(a), 2r_a) \subset \mu_a(U_{\mu_a} \cap W)$. Now set

$$f_a(x) = \begin{cases} e^{1/(|\mu_a(x)-\mu_a(a)|^2 - r_a^2)} & \text{if } x \in \mu_a^{-1}(U(\mu_a(x), r_a)), \\ 0 & \text{in remaining cases.} \end{cases}$$

Then $\{\{f_a > 0\} : a \in K\}$ is a covering of K and taking into account the compactness of K we can select a finite set $\{f_{a_1}, \dots, f_{a_m}\}$ forming an admissible family of functions, whose sum is positive on K. We have yet solved the existence of the partition of unity provided Ω is compact.

Recall that, by our definitions, Ω is a metrizable space. If Ω is connected, then the Topological Lemma 39.23 yields the existence of compact sets K_q and open sets W_q ($q \in \mathbf{N}$) such that $K_q \subset W_q$, $\Omega = \bigcup_{q=1}^{\infty} K_q$, and each point has a neighborhood which intersects only a finite number of sets W_q. For each couple (K_q, W_q) we find an admissible family of functions by the preceding procedure. The union with respect to q will be an admissible family of functions whose sum is positive on Ω. This solves the existence problem if Ω is connected. If Ω is not connected, then its topological components are connected submanifolds Ω. By a simple "union" of partitions of unity on components we obtain the partition of unity on Ω.

39.23. Topological Lemma. *Let (P, ρ) be a connected locally compact metric space. Then there exists a sequence $\{K_q\}$ of compact subsets of P and a sequence $\{W_q\}$ of open subsets of X such that $X = \bigcup_{q=1}^{\infty} K_q$ and each point P has a neighborhood intersecting only finitely many sets W_q.*

Proof. We may assume that P is not compact, for otherwise there is nothing to prove. Further, we consider an equivalent metric in which P is bounded. Given $x \in P$, there is a radius $r(x)$ such that $B(x, 2r(x))$ is compact and $B(x, 4r(x))$ is not compact. We construct recursively a sequence $\{V_q\}$ of open relatively compact subsets of P. Choose $x_0 \in P$ and set $V_1 = U(x_0, 2r_0)$. Assume that $V_1, \dots, V_q$ were already constructed. Thanks to compactness of $\overline{V_q}$ it follows that there is a finite system $\{U(x_j, r_j)\}$ of balls selected from $\{U(x, r(x)) : x \in \overline{V}_q\}$ such that it covers $\overline{V}_q$. Set $V_{q+1} = \bigcup_j U(x_j, 2r_j)$. The resulting sequence satisfies $\overline{V}_q \subset V_{q+1}$. We prove by a contradiction that $V := \bigcup_{q=1}^{\infty} V_q = P$. Suppose $V \neq P$. Since P is connected, there exists $z \in \partial V$. Set $R = r(z)/3$, find $x \in U(z, R) \cap V$ and q such that $x \in V_q$. Further find $y \in \overline{V_q}$ and $r = r(y)$ such that $x \in U(y, r)$ and $U(y, 2r) \subset V_{q+1}$. Since $z \notin V$, it follows that $\rho(z, y) \geq 2r$. We have

$$2r \leq \rho(y, z) \leq \rho(y, x) + \rho(x, z) \leq r + R,$$

thus $r \leq R$. If $t \in B(y, 4r)$, then

$$\rho(t, z) \leq \rho(t, y) + \rho(y, z) \leq 4r + r + R \leq 6R,$$

so that $B(y, 4r(y)) \subset B(z, 2r(z))$. This is a contradiction, because the ball $\overline{B(y, 4y)}$ is not compact and the ball $\overline{B(z, 2r(z)}$ is compact. We have proved that $V = P$. To finish the proof it is enough to set $K_1 = \overline{V}_1$, $W_1 = V_2$, $W_2 = V_3$, $K_q = \overline{V_q} \setminus V_{q-1}$ for $q \geq 2$ and $W_q = V_{q+1} \setminus \overline{V}_{q-2}$ for $q \geq 3$. ■

39.24. Notes. The modern theory of manifolds is based on ideas of G. F. B. Riemann.

The roots of the topics of Chapter G go back to the 19th century and are connected with famous names of outstanding mathematicians. From an extensive bibliography we recommend M. Berger and B. Gostiaux [*1988], L. Boček [Boč], I. Černý and J. Mařík [ČM II], H. Federer [*1969], W. Fleming [*1965], O. Kowalski [Kow], L. Krump, V. Souček and J. A. Těšínský [KST], F. Moran [*1988], R. Sikorski [Sik], L. Simon [*1983].

H. Vector Integration

40. Measurable Functions

In many branches of analysis we need to integrate functions having values in vector spaces. Throughout this chapter we consider the case when $(\Omega, \mathscr{S}, \mu)$ is a measure space and X is a Banach space. Having a mapping f from Ω to X (a"vector function") we would like to define an integral $\int_\Omega f \,\mathrm{d}\mu$ in a reasonable way. In principle, there are two possibilities:

(1) to utilize real (or complex) case arising by a composition $\varphi \circ f$ where φ varies over the set of all functionals of the dual space X^*,

(2) to try to choose an appropriate definition of the Lebesgue integral suited for the vector case.

Note that both methods are also common in other branches of analysis.

In this rather short and informative chapter we will follow both ways of defining vector integrals. Note that basic knowledge of main topics of functional analysis (like the Hahn–Banach theorem, the Riesz-Fréchet representation theorem of bounded linear functionals on Hilbert spaces, or the notion of a reflexive space) is necessary for a good understanding of vector integration.

It seems that the first attempt to define a vector integral of the Riemann type is due to Graves in 1927. His definition is only a suitable modified original Riemann definition and its main idea is explained in Exercise 43.7.

In what follows, X will denote a Banach space and $(\Omega, \mathscr{S}, \mu)$ will be a measure space, where μ is supposed to be a complete probability measure ($\mu\Omega = 1$).

First of all we concentrate on the notion of measurable functions.

40.1. Measurable Functions. A function $f : \Omega \to X$ is termed

– *simple* provided there exist $x_1, ..., x_n \in X$ and $E_1, ..., E_n \in \mathscr{S}$ such that $f = \sum_i x_i c_{E_i}$,

– *measurable* if there exists a sequence $\{f_n\}$ of simple functions such that $\lim f_n(\omega) = f(\omega)$ for μ-almost all $\omega \in \Omega$ (i.e. if $f_n(\omega)$ converge to $f(\omega)$ in the norm of the space X for μ-almost all $\omega \in \Omega$),

– *weakly* or *scalarly measurable* if functions $\varphi \circ f$ are measurable for each continuous linear functional $\varphi \in X^*$.

40.2. Remarks. 1. Note that we defined measurable functions according to the characterization given in Exercise 5.7. It is clear that the common definition of measurability ($\{\omega \in \Omega : f(\omega) < \alpha\} \in \mathscr{S}$ for each $\alpha \in X$) cannot be used in the case of vector functions. There are other equivalent definitions of measurability of real functions like those in Exercise 3.6.a ($\{\omega \in \Omega : f(\omega) \in B\} \in \mathscr{S}$ for each open, or Borel set $B \subset \mathbf{R}$, respectively) suited to the case of vector functions. Having the last definition in mind the class of all "measurable" functions coincides with the class of all measurable functions (and also with the class of all weakly measurable functions according to Pettis' theorem) provided X is separable. In a non-separable case the sum of two "measurable" functions according to the last definition need not be even"measurable".

2. Assume f_n are measurable, $f_n \to f$ μ-almost everywhere and $\lambda \in \mathbf{R}$. Prove that the functions $f_1 + f_2, \lambda f_1, f$ are also measurable. The same is true for weakly measurable functions.

The relationship between measurable and weakly measurable functions is described in the next theorem.

40.3. Pettis' Theorem. *A function $f : \Omega \to X$ is measurable if and only if f is weakly measurable and there is a μ-null set $E \in \mathscr{S}$ such that $f(\Omega \setminus E)$ is a separable subset of X. In particular, if X is separable the notions of measurability and weak measurability coincide.*

Proof. Suppose f_k are simple functions, $\mu(E) = 0$ and $f_k \to f$ on $\Omega \setminus E$. Since $f(\Omega \setminus E)$ is a subset of the closure of $\bigcup f_k(\Omega)$ and $f_k(\Omega)$ are finite sets, it follows that $f(\Omega \setminus E)$ is separable.

The proof that $\varphi \circ f$ is measurable provided $\varphi \in X^*$ is easy. Indeed, the assertion is true if f is a simple function and for the general case we pass to the limit.

Now assume that f is weakly measurable and that the set $f(\Omega \setminus E)$ is separable for $E \in \mathscr{S}$, $\mu E = 0$. In the first step of the proof we show that the function $\omega \mapsto \|f(\omega)\|$ is measurable. To this end let $\{x_n\} \subset f(\Omega \setminus E)$ be a dense countable set. Using the Hahn-Banach theorem there are $\varphi_n \in X^*, \|\varphi_n\| = 1$ such that $\varphi_n(x_n) = \|x_n\|$. It remains to show that

$$\|f(\omega)\| = \sup_n |\varphi_n(f(\omega))|$$

for each $\omega \in \Omega \setminus E$. Obviously, $|\varphi_n(f(\omega)| \leq \|\varphi_n\| \cdot \|f(\omega)\| = \|f(\omega)\|$. To prove the reverse inequality, let $\omega \in \Omega \setminus E$, $n \in \mathbf{N}$ and $\varepsilon > 0$ be given. There is x_n such that $\|f(\omega) - x_n\| < \varepsilon$. Then

$$\begin{aligned} |\, \|f(\omega)\| - \varphi_n(f(\omega))| &\leq |\, \|f(\omega)\| - \|x_n\|\, | + |\, \|x_n\| - \varphi_n(x_n)| \\ &\quad + |\varphi_n(x_n) - \varphi_n(f(\omega))| \\ &\leq \|f(\omega) - x_n\| + |\varphi_n(x_n - f(\omega))| \\ &\leq \varepsilon + \|\varphi_n\| \cdot \|x_n - f(\omega)\| < 2\varepsilon. \end{aligned}$$

Similarly, we can prove that the functions $g_n : \omega \mapsto \|f(\omega) - x_n\|$ are measurable. Now fix $k \in \mathbf{N}$, put $E_n^k = \{\omega \in \Omega : g_n(\omega) < \frac{1}{k}\}$ (obviously $E_n^k \in \mathscr{S}$) and define

$$h_k(\omega) = \begin{cases} x_n & \text{if} \quad \omega \in E_n^k \setminus \bigcup_{j<n} E_j^k\,, \\ 0 & \text{otherwise.} \end{cases}$$

If $\omega \in \Omega \setminus E$, then $\|f(\omega) - h_k(\omega)\| < \frac{1}{k}$. Thus, we have shown that there is a sequence $\{h_k\}$ such that $h_k \to f$ μ-almost everywhere and each function h_k has a countable range. Since such functions are measurable, the assertion easily follows. ∎

To get a good understanding of vector integration we introduce examples. Before proceeding let us agree on the following notations:

- c_0 is the Banach space of all real sequences $x = \{x_n\}$ satisfying $x_n \to 0$ equipped with the sup-norm $\|x\| = \max_n |x_n|$,
- l^p, $1 \leq p < \infty$ is the space of all sequences $x = \{x_n\}$ for which $\|x\|_p := \big(\sum_n |x_n|^p\big)^{1/p} < \infty$,

- l^∞ will denote the space of all bounded sequences $x = \{x_n\}$ equipped with the norm $\|x\|_\infty := \sup_n |x_n|$,
- $l^2([0,1])$ is the Hilbert space of all real functions f on $[0,1]$ vanishing off a countable set such that $\sum_{t\in[0,1]} |f(t)|^2 < \infty$, equipped with the inner product $(f,g) = \sum_t f(t)g(t)$.

