CHAPTER IV.

Derivation of additive functions of a set and of an interval.

§ 1. Introduction. In this chapter we shall study Lebesgue's theory of derivation of additive functions in a Euclidean space of any number of dimensions. When other spaces are considered, or when we specialize our space (to be, say, the straight line $\boldsymbol{R}_1$ or the plane $\boldsymbol{R}_2$), we shall say so explicitly.

In what follows, an essential part is played by Vitali's Covering Theorem (*vide*, below, § 3) which is restricted to the case of Lebesgue measure. For this reason, the theorems of the present chapter have not in general any complete or direct extension to other measures, not even when the latter are determined by additive functions of an interval. In accordance with the conventions of § 5, Chap. III, the terms measure, integral, almost everywhere, etc. will be understood in the Lebesgue sense whenever we do not explicitly assign another meaning to them. Similarly, by additive functions of a set we shall always mean functions of a set ($\mathfrak{B}$) (some of which may of course be continued on to wider classes of sets, cf. Chap. III, § 5).

We have already remarked in § 1, Chap. I, that any additive function of a set Φ in a space $\boldsymbol{R}_m$, may be regarded as a distribution of mass. It is then natural to consider the limit of the ratio $\Phi(S)/|S|$ where S denotes a cube, or a sphere, with a fixed centre a and with diameter tending to 0, as the density of the mass at the point a. By the fundamental theorem of Lebesgue (*vide*, below, Theorem 5.4) this limit exists almost everywhere. Moreover Lebesgue has shown that in the above ratio, S may be taken to denote much more general sets than cubes or spheres. Of these, further details will be given in the next §.

§ 2. Derivates of functions of a set and of an interval. Suppose given a Euclidean space R_m. By *parameter of regularity* $\mathrm{r}(E)$ of a set E lying in this space, we shall mean the lower bound of the numbers $|E|/|J|$ where J denotes any cube containing E. Thus when E is an interval, $l^m/L^m \leqslant \mathrm{r}(E) \leqslant l/L$, where l denotes the smallest and L the largest of the edges of E; in particular, the parameter of regularity of a cube is equal to 1.

A sequence of sets $\{E_n\}$ will be termed *regular*, if there exists a positive number α such that $\mathrm{r}(E_n) > \alpha$ for $n=1, 2, \ldots$.

We shall say that a sequence of sets $\{E_n\}$ *tends* to a point a, if $\delta(E_n)\to 0$ as $n\to\infty$, and the point a belongs to all the sets of the sequence.

Given a function of a set Φ (not necessarily additive) we call *general upper derivate* of Φ at a point a the upper bound of the numbers l such that there exists a regular sequence of closed sets $\{E_n\}$ tending to a, for which $\lim\limits_n \Phi(E_n)/|E_n| = l$. We shall denote this derivate by $\overline{\mathrm{D}}\Phi(a)$. Similarly, merely replacing the closed sets by intervals, we define the *ordinary upper derivate* of Φ at a point a, and we denote it by $\overline{\Phi}(a)$. If we remove the condition of regularity of the sequences of intervals considered, we obtain the definition of strong upper derivate. In other words the *strong upper derivate* of Φ at a point a is the upper limit of the ratio $\Phi(I)/|I|$, where I is any interval containing a, whose diameter tends to zero. This derivate will be denoted by $\overline{\Phi}_s(a)$.

The three *lower* derivates at a point a, general $\underline{\mathrm{D}}\Phi(a)$, ordinary $\underline{\Phi}(a)$, and strong $\underline{\Phi}_s(a)$, have corresponding definitions, and if at a point a the numbers $\underline{\mathrm{D}}\Phi(a)$ and $\overline{\mathrm{D}}\Phi(a)$ are equal, their common value is termed *general derivative* of Φ at the point a and will be denoted by $\mathrm{D}\Phi(a)$. If further $\mathrm{D}\Phi(a) \neq \infty$, the function Φ is said to be *derivable in the general sense* at the point a. Similarly we define the *ordinary derivative* $\Phi'(a)$ and the *strong derivative* $\Phi'_s(a)$, as well as *derivability in the ordinary sense*, and *in the strong sense*, of the function Φ at the point a. Sometimes the derivatives $\mathrm{D}\Phi(a)$, $\Phi'(a)$ and $\Phi_s(a)$ are termed *unique* derivates, while the upper and lower derivates (general, ordinary and strong) are termed *extreme* derivates. At any point a, we clearly have $\underline{\mathrm{D}}\Phi(a) \leqslant \underline{\Phi}(a) \leqslant \overline{\Phi}(a) \leqslant \overline{\mathrm{D}}\Phi(a)$ and similarly $\underline{\Phi}_s(a) \leqslant \underline{\Phi}(a) \leqslant \overline{\Phi}(a) \leqslant \overline{\Phi}_s(a)$; so that the existence either of a general derivative, or of a strong derivative, always implies that of an ordinary derivative. On the other hand, no such relation holds between the general and the strong extreme derivates.

It may be noted that in order that it be possible to determine the general derivates of a function of a set Φ, the latter must be defined at any rate for all closed sets; whereas in order to determine the extreme ordinary derivates, or the extreme strong derivates, we need only have the function Φ defined for the intervals. This is why the process of general derivation is most frequently applied to additive functions of a set, and that of ordinary, or of strong, derivation to functions of an interval. We shall often omit the terms "ordinary", "in the ordinary sense", in expressions such as "ordinary derivate", "derivability in the ordinary sense".

We have seen (Chap. III, § 5) that an additive function of an interval F of bounded variation determines an additive function of a set F^*. Let us mention, in the case in which the function F is non-negative, an almost evident relation between ordinary derivates of F and general derivates of F^*:

(2.1) ***Theorem.*** *If F is an additive non-negative function of an interval, then, at any point x, which belongs to no hyperplane of discontinuity of F, we have* $\underline{\mathrm{D}}F^*(x) \leqslant \underline{F}(x) \leqslant \overline{F}(x) \leqslant \overline{\mathrm{D}}F^*(x)$.

In particular therefore, the ordinary derivative $F'(x) = \mathrm{D}F^*(x)$ exists at almost every point x at which F^* has a general derivative.

Proof. Since $F(I) \leqslant F^*(I)$ for any interval I, the inequality $\overline{F}(x) \leqslant \overline{\mathrm{D}}F^*(x)$ is obvious. On the other hand, let l denote any number exceeding $\underline{F}(x)$. Then there exists a regular sequence of intervals $\{I_n\}$ tending to the point x and such that $\lim\limits_n F(I_n)/|I_n| < l$. Since x does not belong to any plane of discontinuity of F, we may asssume that it is an internal point of all the intervals I_n. Hence we can make correspond to each interval I_n an interval $J_n \subset I_n^\circ$ such that $x \in J_n$, $|J_n| \geqslant (1 - 1/n)\cdot|I_n|$ and $\mathrm{r}(J_n) = \mathrm{r}(I_n)$. We then have $\limsup\limits_n F^*(J_n)/|J_n| \leqslant \lim\limits_n F(I_n)/|I_n| < l$. Now since $\{J_n\}$ is a regular sequence of intervals tending to x, it follows that $\underline{\mathrm{D}}F^*(x) \leqslant l$, and therefore that $\underline{\mathrm{D}}F^*(x) \leqslant \underline{F}(x)$.

Let us note also the following result:

(2.2) ***Theorem.*** *If f is a summable function and Φ is the indefinite integral of f, then $\overline{\mathrm{D}}\Phi(x) \leqslant f(x)$ and $\overline{\Phi}_s(x) \leqslant f(x)$ at any point x at which the function f is upper semi-continuous, and similarly $\underline{\mathrm{D}}\Phi(x) \geqslant f(x)$ and $\underline{\Phi}_s(x) \geqslant f(x)$ at any point x at which the function f is lower semi-continuous.*

In particular therefore, $\Phi'(x) = \Phi_s'(x) = \mathrm{D}\Phi(x) = f(x)$ at any point x at which the function f is continuous.

In R_1 there is no difference between ordinary and strong derivation. If $F(x)$ is a finite function of a real variable, we understand by its *extreme derivates* $\overline{F}(x)$ and $\underline{F}(x)$, and by its *derivative*, or *unique derivate*, $F'(x)$, the corresponding derivates of the function of an interval that $F(x)$ determines (cf. Chap. III, § 13). Besides these derivates, which we shall often term *bilateral*, we also define, for functions of a real variable, unilateral derivatives and derivates. Thus, if $F(x)$ is a finite function of a real variable defined in the neighbourhood of a point x_0, the upper limit of $[F(x)—F(x_0)]/(x—x_0)$ as x tends to x_0 by values of $x>x_0$ is called *right-hand upper derivate* of the function F at the point x_0 and is denoted by $\overline{F}^+(x_0)$. Similarly we define at the point x_0 the *right-hand lower derivate* $\underline{F}^+(x_0)$ and the two *left-hand, upper* and *lower, derivates,* $\overline{F}^-(x_0)$ and $\underline{F}^-(x_0)$. These four derivates are called *unilateral extreme,* or *Dini, derivates.* If the two derivates on one side (right or left) are equal, their common value is called *unilateral* (*right-hand* or *left-hand*) *derivative* of the function F at the point in question. Finally, we shall call *intermediate derivate* of $F(x)$ at the point x_0, any number l such that there exists a sequence $\{x_n\}$ of points distinct from x_0 for which $\lim\limits_n x_n=x_0$ and $\lim\limits_n [F(x_n)—F(x_0)]/(x_n—x_0)=l$.

Let E be a linear set, x_0 a point of accumulation of E, and $F(x)$ a finite function defined on E and at the point x_0. The upper and lower limit of the ratio $[F(x)—F(x_0)]/(x—x_0)$ as x tends to x_0 by values belonging to the set E, are called respectively the *upper* and *lower derivate* of F at x_0, *relative to the set E.* We shall denote them respectively by $\overline{F}_E(x_0)$ and $\underline{F}_E(x_0)$. When they are equal, their common value is termed *derivative* of F at x_0 *relative to the set E*, and is denoted by $F'_E(x_0)$.

Besides this derivation relative to a set, we define also derivation relative to a function. Suppose given two finite functions $F(x)$ and $U(x)$, and let x_0 be a point such that the function U is not identically constant in any interval containing x_0. We then call *upper derivate* $\overline{F}_U(x_0)$ and *lower derivate* $\underline{F}_U(x_0)$ of the function F *with respect to the function U* at the point x_0, the upper limit and the lower limit of the ratio $[F(x)—F(x_0)]/[U(x)—U(x_0)]$ as x tends to x_0 by values other than those for which $F(x)—F(x_0)=$ $=U(x)—U(x_0)=0$. Similarly, considering unilateral limits of the same ratio, we define four *Dini derivates* of F *with respect to U:* $\overline{F}_U^+(x)$, $\underline{F}_U^+(x)$, $\overline{F}_U^-(x)$ and $\underline{F}_U^-(x)$.

