

Chapter 3

Measure Theory. Measurable Functions And Integration

1. MEASURE THEORY

In constructing an abstract measure in this section we carry out the extension of a countably additive measure from a semiring to a ring and also the Lebesgue extension of a measure. The case of measure on \mathbf{R}^n is studied in detail.

Throughout the following if X is a given set, then $2^X = M(X)$ is the set of all its subsets.

DEFINITION 1. A *ring of subsets* of a set X is a family $K(X) \subset M(X)$ that is closed with respect to the operations of union, intersection, and difference, i.e., the relation $A, B \in K$ implies $A \cup B \in K$, $A \cap B \in K$, and $A \setminus B \in K$.

It is obvious that the symmetric difference also belongs to the ring, i.e., if $A, B \in K$, then $A \Delta B \in K$.

An *algebra* of sets is a ring containing an identity E . The set E is the *identity of the ring* if for any $A \in K(X)$ we have $A \subset E$.

A σ -*ring* is a ring that is closed under countable unions; a σ -*algebra* is a σ -ring with identity. It is easy to verify that a σ -algebra is also closed under countable intersections, i.e., forms a so-called δ -algebra.

In what follows $A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n$ will denote the union of disjoint sets.

DEFINITION 2. A *semiring* of subsets of a set X is a family $P(X) \subset M(X)$ closed with respect to the operation of intersection, containing the empty set, and such that if $A, B \in P$, $A \supset B$, then $A \setminus B = C_1 \sqcup C_2 \sqcup \cdots \sqcup C_n$, where $C_i \in P$.

It follows immediately from the definition of a ring of subsets that the intersection $K = \bigcap_{\alpha} K_{\alpha}$ of any set of rings K_{α} is also a ring.

In general a given system of subsets may be contained in many different rings.

A ring containing a given system of subsets and contained in any ring containing the system, is called *minimal*.

The minimal ring is obviously uniquely determined by the given system of subsets. Indeed, if there were two minimal rings, taking their intersection, we would obtain a ring contained in these minimal rings (which would contradict the minimality of the original rings).

PROPOSITION. *If P is a semiring, then among the rings containing P there is a unique minimal ring K_0 ; it coincides with the system Z of sets $\{A\}$ admitting a finite expansion $A = \bigcup_{i=1}^n C_i$, $C_i \in P$.*

PROOF: The existence and uniqueness of a minimal ring is easy to prove. Indeed if $M(P)$ is the set of all subsets of the set $\bigcup_{A \in P} A$, then $K_0 = \bigcap_{K \in \Sigma} K$ is the unique minimal ring containing P . Here Σ is the set of all rings of sets contained in $K(M)$ (the ring of all subsets $M(P)$) and containing P . We now prove that K_0 coincides with the system Z . We first verify that Z is a ring. If $A, B \in Z$, then $A = \bigcup_i A_i$, $B = \bigcup_j B_j$, $A_i, B_j \in P$. Then $A \cap B = \bigcup_{i,j} (A_i \cap B_j)$, and since $A_i \cap B_j \in P$, it follows that $A \cap B \in Z$. If $A \cap B = \emptyset$, then $A \cup B = A \sqcup B = \bigcup_i A_i \sqcup \bigcup_j B_j \in Z$. Furthermore $A \setminus B = \bigcup_i A_i \setminus \left(\bigcup_j B_j \right) = \bigcup_i \left(\bigcup_j (A_i \setminus B_j) \right)$. By definition of a semiring P we have $A_i \setminus B_j = \bigcup_{i=1}^n C_i$, $C_i \in P$, i.e., $A_i \setminus B_j \in Z$. Therefore by what has been proved $A \setminus B \in Z$. Finally $A \cup B = (A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A) \in Z$. Thus Z is a ring, and obviously minimal, i.e., it coincides with K_0 . The proposition is now proved. ■

We now give some examples of a ring of subsets, a semiring, a σ -ring, an algebra, and a σ -algebra.

EXAMPLES

1. The system of all bounded subsets of the real line is a ring. Indeed, it is easy to see that the union, intersection, and difference of two bounded subsets of the real line are again bounded subsets, i.e., belong to the same collection. This ring is obviously not a σ -ring. The system of all subsets of

the real line is an example of a σ -algebra. The identity of the algebra is the line itself.

2. The system of all finite subsets of an arbitrary set X is a ring of sets. This system is an algebra if and only if the set X is itself finite.

3. For any nonempty set X the system $\{\emptyset, X\}$ consisting of the set X and the empty set \emptyset forms an algebra having the identity $E = X$.

4. Consider the plane \mathbf{R}^2 and the set of all rectangles $\{\Pi\} : a_i \leq x_i \leq b_i, i = 1, 2, a_i \leq b_i$. In the inequalities above some or all of the signs \leq may be replaced by the sign $<$. The system $P = \{\Pi\}$ of all such rectangles forms a semiring.

Indeed the intersection of two rectangles of the indicated form (sides parallel to the OX - and OY -axes in a rectangular coordinate system) is obviously again a rectangle of the same type (the case of the "empty" rectangle $a_i = b_i, i = 1, 2$, is not excluded in our system).

To verify that P is a semiring it suffices to verify that the difference of any two rectangles A and B of the system P is representable in the form of a disjoint union of rectangles $C_i, i = 1, 2, \dots, n$, from P . But this fact is geometrically obvious. Thus the system $P = \{\Pi\}$ is a semiring.

5. A set in the plane is called *elementary* if it can be represented in at least one way as the union of a finite number of pairwise disjoint rectangles. It is a fact that the union, intersection, difference, and symmetric difference of elementary sets are again elementary sets.

In fact if Q and G are two elementary sets, i.e., $Q = \bigcup_k A_k, G = \bigcup_j B_j, A_k$ and B_j rectangles, then $Q \cap G = \bigcup_{k,j} (A_k \cap B_j)$ is also an elementary set (since the intersection of two rectangles is a rectangle).

The difference of two rectangles, as can easily be verified, is an elementary set. Consequently if we subtract an elementary set Q from a rectangle Π , then $\Pi \setminus Q = (\Pi \setminus \bigcup_k A_k) = \bigcap_k (\Pi \setminus A_k)$ is again an elementary set (being the intersection of elementary sets). Now let Q and G be two elementary sets. Obviously there is a rectangle Π containing both of them. Then the set $Q \cup G = \Pi \setminus [(\Pi \setminus Q) \cap (\Pi \setminus G)]$ is elementary. From this, and also from the relations

$$Q \setminus G = Q \cap (\Pi \setminus G), \quad Q \Delta G = (Q \cup G) \setminus (Q \cap G)$$

we find that the difference and symmetric difference of elementary sets are elementary sets, which was to be proved.

It follows from what has been proved above that the collection of all elementary sets forms a ring. This ring is obviously the minimal ring containing the rectangles (the semiring $P = \{\Pi\}$).

We now turn to the study of the basic concept of the present section.

DEFINITION 3. A measure μ on a semiring P is a nonnegative function assuming finite values and having the property of additivity, i.e.

$$\mu(A \sqcup B) = \mu(A) + \mu(B)$$

for any A and B belonging to the semiring P such that $A \sqcup B \in P$.

A countably additive measure on a semiring P is a measure possessing the property of countable additivity, i.e.,

$$\mu\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i), \quad A_i \in P, \quad \bigsqcup_{i=1}^{\infty} A_i \in P.$$

An additive nonnegative set function defined on a system of sets has several simple properties of which we shall make frequent use:

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

In fact $\mu(\emptyset \sqcup \emptyset) = \mu(\emptyset) = \mu(\emptyset) + \mu(\emptyset) = 2\mu(\emptyset)$, i.e. $\mu(\emptyset) = 2\mu(\emptyset)$. Therefore $\mu(\emptyset) = 0$.

b) $\mu\left(\bigsqcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu(A_i)$ for any natural number n .

The proof of this relation is easily obtained by induction.

c) $\mu(A \setminus B) = \mu(A) - \mu(B)$ if $B \subset A$.

Indeed, $A = (A \setminus B) \sqcup B$, and therefore

$$\mu(A) = \mu(A \setminus B) + \mu(B).$$

d) $\mu(A_1 \cup A_2) + \mu(A_1 \cap A_2) = \mu(A_1) + \mu(A_2)$.

We shall show this. We write the relation

$$A_1 \cup A_2 = A_1 \sqcup (A_2 \setminus (A_1 \cap A_2)).$$

Therefore by the additivity of μ and property c), we obtain

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2 \setminus (A_1 \cap A_2)) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2),$$

which was to be proved.

e) $\mu(A_1) \leq \mu(A_2)$ if $A_1 \subset A_2$.

In fact $A_2 = A_1 \sqcup (A_2 \setminus A_1)$. Therefore

$$\mu(A_2) = \mu(A_1) + \mu(A_2 \setminus A_1) \geq \mu(A_1).$$

We shall now prove one of the fundamental theorems in the construction of measure theory.

THEOREM 1. *Every countably additive measure defined on a semiring P has a unique extension to a countably additive measure defined on the minimal ring K_0 containing P .*

PROOF: Every element $A \in K_0$ admits an expansion

$$A = \bigsqcup_{i=1}^n C_i, \quad C_i \in P.$$

By definition we set $\mu(A) = \mu\left(\bigsqcup_{i=1}^n C_i\right) = \sum_{i=1}^n \mu(C_i)$. If A is also represented in the form $A = \bigsqcup_{j=1}^m D_j$, $D_j \in P$, then $A = \bigsqcup_{i,j}^{n,m} C_i \cap D_j$, and

$$\mu(A) = \sum_{i,j}^{n,m} \mu(C_i \cap D_j) = \sum_{j=1}^m \mu(D_j) = \sum_{i=1}^n \mu(C_i), \quad C_i \cap D_j \in P,$$

i.e., the measure of the set $A \in K_0$ is independent of the expansion.* We shall prove that this measure is countably additive. Let A and A_i be sets in the ring K_0 . Let $A = \bigsqcup_{i=1}^{\infty} A_i$. Then

$$A = \bigsqcup_{j=1}^n C_j, \quad A_i = \bigsqcup_{k=1}^{m_i} C_{ik}, \quad C_{ik} \cap C_j = C_{ikj}, \quad C_j, C_{ik} \in P;$$

*The uniqueness of the extension of the measure to the ring K_0 follows from the fact that if $\tilde{\mu}$ is another extension, then

$$\tilde{\mu}(A) = \sum_{i=1}^n \tilde{\mu}(C_i) = \sum_{i=1}^n \mu(C_i) = \mu(A).$$

the sets C_{ikj} are pairwise disjoint, and

$$C_j = \bigcup_{i=1}^{\infty} \bigcup_{k=1}^{m_i} C_{ikj}, \quad C_{ik} = \bigcup_{j=1}^n C_{ikj}.$$

By the countable additivity of the measure on P we have

$$\mu(C_j) = \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \mu(C_{ikj}), \quad \mu(C_{ik}) = \sum_{j=1}^n \mu(C_{ikj}).$$

By definition of the measure on the ring K_0 we obtain

$$\mu(A) = \sum_{j=1}^n \mu(C_j), \quad \mu(A_i) = \sum_{k=1}^{m_i} \mu(C_{ik}).$$

Then, since all the terms are nonnegative, it follows that

$$\mu(A) = \sum_{j=1}^n \mu(C_j) = \sum_{j=1}^n \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \mu(C_{ikj}) = \sum_{i=1}^{\infty} \sum_{k=1}^{m_i} \mu(C_{ik}) = \sum_{i=1}^{\infty} \mu(A_i),$$

which was to be proved. ■

We shall now establish an important property of a countably additive measure.

We shall show that a countably additive measure is continuous, i.e., if $A_1 \supset A_2 \supset \dots$, and $A = \bigcap_n A_n$, then $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$.

Consider, for example, the case $A = \emptyset$. Then

$$A_1 = \bigcup_{i=1}^{\infty} (A_i \setminus A_{i+1}), \dots, A_n = \bigcup_{i=n}^{\infty} (A_i \setminus A_{i+1}).$$

Therefore

$$\mu(A_1) = \sum_{i=1}^{\infty} \mu(A_i \setminus A_{i+1}), \dots, \mu(A_n) = \sum_{i=n}^{\infty} \mu(A_i \setminus A_{i+1}).$$

The series for $\mu(A_1)$ converges, and so its remainder $\mu(A_n) \rightarrow 0$ as $n \rightarrow \infty$. Since $\mu(A) = \mu(\emptyset) = 0$, everything is proved. The general case reduces to this one when A_n is replaced by $A_n \setminus A$.

The continuity property of a measure can also be stated as follows: if $A_1 \subset A_2 \subset \dots$, and $A = \bigcup_n A_n$, then $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$. (The proof reduces to the proof presented above by passing to the complements of the sets.)

The following proposition also holds.

PROPOSITION. *A countably additive measure μ on a semiring P is countably subadditive, i.e., if*

$$A \subset \bigcup A_i,$$

then $\mu(A) \leq \sum_n \mu(A_n)$.

PROOF: Indeed let $B_n = (A_n \cap A) \setminus \bigcup_{k=1}^{n-1} A_k$ and $B_1 = A_1 \cap A$. Then $B_i \cap B_j = \emptyset$, $i \neq j$, $B_n \subset A_n$ and $A = \bigcup_n B_n$. The inequality now follows from this by the countable additivity of the measure:

$$\mu(A) = \sum_n \mu(B_n) \leq \sum_n \mu(A_n). \blacksquare$$

Thus we have carried out the following construction above: a countably additive measure μ was given on a semiring $P(X)$. It turned out that it could be uniquely extended to a countably additive measure on the minimal ring K_0 containing the given semiring. (Here X is some initial set.) Therefore we can assume from the very beginning that the measure is defined on a ring.*

In general the ring $K_0(X)$ does not exhaust all the sets of $M(X)$, the set of all subsets of X . The question arises: what is the maximal class to which the function μ can be extended, and how is this extension to be made? The answer is given using the so-called *Lebesgue extension*. We consider the case when $X \in P$ and $\mu(X) < \infty$.

*It follows from the proof of Theorem 1 that if μ is additive on a semiring P , it is uniquely extendible to an additive measure μ on a ring K_0 containing the semiring P . An extension of the measure μ to a ring that is larger than K_0 is a possibility. The corresponding construction is called a *Jordan extension*. The idea of such a construction consists of approximating the set A for which the Jordan measure is being defined by sets A' and A'' (to which a measure has already been assigned) from within and from without, i.e., so that $A' \subset A \subset A''$.

DEFINITION 4. Let the semiring $P(X) \ni X$ and the countably additive measure μ on $P(X)$ be defined, with $\mu(X) < \infty$. The outer measure μ^* on the set $M(X)$ of all subsets of X is the function

$$\mu^*(A) = \inf_{A \subset \bigcup_{i=1}^{\infty} C_i, C_i \in P} \sum \mu(C_i).$$

Here the infimum is taken over all possible coverings of the set A by systems of sets from the semiring P .