40.4. Example. Consider $X = l^2([0,1])$ and the measure space $([0,1], \mathfrak{M}, \lambda)$. Let $\{e_t : t \in [0,1]\}$ be the usual orthonormal base of X ($e_t(x) = 1$ for $x = t$, $e_t(x) = 0$ otherwise). Define the mapping $f\colon [0,1] \to X$ as $f(t) = e_t$. If $\varphi \in X^*$, the Riesz-Fréchet representation theorem on Hilbert spaces implies the existence of a uniquely determined $a \in X$ such that $\varphi(x) = (x,a)$ for every $x \in X$. Hence $\varphi(f(t)) = (e_t, a) = a(t)$ and the set $\{t \in [0,1] : (e_t, a) \neq 0\}$ is countable. We can see that $\varphi \circ f = 0$ almost everywhere and therefore f is a weakly measurable function. On the other hand, $\|e_t - e_s\| = \sqrt{2}$ for $t \neq s$. It follows that $f([0,1] \setminus E) = \{e_t : t \in [0,1] \setminus E\}$ is separable if and only if the set $[0,1] \setminus E$ is countable. Thus, there is no set $E \subset [0,1]$ of measure zero for which $f([0,1] \setminus E)$ would be separable and Pettis' theorem implies that f is not measurable. (Further examples can be found in Exercises 43.7.e and f.)

40.5. Notes. The theory of vector measures and integration was developed in an essential way during 1930's. Famous Theorem 40.3 appears in B.J. Pettis [1938].

41. Vector Measures

In this chapter we touch briefly on measures whose values are in a given Banach space X. Before doing this we take notice of a convergence of series in Banach spaces.

41.1. Absolute and Unconditional Convergence. Let $\{x_n\}$ be a sequence of elements of a Banach space X. We say that a (formal) series $\sum x_i = \sum_{i=1}^{\infty} x_i$ is

- *convergent* if there is a limit $\lim_n \sum_{i=1}^{n} x_i$ (we write $\sum_{i=1}^{\infty} x_i = \lim_n \sum_{i=1}^{n} x_i$),
- *absolutely convergent* if $\sum_{i=1}^{\infty} \|x_i\| < +\infty$,
- *unconditionally convergent* to $x \in X$ if $\sum_n x_{P(n)} = x$ whenever P is a one-to-one mapping of $\mathbf{N}$ onto $\mathbf{N}$.

Each absolutely convergent series converges (X is a complete space !) even unconditionally. On the other hand, every unconditionally convergent series converges absolutely provided $\dim X < +\infty$ (Riemann's theorem) while in infinite dimensional spaces this assertion is no longer true (consider the example $x_n = (0, \dots, 0, \frac{1}{n}, 0, 0, \dots) \in c_0$).

In what follows, $(\Omega, \mathscr{S}, \mu)$ will stand for a fixed measure space and X a given Banach space.

41.2. Vector Measures. A vector-valued set function $F : \mathscr{S} \to X$ is called an *additive* (σ-*additive*) vector measure if $F(\emptyset) = 0$ and

$$F\Big(\bigcup_n E_n\Big) = \sum_n F(E_n)$$

for every finite (countable) sequence of pairwise disjoint sets $E_n \in \mathscr{S}$. In case of σ-additive measures the convergence of a series in the definition is understood in the sense as above. Realize that this convergence is even an unconditional one. Any σ-additive vector measure will be shortly called a *vector measure*.

41.3. Examples. In all examples, $(\Omega, \mathscr{S}, \mu)$ stands for the measure space $([0,1], \mathfrak{M}, \lambda)$.

1. If $X = L^p([0,1])$, $1 \le p \le \infty$, and $F : E \mapsto c_E$ for $E \in \mathfrak{M}$, then F is a σ-additive vector measure.

2. Let L be a continuous linear operator from $L^1[0,1]$ into a Banach space X. If $F(E) := L(c_E)$ for $E \in \mathfrak{M}$, then F is again a σ-additive vector measure. Indeed, a moment's reflection shows that

$$\|F(\bigcup_{j=1}^{\infty} E_j) - \sum_{j=1}^{n} F(E_j)\| = \|F(\bigcup_{j=n+1}^{\infty} E_j)\| \le \lambda(\bigcup_{j=n+1}^{\infty} E_j)\,\|L\|\,.$$

3. Let $T : L^\infty[0,1] \to X$ be a continuous linear operator, $F(E) := T(c_E)$ for $E \in \mathfrak{M}$. Then F is an additive vector measure which may not be σ-additive. To see an example, let T be the Hahn-Banach extension of the functional $\varphi : x \mapsto x(\frac{1}{2})$ from $\mathscr{C}[0,1]$ to $L^\infty[0,1]$ (notice that $X = \mathbf{R}$!)

41.4. Absolute Continuity. An additive vector measure $F : \Omega \to X$ is said to be *absolutely continuous* with respect to a measure μ on $\mathscr{S}$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\|F(E)\| < \varepsilon$ whenever $\mu E < \delta$.

41.5. Theorem (Pettis). *A σ-additive vector measure F on $\mathscr{S}$ is absolutely continuous with respect to a finite measure μ on $\mathscr{S}$ if and only if $F(E) = 0$ whenever $\mu E = 0$.*

Proof. The necessity is obvious. For the proof of the converse we can use analogous reasoning as in Exercise 8.22.b: We assume the existence of an $\varepsilon > 0$ and a sequence $\{E_n\} \subset \mathscr{S}$ for which

$$\|F(E_n)\| \ge \varepsilon \quad \text{and} \quad \mu E_n < 2^{-n}.$$

To reach the contradiction consider compositions $\varphi \circ F$ where $\varphi \in X^*$. Now to complete the proof we need "uniform" estimates with respect to φ and this is more difficult than the conclusion in Exercise 8.22.b. ∎

41.6. Exercise. Let μ be the counting measure on $\mathbf{N}$ and $X = l^2$. Show that $F : E \mapsto \{\frac{1}{n} c_E(n)\}$, $E \subset \mathbf{N}$, is a σ-additive vector measure.

41.7. Variation of Vector Measures. Let $F : \Omega \to X$ be a vector measure. According to Theorem 6.9 we define the *variation* of F as

$$|F(E)| := \sup\{\sum_{k=1}^{n} \|F(A_k)\| : A_k \in \mathscr{S}\,, A_i \cap A_j = \emptyset \text{ for } i \ne j\,,\ \bigcup_{k=1}^{n} A_k = E\}\,.$$

If $|F(\Omega)| < +\infty$, we say that F is a vector measure of *bounded variation*.

(a) Show that the variation of a vector measure is a (nonnegative) measure.

(b) Examine vector measures of the previous examples and decide whether or not they are of bounded variation.

42. The Bochner Integral

42.1. Bochner Integral. We say that a vector function $f : \Omega \to X$ is *Bochner integrable* if it is measurable and there exists a sequence of simple functions $\{f_n\}$ such that $\int_\Omega \|f - f_n\|\, \mathrm{d}\mu \to 0$.

A few remarks should be now added. Recall (cf. the proof of Pettis' theorem, or Exercise 42.5) that (real) functions $\|f - f_n\|$ are measurable. Further, the above limit $\lim \int_B f_n \, \mathrm{d}\mu$ exists for any $B \in \mathscr{S}$ since X is complete and the estimates

$$\left\| \int_B f_n - \int_B f_k \right\| \le \int_B \|f_n - f_k\| \le \int_\Omega \|f_n - f\| + \|f_k - f\|$$

hold. Note also that $\int_B \varphi \, \mathrm{d}\mu := \sum x_i \mu(E_i \cap B)$ when $\varphi = \sum x_i c_{E_i}$ is a simple function and that this definition does not depend on the expression of φ as $\sum x_i c_{E_i}$ which is not unique.

Moreover, if $\int_B \|f - f_n\| \to 0$ and $\int_B \|f - g_n\| \to 0$, $B \in \mathscr{S}$ (where f_n, g_n are simple functions) then $\lim \int_B f_n = \lim \int_B g_n$ (consider the sequence $f_1, g_1, f_2, g_2, \dots$).

Now if f is Bochner integrable, $B \in \mathscr{S}$ and $\{f_n\}$ is a sequence of simple functions satisfying $\int_B \|f - f_n\| \to 0$, the limit $\lim \int_B h_n \, \mathrm{d}\mu$ exists and is independent of the sequence $\{h_n\}$. This limit (which is an element of our Banach space X) is called the *Bochner integral* of f and it is denoted by $\int_B f \, \mathrm{d}\mu$. The fundamental characterization of Bochner integrable functions is given in the next theorem.

42.2. Theorem (Bochner). *A measurable function $f : \Omega \to X$ is Bochner integrable if and only if $\int_\Omega \|f\| \, \mathrm{d}\mu < \infty$ (i.e. exactly when the (real) function $\|f\|$ is Lebesgue integrable).*

Proof. If f is Bochner integrable, then $\|f\|$ is measurable (cf. Exercise 42.5). Now the assertion follows from the estimates

$$\int \|f\| \le \int \|f - f_n\| + \int \|f_n\|$$

where $\{f_n\}$ is a sequence of simple functions satisfying $\int \|f - f_n\| \to 0$.

For the converse, suppose $\|f\|$ is integrable. There are simple functions f_n tending to f μ- almost everywhere. Set

$$g_n(\omega) = \begin{cases} f_n(\omega) & \text{if} \quad \|f_n(\omega)\| \le 2\,\|f(\omega)\|, \\ 0 & \text{otherwise in} \quad \Omega. \end{cases}$$

Obviously, g_n are again simple functions. Moreover,

$$\|g_n(\omega)\| \le 2\,\|f(\omega)\|, \quad \|f(\omega) - g_n(\omega)\| \to 0$$

for μ-almost all $\omega \in \Omega$. Since

$$\|f(\omega) - g_n(\omega)\| \le \|f(\omega)\| + \|g_n(\omega)\| \le 3\,\|f(\omega)\|,$$

an appeal to the Lebesgue dominated convergence theorem 8.13 shows $\int \|f - g_n\| \to 0$. ■

Denoting $\mathscr{L}^1_X$ (or, more precisely, $\mathscr{L}^1_X(\Omega, \mathscr{S}, \mu)$) the space of all Bochner integrable functions we see that $f \in \mathscr{L}^1_X$ if and only if $\|f\| \in \mathscr{L}^1$. Basic properties of the Bochner integral are summarized in the next theorem.