When all these extreme derivates are equal, their common value is denoted by $F'_U(x_0)$ and called *derivative* of F *with respect to the function* U at the point x_0. The most usual case in which this method of derivation is applied, is when U is a monotone increasing function; and it is then easy, by change of variable, to reduce derivation with respect to U to ordinary derivation.

§ 3. Vitali's Covering Theorem. We shall say that a family $\mathfrak{E}$ of sets covers a set A *in the sense of Vitali*, if for every point x of A there exists a regular sequence of sets ($\mathfrak{E}$) tending to x (cf. p. 106).

(3.1) ***Vitali's Covering Theorem.*** *If in the space* $\boldsymbol{R}_m$ *a family of closed sets* $\mathfrak{E}$ *covers in the sense of Vitali a set* A, *then there exists in* $\mathfrak{E}$ *a finite or enumerable sequence* $\{E_n\}$ *of sets no two of which have common points, such that*

(3.2) $$|A-\sum_n E_n|=0.$$

Proof. a) We first prove the theorem in the special case in which (i) the parameters of regularity of all the sets ($\mathfrak{E}$) exceed a fixed number $\alpha>0$ and (ii) the set A is bounded i. e. contained in an open sphere S. We may clearly assume that, in addition, all the sets ($\mathfrak{E}$) are also contained in S.

This being so, we shall define the required sequence $\{E_n\}$ by induction in the following manner.

For E_1 we choose an arbitrary set ($\mathfrak{E}$), and when the first p sets $E_1, E_2, \ldots, E_p$ no two of which have common points, have been defined, we denote by δ_p the upper bound of the diameters of all the sets ($\mathfrak{E}$) which have no points in common with $\sum_{i=1}^{p} E_i$, and by E_{p+1} any one of these sets with diameter exceeding $\delta_p/2$. Such a set must exist, unless the sets $E_1, E_2, \ldots, E_p$ already cover the whole of the set A, in which case they constitute the sequence whose existence was to be established. We may therefore suppose that this induction can be continued indefinitely.

To show that the infinite sequence $\{E_n\}$ thus defined covers A almost entirely, let us write

(3.3) $$B=A-\sum_n E_n$$

and suppose, if possible, that $|B|>0$. On account of condition (i), we can associate with each set E_n a cube J_n such that $E_n \subset J_n$ and

$|E_n| > \alpha \cdot |J_n|$. Let $\tilde{J}_n$ denote the cube with the same centre as J_n and with diameter $(4m+1)\,\delta(J_n)$. The series

$$(3.4)\qquad \sum_n |\tilde{J}_n| = (4m+1)^m \cdot \sum_n |J_n| \leqslant (4m+1)^m \cdot \alpha^{-1} \cdot \sum_n |E_n| \leqslant (4m+1)^m \cdot \alpha^{-1}$$

converges; therefore we can find an integer N such that $\sum_{n=N+1}^{\infty} |\tilde{J}_n| < |B|$. It follows that there exists a point $x_0 \in B$ not belonging to any $\tilde{J}_n$ for $n > N$; and since by (3.3) the point x_0 does not belong to $\sum_n E_n$ and the sets E_n are supposed closed, there must exist in $\mathfrak{E}$ a set E containing x_0 and such that

$$(3.5)\qquad E \cdot E_n = 0 \qquad \textit{for} \quad n = 1, 2, \dots, N.$$

Hence the set E has common points with at least one of the sets E_n for $n > N$; for otherwise we should have $0 < \delta(E) \leqslant \delta_n \leqslant 2\delta(E_{n+1}) \leqslant \leqslant 2\delta(J_{n+1})$ for every positive integer n, and this is clearly impossible since by (3.4) we have $\lim_n \delta(J_n) = 0$. Let n_0 be the smallest of the values of n for which $E \cdot E_n \neq 0$. Then on the one hand, $E \cdot E_n = 0$ for $n = 1, 2, \dots, n_0 - 1$, so that

$$(3.6)\qquad \delta(E) \leqslant \delta_{n_0 - 1};$$

and on the other hand, by (3.5), $n_0 > N$, which implies, by definition of x_0, that x_0 does not belong to $\tilde{J}_{n_0}$. Thus we find that there are both some points outside $\tilde{J}_{n_0}$ and some points belonging to the set $E_{n_0} \subset J_{n_0}$, which are contained in the set E; this set must therefore have diameter exceeding $2\delta(J_{n_0}) \geqslant 2\delta(E_{n_0}) > \delta_{n_0-1}$, in contradiction to (3.6). The assumption that $|B| > 0$ thus leads to a contradiction and this proves the theorem, subject to the additional hypotheses (i) and (ii).

b) Now let $\mathfrak{E}$ be any family of closed sets covering the set A in the sense of Vitali; and let us denote, for any positive integer n, by S_n the sphere $S(0; n)$ and by A_n the set of the points $x \in A \cdot S_n$ for which there exists a sequence of sets tending to x and consisting of sets ($\mathfrak{E}$) whose parameters of regularity exceed $1/n$. The sets A_n constitute an ascending sequence and $A = \lim_n A_n$.

We can now define by induction a sequence of families of sets $\{\mathfrak{T}_n\}$ subject to the following conditions: 1⁰ each family $\mathfrak{T}_n$ consists of a finite number of sets ($\mathfrak{E}$) no two of which have common points and none of which, for $n > 1$, have points in common with the sets of preceding families $\mathfrak{T}_1, \mathfrak{T}_2, \dots, \mathfrak{T}_{n-1}$; 2^0 denoting by T_n the sum of the sets which belong to $\mathfrak{T}_n$,

$$(3.7) \qquad |A_n - \sum_{i=1}^{n} T_i| < 1/n.$$

To see this, suppose that for $n \leqslant p$ we have determined families $\mathfrak{T}_n$ subject to 1^0 and 2^0. Write $\widetilde{A}_{p+1} = A_{p+1} - \sum_{i=1}^{p} T_i$, and consider the family of all the sets ($\mathfrak{E}$) which are contained in the open set $C\sum_{i=1}^{p} T_i$ and whose parameters of regularity exceed $1/(p+1)$. This family evidently covers the set $\widetilde{A}_{p+1} \subset A_{p+1} \subset S_{p+1}$ in the sense of Vitali, and by what we have already proved we can extract from it a sequence $\{\widetilde{E}_i\}$ of sets, no two of which have common points, so as to cover $\widetilde{A}_{p+1}$ almost entirely. Therefore, for a sufficiently large index i_0,

$$|A_{p+1} - \sum_{i=1}^{p} T_i - \sum_{i=1}^{i_0} \widetilde{E}_i| = |\widetilde{A}_{p+1} - \sum_{i=1}^{i_0} \widetilde{E}_i| < 1/(p+1),$$

and, denoting by $\mathfrak{T}_{p+1}$ the family consisting of the sets $\widetilde{E}_1, \widetilde{E}_2, \dots, \widetilde{E}_{i_0}$, we find that conditions 1^0 and 2^0 hold when $n = p+1$.

Let us now write $\mathfrak{T} = \sum_n \mathfrak{T}_n$. The family $\mathfrak{T}$ consists of a finite number, or of an enumerable infinity, of sets ($\mathfrak{E}$) no two of which have common points. Denoting the sum of these sets by T, we find, by (3.7), $|A - T| = 0$ and this completes the proof.

The proof given above is due to S. Banach [3] (for other proofs cf. C. Carathéodory [II, pp. 299—307] and T. Radó [3]).

Theorem 3.1 was proved by G. Vitali [3] in a slightly less general form; he assumed the family $\mathfrak{E}$ to consist of cubes. H. Lebesgue [5] while retaining the line of argument of Vitali, showed that the conclusion drawn by Vitali could be generalized as follows:

(3.8) ***Theorem.*** *Given a set A and a family $\mathfrak{E}$ of closed sets, suppose that with each point $x \in A$ we can associate a number $\alpha > 0$, a sequence $\{X_n\}$ of sets ($\mathfrak{E}$) and a sequence $\{J_n\}$ of cubes such that*

$$x \in J_n, \quad X_n \subset J_n, \quad |X_n| > \alpha \cdot |J_n| \quad \textit{for} \quad n = 1, 2, \dots, \quad \textit{and} \quad \lim_n \delta(J_n) = 0.$$

Then $\mathfrak{E}$ contains a sequence of sets no two of which have common points, that covers the set A almost entirely.

This statement, although apparently more general than that of Theorem 3.1, easily reduces to the latter. For let us denote by $\mathfrak{E}^0$ the family of all the sets of the form $E+(x)$, where E is any set of $\mathfrak{E}$ and x any point of the set A. The family $\mathfrak{E}^0$ clearly covers the set A in the Vitali sense, and by Theorem 3.1 we can therefore extract from it a sequence $\{E_n^0\}$ of sets no two of which have common points and which covers the set A almost entirely. Now each set E_n^0 either already belongs to $\mathfrak{E}$, or becomes a set $(\mathfrak{E})$ as soon as we remove from it a suitably chosen point. Therefore, removing where necessary a point from each E_n^0, we obtain a sequence of sets $(\mathfrak{E})$ no two of which have common points and whose sum, since it differs from $\sum_n E_n^0$ by an at most enumerable set, covers A almost entirely.

For further generalizations of Vitali's theorem (which, again, can be proved without introducing fresh methods), see B. Jessen, J. Marcinkiewicz and A. Zygmund [1, p. 224].

It is easy to see that the hypothesis that the family $\mathfrak{E}$ covers the set A in the Vitali sense (and not merely in the ordinary sense) is essential for the validity of Theorem 3.1. But, as has been shown by S. Banach [1] and H. Bohr (*vide* C. Carathéodory [II, p. 689]), this hypothesis cannot be dispensed with in the theorem even in the case where $\mathfrak{E}$ is a family of intervals such that to each point x of the set A there corresponds a sequence $\{I_n\}$ of intervals belonging to $\mathfrak{E}$, of centre x and diameter $\delta(I_n)$ tending to zero as $n \to \infty$.

For covering theorems similar to that of Vitali and which correspond to linear measure (length, cf. Chap. II, § 9) of sets, *vide* W. Sierpiński [7], A. S. Besicovitch [1] and J. Gillis [1].

§ 4. Theorems on measurability of derivates. Of the two theorems which we shall establish in this §, the first is due to S. Banach [4, p. 174] (cf. also A. J. Ward [2, p. 177]) and concerns the extreme derivates of any function of an interval (not necessarily additive). We begin by proving the following lemma:

(4.1) ***Lemma.*** *Any set expressible as the sum of a family of intervals is measurable.*

Proof. Let $\mathfrak{J}$ be any family of intervals and let S be the sum of the intervals of $\mathfrak{J}$. Let $\mathfrak{C}$ denote the family of cubes each of which is contained in one at least of the intervals $(\mathfrak{J})$. The set S is clearly covered by $\mathfrak{C}$ in the Vitali sense and by Theorem 3.1, S is therefore expressible as the sum of a sequence of cubes $(\mathfrak{C})$ and of a set of measure zero. Therefore the set S is measurable, as asserted.

