CHAPTER I.

The integral in an abstract space.

§ 1. Introduction. Apart from functions having as argument a variable number, or system of n numbers (point in n-dimensional space), we shall discuss in this book functions for which the independent variable is a set of points. Functions of this kind have occurred already in classical Analysis, in several important particular cases. But they only began to be studied in their full generality during the growth of the Theory of Sets, and in close relation to the parts of Analysis directly based on that theory.

If we are, for instance, given a function $f(x)$ integrable on every interval, then by associating with each interval I the value of the integral of $f(x)$ over I, we obtain a function $F(I)$ that is a function of an interval. Similarly, by taking multiple integrals of functions $f(x_1, x_2, ..., x_n)$ of n variables, we are led to consider functions of more general sets lying in spaces of several dimensions, the argument I of our function $F(I)$ being now replaced by any set for which the integral of our given function $f(x_1, x_2, ..., x_n)$ is defined.

We dwell on these examples in order to emphasize the natural connection between the notion of integral (in any sense) and that of function of a set. Needless to say, there are many other examples of functions of a set. Thus in elementary geometry, we have for instance, the length of a segment or the area of a polygon. The class of values of the argument of these two functions (the length and the area) is in the first case, the class of segments and in the second, that of polygons. The problem of extending these classes gave birth to the general theories of measure, in which the notions of length, area, and volume, defined in elementary geometry for a restricted number of figures, are now extended to sets of points

1

of much greater diversity. It is, nevertheless, remarkable that these researches arose far less from problems of Geometry than from their connection with problems of Analysis, above all with the tendency to generalize, and to render more precise, the notion of definite integral. This connection has occasionally found expression even in the terminology. Thus du Bois-Reymond called *integrable* the sets that to-day are said to be of measure zero in the Jordan sense.

The theories of measure have, in the course of their development, been modified in accordance with the changing requirements of the Theory of Functions. In our account, the most important part will be played by the theory of H. Lebesgue.

Lebesgue's theory of measure has made it possible to distinguish in Euclidean spaces a vast class of sets, called *measurable*, in which measure has the property of *complete additivity* — by this we mean that the measure of the sum of a sequence, even infinite, of measurable sets, no two of which have points in common, is equal to the sum of the measures of these sets. The importance of this class of sets is due to the fact that it includes, in particular, (with their classical measures), all the sets of points occurring in problems of classical Analysis, and further, that the fundamental operations applied to measurable sets lead always to measurable sets.

It is nevertheless to be observed that the ground was prepared for Lebesgue's theory of measure by earlier theories associated with the names of Cantor, Stolz, Harnack, du Bois-Reymond, Peano, Jordan, Borel, and others. These earlier theories have, however, to-day little more than historical value. They, too, were suitable instruments for studying and generalizing the notion of integral understood in the classical sense of Riemann, but their results in this direction have been largely artificial and accidental. It is only Lebesgue's theory of measure that makes a decisive step in the development of the notion of integral. This is the more remarkable in that the definition of Lebesgue apparently requires only a very small modification of a formal kind in the definition of integral due to Riemann.

To fix the ideas, let us consider a bounded function $f(x, y)$ of two variables, or what comes to the same thing, a bounded function of a variable point defined on a square K_0. In order to determine its Riemann integral, or more precisely, its lower Riemann-Darboux integral over K_0, we proceed as follows. We divide the

square K_0 into an arbitrary finite number of non-overlapping rectangles R_1, R_2, ..., R_n, and we form the sum

$$(1.1) \qquad \sum_{i=1}^{n} v_i \cdot m(R_i)$$

where v_i denotes the lower bound of the function f on R_i, and $m(R_i)$ denotes the area of R_i. The upper bound of all sums of this form is, by definition, the *lower Riemann-Darboux integral* of the function f over K_0. We define similarly the upper integral of f over K_0. If these two extreme integrals are equal, their common value is called the *definite Riemann integral* of the function f over K_0, and the function f is said to be *integrable in the Riemann sense* over K_0.

The extension of measure to all sets measurable in the Lebesgue sense, has rendered necessary a modification of the process of Riemann-Darboux, it being natural to consider sums of the form (1.1) for which $\{R_i\}_{i=1, 2, ..., n}$ is a subdivision of the square K_0 into a finite number of arbitrary measurable sets, not necessarily either rectangles or elementary geometrical figures. Accordingly, $m(R_i)$ is to be understood to mean the measure of R_i. The v_i retain their former meaning, i. e. represent the lower bounds of f on the corresponding sets R_i. We might call the upper bound of the sums (1.1) interpreted in this way the *lower Lebesgue integral* of the function f over K_0. But actually, this process is of practical importance only for a class of functions, called *measurable*, and for these the number obtained as the upper bound of the sums (1.1) is called simply the *definite Lebesgue integral* of f over K_0. What is important, is that the functions which are measurable in the sense of Lebesgue, and whose definition is closely related to that of the measurable sets, form a very general class. This class includes, in particular, all the functions integrable in the Riemann sense.

Apart from this, the method of Lebesgue is not only more general, but even, from a certain point of view, simpler than that of Riemann-Darboux. For, it dispenses with the simultaneous introduction of two extreme integrals, the lower and the upper. Thanks to this, Lebesgue's method lends itself to an immediate extension to unbounded functions, at any rate to certain classes of the latter, for instance, to all measurable functions of constant sign (cf. below § 10). Finally, the Lebesgue integral renders it permissible to integrate term by term sequences and series of functions in certain general cases where passages to the limit under the in-

tegral sign were not allowed by the earlier methods of integration. The reason for this is to be found in the complete additivity of Lebesgue measure. The fundamental theorems of Lebesgue (cf. below § 12) stating the precise circumstances under which term by term integration is permissible, are justly regarded by Ch. J. de la Vallée Poussin [I, p. 44] as one of the finest results of the theory.

Lebesgue's theory of measure has, in its turn, led naturally to further important generalizations. Instead of starting with area, or volume, of figures, we may imagine a mass distributed in the Euclidean space under consideration, and associate with each set as its measure, (its "weight" according to Ch. J. de la Vallée Poussin [I, Chap.VI; 1]), the amount of mass distributed on the set. This, again, leads to a generalization of the integral, parallel to Lebesgue's, known as the Lebesgue-Stieltjes integral. In order to present a unified account of the latter, we shall consider in this chapter an additive class of measurable sets given *a priori* in an arbitrary abstract space. We shall suppose further, that in this class, a completely additive measure is determined for the measurable sets. These hypotheses determine completely a corresponding method of integration in the Lebesgue sense. All the essential properties of the ordinary Lebesgue integral, except at most those implying the process of derivation, remain valid for this abstract integral. From this point of view, in a more or less general form, the Lebesgue integral has been studied by a number of authors, among whom we may mention J. Radon [1], P. J. Daniell [2], O. Nikodym [2] and B. Jessen [1]. For further generalizations (of a somewhat different kind) see also A. Kolmogoroff [1], S. Bochner [1], G. Fichtenholz and L. Kantorovitch [1], and M. Gowurin [1].

§ 2. Terminology and notation.

Given two sets A and B, we write $A \subset B$ when the set A is a *subset* of the set B, i. e. when every *element* of A is an element of B. When we have both $A \subset B$ and $B \subset A$, i. e. when the sets A and B consist of the same elements, we write $A = B$. Again, $a \epsilon A$ means that a is an element of the set A (*belongs* to A). By the *empty* set, we mean the set without any element; we denote it by 0. A set A is *enumerable* if there exists an infinite sequence of distinct elements $a_1, a_2, ..., a_n, ...$ consisting of all the elements of the set A.

Given a class $\mathfrak{A}$ of sets, we call *sum* of the sets belonging to this class, the set of all the objects each of which is an element of at least one set belonging to the class $\mathfrak{A}$. We call *product,* or *common part,* of the sets belonging to the class $\mathfrak{A}$, the set of all the objects that belong at the same time to all the sets of this class. We call *difference* of two sets A and B, and we denote by $A - B$, the set of all the objects that belong to A without belonging to B.

Given a sequence of sets $\{A_n\}$ — a finite sequence $A_1, A_2, ..., A_n$, or an infinite sequence $A_1, A_2, ..., A_n, ...$ — we denote the sum by $\sum_i A_i$, by $\sum_{i=1}^{n} A_i$, or by $A_1 + A_2 + ... + A_n$, in the finite case, and by $\sum_i A_i$, by $\sum_{i=1}^{\infty} A_i$, or by $A_1 + A_2 + ... + A_n + ...$ in the infinite case. Similarly, merely replacing the sign Σ by Π, we have the expression for the product of a sequence of sets. If the sequence $\{A_n\}$ is infinite, we call *upper limit* of this sequence, the set of all the elements a such that $a \epsilon A_n$ holds for an infinity of values of the index n. The set of all the elements a belonging to all the sets A_n from some n (in general depending on a) onwards, we call *lower limit* of the sequence $\{A_n\}$. The upper and lower limits of the sequence $\{A_n\}_{n=1,2,...}$, we denote by $\lim \inf\limits_n A_n$ and $\lim \inf\limits_n A_n$ respectively. We have

$$(2.1) \qquad \lim_n \inf A_n = \sum_{k=1}^{\infty} \prod_{n=k}^{\infty} A_n \subset \prod_{k=1}^{\infty} \sum_{n=k}^{\infty} A_n = \lim_n \sup A_n.$$

If $\lim\limits_n \sup A_n = \lim\limits_n \inf A_n$, the sequence $\{A_n\}$ is said to be *convergent*; its upper and lower limits are then called simply *limit* and denoted by $\lim\limits_n A_n$.

If, for a sequence $\{A_n\}$ of sets, we have $A_n \subset A_{n+1}$, for each n, the sequence $\{A_n\}$ is said to be *ascending*, or *non-decreasing*; if, for each n, we have $A_{n+1} \subset A_n$, the sequence $\{A_n\}$ is said to be *descending* or *non-increasing*. Ascending and descending sequences are called *monotone*. We see directly that every monotone sequence is convergent, and that we have $\lim\limits_n A_n = \Sigma A_n$ for every ascending sequence $\{A_n\}$, and $\lim\limits_n A_n = \Pi A_n$ for every descending sequence $\{A_n\}$.

Finally, given a class $\mathfrak{C}$ of sets, we shall often call the sets belonging to $\mathfrak{C}$, for short, sets ($\mathfrak{C}$). The class of the sets which are the sums of sequences of sets ($\mathfrak{C}$) will be denoted by $\mathfrak{C}_\sigma$. The class of the sets which are the products of such sequences will be denoted by $\mathfrak{C}_\delta$ (see F. Hausdorff [II, p. 83]).

§ 3. Abstract space X. In the rest of this chapter, a set X will be fixed and called *space*. The elements of X will be called *points*. If A is any set contained in X the set $X - A$ will be called *complement of A with respect to X*; the expression "with respect to X" will, however, generally be omitted, since sets outside the space X will not be considered. The complement of a set A will be denoted by CA. We evidently have, for every pair of sets A and B,

$$(3.1) \qquad A - B = A \cdot CB,$$

and for every sequence $\{X_n\}$ of sets

$$(3.2) \qquad \underset{n}{\Pi} X_n = C\underset{n}{\Sigma}CX_n, \qquad \underset{n}{\Sigma} X_n = C\underset{n}{\Pi}CX_n,$$

$$\underset{n}{\lim\sup} X_n = C\underset{n}{\lim\inf}CX_n, \qquad \underset{n}{\lim\inf} X_n = C\underset{n}{\lim\sup}CX_n.$$

In the space X we shall consider functions of a set, and functions of a point. The values of these functions will always be real numbers, finite or infinite. A function will be called *finite*, when it assumes only finite values.

To avoid misunderstanding, let us agree that when infinite functions are subjected to the elementary operations of addition, subtraction etc., we make the following conventions: $a+(\pm\infty)=(\pm\infty)+a=\pm\infty$ for $a \neq \mp\infty$; $(+\infty)+(-\infty)=(-\infty)+(+\infty)=(\pm\infty)-(\pm\infty)=0$; $a\cdot(\pm\infty)=(\pm\infty)\cdot a=\pm\infty$ and $a\cdot(\pm\infty)=(\pm\infty)\cdot a=\mp\infty$, according as $a>0$ or $a<0$; $0\cdot(\pm\infty)=(\pm\infty)\cdot 0=0$; $a/(\pm\infty)=0$; $a/0=+\infty$.

We call *characteristic function* $c_E(x)$ of a set E, the function (of a point) equal to 1 at all points of the set, and to 0 everywhere else. The following theorem is obvious:

(3.3) *If $E=\underset{n}{\Sigma} E_n$, and $E_i\cdot E_k=0$ whenever $i \neq k$, then $c_E(x)=\underset{n}{\Sigma} c_{E_n}(x)$.*

If $\{E_n\}$ is a monotone sequence of sets, the sequence of their characteristic functions is also monotone, non-decreasing or non-increasing according as the sequence $\{E_n\}$ is ascending or descending.