42.3. Theorem. *(a) If $f \in \mathscr{L}_X^1$ and $E \in \mathscr{S}$, then $\left\|\int_E f\right\| \le \int_E \|f\|$.*

(b) The space $\mathscr{L}_X^1$ equipped with the norm $\|g\|_1 = \int_\Omega \|g\| \,\mathrm{d}\mu$ is complete. In other words, identifying functions which are equal μ-almost everywhere, $\mathscr{L}_X^1$ is a Banach space.

(c) If $F(E) := \int_E f \,\mathrm{d}\mu$ denotes the indefinite Bochner integral, then F is a σ-additive vector measure absolutely continuous with respect to μ. Moreover, if $\{E_n\}$ is a sequence of pairwise disjoint sets from $\mathscr{S}$, then the series $\sum_n F(E_n)$ converges absolutely.

Sketch of the proof. (a) Using the triangle inequality, the assertion is true for simple functions. For the general case, pass to the appropriate limit appealing to the Lebesgue dominated convergence theorem (which is valid even in the vector case).

(b) The proof is the same as in the case of real functions.

(c) Without any difficulty you can show that the indefinite Bochner integral is (finitely) additive. If $E_n \in \mathscr{S}$ are pairwise disjoint, then the series $\sum F(E_n)$ is absolutely convergent ($\sum \|F(E_n)\| \le \sum \int_{E_n} \|f\| = \int_{\bigcup E_n} \|f\| \int_\Omega \|f\| < +\infty$) and

$$\left\|F\Big(\bigcup_{n=1}^{\infty} E_n\Big) - \sum_{n=1}^{k} F(E_n)\right\| = \left\|F\Big(\bigcup_{n=k+1}^{\infty} E_n\Big)\right\|.$$

The assertion now follows since $\lim_k \mu\Big(\bigcup_{n=k+1}^{\infty} E_n\Big) = 0$ and the measure $E \mapsto \int_E \|f\| \,\mathrm{d}\mu$ is absolutely continuous with respect to μ (Exercise 8.22.b). Indeed, given $\varepsilon > 0$ there is $\delta > 0$ such that $\left\|\int_E f\right\| \le \int_E \|f\| < \varepsilon$ whenever $\mu E < \delta$. ■

42.4. Remark. We have just seen that the indefinite Bochner integral $E \mapsto \int_E f \,\mathrm{d}\mu$ is a σ-additive vector measure which is absolutely continuous with respect to μ. Moreover, it is simply checked that this measure is of bounded variation. As in the real case, a question arises whether each σ-additive X-valued vector measure of bounded variation which is absolutely continuous with respect to μ can be expressed as an indefinite Bochner integral of a Bochner integrable function. Thus there is a question whether or not the Radon-Nikodým theorem holds for vector measures. The answer is negative. The vector measure of Example 41.3.1 is absolutely continuous with respect to Lebesgue measure, in case of $p = 1$ it is of bounded variation and still it is not the indefinite Bochner integral of any Bochner integrable function.

We say that a Banach space X has the *Radon-Nikodým property* (shortly, RNP) if the Radon–Nikodým theorem holds for any X-valued vector measure. More precisely, whenever $(\Omega, \mathscr{S}, \mu)$ is a probability measure space and $\nu\colon \mathscr{S} \to X$ is a vector measure of bounded variation which is absolutely continuous with respect to μ, then ν is an indefinite Bochner integral of a Bochner integrable function.

Consequently, the space $L^1[0,1]$ does not have the RNP. Neither do the spaces c_0, l^∞, $\mathscr{C}(K)$ (K infinite compact) have the RNP. On the other hand, any reflexive Banach space has the RNP.

42.5. Exercise. If $f : \Omega \to X$ is measurable, then the real function $\|f\|$ is measurable. (In fact, we proved this assertion in the course of the proof of Pettis' theorem 40.3.) Prove this assertion directly.

Hint. The assertion is obvious for simple functions. Now, if f_n are simple and $f_n \to f$ μ-almost everywhere, it follows that $\|f_n\| \to \|f\|$ μ-almost everywhere.

42.6. Exercise. Let f, g be measurable functions, $\|f\| \le \|g\|$ μ-almost everywhere and let g be Bochner integrable. Show that f is Bochner integrable.

42.7. Exercise. Show that the indefinite Bochner integral is a vector measure of bounded variation.

42.8. Notes. The Bochner integral was studied by S. Bochner [1933] and N. Dunford [1935]. Today, it is used as a tool in function spaces theories, when examining evolution PDE's and differential equations in Banach spaces, or in some problems of the geometry of Banach spaces.

43. The Dunford and Pettis Integrals

Most of this chapter is devoted to weak integrals in Banach spaces. Next Dunford's lemma seems to be of a great importance.

43.1. Dunford's Lemma. *Given a vector function $f : \Omega \to X$ such that $\varphi \circ f \in \mathscr{L}^1(\mu)$ for every $\varphi \in X^*$, then for any set $E \in \mathscr{S}$ there exists an element $L_E \in X^{**}$ such that*

$$L_E(\varphi) = \int_E \varphi \circ f \, \mathrm{d}\mu$$

for each $\varphi \in X^*$.

Proof. Obviously, f is weakly measurable. Fix $E \in \mathscr{S}$ and define $L_E : \varphi \to \int_E \varphi \circ f$ for $\varphi \in X^*$. Without doubt L_E is a linear functional and we have to show that L_E is bounded. To this end, define the mapping $T : \varphi \to \varphi \circ (fc_E)$ $(X^* \to \mathscr{L}^1(\mu))$. We will finish the proof by showing that T has a closed graph. Indeed, then the closed graph theorem implies that T is bounded (the spaces X^* and $L^1(\mu)$ are complete !) and

$$|L_E(\varphi)| = \left| \int_E \varphi \circ f \right| = \left| \int_\Omega \varphi \circ (fc_E) \right| \le \|\varphi \circ (fc_E)\|_1 = \|T\varphi\|_1 \le \|T\| \, \|\varphi\|$$

which shows that $L_E \in X^{**}$.

Now let us see why T is closed. Assume that $\varphi_n \to \varphi$, $T\varphi_n \to g$. By Theorem 12.4 or Exercise 10.9 we can find a subsequence $\{\varphi_{n_k}\}$ such that $T\varphi_{n_k} \to g$ μ-almost everywhere (i.e. $\varphi_{n_k} \circ (fc_E) \to g$ μ-almost everywhere). Since $\varphi_n \circ (fc_E) \to \varphi \circ (fc_E)$ everywhere, it follows that $g = \varphi \circ (fc_E)$ μ-almost everywhere and we see that $T\varphi = g$. ■

43.2. Weak Integrals. A vector function $f \colon \Omega \to X$ is said to be *Dunford integrable* (or, *weakly integrable*) if $\varphi \circ f \in \mathscr{L}^1(\mu)$ for each $\varphi \in X^*$. The element $L_E \in X^{**}$ whose existence is guaranteed by Dunford's lemma is called the *Dunford integral* (according to some authors also the *Gelfand integral*) of f on E. Thus L_E is the Dunford integral of f if

$$L_E(\varphi) = \int_E \varphi \circ f \, \mathrm{d}\mu \quad \text{for each} \quad \varphi \in X^* .$$

If even $L_E \in X$ for each $E \in \mathscr{S}$ (or, more precisely $L_E \in \varepsilon X \subset X^{**}$ where ε denotes the canonical embedding of X into X^{**}), then f is called *Pettis integrable.* Thus $P_E \in X$ is the *Pettis integral* of f over E if

$$\varphi(P_E) = \int_E \varphi \circ f \, \mathrm{d}\mu \quad \text{for each} \quad \varphi \in X^* .$$

Instead of L_E we will use the notation $(D)\int_E f$ (remember that L_E is an element of X^{**}) while P_E (an element of X) will be denoted by $(P)\int_E f$.

Both integrals can be identified provided X is reflexive.

43.3. Example. Let $X = c_0$, $(\Omega, \mathscr{S}, \mu) = ([0,1], \mathfrak{M}, \lambda)$ and $f\colon [0,1] \to c_0$ be defined as

$$f(t) = \{c_{(0,1]}(t),\, 2c_{(0,1/2]}(t),\, 3c_{(0,1/3]}, \dots\}\,.$$

For each $\varphi \in (c_0)^*$ there is a sequence $\{\alpha_n\} \subset l^1$ such that $\varphi(x) = \sum\limits_n \alpha_n x_n$ whenever $x = \{x_n\} \in c_0$. Then one has (the Lebesgue integral in consideration!)

$$\int_0^1 |\varphi \circ f| = \int_0^1 \left|\sum \alpha_n n c_{(0,1/n]}(t)\right| \,\mathrm{d}t \le \sum_n \int_0^1 |\alpha_n|\, n c_{(0,1/n]} = \sum_n |\alpha_n| < +\infty.$$

It follows that f is Dunford integrable. Since

$$\int_0^1 \varphi \circ f = \int_0^1 \sum_n \alpha_n n c_{(0,1/n]} = \sum_n \alpha_n\,,$$

we see that $(D)\int_0^1 f = \{1,1,1,,\dots\}$. Therefore the Dunford integral of f is an element of $l^\infty = (l^1)^* = (c_0)^{**}$ determined as a functional $\varphi \mapsto \sum\limits_n \alpha_n$. Consequently, f is not Pettis integrable.

Show that

$$(D)\int_0^{1/n} f = \int_0^{1/n} \varphi \circ f = \int_0^{1/n} \sum_i \alpha_i i c_{[0,1/i]} = \frac{1}{n}\alpha_1 + \frac{2}{n}\alpha_2 + \dots + \frac{n}{n}\alpha_n + \alpha_{n+1} + \dots\,.$$

It follows that

$$(D)\int_0^{1/n} f = \{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}, 1, 1, 1, \dots\} \in l^\infty$$

and

$$\left\|(D)\int_0^{1/n} f\right\|_\infty = 1.$$

43.4. Remark. Having in mind the last example, notice that the indefinite Dunford integral

$$F : E \mapsto (D)\int_E f\ ,\ E \in \mathfrak{M}$$

is not absolutely continuous with respect to the Lebesgue measure ($\lambda[0,1/n] = 1/n \to 0$ in spite of $\left\|(D)\int_0^{1/n} f\right\| = 1$). Moreover, the vector measure F is not σ-additive.

The next theorem characterizes Pettis integrable functions among those which are Dunford integrable.