(4.2) ***Theorem.*** *If F is a function of an interval, its two extreme ordinary derivates $\overline{F}$ and $\underline{F}$ and its extreme strong derivates $\overline{F}_s$ and $\underline{F}_s$ are measurable.*

Proof. Let us take first the strong derivates of F, say $\overline{F}_s$. Let a be a finite number and P the set of the points x for which $\overline{F}_s(x) > a$. For any pair of positive integers h and k, let us denote by $P_{h,k}$ the sum of all the intervals I for which $\delta(I) \leqslant 1/k$ and $F(I)/|I| \geqslant a+1/h$. We see at once that $P = \sum_h \prod_k P_{h,k}$. Now the sets $P_{h,k}$ are measurable on account of Lemma 4.1 and so is the set P. This proves $\overline{F}_s(x)$ to be a measurable function.

Consider next the ordinary upper derivate $\overline{F}$. As before let a be any finite number, and Q the set of the points x at which $\overline{F}(x) > a$. In order that a point x should belong to Q, it is clearly necessary and sufficient that there should exist a positive number α and a sequence $\{I_n\}$ of intervals tending to x such that $\mathrm{r}(I_n) \geqslant \alpha$ and $F(I_n)/|I_n| \geqslant a + \alpha$ for $n = 1, 2, \ldots$. Hence denoting for any pair of positive integers h, k by $Q_{h,k}$ the sum of the intervals I such that $\mathrm{r}(I) \geqslant 1/h$, $\delta(I) \leqslant 1/k$ and $F(I)/|I| \geqslant a+1/h$, we find easily that $Q = \sum_h \prod_k Q_{h,k}$. Thus, since each set $Q_{h,k}$ is measurable by Lemma 4.1, so is also the set Q. The derivate $\overline{F}(x)$ is therefore measurable.

It follows in particular from Theorem 4.2 that the bilateral extreme derivates of any function of a real variable are measurable. The same is not true of unilateral extreme derivates. Nevertheless, as shown by S. Banach [2] (cf. also H. Auerbach [1]), these derivates are measurable whenever the given function is so. Similarly by a theorem of W. Sierpiński [8], the Dini derivates of a function measurable ($\mathfrak{B}$) are themselves measurable ($\mathfrak{B}$). These two results are included in the following proposition, from which they are obtained by choosing the class $\mathfrak{X}$ to be either $\mathfrak{L}$ or $\mathfrak{B}$.

(4.3) ***Theorem.*** *If $\mathfrak{X}$ is an additive class of sets in $\boldsymbol{R}_1$, which includes the sets measurable ($\mathfrak{B}$), the Dini derivates of any function of a real variable which is finite and measurable ($\mathfrak{X}$), are themselves measurable ($\mathfrak{X}$).*

Proof. If F is any finite function of a real variable, x any point and h, k any pair of positive integers subject to $k > h$, let us write $D_{h,k}(F; x)$ for the upper bound of the ratio $[F(t)-F(x)]/(t-x)$ when $x+1/k < t < x+1/h$. At any point x we clearly have

$$(4.4) \qquad \overline{F}^+(x) = \lim_h \lim_k D_{h,k}(F; x).$$

Now let a be any finite number and consider the set

$$(4.5) \qquad \mathop{\mathrm{E}}_x [D_{h,k}(F; x) > a].$$

We see at once that if the function F is constant on a set E, the set of the points x of E at which $D_{h,k}(F;x) > a$ is open in E (cf. Chap. II, p. 41). Thus the set (4.5), and consequently the expression $D_{h,k}(F;x)$ as function of x, is measurable ($\mathfrak{X}$) whenever the function F is finite, measurable ($\mathfrak{X}$) and assumes at most an enumerable infinity of distinct values.

This being so, let F be any finite function measurable ($\mathfrak{X}$). We can represent it as the limit of a uniformly convergent sequence $\{F_n\}$ of functions measurable ($\mathfrak{X}$) each of which assumes at most an enumerable infinity of distinct values: for instance we may write $F_n(x) = i/n$, when $i/n \leqslant F(x) < (i+1)/n$ for $i = \ldots -2, -1, 0, 1, 2, \ldots$. We then have $D_{h,k}(F;x) = \lim\limits_n D_{h,k}(F_n;x)$, and since by the above the functions $D_{h,k}(F_n;x)$ are measurable ($\mathfrak{X}$) in x, so is $D_{h,k}(F;x)$. It follows at once from (4.4) that the derivate $\overline{F}^+(x)$ is also measurable ($\mathfrak{X}$), and this completes the proof.

§ 5. Lebesgue's Theorem. We shall establish in this § the fundamental theorem of Lebesgue on derivation of additive functions of a set and of additive functions of bounded variation of an interval.

(5.1) ***Lemma.*** *If for a non-negative additive function of a set Φ the inequality $\overline{D}\Phi(x) \geqslant a$ holds at every point x of a set A, then*

$$\Phi(X) \geqslant a \cdot |A| \tag{5.2}$$

holds for every set $X \supset A$, bounded and measurable ($\mathfrak{B}$).

Proof. Let ε be any positive number and b any number less than a. By Theorems 6.5 and 6.10 of Chap. III, there exists a bounded open set G such that

$$X \subset G \qquad \text{and} \qquad \Phi(X) \geqslant \Phi(G) - \varepsilon. \tag{5.3}$$

Let us denote by $\mathfrak{E}$ the family of closed sets $E \subset G$ for each of which $\Phi(E) \geqslant b \cdot |E|$. Since by hypothesis, $\overline{D}\Phi(x) \geqslant a > b$ at any point $x \in A$, the family $\mathfrak{E}$ covers the set A in the Vitali sense, and by Theorem 3.1, we can extract from it a sequence $\{E_n\}$ of sets no two of which have common points, so as to cover almost entirely the set A. Therefore, on account of (5.3),

$$\Phi(X) \geqslant \Phi(G) - \varepsilon \geqslant \sum_n \Phi(E_n) - \varepsilon \geqslant b \cdot \sum_n |E_n| - \varepsilon \geqslant b \cdot |A| - \varepsilon.$$

In this we make $\varepsilon \to 0$ and $b \to a$, and (5.2) then follows at once.

(5.4) ***Lebesgue's Theorem.*** *An additive function of a set is almost everywhere derivable in the general sense. An additive function of bounded variation of an interval is almost everywhere derivable in the ordinary sense.*

Proof. On account of Theorem 2.1 we may restrict ourselves to non-negative additive functions of a set.

Let Φ be such a function and suppose that $\overline{\mathrm{D}}\Phi(x) > \underline{\mathrm{D}}\Phi(x)$ holds at each point x of a set A of positive measure. For any pair of positive integers h and k let us denote by $A_{h,k}$ the set of the points x of A for which $\overline{\mathrm{D}}\Phi(x) > (h+1)/k > h/k > \underline{\mathrm{D}}\Phi(x)$. We have $A = \sum_{h,k} A_{h,k}$, and therefore there exists a pair of integers h_0 and k_0 such that $|A_{h_0,k_0}| > 0$. Let us denote by B any bounded subset of A_{h_0,k_0} of positive outer measure. Let ε be any positive number and G a bounded open set such that

$$(5.5) \qquad B \subset G \qquad \textit{and} \qquad |G| \leqslant |B| + \varepsilon.$$

Consider the family of all closed sets $E \subset G$ for which $\Phi(E) \leqslant \frac{h_0}{k_0} \cdot |E|$. This family covers the set B in the Vitali sense, and therefore contains a sequence $\{E_n\}$ of sets no two of which have common points, which covers the set B almost entirely. Writing $Q = \sum_n E_n$ we find, on account of (5.5),

$$(5.6) \qquad \Phi(Q) = \sum_n \Phi(E_n) \leqslant \frac{h_0}{k_0} \cdot \sum_n |E_n| \leqslant \frac{h_0}{k_0} \cdot |G| \leqslant \frac{h_0}{k_0} \cdot (|B| + \varepsilon).$$

On the other hand, $\overline{\mathrm{D}}\Phi(x) \geqslant (h_0+1)/k_0$ at each point $x \in B$. Therefore, since all but a subset of measure zero of the set B is contained in the set Q, it follows from Lemma 5.1 that $\Phi(Q) \geqslant \frac{h_0+1}{k_0} \cdot |B|$, and therefore, on account of (5.6) that $(h_0+1) \cdot |B| \leqslant h_0 \cdot (|B| + \varepsilon)$. But this is clearly a contradiction since $|B| > 0$ and ε is an arbitrary positive number.

Thus the function Φ has almost everywhere a general derivative $\mathrm{D}\Phi$. It remains to prove that the latter is almost everywhere finite. Suppose the contrary: there would then exist a sphere S such that $\mathrm{D}\Phi(x) = +\infty$ at any point x of a subset of S of positive measure. We should then have, by Lemma 5.1, $\Phi(S) = +\infty$, which is impossible and completes the proof.

The preceding theorem was proved by H. Lebesgue [I, p. 128] first for continuous functions of a real variable and later [5, p. 408—425] for additive functions of a set in $\boldsymbol{R}_n$. Among the many memoirs devoted to simplifying the proof we may mention: G. Faber [1], W. H. and G. C. Young [1], H. Steinhaus [1], Ch. J. de la Vallée Poussin [1, I; p. 103], A. Rajchman and S. Saks [1], J. Ridder [3] (cf. also the direct proof of Lebesgue's theorem for additive functions of bounded variation of an interval in the first edition of this book). More recently F. Riesz [6; 7] has given an elegant proof of Lebesgue's theorem for functions of a real variable. Finally S. Banach [4, p. 177] has extended the theorem in question to a class of functions of an interval which is slightly wider than that of additive functions of bounded variation. The proof given here applies also without any essential modification to the theorem of Banach.

For an extension of the theorem to certain abstract spaces, see J. A. Clarkson [1] (cf. also S. Bochner [3]).

Another application of Lemma 5.1 is the proof of the following theorem *on term by term derivation of monotone sequences of additive functions*:

(5.7) ***Theorem.*** *If an additive function of a set Φ is the limit of a monotone sequence $\{\Phi_n\}$ of additive functions of a set, then almost everywhere* $D\Phi(x) = \lim\limits_n D\Phi_n(x)$.

In the same way, if an additive function of bounded variation of an interval, F, is the limit of a monotone sequence $\{F_n\}$ of additive functions of bounded variation of an interval, then almost everywhere $F'(x) = \lim\limits_n F_n'(x)$.

Proof. Suppose, to fix the ideas, that the sequence $\{\Phi_n\}$ is non-decreasing and write $\Theta_n = \Phi - \Phi_n$. To establish the first part of the theorem, we need only show that

$$\lim_n \overline{D}\Theta_n(x) = 0 \tag{5.8}$$

almost everywhere.