If $\{E_n\}$ is any sequence of sets, A and B denoting its upper and lower limits respectively, we have

$$c_A(x)=\lim_{n}\sup c_{E_n}(x), \quad and \quad c_B(x)=\lim_{n}\inf c_{E_n}(x);$$

so that, in order that a sequence of sets $\{E_n\}$ converge to a set E, it is necessary and sufficient that the sequence of their characteristic functions $\{c_{E_n}(x)\}$ converge to the function $\{c_E(x)\}$.

A function assuming only a finite number of different values on a set E is called a *simple function* on E. If $v_1, v_2,..., v_n$ are all the distinct values of a simple function $f(x)$ on a set E, the function $f(x)$ may on E be written in the form

$$f(x)=\sum_{k=1}^{n} v_k c_{E_k}(x) \quad where \quad E=E_1+E_2+...+E_k \quad and \quad E_i\cdot E_j=0 \quad for \quad i \neq j.$$

The function f given by this formula over the set E will be denoted by $\{v_1, E_1; v_2, E_2; ...; v_n, E_n\}$.

The notion of characteristic function is due to Ch. J. de la Vallée Poussin [1] and [I, p. 7].

§ 4. Additive classes of sets. A class $\mathfrak{X}$ of sets in the space X will be called *additive* if (i) the empty set belongs to $\mathfrak{X}$, (ii) when a set X belongs to $\mathfrak{X}$ so does its complement CX, and (iii) the sum of a sequence $\{X_n\}$ of sets selected from the class $\mathfrak{X}$, belongs also to the class $\mathfrak{X}$.

The classes of sets, additive according to this definition, are sometimes termed *completely* additive. We get the definition of a class of sets additive *in the weak sense* if we replace the condition (iii) of the preceding definition by the following: (iii-bis) *the sum of two sets belonging to $\mathfrak{X}$ also belongs to $\mathfrak{X}$.*

The sets of an additive class $\mathfrak{X}$ will be called *sets measurable $(\mathfrak{X})$*, or, in accordance with the definition given in § 2 (p. 5), simply *sets $(\mathfrak{X})$*. We see at once that, on account of the conditions (i) and (ii), the space X, as complement of the empty set, belongs to every additive class of sets. Making use of the relations (2.1), (3.1), and (3.2), we obtain immediately the following:

(4.1) ***Theorem.*** *If $\mathfrak{X}$ is an additive class of sets, the sum, the product, and the two limits, upper and lower, of every sequence of sets measurable $(\mathfrak{X})$, and the difference of two sets measurable $(\mathfrak{X})$, are also measurable $(\mathfrak{X})$.*

In later chapters we shall consider certain additive classes of sets that present themselves naturally to us, in connection with the theory of measure, in metrical or in Euclidean spaces. In the abstract space X, about which we have made practically no hypothesis, we can only mention a few trivial examples of additive classes of sets, such as the class of all sets in X, or the class of all finite or enumerable sets and their complements. Let us still mention one further general theorem:

(4.2) ***Theorem.*** *Given any class $\mathfrak{M}$ of sets in X, there exists always a smallest additive class of sets containing $\mathfrak{M}$, i. e. an additive class $\mathfrak{N}_0 \supset \mathfrak{M}$ contained in every other additive class that contains $\mathfrak{M}$.*

For let $\mathfrak{N}_0$ be the product of all the additive classes that contain $\mathfrak{M}$. Such classes evidently exist, one such class being the class of all sets in X. We see at once that the class $\mathfrak{N}_0$ thus defined has the required properties.

§ 5. Additive functions of a set.

In the rest of this chapter we suppose that *a definite additive class $\mathfrak{X}$ of sets is fixed in the space X.* In accordance with this hypothesis, we may often omit the symbol $\mathfrak{X}$ in our statements, without causing any ambiguity.

A function of a set, $\Phi(X)$, will be called *additive function of a set* $(\mathfrak{X})$ *on a set E*, if (i) E is a set $(\mathfrak{X})$, (ii) the function $\Phi(X)$ is defined and finite for each set $X \subset E$ measurable $(\mathfrak{X})$, and if (iii) $\Phi(\sum_n X_n) = \sum_n \Phi(X_n)$ for every sequence $\{X_n\}$ of sets $(\mathfrak{X})$ contained in E and such that $X_i \cdot X_k = 0$ whenever $i \neq k$. For simplicity, we shall speak of an "additive function" instead of an "additive function of a set $(\mathfrak{X})$" whenever there is no mistaking the meaning. An additive function of a set $(\mathfrak{X})$ will be called *monotone* on E if its values for the subsets $(\mathfrak{X})$ of E are of constant sign. A non-negative function $\Phi(\mathfrak{X})$ additive and monotone, will also be termed *non-decreasing*, on account of the fact that, for each pair of sets A and B measurable $(\mathfrak{X})$, the inequality $A \subset B$ implies $\Phi(B) = \Phi(A) + \Phi(B - A) \geqslant \Phi(A)$. For the same reason, non-positive monotone functions will be termed *non-increasing*.

(5.1) ***Theorem.*** *If $\Phi(x)$ is an additive function on a set E, then*

$$(5.2) \qquad \Phi(\lim_n X_n) = \lim_n \Phi(X_n)$$

for every monotone sequence $\{X_n\}$ of sets $(\mathfrak{X})$ contained in E. If $\Phi(X)$ is a non-negative monotone function, then

$$(5.3) \quad \Phi(\liminf_n X_n) \leqslant \liminf_n \Phi(X_n) \;\; and \;\; \Phi(\limsup_n X_n) \geqslant \limsup_n \Phi(X_n)$$

for every sequence $\{X_n\}$ of sets $(\mathfrak{X})$ in E.

Proof. Let $\{X_n\}_{n=1,2,\dots}$ be a sequence of sets $(\mathfrak{X})$ contained in E. If $\{X_n\}$ is an ascending monotone sequence, then

$$\lim_n X_n = \sum_{n=1}^{\infty} X_n = X_1 + \sum_{n=1}^{\infty} (X_{n+1} - X_n),$$

and consequently, $\Phi(X)$ being an additive function on E,

$$\Phi(\lim_n X_n) = \Phi(X_1) + \sum_{n=1}^{\infty} \Phi(X_{n+1} - X_n) =$$
$$= \lim_n [\Phi(X_1) + \sum_{k=1}^{n-1} \Phi(X_{k+1} - X_k)] = \lim_n \Phi(X_n).$$

If $\{X_n\}$ is a descending sequence, the sequence $\{E - X_n\}$ is ascending, and, by the result already proved,

$$\Phi(E) - \Phi(\lim_n X_n) = \Phi[\lim_n (E - X_n)] = \lim_n \Phi(E - X_n) = \Phi(E) - \lim_n \Phi(X_n),$$

from which (5.2) follows at once.

Finally, if $\{X_n\}$ is any sequence, but $\Phi(X)$ is a non-negative monotone function, we put

$$(5.4) \qquad Y_n = \prod_{k=n}^{\infty} X_k \quad for \quad n = 1, 2, \dots$$

The sets Y_n are measurable $(\mathfrak{X})$ on account of (4.1), and form an ascending sequence. We therefore have, by the part of our theorem proved already,

$$(5.5) \qquad \Phi(\lim_n Y_n) = \lim_n \Phi(Y_n).$$

Now, it follows from (5.4) that $Y_n \subset X_n$, and so, $\Phi(Y_n) \leqslant \Phi(X_n)$, for each n. On the other hand, $\liminf_n X_n = \lim_n Y_n$, and therefore the first of the relations (5.3) is an immediate consequence of (5.5). We establish similarly (or, if preferred, by changing X_n to $E - X_n$) the second of these relations, and this completes the proof of the theorem.

Every function of a set $\Phi(X)$, additive on a set E, can easily be extended to the whole space X. In fact, if we write, for instance, $\Phi_1(X) = \Phi(X \cdot E)$ for every set X measurable $(\mathfrak{X})$, we see at once that $\Phi_1(X)$ is a function additive on the whole space X, that coincides with $\Phi(X)$ for measurable subsets of E and vanishes for measurable sets containing no points of E. We shall call the function $\Phi_1(X)$, thus defined, the *extension* of $\Phi(X)$ from the set E to the space X.

§ 6. The variations of an additive function.

The upper and lower bounds of the values that a function of a set $\Phi(X)$, additive on a set E assumes for the measurable subsets of this set E, will be called *upper variation* and *lower variation* of the function Φ over E, and denoted by $\overline{W}(\Phi; E)$ and $\underline{W}(\Phi; E)$ respectively. Since every additive function vanishes for the empty set, we evidently have $\underline{W}(\Phi; E) \leqslant 0 \leqslant \overline{W}(\Phi; E)$. The number $\overline{W}(\Phi; E) + |\underline{W}(\Phi; E)|$ will be called *absolute variation* of the function Φ on E and denoted by $W(\Phi; E)$.

(6.1) **Theorem.** *If $\Phi(X)$ is an additive function on a set E, its variations over E are always finite.*

Proof. Suppose that $W(\Phi; E) = +\infty$. We shall show firstly that there then exists a sequence $\{E_n\}_{n=1,2,...}$ of sets $(\mathfrak{X})$ such that

(6.2) $E_n \subset E_{n-1}$ for $n > 1$; $W(\Phi; E_n) = \infty$; $|\Phi(E_n)| \geqslant n - 1$.

For let us choose $E_1 = E$ and suppose the sets E_n for $n = 1, 2, ..., k$ defined so as to satisfy the conditions (6.2). By the second of these conditions with $n = k$, there exists a measurable set $A \subset E_k$ such that

(6.3) $|\Phi(A)| \geqslant |\Phi(E_k)| + k$.

If $W(\Phi; A) = \infty$, we have only to choose $E_{k+1} = A$ in order to satisfy the conditions (6.2) for $n = k + 1$. If, on the other hand, $W(\Phi; A)$ is finite, we must have $W(\Phi; E_k - A) = +\infty$, and, by (6.3), $|\Phi(E_k - A)| \geqslant |\Phi(A)| - |\Phi(E_k)| \geqslant k$, so that the conditions (6.2) will be satisfied for $n = k + 1$, if we choose $E_{k+1} = E_k - A$. The sequence $\{E_n\}$ is thus obtained by induction.

Now, on account of Theorem 5.1 and of the third of the conditions (6.2), we should have the equality $\Phi(\lim_n E_n) = \lim_n \Phi(E_n) = \infty$, and since every additive function of a set is, by definition, finite, this is evidently impossible. Q. E. D.

It follows from the theorem just proved that every function $\Phi(X)$ additive on a set E is not only finite for the subsets $(\mathfrak{X})$ of E, but also *bounded*; in fact, the values it assumes are bounded in modulus by the finite number $W(\Phi; E)$.

Theorem 6.1 can be further completed as follows:

(6.4) **Theorem.** *For every function $F(X)$ additive on a set E, the variations $\overline{W}(\Phi; X)$, $\underline{W}(\Phi; X)$ and $W(\Phi; X)$ are also additive functions of a set $(\mathfrak{X})$ on E, and we have, for every measurable set $X \subset E$*

(6.5) $$\Phi(X) = \overline{W}(\Phi; X) + \underline{W}(\Phi; X).$$

Proof. To fix the ideas, consider the function $\Omega_1(X) = \overline{W}(\Phi; X)$. Since this function is finite by Theorem 6.1, we have to show that for every sequence $\{X_n\}$ of measurable sets contained in E, and such that $X_i \cdot X_k = 0$ whenever $i \neq k$,

(6.6) $$\Omega_1(\underset{n}{\Sigma} X_n) = \underset{n}{\Sigma} \Omega_1(X_n).$$

For this purpose, let us observe that for every measurable set $X \subset \underset{n}{\Sigma} X_n$ we have $\Phi(X) = \underset{n}{\Sigma} \Phi(X \cdot X_n) \leqslant \underset{n}{\Sigma} \Omega_1(X_n)$,   and hence

(6.7) $$\Omega_1(\underset{n}{\Sigma} X_n) \leqslant \underset{n}{\Sigma} \Omega_1(X_n).$$

On the other hand, denoting generally by Y_n any measurable set variable in X_n, we have $\Omega_1(\underset{n}{\Sigma} X_n) \geqslant \Phi(\underset{n}{\Sigma} Y_n) = \underset{n}{\Sigma} \Phi(Y_n)$,   and therefore also $\Omega_1(\underset{n}{\Sigma} X_n) \geqslant \underset{n}{\Sigma} \Omega_1(X_n)$. Combining this with (6.7) we get the equality (6.6).