43.5. Pettis' Theorem. *For a measurable Dunford integrable function $f\colon \Omega \to X$ the following assertions are equivalent:*

(i) *f is Pettis integrable,*
(ii) *the indefinite Dunford integral of f is a σ-additive vector measure,*
(iii) *the indefinite Dunford integral of f is absolutely continuous with respect to μ.*

The proof of this theorem makes use of deeper theorems of functional analysis and is beyond the scope of this manuscript. ■

Many problems dealing with vector integration are subtle and need a deeper knowledge of Banach spaces theory. This is also the case of the next remarks. The reader is referred, for example, to J. Diestel and J.J. Uhl [*1977] or L. Mišík [*1989].

43.6. Remarks. 1. Any Bochner integrable function is also Pettis integrable and both integrals coincide on sets of $\mathscr{S}$.

2. Moreover, the following theorem holds:

Let $f : \Omega \to X$ be a measurable Pettis integrable function. Then f is Bochner integrable if and only if the indefinite Pettis integral of f is a vector measure of bounded variation.

3. The space c_0 is exceptional: If X does not contain a copy of c_0 (i.e. there is no subspace of X topologically and algebraically isomorphic with c_0), then any measurable Dunford integrable function is Pettis integrable.

4. Let $f : \Omega \to X$ be a measurable function. There are $x_n \in X$ and $E_n \in \mathscr{S}$ such that $f = \sum\limits_{n=1}^{\infty} x_n c_{E_n}$ almost everywhere. Then

f is Pettis integrable if and only if the series $\sum x_n \mu(E_n \cap E)$ is unconditionally convergent for any $E \in \mathscr{S}$,

f is Bochner integrable if and only if the series $\sum x_n \mu(E_n \cap E)$ is absolutely convergent for any $E \in \mathscr{S}$.

5. Let C be a compact subset of a Banach space X. Let $\mu \in \mathscr{M}^1(C)$ be a probability Radon measure on C. (For any $A \subset X$ we denote by $\overline{\text{co}}A$ the smallest closed convex subset of X containing A. It is simply checked that $\overline{\text{co}}A$ equals the closure of the convex hull of A.) There exists a unique $z \in \overline{\text{co}}C$ so that $\varphi(z) = \int\limits_C \varphi \, \mathrm{d}\mu$ for any $\varphi \in X^*$. This z is called the *barycenter* of μ. How is this related to the Pettis integral? The answer is simple. When defining $f(x) = x$ for $x \in C$ then $z = (P) \int\limits_C f \, \mathrm{d}\mu$, or $z = (P) \int\limits_C x \, \mathrm{d}\mu$. Illustrate for the case $X = \mathbf{R}$, $C = [0,1]$ and $\mu = \lambda$!

6. There is a generalization of the previous example giving a criterion on the existence of the Pettis integral.

Let K be a compact metric space, X a Banach space and μ a probability Radon measure on K. If the mapping $f \colon K \to X$ is continuous, then there exists the Pettis integral of f and $(P) \int_K f \, \mathrm{d}\mu \in \overline{\text{co}} f(K)$.

43.7. The Graves Integral. There is a straightforward analogy of Riemann integration theory for functions having their values in a Banach space. Remember that the Riemann integral can be defined using Darboux upper and lower sums (and upper and lower integrals), or for its definition an original Riemann's approach can be used. It is clear that any definition (like Darboux's one) making use of the ordering of the real line (and the notion of the least upper bound) cannot be immediately carried over for a vector case. Of course, in general Banach spaces there is no ordering. Nevertheless, in the sequel we will touch briefly analogues of both Riemann's and Darboux's approaches.

Let X again be a Banach space, f a mapping from $[0,1]$ into X. Given a partition $D := \{0 = x_0 < x_1 < ... < x_n = 1\}$ of $[0,1]$, $I(D) = \{\xi = \{\xi_i\} : \xi_i \in [x_{i-1}, x_i]\}$ set

$$\Xi(f, D, \xi) := \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).$$

The real number $\Xi(f, D, \xi)$ is called the *Riemann sum* of f. Further, define the *norm* of a partition D as $\nu D := \max\{x_i - x_{i-1} : i = 1, ..., n\}$.

(a) We will say that f is *Riemann integrable* provided there exists $z \in X$ with the following property: For any $\varepsilon > 0$ there is $\delta > 0$ such that

$$\|\Xi(f, D, \xi) - z\| < \varepsilon$$

whenever $\nu D < \delta$ and $\xi \in I(D)$.

If f is Riemann integrable, then the element z of the definition is uniquely determined. It will be termed the *Graves* (sometimes also the *Riemann–Graves*) integral of f and denoted by $(RG)\int_0^1 f$. As in the real case, a mapping f is Riemann integrable if and only if there exists $w \in X$ with the property: For any $\varepsilon > 0$ there is a partition D_0 such that

$$\|\Xi(f, D, \xi) - w\| < \varepsilon$$

whenever a partition D is *finer* than D_0 (i.e. $D \subset D_0$) and $\xi \in I(D)$. Of course, w is then the Graves integral of f.

(b) Realize that the Graves integral is a (Moore-Smith) limit of a "generalized" sequence $\{\Xi(f, D, \xi)\}$ when "ordered" either by a norm or an inclusion.

(c) Any Riemann integrable X-valued function f is bounded (there exists $K > 0$ so that $\|f(t)\| < K$ whenever $t \in [0, 1]$).

In what follows, we make use of the vector integration theory for the case of the special measure space $([0, 1], \mathfrak{M}, \lambda)$.

(d) If $f : [0, 1] \to X$ is Riemann integrable and $\varphi \in X^*$, then the (real) function $\varphi \circ f$ is Riemann integrable. In particular, f is weakly measurable and Pettis integrable.

Hint. Almost all assertions are obvious, even that f is Dunford integrable and $(RG)\int_0^1 = (D)\int_0^1$. Since f is bounded it follows that $(D)\int_I f \in X$ for any interval $I \subset [0, 1]$ and one can see that f is Pettis integrable.

(e) Define a (vector) function $f : [0, 1] \to l^\infty[0, 1]$ (=the space of all bounded functions on $[0, 1]$ equipped with the sup -norm) as $f(t) = c_{[0,t]}$. With the aid of the "Bolzano–Cauchy condition" show that f is Riemann integrable. Now (d) implies that f is weakly measurable. (Check the last assertion also directly: If φ is a continuous linear form on $l^\infty[0, 1]$, then it is easy to see that the function $t \mapsto \varphi(f(t))$ is of bounded variation, and consequently it is measurable.) On the other hand, f is not measurable (use Pettis' theorem 40.3 and realize that $\|f(s) - f(t)\| = 1$ for $s \neq t$).

(f) Another example. Let $E \subset [0, 1]$ and define $g_E \colon [0, 1] \to l^\infty[0, 1]$ as follows: Put $g_E(t) = c_{\{t\}}$ if $t \in E$ and $g_E(t)$ equals zero function for $t \notin E$. Show that g_E is Riemann integrable. If E is a nonmeasurable set, then the function $t \mapsto \|g_E(t)\|$ cannot be measurable. Accordingly, in this case g_E is not measurable (cf. Exercise 42.5 or Pettis' theorem 40.3). State conditions on a set E under which g_E will be measurable.

(g) Any measurable Riemann integrable X-valued function f is Bochner integrable (and both integrals equal).

Hint. Recall that a (real) function $\|f\|$ is bounded and measurable. Now use Bochner's characterization 42.2.

(h) Given again a partition $D := \{0 = x_0 < x_1 < ... < x_n = 1\}$ of $[0, 1]$, set

$$\mathscr{D}(f, D) = \sum_{i=1}^{n} \sup\{\|f(s) - f(t)\| : s, t \in [x_{i-1}, x_i]\}(x_i - x_{i-1}).$$

We say that $f : [0, 1] \to X$ is *Darboux* integrable if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\mathscr{D}(f, D) < \varepsilon$ whenever D is a partition and $\nu D < \delta$. Show again that instead of "ordering" given by a norm an equivalent definition can be formulated using an "ordering" determined by an inclusion. Prove also that any Darboux integrable function is Riemann integrable.

(i) A function $f : [0,1] \to X$ is Darboux integrable if and only if f is bounded and continuous in almost all points of $[0,1]$ (cf. 7.9.d).

Hint. Define the *oscillation* of a X-valued function g as

$$t \mapsto \omega_g(t) := \lim_{\delta \to 0_+} \sup\{\|g(u) - g(v)\| : u, v \in (t-\delta, t+\delta)\}.$$

First show that f is continuous at t_0 if and only if $\omega_f(t_0) = 0$. Further prove that the set $\{t \in [0,1] : \omega_f(t) \geq \alpha\}$ is always closed.

Any Darboux integrable function is obviously bounded. To complete the proof check that

$$\lambda\{t \in [0,1] : \omega_f(t) \geq \frac{1}{n}\} = 0$$

for any n.

For the converse, fix $\varepsilon > 0$ and choose an open set $G \subset [0,1]$ such that $\lambda G \sup \|f\| < \varepsilon$ and f is continuous at all points of the compact set $K := [0,1] \backslash G$. For each $t \in K$ find the greatest $\delta_t > 0$ for which $\|f(s) - f(t)\| \leq \frac{1}{2}\varepsilon$ for any $s \in [0,1] \cap (t-\delta_t, t+\delta_t)$. A routine compactness argument establishes the existence of $\delta > 0$ with the property: If $D := \{0 = x_0 < x_1 < ... < x_n = 1\}$ is a partition of $[0,1]$ and $\nu D < \delta$, then each interval $[x_{i-1}, x_i]$ is either contained in G or, it satisfies $\|f(s) - f(t)\| \leq \varepsilon$ for any couple $s, t \in [x_{i-1}, x_i]$. Having such a partition D, we get $\mathscr{D}(f, D) \leq 3\varepsilon$.

(j) It follows from the above considerations that any Darboux integrable function is measurable (use Pettis' theorem 40.3 and the fact that a continuous image of a separable space is separable), and therefore with the aid of Bochner's theorem 42.2 it is also Bochner integrable.

(k) Each bounded function which is continuous at almost all points is Riemann integrable. When X is a general Banach space, the converse may not be the case.

Choose $E = \mathbf{Q}$ in (f). Then $g_{\mathbf{Q}}$ is Riemann integrable while $\|g_{\mathbf{Q}}\|$ is the Dirichlet function of Example 7.5 which is nowhere continuous. Nor is $g_{\mathbf{Q}}$ continuous at any point of $[0,1]$ (remember that the norm is a continuous function!).

(l) We can see that the classes of Darboux and Riemann integrable functions can differ. It seems that there is no known reasonable characterization of Banach spaces where these classes coincide.

43.8. Exercise. Let $X = c_0$, $(\Omega, \mathscr{S}, \mu) = ([0,1], \mathfrak{M}, \lambda)$ and define

$$\begin{aligned} f(t) &= \{c_{(0,1]}(t), 2c_{(0,1/2]}(t), 3c_{(0,1/3]}(t), ...\}, \\ g(t) &= \{c_{(1/2,1]}(t), 2c_{(1/3,1/2]}(t), 3c_{(1/4,1/3]}(t), ...\}, \\ h(t) &= \{\sum_{n=1}^{\infty} n c_{(\frac{1}{n+1}, \frac{1}{n}]}(t), 0, 0, ...\}. \end{aligned}$$

Show that:
(a) g is Pettis integrable ,
(b) f is Dunford but not Pettis integrable,
(c) h is not even Dunford integrable but it is measurable,
(d) $\|f(t)\| = \|g(t)\| = \|h(t)\|$ for any $t \in [0,1]$.