If $A \in K_0$, then $\mu^*(A) = \mu(A)$. Indeed, if $A = \bigcup_{i=1}^n A_i$, $A_i \in P$, then $\mu^*(A) \leq \sum_{i=1}^n \mu(A_i) = \mu(A)$; if $A \subset \bigcup_i C_i$, then $\mu(A) \leq \sum_i \mu(C_i)$, since the measure is countably subadditive. It follows from this that $\mu^*(A) = \mu(A)$, if $A \in K_0$.

We now give the following definition, which we shall use below.

DEFINITION 5. The distance between the sets A and B is the number

$$\rho(A, B) = \mu^*(A \Delta B).$$

Since $A \Delta B \subset (A \Delta C) \cup (B \Delta C)$, it follows that $\rho(A, B) \leq \rho(A, C) + \rho(B, C)$, i.e., the triangle axiom holds.

It is easy to verify that $\rho(A, B) = \rho(B, A)$ and $\rho(A, A) = 0$. Nevertheless the equality $\rho(A, B) = 0$ does not imply $A = B$. However, if we consider the sets A and B equivalent in the case when $\rho(A, B) = 0$, then the given distance turns the set of equivalence classes of sets into a metric space $R = (X, \mu, \rho)$.

We now turn to the definition of the most important concept, that of a Lebesgue-measurable set.

DEFINITION 6. A set $A \in M(X)$ is called *Lebesgue-measurable* if for any $\varepsilon > 0$ there exists a set $B \in K_0(X)$ such that

$$\rho(A, B) < \varepsilon.$$

Here $K_0(X)$ is the minimal ring to which the measure μ defined on the semiring $P(X) \subset M(X)$ is extended, where X is the initial set and $X \in P(X)$.

Thus a set A is measurable if it can be approximated with arbitrary accuracy (in the distance just introduced) by sets of the ring.

The class of Lebesgue-measurable sets will be denoted $\mathcal{L}(X)$. It is clear that $\mathcal{L}(X)$ is a subset of $M(X)$, the set of all subsets of X . The function μ^*

considered on $\mathcal{L}(X)$ is called *Lebesgue measure* and denoted by the symbol μ . Thus a countably additive function μ defined first on the semiring $P(X)$ can be extended to the ring $K_0(X)$, and even to the set $\mathcal{L}(X)$. We shall study the properties of this extension.

The fundamental result is the following theorem.

THEOREM 2. *The class of Lebesgue-measurable sets $\mathcal{L}(X)$ is a σ -algebra, and the function μ is a countably additive measure on $\mathcal{L}(X)$.*

PROOF: We first show that $\mathcal{L}(X)$ is a ring. Let the sets A_i be measurable and $B_i \in K_0(X)$. By direct verification we establish the following inclusions:

$$\begin{aligned} \left(\bigcup_{i=1}^n A_i \right) \Delta \left(\bigcup_{i=1}^n B_i \right) &\subset \bigcup_{i=1}^n (A_i \Delta B_i), \\ \left(\bigcap_{i=1}^n A_i \right) \Delta \left(\bigcap_{i=1}^n B_i \right) &\subset \bigcup_{i=1}^n (A_i \Delta B_i), \\ (A_1 \setminus A_2) \Delta B_1 \setminus B_2 &\subset (A_1 \Delta B_1) \bigcup (A_2 \Delta B_2). \end{aligned}$$

But then by the definition of $\rho(A, B)$ and the subadditivity of μ we have the inequalities

$$\begin{aligned} \rho \left(\bigcup_{i=1}^n A_i, \bigcup_{i=1}^n B_i \right) &\leq \sum_{i=1}^n \rho(A_i, B_i) \\ \rho \left(\bigcap_{i=1}^n A_i, \bigcap_{i=1}^n B_i \right) &\leq \sum_{i=1}^n \rho(A_i, B_i), \\ \rho(A_1 \setminus A_2, B_1 \setminus B_2) &\leq \rho(A_1, B_1) + \rho(A_2, B_2). \end{aligned} \quad (*)$$

The right-hand sides can be made arbitrarily small by virtue of the measurability of the sets occurring in it, and consequently $\mathcal{L}(X)$ is closed under the finite operations of union, intersection, and difference, i.e., the sets $\bigcup_{i=1}^n A_i$,

$\bigcap_{i=1}^n A_i$, and $A_1 \setminus A_2$ are measurable: they can be approximated with arbitrary accuracy by the following elements of the ring: $\bigcup_{i=1}^n B_i$, $\bigcap_{i=1}^n B_i$, $B_1 \setminus B_2$,

$B_i \in K_0(X)$. Thus it has been shown that $\mathcal{L}(X)$ is a ring. Since it contains the identity, the ring $\mathcal{L}(X)$ is an algebra.

We shall prove that $\mathcal{L}(X)$ is a σ -ring. Let $A_i, i = 1, 2, \dots$ be measurable, and choose $B_i \in K_0(X)$ such that $\rho(A_i \setminus B_i) < \varepsilon/2^n$. Then, passing to the limit as $n \rightarrow \infty$ in the inequality (*), we obtain

$$\rho(A, B) \leq \varepsilon,$$

where

$$A = \bigcup_{i=1}^{\infty} A_i, \quad B = \bigcup_{i=1}^{\infty} B_i.$$

We now set $C_i = B_i \setminus \bigcup_{k=1}^{i-1} B_k$. Then $B = \bigcup_{i=1}^{\infty} C_i, C_i \in K_0(X)$. The series $\sum_i \mu(C_i)$ converges (we are considering a semiring with identity and $\mu(X) < \infty$). Consequently there exists a number N such that $\sum_{i=N+1}^{\infty} \mu(C_i) < \varepsilon$.

Then $\rho\left(B, \bigcup_{i=1}^N C_i\right) < \varepsilon$ and

$$\rho\left(A, \bigcup_{i=1}^N C_i\right) \leq \rho\left(B, \bigcup_{i=1}^N C_i\right) + \rho(A, B) < 2\varepsilon.$$

Since $C_i \in K_0(X)$, the set A is measurable and $\mathcal{L}(X)$ is a σ -ring.

We now prove that μ^* is finitely additive on $\mathcal{L}(X)$. The following two inclusions are obvious:

$$A \subset B \cup (A \Delta B), \quad B \subset A \cup (A \Delta B).$$

By the monotonicity of μ^* (Property e) above, with μ^* in place of μ) and the definition of $\rho(A, B)$ we obtain from these inclusions

$$|\mu^*(A) - \mu^*(B)| \leq \rho(A, B).$$

Let A_1 and A_2 belong to $\mathcal{L}(X)$ and $A = A_1 \sqcup A_2$. Choose $\varepsilon > 0$ and sets B_1 and B_2 in $K_0(X)$ such that $\rho(A_i, B_i) < \varepsilon/6, i = 1, 2$. Then

$$\rho(A_1 \sqcup A_2, B_1 \cup B_2) \leq \rho(A_1, B_1) + \rho(A_2, B_2) < \varepsilon/3.$$

Also,

$$|\mu^*(A_1 \sqcup A_2) - \mu^*(B_1 \cup B_2)| \leq \rho(A_1 \sqcup A_2, B_1 \cup B_2) < \varepsilon/3.$$

By the additivity of μ^* on $K_0(X)$ we have (since the measure μ^* coincides with μ on K_0)

$$\mu^*(B_1 \cup B_2) = \mu(B_1 \cup B_2) = \mu(B_1) + \mu(B_2) - \mu(B_1 \cap B_2).$$

Let us compute the measure $\mu(B_1 \cap B_2)$. We have

$$\begin{aligned} \mu(B_1 \cap B_2) &= \rho(B_1 \cap B_2, \emptyset) = \rho(B_1 \cap B_2, A_1 \cap A_2) \\ &\leq \rho(B_1, A_1) + \rho(B_2, A_2) < \varepsilon/3. \end{aligned}$$

Then

$$|\mu^*(A_i) - \mu^*(B_i)| \leq \rho(A_i, B_i) < \varepsilon/6, \quad i = 1, 2.$$

Finally we obtain

$$\begin{aligned} &|\mu^*(A_1 \sqcup A_2) - \mu^*(A_1) - \mu^*(A_2)| \\ &= |\mu^*(A_1 \sqcup A_2) - \mu^*(B_1 \cup B_2) + \mu^*(B_1 \cup B_2) - \mu^*(A_1) - \mu^*(A_2)| \\ &\leq |\mu^*(A_1 \sqcup A_2) - \mu^*(B_1 \cup B_2)| + |\mu^*(B_1) - \mu^*(A_1)| \\ &\quad + |\mu^*(B_2) - \mu^*(A_2)| + \mu^*(B_1 \cap B_2) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Since ε is arbitrary,

$$\mu^*(A_1 \sqcup A_2) = \mu^*(A_1) + \mu^*(A_2),$$

i.e., the measure μ^* is additive on $\mathcal{L}(X)$.

It turns out that the measure μ^* is also countably additive. Indeed let $A = \bigcup_{k=1}^{\infty} A_k$. Then it follows from the countable monotonicity that $\mu^*(A) \leq \sum_{k=1}^{\infty} \mu^*(A_k)$.^{*} On the other hand

$$\mu^*(A) = \mu^*\left(\bigcup_{k=1}^N A_k\right) + \mu^*\left(\bigcup_{k=N+1}^{\infty} A_k\right) \geq \sum_{k=1}^N \mu^*(A_k)$$

^{*}Indeed if $A \subset \bigcup_k A_k$, then for any $\varepsilon > 0$ there exists a system $\{A_{k_j}\}$,

$A_{k_j} \in \mathcal{P}$ such that $A_k \subset \bigcup_j A_{k_j}$ and $\mu^*(A_k) > \sum_j \mu(A_{k_j}) - \frac{\varepsilon}{2^k}$. Then

$$A \subset \bigcup_{k,j} A_{k_j} \quad \text{and} \quad \mu^*(A) \leq \sum_{k,j} \mu^*(A_{k_j}) < \sum_k \mu^*(A_k) + \varepsilon.$$

Since ε is arbitrary, we find that $\mu^*(A) \leq \sum_{k=1}^{\infty} \mu^*(A_k)$.

for any N , i.e., $\mu^*(A) \geq \sum_{k=1}^N \mu^*(A_k)$. Thus $\mu^*(A) = \sum_{k=1}^{\infty} \mu^*(A_k)$. ■

We note that the Lebesgue extension process for a measure is closely connected with the process of completing the metric space $R = (X, \mu, \rho)$, $\rho(A, B) = \mu^*(A \Delta B)$. It turns out that $\mathcal{L}(X)$ coincides with the completion of this metric space in the metric ρ . (Sets A and B for which $\rho(A, B) = 0$ are not distinguished.)

EXAMPLES

1. Let the semiring P consist of the rectangles $\{\Pi\}$ in \mathbf{R}^n : $a_i \leq x_i \leq b_i$, $i = 1, 2, \dots, n$, $a_i \leq b_i$, where some or all of the signs \leq may be replaced by the $<$ sign. Just as in the plane it can be verified that P is a semiring. Then the minimal ring K_0 over this semiring will coincide with the set of all elementary subsets of \mathbf{R}^n . (As in the case of the plane, a set is *elementary* if it is the union of a finite number of rectangles.) We define an additive measure μ on P as follows: the measure of a rectangle Π is its volume $S = \prod_{i=1}^n (b_i - a_i) = \mu(\Pi)$. This measure can be extended to a countably additive measure on the ring K_0 . Then the measure μ^* can be defined as in the abstract case, and the measure μ can be extended to a countably additive measure on the class \mathcal{L} as in Theorem 2. The class \mathcal{L} is called the class of Lebesgue-measurable subsets of \mathbf{R}^n .

The only difference between the construction of Lebesgue measure on \mathbf{R}^n and the construction in the abstract case is that in the case of \mathbf{R}^n we do not have to assume that the measure is countably additive on the semiring P . The countable additivity of the measure on K_0 and on the class \mathcal{L} follows from additivity and the topological properties of sets in \mathbf{R}^n (the Heine-Borel theorem).

In fact, let us return once again to the proof of countable additivity of the measure μ^* in the abstract case. The crucial point in that proof is the assertion of countable monotonicity, i.e., the assertion that if $A \subset \bigcup_{k=1}^{\infty} A_k$, then $\mu^*(A) \leq \sum_{k=1}^{\infty} \mu^*(A_k)$. This inequality in turn, as we have shown in the proof of Theorem 2, follows from the countable monotonicity of the measure μ defined on the semiring.

Consequently to show the countable additivity of the measure μ^* in the case of \mathbf{R}^n we need only show that the measure μ defined on rectangles and extended to elementary sets by the rule

$$\mu(A) = \sum_k \mu(\Pi_k) \quad \text{if} \quad A = \bigsqcup_k \Pi_k$$

is countably subadditive (on the ring K_0). Here A is an elementary set and Π_k are rectangles.

We now establish that the measure μ is countably subadditive for elementary sets in the case of \mathbf{R}^n . Let A be an elementary set and $\{A_k\}$ a finite or countable system of elementary sets such that $A \subseteq \bigcup_k A_k$. We shall then show that $\mu(A) \leq \sum_k \mu(A_k)$. Choose $\varepsilon > 0$ and a closed elementary set $F \subset A$ such that $\mu(F) \geq \mu(A) - \frac{\varepsilon}{2}$. To do this it obviously suffices to replace each of the rectangles Π_i comprising A with a closed rectangle of volume larger than $\mu(\Pi_i) - \frac{\varepsilon}{2^i}$.

Next, for each A_k we choose an open elementary set G_k containing A_k and satisfying the condition

$$\mu(G_k) \leq \mu(A_k) + \frac{\varepsilon}{2^{k+1}}.$$

Obviously $F \subset \bigcup_k G_k$. By the Heine-Borel theorem we choose a finite subsystem G_{k_1}, \dots, G_{k_p} covering F . When this is done, we obviously have

$$\mu(F) \leq \sum_{i=1}^p \mu(G_{k_i}),$$

since otherwise the set F would be covered by a finite number of rectangles with total volume less than $\mu(F)$.

Consequently

$$\begin{aligned} \mu(A) &\leq \mu(F) + \frac{\varepsilon}{2} \leq \sum_{i=1}^p \mu(G_{k_i}) + \frac{\varepsilon}{2} \leq \sum_k \mu(G_k) + \frac{\varepsilon}{2} \\ &\leq \sum_k \mu(A_k) + \sum_k \frac{\varepsilon}{2^{k+1}} + \frac{\varepsilon}{2} = \sum_k \mu(A_k) + \varepsilon, \end{aligned}$$

from which, since ε is arbitrary, it follows that the measure μ is countably subadditive for elementary sets:

$$\mu(A) \leq \sum_k \mu(A_k).$$

Hence, as already stated, just as in the abstract case, we obtain the countable subadditivity and countable additivity of the measure μ^* extended to the class \mathcal{L} in \mathbf{R}^n .