Finally, to establish (6.5), we observe that for every measurable subset Y of X we have $\Phi(Y) = \Phi(X) - \Phi(X - Y) \leqslant \Phi(X) - \underline{W}(\Phi; X)$, and so $\overline{W}(\Phi; X) \leqslant \Phi(X) - \underline{W}(\Phi; X)$. Similarly $\underline{W}(\Phi; X) \geqslant \Phi(X) - \overline{W}(\Phi; X)$. These two inequalities give together the equality (6.5), and the proof of Theorem 6.4 is complete.

It follows from this theorem that every function of a set $\Phi(X)$ additive on a set E is, on E, the difference of two non-negative additive functions. The formula (6.5) expresses, in fact, $\Phi(X)$ as the sum of two variations of $\Phi(X)$, of which the one is non-negative and the other non-positive; this particular decomposition of an additive function of a set will be termed the *Jordan decomposition*.

We can now complete Theorem 5.1 as follows:

(6.8) *Theorem. If* $\dot{\Phi}(X)$ *is additive on a set* E, *we have* $\Phi(\lim_n X_n) = \lim_n \Phi(X_n)$ *for every convergent sequence* $\{X_n\}$ *of sets* $(\mathfrak{X})$ *contained in* E.

In fact, making use of the Jordan decomposition, we may restrict ourselves to non-negative functions $\Phi(X)$, and for these Theorem 6.8 follows at once from the second part of Theorem 5.1.

§ 7. Measurable functions. Given an arbitrary condition, or property, (V) of a point x, let us denote generally by $\mathop{\mathrm{E}}\limits_{x}[(V)]$ the set of all the points x of the space considered that fulfill this condition, or have this property. Thus, for instance, if $f(x)$ denotes a function of a point defined on a set E and a is a real number, the symbol

(7.1) $$\mathop{\mathrm{E}}\limits_{x}[x \,\epsilon\, E; f(x) > a]$$

denotes the set of the points x of E at which $f(x) > a$.

A function of a point, $f(x)$, defined on a set E, will be termed *measurable* $(\mathfrak{X})$, or simply *function* $(\mathfrak{X})$, if the set E, and the set (7.1) for each finite a, are measurable $(\mathfrak{X})$. It is easy to see that

(7.2) *In order that a function $f(x)$ be measurable on a measurable set E, it suffices that the set (7.1) should be so for all values of a belonging to an arbitrary everywhere dense set R of real numbers (the same holds with the set (7.1) replaced by the set* $\mathop{\mathrm{E}}\limits_{x}[x\,\epsilon\,E; f(x) \geqslant a]$).

In fact, for every real a, the set R contains a decreasing sequence of numbers $\{r_n\}$ converging to a. We therefore have $\mathop{\mathrm{E}}\limits_{x}[x\,\epsilon\,E; f(x) > a] = \sum\limits_{n=1}^{\infty} \mathop{\mathrm{E}}\limits_{x}[x\,\epsilon\,E; f(x) > r_n]$ and, each term of the sum on the right being measurable by hypothesis, the same holds for the sum itself (cf. Theorem 4.1).

Every function $f(x)$ measurable on a set E, can be continued in various ways, so as to become a measurable function on the whole space X. For definiteness, we shall understand by the *extension* of the function $f(x)$ from the set E to the space X, the function $f_0(x)$ equal to $f(x)$ on E and to 0 everywhere else. For brevity, we shall often deal only with functions measurable on the whole

space X, but it is easy to see that all the theorems and the reasonings of this, and of the succeeding, § could be taken relative to an arbitrary set $(\mathfrak{X})$.

The equations

$$\mathop{\mathrm{E}}\limits_{x}[f(x) \leqslant a] = \mathop{\mathrm{C}} \mathop{\mathrm{E}}\limits_{x}[f(x) > a], \qquad \mathop{\mathrm{E}}\limits_{x}[f(x) \geqslant a] = \prod_{n=1}^{\infty} \mathop{\mathrm{E}}\limits_{x}\left[f(x) > a - \frac{1}{n}\right],$$

$$\mathop{\mathrm{E}}\limits_{x}[f(x) < a] = \mathop{\mathrm{C}} \mathop{\mathrm{E}}\limits_{x}[f(x) \geqslant a], \qquad \mathop{\mathrm{E}}\limits_{x}[f(x) = a] = \mathop{\mathrm{E}}\limits_{x}[f(x) \geqslant a] \cdot \mathop{\mathrm{E}}\limits_{x}[f(x) \leqslant a],$$

$$\mathop{\mathrm{E}}\limits_{x}[f(x) < +\infty] = \sum_{n=1}^{\infty} \mathop{\mathrm{E}}\limits_{x}[f(x) < n], \qquad \mathop{\mathrm{E}}\limits_{x}[f(x) > -\infty] = \sum_{n=1}^{\infty} \mathop{\mathrm{E}}\limits_{x}[f(x) > -n],$$

$$\mathop{\mathrm{E}}\limits_{x}[f(x) = +\infty] = \mathop{\mathrm{C}}\mathop{\mathrm{E}}\limits_{x}[f(x) < +\infty], \qquad \mathop{\mathrm{E}}\limits_{x}[f(x) = -\infty] = \mathop{\mathrm{C}} \mathop{\mathrm{E}}\limits_{x}[f(x) > -\infty]$$

show that for every measurable function $f(x)$ and for every number a, the left hand sides are measurable sets. Conversely, in the definition of measurable function, we may replace the set (7.1) by any one of the sets $\mathop{\mathrm{E}}\limits_{x}[f(x) \geqslant a]$, $\mathop{\mathrm{E}}\limits_{x}[f(x) \leqslant a]$ or $\mathop{\mathrm{E}}\limits_{x}[f(x) < a]$; this follows at once from the identity

$$\mathop{\mathrm{E}}\limits_{x}[f(x) > a] = \sum_{n=1}^{\infty} \mathop{\mathrm{E}}\limits_{x}\left[f(x) \geqslant a + \frac{1}{n}\right] = \mathop{\mathrm{C}} \mathop{\mathrm{E}}\limits_{x}[f(x) \leqslant a] = \mathop{\mathrm{C}} \prod_{n=1}^{\infty} \mathop{\mathrm{E}}\limits_{x}\left[f(x) < a + \frac{1}{n}\right].$$

To any function $f(x)$ on a set E, we attach two functions $\overset{\circ}{f}(x)$ and $f(x)$ on E, called, respectively, the *non-negative part* and the *non-positive part* of $f(x)$ and defined as follows:

$$\overset{\circ}{f}(x) = f(x) \quad or \quad 0 \quad according \ as \quad f(x) \geqslant 0 \quad or \quad f(x) < 0,$$

$$\underset{\circ}{f}(x) = f(x) \quad or \quad 0 \quad according \ as \quad f(x) < 0 \quad or \quad f(x) \geqslant 0.$$

We see at once, *that in order that a function be measurable on a set E, it is necessary and sufficient that its two parts, the non-negative and the non-positive, be measurable.*

Returning now to the notions of characteristic function, and simple function introduced in § 3, we have the theorem:

(7.3) *Theorem. In order that a set E be measurable $(\mathfrak{X})$, it is necessary and sufficient that its characteristic function be measurable. More generally, in order that, on a set E, a simple function $f(x)$ be measurable $(\mathfrak{X})$, it is necessary and sufficient that, for each value of $f(x)$, the points at which this value is assumed on E, should constitute a measurable subset of E.*

Another theorem, of great utility in applications, is the following:

(7.4) **Theorem.** *Every function $f(x)$ that is measurable $(\mathfrak{X})$ and non-negative on a set E, is the limit of a non-decreasing sequence of simple functions, finite, measurable and non-negative on E.*

In fact, if we write for each positive integer n and for $x \in E$,

$$f_n(x) = \begin{cases} \dfrac{i-1}{2^n}, & if \quad \dfrac{i-1}{2^n} \leqslant f(x) < \dfrac{i}{2^n}, \quad 1 \leqslant i \leqslant 2^n \cdot n, \\ n, & if \quad f(x) \geqslant n, \end{cases}$$

the functions $f_n(x)$ thus defined are evidently simple and non-negative, and, on account of Theorem 7.3, measurable on E. Further, as is easily seen, the sequence $\{f_n(x)\}$ is non-decreasing. Finally $\lim_n f_n(x) = f(x)$ for every $x \in E$; for, if $f(x) < +\infty$, we have, as soon as n exceeds the value of $f(x)$, the inequalities $0 \leqslant f(x) - f_n(x) \leqslant 1/2^n$, while, if $f(x) = +\infty$, we have $f_n(x) = n$ for $n = 1, 2, \ldots$, and so $\lim_n f_n(x) = +\infty = f(x)$.

§ 8. Elementary operations on measurable functions.

We shall now show that elementary operations effected on measurable functions always lead to measurable functions.

(8.1) **Theorem.** *Given two measurable functions $f(x)$ and $g(x)$, the sets*

$$\underset{x}{E}\,[f(x) > g(x)], \quad \underset{x}{E}\,[f(x) \geqslant g(x)] \quad and \quad \underset{x}{E}\,[f(x) = g(x)],$$

are measurable.

The proof follows at once from the identities

$$\underset{x}{E}\,[f(x) > g(x)] = \sum_{n=-\infty}^{+\infty} \sum_{m=1}^{+\infty} \underset{x}{E}\left[f(x) > \frac{n}{m}\right] \cdot \underset{x}{E}\left[g(x) < \frac{n}{m}\right],$$

$$\underset{x}{E}\,[f \geqslant g] = C\,\underset{x}{E}\,[g > f] \quad and \quad \underset{x}{E}\,[f = g] = \underset{x}{E}\,[f \geqslant g] \cdot \underset{x}{E}\,[g \geqslant f].$$

(8.2) **Theorem.** *If the function $f(x)$ is measurable, $|f(x)|^a$ is also a measurable function.*

For $a > 0$, the proof is a consequence of the identity

$$\underset{x}{E}\,[|f(x)|^a > a] = \underset{x}{E}\,[f(x) > a^{1/a}] + \underset{x}{E}\,[f(x) < -a^{1/a}],$$

which is valid for every $a \geqslant 0$, while for $a < 0$ its left hand side coincides with the whole space and therefore constitutes a measurable set. For $a < 0$, the proof is similar.

(8.3) **Theorem.** *Every linear combination of measurable functions with constant coefficients represents a measurable function.*

The identities

$$\underset{x}{E}\,[a \cdot f(x) + \beta > a] = \underset{x}{E}\left[f(x) > \frac{a - \beta}{a}\right] \qquad for \quad a > 0,$$

$$\underset{x}{E}\,[a \cdot f(x) + \beta > a] = \underset{x}{E}\left[f(x) < \frac{a - \beta}{a}\right] \qquad for \quad a < 0,$$

valid for every function $f(x)$ and for all numbers a, $a \neq 0$, and β, show, in the first place, that $a \cdot f(x) + \beta$ is a measurable function, if $f(x)$ is measurable. It follows further, from Theorem 8.1 and from the identities:

$$\underset{x}{E}\,[a \cdot f + \beta \cdot g > a] = \underset{x}{E}\left[f > -\frac{\beta}{a} g + \frac{a}{a}\right] \qquad for \quad a > 0$$

$$\underset{x}{E}\,[a \cdot f + \beta \cdot g > a] = \underset{x}{E}\left[f < -\frac{\beta}{a} g + \frac{a}{a}\right] \qquad for \quad a < 0$$

that if $f(x)$ and $g(x)$ are measurable functions, so is $a \cdot f(x) + \beta \cdot g(x)$.

(8.4) **Theorem.** *The product of two measurable functions $f(x)$ and $g(x)$ is a measurable function.*

Measurability of the product $f \cdot g$ is derived by applying Theorems 8.2 and 8.3 to the identity $fg = \frac{1}{4}[(f+g)^2 - (f-g)^2]$, the completion of the proof, by taking into account possible infinities of f and g, being trivial.

(8.5) **Theorem.** *Given a sequence of measurable functions $\{f_n(x)\}$, the functions*

$$\text{upper bound}_n f_n(x), \quad \text{lower bound}_n f_n(x), \quad \limsup_n f_n(x) \quad and \quad \liminf_n f_n(x)$$

are also measurable.

The measurability of $h(x) = \text{upper bound}_n f_n(x)$ follows from the identity $\underset{x}{E}[h(x) > a] = \sum_n \underset{x}{E}[f_n(x) > a]$. For the lower bound, the corresponding proof is derived by change of sign.

Hence, the functions $h_n(x) = \text{upper bound}\,[f_{n+1}(x), f_{n+2}(x), \ldots]$ are measurable, and the same is therefore true of the function $\limsup_n f_n(x) = \lim_n h_n(x) = \text{lower bound}_n h_n(x)$. By changing the sign of $f_n(x)$, we prove the same for lim inf.