43.9. Exercise. Define the mapping f from the interval $(0,1)$ (Lebesgue measure in consideration) into the Hilbert space l^2 as

$$f : x \mapsto (\frac{1}{x+1}, \frac{1}{x+2}, \frac{1}{x+3}, ...), \quad x \in (0,1).$$

Show that $(P) \int_0^1 f = \{\log(1 + \frac{1}{n})\}_n$.

43.10. Exercise. Consider again the measure space $([0,1], \mathfrak{M}, \lambda)$. Let $e_n := \{0, ..., 0, 1, 0, ...\}$ $e_0 = \{0, 0, ...\}$.

(a) Put $f(r_n) = e_n$ and $f = e_0$ otherwise (here $\{r_n\}$ is a sequence of all rational numbers of $[0,1]$). Show that $f : [0,1] \to c_0$ is measurable and Bochner integrable (f is even Riemann integrable).

(b) Examine the measurability and integrability of the mapping f of (a) replacing the space c_0 by l^p, $1 \le p < \infty$.

(c) Let $\alpha \in \mathbf{R}$ and define

$$h(t) = \{n^\alpha c_{(0,\frac{1}{n}]}(t)\}_n \text{ whenever } t \in [0,1].$$

If X is one of the Banach spaces c_0, l^p, $1 \le p \le \infty$ show that h is measurable. There is no difficulty to prove that h is Dunford integrable if $X = c_0$ and it is Dunford integrable for $X = l^\infty$ if and only if $\alpha \le 1$. Further, h is Pettis integrable if and only if it is Bochner integrable, and this is the case exactly when $\alpha < 1$. For other cases consult I. Chitescu [1990].

43.11. Notes. Fundamental properties of the Pettis integral appeared in B.J. Pettis [1938], also N. Dunford [1936] studied this integral. Another weak integral (for functions having values in duals of Banach spaces) was introduced by I.M. Gelfand in [1936]. Nowadays, the Pettis integral plays an important role in many branches of functional analysis.

L.M. Graves gave the Riemann-type definition for X-valued mappings on $[0,1]$ in [1927].

Historical comments on vector integration can be found in T.H. Hildebrandt [1953], or J. Diestel and J. J. Uhl [*1977].

Appendix on Topology

In this appendix, we mention briefly some topological notions used in the manuscript which may not be familiar to the reader.

Let us remind that by a *topology* we always mean a family τ of subsets of a set X, designated as *open* or τ*-open* sets, which has the following properties:

(a) τ contains $\emptyset$ and X;
(b) $A \cap B \in \tau$ for any $A, B \in \tau$;
(c) $\bigcup_\alpha A_\alpha \in \tau$ if $A_\alpha \in \tau$.

Complements of open sets are called *closed* sets.

A family $\mathscr{B} \subset \tau$ of open sets is a *base for the topology* τ if every open set can be expressed as a union of members of $\mathscr{B}$. Since a topology is often defined with help of a base $\mathscr{B}$ by the formula $\tau = \{\bigcup\{B : B \in \mathscr{X}\} : \mathscr{X} \subset \mathscr{B}\}$, it is useful to know when a family of sets determines a topology (in this way). The answer is as follows:

Proposition. *A collection of sets* $\mathscr{B}$ *is a base for a topology on* X *if and only if* $X = \bigcup \mathscr{B}$, *and if for any* $z \in B_1 \cap B_2$ $(B_1, B_2 \in \mathscr{B})$ *there is* $B \in \mathscr{B}$ *with* $z \in B \subset B_1 \cap B_2$.

An important example is the topology of a metric space formed by the family of all open sets. A base for this topology is, for instance, the set of all open balls. The *discrete* topology on X consists of all subsets of X and the singletons form a base for this topology.

General topological spaces may enjoy a very complicated structure. In whole of this manuscript we assume that all spaces are Hausdorff: (X, τ) is a *Hausdorff space* whenever x, y are distinct points of X, there exist disjoint open sets $G_x, G_y \in \tau$ with $x \in G_x$, $y \in G_y$.

A *neighborhood* of a point $z \in X$ is every set whose interior contains z. The *interior* of a set M is defined as the largest open set contained in M. It is the union of all open sets contained in M.

The *closure* $\overline{A}$ of a set A is the smallest closed set containing A (it exists!). A set E is *dense* in X if $\overline{E} = X$ and *nowhere dense* if the interior of its closure is empty.

A topological space is said to be *separable* if it contains a countable dense subset. A metric space X is separable if and only if it has a *countable base* for its topology, and this is the case exactly when X has the *Lindelöf property*: Any open cover of X contains a countable subcover.

If X, Y are topological spaces, a mapping $f : X \to Y$ is *continuous* if the pre-images of open sets in Y are open in X. In a usual way, we can define the *continuity at a point*. A mapping is continuous if and only if it is continuous at each point.

A function f is continuous on X if and only if the level sets

$$\{x \in X : f(x) > \alpha\} \text{ and } \{x \in X : f(x) < \alpha\}$$

are open for every $\alpha \in \mathbf{R}$. A function f is said to be *lower semicontinuous* if the set $\{x \in X\colon f(x) > \alpha\}$ is open for every $\alpha \in \mathbf{R}$.

A one-to-one mapping $f\colon X \to Y$ is called a *homeomorphism* if it is continuous and the inverse mapping $f^{-1}\colon f(X) \to X$ is also continuous.

A base for the topology of the Cartesian product of topological spaces X_1, X_2 is formed by the collection of all sets of the form $G_1 \times G_2$, where G_j are open in X_j.

A topological space is *normal* if every pair of disjoint closed sets can be separated by (disjoint) open sets. The normal topological spaces are exactly the spaces where *Urysohn's lemma* and *Tietze's extension theorem* hold.

Urysohn's Lemma for Normal Spaces. *If F_1 and F_2 are disjoint closed subsets of a normal topological space X, then there exists a continuous function f on X such that*

$$0 \le f \le 1, \quad f = 0 \text{ on } F_1, \quad f = 1 \text{ on } F_2.$$

Tietze's Extension Theorem. *If f is a continuous function on a closed subset Z of a normal topological space X, then there exists a continuous function F on X such that*

$$f = F \text{ on } Z \text{ and } \sup_X |F| = \sup_Z |f|\,.$$

Any metric space is normal.

An interesting and from the point of view of the measure theory important class of topological spaces is formed by *locally compact spaces*, i.e. spaces in which every point has a compact neighborhood. A set is *compact* if every its open cover (cover by open sets) contains a finite subcover. The locally compact spaces fails to be normal but for them a version of Urysohn's lemma holds.

Urysohn's Lemma for Locally Compact Spaces. *If K is a compact set and U an open subset of a locally compact space X, $K \subset U \subset X$, then there exists a continuous function f and a compact set L with*

$$K \subset L \subset U, \quad 0 \le f \le 1, \quad f = 1 \text{ on } K, \quad f = 0 \text{ on } X \setminus L.$$

Every locally compact space with a countable base is metrizable. A finite Cartesian product of locally compact spaces is a locally compact space.

Let $\mathscr{C}(X)$ be the space of all continuous (real- or complex-valued) functions on a compact set X. If

$$\|f\| := \max\{|f(t)| : t \in X\} \quad \text{for} \quad f \in \mathscr{C}(X)\,,$$

then $\mathscr{C}(X)$ equipped with this norm is a Banach space.

The convergence in the space $\mathscr{C}(X)$ is the uniform convergence. The pointwise convergence of special sequences of continuous functions can guarantee the convergence in $\mathscr{C}(X)$ as following Dini's theorem shows.

Dini's Theorem. *If $\{f_n\}$ is a monotone sequence of continuous functions on a compact space X which converges pointwise to a continuous function, then the convergence of $\{f_n\}$ is uniform on X.*

Let $\mathscr{A} \subset \mathscr{C}(X)$. We say that

- $\mathscr{A}$ is an *algebra* if $fg \in \mathscr{A}$ for $f, g \in \mathscr{A}$;
- $\mathscr{A}$ is a *lattice* if $\max(f, g)$, $\min(f, g) \in \mathscr{A}$ whenever $f, g \in \mathscr{A}$;
- $\mathscr{A}$ *separates points of* X if for any $x, y \in X$, $x \neq y$ there is $\varphi \in \mathscr{A}$ such that $\varphi(x) \neq \varphi(y)$.

The following theorem is useful in many branches of modern analysis.

Stone-Weierstrass Theorem. *Let X be a compact space and $\mathscr{A}$ a linear subspace of $\mathscr{C}(X)$. If $\mathscr{A}$ is an algebra or a lattice, if $\mathscr{A}$ separates points of X and contains the constant functions, then $\mathscr{A}$ is dense in $\mathscr{C}(X)$.*

Let P be a metric compact space. Then there exists a countable base $\{V_n\}$ for the topology on P. Put $f_n(x) = \operatorname{dist}(x, P \setminus V_n)$ and consider the algebra generated by $\{f_n\}$. The Stone-Weierstrass theorem yields the following proposition.

Theorem. *If P is a metric compact space, then the space $\mathscr{C}(P)$ is separable.*

References

Czech Books and Lecture Notes

L. Boček

[Boč] *Tenzorový počet*, SNTL Praha, 1976.

I. Černý and J. Mařík

[ČM I] *Integrální počet I*, lecture notes, SPN Praha 1960.

[ČM II] *Integrální počet II*, lecture notes, SPN Praha 1961.

S. Fučík and J. Milota

[FM] *Matematická analýza II. Diferenciální počet funkcí více proměnných*, lecture notes, SPN Praha 1975.

O. Kowalski

[Kow] *Základy matematické analýzy na varietách*, lecture notes, UK Praha 1975.

J. Král, I. Netuka and J. Veselý

[KNV] *Teorie potenciálu III*, lecture notes, SPN Praha 1976.

L. Krump, V. Souček and J. A. Těšínský

[KST] *Matematická analýza na varietách*, Karolinum, Praha 1998.

J. Kučera and Š. Schwabik

[KS] *Integrální transformace*, lecture notes, SPN Praha 1969.

J. Lukeš

[L-Př] *Příklady z matematické analýzy I. Příklady k teorii Lebesgueova integrálu*, lecture notes, SPN Praha 1968 (1972, 1984).

[L-T] *Teorie míry a integrálu I*, lecture notes, SPN Praha 1972 (1974, 1980).

[Zap] *Zápisky z funkcionální analýzy*, Karolinum, Praha 1998 (2002, 2003).

[Uvod] *Úvod do funkcionální analýzy*, Karolinum, Praha 2005.

J. Lukeš et al.

[Pr] *Problémy z matematické analýzy*, lecture notes, SPN Praha 1972 (1974, 1977, 1982).