For this purpose, let A denote the set of the points x at which (5.8) is not satisfied and suppose that $|A| > 0$. For any positive integer k we write A_k for the set of the points x at which $\lim\limits_n \overline{D}\Theta_n(x) \geqslant 1/k$. Since $A = \sum\limits_k A_k$, there is an index $k = k_0$ such that $|A_{k_0}| > 0$. Let B denote any bounded subset of A_{k_0} of positive measure and I an interval containing B. Since the sequence $\{\Theta_n\}$ is non-increasing, so is the sequence $\{\overline{D}\Theta_n\}$ and therefore we must have $\overline{D}\Theta_n(x) \geqslant 1/k_0$ for $n = 1, 2, \ldots$ at any point $x \in B \subset A_{k_0}$. Hence by Lemma 5.1, we find

$\Theta_n(I) \geqslant |B|/k_0$ for every positive integer n, and this is clearly a contradiction since $|B| > 0$ and $\lim_n \Theta_n(I) = \Phi(I) - \lim_n \Phi_n(I) = 0$.

We might prove similarly the second part of the theorem, but actually the latter can be reduced at once to the first part. In fact if we suppose the given sequence $\{F_n\}$ non-decreasing and write $T_n = F - F_n$, then the functions of an interval, T_n, are non-negative and the sequence $\{T_n(I)\}$ converges to 0 in a non-increasing manner for every interval I. The sequence $\{T_n^*\}$ of additive and non-negative functions of a set is then also non-increasing and converges to 0 (cf. Chap. III, Theorem 6.2). Hence, by the first part of our theorem and by Theorem 2.1, we have $\lim_n T_n'(x) = \lim_n \mathrm{D} T_n^*(x) = 0$, and therefore $\lim_n F_n'(x) = F'(x)$, for almost all x, which completes the proof.

For functions of a real variable, Theorem 5.7 may be stated in the following form (*vide* G. Fubini [2]; cf. also L. Tonelli [3] and F. Riesz [6; 7]):

If $F(x) = \sum_n F_n(x)$ *is a convergent series of monotone non-decreasing functions, then the relation* $F'(x) = \sum_n F_n'(x)$ *holds almost everywhere.*

§ 6. Derivation of the indefinite integral. Given a set A, let us write $\mathrm{L}_A(X) = |A \cdot X|$ for every measurable set X. The function L_A of a measurable set, thus defined, is termed *measure-function* for the set A. Considered as function of a measurable set, or as function of an interval, L_A is additive and absolutely continuous; and, if further the set A is measurable, we have $\mathrm{L}_A(X) = \int_X \mathrm{c}_A(x)\,dx$ for every measurable set X, i. e. the function L_A is the indefinite integral of the characteristic function of the set A.

(6.1) ***Theorem.*** *For any set A we have*

$$\mathrm{D}\mathrm{L}_A(x) = \mathrm{c}_A(x) \tag{6.2}$$

at almost all x of A; and if further, the set A is measurable, then (6.2) *holds almost everywhere in the whole space.*

Proof. By Theorem 6.7 of Chap. III the set A can be enclosed in a set $H \in \mathfrak{G}_\delta$ for which the measure-function is the same as for the set A. Let us write $H = \prod_n G_n$, where $\{G_n\}$ is a descending sequence of open sets. We clearly have $\mathrm{DL}_{G_n}(x)=1$ at any point $x \in G_n$, and so *a fortiori* at any point $x \in A \subset H \subset G_n$. Hence, remembering that the sequence $\{\mathrm{L}_{G_n}\}$ of functions of a set is non-increasing and converges to the function $\mathrm{L}_A = \mathrm{L}_H$, it follows by Theorem 5.7 that $\mathrm{DL}_A(x) = \lim_n \mathrm{DL}_{G_n}(x) = 1 = \mathrm{c}_A(x)$ almost everywhere in A.

Now suppose the set A measurable. Then $\mathrm{L}_A(X) + \mathrm{L}_{\mathrm{C}A}(X) = |X|$ for every measurable set X, and consequently $\mathrm{DL}_A(x) + \mathrm{DL}_{\mathrm{C}A}(x)=1$ at any point x at which the two derivatives $\mathrm{DL}_A(x)$ and $\mathrm{DL}_{\mathrm{C}A}(x)$ exist, i. e. almost everywhere by Lebesgue's Theorem 5.4. Now, by what has been proved already, $\mathrm{DL}_{\mathrm{C}A}(x)=1$ almost everywhere in $\mathrm{C}A$. Therefore $\mathrm{DL}_A(x) = 0 = \mathrm{c}_A(x)$ at almost all x of $\mathrm{C}A$ and this shows that (6.2) holds almost everywhere in the whole space.

(6.3) ***Theorem.*** *If Φ is the indefinite integral of a summable function f, then*

$$\mathrm{D}\Phi(x) = f(x) \tag{6.4}$$

at almost all points x of the space.

Proof. We may clearly assume that f is a non-negative function. If f is the characteristic function of a measurable set, the relation (6.4) holds almost everywhere by the second part of Theorem 6.1. The relation therefore remains valid when f is a finite simple function, i. e. the linear combination of a finite number of characteristic functions. Finally, in the general case, any non-negative summable function f is the limit of a non-decreasing sequence $\{f_n\}$ of finite simple measurable functions; therefore, denoting by Φ_n the indefinite integral of f_n, it follows from Theorem 5.7 that $\mathrm{D}\Phi(x) = \lim_n \mathrm{D}\Phi_n(x) = \lim_n f_n(x) = f(x)$ almost everywhere, and this completes the proof.

§ 7. The Lebesgue decomposition. In this § we shall give for additive functions of a set ($\mathfrak{B}$), a more precise form to the Lebesgue Decomposition Theorem 14.6, Chap. I. We shall prove in fact that the absolutely continuous function which occurs in this theorem is the indefinite integral of the general derivative of the given function. At the same time we shall establish the corresponding decomposition for additive functions of bounded variation.

(7.1) ***Lemma.*** *If Φ is a singular additive function of a set, then* $D\Phi(x)=0$ *almost everywhere.*

Proof. We may assume (cf. Chap. I, Theorem 13.1 (1^0)) that the function Φ is non-negative.

The function Φ being singular, there exists a set E_0 measurable ($\mathfrak{B}$) and of measure zero, such that

(7.2) $\Phi(X \cdot CE_0)=0$ *for every set X bounded and measurable* ($\mathfrak{B}$).

Suppose that the set of the points x at which $D\Phi(x)>0$ has positive measure. Then denoting by Q_n the set of the points $x \in CE_0$ at which $D\Phi(x)>1/n$, there exists a positive integer N such that $|Q_N|>0$. Consequently, there also exists an interval I such that $|I \cdot Q_N|>0$, and by Lemma 5.1 we find $\Phi(I \cdot CE_0) \geqslant \Phi(I \cdot Q_N) \geqslant$ $\geqslant |I \cdot Q_N|/N>0$, which clearly contradicts (7.2).

(7.3) ***Theorem.*** *If Φ is an additive function of a set, the derivative $D\Phi$ is summable, and the function Φ is expressible as the sum of its function of singularities and of the indefinite integral of its general derivative.*

Proof. By Theorem 14.6, Chap. I, we have $\Phi=\Theta+\Psi$ where Θ is a singular additive function of a set and Ψ is the indefinite integral of a summable function f. Hence, making use of Theorem 6.3 and of Lemma 7.1 we find almost everywhere $D\Phi(x)=D\Theta(x)+D\Psi(x)=f(x)$ and this proves the theorem.

We can extend the theorem to additive functions of bounded variation of an interval. We have in fact:

(7.4) ***Theorem.*** *If F is an additive function of bounded variation of an interval, the derivative F' is summable, and the function F is the sum of a singular additive function of an interval and of the indefinite integral of the derivative F'.*

Moreover, if the function F is non-negative, we have for every interval I_0

$$F(I_0) \geqslant \int_{I_0} F'(x)\,dx, \tag{7.5}$$

equality holding only in the case in which the function F is absolutely continuous on I_0.

Proof. We may clearly assume the function F to be non-negative in both parts of the theorem. The corresponding function of a set F^*, together with its function of singularities, will then also be non-negative; and on account of Theorems 7.3 and 2.1 we shall have

$$(7.6) \qquad F^*(X) \geqslant \int_X \mathrm{D}F^*(x)\,dx = \int_X F'(x)\,dx \quad \textit{for every bounded set } X \in \mathfrak{B}.$$

Let us write for any interval I

$$(7.7) \qquad T(I) = F(I) - \int_I F'(x)\,dx.$$

The function of an interval thus defined is clearly non-negative, since by (7.6)

$$F(I) \geqslant F^*(I^\circ) \geqslant \int_{I^\circ} F'(x)\,dx = \int_I F'(x)\,dx$$

for every interval I. Moreover, if we take the derivative of both sides of (7.7) in accordance with Theorem 6.3, we find $\mathrm{D}T^*(x)=T'(x)=0$ almost everywhere. It therefore follows from Theorem 7.3 that the function of a set T^* — and so by Theorem 12.6, Chap. III, the function of an interval T — are singular. The relation (7.7) therefore provides the required decomposition for the function F.

Finally, since the function T is non-negative, it follows from (7.7) that the inequality (7.5) holds for every interval I_0, and reduces to an equality if, and only if, $T(I)=0$ for every interval $I \subset I_0$. In other words, in order that there be equality in (7.5) it is necessary and sufficient that the function F be on I_0 the indefinite integral of its derivative, i. e. be absolutely continuous on I_0.

Theorem 7.4 provides a decomposition of an additive function of bounded variation of an interval into two additive functions one of which is absolutely continuous and the other singular. Just as for functions of a set, this decomposition is termed *Lebesgue decomposition* and is uniquely determined for any additive function of bounded variation. For suppose that $G_1+T_1=G_2+T_2$ where G_1 and G_2 are absolutely continuous functions and T_1 and T_2 are singular functions; then $G_2-G_1=T_1-T_2$ and by Theorem 12.1 (2^0, 6^0), Chap. III, this requires $G_1=G_2$ and $T_1=T_2$. The absolutely continuous function and the singular function occurring in the Lebesgue decomposition of a function of bounded variation F are called, respectively, the ***absolutely continuous part*** and the ***function of singularities*** of the function F.

As a special case of Theorems 7.3 and 7.4, let us mention the following result, in which part 2^0 includes Lemma 7.1 and its converse.

(7.8) ***Theorem.*** 1^0 *An additive function of a set, or an additive function of an interval of bounded variation, is absolutely continuous, if, and only if, it is the indefinite integral of its derivative.*

2^0 *An additive function of a set, or an additive function of an interval of bounded variation, is singular if, and only if, its derivative vanishes almost everywhere.*

Finally, let us mention also an almost immediate consequence of Theorems 7.3 and 7.4:

(7.9) ***Theorem.*** *The derivative of the absolute variation of an additive function of a set, or of an additive function of an interval of bounded variation, is almost everywhere equal to the absolute value of the derivative of the given function.*

Proof. Consider to fix the ideas, an additive function of bounded variation of an interval, F. Let T be the function of singularities of F, and let W and V be the absolute variations of the functions F and T respectively. In virtue of Theorem 7.4 the relation $F(I)=\int_I F'(x)\,dx+T(I)$ holds for every interval I, and hence also

$$W(I)\leqslant\int_I |F'(x)|\,dx+V(I). \tag{7.10}$$

Now the function V is singular together with T, so that its derivative vanishes almost everywhere by Lemma 7.1. Hence taking the derivative of (7.10), we find on account of Theorem 6.3 that $W'(x)\leqslant|F'(x)|$ almost everywhere, and this completes the proof since the opposite inequality is obvious.