§ 9. Measure. A function of a set $\mu(X)$ will be called a *measure* ($\mathfrak{X}$), if it is defined and non-negative for every set ($\mathfrak{X}$), and if

$$\mu\left(\underset{n}{\Sigma} X_n\right) = \underset{n}{\Sigma}\,\mu(X_n)$$

for every sequence $\{X_n\}$ of sets ($\mathfrak{X}$) no two of which have points in common. The number $\mu(X)$ is then termed, for every set X measurable ($\mathfrak{X}$), the *measure* (μ) of X. If every point of a set E, except at most the points belonging to a subset of E of measure (μ) zero, possesses a certain property V, we shall say that the condition V is satisfied *almost everywhere* (μ) in E, or, that *almost every* (μ) point of E has the property V. We shall suppose, in the sequel of this Chapter, that, just as the class $\mathfrak{X}$ was chosen once for all, a measure μ corresponding to this class is also kept fixed. Accordingly, we shall often omit the symbol (μ) in the expressions "measure (μ)", "almost everywhere (μ)", etc. Clearly $\mu(X) \leqslant \mu(Y)$ for any pair of sets X and Y measurable ($\mathfrak{X}$) such that $X \subset Y$, and $\mu\left(\underset{n}{\Sigma} X_n\right) \leqslant \underset{n}{\Sigma}\,\mu(X_n)$ for every sequence of measurable sets $\{X_n\}$.

A measure may also assume infinite values, and is therefore not in general an additive function according to the definition of § 5.

The results established in this chapter concerning perfectly arbitrary measures will be interpreted in the sequel for more special theories of measure, (for instance, those of Lebesgue and Carathéodory). For the present, we shall be content mentioning a few examples.

Let us take for $\mathfrak{X}$, the class of all sets in a space X. We obtain a trivial example of measure ($\mathfrak{X}$) by writing $\mu(X) = 0$ identically, (or else $\mu(X) = +\infty$) for every set $X \subset X$. Another example consists in choosing an element a in X and writing $\mu(X) = 1$ or $\mu(X) = 0$, according as $a \in X$ or not. In the case of an enumerable space X, consisting of elements $a_1, a_2, ..., a_n, ...$, the general form of a measure $\mu(X)$ defined for all subsets X of X is $\mu(X) = \underset{n}{\Sigma} k_n f_n(X)$ where $\{k_n\}$ is a sequence of non-negative real numbers and $f_n(X)$ is equal to 1 or 0 according as $a_n \in X$ or not. It follows that every measure defined for all subsets of an enumerable space, and vanishing for the sets that consist of a single point, vanishes identically. The similar problem for spaces of higher potencies is much more difficult (see S. Ulam [1]). For a space of the potency of the continuum see also S. Banach and C. Kuratowski [1], E. Szpilrajn [1], W. Sierpiński [I, p. 60], W. Sierpiński and E. Szpilrajn [1].

We shall now prove the following theorem analogous to Theorem 5.1:

(9.1) **Theorem.** *If* $\{X_n\}$ *is a monotone ascending sequence of measurable sets, then* $\underset{n}{\lim}\,\mu(X_n) = \mu(\underset{n}{\lim} X_n)$. *The same holds for monotone descending sequences provided, however, that* $\mu(X_1) \nmid \infty$.

More generally, for every sequence $\{X_n\}$ *of measurable sets,*

$$(9.2) \qquad \mu(\underset{n}{\lim\inf} X_n) \leqslant \underset{n}{\lim\inf}\,\mu(X_n)$$

and, if further $\mu(\underset{n}{\Sigma} X_n) \neq \infty$,

$$(9.3) \qquad \mu(\underset{n}{\lim\sup} X_n) \geqslant \underset{n}{\lim\sup}\,\mu(X_n),$$

so that, in particular, if the sequence $\{X_n\}$ *converges and its sum has finite measure,* $\underset{n}{\lim}\,\mu(X_n) = \mu(\underset{n}{\lim} X_n)$.

Proof. For an ascending sequence $\{X_n\}_{n=1,2,...}$ the equation $\underset{n}{\lim}\,\mu(X_n) = \mu(\underset{n}{\lim} X_n)$ follows at once from the relation

$$\underset{n}{\lim} X_n = \overset{\infty}{\underset{n=1}{\Sigma}} X_n = X_1 + \overset{\infty}{\underset{n=1}{\Sigma}}(X_{n+1} - X_n),$$

and if the sequence $\{X_n\}$ is descending and $\mu(X_1) \neq \infty$, then the measure $\mu(X)$ is an additive function on the set X_1 and consequently the required result follows from Theorem 5.1.

In exactly the same way, if for an arbitrary sequence $\{X_n\}$ of measurable sets, $\underset{n}{\Sigma} X_n$ is of finite measure, the measure $\mu(X)$ is an additive function on this set, and the two inequalities (9.2) and (9.3) follow from Theorem 5.1. To establish the first of these inequalities without assuming that the sum of the sets X_n has finite measure, we write as in the proof of Theorem 5.1

$$Y_n = \overset{\infty}{\underset{k=n}{\prod}} X_k.$$

Since the sequence is ascending, and $Y_n \subset X_n$ for every n, we have $\mu(\underset{n}{\lim\inf} X_n) = \mu(\underset{n}{\lim} Y_n) = \underset{n}{\lim}\,\mu(Y_n) \leqslant \underset{n}{\lim\inf}\,\mu(X_n)$.

We conclude this § with an important theorem due to D. Egoroff, concerning sequences of measurable functions (cf. D. Egoroff [1], and also W. Sierpiński [3], F. Riesz [2; 3], H. Hahn [I, pp. 556—8]). We shall first prove the following lemma:

(9.4) **Lemma.** *If* E *is a measurable set of finite measure* (μ) *and if* $\{f_n(x)\}$ *is a sequence of finite measurable functions on* E, *converging on this set to a finite measurable function* $f(x)$, *there exists, for each pair of positive numbers* ε, η, *a positive integer* N *and a measurable subset* H *of* E *such that* $\mu(H) < \eta$ *and*

$$(9.5) \qquad |f_n(x) - f(x)| < \varepsilon$$

for every $n > N$ *and every* $x \in E - H$.

Proof. Let us denote generally by E_m the subset of E consisting of the points x for which (9.5) holds whenever $n > m$. Thus defined, the sets E_m are measurable and form a monotone ascending sequence, since for each integer m, we have

$$E_m = \prod_{n=m+1}^{\infty} E[x \in E; |f(x) - f_n(x)| < \varepsilon].$$

Further, since $\{f_n(x)\}$ converges to $f(x)$ on the whole of E, we have $E = \sum_m E_m$, and so, by Theorem 9.1, $\mu(E) = \lim_m \mu(E_m)$, i. e. $\lim_m \mu(E - E_m) = 0$, and therefore, from a sufficiently large m_0 onwards, $\mu(E - E_m) < \eta$. We have now only to choose $N = m_0$ and $H = E - E_{m_0}$, and the lemma is proved.

(9.6)   *Egoroff's Theorem. If E is a measurable set of finite measure (μ) and if $\{f_n(x)\}$ is a sequence of measurable functions finite almost everywhere on E, that converges almost everywhere on this set to a finite measurable function $f(x)$, then there exists, for each $\varepsilon > 0$, a subset Q of E such that $\mu(E - Q) < \varepsilon$ and such that the convergence of $\{f_n(x)\}$ to $f(x)$ is uniform on Q.*

Proof. By removing from E, if necessary, a set of measure (μ) zero, we may suppose that on E, the functions $f_n(x)$ are everywhere finite, and converge everywhere to $f(x)$. By the preceding lemma, we can associate with each integer $m > 0$ a set $H_m \subset E$ such that $\mu(H_m) < \varepsilon/2^m$ and an index N_m such that

(9.7) $|f_n(x) - f(x)| < 1/2^m$ for $n > N_m$ and for $x \in E - H_m$.

Let us write $Q = E - \sum_{m=1}^{\infty} H_m$. We find

$$\mu(E - Q) \leqslant \sum_{m=1}^{\infty} \mu(H_m) \leqslant \sum_{m=1}^{\infty} \varepsilon/2^m = \varepsilon,$$

and since the sequence $f_n(x)$ converges uniformly to $f(x)$ on the set Q on account of (9.7), the theorem is proved.

The theorem of Egoroff can be given another form (cf. N. Lusin [I, p. 20]), and, at the same time, the hypothesis concerning finite measure of E can be slightly relaxed.

(9.8)   *If E is the sum of a sequence of measurable sets of finite measure (μ) and if $\{f_n(x)\}$ is a sequence of measurable functions finite almost everywhere on this set, converging almost everywhere on E to a finite function, then the set E can be expressed as the sum of a sequence of measurable sets $H, E_1, E_2, \ldots$ such that $\mu(H) = 0$ and that the sequence $\{f_n(x)\}$ converges uniformly on each of the sets E_n.*

For the proof, it suffices to take the case in which the set E is itself of finite measure. With this hypothesis, we can, on account of Theorem 9.6, define by induction a sequence $\{E_k\}_{k=1,2,\ldots}$ of measurable sets such that $\mu(E - \sum_{k=1}^{n} E_k) \leqslant 1/n$, and that the sequence $\{f_n(x)\}$ converges uniformly on the set E_k for each k. Choosing $H = E - \sum_{k=1}^{\infty} E_k$, we have $\mu(H) = 0$, and the theorem is proved.

As we may observe, the hypothesis that the set E is the sum of a sequence of sets of finite measure, is essential for the validity of Theorem 9.8. For this purpose, let us take as a space X_0 the interval $[0, 1]$, and as an additive class $\mathfrak{X}_0$ of sets, that of all subsets of X_0. Further, let us define a measure μ_0 by writing $\mu_0(X) = \infty$ whenever the set $X \in \mathfrak{X}_0$ is infinite and $\mu_0(X) = n$, if X is a finite set and n denotes the number of its elements. The sets of measure (μ_0) zero then coincide with the empty set. Finally, let $\{g_n(x)\}$ be an arbitrary sequence of functions, continuous on the interval $[0, 1]$, converging everywhere on this interval, but not uniformly on any subinterval of $[0, 1]$.

To justify our remark concerning Theorem 9.8, it suffices to show that the interval $X_0 = [0, 1]$ is not representable as the sum of a sequence $\{E_n\}$ of sets such that the sequence of functions $\{g_n(x)\}$ converges uniformly on each of them. But if such a decomposition were to exist, we might suppose firstly — since the functions $g_n(x)$ are continuous — all the sets E_i closed. Then, however, by the theorem of Baire (cf. Chap. II, § 9) one of them at least would contain a subinterval of $[0, 1]$. This gives a contradiction, since by hypothesis, the sequence $\{g_n(x)\}$ does not converge uniformly on any interval whatsoever.

§ 10. Integral. If we are given in the space X an additive class of sets $\mathfrak{X}$ and a measure μ defined for the sets of this class, we can attach to them a process of integration for functions of a point. In fact:

(i) If $f(x)$ is a function $(\mathfrak{X})$ non-negative on a set E, we shall understand by the *definite integral* $(\mathfrak{X}, \mu)$ of $f(x)$ over E the upper bound of the sums

$$\sum_{k=1}^{n} v_k \mu(E_k),$$

where $\{E_k\}_{k=1,2,\ldots,n}$ is an arbitrary finite sequence of sets $(\mathfrak{X})$ such that $E = E_1 + E_2 + \ldots + E_n$ and $E_i \cdot E_k = 0$ for $i \neq k$, and where v_k, for $k = 1, 2, \ldots, n$, denotes the lower bound of $f(x)$ on E_k.

(ii) If $f(x)$ is an arbitrary function measurable ($\mathfrak{X}$) on a set E, we shall say that $f(x)$ *possesses a definite integral* ($\mathfrak{X}, \mu$) *over E*, if one at least of the non-negative functions $\overset{0}{f}(x)$ and $-\overset{0}{f}(x)$ (cf. § 3) possesses a finite integral over E according to definition (i). And, if this condition is satisfied, we shall understand by the *definite integral* ($\mathfrak{X}, \mu$) of the function $f(x)$ over E the difference between the integral of $\overset{0}{f}(x)$ and that of $-\overset{0}{f}(x)$ over E. The definite integral ($\mathfrak{X}, \mu$) of $f(x)$ over E will be written $(\mathfrak{X}) \int_{E} f(x) \, d\mu(x)$. If this integral is finite, the function $f(x)$ is said to be *integrable* ($\mathfrak{X}, \mu$). For every function $f(x)$ possessing a definite integral over a set E, we evidently have

$$(\mathfrak{X}) \int_{E} f \, d\mu = (\mathfrak{X}) \int_{E} \overset{0}{f} \, d\mu - (\mathfrak{X}) \int_{E} (-\overset{0}{f}) \, d\mu = (\mathfrak{X}) \int_{E} \overset{0}{f} \, d\mu + (\mathfrak{X}) \int_{E} \overset{0}{f} \, d\mu.$$

We see at once that the two definitions (i) and (ii) are compatible, i. e. that they give the same value of the integral to any non-negative measurable function. Moreover:

(10.1) *If $g = \{v_1, X_1; v_2, X_2; ...; v_m, X_m\}$ is a simple non-negative function on the set $E = X_1 + X_2 + ... + X_m$, the sets X_l being measurable ($\mathfrak{X}$), then*

$$\int_{E} g \, d\mu = \sum_{l=1}^{m} v_l \mu(X_l).$$

For, if $\{E_j\}_{j=1,2,...,n}$ is an arbitrary subdivision of E into a finite number of sets ($\mathfrak{X}$) without points in common, and if w_j denotes the lower bound of $g(x)$ on E_j, we have $w_j \leqslant v_l$ whenever $E_j \cdot X_l \neq 0$.