J. Lukeš and J. Malý

[LM] *Míra a integrál*, lecture notes, Univerzita Karlova, Praha 1993.

S. Marcus

[Mar] *Matematická analýza čtená podruhé*, Academia Praha 1976.

I. Netuka and J. Veselý

[NV] *Příklady z matematické analýzy. Míra a integrál*, lecture notes, UK Praha 1982.

W. Rudin

[Ru] *Analýza v reálném a komplexním oboru*, Academia Praha 1977.

R. Sikorski

[Sik] *Diferenciální a integrální počet. Funkce více proměnných*, Academia Praha 1973.

J. Stará and O. John

[SJ] *Funkcionální analýza. Nelineární úlohy*, lecture notes, SPN Praha 1986.

Other books

C.D. Aliprantis and O. Burkinshaw

[*1981] *Principles of real analysis*, North-Holland.

E. Asplund and L. Bungart

[*1966] *A first course in integration*, Holt, Reinhart and Winston.

G. Bachman
[*1964] *Elements of abstract harmonic analysis*, Academic Press.

S. Banach
[*1932] *Théorie des opérations linéaires*, Warszava.

R.G. Bartle
[*1966] *The elements of integration*, Wiley.

H. Bauer
[*1990] *Maß- und Integrationstheorie*, Walter de Gruyter.

E. Behrends
[*1987] *Maß- und Integrationstheorie*, Springer-Verlag.

S.K. Berberian
[*1965] *Measure and integration*, Macmillan.

M. Berger and B. Gostiaux
[*1988] *Differential geometry: Manifolds, Curves, and Surfaces*, Springer-Verlag.

K.P.S. Bhaskara Rao and M. Bhaskara Rao
[*1983] *Theory of charges*, Academic Press.

P. Billingsley
[*1968] *Convergence of probability measures*, John Wiley 1968, Russian translation 1977.
[*1979] *Probability and measure*, Wiley.

E. Borel
[*1898] *Leçons sur la théorie des fonctions*, Gauthier-Villars, Paris.

T.A. Botts and E.J. McShane
[*1959] *Real Analysis*, Van Nostrand.

N. Bourbaki
[*1952] *Intégration*, Herman et Cie, Paris 1952, 2nd edition 1965.

A.M. Bruckner
[*1978] *Differentiation of real functions*, Springer-Verlag.

C. Carathéodory
[*1918] *Vorlesungen über reelle Funktionen*, Teubner Leipzig 1918, 2nd edition 1927, 3rd edition Chelsea 1948.

A.L. Cauchy
[*1821] *Cours d'analyse de l'École Royale Polytechnique*, Paris.

G. Choquet
[*1969] *Lectures on analysis, Vol. I*, W.A. Benjamin.

D.L. Cohn
[*1980] *Measure theory*, Birkhäuser.

C. Constantinescu, K. Weber and A. Sontag
[*1985] *Integration theory, Vol. 1: Measure and integral*, John Wiley & Sons.

M. Davis
[*1977] *Applied nonstandard analysis*, Wiley-Interscience.

K. Deimling
[*1985] *Nonlinear functional analysis*, Springer–Verlag.

J. Diestel and J.J. Uhl
[*1977] *Vector measures*, AMS.

J.L. Doob
[*1994] *Measure theory*, Springer-Verlag.

P. Drábek and J. Milota
[*2004] *Lectures on Nonlinear Analysis*, Vydavatelský servis, Plzeň.

R.M. Dudley
[*1989] *Real analysis and probability*, Wadsworth&Brooks/Cole.

G.A. Edgar
[*1990] *Measure, topology, and fractal geometry*, Springer–Verlag.

L.C. Evans and R.F. Gariepy
[*1992] *Measure theory and fine properties of functions*, CRC Press, Boca Raton.

K.J. Falconer
[*1985] *The geometry of fractal sets*, Cambridge University Press, Cambridge.

H. Federer
[*1969] *Geometric measure theory*, Springer–Verlag. (Second edition Springer 1996).

W.H. Fleming
[*1965] *Functions of several variables*, Addison–Wesley.

K. Floret
[*1981] *Maß- und Integrationstheorie*, B.G. Teubner-Verlag, Stuttgart.

G.B. Folland
[*1984] *Real analysis*, John Wiley.
[*1995] *A course in abstract harmonic analysis*, CRC Press.

I. Fonseca and W. Gangbo
[*1995] *Degree theory in analysis and applications.*, The Clarendon Press, Oxford University Press, New York.

J. Foran
[*1991] *Fundamentals of real analysis*, Marcel Dekker.

J. B. F. Fourier
[*1822] *Théorie analytique de la chaleur*, F. Didot, Paris.

D.H. Fremlin
[*1974] *Topological Riesz spaces and measure*, Cambridge University Press.

S. Fučík, J. Nečas, J. Souček and V. Souček
[*1973] *Spectral analysis of nonlinear operators*, Lecture Notes in Math. 346, Springer–Verlag 1973.

R. F. Gariepy and W. P. Ziemer
[*1995] *Modern Real Analysis*, PWS Publishing Company, Boston.

R.A. Gordon
[*1994] *The integrals of Lebesgue, Denjoy, Perron, and Henstock*, American Mathematical Society, Graduate studies in mathematics, vol. 4.

M. de Guzmán
[*1975] *Differentiation of integrals in* $\mathbf{R}^n$, Lecture Notes in Math. 481, Springer–Verlag.

H. Hahn
[*1921] *Theorie der reellen Funktionen, "I. Band"*, Julius Springer, Berlin.

H. Hahn and A. Rosenthal
[*1948] *Set functions*, Univ. New Mexico Press.

P.R. Halmos
[*1950] *Measure theory*, Van Nostrand 1950, 1966, Springer 1974.

T. Hawkins
[*1970] *Lebesgue's theory of integration (Its origins and development)*, Wisconsin Press.

R. Henstock
[*1988] *Lectures on the theory of integration*, World Scientific Publishing, Singapore.
[*1991] *The general theory of integration*, Clarendon Press, Oxford.

E. Hewitt and K.A. Ross
[*1963] *Abstract harmonic analysis I*, Academic Press 1963 (Russian translation 1975).
[*1970] *Abstract harmonic analysis II*, Academic Press 1970 (Russian translation 1975).

E. Hewitt and K. Stromberg
[*1965] *Real and abstract analysis*, Springer Verlag.

K. Hoffman
[*1975] *Analysis in Euclidean spaces*, Prentice-Hall Inc.

J. Horváth
[*1966] *Topological vector spaces and distributions I*, Addison Wesley.

K. Jacobs
[*1978] *Measure and integral*, Academic Press.

A.N. Kolmogorov
[*1933] *Grundbegriffe der Wahrscheinlichkeitsrechnung [Foundations of the theory of probability]*, Springer-Verlag 1933, Chelsea 1950.

J. Kurzweil
[*1980] *Nichtabsolut konvergente Integrale*, B.S.B.G. Teubner, Leipzig.

S. Lang
[*1993] *Real and functional analysis*, Springer-Verlag (3rd edition).

H.L. Lebesgue
[*1904] *Leçons sur l'intégration et la recherche des fonctions primitives*, Paris, 2nd edition 1928.

Peng Yee Lee
[*1989] *Lanzhou lectures on Henstock integration*, World Scientific Publishing Co..

S. Łojasiewicz
[*1988] *An introduction to the theory of real functions*, John Wiley & Sons.

J. Lukeš, J. Malý and L. Zajíček
[*1986] *Fine topology methods in real analysis and potential theory*, Lecture Notes in Math. 1189, Springer–Verlag.

J. Lützen
[*1982] *The prehistory of the theory of distributions*, Springer–Verlag.

N.N. Luzin
[*1915] *Integral i trigonometričeskie rjady*, Moskva 1915, 2nd edition 1950.

P. Mattila
[*1986] *Lecture notes on geometric measure theory*, Universidad de Extremadura.
[*1995] *Geometry of sets and measures in Euclidean spaces. Fractals and rectifiability*, Cambridge University Press, Cambridge.

J. Mawhin
[*1992] *Analyse (Fondaments, techniques, évolution)*, DeBoeck Université, Bruxelles.

J. Mikusiński
[*1978] *The Bochner integral*, Birkhäuser.

H. Minkowski
[*1907] *Diophantische Approximationen*, Teubner, Leipzig.

L. Mišík
[*1989] *Funkcionálna analýza*, Alfa.

F. Moran
[*1988] *Geometric measure theory*, Academic Press.

A. Mukherjea and K. Pothoven
[*1986] *Real and functional analysis, Part A and B*, Plenum Press.

M.E. Munroe
[*1971] *Measure and integration*, Addison-Wesley.

L. Nachbin
[*1965] *The Haar integral*, Van Nostrand.

T. Neubrunn and B. Riečan
[*1981] *Miera a integrál*, Veda.

J.C. Oxtoby
[*1971] *Measure and category*, Springer-Verlag 1971, 1980, Moskva 1974.

K.R. Parthasarathy
[*1978] *Introduction to probability and measure*, Springer-Verlag.

G.K. Pedersen
[*1989] *Analysis now*, Springer-Verlag.

W.F. Pfeffer
[*1977] *Integrals and measures*, Marcel Dekker, New York.

M.M. Rao
[*1987] *Measure theory and integration*, John Wiley & Sons. (Second revised edition Marcel Dekker, Inc., New York, 2004).

B. Riečan and T. Neubrunn
[*1992] *Téoria miery*, Veda.

C.A. Rogers
[*1970] *Hausdorff measures*, Cambridge University Press.

A.C.M. van Rooij and W.H. Schikhof
[*1982] *A second course on real functions*, Cambridge University Press.

W. Rudin
[*1974] *Real and complex analysis*, McGraw-Hill (2nd ed.).

H.L. Royden
[*1968] *Real analysis*, Macmillan.

S. Saks
[*1937] *Theory of the integral*, Stechert 1937.

Š. Schwabik
[*1992] *Generalized ordinary differential equations*, World Scientific, Singapore.

Š. Schwabik and Guoju Ye
[*2005] *Topics in Banach space integration.*, World Scientific Publishing Co. Pte. Ltd., Hackensack.

J.T. Schwartz
[*1969] *Nonlinear functional analysis*, Gordon and Breach, New York.

I. Segal and R.A. Kunze
[*1968] *Integrals and operators*, McGraw-Hill.

L. Simon
[*1983] *Lectures on geometric measure theory*, Proc. of the Centre for mathematical analysis, Australian National University, vol.3.

K.T. Smith
[*1983] *Primer on modern analysis*, Springer-Verlag.

K. Stromberg
[*1984] *An introduction to classical real analysis*, Wadsworth International.

M. Švec, T. Šalát and T. Neubrunn
[*1987] *Matematická analýza funkcí reálnej premennej*, Alfa.

A.E. Taylor
[*1965] *General theory of functions and integration*, Blaisdell.