§ 8. Rectifiable curves. By a curve in a space $\boldsymbol{R}_m$ we shall mean any system C of m equations $x_i=X_i(t)$ where $i=1,2,\dots,m$ and the $X_i(t)$ are arbitrary finite functions defined on a linear interval or on the whole straight line $\boldsymbol{R}_1$. The variable t will be termed *parameter* of the curve. The point $(X_1(t), X_2(t), \dots, X_m(t))$ will be called *point of the curve* corresponding to the value t of the parameter, and denoted by $\mathrm{p}(C;t)$. If E is a set in $\boldsymbol{R}_1$, the set of the points $\mathrm{p}(C;t)$ for $t\in E$ will be called *graph* of the curve C on E and denoted by $\mathrm{B}(C;E)$ (cf. the similar notation for graphs of functions, Chap. III, p. 88).

For simplicity of wording we consider in the rest of this §, only curves in the plane $\boldsymbol{R}_2$; we shall suppose also that the functions determining these curves are defined in the whole straight line $\boldsymbol{R}_1$. But needless to say, these restrictions are not essential for the validity of the proofs that follow.

Let therefore C be a curve in the plane, defined by the equations $x=X(t)$, $y=Y(t)$. Given any two points a and $b>a$, a finite sequence $\tau=\{t_j\}_{j=0,1,\ldots,n}$ of points such that $a=t_0\leqslant t_1\leqslant\ldots\leqslant t_n=b$ will be called *chain* between the points a and b, and the number

$$\sum_{j=1}^{n}\varrho(p_{j-1},p_j),\qquad \textit{where}\qquad p_j=\mathrm{p}(C;t_j),$$

will be denoted by $\sigma(C;\tau)$. (We may regard this number as the length of the polygon inscribed in the curve C and whose vertices correspond to the values $a=t_0, t_1, \ldots, t_n=b$ of the parameter.) The upper bound of the numbers $\sigma(C;\tau)$ when τ is any chain between two fixed points a and b, will be called *length of arc* of the curve C on the interval $I=[a,b]$, and will be denoted by $\mathrm{S}(C;I)$ or $\mathrm{S}(C;a,b)$. If $\mathrm{S}(C;I)\neq\infty$ the curve C is said to be *rectifiable on the interval* I; and if this is the case on every interval we say simply that the curve C is *rectifiable.*

(8.1) *For any curve* C *we have* $\mathrm{S}(C;a,b)+\mathrm{S}(C;b,c)=\mathrm{S}(C;a,c)$ *whenever* $a<b<c$.

It is enough to prove that $\mathrm{S}(C;a,b)+\mathrm{S}(C;b,c)\geqslant\mathrm{S}(C;a,c)$, since the opposite inequality is obvious. Let $\tau=\{a=t_0,t_1,\ldots,t_n=c\}$ be any chain between a and c, and let h be the index for which $t_{h-1}\leqslant b<t_h$. Writing $\tau_1=\{a=t_0,t_1,\ldots,t_{h-1},b\}$ and $\tau_2=\{b,t_h,\ldots,t_n=c\}$, we have $\sigma(C;\tau)\leqslant\sigma(C;\tau_1)+\sigma(C;\tau_2)\leqslant\mathrm{S}(C;a,b)+\mathrm{S}(C;b,c)$, and so $\mathrm{S}(C;a,c)\leqslant\mathrm{S}(C;a,b)+\mathrm{S}(C;b,c)$.

It follows from (8.1) that if a curve C is rectifiable, the length of arc $\mathrm{S}(C;I)$ is an additive function of the interval I. We shall call this function *length of the curve* C. Any function of a real variable that corresponds to this function of an interval, i. e. any function $S(t)$ such that $S(b)-S(a)=\mathrm{S}(C;a,b)$ for every interval $[a,b]$, will also be termed *length of the curve* C.

(8.2) ***Jordan's Theorem.*** *If C is a curve given by the equations $x=X(t)$, $y=Y(t)$, we have*

$$\text{(8.3)}\qquad \begin{gathered} \mathrm{W}(X; I)\leqslant \mathrm{S}(C; I), \qquad \mathrm{W}(Y; I)\leqslant \mathrm{S}(C; I), \\ \mathrm{W}(X; I)+\mathrm{W}(Y; I)\geqslant \mathrm{S}(C; I) \end{gathered}$$

for any interval I; and therefore, in order that the curve C be rectifiable on an interval I_0, it is necessary and sufficient that the functions X and Y be of bounded variation on I_0.

Proof. Given an interval $I=[a, b]$, we easily find that for any chain $\tau=\{a=t_0, t_1, \dots, t_n=b\}$ between the points a and b,

$$\sum_{j=1}^{n}|X(t_j)-X(t_{j-1})|\leqslant\sigma(C; \tau), \qquad \sum_{j=1}^{n}|Y(t_j)-Y(t_{j-1})|\leqslant\sigma(C; \tau),$$

$$\sigma(C; \tau)\leqslant\sum_{j=1}^{n}|X(t_j)-X(t_{j-1})|+\sum_{j=1}^{n}|Y(t_j)-Y(t_{j-1})|\leqslant \mathrm{W}(X; I)+\mathrm{W}(Y; I),$$

from which the inequalities (8.3) follow at once.

(8.4) ***Theorem.*** *If C is a rectifiable curve given by the equations $x=X(t)$, $y=Y(t)$, and the function $S(t)$ is its length, then*

(i) *in order that $S(t)$ be continuous at a point [absolutely continuous on an interval] it is necessary and sufficient that the two functions $X(t)$ and $Y(t)$ should both be so;*

(ii) *we have $\Lambda\{\mathrm{B}(C; E)\}\leqslant|S[E]|$ for any linear set E (i.e. on any set E the length of the graph of the curve C does not exceed the measure of the set of values taken by the function $S(t)$ for $t\in E$);*

(iii) $[S'(t)]^2=[X'(t)]^2+[Y'(t)]^2$ *for almost every t;*

(iv) $\mathrm{S}(C; a, b)\geqslant\int_a^b\sqrt{[X'(t)]^2+[Y'(t)]^2}\,dt$ *for every interval $[a, b]$,*

and the sign of equality holds if, and only if, both the functions $X(t)$ and $Y(t)$ are absolutely continuous on $[a, b]$.

Proof. *re* (i): This part of the theorem is an immediate consequence of the relations (8.3), since by Theorems 4.8 and 12.1 (1^0) of Chap. III, a function of bounded variation is continuous at a point [absolutely continuous on an interval] if, and only if, its absolute variation is so.

re (ii): Let ε be any positive number and let $\{I_n\}$ be a sequence of intervals such that

(8.5) $S[E]\subset\sum_n I_n$ and (8.6) $\sum_n |I_n|\leqslant|S[E]|+\varepsilon.$

We may clearly assume that no interval I_n has a diameter exceeding ε.

This being so, we write $\widetilde{E}=\mathrm{B}(C;E)$ and we denote by $\widetilde{E}_n$ the set of the points $\mathrm{p}(C;t)$ of the curve C for each of which $S(t)\in I_n$. It is easy to see, on account of (8.5), that $\widetilde{E}\subset\sum_n \widetilde{E}_n$. On the other hand $\delta(\widetilde{E}_n)\leqslant\delta(I_n)=|I_n|\leqslant\varepsilon$ for any n, and it follows from (8.6) (cf. Chap. II, § 8) that $\Lambda^{(\varepsilon)}(\widetilde{E})\leqslant|S[E]|+\varepsilon$, and hence by making ε tend to zero, that $\Lambda(\widetilde{E})\leqslant|S[E]|$, as asserted.

re (iii): Let $I_0=[a_0,b_0]$ be any interval. We shall show successively that both the relations

(8.7) $[S'(t)]^2\geqslant[X'(t)]^2+[Y'(t)]^2$ and (8.8) $[S'(t)]^2\leqslant[X'(t)]^2+[Y'(t)]^2$

hold almost everywhere in I_0; here the derivatives $S'(t)$, $X'(t)$ and $Y'(t)$ exist almost everywhere on account of Theorem 8.2 and of Lebesgue's Theorem 5.4.

We have $S(t+h)-S(t)\geqslant\{[X(t+h)-X(t)]^2+[Y(t+h)-Y(t)]^2\}^{1/2}$ for any point t and any $h>0$, and if we divide both sides by h and make $h\to 0$, this implies the relation (8.7) at any point t for which all three functions are derivable at the same time, i. e. almost everywhere.

Now let A denote the set of the points $t\in I_0$ at which the three derivatives $X'(t)$, $Y'(t)$ and $S'(t)$ exist without satisfying the inequality (8.8); and for any positive integer n, let A_n denote the set of the points $t\in A$ for each of which the inequality

$$S(I)/|I|\geqslant\{[X(I)/|I|]^2+[Y(I)/|I|]^2\}^{1/2}+1/n$$

holds for all the intervals I containing t, whose diameters are less than $1/n$. Clearly $A=\sum_n A_n$.

Keeping n fixed for the moment, let ε be any positive number. There exists a chain $\tau=\{a_0=t_0,t_1,\dots,t_p=b_0\}$ such that $t_k-t_{k-1}<1/n$ for $k=1,2,\dots,p$, and such that $\mathrm{S}(C;I_0)\leqslant\sigma(C;\tau)+\varepsilon$. Consequently, writing for brevity $J_k=[t_{k-1},t_k]$, $p_k=\mathrm{p}(C;t_k)$, $\varrho_k=\varrho(p_{k-1},p_k)$, we have

(8.9) $S(J_k)\geqslant\varrho_k+|J_k|/n$ *for* $k=1,2,\dots,p$, *whenever* $J_k\cdot A_n\neq 0$,

and on the other hand

$$\sum_{k=1}^{p}S(J_k)\leqslant\sum_{k=1}^{p}\varrho_k+\varepsilon. \tag{8.10}$$

Therefore if $\sum_k^{(n)}$ stands for a summation over all the indices k for which $J_k \cdot A_n \neq 0$, we find on account of (8.9) and (8.10) that

$$|A_n| \leqslant \sum_k^{(n)} |J_k| \leqslant n \cdot \sum_k^{(n)} [S(J_k) - \varrho_k] \leqslant n \cdot \sum_{k=1}^{p} [S(J_k) - \varrho_k] \leqslant n\varepsilon.$$

Now since ε is an arbitrary positive number, it follows that $|A_n| = 0$ for any n, and therefore also that $|A| = 0$. Thus the inequality (8.8) holds almost everywhere, as well as the inequality (8.7), and this completes the proof of part (iii) of the theorem.

Finally, since $\mathrm{S}(C; a, b) = S(b) - S(a)$, part (iv) reduces on account of (i) and (iii) to an immediate consequence of Theorem 7.4.

Theorem 8.4 (in particular its parts (iii) and (iv)) is due to L. Tonelli [1; 4]; cf. also F. Riesz [6; 7].