Hence $\sum_{j=1}^{n} w_j \mu(E_j) = \sum_{j=1}^{n} \sum_{i=1}^{m} w_j \mu(E_j \cdot X_l) \leqslant \sum_{j=1}^{n} \sum_{i=1}^{m} v_l \mu(E_j \cdot X_l) = \sum_{l=1}^{m} v_l \mu(X_l)$, and therefore $\int_{E} g \, d\mu \leqslant \sum_{l} v_l \mu(X_l)$. The opposite inequality is obvious, since the sets $X_1, X_2, ..., X_m$ themselves constitute a subdivision of E into a finite sequence of sets ($\mathfrak{X}$) on which the values of $g(x)$ are $v_1, v_2; ...; v_m$ respectively.

§ 11. Fundamental properties of the integral.

We shall begin with a few lemmas concerning integration of simple functions. As in the preceding §§, the symbols $\mathfrak{X}$, μ etc. will often be omitted.

(11.1) *Lemma.* 1⁰ *For every pair of functions $g(x)$ and $h(x)$, simple, non-negative, and measurable ($\mathfrak{X}$) on a set E, we have*

$$(11.2) \qquad \int_{E} [g(x) + h(x)] \, d\mu(x) = \int_{E} g(x) \, d\mu(x) + \int_{E} h(x) \, d\mu(x).$$

2⁰ *If the function $f(x)$ is simple, non-negative, and measurable ($\mathfrak{X}$) on the set $A+B$ where A and B are sets ($\mathfrak{X}$) without common points, then*

$$(11.3) \qquad \int_{A+B} f(x) \, d\mu(x) = \int_{A} f(x) \, d\mu(x) + \int_{B} f(x) \, d\mu(x).$$

Proof. As regards 1⁰, let

$$g = \{g_1, G_1; g_2, G_2; ...; g_n, G_n\} \qquad \text{and} \qquad h = \{h_1, H_1; h_2, H_2; ...; h_m, H_m\},$$

where $\quad E = G_1 + ... + G_n = H_1 + ... + H_m$.

We then have, by (10.1),

$$\int_{E} [g(x) + h(x)] \, d\mu(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} (g_i + h_j) \mu(G_i \cdot H_j) =$$
$$= \sum_{i=1}^{n} g_i \sum_{j=1}^{m} \mu(G_i \cdot H_j) + \sum_{j=1}^{m} h_j \sum_{i=1}^{n} \mu(G_i \cdot H_j) =$$
$$= \sum_{i=1}^{n} g_i \mu(G_i) + \sum_{j=1}^{m} h_j \mu(H_j) = \int_{E} g(x) \, d\mu(x) + \int_{E} h(x) \, d\mu(x).$$

As regards 2⁰, if $E = A + B$ and $f = \{f_1, Q_1; f_2, Q_2; ...; f_n, Q_n\}$, where $E = Q_1 + Q_2 + ... + Q_n$, we have

$$\int_{E} f(x) \, d\mu(x) = \sum_{i=1}^{n} f_i \mu(Q_i) = \sum_{i=1}^{n} f_i \mu(A \cdot Q_i) + \sum_{i=1}^{n} f_i \mu(B \cdot Q_i) =$$
$$= \int_{A} f(x) \, d\mu(x) + \int_{B} f(x) \, d\mu(x).$$

(11.4) *Lemma.* If $\{g_n(x)\}$ is a non-decreasing sequence of functions that are simple, non-negative, and measurable ($\mathfrak{X}$) on a set E, and if, for a function $h(x)$, simple, non-negative, and measurable, on E, we have $\lim_{n} g_n(x) \geqslant h(x)$ on E, then

$$(11.5) \qquad \lim_{n} \int_{E} g_n(x) \, d\mu(x) \geqslant \int_{E} h(x) \, d\mu(x).$$

Proof. Let $h = \{v_1, E_1; v_2, E_2; ...; v_m, E_m\}$, where

$$0 \leqslant v_1 < v_2 < ... < v_m \quad and \quad E = E_1 + E_2 + ... + E_m.$$

We may suppose $v_1 > 0$, for, otherwise, we should have $\int_E h \, d\mu = \int_{E-E_1} h \, d\mu$, and, since $\int_E g_n \, d\mu \geqslant \int_{E-E_1} g_n \, d\mu$, we could replace the set E by the set $E - E_1$ on which $h(x)$ does not vanish anywhere. Further we shall assume first that $v_m < +\infty$.

Let us choose an arbitrary positive number $\varepsilon < v_1$, and let us denote, for each positive integer n, by Q_n the set of the points x of E for which $g_n(x) > h(x) - \varepsilon$. The sets Q_n evidently form an ascending sequence converging to E, and, by Theorem 9.1, we have $\mu(Q_n) \to \mu(E)$. This being so we have two cases to distinguish:

(i) $\mu(E) \neq \infty$. We then can find an integer n_0 such that for $n > n_0$ we have $\mu(E - Q_n) < \varepsilon$, and therefore, by Lemma 11.1,

$$\int_E g_n \, d\mu \geqslant \int_{Q_n} g_n \, d\mu \geqslant \int_{Q_n} [h(x) - \varepsilon] \, d\mu(x) =$$

$$= \int_{Q_n} h \, d\mu - \varepsilon \mu(Q_n) \geqslant \int_E h \, d\mu - v_m \cdot \mu(E - Q_n) - \varepsilon \mu(Q_n) \geqslant \int_E h \, d\mu - [v_m + \mu(E)] \cdot \varepsilon;$$

and, passing to the limit, making first $n \to \infty$, and then $\varepsilon \to 0$, we obtain the inequality (11.5).

(ii) $\mu(E) = \infty$. Then, since $\int_E g_n \, d\mu \geqslant (v_1 - \varepsilon) \mu(Q_n)$, we obtain $\lim_n \int_E g_n \, d\mu = \infty$, so that the inequality (11.5) is evidently satisfied.

Suppose now $v_m = +\infty$. Then by (10.1) and by what has already been proved, $\lim_n \int_E g_n \, d\mu \geqslant v \cdot \mu(E_m) + \sum_{l=1}^{m-1} v_l \cdot \mu(E_l)$ for any finite number v, and consequently for $v = +\infty = v_m$ also; whence, in virtue of (10.1) the inequality (11.5) follows at once.

(11.6) Lemma. *If the functions of a non-decreasing sequence $\{g_n(x)\}$ are simple, non-negative, and measurable $(\mathfrak{X})$ on a set E, and if $g(x) = \lim_n g_n(x)$, then $\lim_n \int_E g_n(x) \, d\mu(x) = \int_E g(x) \, d\mu(x)$.*

Proof. Let $E_1, E_2, ..., E_m$ be an arbitrary subdivision of E into a finite number of measurable sets, and let $v_1, v_2, ..., v_m$ be the lower bounds of $g(x)$ on these sets respectively. Let us write $v = \{v_1, E_1; v_2, E_2; ...; v_m, E_m\}$. We evidently have $\lim_n g_n(x) = g(x) \geqslant v(x)$ on E, and hence, by Lemma 11.4 and by Theorem 10.1

$$\lim_n \int_E g_n \, d\mu \geqslant \int_E v \, d\mu = \sum_{i=1}^m v_i \mu(E_i).$$

It follows that $\lim_n \int_E g_n \, d\mu \geqslant \int_E g \, d\mu$, and since the opposite inequality is obvious, the proof is complete.

We are now in a position to generalize Lemma 11.1 as follows:

(11.7) Theorem. *The relation* (11.2) *holds for every pair of functions, $g(x)$ and $h(x)$, non-negative and measurable $(\mathfrak{X})$ on the set E, and the relation* (11.3) *holds for every function $f(x)$ non-negative and measurable $(\mathfrak{X})$ on the set $A + B$, where A and B are sets $(\mathfrak{X})$ without points in common.*

Proof. By Theorem 7.4 there exist two non-decreasing sequences $\{g_n(x)\}$ and $\{h_n(x)\}$ of simple non-negative functions measurable $(\mathfrak{X})$ on E, such that $g(x) = \lim_n g_n(x)$ and $h(x) = \lim_n h_n(x)$. Now, by Lemma 11.1 (1^0), we have $\int_E (g_n + h_n) \, d\mu = \int_E g_n \, d\mu + \int_E h_n \, d\mu$ and hence, making $n \to \infty$, we obtain, on account of Lemma 11.6, the relation (11.2). Similarly, if we approximate to $f(x)$ on $A + B$ by a non-decreasing sequence of simple non-negative functions and make use of Lemma 11.1 (2^0), we obtain the relation (11.3).

(11.8) Theorem. 1^0 *For any function measurable $(\mathfrak{X})$, the integral over a set of measure zero is equal to zero. 2^0 If the functions $g(x)$ and $h(x)$ measurable on a set E are almost everywhere equal on E, and if one of the two is integrable on E, so is the other, and their integrals over E have the same value. 3^0 If a function $f(x)$ measurable $(\mathfrak{X})$ on a set E has an integral over E different from $+\infty$, the set of the points x of E at which $f(x) = +\infty$ has measure zero. In particular, if the integral of $f(x)$ over E is finite, the function $f(x)$ is finite almost everywhere on E.*

Proof. We obtain at once part 1^0 of this theorem by making successive use of the definitions (i) and (ii) of § 10.

As regards 2^0, it is evidently sufficient to consider the case of non-negative functions $g(x)$ and $h(x)$. If we denote by E_1 the set of the points x of E at which $g(x) \neq h(x)$, we have by hypothesis $\mu(E_1) = 0$, and, on account of (1^0) and of Theorem 11.7, we obtain

$$\int_E g \, d\mu = \int_{E-E_1} g \, d\mu = \int_{E-E_1} h \, d\mu = \int_E h \, d\mu, \quad \text{as required.}$$

Finally, as regards 3^0, let us suppose that for a function $f(x)$ measurable ($\mathfrak{X}$) on E we have $f(x) = +\infty$ on a set $E_0 \subset E$ of positive measure. We then have $\int\limits_E \overset{\circ}{f}\, d\mu \geqslant \int\limits_{E_0} \overset{\circ}{f}\, d\mu \geqslant n \cdot \mu\,(E_0)$ for every n, and so $\int\limits_E \overset{\circ}{f}\, d\mu = +\infty$. Consequently, the integral of $f(x)$ over E, if it exists, is positively infinite, and this completes the proof.

We now generalize Lemma 11.1 (1^0) and also complete Theorem 11.7, as follows:

(11.9) **Theorem of distributivity of the integral.** *Every linear combination with constant coefficients, $a \cdot g(x) + b \cdot h(x)$ of two functions $g(x)$ and $h(x)$, integrable ($\mathfrak{X}, \mu$) over a set E, is also integrable over E, and we have*

(11.10) $$\int\limits_E (ag + bh)\, d\mu = a \int\limits_E g\, d\mu + b \int\limits_E h\, d\mu.$$

Proof. By Theorem 11.8 (3^0), the set of the points at which either of the functions $g(x)$ and $h(x)$ is infinite, has measure zero, and if we replace on this set the values of both functions by 0, we shall not affect the values of the integrals appearing in the relation (11.10). We may therefore suppose that the given functions g and h are finite on E. Further, the relations

$$\int\limits_E ag\, d\mu = a \int\limits_E g\, d\mu, \qquad \int\limits_E bh\, d\mu = b \int\limits_E h\, d\mu$$

being obvious, we need only prove the formula (11.10) for the case $a = b = 1$. Finally, the set E can be decomposed into four sets on each of which the two functions $g(x)$ and $h(x)$ are of constant sign. So that, on account of Theorem 11.7, we may assume that the functions $g(x)$ and $h(x)$ are of constant sign on the whole set E. Now, by the same theorem, the relation

(11.11) $$\int\limits_E (g + h)\, d\mu = \int\limits_E g\, d\mu + \int\limits_E h\, d\mu$$

holds whenever the functions g and h are both non-negative or both non-positive on E, and it only remains, therefore, to show that this relation is valid when g and h have, on E, opposite signs, the one, $g(x)$ say, being non-negative, the other, $h(x)$, non-positive.