F. Topsøe
[*1970] *Topology and measure*, Lecture Notes in Math. 133, Springer-Verlag.

A. Torchinsky
[*1988] *Real variables*, Addison-Wesley.

G. Vitali
[*1905] *Sul problema della misura dei gruppi di punta di una retta*, Bologna.

S. Wagon
[*1985] *The Banach-Tarski paradox*, Cambridge University Press.

A. Weil
[*1940] *L'Intégration dans les Groupes Topologiques et ses Applications*, Hermann et Cie, Paris 1940, 2nd edition 1965.

A.J. Weir
[*1973] *Lebesgue integration and measure*, Cambridge University Press.

R.L. Wheeden and A. Zygmund
[*1977] *Measure and integral*, Marcel Dekker.

H. Widom
[*1969] *Lectures on measure and integration*, Van Nostrand.

J.H. Williamson
[*1962] *Lebesgue integration*, Holt,Rinehart and Winston.

A.C. Zaanen
[*1967] *Integration*, North Holland.

Papers

S. Banach
[1922] *Sur les opérations dans les ensembles abstraits et leurs applications aux équations intégrales*, Fund. Math. **3**, 133-181.
[1923] *Sur le problème de la mesure*, Fund. Math. **4**, 7-33.
[1924] *Sur un theorème de M. Vitali*, Fund. Math. **5**, 130-136.
[1925] *Sur les lignes rectifiables et les surfaces dont l'aire est finie*, Fund. Math. **7**, 225-237.

S. Banach and A. Tarski

[1924] *Sur la décomposition des ensembles de points en parties respectivement congruentes*, Fund. Math. **6**, 244-277.

H. Bauer

[1957] *Sur l'équivalence des théories de l'intégration selon N. Bourbaki et selon M.H. Stone*, Bull. Soc. math. France **85**, 51-75.

A.S. Bezikovitch

[1945] *A general form of the covering principle and relative differentiation of additive functions*, Proc. Cambridge Philos. Soc. **41**, 103-110.

[1946] *A general form of the covering principle and relative differentiation of additive functions*, Proc. Cambridge Philos. Soc. **42**, 1-10.

S. Bochner

[1933] *Integration von Funktionen deren Werte die Elemente eines Vectorraumes sind*, Fund. Math. **20**, 262-276.

E. Borel

[1895] *Sur quelques points de la théorie des fonctions*, Ann. Ecole Normale Sup. **12**, 9-55.

G.E. Bredon

[1963] *A new treatment of the Haar integral*, Michigan Math. J. **10**, 365-373.

C. Carathéodory

[1914] *Über das lineare Mass von Punktmengeneine Verallgemeinerung des Längenbegriffs*, Nach. Ges. Wiss. Göttingen, 404-426.

H. Cartan

[1940] *Sur la mesure de Haar*, C. R. Acad. Sci. Paris **211**, 759-762.

I. Chitescu

[1990] *A parametrical example of Dunford, Pettis and Bochner integration*, Stud. Cerc. Mat. **42**, 405-418.

G. Choquet

[1986] *La naissance de la théorie des capacités: réflexion sur une expérience personelle*, La Vie des Sciences, Comptes rendus, sér. générale **3,4**, 385-397.

[1989] *Vznik teorie kapacit: zamyšlení nad vlastní zkušeností*, Pokroky matematiky, fyziky a astronomie **34**, 71-83.

P.J. Daniell

[1918] *A general form of integral*, Ann. of Math. **19**, 279-284.

R. Doss

[1980] *The Hahn decomposition theorem*, Proc. Amer. Math. Soc. **80**, 377.

N. Dunford

[1935] *Integration in general analysis*, Trans. Amer. Math. Soc. **37**, 441-453.

D. Egorov

[1911] *Sur les suites de fonctions mesurables*, C. R. Acad. Sci. Paris **152**, 244-246.

[1936] *Integration and linear operation*, Trans. Amer. Math. Soc. **40**, 474-494.

G. Faber

[1910] *Über stetige Funktionen. II*, Math. Annalen **69**, 372-433.

P.J.L. Fatou

[1906] *Séries trigonométriques et séries de Taylor*, Acta Math. **30**, 335-400.

M.B. Feldman

[1981] *A proof of Lusin's theorem*, Amer. Math. Monthly **88**, 191-192.

E. Fischer
[1907] *Sur la convergence en moyenne*, Comptes Rendus Acad. Sci. Paris **144**, 1022-1024.

M. Foreman and F. Wehrung
[1991] *The Hahn-Banach theorem implies the existence of a non Lebesgue-measurable set*, Fund. Math. **138**, 13-19.

M. Fréchet
[1915] *Sur l'intégrale d'une fonctionnelle étendue à un ensemble abstrait*, Bull. Soc. Math. France **43**, 248-265.
[1924] *Des familles et fonctions additives d'ensembles abstraits*, Fund. Math. **5**, 206-251.

G. Fubini
[1907] *Sugli integrali multipli*, Rendiconti Accad. Nazionale dei Lincei (Roma) **16**, 608-614.
[1915] *Sulla derivazione per serie*, Rendiconti Accad. Nazionale dei Lincei (Roma) **24**, 204-206.

I.M. Gel'fand
[1936] *Sur un lemme de la théorie des espaces linéaires*, Comm. Ins. Sci. Math. Méc. Univ. de Kharkov et Soc. Mat. Kharkov **13**, 35-40.

H.H. Goldstine
[1941] *Linear functionals and integrals in abstract spaces*, Bull. Amer. Math. Soc. **47**, 615-620.

L.M. Graves
[1927] *Riemann integration and Taylor's theorem in general analysis*, Trans. Amer. Math. Soc. **29**, 163-177.

A. Haar
[1933] *Der Maßbegriff in der Theorie der kontinuierlichen Gruppen*, Ann. of Math. **34**, 147-169.

H. Hahn
[1933] *Über die multiplikation total-additiver Mengenfunktionen*, Annali Scuola Norm. Sup. Pisa **2**, 429-452.

O. Hájek
[1957] *Note sur la mesurabilité B de la dérivée supérieure*, Fund. Math. **44**, 238-240.

D.G. Hartig
[1983] *The Riesz representation theorem revisited*, Amer. Math. Monthly **90**, 277-280.

F. Hausdorff
[1919] *Dimension und äusseres Mass*, Math. Ann. **79**, 157-179.

R. Henstock
[1961] *Definitions of Riemann type of the variational integrals*, Proc. London Math. Soc. **13,3**, 305-321.
[1988] *A short history of integration theory*, SEA Bull. Math. **12**, 75-95.

T.H. Hildebrandt
[1953] *Integration in abstract spaces*, Bull. Amer. Math. Soc. **59**, 111-139.

O. Hölder
[1889] *Über einen Mittelwerthssatz*, Nachr. Akad. Wiss. Göttingen Math.-Phys., 38-47.

J. Horváth
[1970] *An introduction to distributions*, Amer. Math. Monthly **77**, 227-240.

C. Jordan
[1881] *Sur la série de Fourier*, Comptes Rendus Acad. Sci. Paris **92**, 228-230.

S. Kakutani
[1941] *Concrete representation of abstract (M)-spaces*, Annals of Math. **42**, 994-1024.
[1948] *A proof of the uniqueness of Haar's measure*, Ann. Math. **49**, 225-226.

S. Kakutani and J.C. Oxtoby
[1950] *Construction of a non-separable invariant extension of the Lebesgue measure space*, Annals of Math. **52**, 580-590.

M.D. Kirszbraun
[1934] *Über die zussammenziehenden und Lipschitzchen Transformationen*, Fund. Math. **22**, 77-108.

A. Kolmogorov
[1932] *Beiträge zur Masstheorie*, Math. Ann. **107**, 351-366.

J. Král
[1985] *Note on generalized multiple Perron integral*, Čas. Pěst. Mat. **110**, 371-374.

J. Kurzweil
[1957] *Generalized ordinary differential equations and continuous dependence on a parameter*, Czechoslovak Math. J. **82**, 418-446.

H. Lebesgue
[1901] *Sur une généralisation de l'intégrale définie*, Comptes Rendus Acad. Sci. Paris **132**, 1025-1028.
[1902] *Integrale, longuer, aire*, Annali Mat. Pura Appl. **7**, 231-359.
[1903] *Sur une propriétè des fonctions*, Comptes Rendus Acad. Sci. Paris **137**, 1228-1230.
[1910] *Sur l'intégration des fonctions discontinues*, Ann. Sci. Ecole Norm. Sup. **27**, 361-450.

B. Levi
[1906] *Sopra l'integrazione delle serie*, Rend. Instituto Lombardo di Sci. e Lett. **39**, 775-780.

A. Ljapunov
[1940] *Sur les fonctions–vecteurs complètement additives*, Bull. Acad. Sci. URSS **6**, 465-478.

N.N. Luzin
[1912] *Sur les propriétés des fonctions mesurables*, C. R. Acad. Sci. Paris **154**, 1688-1690.

J. Mařík
[1952] *Základy theorie integrálu v Euklidových prostorech*, Časopis Pěst. Mat. **77**, 1-51, 125-145, 267-301.

E.J. McShane
[1934] *Extension of range of functions*, Bull. Amer. Math. Soc. **40**, 837-842.

F.A. Medvedev
[1975] *The work of Henri Lebesgue on the theory of functions (on the occasion of his centenary) Transl. from Uspechi Mat. Nauk, 30(1975), 227-238*, Russian Math.Surveys **30**, 179-191.

F. Mignot
[1976] *Côntrole dans les inéquations variationelles elliptiques*, J. Functional Analysis **22**, 130-185.

G.J. Minty
[1970] *On the extension of Lipschitz, Lipschitz–Hölder continuous and monotone functions*, Bull. Amer. Math. Soc. **76**, 334-339.

A.P. Morse
[1939] *The behaviour of a function on its critical set*, Ann. of Math. **40**, 62–70.
[1947] *Perfect blankets*, Trans. Amer. Math. Soc. **6**, 418-442.

A. Nekvinda and L. Zajíček

[1988] *A simple proof of the Rademacher theorem*, Časopis Pěst. Mat. **113**, 337-341.

J. von Neumann

[1934] *Zum Haarschen Maß in topologischen Gruppen*, Compositio Math. **1**, 106-114.

[1936] *The uniqueness of Haar's measure*, Matem. sbornik **43**, 721-734.

O. Nikodym

[1930] *Sur une généralisation des intégrales de M.J. Radon*, Fund. Math. **15**, 131-179.

O. Perron

[1914] *Über den Integralbegriff*, Sitzungsber. Heidelberg Akad. Wiss. **A16**, 1-16.

F. Peter and H. Weyl

[1927] *Die Vollständigkeit der primitiven Darstellungen einer geschlossenen kontinuierlichen Gruppe*, Math. Ann. **97**, 737-755.

B.J. Pettis

[1938] *On integration in vector spaces*, Trans. Amer. Math. Soc. **44**, 277-304.

M. Plancherel

[1910] *Contributions à l'étude de la représentation d'une fonction arbitrire par des intǵrales définies*, Rend. Circ. mat. Palermo **30**, 289-335.