Thus, in contrast to the construction of a measure on an abstract system of sets forming a semiring, in the case of the system of subsets in \mathbf{R}^n the countable additivity of the measure on the semiring need not be assumed—it follows from additivity and the Heine-Borel theorem.

2. The following proposition is frequently used. It turns out that for the covers in the definition of $\mu^*(A)$ in the case of \mathbf{R}^n we can restrict ourselves to open rectangles Q of the semiring P , $Q = \{x : a_i < x < b_i, i = 1, 2, \dots, n\}$.

In fact if

$$\mu^*(A) = \inf_{A \subset \bigcup_k \Pi_k, \Pi_k \in P} \sum_k \mu(\Pi_k), \quad \text{and} \quad \mu_0^*(A) = \inf_{A \subset \bigcup_k Q_k, Q_k \in P} \sum_k \mu(Q_k),$$

where Π_k are arbitrary rectangles and Q_k are open rectangles, then we always have $\mu^*(A) \leq \mu_0^*(A)$. We shall show that $\mu^*(A) = \mu_0^*(A)$. Suppose, to the contrary, that $\mu^*(A) < \mu_0^*(A)$, and $\mu_0^*(A) - \mu^*(A) = \varepsilon$. Choose a cover of the set A such that

$$\sum_k \mu(\Pi_k) - \mu^*(A) \leq \frac{\varepsilon}{2}.$$

We replace each rectangle Π_k with an open rectangle Q_k , $Q_k \supset \Pi_k$, and such that

$$\mu(Q_k) - \mu(\Pi_k) \leq \frac{\varepsilon}{4 \cdot 2^k}.$$

Then

$$\sum_k \mu(Q_k) - \mu^*(A) \leq \frac{3\varepsilon}{4},$$

i.e.,

$$\sum_k \mu(Q_k) - \mu_0^*(A) + \varepsilon \leq \frac{3\varepsilon}{4}, \quad \sum_k \mu(Q_k) \leq \mu_0^*(A) - \frac{\varepsilon}{4},$$

contradicting the definition of $\mu^*(A)$. Thus $\mu_0^*(A) = \mu^*(A)$. Therefore if the set A is measurable, then for any $\varepsilon > 0$ there exists an open set $G \supset A$, $G = \bigcup_k Q_k$, such that $\mu^*(G \setminus A) < \varepsilon$. Passing to complements, we conclude that for any $\varepsilon > 0$ there exists a closed set F such that

$$\mu(A \setminus F) < \varepsilon, \quad F \subset A.$$

3. Open sets are measurable, i.e., belong to the σ -ring $\mathcal{L}(X)$ with identity X (in this case we have called the σ -ring an algebra). Indeed, as we know

from the considerations of Sec. 1.2, a countable basis for the topology of \mathbf{R}^n can be chosen consisting of open rectangles. Therefore, since every open set A is the union of a countable system of open rectangles (and the latter are obviously measurable), an open set is also measurable by Theorem 2. A closed set, being the complement of an open set, is also measurable. A set is called a *Borel set* (cf. Example 4 of Sec. 1.4.2) if it can be obtained using a countable number of operations of union, intersection, and complementation starting from open sets. Every Borel set, for example a unit square, is measurable. The Borel sets form the smallest σ -ring containing the open sets.

4. Countable subsets of \mathbf{R}^n are measurable and have measure zero. Indeed, let $A = \{x_k\}$ be a countable set. Cover the point x_k by an open rectangle Q_k such that $\mu(Q_k) = \varepsilon/2^k$. Then $\mu^*(A) \leq \varepsilon$ for any $\varepsilon > 0$, i.e., $\mu^*(A) = 0$. A set A whose outer measure is zero is always measurable. In fact, we remark that the empty set always belongs to the ring $K_0(X)$ and also that

$$\rho(A, \emptyset) = \mu^*(A \Delta \emptyset) = \mu^*(A) = \mu(A) = 0.$$

5. In constructing the measure on \mathbf{R}^n it is convenient to consider first only subsets of a unit square E . Then the semiring P has E as an identity and $\mu(E) = 1$. This restriction can be removed by covering all of \mathbf{R}^n by a countable system of unit squares E_k and calling a set A measurable if its intersection with each of the unit squares is measurable and $\sum_k \mu(A \cap E_k) < \infty$. When this is done, $\mu(A) = \sum_k \mu(A \cap E_k)$.

6. If the semiring of arbitrary sets on which the initial measure μ is defined has no identity, then all the constructions remain valid, but the measure μ^* is then defined only on the system of sets that have a cover $\bigcup_k C_k$ by sets of P with the sum $\sum_k \mu(C_k)$ finite.

7. The measure μ is called *complete* if the relations $\mu(A) = 0$ and $A_1 \subset A$ imply that $\mu(A_1) = 0$. The Lebesgue extension of a measure is complete. Indeed, if $A_1 \subset A$ and $\mu(A) = 0$, then $\mu^*(A_1) = 0$, and any set B for which $\mu^*(B) = 0$, as shown in Example 4, is measurable.

8. It can be shown that if the initial measure μ is defined not on a semiring but on an arbitrary system of sets, then it may have more than one extension to $\mathcal{L}(X)$.

9. Let X be an arbitrary set. Then (X, K_σ, μ) is called a *measure space* if there exists a σ -ring K_σ of subsets of X and a countably additive measure

μ on K_σ . If $X \in K_\sigma$, then X is called a measurable space. For example let X be a unit square in \mathbf{R}^n and let μ be Lebesgue measure on \mathbf{R}^n . Then (X, K_σ, μ) is a measurable space. Let X be the set of positive numbers, K_σ the set of all subsets of X , and $\mu(A)$ the number of elements in the set A . Then (X, K_σ, μ) is a measurable space. In probability theory an event is a set, and the probability of the event is an additive or countably additive set function (the measure of the set). When it is clear what is meant, we shall often denote a measurable space (or a measure space) simply by the symbol X .

10. Let $X = \{x_1, x_2, \dots\}$ be an arbitrary countable set. To each element we assign a "weight"—a positive number α_n —and require that $\sum_{k=1}^{\infty} \alpha_k = 1$. Let P be the semiring of all subsets of X . For each $A \subset X$ we set $\mu(A) = \sum_{x_k \in A} \alpha_k$. It is obvious that $\mu(X) = 1$ and also easy to verify that μ is a countably additive measure.

11. We give one more example of a measure. Let $\Phi(t)$ be a nondecreasing function on the line that is continuous from the left. We set

$$\begin{aligned}\mu(a, b) &= \Phi(b) - \Phi(a+0), & \mu[a, b] &= \Phi(b+0) - \Phi(a), \\ \mu(a, b] &= \Phi(b+0) - \Phi(a+0), & \mu[a, b) &= \Phi(b) - \Phi(a).\end{aligned}$$

The function μ so defined is nonnegative and additive. Applying the general procedure for extending a measure to it, we construct a class \mathcal{L} of measurable sets on which the measure μ is countably additive. Measures obtained using such functions $\Phi(t)$ are called Lebesgue-Stieltjes measures. In particular if $\Phi(t) = t$, the corresponding measure is ordinary Lebesgue measure on the line.

12. Throughout the considerations above we have been dealing with measurable sets. The class $\mathcal{L}(X)$ of measurable sets contains all open, closed, and Borel sets and is quite extensive.

We shall now construct an example of a nonmeasurable set on a circle C of length 1. Let α be an irrational number. Partition the points of the circle into classes, including in a given class those points of the circle C that can be moved into one another by rotation through an angle $k\alpha\pi$, where k is some integer. Each such class will consist of a countable set of points. We now choose one "representative" of each class. The collection of all such points forms a nonmeasurable set Ψ_0 .

We denote by Ψ_k the set obtained from Ψ_0 by rotating through the angle $\alpha\pi k$.

The sets Ψ_k are pairwise disjoint and their union is the entire circle.

If the set Ψ_0 were measurable, the sets Ψ_k , which are congruent to it, would also be measurable. Since $C = \bigcup_{k=-\infty}^{\infty} \Psi_k$ and $\Psi_k \cap \Psi_l = \emptyset$ if $k \neq l$, we would find that

$$\mu(C) = 1 = \sum_{k=-\infty}^{\infty} \mu(\Psi_k).$$

But all the sets Ψ_k must have the same measure, i.e., $\mu(\Psi_k) = \gamma$. Consequently we have obtained a contradiction; for if $\mu(\Psi_k) = \gamma = 0$, then $\sum_{k=-\infty}^{\infty} \mu(\Psi_k) = 0 \neq 1$, while if $\gamma > 0$, then the series $\sum_{k=-\infty}^{\infty} \mu(\Psi_k)$ diverges. Consequently the set Ψ_0 and each Ψ_k must be nonmeasurable.

13. The condition $X \in K_0(X)$ turns out to be too strong in many cases. Therefore we often consider the weaker condition when X belongs to $K_\sigma^0(X)$, the minimal σ -ring containing the initial semiring $P(X)$. Then the whole set X is a countable union of sets of the semiring: $X = \bigcup_k X_k$, $X_k \in P(X)$.

The measure μ is called σ -finite in this case.

The set A is called Lebesgue-measurable with respect to a σ -finite measure μ if all the sets $A \cap X_k$, $k = 1, 2, \dots$ are measurable.

The measure of the set A is then defined to be the sum of the series $\sum_{i=1}^{\infty} \mu^*(A \cap X_i)$ if it converges and $+\infty$ otherwise.

It is not difficult to verify that the measurable sets form a σ -algebra as before, and the outer measure so defined is countably additive. In this case both sides of the equality

$$\mu\left(\bigcup_k A_k\right) = \sum_k \mu(A_k)$$

may be infinite.

2. MEASURABLE FUNCTIONS

We now turn to the study of an important class of functions called measurable functions. These are the functions for which the theory of integration will be constructed in the next section.

Consider a measurable space X , i.e., an arbitrary set with a distinguished σ -algebra of subsets (whose identity is the space X itself) and a countably additive measure defined on the σ -algebra. (Sets are called measurable if they belong to this σ -algebra.)

We give the following definition.

DEFINITION 7. A real-valued function $f(x)$ defined on a measurable space X is called *measurable* if for any number $a \in \mathbf{R}^1$ the set $\{x : f(x) > a\}$ is measurable.

The following simple proposition holds.

PROPOSITION. A function $f(x)$ is measurable if one of the following sets is measurable for all numbers a :

$$\{x : f(x) \geq a\}, \quad \{x : f(x) < a\}, \quad \{x : f(x) \leq a\}.$$

PROOF: The proof of this proposition follows from the representations

$$\{x : f(x) \geq a\} = \bigcap_{k=1}^{\infty} \{x : f(x) > a - \frac{1}{k}\},$$

$$\{x : f(x) < a\} = X \setminus \{x : f(x) \geq a\},$$

$$\{x : f(x) \leq a\} = \bigcap_{k=1}^{\infty} \{x : f(x) < a + \frac{1}{k}\},$$

$$\{x : f(x) > a\} = X \setminus \{x : f(x) \leq a\},$$

along with the fact that X itself is measurable, as is the intersection of measurable sets and the complement of a measurable set. (Recall that X is a measurable space.) ■

Measurable functions have the following properties.

PROPERTY 1. If the function $f(x)$ is measurable, so is $|f(x)|$.

PROOF: Indeed,

$$\{x : |f(x)| > a\} = \{x : f(x) > a\} \cup \{x : f(x) < -a\}. \blacksquare$$

PROPERTY 2. If the functions $f_n(x)$ are measurable, then the functions

$$f_{\sup}(x) = \sup_n f_n(x), \quad f_{\inf}(x) = \inf_n f_n(x);$$

$$\overline{f}(x) = \overline{\lim}_{n \rightarrow \infty} f_n(x), \quad \underline{f}(x) = \underline{\lim}_{n \rightarrow \infty} f_n(x),$$

along with the functions $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = -\min\{f(x), 0\}$, are all measurable.

PROOF: Indeed,

$$\{x : f_{\sup}(x) > a\} = \bigcup_{n=1}^{\infty} \{x : f_n(x) > a\},$$

$$\overline{f}(x) = \inf_m g_m(x), \quad g_m(x) = \sup_{n \geq m} f_n(x).$$

For the functions $f_{\inf}(x)$ and $f(x)$ the proof is analogous. In particular the limit of measurable functions is measurable. The functions $f^+(x)$ and $f^-(x)$ are obviously measurable. ■

PROPERTY 3. *If $f(x)$ and $g(x)$ are measurable finite-valued (real-valued) functions defined on the set X and the function $F(u, v)$ is real-valued and continuous on \mathbf{R}^2 , then the function $F(f(x), g(x))$ is measurable. In particular the functions $f + g$, $f \cdot g$, and f^m , where m is a natural number, are measurable.*

PROOF: Indeed, let I_n be a rectangle:

$$I_n = \{(u, v) : a_n < u < b_n, c_n < v < d_n\}, \quad a_n < b_n, c_n < d_n.$$

The rectangles I_n form a basis for the topology of \mathbf{R}^2 . Since $F(u, v)$ is continuous on \mathbf{R}^2 , so that the set on which $F(u, v) > a$ for a real is an open subset of \mathbf{R}^2 , we can write

$$\{(u, v) : F(u, v) > a\} = \bigcup_{n=1}^{\infty} I_n.$$

The set $\{x : (f(x), g(x)) \in I_n\} = \{x : a_n < f(x) < b_n\} \cap \{x : c_n < g(x) < d_n\}$ is measurable.

Hence the set

$$\{x : F(f(x), g(x)) > a\} = \bigcup_{n=1}^{\infty} \{x : (f(x), g(x)) \in I_n\}$$

is measurable, which was to be proved. ■

A complex-valued function is called measurable if its real and imaginary parts are measurable.

Several kinds of convergence can be defined for measurable functions:

a) uniform convergence: $f_n \xrightarrow{\text{unif.}} f$ if $\sup_{x \in X} |f_n(x) - f(x)| \rightarrow 0$ as $n \rightarrow \infty$;

b) convergence almost everywhere: $f_n \xrightarrow{\text{a.e.}} f$ if $f_n(x)$ converges to $f(x)$ as $n \rightarrow \infty$ for all x except those x in a subset of measure zero.;

c) convergence in measure: $f_n \xrightarrow{\text{meas.}} f$ if for any $\varepsilon > 0$ the sequence $\mu\{x : |f_n(x) - f(x)| > \varepsilon\}$ tends to zero as $n \rightarrow \infty$.

Uniform convergence obviously implies convergence almost everywhere and, when $\mu(X) < +\infty$, in measure. Convergence almost everywhere on a

measurable space X (with $\mu(X) < \infty$) implies convergence in measure. This follows from the fact that the sets $E_{n,\varepsilon} = \bigcup_{k \geq n} \{x : |f_k(x) - f(x)| > \varepsilon\}$ have the property $E_{1,\varepsilon} \supset E_{2,\varepsilon} \supset \dots$ and $\mu(E_{n,\varepsilon}) \rightarrow 0$ as $n \rightarrow \infty$.