As regards part (ii) of the theorem, it may be observed that in the case in which the curve C has no multiple points (i. e. when every point of the curve corresponds to a single value of the parameter) the inequality $\Lambda\{\mathrm{B}(C; E)\} \leqslant |S[E]|$ can easily be shown to reduce to an equality.

As proved by T. Ważewski [1], any bounded continuum P of finite length may be regarded as the set of the points of a curve C on the interval $[0, 1]$, such that $\mathrm{S}(C; 0,1) = 2 \cdot \Lambda(P)$.

§ 9. De la Vallée Poussin's theorem. With the help of the results established in the preceding § we can complete further, for continuous functions of bounded variation of a real variable, the decomposition formula of Lebesgue.

We shall begin with the following theorem, which itself completes, in part at any rate, the second half of Lebesgue's Theorem 5.4.

(9.1) ***Theorem.*** *If $F(x)$ is a function of bounded variation and $W(x)$ denotes its absolute variation, then for the set N of the points at which the function $F(x)$ is continuous but has no derivative finite or infinite, we have*

(9.2) $|F^*(N)| = W^*(N) = |N| = 0$ *and* (9.3) $\Lambda\{\mathrm{B}(F; N)\} = 0$.

Proof. Consider the curve C: $x = x$, $y = F(x)$. Let $S(x)$ be the length of this curve and let E be the set of the values assumed by the function $S(x)$. For any $s \in E$, denote by $X(s)$ the value of x for which $S(x) = s$ and write $Y(s) = F(X(s))$. Since (cf. Theorem 8.2) $|X(s_2) - X(s_1)| \leqslant |s_2 - s_1|$ and $|Y(s_2) - Y(s_1)| \leqslant |s_2 - s_1|$ for any pair of points s_1, s_2 of E, the functions $X(s)$ and $Y(s)$ are continuous on the set E, and moreover may be continued on to the closure $\bar{E}$ of this set by continuity. If now $[a, b]$ is an interval contiguous to $\bar{E}$, the function $X(s)$ assumes equal values at the ends of this interval, and the point $x=X(a)=X(b)$ is a point of discontinuity of the function $F(x)$. We shall complete further the definition of the functions $X(s)$ and $Y(s)$ on the whole straight line $\boldsymbol{R}_1$ so as to make the former constant and the second linear, on each interval contiguous to $\bar{E}$.

This being so, consider the curve C_1 given by the equations $x=X(s)$, $y=Y(s)$. We verify easily that the parameter s of this curve is its length. (Actually we see easily that the graph of the curve C_1 is derived from that of the curve C by adding to the latter at most an enumerable infinity of segments situated on the lines $x=c_i$ where c_i are the points of discontinuity of the function F.) By Theorem 8.4 (iii), we therefore have $[X'(s)]^2+[Y'(s)]^2=1$ for almost all s, and therefore the set H of the points s for which either one of the derivatives $X'(s)$ and $Y'(s)$ does not exist, or both exist and vanish, is of measure zero.

Now we see at once that if $s \in E-H$, then at the point $x=X(s)$, the derivative $F'(x)$ exists (with the value $Y'(s)/X'(s)$ if $X'(s)\neq 0$, or with the value $\pm\infty$ if $X'(s)=0$ and $Y'(s)\gtrless 0$). Therefore $N \subset X[E\cdot H]$, or what amounts to the same, $S[N]\subset E\cdot H$, and hence $|S[N]|\leqslant|H|=0$.

From this we derive at once with the help of Theorem 8.4 (ii) the relation (9.3), since $\mathrm{B}(F; N)=\mathrm{B}(C; N)$. Finally, the function $S(x)$ is continuous (cf. Theorem 8.4 (i)) at any point at which the function $F(x)$ is continuous, and so at any point of the set N, and therefore it follows from Theorem 13.3, Chap. III, and from Theorem 8.2, that $|F^*(N)|\leqslant W^*(N)\leqslant S^*(N)=|S[N]|=0$; this completes the proof.

(9.4) ***Lemma.*** *If $F(x)$ is a function of bounded variation, then*

(i) $F^*(A)\geqslant k\cdot|A|$ *for any bounded set A and any finite number k whenever the inequality $F'(x)\geqslant k$ holds at every point x of A (and the assertion obtained by changing the direction of both inequalities is then evidently also true),*

(ii) $F^*(B)=0$ *for any bounded set B of measure zero throughout which the derivative $F'(x)$ exists and is finite.*

Proof. *re* (i). Let ε be any positive number. Denote, for any positive integer n, by $A^{(n)}$ the set of the points $x \in A$ such that $F(I) \geqslant (k-\varepsilon)\cdot|I|$ holds whenever I is an interval containing x and of diameter less than $1/n$. Clearly $A = \lim\limits_n A^{(n)}$.

Keeping the index n fixed for the moment, let us denote by $G^{(n)}$ a bounded open set containing $A^{(n)}$, such that (cf. Theorem 6.9, Chap. III)

$$(9.5) \qquad |F^*(X) - F^*(A^{(n)})| < \varepsilon \quad \textit{whenever} \quad A^{(n)} \subset X \subset G^{(n)},$$

and let us represent $G^{(n)}$ as the sum of a sequence $\{I_p^{(n)}\}_{p=1,2,\ldots}$ of non-overlapping intervals. We may clearly suppose that all the intervals $I_p^{(n)}$ are of diameter less than $1/n$ and that their extremities are not points of discontinuity of $F(x)$. So that if $\sum\limits_p^{(n)}$ stands for summation over the indices p for which $I_p^{(n)} \cdot A^{(n)} \neq 0$, and $X^{(n)}$ denotes the sum of the intervals $I_p^{(n)}$ corresponding to these indices, we find $F^*(X^{(n)}) = \sum\limits_p^{(n)} F(I_p^{(n)}) \geqslant (k-\varepsilon)\cdot \sum\limits_p^{(n)} |I_p^{(n)}| \geqslant (k-\varepsilon)\cdot|A^{(n)}|$, and hence on account of (9.5), $F^*(A^{(n)}) \geqslant (k-\varepsilon)\cdot|A^{(n)}| - \varepsilon$. Making $\varepsilon \to 0$ and $n \to \infty$, we obtain in the limit $F^*(A) \geqslant k\cdot|A|$.

re (ii). Let $B^{(n)}$ denote the set of the points $x \in B$ for which $F'(x) \geqslant -n$. By (i) we have $F^*(B^{(n)}) \geqslant -n\cdot|B^{(n)}| = 0$ for any positive integer n, and so $F^*(B) \geqslant 0$. By symmetry we must have also $F^*(B) \leqslant 0$, and therefore $F^*(B) = 0$.

(9.6) ***De la Vallée Poussin's Decomposition Theorem.*** *If $F(x)$ is a function of bounded variation with $W(x)$ for its absolute variation, and if $E_{+\infty}$ and $E_{-\infty}$ denote the sets in which $F(x)$ has a derivative equal to $+\infty$ and to $-\infty$ respectively, then*

(i) *for any bounded set X measurable ($\mathfrak{B}$) at each point of which the function $F(x)$ is continuous, we have the relations:*

$$(9.7) \qquad F^*(X) = F^*(X\cdot E_{+\infty}) + F^*(X\cdot E_{-\infty}) + \int_X F'(x)\,dx,$$

$$(9.8) \qquad W^*(X) = F^*(X\cdot E_{+\infty}) + |F^*(X\cdot E_{-\infty})| + \int_X |F'(x)|\,dx;$$

(ii) *the two derivatives $F'(x)$ and $W'(x)$ exist and fulfil the relation $W'(x) = |F'(x)|$ at any point x of continuity of F, except at most at the points of a set N such that $W^*(N) = |N| = 0$ (i. e. a set which is at the same time of measure* (L) *zero and of measure* (W) *zero).*

Proof. *re* (i). On account of Theorem 7.3 there exists a set A of measure zero such that for any set X bounded and measurable ($\mathfrak{B}$)

$$F^*(X) = F^*(X \cdot A) + \int_X F'(x)\, dx. \tag{9.9}$$

Supposing further that the function F is continuous at every point of X, we may assume, in virtue of Theorem 9.1, that the function $F(x)$ has everywhere in A, a derivative, finite or infinite. Moreover since by Lemma 9.4 (ii) the function $F^*(X)$ vanishes for any bounded subset of $A-(E_{+\infty}+E_{-\infty})$, we may assume that $A \subset E_{+\infty}+E_{-\infty}$. Finally, we obtain directly from (9.9) that $F^*(X)$ vanishes when $X \subset CA$ and $|X|=0$. We may therefore choose simply $A=E_{+\infty}+E_{-\infty}$ and formula (9.9) becomes (9.7).

Formula (9.8) is, by Theorem 13.2, Chap. III, an immediate consequence of (9.7), since Lemma 9.4 (i) shows that $F^*(X \cdot E_{+\infty}) \geqslant 0$ and $F^*(X \cdot E_{-\infty}) \leqslant 0$ for every set X bounded and measurable ($\mathfrak{B}$).

re (ii). Let N be the set of the points x at which the function $F(x)$ is continuous and either one at least of the derivatives $F'(x)$ and $W'(x)$ does not exist, or both exist but do not satisfy the relation $W'(x)=|F'(x)|$. We then have $N \cdot E_{+\infty} = N \cdot E_{-\infty} = 0$, since evidently $W'(x) = +\infty = |F'(x)|$ at any point where $F'(x) = \pm\infty$. Therefore, since the set N is further, by Theorem 7.9, of measure (L) zero, it follows from formula (9.8) that $W^*(N)=0$ and this completes the proof.

Let us mention an immediate consequence of Theorem 9.6. *In order that a continuous function of bounded variation $F(x)$ be absolutely continuous, it is necessary and sufficient that the function of a set $F^*(X)$ should vanish identically on the set of the points at which $F(x)$ has an infinite derivative. In particular therefore, any continuous function of bounded variation which is not absolutely continuous has an infinite derivative on a non-enumerable set.*

Let us remark further that the theorems of this § cannot be extended directly to additive functions of an interval in the plane. Thus if $F(I)$ denotes the continuous singular function of an interval in $\boldsymbol{R}_2$, which for any interval I equals the length of the segment of the line $x=y$ contained in I, we have $\underline{F}(x)=0$ for every point x, so that $F'(x)=\infty$ does not hold at any point.

§ 10. Points of density for a set. Given a set E in a space $\boldsymbol{R}_m$, the strong upper and lower derivates of the measure-function of E (cf. § 6, p. 117) at a point x will be called respectively the *outer upper* and *outer lower density* of E at x. The points at which these two densities are equal to 1 are termed *points of outer density*, and the points at which they are equal to 0, *points of dispersion*, for the set E.

If the set E is measurable we suppress the word "outer" in these expressions. We see further, that if the set E is measurable, any point of density for E is a point of dispersion for CE, and vice-versa.

We shall show in this § (cf. below Theorem 10.2) that almost all points of any set E are points of outer density for E, or what amounts to the same that

(10.1) *For almost all points x of E, if $\{I_n\}$ is any sequence of intervals tending to x (in the sense of § 2, p. 106), we have* $|E \cdot I_n|/|I_n| \to 1$.