H. Rademacher

[1919] *Über partielle und totale Differenzierbarkeit*, Math. Ann. **89**, 340-359.

J. Radon

[1913] *Theorie und Anwendungen der absolut additiv Mengenfunktionen*, S.-B. Math. Natur. Kl. Kais. Akad. Wiss. Wien **122.IIa**, 1295-1438.

F. Riesz

[1906] *Sur les ensembles de fonctions*, Comptes Rendus Acad. Sci. Paris **143**, 738-741.

[1909a] *Sur les suites de fonctions mesurables*, Comptes Rendus Acad. Sci. Paris **148**, 1303-1305.

[1909b] *Sur les opérations fonctionnelles linéaires*, Comptes Rendus Acad. Sci. Paris **149**, 974-977.

[1910] *Untersuchungen über Systeme integrirbarer Funktionen*, Math. Annalen **69**, 449-497.

[1912] *Sur quelques points de la théorie des fonctions sommables*, Comptes Rendus Acad. Sci. Paris **154**, 641-643.

[1920] *Sur l'intégrale de Lebesgue*, Acta Math. **42**, 191-205.

[1930-32] *Sur l'existence de la dérivée des fonctions monotones et sur quelques problèmes qui s'y rattachent*, Acta Sci. Math. Szeged **5**, 208-221.

L.J. Rogers

[1888] *An extension of a certain theorem in inequalities*, Messenger of Math. **17**, 145-150.

S. Saks

[1938] *Integration in abstract metric spaces*, Duke Math. J. **4**, 408-411.

A. Sard

[1942] *The measure of the critical set values of differentiable mappings*, Bull. Amer. Math. Soc. **48**, 883–890.

J. Schwartz

[1951] *A note on the space* L_p^*, Proc. Amer. Math. Soc. **2**, 270-275.

L. Schwartz

[1948] *Thorie des distributions et transformation de Fourier (French)*, Ann. Univ. Grenoble. Sect. Sci. Math. Phys. (N.S.) **23**, 7-24.

I.E. Segal
[1954] *Equivalence of measure spaces*, Am. J. Math. **73**, 275-313.

W. Sierpiński
[1928] *Un théorème général sur les familles d'ensembles*, Fund. Math. **12**, 206-210.

L. S. Sobolev
[1936] *Méthode nouvelle à résoudre le problème de Cauchy pour les équations hyperboliques normales*, Mat. Sb. **1 (43)**, 39-71.
L. S. Sobolev
Méthode nouvelle à résoudre le problème de Cauchy pour les équations hyperboliques normales
Mat. Sb. 1 (43) (1936), 39 – 71
R.M. Solovay
[1970] *A model of set theory in which every set of reals is Lebesgue measurable*, Annals of Math. **92**, 1-56.

H. Steinhaus
[1919] *Additive und stetige Funktionaloperationen*, Math. Z. **5**, 186-221.

M.H. Stone
[1948] *Notes on integration I - III*, Proc. Nat. Acad. Sci. **34**, 336-342, 447-455, 483-490.
[1949] *Notes on integration IV*, Proc. Nat. Acad. Sci. **35**, 50-58.

R. Thomas
[1985] *A combinatorial construction of a nonmeasurable set*, Amer. Math. Monthly **92**, 421-422.

L. Tonelli
[1909] *Sull'integrazione per parti*, Rendiconti Accad. Nazionale dei Lincei **18**, 246-253.

E.B. van Vleck
[1908] *On non-measurable sets of points with an example*, Trans. Amer. Math. Soc. **9**, 237-244.

N. Wiener
[1922] *Limit in terms of continuous transformation*, Bull. Soc. Math. France **50**, 119-134.
[1939] *The ergodic theorem*, Duke Math. J. **1-18**.

C.G. Young and W.H. Young
[1911] *On the existence of a differential coefficient*, Proc. London Math. Soc. **9**, 325-335.

W.H. Young
[1904] *On upper and lower integration*, Proc. London Math. Soc. **2**, 52-66.

L. Zajíček
[1983] *On differentiation of metric projections in finite dimensional Banach spaces*, Czech. Math. J. **33,3**, 325-336.

A Short Guide to the Notation

$\mathfrak{M}$, $\mathfrak{M}(\gamma)$, $\mathfrak{M}_n$ ... measurable sets 1.3, 4.4, 26
λ, λ_n ... Lebesgue measure 1.3, 1.15, 26
$\sigma(\mathscr{A})$ ... σ-algebra generated by $\mathscr{A}$ 2.3
$\mathscr{B}$, $\mathscr{B}(P)$ ... Borel sets 2.3
$\mu_{|A}$, μ_A, $\mu_{|\mathscr{T}}$ ... restrictions of measures 2.4
$\mathscr{S}_A$ ... restriction of a σ-algebra 2.4
ε_x ... Dirac measure 2.5
$\overline{\mu}$ ... completion of a measure 2.7
μ^*, μ_* ... outer and inner measure 1.3, 4.8
$\mathbf{M}(\mathscr{S})$ ... space of all (signed, complex) measures on $(X, \mathscr{S})$ 6.17
$\operatorname{vol} I$ ... volume of an interval 1.15
μ^+, μ^-, $|\mu|$ ... variations of a measure 6.6, 6.10
$f(\mu)$ ... image of a measure 8.23
$\mathscr{L}^*$ ... the set of all functions for which the integral exists 8.3
$\mathscr{L}^p$, L^p ... L^p-spaces 8.3, 10.1
$\|f\|_p$ ... L^p-norm of a function f 10.1, 10.5
l^p ... l^p-spaces 10.7, 40.3-4
$\mathscr{S} \otimes \mathscr{T}$ ... product σ-algebra 11.1
M^x, M_x ... sections 11.1
$\mu \otimes \nu$ ... product measure 11.6
w-lim f_j ... weak limit 12.12
w^*-lim f_j ... weak* limit 12.13
X^* ... (topological) dual 12.4
μ_f ... measure having a density f, 8.19, 13.1
$\frac{d\nu}{d\mu}$ ... Radon-Nikodým derivative 13.1
$\nu \ll \mu$ ... absolute continuity of measures 13.1
$\nu \perp \mu$ ... mutually singular measures 13.8
$\operatorname{supt} f$ ... support 14.1
$\mathscr{C}_K(P)$ ... continuous functions having compact support in K 14.1
$\mathscr{C}_c(P)$ ... continuous functions having compact support 14.1
$C_c^{\uparrow}(P)$, $C_c^{\downarrow}(P)$ ... semicontinuous functions 14.4
A^+, A^-, $|A|$ ... variation of a (signed) Radon integral 14.11, 14.12
μ_A, μ_A^* ... measure (outer measure) corresponding to a Radon integral 16.1, 16.4
$\mathscr{C}_b(P)$ ... bounded continuous functions 16.7
$\mathscr{C}_0(P)$ ... continuous functions vanishing at infinity 16.7
$\mu_j \xrightarrow{\text{v}} \mu$ ... vague convergence 17.1
$\mathscr{M}(P)$ ... signed (complex) Radon measures on P 17.1
$\mathscr{M}_b(P)$ ... finite signed (complex) Radon measures 17.1
$\mathbf{e}$ ... unit of a group 19.2
Δ ... modular function 19.9; also a set of all positive functions 25.2
$f * g$ ... convolution 19.19, 26.21
$\operatorname*{V}_a^b f$ ... variation 21.1

D^+f, D^-f, D_+f, D_-f $\overline{D}f$, $\underline{D}f$ ... extreme derivatives 22.3
$\nabla\varphi$ ... Jacobi matrix 26.12
φ' ... (Frchet) derivative 26.12
$\mathscr{C}^\ell$ ... contiuously differentiable functions of order l 26.12
J_φ ... Jacobian 26.12 2
D_μ ... symmetric derivative of a measure 28.2
$\mathscr{D}(\Omega)$ ... (infinitely) smooth functions having compact support 31.1
χ_k ... smoothing convolution kernel 31.1
T_f ... regular distribution 32.1
T_μ ... distribution determined by a measure μ 32.3
δ ... Dirac distribution 32
$f_k \stackrel{\mathscr{D}}{\to} 0$ ... convergence in $\mathscr{D}$ 32.1
$D^\alpha T$ ... multiindex derivative 32.4
$Z_k \to Z$ ... convergence of distributions 32.9
$\hat{f}$, $\mathscr{F}f$ ... Fourier transform 33.1, 33.5, 33.8
$\tilde{f}$ ... inverse Fourier transform 33.5
$\mathscr{S}$ ... Schwartz space 33.5
$\mathbf{M}_{n,k}$ ... space of all matrices 34
L^* ... adjoint mapping 34
A^T ... transpose matrix 34
$\|L\|$ ... norm of a linear mapping 34
$\det A$ ... determinant 34
σ ... k-dimensional measure in $\mathbf{R}^n$ 34.8
$\operatorname{vol} L$, $\operatorname{vol}(u_1,\dots,u_k)$ ... volume 34.10
$N(x,\varphi,E)$ ... Banach indicatrix 34.19
$\deg(y.\varphi,G)$ ... degree of a mapping 35.7
$\mathscr{H}_p(A)$, $\mathscr{H}_p(A,\delta)$ ... Hausdorff measure 36.1
s, S ... one-dimensional ($n-1$-dimensiona) measure 37
$\operatorname{grad} g$, $\operatorname{div} f$, $\operatorname{curl} f$ ... gradient, divergence, curl 37.1
$u_1 \times \cdots \times u_{k-1}$ ... vector product 37.5
$\mathbf{t}$, $\mathbf{n}$ ... unit tangent (normal) vector 37.10, 37.13
$\mathbf{H}^k_-$ ... half-space 37.17
$\mathbf{i}$ ... embbeding of $\mathbf{R}^{n-1}$ into $\partial\mathbf{H}^k_-$ 37.17
$<\cdot,\cdot>$ ... duality 38.1
$\Lambda^k(\mathbf{V})$, $\Lambda_k(\mathbf{V})$ ... k-covectors and vectors 38.1
$V_1 \wedge \cdots \wedge V_k$ ... exterior product 38.3
$I(k,n)$ ... set of all increasing multiindices 38.3
X_i ... coordinate form 38.3
$d\omega$ ... diferential 38.11
$\xi(x)$ ... unit tangent k-vector 38.14
U_μ ... domain of a chart μ 39.1
T_x ... tangent space 39.4
c_0 ... space of all sequences converging to zero 40.3-4

Subject Index

Jaroslav Lukeš, Jan Malý

MEASURE AND INTEGRAL

Published by
MATFYZPRESS
publishing house
of the Faculty of Mathematics and Physics
Charles University in Prague
Sokolovská 83, CZ - 186 75 Praha 8
as the 162. publication

Reviewer: Prof. RNDr. Ivan Netuka, DrSc.

This volume was typeset by the authors using AMS-TEX
the macro system of the American Mathematical Society

Printed by Reproduction center UK MFF
Sokolovská 83, CZ - 186 75 Praha 8

Second edition

Prague 2005

ISBN 80-86732-68-1
ISBN 80-85863-06-5 (First edition)