It is not difficult to verify that convergence in measure of a sequence does not in general imply convergence almost everywhere. For example, the sequence $f_1^{(k)}, f_2^{(k)}, \dots, f_k^{(k)}$, $k = 1, 2, \dots$ given by

$$f_i^{(k)}(x) = \begin{cases} 1, & \text{if } \frac{i-1}{k} < x \leq \frac{i}{k}, i = 1, 2, \dots, k, \\ 0, & \text{for all other } x. \end{cases}$$

converges to zero in measure, but does not converge to zero at any point.

However if we alter the initial sequence slightly, the converse assertion is true: if $f_n \rightarrow f$ in measure, there exists a subsequence $\{f_{n(k)}\}$ converging to f almost everywhere. In fact let $P_{n,\varepsilon} = \{x : |f_n(x) - f(x)| > \varepsilon\}$. By hypothesis $\lim_{n \rightarrow \infty} \mu(P_{n,\varepsilon}) = 0$ for $\varepsilon > 0$. Therefore for any natural number k there exists an index $n(k)$ such that $\mu(P_{n(k), 1/k}) < 1/2^k$. We claim that the desired subsequence is $f_{n(k)}$. The set of x where $f_{n(k)}(x)$ does not converge to $f(x)$ is contained in the set $\overline{\lim}_{k \rightarrow \infty} P_{n(k), 1/k}$ and therefore has measure zero (here $\overline{\lim}$ denotes the upper limit of the sequence of sets).

The following two theorems play an important role in the theory of measurable functions and clarify their structure.

THEOREM (Egorov). *Let X be a measurable space ($\mu(X) < \infty$) and $f_n \xrightarrow{\text{a.e.}} f$ on X . Then for any $\delta > 0$ there exists a set $E_\delta \subset X$ such that $\mu(E_\delta) < \delta$ and f_n converges uniformly to f on $X \setminus E_\delta$.*

PROOF: Let $E_{n,\varepsilon} = \bigcup_{k \geq n} \{x : |f_k(x) - f(x)| > \varepsilon\}$. For any natural number p there exists an index $n(p)$ such that $\mu(E_{n(p), 1/p}) < \delta/2^p$. Let $E_\delta = \bigcup_{p=1}^{\infty} E_{n(p), 1/p}$. Then $\mu(E_\delta) < \delta$, and for $k > n(p)$ we have $|f_k(x) - f(x)| \leq 1/p$ if $x \notin E_\delta$, which was to be proved. ■

THEOREM (Luzin). *Let $f(x)$ assume finite values and be measurable on a measurable space X , $\mu(X) < \infty$, where X is a subset of \mathbb{R}^n . Then for any $\varepsilon > 0$ there exists a measurable set $A_\varepsilon \subset X$ such that $\mu(X \setminus A_\varepsilon) < \varepsilon$ and $f(x)$ is continuous on A_ε (i.e., continuous as a function defined only on A_ε).*

PROOF: Since $f = f^+ - f^-$, we may assume $f \geq 0$. We first assume that f is a simple function, i.e., that $f(x) = c_i$ for $x \in X_i$, $\bigsqcup_{i=1}^m X_i = X$,

$X_i \cap X_j = \emptyset$ for $i \neq j$, and X_i are measurable. According to Example 1 for each X_i we can find a closed set $F_i \subset X_i$ such that $\mu(X_i \setminus F_i) < \varepsilon/m$. Let $F = \bigcup_{i=1}^m F_i$, so that $F_i \cap F_j = \emptyset$ if $i \neq j$. Then $\mu(X \setminus F) < \varepsilon$ and $f(x)$ is continuous on F . In fact let $x_0 \in F$. Then $x_0 \in F_j$ for some j . All F_j are closed, and their number is finite, so that x_0 is not a limit point of any F_i with $i \neq j$. The function $f(x)$ is constant on F_j , and therefore continuous on F_j . From what has been said, it is therefore continuous on F also.

Now let $f(x)$ be an arbitrary finite-valued nonnegative and measurable function. We form the sequence of simple functions

$$f_k(x) = \begin{cases} \frac{\nu-1}{k}, & \text{if } \frac{\nu-1}{k} \leq f(x) < \frac{\nu}{k}, 1 \leq \nu \leq k^2, \\ k, & \text{if } f(x) \geq k. \end{cases}$$

We verify easily that f_1, f_2, \dots converges to f on X . By Egorov's theorem there exists a set $E_{\varepsilon/2} \subset X$ such that $\mu(E_{\varepsilon/2}) < \varepsilon/2$ and f_k converges uniformly to f on $X \setminus E_{\varepsilon/2}$. From what has been proved, there is a set $F_\varepsilon^k \subset X$ for each f_k such that $\mu(X \setminus F_\varepsilon^k) < \frac{\varepsilon}{2^{k+1}}$ and $f_k(x)$ is continuous on F_ε^k . On the set

$$A_\varepsilon = X \setminus (E_{\varepsilon/2} \cup \bigcup_{k=1}^{\infty} (X \setminus F_\varepsilon^k))$$

the functions $f_k(x)$ are continuous and the sequence f_1, f_2, \dots converges uniformly to f , i.e., f is continuous on A_ε . Furthermore

$$\mu(X \setminus A_\varepsilon) < \varepsilon,$$

which was to be proved. ■

3. THE LEBESGUE INTEGRAL

In this section we shall construct the Lebesgue integral with respect to an abstract measure.

DEFINITION 8. A real-valued function $s(x)$ defined on a set X is called a *simple function* if its set of values is finite.

Let the set of values of the function s consist of the distinct numbers c_1, c_2, \dots, c_n and let

$$E_i = \{x : s(x) = c_i\}, \quad i = 1, 2, \dots, n.$$

Then it is obvious that the simple function $s(x)$ can be represented as a (finite) linear combination of characteristic functions $\chi_{E_i}(x)$ of the sets E_i , $i = 1, 2, \dots, n$, i.e.,

$$s(x) = \sum_{i=1}^n c_i \chi_{E_i}(x), \quad \chi_{E_i}(x) = \begin{cases} 1, & x \in E_i, \\ 0, & x \notin E_i. \end{cases}$$

The following lemma on approximation by simple functions will be frequently used in what follows.

LEMMA 1. *Let f be a real-valued function defined on a set X . Then there exists a sequence of simple functions $\{s_n(x)\}$ such that $s_n(x) \rightarrow f(x)$ as $n \rightarrow \infty$ for all $x \in X$. If f is measurable, there exists a sequence of measurable simple functions converging pointwise to $f(x)$. If $f \geq 0$ the sequence $\{s_n\}$ can be taken to be monotonically increasing. The sequence $\{s_n\}$ converges to f uniformly if f is bounded.*

PROOF: If $f \geq 0$,* We set

$$E_{n,i} = \left\{ x : \frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n} \right\}, \quad F_n = \{x : f(x) \geq n\}, \\ n = 1, 2, \dots, \quad i = 1, 2, \dots, n \cdot 2^n.$$

We set

$$s_n = \sum_{i=1}^{n \cdot 2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} + n \chi_{F_n}.$$

It is obvious that s_n satisfies all the conditions of the theorem for functions $f \geq 0$. Indeed, fix $x = x_0$ and consider the case $f(x_0) < \infty$:

$$|f(x_0) - s_n(x_0)| \leq \left| f(x_0) - \frac{j-1}{2^n} \right| + n \chi_{F_n}(x_0),$$

if

$$\frac{j-1}{2^n} \leq f(x_0) < \frac{j}{2^n},$$

i.e.,

$$|f(x_0) - s_n(x_0)| < \frac{1}{2^n} + n \chi_{F_n}(x_0) \rightarrow 0, \quad n \rightarrow \infty,$$

*In the case of an arbitrary function $f(x)$ one must use the representation $f(x) = f^+(x) - f^-(x)$.

since if $f(x_0) < \infty$, then for sufficiently large n ($n \geq N$), we have $f(x_0) < n$ and so $x_0 \notin F_n$ and $\chi_{F_n}(x_0) = 0$ for $n \geq N$. It is also obvious that as $n \rightarrow \infty$ we can always arrange to have $\frac{j-1}{2^n} \leq f(x_0) < \frac{j}{2^n}$ for some j with $1 \leq j \leq n \cdot 2^n$. If f is bounded on the set X , then the quantity $n\chi_{F_n}(x)$ will be zero for all $x \in X$ if n is sufficiently large, and so by the estimate carried out above $s_n(x)$ will tend uniformly to $f(x)$ for all $x \in X$. If $f(x_0) = +\infty$, then obviously $s_n(x_0) \rightarrow +\infty$ also, since in this case $\chi_{F_n}(x_0) = 1$. The sequence $s_n(x)$ is monotone increasing by construction.

If the function f is measurable, then all the sets E_{n_i} and F_n are measurable, hence the functions $\chi_{E_{n_i}}$ and χ_{F_n} are measurable, along with linear combinations of them. ■

3.1. Definition of the Lebesgue Integral

We shall now define* integration of functions on a measurable space X with a σ -ring K_σ and a measure μ .

DEFINITION 9. Let $s(x) = \sum_{i=1}^n c_i \chi_{E_i}(x)$, $x \in X$, $c_i > 0$, be a nonnegative measurable simple function. Let the set E belong to $\mathcal{L}(X)$, i.e., E is a measurable set. The *Lebesgue integral over the set E* of such a function is the number

$$I_E(s) = \sum_{i=1}^n c_i \mu(E \cap E_i).$$

If the function f is measurable and $f \geq 0$, then the *Lebesgue integral of the function f* over the measurable set E with respect to the measure μ is defined as the number

$$\int_E f d\mu = \sup_s I_E(s),$$

where the upper bound is taken over all simple functions s such that $0 \leq s \leq f$. The number $\int_E f d\mu$ may be infinite.

If f is a simple function, then $\int_E f d\mu = I_E(f)$. Indeed $\int_E f d\mu \geq I_E(f)$.

We shall show that

$$\int_E f d\mu \leq I_E(f).$$

*This method of constructing the Lebesgue integral is based on the exposition in the book by W. Rudin, *Foundations of Mathematical Analysis*, McGraw-Hill, 1974.

Let s be a simple function different from $f : 0 \leq s \leq f$. Choose some $E_i = \{x : f(x) = c_i\}$ and consider the integrals of f and s over $E \cap E_i$. Then if $E_{ji} = \{x \in E_i : s(x) = k_{ij}\}$, $\bigcup_j E_{ji} = E \cap E_i$, it follows that $I_{E_i}(f) = c_i \mu(E \cap E_i)$ and

$$I_{E_i}(s) = \sum_j k_{ij} \mu((E \cap E_i) \cap E_{ji}).$$

We have

$$I_{E_i}(s) \leq \max_j k_{ij} \sum_{j=1}^m \mu((E \cap E_i) \cap E_{ji}) \leq c_i \sum_{j=1}^m \mu((E \cap E_i) \cap E_{ji}).$$

But

$$(E \cap E_i) \cap E_{ji} \cap (E \cap E_i) \cap E_{ki} = \emptyset \quad \text{for } j \neq k.$$

Therefore

$$\sum_{j=1}^m \mu((E \cap E_i) \cap E_{ji}) = \mu\left(\bigcup_j (E \cap E_i) \cap E_{ji}\right).$$

Further,

$$\bigcup_{j=1}^m (E \cap E_i) \cap E_{ji} = E \cap E_i.$$

Consequently,

$$I_{E_i}(s) \leq c_i \mu(E \cap E_i) = I_{E_i}(f).$$

Thus if $s \leq f$, then

$$I_E(s) \leq I_E(f), \quad \sup_s I_E(s) \leq I_E(f), \quad \int_E f \, d\mu \leq I_E(f).$$

From this the equality $\int_E f \, d\mu = I_E(f)$ follows if f is a simple function.

DEFINITION 10. The *Lebesgue integral* of an arbitrary measurable function f over the measurable set E is the number

$$\int_E f \, d\mu = \int_E f^+ \, d\mu - \int_E f^- \, d\mu,$$

provided at least one of the integrals on the right is finite.

DEFINITION 11. A function f is *Lebesgue-integrable* (or *Lebesgue-summable*) over the measurable set E with respect to the measure μ if both of the integrals $\int_E f^+ d\mu$ and $\int_E f^- d\mu$ are finite. The class of functions that are Lebesgue-integrable over the set E with respect to the measure μ will be denoted $L^1(\mu)$.

It should be emphasized that under these definitions the Lebesgue integral of a function may be defined and equal to $+\infty$ or $-\infty$ while the function is not Lebesgue-integrable. A function is Lebesgue-integrable if and only if its Lebesgue integral is finite.

3.2. Properties of the Lebesgue Integral

Let us prove the following properties of the Lebesgue integral.

PROPERTY 1. If the function f is measurable and bounded ($|f| \leq K$) on a set E and $\mu(E) < +\infty$, then f is Lebesgue-integrable over E , i.e., $f \in L^1(\mu)$.

PROOF: If $f \geq 0$, is a bounded measurable simple function, then

$$\begin{aligned} \int_E f d\mu &= I_E(f) = \sum_{i=1}^n c_i \mu(E \cap E_i) \leq K \cdot \sum_{i=1}^n \mu(E \cap E_i) \\ &= K \cdot \mu\left(\bigcup_{i=1}^n E \cap E_i\right) \leq K \cdot \mu(E) < \infty, \quad (f \leq K). \end{aligned}$$

If f is simple and of variable sign, then the assertion follows from the representation

$$f = f^+ - f^-.$$

Thus the property is established for simple functions. In the general case for $f \geq 0$ it follows from the relation $\int_E f d\mu = \sup_s I_E(s)$, where s are simple functions such that $0 \leq s \leq f$ and therefore all the s are bounded by the single number K , so that $I_E(s)$ are all bounded by the single number $K \cdot \mu(E)$. The representation $f = f^+ - f^-$ finishes the proof. ■

PROPERTY 2. If f is measurable on E , $\mu(E) < +\infty$, and $a \leq f(x) \leq b$, then

$$a \cdot \mu(E) \leq \int_E f d\mu \leq b \cdot \mu(E).$$

PROOF: The assertion follows from Property 1:

$$f(x) \leq b, \quad -f(x) \leq -a. \blacksquare$$

PROPERTY 3. If f and g belong to $L^1(\mu)$ and $f(x) \leq g(x)$ for all $x \in E$, then

$$\int_E f d\mu \leq \int_E g d\mu.$$

PROOF: For simple functions the inequality is verified directly. Let f and g be nonnegative. Then

$$\begin{aligned} \int_E f d\mu &= \sup_s I_E(s), \quad s \leq f, \\ \int_E g d\mu &= \sup_t I_E(t), \quad t \leq g, \end{aligned}$$

where s and t are simple functions. Since $s \leq f \leq g$, it is clear that $\sup_s I_E(s) \leq \sup_t I_E(t)$. In the general case we must make use of the representations

$$f = f^+ - f^-, \quad g = g^+ - g^-, \quad f^+ \leq g^+, \quad g^- \leq f^- . \blacksquare$$

PROPERTY 4. If $f \in L^1(\mu)$, then $cf \in L^1(\mu)$ on E for any finite number c , and

$$\int_E cf d\mu = c \int_E f d\mu.$$

PROOF: This property follows from the analogous property for simple functions upon taking upper bounds. \blacksquare

PROPERTY 5. If $\mu(E) = 0$ and f is measurable, then

$$\int_E f d\mu = 0.$$

PROOF: For simple functions we obtain from the definition of the Lebesgue integral that

$$I_E(s) = \sum_{i=1}^n c_i \mu(E \cap E_i) = 0.$$

Therefore for $f \geq 0$ we have

$$\int_E f d\mu = \sup_s I_E(s) = 0, \quad s \leq f.$$

Hence

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu = 0,$$

which was to be proved. ■

PROPERTY 6. If $f \in L^1(\mu)$ on E and $A \in \mathcal{L}(X)$ (A is measurable), $A \subset E$, then $f \in L^1(\mu)$ on A .