This being so, let E_1 and E_2 be the sets consisting of the points x of E for which we have $g(x) + h(x) \geqslant 0$ and $g(x) + h(x) < 0$, respectively. The functions g, $g + h$, and $-h$ are non-negative on E_1 and we therefore have, by Theorem 11.7,

$$\int\limits_{E_1} g\, d\mu = \int\limits_{E_1} (g + h)\, d\mu + \int\limits_{E_1} (-h)\, d\mu = \int\limits_{E_1} (g + h)\, d\mu - \int\limits_{E_1} h\, d\mu.$$

Similarly

$$-\int\limits_{E_2} h\, d\mu = \int\limits_{E_2} (-h)\, d\mu = \int\limits_{E_2} (-g - h)\, d\mu + \int\limits_{E_2} g\, d\mu = -\int\limits_{E_2} (g + h)\, d\mu + \int\limits_{E_2} g\, d\mu.$$

Therefore, for $i = 1, 2$, we have $\int\limits_{E_i} (g + h)\, d\mu = \int\limits_{E_i} g\, d\mu + \int\limits_{E_i} h\, d\mu$, and by Theorem 11.7 we obtain the relation (11.11).

(11.12) **Theorem on absolute integrability.** 1^0 *In order that a function $f(x)$ measurable ($\mathfrak{X}$) on a set E should be integrable ($\mathfrak{X}, \mu$) on E, it is necessary and sufficient that its absolute value should be so. 2^0 If, for a function $g(x)$ measurable ($\mathfrak{X}$) on a set E, there exists a function $h(x)$, integrable ($\mathfrak{X}, \mu$) and such that $|g(x)| \leqslant h(x)$ on E, then the function $g(x)$ also is integrable on E; in particular, every function measurable ($\mathfrak{X}$) and bounded on a set E of finite measure (μ) is integrable ($\mathfrak{X}, \mu$) on E.*

Proof. As regards 1^0, we have by Theorem 11.7

$$\int\limits_E |f|\, d\mu = \int\limits_E \overset{\circ}{f}\, d\mu + \int\limits_E (-f)\, d\mu,$$

and integrability of $|f|$ is therefore equivalent to that of $\overset{\circ}{f}$ and that of $-f$ holding together, i. e. to integrability of f.

As regards 2^0, we have the inequalities $g(x) \leqslant |g(x)| \leqslant h(x)$ and $-g(x) \leqslant |g(x)| \leqslant h(x)$ on E, and, since $h(x)$ is, by hypothesis, integrable on E, it follows that the same is true of the non-negative functions $\overset{\circ}{g}$ and $-g$, and therefore of the function $g(x)$.

As an immediate consequence of Theorem 11.12 we have the following theorem, known as the

(11.13) *First Mean Value Theorem. Given, on a set E, a function $f(x)$ bounded and measurable $(\mathfrak{X})$ on E and a function $g(x)$ integrable $(\mathfrak{X}, \mu)$ on E, the function $f(x) \cdot g(x)$ is integrable on E and there exists a number γ lying between the bounds of $f(x)$ on E, such that*

$$(11.14) \qquad \int_E |f(x)| |g(x)| \, d\mu(x) = \gamma \cdot \int_E |g(x)| \, d\mu(x).$$

Proof. If we denote by m and M respectively the lower and the upper bound of $f(x)$ on E, and make use of Theorem 11.12, we verify successively, that the functions $(|M|+|m|) \cdot |g(x)|$, $|f(x) g(x)|$, $f(x) |g(x)|$ and $f(x) g(x)$ are integrable on E. Further, we have $m |g(x)| \leqslant f(x) \cdot |g(x)| \leqslant M |g(x)|$ over E, and, therefore also, $m \int_E |g| \, d\mu \leqslant \int_E f \cdot |g| \, d\mu \leqslant M \int_E |g| \, d\mu$, and so choosing $\gamma = [\int_E f \cdot |g| \, d\mu] : [\int_E |g| \, d\mu]$

(or, if the denominator vanishes, an arbitrary γ between m and M), we obtain the formula (11.14) with $m \leqslant \gamma \leqslant M$.

§ 12. Integration of sequences of functions.
In this §, we shall establish some theorems on term by term integration of sequences and series of functions.

(12.1) *Theorem. If the functions of a sequence $\{g_n(x)\}$ are finite and integrable $(\mathfrak{X}, \mu)$ on a set E of finite measure, and the sequence converges uniformly on E to a function $g(x)$, then the function $g(x)$ also is integrable over E, and we have*

$$(12.2) \qquad \lim_n \int_E g_n(x) \, d\mu(x) = \int_E g(x) \, d\mu(x).$$

Proof. By Theorem 8.5, the function $g(x)$ is measurable $(\mathfrak{X})$ on E. The functions $g(x) - g_n(x)$ are therefore all measurable also, and, further, since the sequence $\{g_n(x)\}$ converges uniformly to $g(x)$ the functions $g(x) - g_n(x)$ are all bounded, at any rate from some value of the index n onwards. These functions are thus, by Theorem 11.12 (2^0), integrable on E, and it follows, by Theorem 11.9, that the function $g(x) = [g(x) - g_n(x)] + g_n(x)$ is integrable too. Finally, denoting by ε_n the upper bound of $|g(x) - g_n(x)|$ on E, we have

$$\left| \int_E g(x) \, d\mu(x) - \int_E g_n(x) \, d\mu(x) \right| \leqslant \int_E |g(x) - g_n(x)| \, d\mu(x) \leqslant \varepsilon_n \mu(E),$$

and this establishes the relation (12.2) since, by hypothesis, $\varepsilon_n \to 0$ and $\mu(E) \neq \infty$.

Neither the theorem thus established, nor its proof, contains, at bottom, anything new, as compared with the similar result for the classical processes of integration of Cauchy, or of Riemann. We now pass on to the proof of theorems more closely related to Lebesgue integration. Among these theorems, a fundamental part is played by the following one, which is due to Lebesgue:

(12.3) *Theorem. Let $f(x) = \sum_{n=1}^{\infty} f_n(x)$ be a series of non-negative functions measurable $(\mathfrak{X})$ on a set E. Then*

$$(12.4) \qquad \int_E f(x) \, d\mu(x) = \sum_{n=1}^{\infty} \int_E f_n(x) \, d\mu(x).$$

Proof. From Theorem 11.7, we derive in the first place, that $\int_E f \, d\mu \geqslant \int_E [\sum_{n=1}^{m} f_n \, d\mu] = \sum_{n=1}^{m} \int_E f_n \, d\mu$ for every m, and so

$$(12.5) \qquad \int_E f \, d\mu \geqslant \sum_{n=1}^{\infty} \int_E f_n \, d\mu.$$

To establish the opposite inequality, let us attach, in accordance with Theorem 7.4, to each function $f_n(x)$ a non-decreasing sequence $\{g_n^{(k)}(x)\}_{k=1,2,\dots}$ of simple functions measurable and non-negative on E, in such a manner that $\lim_k g_n^{(k)}(x) = f_n(x)$ for $n=1, 2,\dots$ Let us write $s_k(x) = \sum_{i=1}^{k} g_i^{(k)}(x)$. The functions $s_k(x)$ are clearly simple, measurable, and non-negative, on E, and they form a non-decreasing sequence. Further, for each m, and for $k \geqslant m$, we have $\sum_{i=1}^{m} g_i^{(k)}(x) \leqslant s_k(x) \leqslant f(x)$. Making $k \to \infty$, we derive $\sum_{i=1}^{m} f_i(x) \leqslant \lim_k s_k(x) \leqslant f(x)$ for every m, and so, $f(x) = \lim_k s_k(x)$. Therefore, by Lemmas 11.6 and 11.1 (1^0),

$$\int_E f \, d\mu = \lim_k \int_E s_k \, d\mu = \lim_k \sum_{i=1}^{k} \int_E g_i^{(k)} \, d\mu \leqslant \sum_{i=1}^{\infty} \int_E f_i \, d\mu,$$

and this, combined with (12.5), gives the equality (12.4).

Theorem 12.3 may also be stated in the following form:

(12.6) *Lebesgue's Theorem on integration of monotone sequences of functions.* *If $\{f_n(x)\}$ is a non-decreasing sequence of non-negative functions measurable $(\mathfrak{X})$ on a set E, and $f(x) = \lim_n f_n(x)$,*
then $\int_E f(x)\, d\mu(x) = \lim_n \int_E f_n(x)\, d\mu(x)$.

Proof. If we write $g_n(x) = f_{n+1} - f_n(x)$, we obtain

$$f(x) = f_1(x) + \sum_{n=1}^{\infty} g_n(x),$$

and the functions $g_n(x)$ will be non-negative and measurable on E, so that by Theorem 12.3

$$\int_E f\, d\mu = \int_E f_1\, d\mu + \sum_{n=1}^{\infty} \int_E g_n\, d\mu = \lim_k \int_E [f_1 + \sum_{n=1}^{k-1} g_n]\, d\mu = \lim_k \int_E f_k\, d\mu.$$

Q. E. D.

(12.7) *Theorem of additivity for the integral.* *If $\{E_n\}$ is a sequence of sets measurable $(\mathfrak{X})$ no two of which have common points, and $E = \sum_n E_n$, then*

(12.8) $$\int_E f\, d\mu = \sum_n \int_{E_n} f\, d\mu$$

for every function $f(x)$ possessing a definite integral (finite or infinite) over E.

Proof. It is clearly sufficient to prove (12.8) in the case of a function $f(x)$ non-negative on E. Supposing this to be the case, let us write $f_n(x) = f(x)$ for $x \epsilon E_n$, and $f_n(x) = 0$ for $x \epsilon E - E_n$. We then have $f(x) = \sum_n f_n(x)$ on E, and, the functions f_n being measurable and non-negative, we may apply Lebesgue's Theorem 12.3. This gives, by Theorem 11.7,

$$\int_E f\, d\mu = \sum_n \int_E f_n\, d\mu = \sum_n \int_{E_n} f_n\, d\mu = \sum_n \int_{E_n} f\, d\mu.$$ Q. E. D.

If a function $f(x)$ has a definite integral $(\mathfrak{X}, \mu)$ over a set E, then $f(x)$ also has a definite integral over any subset of E measurable $(\mathfrak{X})$. We may therefore associate with it the function of a set $(\mathfrak{X})$ defined as follows:

(12.9) $$F(X) = \int_X f(x)\, d\mu(x) \quad \text{where} \quad X \subset E \quad \text{and} \quad X \epsilon \mathfrak{X}.$$

The latter will be called the *indefinite integral* $(\mathfrak{X}, \mu)$ of $f(x)$ on E. It follows from Theorem 12.7 that, whenever the function $f(x)$ is integrable $(\mathfrak{X}, \mu)$ on E, its indefinite integral is an additive function of set $(\mathfrak{X})$ on E.

We end this § with two simple but important theorems. The first is known as *Fatou's lemma*, and appears for the first time (in a slightly less general form) in the classical memoir of P. Fatou [1, p. 375] on trigonometric series. The second is due to Lebesgue [5, in particular p. 375], and is called the *theorem on term by term integration of sequences of functions*; cf. also Ch. J. de la Vallée Poussin [1, p. 445—453], R. L. Jeffery [1] and T. H. Hildebrandt [2].

(12.10) *Theorem (Fatou's Lemma).* *If $\{f_n(x)\}$ is any sequence of non-negative functions measurable $(\mathfrak{X})$ on a set E, we have*

$$\int_E \liminf_n f_n(x)\, d\mu(x) \leqslant \liminf_n \int_E f_n(x)\, d\mu(x).$$

Proof. Let us write $g_i(x) = \text{lower bound } [f_i(x), f_{i+1}(x), f_{i+2}(x), \ldots]$ where $i = 1, 2, \ldots$. Thus defined $\{g_i(x)\}$ is a non-decreasing sequence of non-negative functions measurable on E, and converges on the set E to $\liminf_i f_i(x)$. We therefore have, by Lebesgue's Theorem 12.6,

$$\int_E \liminf_i f_i(x)\, d\mu(x) = \lim_i \int_E g_i(x)\, d\mu(x) \leqslant \liminf_i \int_E f_i(x)\, d\mu(x).$$

(12.11) *Lebesgue's Theorem on term by term integration.* *Let $\{f_n(x)\}$ be a sequence of functions measurable $(\mathfrak{X})$ on a set E, fulfilling, for a function $s(x)$ integrable $(\mathfrak{X})$ on E, the inequality $|f_n(x)| \leqslant s(x)$ for $n = 1, 2, \ldots$ Then*

(12.12)
$$\liminf_n \int_E f_n\, d\mu \geqslant \int_E \liminf_n f_n\, d\mu,$$

$$\limsup_n \int_E f_n\, d\mu \leqslant \int_E \limsup_n f_n\, d\mu.$$

If, further, the sequence $\{f_n\}$ converges on E to a function f, the sequence is integrable term by term, i. e. we have

(12.13) $$\lim_n \int_E f_n\, d\mu = \int_E f\, d\mu.$$

Proof. Let $g(x) = \liminf_n f_n(x)$ and let $h(x) = \limsup_n f_n(x)$. We may clearly suppose $s(x) < +\infty$ throughout E. We then derive from Fatou's Lemma 12.10, $\liminf_n \int_E (s+f_n)\,d\mu \geqslant \int_E (s+g)\,d\mu$ and

$$\liminf_n \int_E (s-f_n)\,d\mu \geqslant \int_E (s-h)\,d\mu, \quad \text{which gives at once the rela-}$$

tions (12.12).