This proposition presents an obvious analogy to Theorem 6.1 and it is in the form (6.1) that the "density theorem" is often stated and proved either with the help of Vitali's Covering Theorem or by more or less equivalent means (*vide*, for instance, E. W. Hobson [I], Ch. J. de la Vallée Poussin [I, p. 71] and W. Sierpiński [10]). Theorem 10.2 will be however, so to speak, independent of Vitali's theorem, because the sequences of intervals occurring in (10.1) are not supposed regular.

Also, Theorem 10.2 will be more precise than Theorem 6.1, because at any point x of outer density for a set E we have *a fortiori* $\mathrm{DL}_E(x)=1$, and indeed, at any x the relations $\mathrm{L}'_E(x)=1$ and $\mathrm{DL}_E(x)=1$ are equivalent. To see this it is enough to show that the former of these relations implies the latter, the converse being obvious from Theorem 2.1. We may assume moreover, on account of Theorem 6.7, Chap. III, that the set E is measurable.

Let therefore x_0 be a point such that $\mathrm{L}'_E(x_0)=1$ and let $\{X_n\}$ be a regular sequence of measurable sets tending to x_0. Then there exists a sequence $\{J_n\}$ of cubes such that $X_n \subset J_n$ and $|X_n|/|J_n|>\alpha$, where $n=1, 2, \ldots$ and α is a fixed positive number. Since by hypothesis, $|E \cdot J_n|/|J_n| \to 1$, and so $|CE \cdot J_n|/|J_n| \to 0$, it follows that $|CE \cdot X_n|/|X_n| \to 0$, or what amounts to the same that $|E \cdot X_n|/|X_n| \to 1$. Therefore $\mathrm{DL}_E(x_0)=1$.

It is, of course, only for spaces $\boldsymbol{R}_m$ of dimension number $m \geqslant 2$ that Theorem 10.2 will differ from Theorem 6.1. The two statements are equivalent for $\boldsymbol{R}_1$.

For the various proofs of Theorem 10.2 *vide* F. Riesz [8] and H. Busemann and W. Feller [1]. In the second of these memoirs will be found a general discussion of the different forms of "density theorems".

It is of interest to observe that the proposition (10.1) ceases to be true even for closed sets E, if the intervals I_n are replaced by arbitrary rectangles with sides not necessarily parallel to the axes of coordinates. This remarkable fact has been established by O. Nikodym and A. Zygmund (*vide* O. Nikodym [1, p. 167]) and by H. Busemann and W. Feller [1, p. 243].

(10.2) ***Density Theorem.*** *Almost all the points of an arbitrary set E are points of outer density for E; and if further the set E is measurable, almost all the points of* CE *are points of dispersion for E.*

Proof. For simplicity of notation, we shall state the proof for sets which lie in the plane; the corresponding discussion in any space $\boldsymbol{R}_m$ is however essentially the same (except that in $\boldsymbol{R}_1$, as already remarked, the theorem reduces to Theorem 6.1).

By Theorem 6.7, Chap. III, any set can be enclosed in a measurable set having the same measure-function. On the other hand any measurable set is the sum of a set of zero measure and of a sequence of bounded closed sets. We may therefore suppose that the set considered is bounded and closed.

Let ε be any positive number. We shall begin by defining a positive number σ and a closed subset A of E such that $|E-A|<\varepsilon$ and that, for any point (ξ, η) in the plane,

$$\text{(i)} \qquad \begin{gathered} |\underset{x}{\mathrm{E}}[(x,\eta)\,\epsilon\, E;\; a\leqslant x\leqslant b]| \geqslant (1-\varepsilon)(b-a) \\ \textit{whenever} \quad (\xi,\eta)\,\epsilon\, A, \quad a\leqslant \xi\leqslant b \quad \textit{and} \quad b-a<\sigma. \end{gathered}$$

To do this, let us write for brevity, when Q is any set in the plane, $Q^{[\eta]}=\underset{x}{\mathrm{E}}[(x,\eta)\,\epsilon\, Q]$, and let us denote, for any positive integer n, by A_n the set of the points (ξ, η) of E such that $|E^{[\eta]}\cdot I| \geqslant (1-\varepsilon)\cdot|I|$ whenever I is a linear interval containing ξ and of diameter less than $1/n$. The sequence $\{A_n\}$ is evidently ascending. Let us write

$$(10.3) \qquad N = E - \lim_n A_n.$$

For any η, if ξ is a point of the set $N^{[\eta]}$, there will then exist, for each positive integer n, a linear interval I such that $\xi\,\epsilon\, I$, $\delta(I)<1/n$ and $|N^{[\eta]}\cdot I|\leqslant|E^{[\eta]}\cdot I|<(1-\varepsilon)\cdot|I|$. Therefore the lower derivate of the measure-function for the linear set $N^{[\eta]}$ cannot exceed $1-\varepsilon$ at any point of this set, whence by Theorem 6.1,

$$(10.4) \qquad |N^{[\eta]}|=|\underset{x}{\mathrm{E}}[(x,\eta)\,\epsilon\, N]|=0 \quad \textit{for every real number } \eta.$$

Let us now remark that the sets A_n are closed. For, keeping for a moment an index n fixed, let (ξ_0, η_0) be the limit of a sequence $\{(\xi_k, \eta_k)\}_{k=1,2,\ldots}$ of points of A_n. Let I_0 be a linear interval such that $\xi_0\,\epsilon\, I_0$ and $\delta(I_0)<1/n$, and let I be any linear interval containing I_0 in its interior, whose diameter is less than $1/n$. Then for any sufficiently large k, $\xi_k\,\epsilon\, I$ and therefore $|E^{[\eta_k]}\cdot I|\geqslant(1-\varepsilon)\cdot|I|$. On the other hand, the set E being closed, we easily see that $E^{[\eta_0]}\supset\limsup_k E^{[\eta_k]}$; so that $|E^{[\eta_0]}\cdot I|\geqslant\limsup_k|E^{[\eta_k]}\cdot I|\geqslant(1-\varepsilon)\cdot|I|$, and therefore also $|E^{[\eta_0]}\cdot I_0|\geqslant(1-\varepsilon)\cdot|I_0|$. Hence $(\xi_0, \eta_0)\,\epsilon\, A_n$, i.e. each set A_n is closed.

It follows, according to (10.3), that the set N is measurable.

Therefore, applying Fubini's theorem in the form (8.6), Chap. III, we conclude, on account of (10.4), that the plane set N is of measure zero. Consequently $|E-A_{n_0}|<\varepsilon$ for a sufficiently large index n_0, and writing $\sigma=1/n_0$ and $A=A_{n_0}$ we find that the inequality $|E-A|<\varepsilon$ and condition (i) are both satisfied.

In exactly the same way, but replacing the set E by A and interchanging the rôle of the coordinates x and y, we determine now a positive number $\sigma_1<\sigma$ and a closed subset B of A such that $|A-B|<\varepsilon$ and that, for any point (ξ,η) in the plane

$$\text{(ii)}\qquad |\underset{y}{\mathrm{E}}[(\xi,y)\in A;\ a\leqslant y\leqslant b]|\geqslant(1-\varepsilon)(b-a)$$
$$\textit{whenever}\quad (\xi,\eta)\in B,\quad a\leqslant\eta\leqslant b\quad \textit{and}\quad b-a<\sigma_1.$$

This being so, let (ξ_0,η_0) be any point of B. Let $J=[\alpha_1,\beta_1;\ \alpha_2,\beta_2]$ denote any interval such that $(\xi_0,\eta_0)\in J$ and $\delta(J)<\sigma_1<\sigma$. By Fubini's theorem (in the form (8.6), Chap. III) we have

$$(10.5)\qquad |E\cdot J|=\int_{\alpha_2}^{\beta_2}|\underset{x}{\mathrm{E}}[(x,y)\in E;\ \alpha_1\leqslant x\leqslant\beta_1]|\,dy.$$

Since $(\xi_0,\eta_0)\in B$ and $\alpha_2\leqslant\eta_0\leqslant\beta_2$, it follows from (ii) that the set of the y such that $(\xi_0,y)\in A$ and $\alpha_2\leqslant y\leqslant\beta_2$ is of measure at least equal to $(1-\varepsilon)(\beta_2-\alpha_2)$. On the other hand, since $\alpha_1\leqslant\xi_0\leqslant\beta_1$, it follows from (i) that $|\underset{x}{\mathrm{E}}[(x,y)\in E;\ \alpha_1\leqslant x\leqslant\beta_1]|\geqslant(1-\varepsilon)(\beta_1-\alpha_1)$ whenever $(\xi_0,y)\in A$. Hence, formula (10.5) gives

$$|E\cdot J|\geqslant(1-\varepsilon)^2(\beta_1-\alpha_1)(\beta_2-\alpha_2)=(1-\varepsilon)^2\cdot|J|,$$

i. e. the lower density of E is at least equal to $(1-\varepsilon)^2$ at any point (ξ_0,η_0) of B. Therefore, since $|E-B|\leqslant|E-A|+|A-B|\leqslant 2\varepsilon$ and since ε is an arbitrary positive number, the lower density of E is exactly equal to 1 at almost all the points of the set E, i. e. almost all points of this set are points of density for it.

The second part of the theorem is an immediate consequence of the first part. In fact, if the set E is measurable, so is the set CE, and almost all points of CE, since they are points of density for CE, are points of dispersion for E.

In connection with the definition of points of density, Denjoy introduced the important notion of approximate continuity of a function. We call a function of a point $f(x)$ (in any space $\boldsymbol{R}_m$), *approximately continuous* at a point x_0, if $f(x_0) \neq \infty$ and $f(x) \to f(x_0)$ as x tends to x_0 on a measurable set E for which x_0 is a point of density.

(10.6) ***Theorem.*** *If f is a measurable function almost everywhere finite on a set E, then the function f is approximately continuous at almost all points of E.*

Proof. On account of Lusin's theorem (Chap. III, § 7), given any $\varepsilon > 0$, we can represent the set E as the sum of a closed set F on which the function f is continuous and of a set of measure less than ε. The function f is clearly approximately continuous at any point of density for the set F, and so, by Theorem 10.2, at almost all points of F. This implies, as ε is an arbitrary positive number, that the function f is aproximately continuous at almost all points of the set E itself.

Theorem 10.6 is due to A. Denjoy [5] (cf. also W. Sierpiński [6; 9]). It is easy to see that the converse holds also, i. e. that *every function which is approximately continuous at almost all the points of a measurable set E is measurable on E* (*vide* W. Stepanoff [2] and E. Kamke [1]).