PROOF: If $f \geq 0$ is a simple function on E , then for its restriction \hat{f} to A we have

$$I_A(\hat{f}) = \sum_{i=1}^n c_i \mu(A \cap A_i) \leq \sum_{i=1}^n c_i \mu(E \cap E_i),$$

where

$$A_i = \{x \in A : \hat{f}(x) = c_i\}.$$

Now let $f \geq 0$ be an arbitrary function, $f \in L^1(\mu)$. Then $\sup_{\hat{s} \leq \hat{f}} I_A(\hat{s}) \leq$

$\sup_{s \leq f} I_E(s)$, where \hat{s} and \hat{f} are the restrictions of the functions s and f to $A \subset E$. As above, the representation $f = f^+ - f^-$ finishes the proof. ■

We now prove the basic theorems of Lebesgue integration theory.

THEOREM 3. Let f be measurable and nonnegative on the set X . For $A \in \mathcal{L}(X)$ we define $\varphi(A) = \int_A f d\mu$. Then the function φ is a countably additive function on $\mathcal{L}(X)$. The same is true if $f \in L^1(\mu)$ on X .

PROOF: We must show that $\varphi(A) = \sum_{n=1}^{\infty} \varphi(A_n)$ if $A_n \in \mathcal{L}(X)$ for all n , $A_i \cap A_j = \emptyset$ when $i \neq j$, $A = \bigcup_{n=1}^{\infty} A_n$. If f is a characteristic function the countable additivity of the function φ follows from the countable additivity of the measure μ . Indeed,

$$\int_A \chi_E d\mu = \mu(A \cap E).$$

If f is a simple function, i.e.,

$$f = \sum_{i=1}^m c_i \chi_{E_i},$$

then

$$\varphi(A) = \int_A f d\mu = I_A(f) = \sum_{i=1}^m c_i \mu(A \cap E_i).$$

By the countable additivity of μ we find that

$$\varphi(A) = \sum_{n=1}^{\infty} \varphi(A_n).$$

Furthermore, for each measurable simple function s such that $0 \leq s \leq f$ we have

$$I_A(s) = \int_A s d\mu = \sum_{n=1}^{\infty} \int_{A_n} s d\mu \leq \sum_{n=1}^{\infty} \varphi(A_n),$$

since $s \leq f$.

Therefore

$$\varphi(A) = \sup_{0 \leq s \leq f} I_A(s) \leq \sum_{n=1}^{\infty} \varphi(A_n).$$

We shall prove the converse inequality.

If $\varphi(A_n) = +\infty$ for some n , then, since $\varphi(A) \geq \varphi(A_n)$ for all n , we have $\varphi(A) = \int_A f d\mu \geq \int_{A_n} f d\mu$. Therefore suppose $\varphi(A_n) < +\infty$ for any n . For a given $\varepsilon > 0$ we choose a measurable simple function s such that $0 \leq s \leq f$ and

$$\int_{A_1} s d\mu \geq \int_{A_1} f d\mu - \varepsilon, \quad \int_{A_2} s d\mu \geq \int_{A_2} f d\mu - \varepsilon.$$

Then

$$\varphi(A_1 \sqcup A_2) \geq \int_{A_1 \sqcup A_2} s d\mu = \int_{A_1} s d\mu + \int_{A_2} s d\mu \geq \varphi(A_1) + \varphi(A_2) - 2\varepsilon.$$

Consequently

$$\varphi(A_1 \sqcup A_2) \geq \varphi(A_1) + \varphi(A_2).$$

Therefore for any n

$$\varphi(A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n) \geq \varphi(A_1) + \cdots + \varphi(A_n).$$

Since $A \supset A_1 \sqcup \cdots \sqcup A_n$, it follows that $I_A(s) \geq I_{A_1 \sqcup \cdots \sqcup A_n}(s)$ and

$$\sup_s I_A(x) \geq \sup_s I_{A_1 \sqcup \cdots \sqcup A_n}(s),$$

i.e.,

$$\varphi(A) \geq \sum_{i=1}^n \varphi(A_i).$$

Hence

$$\varphi(A) \geq \sum_{i=1}^{\infty} \varphi(A_i).$$

If f is an arbitrary function, one must consider the representation $f = f^+ - f^-$ and apply the assertion already proved. ■

COROLLARY. If B and A belong to $\mathcal{L}(X)$, $B \subset A$, $\mu(A \setminus B) = 0$, and the function f is measurable, then

$$\int_A f d\mu = \int_B f d\mu.$$

PROOF: Indeed,

$$A = B \sqcup (A \setminus B), \quad \int_A f d\mu = \int_B f d\mu + \int_{A \setminus B} f d\mu = \int_B f d\mu,$$

since $\mu(A \setminus B) = 0$. ■

This corollary shows that sets of measure zero can be neglected in Lebesgue integration.

In this connection we give the following definition.

DEFINITION 12. If a property T holds for all $x \in E \setminus A$, and $\mu(A) = 0$, we shall say that property T holds *almost everywhere* on the measurable set E , or for *almost every* $x \in E$ under the given measure μ .

If $\mu(\{x : f(x) \neq g(x)\} \cap E) = 0$, then f is *equivalent* to g on E , and we write $f \sim g$ on E . Thus $f \sim f$; $f \sim g$ implies $g \sim f$; and if $f \sim g$ and $g \sim h$, then $f \sim h$:

$$\{x : f \neq h\} \subset \{x : f \neq g\} \cup \{x : g \neq h\}.$$

If $f \sim g$ on E , then for any $A \subset E$ we have $\int_A f d\mu = \int_A g d\mu$, if either one of these integrals exists. Indeed, let A' be the set where $f \neq g$. Then

$$\mu(A') = 0, \quad \int_A f d\mu = \int_{A \setminus A'} f d\mu + \int_{A'} f d\mu = \int_{A \setminus A'} g d\mu + \int_{A'} g d\mu = \int_A g d\mu.$$

THEOREM 4. If $f \in L^1(\mu)$, then $|f| \in L^1(\mu)$ also, and

$$\left| \int_E f d\mu \right| \leq \int_E |f| d\mu.$$

PROOF: Let $E = A \sqcup B$, where $f \geq 0$ on A and $f < 0$ on B . Since the Lebesgue integral is an additive set function,

$$\int_E |f| d\mu = \int_A |f| d\mu + \int_B |f| d\mu = \int_A f^+ d\mu + \int_B f^- d\mu < +\infty.$$

Therefore $|f| \in L^1(\mu)$. Further $f \leq |f|$, $-f \leq |f|$, and

$$\int_E f d\mu \leq \int_E |f| d\mu, \quad -\int_E f d\mu \leq \int_E |f| d\mu,$$

and therefore $\left| \int_E f d\mu \right| \leq \int_E |f| d\mu$, which was to be proved. ■

THEOREM 5. Let the function f be measurable on E , $|f| \leq g$, $g \in L^1(\mu)$. Then $f \in L^1(\mu)$.

PROOF: We always have $f^+ \leq g$ and $f^- \leq g$. Hence, if $0 \leq s \leq f^+$, and s is a simple function, then $s \leq g$. Since $\int_E g d\mu = \sup_s \int_E s d\mu < +\infty$, we also have $\int_E f^+ d\mu < +\infty$. The reasoning for f^- is similar. ■

THEOREM (Chebyshev). If the function $f(x) \geq 0$ on A belongs to $\mathcal{L}(X)$, then $\mu\{x : x \in A, f(x) \geq c\} \leq \frac{1}{c} \int_A f d\mu$ for any number $c > 0$.

PROOF: Let $A' = \{x : x \in A, f(x) \geq c\}$. Then

$$\int_A f d\mu = \int_{A'} f d\mu + \int_{A \setminus A'} f d\mu \geq \int_{A'} f d\mu \geq c\mu(A'),$$

which was to be proved. ■

COROLLARY. If $\int_A |f| d\mu = 0$, then $f = 0$ almost everywhere on A .

PROOF: By Chebyshev's theorem

$$\mu\left\{x : x \in A, |f| \geq \frac{1}{n}\right\} \leq n \int_A |f| d\mu = 0 \quad \text{for } n = 1, 2, \dots$$

Therefore

$$\mu\{x : x \in A, f \neq 0\} \leq \sum_{n=1}^{\infty} \mu\left\{x : x \in A, |f| \geq \frac{1}{n}\right\} = 0,$$

since

$$\{x : x \in A, f \neq 0\} = \bigcup_{n=1}^{\infty} \left\{x : x \in A, |f| \geq \frac{1}{n}\right\}. \blacksquare$$

Property 5 of the Lebesgue integral asserted that the integral of any measurable function over a set of measure zero is zero. We also have the following theorem, which asserts the property known as the absolute continuity of the Lebesgue integral.

THEOREM 6. If the function f is integrable over the measurable set E with respect to the measure μ , then for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left| \int_e f d\mu \right| < \varepsilon$$

for any measurable set $e \subset E$ such that $\mu(e) < \delta$.

PROOF: If $|f|$ is bounded by the number $K < +\infty$, then by what has been proved

$$\left| \int_e f d\mu \right| \leq K\mu(e),$$

and the assertion of the theorem is obvious. Let $E_n = \{x : x \in E, n \leq |f(x)| < n+1\}$, $B_N = \bigcup_{n=0}^N E_n$, $C_N = E \setminus B_N$. Then by the countable additivity of the integral

$$\int_E |f| d\mu = \sum_{n=0}^{\infty} \int_{E_n} |f| d\mu.$$

We choose N such that

$$\sum_{n=N+1}^{\infty} \int_{E_n} |f| d\mu = \int_{C_N} |f| d\mu < \frac{\varepsilon}{2}.$$

and let

$$0 < \delta < \frac{\varepsilon}{2(N+1)}.$$

Now if $\mu(e) < \delta$, then

$$\left| \int_e f d\mu \right| \leq \int_e |f| d\mu = \int_{e \cap B_N} |f| d\mu + \int_{e \cap C_N} |f| d\mu \leq (N+1)\mu(e) + \frac{\varepsilon}{2} < \varepsilon,$$

which was to be proved. ■

3.3. Passage to the Limit under the Integral Sign

Theorems on passage to the limit under the integral sign in a Lebesgue integral play an important role in analysis.

THEOREM 7. (Monotone convergence theorem). *Let the set E be measurable and let $\{f_n\}$ be a sequence of measurable functions such that $0 \leq f_1(x) \leq f_2(x) \leq \dots$ for $x \in E$. Let $f(x)$ be defined by the equality*

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad \text{for any } x \in E.$$

Then

$$\int_E f_n d\mu \rightarrow \int_E f d\mu \quad \text{as } n \rightarrow \infty.$$

PROOF: Using Property 3 of the Lebesgue integral, we find that $0 \leq \int_E f_1 d\mu \leq \int_E f_2 d\mu \leq \dots$. Therefore there exists α such that $\int_E f_n d\mu \rightarrow \alpha$

as $n \rightarrow \infty$; and since $\int_E f_n d\mu \leq \int_E f d\mu$, we have

$$\alpha \leq \int_E f d\mu.$$

Let c be a number satisfying $0 < c < 1$ and s a measurable simple function such that $0 \leq s \leq f$. Let $E_n = \{x : f_n(x) \geq cs(x)\}$ for $n = 1, 2, 3, \dots$. By

the property of monotonicity of the sequence $\{f_n\}$ we obtain: $E_1 \subset E_2 \subset E_3 \subset \dots$ and, since $f = \lim_{n \rightarrow \infty} f_n$, we have $E = \bigcup_{n=1}^{\infty} E_n$.

For any n we have $\int_E f_n d\mu \geq \int_{E_n} f_n d\mu \geq c \int_{E_n} s d\mu$. Let $\varphi(E_n) = \int_{E_n} s d\mu$ and $\varphi(E) = \int_E s d\mu$. Then by the continuity of a countably additive set function it follows that $\varphi(E_n) \rightarrow \varphi(E)$ as $n \rightarrow \infty$. Consequently $\alpha \geq c \int_E s d\mu$.

We now let c tend to 1. Then $\alpha \geq \int_E s d\mu$ and $\alpha \geq \sup_{0 \leq s \leq f} \int_E s d\mu = \int_E f d\mu$.

Thus $\alpha = \int_E f d\mu$, i.e., $\int_E f_n d\mu \rightarrow \int_E f d\mu$, which was to be proved. ■

COROLLARY 1. Let $f = f_1 + f_2$, where $f_i \in L^1(\mu)$, $i = 1, 2$. Then $f \in L^1(\mu)$ and

$$\int_E f d\mu = \int_E f_1 d\mu + \int_E f_2 d\mu.$$

PROOF: If $f_1 \geq 0$ and $f_2 \geq 0$ are simple functions, the assertion follows from the definition of the integral for a simple function. If $f_1 \geq 0$ and $f_2 \geq 0$ are arbitrary functions, we choose monotonically increasing sequences $\{s'_n\}$ and $\{s''_n\}$ of nonnegative measurable functions that converge to f_1 and f_2 respectively. Let

$$s_n = s'_n + s''_n.$$

Then

$$\int_E s_n d\mu = \int_E s'_n d\mu + \int_E s''_n d\mu.$$

We now let $n \rightarrow \infty$ and apply Theorem 7 on passage to the limit. This finishes the proof for this case.