Further, if $\lim_n f_n(x) = f(x)$, we derive from (12.12) the relation

$$\liminf_n \int_E f_n\,d\mu \geqslant \int_E f\,d\mu \geqslant \limsup_n \int_E f_n\,d\mu \quad \text{which gives the equality}$$

(12.13).

§ 13. Absolutely continuous additive functions of a set.

The fact that the indefinite integral of a function integrable $(\mathfrak{X}, \mu)$ on a set E is, on E, an additive function of a set $(\mathfrak{X})$, raises the problem of characterizing directly the additive functions expressible as indefinite integrals.

If we restrict ourselves to the Lebesgue integral of functions of a real variable, we may regard indefinite integrals as functions of an interval, or, what comes to the same thing, as functions of a real variable. In that case, a necessary and sufficient condition for a function to be expressible as the indefinite integral of a real function was given, in 1904, still by Lebesgue [I, p. 129, footnote]. A little later (in 1905), G. Vitali [1] explicitly distinguished the class of functions possessing the Lebesgue property by introducing the name of "absolutely continuous functions".

The condition of Lebesgue and Vitali was later extended to functions of a set by J. Radon [1] (cf. also P. J. Daniell [2]). But Radon considered only additive functions of sets measurable in the Borel sense in Euclidean spaces, and only measures determined by additive functions of intervals (cf. below Chapter III). The final form of the condition of Lebesgue-Vitali, as given in Theorem 14.11 below, is due to O. Nikodym [2].

An additive function of a set $(\mathfrak{X})$ on a set E, will be said to be *absolutely continuous* $(\mathfrak{X}, \mu)$ on E, if the function vanishes for every subset $(\mathfrak{X})$ of E whose measure (μ) is zero. An additive function $\Phi(X)$ of a set $(\mathfrak{X})$ on a set E will be termed *singular* $(\mathfrak{X}, \mu)$ on E, if there exists a subset $E_0 \subset E$ measurable $(\mathfrak{X})$, of measure (μ) zero, such that $\Phi(X)$ vanishes identically on $E - E_0$, i. e. $\Phi(X) = \Phi(E_0 \cdot X)$ for every subset X of E measurable $(\mathfrak{X})$. The following statements are at once obvious:

(13.1) Theorem. 1^0 *In order that an additive function of a set $(\mathfrak{X})$ on a set E should be absolutely continuous $(\mathfrak{X}, \mu)$ [or should be singular] it is necessary and sufficient that its two variations, the upper and the lower, should both be so.* 2^0 *Every linear combination, with constant coefficients, of two additive functions absolutely continuous [or singular] on a set E is itself absolutely continuous [or singular] on E.* 3^0 *If a sequence $\{\Phi_n(X)\}$ of additive functions, absolutely continuous [singular] on a set E, converges to an additive function $\Phi(X)$ for each measurable subset X of E, then the function $\Phi(X)$ is also absolutely continuous [singular].* 4^0 *If a function of a set $(\mathfrak{X})$ is additive and absolutely continuous [singular] on a set E, the function is so on every measurable subset of E.* 5^0 *If $E = \sum_n E_n$, where $\{E_n\}$ is a sequence of sets $(\mathfrak{X})$, and if an additive function $\Phi(X)$ on E is absolutely continuous [singular] on each of the sets E_n, the function is absolutely continuous [singular] on the whole set E.* 6^0 *An additive function of a set cannot be both absolutely continuous and singular on a set E, without vanishing identically on E.*

For sets of finite measure, it is sometimes convenient to apply the following test for absolute continuity:

(13.2) Theorem. *In order that a function $\Phi(X)$ additive on a set E of finite measure, be absolutely continuous $(\mathfrak{X}, \mu)$ on E, it is necessary and sufficient that to each $\varepsilon > 0$ there correspond an $\eta > 0$, such that $\mu(X) < \eta$ imply $|\Phi(X)| < \varepsilon$ for every set $X \subset E$ measurable $(\mathfrak{X})$.*

Proof. It is evident that the condition is sufficient. To prove it also necessary, let us suppose the function $\Phi(X)$ absolutely continuous in E. We may assume, replacing if necessary, $\Phi(X)$ by its absolute variation, that $\Phi(X)$ is a non-negative monotone function on E. This being so, let us suppose, if possible, that there exists a sequence $\{E_n\}_{n=1,2,\ldots}$ of measurable subsets of E, such that $\mu(E_n) < 1/2^n$ and that $\Phi(E_n) > \eta_0$, where η_0 is a fixed positive number. Let us write $E_0 = \limsup_n E_n$. For every n, we then have

$$\mu(E_0) \leqslant \sum_{k=n}^{\infty} \mu(E_k) \leqslant 1/2^{n-1}, \quad \text{and therefore} \quad \mu(E_0) = 0. \quad \text{On the other}$$

hand, by Theorem 5.1, we have $\Phi(E_0) \geqslant \limsup_n \Phi(E_n) \geqslant \eta_0$. This is a contradiction, since $\Phi(X)$ is absolutely continuous, and the proof is complete.

(13.3) **Theorem.** *In order that a function $\Phi(X)$, additive on a set E, be singular $(\mathfrak{X}, \mu)$ on E, it is necessary and sufficient that for each $\varepsilon > 0$ there exist a set $X \subset E$ measurable $(\mathfrak{X})$ and fulfilling the two conditions $\mu(X) < \varepsilon$, $\mathrm{W}(\Phi; E - X) < \varepsilon$.*

Proof. The condition is clearly necessary. To prove it sufficient, let us suppose that for each n there is a set $X_n \subset E$ measurable $(\mathfrak{X})$ such that $\mu(X_n) < 1/2^n$ and $\mathrm{W}(\Phi; E - X_n) < 1/2^n$, and let us write $E_0 = \limsup\limits_n X_n$. We then have $\mu(E_0) \leqslant \sum\limits_{k=n}^{\infty} \mu(X_k) \leqslant 1/2^{n-1}$ for each n, and so $\mu(E_0) = 0$. On the other hand, by Theorem 5.1 we have $\mathrm{W}(\Phi; E - E_0) \leqslant \liminf\limits_n \mathrm{W}(\Phi; E - X_n) = 0$. The function $\Phi(X)$ is therefore singular on E.

§ 14. The Lebesgue decomposition of an additive function.

Before proving the result announced in the preceding §, we shall establish some auxiliary theorems. We begin with the following theorem due to H. Hahn [I, p. 404] (cf. also W. Sierpiński [11]):

(14.1) **Theorem.** *If $\Phi(X)$ is an additive function of a set $(\mathfrak{X})$ on a set E, there exists always a set $P \subset E$ measurable $(\mathfrak{X})$, such that $\underline{\mathrm{W}}(\Phi; P) = 0 = \overline{\mathrm{W}}(\Phi; E - P)$, or, what comes to the same thing, such that $\Phi(X) \geqslant 0$ for every measurable set $X \subset P$ and $\Phi(X) \leqslant 0$ for every measurable set $X \subset E - P$.*

Proof. For each positive integer n, we choose a set E_n such that $\Phi(E_n) \geqslant \overline{\mathrm{W}}(\Phi; E) - 1/2^n$. By Theorem 6.4 we then have,

$$(14.2) \qquad \underline{\mathrm{W}}(\Phi; E_n) \geqslant -1/2^n \quad and \quad \overline{\mathrm{W}}(\Phi; E - E_n) \leqslant 1/2^n.$$

Writing $P = \liminf\limits_n E_n$, we see that $E - P = \limsup\limits_n (E - E_n) \subset \sum\limits_{n=m}^{\infty}(E - E_n)$ for every m, and therefore, by (14.2),

$$\overline{\mathrm{W}}(\Phi; E - P) \leqslant \sum_{n=m}^{\infty} \overline{\mathrm{W}}(\Phi; E - E_n) \leqslant \frac{1}{2^{m-1}}$$

which gives $\overline{\mathrm{W}}(\Phi; E - P) = 0$. On the other hand, the lower variation $\underline{\mathrm{W}}(\Phi; X)$ is a non-positive monotone function of a measurable set $X \subset E$, and, by Theorem 5.1 and the first inequality (14.2), we must have the relation $|\underline{\mathrm{W}}(\Phi; P)| \leqslant \liminf\limits_n |\underline{\mathrm{W}}(\Phi; E_n)| = 0$, which gives $\underline{\mathrm{W}}(\Phi; P) = 0$ and completes the proof.

(14.3) **Lemma.** *If $\Phi(X)$ is a non-negative additive function of a set $(\mathfrak{X})$ on a set E, there exists, for each $a > 0$, a decomposition of E into a sequence of measurable sets without common points, $H, E_1, E_2, ..., E_n, ...$ such that $\mu(H) = 0$ and that*

$$(14.4) \qquad a \cdot (n-1) \cdot \mu(X) \leqslant \Phi(X) \leqslant an \cdot \mu(X)$$

for every set $X \subset E_n$ measurable $(\mathfrak{X})$.

Proof. By Theorem 14.1, there exists, for each positive integer n, a measurable set A_n such that $\Phi(X) - an \cdot \mu(X) \geqslant 0$ for every measurable set $X \subset A_n$ and $\Phi(X) - an \cdot \mu(X) \leqslant 0$ for every measurable set $X \subset E - A_n$. Write $B_n = \sum\limits_{k=n}^{\infty} A_k$. Any measurable subset X of B_n may be represented in the form $X = \sum\limits_{k=n}^{\infty} X_k$, where X_k are measurable sets, $X_k \subset A_k$ for $k = n$, $n+1$, ..., and $X_i \cdot X_j = 0$ for $i \neq j$; and so $\Phi(X) = \sum\limits_{k=n}^{\infty} \Phi(X_k) \geqslant \sum\limits_{k=n}^{\infty} ak \cdot \mu(X_k) \geqslant an \cdot \mu(X)$. We obtain thus a descending sequence of measurable sets $\{B_n\}$ such that

$$(14.5) \qquad \begin{aligned} \Phi(X) &\geqslant an \cdot \mu(X) \quad if \quad X \subset B_n, \qquad X \in \mathfrak{X}, \\ \Phi(X) &\leqslant an \cdot \mu(X) \quad if \quad X \subset E - B_n, \quad X \in \mathfrak{X}, \end{aligned}$$

the second relation being obvious, since $E - B_n$ is a subset of $E - A_n$.

Let us now write $E_1 = E - B_1$, $E_n = B_{n-1} - B_n$ for $n = 2, 3, ...$, and $H = \lim\limits_n B_n$. Thus defined the sets $H, E_1, E_2, ..., E_n, ...$ are measurable and without common points, and $E = H + \sum\limits_{n=1}^{\infty} E_n$. Taking into account the relations (14.5), we see at once that the inequality (14.4) holds whenever X is a measurable subset of E_n. Finally, $H \subset B_n$ for each positive integer n, and therefore, by the first of the relations (14.5), we get $\Phi(H) \geqslant an \cdot \mu(H)$, which requires $\mu(H) = 0$ and completes the proof.

(14.6) **Theorem.** *If E is a set $(\mathfrak{X})$ of finite measure (μ), or, more generally, a set expressible as the sum of a sequence of sets $(\mathfrak{X})$ of finite measure, every additive function of a measurable set $\Phi(X)$ on E is expressible as the sum of an absolutely continuous additive function $\Psi(X)$ and a singular additive function $\Theta(X)$ on E. Such a decomposition of $\Phi(X)$ on E is unique, and the function $\Psi(X)$ is, on E,*

the indefinite integral of a function integrable $(\mathfrak{X}, \mu)$ *on* E. *If* $\Phi(X)$ *is a non-negative monotone function on* E, *so are the corresponding functions* $\Psi(X)$ *and* $\Theta(X)$.