Now let $f_1 \geq 0$ and $f_2 \leq 0$. Set

$$A = \{x : f(x) \geq 0\}, \quad B = \{x : f(x) < 0\}.$$

Then the functions f , f_1 , and $-f_2$ are nonnegative on A and so

$$\int_A f_1 d\mu = \int_A f d\mu + \int_A (-f_2) d\mu = \int_A f d\mu - \int_A f_2 d\mu.$$

Similarly the functions $-f$, f_1 , and $-f_2$ are nonnegative on B , so that

$$\int_B (-f_2) d\mu = \int_B f_1 d\mu + \int_B (-f) d\mu, \quad \int_B f_1 d\mu = \int_B f d\mu - \int_B f_2 d\mu.$$

In the general case, one must take four sets E_i on each of which f_1 and f_2 are of constant sign. By what has been proved

$$\int_{E_i} f d\mu = \int_{E_i} f_1 d\mu + \int_{E_i} f_2 d\mu, \quad i = 1, 2, 3, 4. \blacksquare$$

COROLLARY 2. *Let E be a measurable set. If $\{f_n\}$ is a sequence of nonnegative measurable functions and*

$$f(x) = \sum_{n=1}^{\infty} f_n(x), \quad x \in E,$$

then

$$\int_E f d\mu = \sum_{n=1}^{\infty} \int_E f_n d\mu.$$

PROOF: In fact the partial sums of the series for $f(x)$ form a monotonically increasing sequence. All that remains is to apply the preceding theorem. \blacksquare

THEOREM (Fatou's lemma). *Suppose the set E is measurable, $\{f_n\}$ is a sequence of nonnegative measurable functions, and*

$$f(x) = \liminf_{n \rightarrow \infty} f_n(x)$$

for all $x \in E$. Then

$$\int_E f d\mu \leq \liminf_{n \rightarrow \infty} \int_E f_n d\mu.$$

PROOF: Set $g_n(x) = \inf_{i \geq n} f_i(x)$ for any $n = 1, 2, \dots$ and $x \in E$. Then the functions g_n are measurable on E and

$$0 \leq g_1(x) \leq g_2(x) \leq \dots, \\ g_n \leq f_n(x), \quad g_n(x) \rightarrow f(x), \quad \text{as } n \rightarrow \infty.$$

By Theorem 7 on passage to the limit we find that $\int_E g_n d\mu \rightarrow \int_E f d\mu$. Since $\int_E g_n d\mu \leq \int_E f_n d\mu$, it follows that $\int_E f d\mu \leq \varliminf_{n \rightarrow \infty} \int_E f_n d\mu$. ■

THEOREM. (Lebesgue's dominated convergence theorem). Let $E \in \mathcal{L}(X)$ and let $\{f_n\}$ be a sequence of measurable functions such that $f_n(x) \rightarrow f(x)$ for $x \in E$ and $n \rightarrow \infty$. If there exists a function $g \in L^1(\mu)$ such that

$$|f_n(x)| \leq g(x), \quad n = 1, 2, \dots, \quad x \in E,$$

then

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu = \int_E f d\mu.$$

The theorem continues to hold if $f_n \rightarrow f$ almost everywhere on E .

PROOF: We remark that Theorem 5 implies that $f_n \in L^1(\mu)$ and also that $f \in L^1(\mu)$. By Fatou's lemma $\int_E (f + g) d\mu \leq \varliminf_{n \rightarrow \infty} \int_E (f_n + g) d\mu$ since $f_n + g \geq 0$. Therefore $\int_E f d\mu \leq \varliminf_{n \rightarrow \infty} \int_E f_n d\mu$. Similarly $\int_E (g - f) d\mu \leq \varliminf_{n \rightarrow \infty} \int_E (g - f_n) d\mu$, since $g - f_n \geq 0$. Consequently

$$-\int_E f d\mu \leq \varliminf_{n \rightarrow \infty} \left[-\int_E f_n d\mu \right],$$

i.e.,

$$\int_E f d\mu \geq \overline{\lim}_{n \rightarrow \infty} \int_E f_n d\mu.$$

Therefore the limit $\lim_{n \rightarrow \infty} \int_E f_n d\mu$ exists, and $\int_E f d\mu = \lim_{n \rightarrow \infty} \int_E f_n d\mu$. ■

COROLLARY. If $\mu(E) < +\infty$, the sequence $\{f_n\}$ is uniformly bounded on E , and $f_n \rightarrow f$ for all $x \in E$, then

$$\int_E f d\mu = \lim_{n \rightarrow \infty} \int_E f_n d\mu.$$

In the considerations above we have assumed that f is a real-valued function. If f is complex-valued, $f = \varphi_1 + i\varphi_2$, then as already mentioned, f is measurable if φ_1 and φ_2 are measurable. It is easy to see that the sum and product of complex-valued measurable functions are measurable, and the absolute value of a measurable function ($|f| = (\varphi_1^2 + \varphi_2^2)^{1/2}$) is also measurable.

A complex-valued function f is *Lebesgue-integrable* on a measurable set E ($f \in L^1(\mu)$) if f is measurable and $\int_E |f| d\mu < \infty$, and here by definition

$$\int_E f d\mu = \int_E \varphi_1 d\mu + i \int_E \varphi_2 d\mu.$$

It is clear that $f \in L^1(\mu)$ if and only if φ_1 and φ_2 belong to $L^1(\mu)$, since $|\varphi_1| \leq |f|$, $|\varphi_2| \leq |f|$, and $|f| \leq |\varphi_1| + |\varphi_2|$. Obviously all the basic facts of the theory of integration for real-valued functions remain true for complex-valued functions also.

3.4. The Connection between the Riemann and Lebesgue Integrals

It is an important question how the Lebesgue integral is related to the Riemann integral. The following theorem holds.

THEOREM 8. *If a function f is Riemann-integrable on a closed interval $[a, b]$ (we denote this fact by writing $f \in R$), then it is also Lebesgue-integrable over the interval $[a, b]$, and*

$$\int_{[a,b]} f dx = \int_a^b f(x) dx = R \int_a^b f(x) dx.$$

PROOF: We recall that a partition P of the closed interval $[a, b]$ is a set of points $a = x_0 < x_1 < \dots < x_n = b$. Let $\Delta x_i = x_i - x_{i-1}$, $i = 1, 2, \dots, n$, $M_i = \sup_{x_{i-1} \leq x \leq x_i} f(x)$, $m_i = \inf_{x_{i-1} \leq x \leq x_i} f(x)$, $U(P, f) = \sum_{i=1}^n M_i \Delta x_i$, $L(P, f) = \sum_{i=1}^n m_i \Delta x_i$, and let $d(P) = \max_{1 \leq i \leq n} \Delta x_i$ be the diameter of the partition P .

A function $f(x)$ is Riemann-integrable if and only if for each $\varepsilon > 0$ there exists a partition P such that $U(P, f) - L(P, f) < \varepsilon$. A partition \tilde{P} is called

a refinement of the partition P if each point of the partition P belongs to the partition \tilde{P} . Now let $\{P_k\}$ be a sequence of partitions of the closed interval $[a, b]$ such that P_{k+1} is a refinement of the partition P_k and the diameter of the partition P_k tends to zero as $k \rightarrow \infty$ ($d(P_k) \rightarrow 0$ as $k \rightarrow \infty$). Let

$$U_k(a) = L_k(a) = f(a), \quad U_k(x) \equiv M_i, \quad L_k(x) \equiv m_i, \\ x_{i-1} \leq x \leq x_i, \quad i = 1, 2, \dots, n.$$

Then $U(P_k, f) = \int_{[a,b]} U_k(x) dx$, $L(P_k, f) = \int_{[a,b]} L_k(x) dx$. Since P_{k+1} is a refinement of the partition P_k , it follows that $U_1(x) \geq U_2(x) \geq \dots \geq f(x) \geq \dots \geq L_2(x) \geq L_1(x)$ for $a \leq x \leq b$. We set $U(x) = \lim_{k \rightarrow \infty} U_k(x)$ and $L(x) = \lim_{k \rightarrow \infty} L_k(x)$.

Now let $f \in R$. Since $d(P_k) \rightarrow 0$, it follows that

$$U(P_k, f) \rightarrow R \int_a^b f(x) dx, \quad L(P_k, f) \rightarrow R \int_a^b f(x) dx.$$

By the theorem on passage to the limit under the integral sign in the Lebesgue integral we have

$$\int_{[a,b]} U_k dx \rightarrow \int_{[a,b]} U dx, \quad \int_{[a,b]} L_k dx \rightarrow \int_{[a,b]} L dx.$$

Therefore

$$\int_{[a,b]} U dx = \int_{[a,b]} L dx = R \int_a^b f(x) dx.$$

But $L(x) \leq f(x) \leq U(x)$. Consequently

$$\int_{[a,b]} L(x) dx \leq \int_{[a,b]} f(x) dx \leq \int_{[a,b]} U(x) dx = \int_{[a,b]} L(x) dx.$$

Therefore $L(x) = f(x) = U(x)$ almost everywhere on $[a, b]$. Indeed $\int_{[a,b]} (U - f) dx = \int_{[a,b]} \varphi(x) dx = 0$ and $\varphi(x) \geq 0$. Then by the corollary to Chebyshev's

theorem we find that $\varphi(x) = 0$ almost everywhere. Consequently $U = f = L$ almost everywhere.

The functions U and L are measurable, and so the function f is also measurable. Thus $f \in L^1(\mu)$ and

$$\int_{[a,b]} L dx = \int_{[a,b]} f dx = \int_{[a,b]} U dx = R \int_a^b f(x) dx. \blacksquare$$

3.5. The Space L^p

The apparatus of measure theory and Lebesgue integration just developed makes it possible to introduce the class of spaces $L^p(\mu)$, $p \geq 1$, which are important in applications. (For $p = 1$ we obtain the space $L^1(\mu)$ of Lebesgue-integrable functions.)

DEFINITION 13. A function $f(x)$ belongs to the space $L^p(\mu)$ on a measurable space X with a countably additive measure μ if $f(x)$ is measurable and $|f(x)|^p \in L^1(\mu)$.

We remark that since a set of measure zero can be neglected in Lebesgue integration, the elements of the space $L^p(\mu)$ are actually equivalence classes of functions (two functions in a given class coincide almost everywhere). It follows from Young's inequality (cf. Example 4 of Sec. 1.2.1)

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \quad a \geq 0, b \geq 0, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad p > 1,$$

that if $f \in L^p(\mu)$ and $g \in L^p(\mu)$, $p \geq 1$, then $f + g \in L^p(\mu)^*$ and

$$\left(\int_X |f(x) + g(x)|^p d\mu \right)^{1/p} \leq \left(\int_X |f(x)|^p d\mu \right)^{1/p} + \left(\int_X |g(x)|^p d\mu \right)^{1/p}.$$

If we set $\|f\|_{L^p} = \left(\int_X |f|^p d\mu \right)^{1/p}$, then the space $L^p(\mu)$, $p \geq 1$ becomes a normed space with the usual operations on functions, and the metric

$$\rho(f, g) = \|f - g\|_{L^p}$$

makes it also a metric space. The axioms for a norm and a distance are verified immediately using the inequality just proved.

*For $p = 1$ the inequality $|f + g| \leq |f| + |g|$ is obvious.

The following theorem is important.

THEOREM 9. *The space $L^p(\mu)$, ($p \geq 1$) is a complete normed (metric) space.*

PROOF: Let $\{f_m\}$ be a fundamental sequence:

$$\lim_{n,m \rightarrow \infty} \|f_n - f_m\|_{L^p} = 0.$$

We shall prove that this sequence converges to an element of $L^p(\mu)$. We can find an index N_k such that

$$\|f_n - f_m\|_{L^p} \leq \frac{1}{2^k}, \quad \text{if } n, m \geq N_k.$$

Then

$$\|f_{N_k} - f_{N_{k+1}}\| \leq \frac{1}{2^k}.$$

Consider the partial sums $s_n(x)$ of the following series:

$$\begin{aligned} f_{N_1} + \sum_{k=1}^{\infty} (f_{N_{k+1}}(x) - f_{N_k}(x)), \\ s_n(x) = f_{N_1} + \sum_{k=1}^n (f_{N_{k+1}}(x) - f_{N_k}(x)) = f_{N_{n+1}}(x). \end{aligned}$$

We shall prove that the series introduced above converges absolutely almost everywhere.

We choose an arbitrary function $g(x) \in L^q(\mu)$, $\frac{1}{p} + \frac{1}{q} = 1$. Then by the easily established inequality*

$$\begin{aligned} \int_X |f \cdot g| d\mu &\leq \left(\int_X |f|^p d\mu \right)^{1/p} \cdot \left(\int_X |g|^q d\mu \right)^{1/q}, \\ \frac{1}{p} + \frac{1}{q} &= 1, \quad p > 1, q > 1, \end{aligned}$$

we obtain the relation

$$\begin{aligned} \int_X |g \cdot (f_{N_{k+1}} - f_{N_k})| d\mu &\leq \left(\int_X |f_{N_{k+1}} - f_{N_k}|^p d\mu \right)^{1/p} \cdot \left(\int_X |g|^q d\mu \right)^{1/q} \\ &\leq \frac{1}{2^k} \|g\|_{L^q}. \end{aligned}$$

*This inequality is established just like the corresponding inequality for l^p (cf. Sec. 1.2).

Thus

$$\sum_{k=1}^{\infty} \int_X |g(f_{N_{k+1}} - f_{N_k})| d\mu \leq \|g\|_{L^q}.$$

Applying Corollary 2 of Theorem 7, we reverse the order of summing and integrating. The result is

$$\int_X F(x) d\mu < \infty, \quad F(x) \geq 0 \quad \text{on } X,$$

and

$$F(x) = \sum_{k=1}^{\infty} |g(f_{N_{k+1}} - f_{N_k})|.$$

Consequently $F(x) < +\infty$ almost everywhere on X . It now follows that

$$|f_{N_1}| + \sum_{k=1}^{\infty} |f_{N_{k+1}} - f_{N_k}| < +\infty$$

almost everywhere. Otherwise, choosing g different from 0 on a set of positive measure (the set where the series diverges), we would arrive at a contradiction to the relation $F(x) < +\infty$.^{*} Since $s_n(x) = f_{N_{n+1}}(x)$, it follows that

$$\lim_{n \rightarrow \infty} s_n(x) = \lim_{n \rightarrow \infty} f_{N_{n+1}}(x) = f(x)$$

exists almost everywhere. We shall show that the function $f(x)$ so defined, first of all, belongs to $L^p(\mu)$ and, second, is the limit in the L^p -norm of the original sequence $\{f_m(x)\}$.

In fact let $\varepsilon > 0$ and let $N(\varepsilon)$ be such that

$$\|f_{N_m} - f_{N_n}\|_{L^p} \leq \varepsilon, \quad \text{for } N_n, N_m \geq N(\varepsilon).$$

Then by Fatou's lemma

$$\|f - f_{N_n}\|_{L^p} \leq \varliminf_{m \rightarrow \infty} \|f_{N_m} - f_{N_n}\|_{L^p} \leq \varepsilon, \quad N_n > N(\varepsilon).$$

^{*}For $p = 1$ we can use the inequality

$$\int_X \left[|f_{N_1}| + \sum_{k=1}^{\infty} |f_{N_{k+1}} - f_{N_k}| \right] d\mu \leq \|f_{N_1}\|_{L^1} + 1$$

and conclude from this that the series converges almost everywhere.

Thus

$$f - f_{N_n} \in L^p(\mu), \quad f = (f - f_{N_n}) + f_{N_n} \in L^p(\mu).$$

We now write the inequality

$$\|f - f_m\|_{L^p} \leq \|f - f_{N_n}\|_{L^p} + \|f_{N_n} - f_m\|_{L^p}.$$

We choose m and N_n so large that neither term on the right-hand side is larger than $\varepsilon/2$. Then $\|f - f_m\|_{L^p} \leq \varepsilon$, and consequently f_m tends to f in the L^p -norm. ■

In several branches of analysis frequent use is made of the fact that the continuous functions form an everywhere-dense subset in the space $L^p(\mu)$ when $X \subset \mathbf{R}^n$ and $\mu(X) < \infty$. To prove this it obviously suffices to prove that any element of the space $L^p(\mu)$ can be approximated in the L^p -norm to any degree of accuracy by a continuous function. We first show that the characteristic function $\chi_M(x)$ of a measurable set M with $\mu(M) < \infty$ can be approximated by continuous functions. By Example 1 of Sec. 3.1 for a given set M there exist a closed set F_M and an open set G_M such that

$$F_M \subset M \subset G_M, \quad \mu(G_M) - \mu(F_M) < \varepsilon.$$

Let

$$f_\varepsilon(x) = \frac{\rho(x, X \setminus G_M)}{\rho(x, X \setminus G_M) + \rho(x, F_M)},$$

where $\rho(x, A)$ is the distance from the point x to the set A . The function $f_\varepsilon(x)$ is 0 for $x \in X \setminus G_M$, 1 for $x \in F_M$, and continuous, since all the functions occurring in its definition are continuous and the denominator is never zero. The inequality $|\chi_M(x) - f_\varepsilon(x)| \leq 1$ holds for $x \in G_M \setminus F_M$ and $\chi_M(x) - f_\varepsilon(x) = 0$ for x outside $G_M \setminus F_M$. Consequently

$$\|\chi_M(x) - f_\varepsilon(x)\|_{L^p} = \left(\int_X |\chi_M(x) - f_\varepsilon(x)|^p d\mu \right)^{1/p} < \varepsilon^{1/p}.$$

It follows from what has been proved that any measurable simple function can be approximated in $L^p(\mu)$ by a sequence of continuous functions (since a simple function is a finite linear combination of characteristic functions).

Now let $f \in L^p(\mu)$ and $f \geq 0$. Let $\{s_n\}$ be a monotonically increasing sequence of nonnegative measurable simple functions such that $s_n(x) \rightarrow f(x)$ for all x . Now $|f - s_n|^p \leq |f|^p$, and so by the Lebesgue dominated convergence theorem $\|f - s_n\|_{L^p} \rightarrow 0$ as $n \rightarrow \infty$, i.e., s_n tends to f in $L^p(\mu)$. The representation $f = f^+ - f^-$ finishes the proof for any function $f \in L^p(\mu)$.

If we set $p = 1$ and $p = 2$ in the class of spaces $L^p(\mu)$ we have studied, $p \geq 1$, we arrive at two important examples of Banach spaces $L^1(\mu)$ and $L^2(\mu)$. The space $L^1(\mu)$ is called the space of integrable functions; the convergence defined by its norm is called *convergence in mean*.

The space $L^2(\mu)$ consists of equivalence classes of functions whose absolute values are square-integrable. The convergence defined by the norm of this space is called *mean-square convergence*.

For the case of a closed interval $X = [a, b]$ and Lebesgue measure on the interval, this space is usually denoted $L^2[a, b]$.

4. ABSOLUTELY CONTINUOUS SET FUNCTIONS. THE RADON-NIKODÝM THEOREM

4.1. Absolutely Continuous Set Functions

Suppose that on a measurable space (X, K_σ, μ) another countably additive measure $\nu(A) < \infty$ is defined for sets $A \in K_\sigma$.

DEFINITION 1. The measure ν is *absolutely continuous* with respect to the measure μ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\nu(A) < \varepsilon$ whenever $\mu(A) < \delta$.

DEFINITION 1'. The measure ν is *absolutely continuous* with respect to the measure μ if $\nu(A) = 0$ whenever $\mu(A) = 0$.

We shall prove that these definitions are equivalent.

Let ν be absolutely continuous in the sense of Definition 1. If $\mu(A) = 0$, then $\nu(A) < \varepsilon$ for any $\varepsilon > 0$, i.e., $\nu(A) = 0$. Hence ν is absolutely continuous in the sense of Definition 1'.

Now suppose $\nu(A) = 0$ whenever $\mu(A) = 0$. We must show that ν is absolutely continuous in the sense of Definition 1. Suppose the contrary. Then there exists $\varepsilon_0 > 0$ and a sequence of sets $A_n \in K_\sigma$ such that $\mu(A_n) < 1/2^n$ and $\nu(A_n) \geq \varepsilon_0$.

We set

$$E_n = \bigcup_{k=n}^{\infty} A_k$$

and

$$E = \bigcap_{n=1}^{\infty} E_n.$$

We have $\mu(E) = \lim_{n \rightarrow \infty} \mu(E_n)$ and $\nu(E) = \lim_{n \rightarrow \infty} \nu(E_n)$, since the measures μ and ν are countably additive and $E_1 \supset E_2 \supset E_3 \supset \dots$.

However

$$\mu(E_n) \leq \sum_{k=n}^{\infty} \mu(A_k) \leq \frac{1}{2^{n-1}} \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

while

$$\nu(E_n) \geq \nu(A_n) \geq \varepsilon_0.$$

Therefore $\mu(E) = 0$, $\nu(E) \geq \varepsilon_0$, in contradiction to Definition 1'. ■

We now give some examples.

EXAMPLES

1. Let $f(x) \geq 0$ and $f(x) \in L^1(\mu)$. For $A \in K_\sigma$ we set

$$\nu(A) = \int_A f d\mu.$$

Then ν is an example of a measure that is absolutely continuous with respect to the measure μ .

2. Let $X = [0, 1]$, let K_σ be the Lebesgue-measurable sets, and let μ be Lebesgue measure on the closed interval $[0, 1]$. We define the measure $\nu(A)$ as follows:

$$\nu(A) = \begin{cases} 1, & \text{if } \frac{1}{2} \in A, \\ 0, & \text{if } \frac{1}{2} \notin A. \end{cases}$$

It is easy to verify that $\nu(A)$ is an additive measure and

$$\nu(\{1/2\}) = 1.$$

Therefore $\nu(A)$ is not absolutely continuous with respect to Lebesgue measure.

4.2 The Radon-Nikodým Theorem

We shall prove the following lemma.

LEMMA. *Let ν be a measure that is absolutely continuous with respect to the measure μ . If $\nu(X) \neq 0$, then there exists a function $f(x) \geq 0$ not equivalent to zero, $f(x) \in L^1(\mu)$, with $\nu(A) \geq \int_A f d\mu$ for any $A \in K_\sigma$.*

PROOF: Fix $\varepsilon > 0$. Set $f_0(x) \equiv \varepsilon$. Consider all the sets $A \in K_\sigma$ such that

$$\nu(A) < \int_A f_0 d\mu.$$

Let $\lambda_1 = \sup \mu(A)$, the supremum being taken over all such sets A . Consider two cases:

- a) $\lambda_1 = \mu(X)$;
- b) $\lambda_1 < \mu(X)$.

We shall show that in the second case the desired function $f(x)$ can be found.

We choose A_1 so that $\mu(A_1) > \lambda_1/2$. On $X \setminus A_1$ consider again all the set $A \in K_\sigma$ for which

$$\nu(A) < \int_A f_0 d\mu.$$

Let $\lambda_2 = \sup \mu(A)$ over all such sets A . We shall show that

$$\lambda_2 \leq \lambda_1/2.$$

Indeed, if $\lambda_2 > \lambda_1/2$, then there would exist a set $A' \subseteq X \setminus A_1$ such that

$$\nu(A') < \int_{A'} f_0 d\mu, \quad \mu(A') > \frac{\lambda_1}{2}.$$

Then $\mu(A' \cup A_1) > \lambda_1$ and

$$\nu(A' \cup A_1) < \int_{A' \cup A_1} f_0 d\mu,$$

contradicting the choice of the number λ_1 .

We choose a set $A_2 \subseteq X \setminus A_1$ such that $\mu(A_2) > \lambda_2/2$ and continue this proces indefinitely. We obtain a sequence of numbers $\{\lambda_n\}$ and sets $\{A_n\}$ such that

$$\lambda_{n+1} \leq \lambda_n/2, \quad \mu(A_n) > \lambda_n/2$$

and $\lambda_n = \sup \mu(A)$, where the supremum is taken over all sets $A \in K_\sigma$, $A \subset X \setminus A_1 \setminus A_2 \setminus \dots \setminus A_{n-1}$ such that

$$\nu(A) < \int_A f_0 d\mu.$$

It follows from these conditions that

$$\lambda_{n+1} \leq \lambda_1/2^n.$$

Let $E = A_1 \cup A_2 \cup A_3 \cup \dots$. We shall prove that for any set $A \subset X \setminus E$ such that $\mu(A) > 0$, the equality

$$\nu(A) \geq \int_A f_0 d\mu$$

holds. Indeed, if for some A_0 the relations

$$\nu(A_0) < \int_{A_0} f_0 d\mu \quad \text{and} \quad A_0 \subset X \setminus E$$

hold, then for any n we have

$$\lambda_n \geq \mu(A_0).$$

This relation contradicts the inequality $\lambda_{n+1} \leq \lambda_1/2^n$. We remark further that $\mu(E) \leq \lambda_1 < \mu(X)$ since

$$\nu(E) < \int_E f_0 d\mu.$$

Therefore the function

$$f(x) = \begin{cases} 0, & x \in E, \\ f_0(x), & x \notin E. \end{cases}$$

is not equivalent to 0. It is obvious that $f(x) \geq 0$ and $f(x) \in L^1(\mu)$.

Let $A \in K_\sigma$ be an arbitrary measurable set. Then

$$\nu(A) = \nu(A \cap E) + \nu(A \setminus E) \geq \nu(A \setminus E) \geq \int_{A \setminus E} f d\mu = \int_A f d\mu.$$

Thus in the second case the desired function $f(x)$ exists.

Now if $\lambda_1 = \mu(X)$, then there exists a sequence of sets $A_n \in K_\sigma$ with $\mu(A_n) \rightarrow \mu(X)$ such that

$$\nu(A_n) \leq \int_{A_n} f_0 d\mu = \varepsilon \mu(A_n).$$

By the absolute continuity of the measure ν

$$\nu(X \setminus A_n) \rightarrow 0$$

as $n \rightarrow \infty$. Hence

$$\nu(X) \leq \varepsilon \mu(X).$$

Therefore if $\varepsilon < \nu(X)/\mu(X)$, the case a) is impossible.

The lemma is now proved. ■

Suppose, as above, (X, K_σ, μ) is a measurable space. We shall prove the following important theorem.

THEOREM (Radon-Nikodým). *Let ν be a measure that is absolutely continuous with respect to the measure μ . Then there exists a function $f(x) \in L^1(\mu)$, $f(x) \geq 0$, such that*

$$\nu(A) = \int_A f d\mu$$

for any set $A \in K_\sigma$. (The function f can be called the derivative of the measure ν with respect to the measure μ .)

PROOF: Let F^+ be the set of functions $f(x) \in L^1(\mu)$ such that $f(x) \geq 0$ and

$$\int_A f d\mu \leq \nu(A)$$

for any $A \in K_\sigma$. Let

$$M = \sup_{f \in F^+} \int_X f d\mu$$

and $f_n(x)$ such that

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = M.$$

We set

$$g_n(x) = \max[f_1(x), f_2(x), \dots, f_n(x)].$$

For any $A \in K_\sigma$

$$\int_A g_n d\mu \leq \nu(A).$$

Indeed the set A can be represented in the form $\bigcup_{k=1}^n A_k$, where the A_k are pairwise disjoint and $g_n(x) = f_k(x)$ for $x \in A_k$. Therefore

$$\int_A g_n d\mu = \sum_{k=1}^n \int_{A_k} g_n d\mu = \sum_{k=1}^n \int_{A_k} f_k d\mu \leq \sum_{k=1}^n \nu(A_k) = \nu(A).$$

We set

$$f_0(x) = \lim_{n \rightarrow \infty} g_n(x),$$

where the limit exists for almost every $x \in X$ (with respect to the measure μ) by the monotone convergence theorem. Then

$$\int_X f_0 d\mu = M$$

and

$$\int_A f_0 d\mu \leq \nu(A) \quad \text{for } A \in K_\sigma.$$

We shall show that

$$\nu(A) - \int_A f_0 d\mu \equiv 0.$$

Indeed, $\lambda(A) = \nu(A) - \int_A f_0 d\mu$ is a measure that is absolutely continuous with respect to μ . If $\lambda(A) \not\equiv 0$, then by the lemma there exists $f(x) \geq 0$, $f(x) > 0$ on a set of positive measure, $f \in L^1(\mu)$, such that

$$\lambda(A) - \int_A f d\mu \geq 0.$$

That is, $f + f_0 \in F^+$ and $\int_X (f + f_0) d\mu > M$. This contradiction proves the theorem. ■

5. PRODUCT MEASURES. FUBINI'S THEOREM

In many problems of analysis theorems on the reduction of a double (or multiple) integral to an iterated integral play an important role. In the case of multiple Lebesgue integrals the fundamental result in this area is Fubini's theorem.

5.1. Product Measures

If Y_1, Y_2, \dots, Y_n are systems of subsets of the sets X_1, X_2, \dots, X_n , then

$$Y = Y_1 \times Y_2 \times \cdots \times Y_n = \prod_{k=1}^n Y_k$$

denotes the system of subsets of the set

$$X = X_1 \times X_2 \times \cdots \times X_n = \prod_{k=1}^n X_k,$$

that are representable in the form

$$A = A_1 \times A_2 \times \cdots \times A_n = \prod_{k=1}^n A_k,$$

where $A_k \in Y_k$.

The following proposition holds.

PROPOSITION. If $P_1(X_1), P_2(X_2), \dots, P_n(X_n)$ are semirings, then the class $P(X) = P_1(X_1) \times P_2(X_2) \times \cdots \times P_n(X_n) = \prod_{k=1}^n P_k(X_k)$ is a semiring.

PROOF: For simplicity we consider the case $n = 2$. Let $A, B \in P(X)$. If $A = A_1 \times A_2$, $B = B_1 \times B_2$, where $A_i, B_i \in P_i(X_i)$, $i = 1, 2$, then

$$A \cap B = (A_1 \cap B_1) \times (A_2 \cap B_2) \in P_1(X_1) \times P_2(X_2) = P(X).$$

In addition suppose $B_1 \subset A_1$, $B_2 \subset A_2$. Then there exist sets $B_1^{(i)} \in P_1(X_1)$, $B_2^{(j)} \in P_2(X_2)$, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, m$ such that

$$A_1 \setminus B_1 = \bigsqcup_{i=1}^n B_1^{(i)}, \quad A_2 \setminus B_2 = \bigsqcup_{j=1}^m B_2^{(j)}.$$

Therefore

$$(A_1 \times A_2) \setminus (B_1 \times B_2) = \bigsqcup_{i=1}^n \bigsqcup_{j=1}^m (B_1^{(i)} \times B_2^{(j)}),$$

where $B_1^{(i)} \times B_2^{(j)} \in P_1(X_1) \times P_2(X_2)$. ■

We remark that it does not in general follow that if $K_1(X_1), K_2(X_2), \dots, K_n(X_n)$ are rings, then their product is a ring.