Proof. Since every additive function of a set is the difference of two non-negative functions of the same kind (cf. § 6), we may restrict ourselves to the case of a non-negative $\Phi(X)$. Further, we shall assume to begin with that the set E has finite measure. By the preceding lemma, there exists, for every positive integer m, a decomposition of E into a sequence of measurable sets $H_1^{(m)}, E_1^{(m)}, E_2^{(m)}, ...$, without common points and subject to the conditions:

$$(14.7) \qquad E = H^{(m)} + E_1^{(m)} + ... + E_n^{(m)} + ..., \qquad \mu(H^{(m)}) = 0,$$

$$(14.8) \quad 2^{-m} \cdot (n-1) \cdot \mu(X) \leqslant \Phi(X) \leqslant 2^{-m} \cdot n \cdot \mu(X), \ \text{if} \ X \subset E_n^{(m)}, \ X \epsilon \mathfrak{X}.$$

We therefore have, for all positive integers m, n, and k,

$$2^{-m} \cdot n \cdot \mu(E_n^{(m)} \cdot E_k^{(m+1)}) \geqslant \Phi(E_n^{(m)} \cdot E_k^{(m+1)}) \geqslant 2^{-m-1} \cdot (k-1) \cdot \mu(E_n^{(m)} \cdot E_k^{(m+1)}),$$

and

$$2^{-m-1} \, k \cdot \mu(E_n^{(m)} \cdot E_k^{(m+1)}) \geqslant \Phi(E_n^{(m)} \cdot E_k^{(m+1)}) \geqslant 2^{-m} \cdot (n-1) \cdot \mu(E_n^{(m)} \cdot E_k^{(m+1)}),$$

from which it follows that $(2n-k+1) \, \mu(E_n^{(m)} \cdot E_k^{(m+1)}) \geqslant 0$, and that $(k-2n+2) \, \mu(E_n^{(m)} \cdot E_k^{(m+1)}) \geqslant 0$. Hence $\mu(E_n^{(m)} \cdot E_k^{(m+1)}) = 0$ whenever either $k > 2n+1$, or $k < 2n-2$.

We may therefore write

$$(14.9) \quad E_n^{(m)} \subset E_{2n-2}^{(m+1)} + E_{2n-1}^{(m+1)} + E_{2n}^{(m+1)} + E_{2n+1}^{(m+1)} + Q_n^{(m)} \ \text{where} \ \mu(Q_n^{(m)}) = 0.$$

This being so, let $H = \sum\limits_{m=1}^{\infty} H^{(m)} + \sum\limits_{m,n=1}^{\infty} Q_n^{(m)}$. We write $f^{(m)}(x) = 2^{-m} \cdot (n-1)$ for $x \epsilon E_n^{(m)} - H$, $n = 1, 2, ...$, and $f^{(m)}(x) = 0$ for $x \epsilon H$. We thus obtain a sequence $\{f^{(m)}(x)\}$ of non-negative functions measurable $(\mathfrak{X})$ on the set E. By (14.9) we have clearly $|f^{(m+1)}(x) - f^{(m)}(x)| \leqslant 2^{-m}$ on E, so that the sequence $\{f^{(m)}(x)\}$ converges uniformly on E to a non-negative measurable function $f(x)$.

The set H being of measure zero, we have, by (14.7) and (14.8), for every measurable set $X \subset E$ and for every positive integer m,

$$\Phi(X) \geqslant \Phi(X \cdot H) + \sum\limits_{n} 2^{-m} \cdot (n-1) \cdot \mu(X \cdot E_n^{(m)}) = \Phi(X \cdot H) + \int\limits_{X} f^{(m)} d\mu,$$

and

$$\Phi(X) \leqslant \Phi(X \cdot H) + \sum\limits_{n} 2^{-m} \cdot n \cdot \mu(X \cdot E_n^{(m)}) = \Phi(X \cdot H) + \int\limits_{X} f^{(m)} d\mu + 2^{-m} \cdot \mu(X).$$

Hence, making $m \to \infty$, we derive $\Phi(X) = \int\limits_{X} f(X) \, d\mu(x) + \Phi(X \cdot H)$.

This decomposition, so far established subject to the hypothesis that the set E is of finite measure, extends at once to sets expressible as the sum of an enumerable infinity of sets of finite measure. In fact, if $E = \sum\limits_{n=1}^{\infty} A_n$, where the A_n are sets $(\mathfrak{X})$ without common points and of finite measure, then, by what we have already proved, there exists on A_n a non-negative function $f_n(x)$ integrable on A_n, and a measurable set $H_n \subset A_n$ having measure zero, such that $\Phi(X \cdot A_n) = \int\limits_{X \cdot A_n} f_n \, d\mu + \Phi(X \cdot H_n)$, for $n = 1, 2, ...$ If we now write $H = \sum\limits_{n} H_n$, and $f(x) = f_n(x)$ for $x \epsilon A_n$, we obtain a measurable set $H \subset E$ of measure zero, and a function $f(X)$, non-negative and integrable on E, such that, by Theorem 12.7, for every measurable set $X \subset E$

$$(14.10) \quad \Phi(X) = \sum\limits_{n} \int\limits_{X \cdot A_n} f \, d\mu + \sum\limits_{n} \Phi(X \cdot H_n) = \int\limits_{X} f \, d\mu + \Phi(X \cdot H).$$

Now, the indefinite integral vanishes for every set of measure zero, and therefore is an absolutely continuous function; on the other hand, we have $\Phi(X \cdot H) = 0$ for every measurable set $X \subset E - H$. Thus, since the set H has measure zero, formula (14.10) provides a decomposition of $\Phi(X)$ into an absolutely continuous function and a singular function. Finally, to establish the unicity of such a decomposition, suppose that $\Phi(X) = \Psi_1(X) + \Theta_1(X) = \Psi_2(X) + \Theta_2(X)$ on E, the functions $\Psi_1(X)$ and $\Psi_2(X)$ being absolutely continuous, and the functions $\Theta_1(X)$ and $\Theta_2(X)$ being singular. Then $\Psi_1(X) - \Psi_2(X) = \Theta_2(X) - \Theta_1(X)$ identically on E, whence by Theorem 13.1 (2^0 and 6^0), we have $\Psi_1(X) = \Psi_2(X)$ and $\Theta_1(X) = \Theta_2(X)$, and this completes the proof of our theorem.

The expression of an additive function as the sum of an absolutely continuous function and of a singular function will be termed the *Lebesgue decomposition*. The singular function that appears in it is often called the *function of the singularities* of the given function. From Theorem 14.6, we derive at once

(14.11) *Theorem of Radon-Nikodym.* *If E is a set of finite measure, or, more generally the sum of a sequence of sets of finite measure (μ), then, in order that an additive function of a set (𝔛) on E be absolutely continuous on E, it is necessary and sufficient that this function of a set be the indefinite integral of some integrable function of a point on E.*

The hypothesis that the set E is the sum of an at most enumerable infinity of sets of finite measure, plays an essential part in the assertion of the theorem of Radon-Nikodym, just as in Theorem 9.8. To see this, let us take again the interval [0, 1] as our space X_1, and let the class $\mathfrak{X}_1$ of all subsets of [0, 1] that are measurable in the Lebesgue sense (cf. below Chap. III) be our fixed additive class of sets in the space X_1. A measure μ_1 will be defined by taking $\mu_1(X)=\infty$ for infinite sets and $\mu_1(X)=n$ for finite sets with n elements. This being so, the sets $(\mathfrak{X}_1)$ of measure (μ_1) zero coincide with the empty set, and therefore, every additive function of a set $(\mathfrak{X}_1)$ on X_1 is absolutely continuous $(\mathfrak{X}_1, \mu_1)$. In particular, denoting by $\Lambda(X)$ the Lebesgue measure for every set $X \in \mathfrak{X}_1$, we see that $\Lambda(X)$ is absolutely continuous $(\mathfrak{X}_1, \mu_1)$ on X_1. We shall show that $\Lambda(X)$ is not an indefinite integral $(\mathfrak{X}_1, \mu_1)$ on X_1. Suppose indeed, if possible, that

$$\Lambda(X) = \int_X g(x)\, d\mu_1(x)$$

for every set $X \in \mathfrak{X}_1$, the function $g(x)$ being integrable $(\mathfrak{X}_1, \mu_1)$ on X_1. Since $\Lambda(X)$ is non-negative, we may suppose that $g(x)$ is so too. Let $E = \underset{x}{\mathrm{E}}[g(x) > 0]$ and $E_n = \underset{x}{\mathrm{E}}[g(x) > 1/n]$ for $n = 1, 2, \ldots$ We have $\Lambda(X_1 - E) = \int_{X_1 - E} g\, d\mu_1 = 0$, so that $\Lambda(E) = \Lambda(X_1) = 1$ and this requires the set E to be non-enumerable. Since $E = \underset{n}{\Sigma} E_n$, the same must be true of E_{n_0} for some positive integer n_0. Thus

$$\Lambda(E_{n_0}) = \int_{E_{n_0}} g(x)\, d\mu_1(x) \geqslant \mu_1(E_{n_0})/n_0 = \infty,$$

which is evidently a contradiction.

§ 15. Change of measure. Any non-negative additive function $\nu(X)$ of a set $(\mathfrak{X})$ may clearly be regarded as a measure corresponding to the given additive class $\mathfrak{X}$. When such a function $\nu(X)$ is defined only on a set E, we can always continue it (cf. §5, p.9) on to the whole space. The terms m e a s u r e (ν), i n t e g r a l $(\mathfrak{X}, \nu)$ etc. are then completely determined for all sets $(\mathfrak{X})$, but, in this case, it is most natural to consider only the subsets of E for which the function $\nu(X)$ was originally given.

(15.1) *Theorem.* *Whenever, on a set E measurable (𝔛), we have*

$$(15.2) \qquad \nu(X) = (\mathfrak{X}) \int_X g(x)\, d\mu(x) + \vartheta(X)$$

where $\vartheta(X)$ is a non-negative function, additive and singular $(\mathfrak{X}, \mu)$ on E, and where $g(x)$ is a non-negative function integrable $(\mathfrak{X}, \mu)$ over E, then also

$$(15.3) \quad (\mathfrak{X}) \int_X f(x)\, d\nu(x) = (\mathfrak{X}) \int_X f(x)\, g(x)\, d\mu(x) + (\mathfrak{X}) \int_X f(x)\, d\vartheta(x)$$

for every set $X \subset E$ measurable $(\mathfrak{X})$ and for every function $f(x)$ that possesses a definite integral $(\mathfrak{X}, \nu)$ over X. If, further, the function $f(x)$ is integrable $(\mathfrak{X}, \nu)$ over E, the formula (15.3) expresses the Lebesgue decomposition of the indefinite integral $\int_X f\, d\nu$ on E, corresponding to the measure μ, the function $\theta(X) = \int_X f\, d\vartheta$ being the function of singularities $(\mathfrak{X}, \mu)$ of the indefinite integral $\int_X f\, d\nu$.

Proof. We may clearly assume that $f(x)$ is defined and non-negative on the whole of the set E. We see at once that, for each set $Y \subset E$ measurable $(\mathfrak{X})$, $\int_Y c_Y(x)\, d\nu(x) = \nu(Y) = \int_Y g(x)\, d\mu(x) + \vartheta(Y) = \int_Y c_Y(x)\, g(x)\, d\mu(x) + \int_Y c_Y(x)\, d\vartheta(x)$, and hence also that for every finite function $h(x)$ simple and measurable $(\mathfrak{X})$ on a set $X \subset E$,

$$(15.4) \qquad \int_X h(x)\, d\nu(x) = \int_X h(x)\, g(x)\, d\mu(x) + \int_X h(x)\, d\vartheta(x).$$

Let now $\{h_n(x)\}$ be a non-decreasing sequence of finite simple functions measurable $(\mathfrak{X})$ and non-negative on X, converging to the given function $f(x)$. Substituting $h_n(x)$ for $h(x)$ in (15.4) and making $n \to \infty$, we obtain (15.3), on account of Lebesgue's Theorem 12.6. If, further, $f(x)$ is integrable $(\mathfrak{X}, \nu)$ over E, the identity just established shows at once that the product $f(x)\, g(x)$ is integrable $(\mathfrak{X}, \mu)$ over E and hence, that the indefinite integral $\int_X f g\, d\mu$ is absolutely continuous $(\mathfrak{X}, \mu)$ on E. On the other hand, the function $\theta(X)$ vanishes on every set on which the function $\vartheta(X)$ vanishes, and therefore, is singular $(\mathfrak{X}, \mu)$ on E together with $\vartheta(X)$. This completes the proof.

The wide scope of Theorem 15.1 is due to the fact that, if $\mu(X)$ and $\nu(X)$ are any two measures associated with the same class $\mathfrak{X}$ of measurable sets and we have at the same time $\mu(E) < +\infty$, and $\nu(E) < +\infty$, for a set $E \epsilon \mathfrak{X}$, then the measure ν can be represented on E in the form (15.2), where $g(x)$ is a function integrable $(\mathfrak{X}, \mu)$ over the set E and $\vartheta(X)$ is a non-negative function, additive and singular $(\mathfrak{X}, \mu)$ on the same set (cf. Th. 14.6). Hence, with the above hypotheses and notation, in order that $\int\limits_{X} f \, d\nu = \int\limits_{X} f g \, d\mu$ should hold identically on E, it is necessary and sufficient that the indefinite integral $\int\limits_{X} f \, d\nu$ be absolutely continuous $(\mathfrak{X}, \mu)$ on E. This condition is clearly satisfied whenever the measure $\nu(X)$ is itself absolutely continuous $(\mathfrak{X}, \mu)$.