Let $P_1(X_1)$ and $P_2(X_2)$ be two semirings and μ and ν measures on $P_1(X_1)$ and $P_2(X_2)$. Consider the semiring $P(X) = P_1(X_1) \times P_2(X_2)$ and define a function $\mu \times \nu$ on it by setting $(\mu \times \nu)(A \times B) = \mu(A) \cdot \nu(B)$. This function, as one can verify, is additive, and is called the *product* of the measures μ and ν .

We prove the following lemma.

LEMMA. If the measures μ and ν are countably additive, then the measure $\mu \times \nu$ has the same property.

PROOF: For any set $C = A \times B$ of the semiring $P_1(X_1) \times P_2(X_2)$ we define the function

$$f_C(x_1) = \chi_A(x_1)\nu(B),$$

where $\chi_A(x_1)$ is the characteristic function of the set A . Let $C = \bigsqcup_{k=1}^{\infty} C_k$, $C_k \in P(X) = P_1(X_1) \times P_2(X_2)$. It follows from the countable additivity of the measure ν that

$$f_C(x_1) = \sum_{k=1}^{\infty} f_{C_k}(x_1).$$

By Lebesgue's dominated convergence theorem we have

$$\int_X f_C d\mu = \sum_{k=1}^{\infty} \int_X f_{C_k} d\mu.$$

Therefore

$$(\mu \times \nu)(C) = \sum_{k=1}^{\infty} (\mu \times \nu)(C_k),$$

which was to be proved. ■

Consider the mapping $C \rightarrow f_C$ defined by the rule

$$f_C(x_1) = \chi_A(x_1) \cdot \nu(B), \quad C = A \times B.$$

Extend it to the minimal ring $K_0(X)$ containing the semiring $P(X)$ by the rule

$$f_{\bigsqcup_{k=1}^n C_k} = \sum_{k=1}^n f_{C_k}.$$

We remark that if $C_1 = A_1 \sqcup B$, and $C_2 = A_2 \sqcup B$, where $B = C_1 \cap C_2$, then $f_{C_1} - f_{C_2} = f_{A_1} - f_{A_2}$ and $f_{C_1 \Delta C_2} = f_{A_1} + f_{A_2}$.

Therefore

$$\left| \int_X (f_{C_1} - f_{C_2}) d\mu \right| \leq \|f_{C_1} - f_{C_2}\|_{L^1(\mu)} \leq (\mu \times \nu)(C_1 \Delta C_2) = \rho(C_1, C_2).$$

Therefore the mapping $C \rightarrow f_C$ extends to a mapping of the entire σ -algebra $\mathcal{L}(X_1 \times X_2)$ of sets that are measurable with respect to $(\mu \times \nu)$ into the space

$L^1(\mu)$ of functions on X_1 that are integrable with respect to the measure μ . The formula for the extension is

$$f_{\lim_n C_n} = \lim_n f_{C_n},$$

where the first limit is taken in $\mathcal{L}(X_1 \times X_2)$ and the second in $L^1(\mu)$. Indeed the inequalities given above provide an estimate of the distance $\rho(f_{C_1}, f_{C_2})$ in the space $L^1(\mu)$ in terms of the distance in the space $\mathcal{L}(X_1 \times X_2)$ of measurable functions. Therefore it is permissible to pass to the limit, and the mapping $C \rightarrow f_C$ extends to the entire σ -algebra.

The following lemmas hold.

LEMMA. Let the set C belong to $\mathcal{L}(X_1 \times X_2)$. Then for almost all $x \in X_1$ the set $C_{x_1} \subset X_2$ given by the formula $C_{x_1} = \{x_2 \in X_2 : (x_1, x_2) \in C\}$ is measurable with respect to ν and $\nu(C_{x_1}) = f_C(x_1)$.

PROOF: For the sets in the minimal ring $K_0(X)$, $X = X_1 \times X_2$, this follows from the definition of f_C .

It $\{C^{(n)}\}$ is a monotonic sequence of sets,* then by the countable additivity of the measure ν we have the relation

$$\nu(\lim_n C_{x_1}^{(n)}) = \lim_n \nu(C_{x_1}^{(n)}).$$

Therefore the equality $\nu(C_{x_1}) = f_C(x_1)$ remains valid under a monotonic passage to the limit.

We shall now show that every measurable set C ($C \in \mathcal{L}(X_1 \times X_2)$) can be obtained up to a set of measure zero from sets of the minimal ring $K_0(X) = K_0(X_1 \times X_2)$ by two monotonic passages to the limit. Indeed, let

$$C_n \in K_0(X_1 \times X_2) \quad \text{and} \quad \rho(C_n, C) \leq 2^{-n},$$

where the approximation is taken using the measure $\mu \times \nu$. We now let $D = \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} C_{n+k}$. Then $(\mu \times \nu)(C \setminus \bigcup_{k=1}^{\infty} C_{n+k}) = 0$ and $(\mu \times \nu)(\bigcup_{k=1}^{\infty} C_{n+k} \setminus C) \leq \frac{1}{2^{n-1}}$, i.e., $(\mu \times \nu)(D \Delta C) = 0$. Therefore f_C and f_D coincide almost everywhere.

Consequently for almost all $x_1 \in X_1$

$$f_C(x_1) = f_D(x_1) = \nu(D_{x_1}) = \nu(C_{x_1}). \blacksquare$$

*Cf. Sec. 1.1.1.1.

Now consider the two measure spaces (X_1, K_σ^1, μ) and (X_2, K_σ^2, ν) . The following lemma holds.

LEMMA. Let μ and ν be σ -finite measures and C a subset of $X_1 \times X_2$ that is measurable with respect to the measure $(\mu \times \nu)$. Then for almost all $x_1 \in X_1$ (with respect to the measure μ) the set C_{x_1} is measurable with respect to the measure ν . The function $f_C(x_1) = \nu(C_{x_1})$ is measurable with respect to the measure μ , and

$$(\mu \times \nu)(C) = \int_{X_1} f_C d\mu$$

(here both sides may be simultaneously infinite.)

PROOF: If the set C has finite measure, the lemma follows from the assertion contained in the preceding lemma, and from the fact that the relation $(\mu \times \nu)(C) = \int_{X_1} f_C d\mu$ is preserved under passage to the limit on the left in the space $\mathcal{L}(X_1 \times X_2)$ and on the right in the space $L^1(\mu)$.

If the measure of the set C is infinite, there exists an increasing family of subsets of finite measure $C_n \subset C$ such that $\cup C_n = C$ and $(\mu \times \nu)(C_n) \rightarrow \infty$. Then

$$f_C(x_1) = \lim_n f_{C_n}(x_1) \quad \text{and} \quad \int_{X_1} f_{C_n} d\mu = (\mu \times \nu)(C_n) \rightarrow \infty. \blacksquare$$

REMARK 1. By virtue of the fact that the measures μ and ν , as well as the sets X_1 and X_2 , occur symmetrically throughout, we can write

$$(\mu \times \nu)(C) = \int_{X_2} \mu(C'_{x_2}) d\nu,$$

where $C'_{x_2} = \{x_1 \in X_1 : (x_1, x_2) \in C\}$. It follows also from this that

$$\int_{X_1} \nu(C_{x_1}) d\mu = \int_{X_2} \mu(C'_{x_2}) d\nu.$$

REMARK 2. For the product of three measure spaces

$$(X_1, K_\sigma^1(X_1), \mu), \quad (X_2, K_\sigma^2(X_2), \nu), \quad (X_3, K_\sigma^3(X_3), \lambda),$$

we can write similarly

$$(\mu \times \nu \times \lambda)(C) = \int_{X_1 \times X_2} \lambda(C_{x_1, x_2}) d(\mu \times \nu) = \int_{X_3} (\mu \times \nu)(C_{x_3}) d\lambda,$$

where

$$C_{x_1, x_2} = \{x_3 \in X_3 : (x_1, x_2, x_3) \in C\},$$

$$C_{x_3} = \{(x_1, x_2) \in X_1 \times X_2 : (x_1, x_2, x_3) \in C\}.$$

5.2. Fubini's Theorem

The following important theorem holds.

THEOREM (Fubini). *Let $f(x_1, x_2)$ be a function on the product of the spaces $(X_1, K_\sigma^1(X_1), \mu)$ and $(X_2, K_\sigma^2(X_2), \nu)$ that is integrable with respect to $(\mu \times \nu)$. Then*

- a) *for almost all $x_1 \in X_1$ (with respect to the measure μ) the function $f(x_1, x_2)$ is integrable on X_2 (with respect to the measure ν) and its integral over X_2 is an integrable function on X_1 ;*
- b) *for almost all $x_2 \in X_2$ (with respect to the measure ν) the function $f(x_1, x_2)$ is integrable on X_1 (with respect to the measure μ) and its integral over X_1 is an integrable function on X_2 ;*
- c) *the following equalities hold:*

$$\int_{X_1 \times X_2} f(x_1, x_2) d(\mu \times \nu) = \int_{X_1} \left(\int_{X_2} f(x_1, x_2) d\nu \right) d\mu = \int_{X_2} \left(\int_{X_1} f(x_1, x_2) d\mu \right) d\nu;$$

- d) *for nonnegative functions that are measurable with respect to the measure $(\mu \times \nu)$ the existence of either of the iterated integrals implies the existence of the double integral, i.e., the integrability of $f(x_1, x_2)$ over $X_1 \times X_2$.*

PROOF: Consider the case of a nonnegative function f . Let the set C be contained in $X_1 \times X_2 \times X_3$, where $X_3 = \mathbf{R}^1$ is the real axis with the usual Lebesgue measure $\lambda = dx_3$, and

$$C = \{(x_1, x_2, x_3) \in (X_1 \times X_2 \times \mathbf{R}^1) : 0 \leq x_3 \leq f(x_1, x_2)\}.$$

We apply the relations written out in Remark 2 above to this case. We obtain

$$\begin{aligned} C_{x_1, x_2} &= \{x_3 \in \mathbf{R}^1 : 0 \leq x_3 \leq f(x_1, x_2)\}; \quad \lambda(C_{x_1, x_2}) = f(x_1, x_2); \\ C_{x_1} &= \{(x_1, x_2) \in X_2 \times \mathbf{R}^1 : 0 \leq x_3 \leq f(x_1, x_2)\}; \\ (\mu \times \lambda)(C_{x_1}) &= \int_{X_2} f(x_1, x_2) d\nu. \end{aligned}$$

All the assertions of the theorem follow from this for a nonnegative function. The decomposition $f = f^+ - f^-$ finishes the proof in the general case. ■

NOTE. Let $K_\sigma(X)$ be a σ -ring. A real-valued (resp. complex-valued) function ν on $K_\sigma(X)$ is called a *signed measure* (resp. *complex measure*) if it is countably additive in the following sense: for any $A_k \in K_\sigma(X)$, if $A = \bigcup_{k=1}^{\infty} A_k$ (so that A belongs to $K_\sigma(X)$), then the series $\sum_{k=1}^{\infty} \nu(A_k)$ converges and its sum is $\nu(A)$.

The concept of a signed measure is a natural extension of the concept of a measure, and many facts discussed above remain true for signed measures.

EXERCISES

1. Show that a system of sets that is closed with respect to the operations of union and intersection is in general not a ring.
2. Prove that a system of sets closed with respect to the operations of union and difference is a ring.
3. Denote by $X = \{a, b, c\}$ a set of three elements, and let 2^X denote the set of all its subsets. Describe all the semirings and rings that can be constructed from elements of 2^X .
4. Prove that the cardinality of the set of Lebesgue-measurable subsets of the closed interval $[0, 1]$ is larger than that of the continuum.
5. Prove that every Lebesgue-measurable set on the line is the union of a Borel set and a set of measure zero.
6. Find the Lebesgue measure of the subset of the unit square in the plane consisting of the points (x, y) such that $|\sin x| < 1/2$ and $\cos(x + y)$ is irrational.
7. Let the function f be Lebesgue-integrable on the set X and assume $\mu(X) < \infty$. Then

$$\int_X f d\mu = \lim_{d(T) \rightarrow 0} \sum_k \xi_k \mu(\{x \in X, t_k \leq f \leq t_{k+1}\}),$$

where $T = \{t_k\}$ is a partition of the real axis, $d(T)$ is the diameter of the partition, and ξ_k any point such that $\xi_k \in [t_k, t_{k+1}]$. Thus the Lebesgue integral can be calculated as a limit of sums, but, in contrast to the Riemann integral, the partition is taken on the range of values of the function f . Prove this.

(The assertion of this exercise remains true even when $\mu(X) = \infty$, if we assume in addition that $\xi_k = 0$ for those k for which the interval $[t_k, t_{k+1}]$ contains the point 0.)

8. Let φ be a monotonically increasing smooth function on the closed interval $[a, b]$ and $\psi = \varphi^{-1}$ the function inverse to it.

Regarding the integral as the limit of sums, prove that

$$\int_{[a,b]} \varphi \, dx = \int_{[\varphi(a), \varphi(b)]} y \psi' \, dy.$$

9. For which values of the parameters α and β is the function $x^\alpha \sin x^\beta$ defined on $[0, 1]$ Lebesgue-integrable?

10. Let $f \geq 0$ be a measurable function on X , $\mu(X) < \infty$. Then f is integrable if and only if the series $\sum_{k=0}^{\infty} 2^k \mu(\{x \in X : f \geq 2^k\})$ converges. Prove this.

11. Calculate the Lebesgue integral over the closed interval $[0, \pi]$ of the function

$$f(x) = \begin{cases} \cos x, & \text{if } x \text{ is rational,} \\ \sin x, & \text{if } x \text{ is irrational.} \end{cases}$$

12. Let g be a measurable function on the real axis and f a continuous real-valued function. Show that $h = g(f)$ is in general not measurable.

13. Let f be a real-valued function. For which n does it follow that f is measurable if $|f|^n$ is measurable?

14. Let f be differentiable on the closed interval $[0, 1]$. Prove that f' is Lebesgue-measurable.

15. Investigate the following sequences of functions for convergence and uniform convergence:

$$f_n(x) = \frac{nx}{n^2 + x^2}, \quad x \in \mathbf{R}^1; \quad f_n(x) = x^n, \quad x \in [0, 1].$$

16. Let $f_n(x) = \frac{n \sin nx}{1 + n^2 \sin^2 x}$, $x \in [0, \pi]$. Let $\delta > 0$ be given. Exhibit explicitly a set E_δ such that $f_n(x)$ converges uniformly on the set $[0, \pi] \setminus E_\delta$ (cf. Egorov's theorem).