Let us mention also the following theorem, an almost immediate consequence of Theorem 10.6, which completes, in part, Theorem 2.2 on derivation of an indefinite integral:

(10.7) ***Theorem.*** *If Φ is the indefinite integral of a bounded measurable function f, then $\Phi'_s(x) = f(x)$ at almost all points x, and in fact, at any point x at which the function f is approximately continuous.*

Proof. Let x_0 be a point at which the function f is approximately continuous and let E be a measurable set for which x_0 is a point of density, while $f(x) \to f(x_0)$ as x tends to x_0 on the set E. We may suppose (by subtracting, if necessary, a constant from $f(x)$) that $f(x_0) = 0$. Therefore, given any positive number ε, we have for any interval I of sufficiently small diameter containing x_0, (i) $|I \cdot CE| \leqslant \varepsilon \cdot |I|$ and (ii) $|f(x)| < \varepsilon$ for every $x \in I \cdot E$. Denoting by M the upper bound of $|f(x)|$, conditions (i) and (ii) imply $|\Phi(I)| \leqslant |\Phi(I \cdot CE)| + |\Phi(I \cdot E)| \leqslant M\varepsilon \cdot |I| + \varepsilon \cdot |I| \leqslant \varepsilon \cdot (M+1) \cdot |I|$, whence $\Phi'_s(x_0) = 0 = f(x_0)$.

If Φ is the indefinite integral of a function f which is summable but unbounded, it may happen that the relation $\Phi'_s(x) = f(x)$ is not fulfilled at any point. In virtue of Theorem 6.3, this relation clearly holds at almost all points at which the strong derivative $\Phi'_s(x)$ exists; but if the function f is unbounded, its indefinite integral Φ may have no finite strong derivative at any point (cf. on this point Busemann and Feller [1, p. 256]; the result of Banach and Bohr mentioned above, p. 112, follows as a particular case).

Nevertheless, the result contained in Theorem 10.7 may be generalized considerably. In fact, according to a theorem of B. Jessen, J. Marcinkiewicz and A. Zygmund [1] (see below § 13) *the indefinite integral of a function f in a space $\boldsymbol{R}_m$ is almost everywhere derivable in the strong sense whenever the function* $|f|\cdot(\log^+|f|)^{m-1}$ *is summable* (in a less general form, for functions f of which the power $p>1$ is summable, this theorem was established a little earlier by A. Zygmund [1]). On the other hand however, *given an arbitrary function $\sigma(t)$ positive for $t>0$ and such that* $\liminf_{t\to+\infty}\sigma(t)=0$, *there always exists a function $f(x)$ in $\boldsymbol{R}_m$ such that the function* $\sigma(|f|)\cdot|f|\cdot(\log^+|f|)^{m-1}$ *is summable and such that the indefinite integral of f is not derivable in the strong sense (has the strong upper derivate $+\infty$) at any point of $\boldsymbol{R}_m$.*

*** § 11. Ward's theorems on derivation of additive functions of an interval.** In the preceding §§ of this Chapter, we have treated the Lebesgue theory of derivation of additive functions of an interval of bounded variation. As regards functions of a real variable, this theory has been extended to arbitrary functions by Montel, Lusin and especially by Denjoy. Recently Denjoy's theorems, which already belong to the classical results of the theory, have been generalized still further. On the one hand they have been given a geometrical form by which they become theorems on certain metrical properties of sets, and in this form an account will be given of them in Chapter IX. On the other hand, recent researches of Besicovitch and Ward have made it possible to extend an essential part of the Denjoy results, particularly the relations between the extreme bilateral derivates, to additive functions of an interval in a space $\boldsymbol{R}_m$ of any number of dimensions. These researches will form the subject of the present §.

It was A. S. Besicovitch [5] who started these researches, by establishing between the extreme strong and ordinary derivates of absolutely continuous functions of an interval, relations analogous to those proved by Denjoy for derivates of functions of a single variable. A. J. Ward [2; 5] has extended this result to quite arbitrary additive functions of an interval. Of the two theorems of Ward (*vide*, below, Theorems 11.15 and 11.21) one concerns only ordinary derivates, while the other applies also to strong derivates. It is the latter that generalizes the result of Besicovitch; this second theorem is one which can be proved fairly simply for functions of an interval in the plane; it is rather curious that it requires much more delicate methods in an arbitrary space $\boldsymbol{R}_m$.

We shall make use in this § of some auxiliary notations. If F is an additive function of an interval and $\alpha \leqslant 1$ is a positive number, $\overline{F}_{(\alpha)}(x)$ and $\underline{F}_{(\alpha)}(x)$ will denote at any point x the upper and lower limit of the ratio $F(I)/|I|$ where I is any interval containing x, which is subject to the condition $\mathrm{r}(I) \geqslant \alpha$, and which has diameter tending to zero. We see at once that at any point x, $\overline{F}_{(\alpha)}(x)$ and $\underline{F}_{(\alpha)}(x)$ tend to $\overline{F}(x)$ and $\underline{F}(x)$ respectively as $\alpha \to 0$.

We shall suppose fixed a Euclidean space $\boldsymbol{R}_m$, and in it we define a regular sequence of nets of cubes $\{\mathfrak{Q}_k\}_{k=1,2,\ldots}$, denoting by $\{\mathfrak{Q}_k\}$ the family of all the cubes of the form

$$[p_1 2^{-k}, (p_1+1)2^{-k};\quad p_2 2^{-k}, (p_2+1)2^{-k}; \ldots;\quad p_m 2^{-k}, (p_m+1)2^{-k}]$$

where $p_1, p_2, \ldots, p_m$ are arbitrary integers.

(11.1) ***Lemma.*** *Given an additive function of an interval G, and positive numbers $\alpha \leqslant 2^{-m}$ and a, suppose that the inequalities $0 < \underline{G}_{(\alpha)}(x) < a$ hold at every point x of a set E having positive outer measure; then there exists for each $\varepsilon > 0$, a cube Q which belongs to one of the nets $\mathfrak{Q}_k$ and for which we have*

(11.2) $\quad \delta(Q) < \varepsilon, \quad |E \cdot Q| > (1-\varepsilon) \cdot |Q| \quad$ *and* $\quad G(Q) < 8^m \cdot \alpha^{-m} \cdot a \cdot |Q|.$

Proof. By replacing, if necessary, the set E by a suitable subset of E having positive outer measure, we may suppose that there exists a positive number σ such that for every interval I,

(11.3) $\quad G(I) > 0 \quad$ *whenever* $\quad I \cdot E \neq 0, \quad \mathrm{r}(I) \geqslant \alpha \quad$ *and* $\quad \delta(I) < \sigma.$

We may further clearly assume that σ is less than both ε and $\alpha^m/8^m$.

This being so, let $x_0 \in E$ be a point of outer density for the set E (cf. § 10). Since $\underline{G}_{(\alpha)}(x_0) < a$, we can determine an interval $J = [a_1, b_1; \ldots; a_m, b_m]$ containing x_0 and such that

(11.4) $\quad \delta(J) < \sigma, \quad \mathrm{r}(J) \geqslant \alpha, \quad |E \cdot J| > (1-\sigma^2) \cdot |J| \quad$ *and* $\quad G(J) < a \cdot |J|.$

It follows in particular that

(11.5) $\quad |E \cdot I| > (1-\sigma) \cdot |I| \quad$ *for any interval* $\quad I \subset J \quad$ *such that* $\quad |I| > \sigma \cdot |J|.$

Let l be the smallest of the edges of J. Since $\mathrm{r}(J) \geqslant \alpha$, no edge of J can exceed l/α, and therefore we have $|J| \leqslant l^m/\alpha^m$. Finally let k be the positive integer given by

(11.6) $$1/2^k < l/4 \leqslant 1/2^{k-1},$$

and let $Q=[a'_1, b'_1; ...; a'_m, b'_m]$ be a cube which belongs to the net $\mathfrak{Q}_k$ and which contains the centre of the interval J. By (11.6) we find that $Q \subset J$ and that $a'_i - a_i \geqslant l/4$, $b_i - b'_i \geqslant l/4$ and $b'_i - a'_i = 1/2^k \geqslant l/8$. It follows easily that the figure $J \ominus Q$ can be subdivided into a finite number of non-overlapping intervals with no edge smaller than $l/8$.

Now any such interval can clearly be further subdivided into a finite number of non-overlapping subintervals whose edges all lie between $l/8$ and $l/4$. We thus obtain a subdivision of the figure $J \ominus Q$ into a finite number of non-overlapping intervals, whose parameters of regularity are greater than, or equal to, $2^{-m} \geqslant \alpha$, and whose volumes are greater than, or equal to, $8^{-m} l^m \geqslant 8^{-m} \alpha^m \cdot |J| > \sigma \cdot |J|$. It therefore follows from (11.5) and (11.3) that

(11.7) $$G(J \ominus Q) > 0.$$

Similarly, it follows from (11.6) that $|Q| = 2^{-km} \geqslant 8^{-m} l^m \geqslant \geqslant 8^{-m} \alpha^m \cdot |J| > \sigma \cdot |J|$, whence by (11.5) we derive at once the second of the relations (11.2); at the same time, by the relations (11.7) and (11.4), $G(Q) < G(J) < a \cdot |J| \leqslant 8^m \alpha^{-m} a \cdot |Q|$, and this gives the third of the relations (11.2) and completes the proof.

(11.8) ***Lemma.*** *Let G be an additive function of an interval in $\boldsymbol{R}_m$, E a set in $\boldsymbol{R}_m$, Q a cube belonging to one of the nets $\mathfrak{Q}_k$, and $\alpha > 0$, $\varepsilon > 0$ and b arbitrary fixed numbers. Suppose that*

(i) $|E \cdot Q| > (1 - \varepsilon) \cdot |Q|$,

(ii) $G(I) > 0$ *for every interval I such that $I \subset Q$, $I \cdot E \neq 0$ and* $\mathrm{r}(I) \geqslant 2^{-m}$,

(iii) $\overline{G}_{(\alpha)}(x) > b$ *at any point $x \in E$;*

then $G(Q) > 12^{-m} \cdot \alpha^m b \cdot (1 - 2^m \varepsilon) \cdot |Q|$.

Proof. We may clearly assume that the set E is contained in the interior of Q and that every point of the set is a point of density.

This being so, we shall begin by establishing the following result:

(11.9) *Given any $\eta > 0$, we can associate with any point $x \in E$ a cube P, belonging to one of the nets $\mathfrak{Q}_k$, and a cube $J \supset P$, such that* (a) $G(P) > \alpha^m \cdot 12^{-m} \cdot b \cdot |P|$ *and* (b) $x \in J$, $\delta(J) < \eta$ *and* $|J| = 3^m \cdot |P|$.

For this purpose, let us associate with the point x an interval S such that $x \in S$, $\delta(S) \leqslant 1/4$, $r(S) \geqslant \alpha$ and $G(S) > b \cdot |S|$. Let h denote the largest edge of S, and let k_1 be the positive integer satisfying the inequality $1/2^{k_1} > 2h \geqslant 1/2^{k_1+1}$. Let S_1 be a cube of the net $\mathfrak{Q}_{k_1}$ having points in common with S, and let J denote the cube formed by the 3^m cubes of the same net (including the cube S_1 itself) which have points in common with S_1.

The cube J clearly contains the interval S, and since no edge of S can be less than αh, we find that

$$(11.10) \qquad |J| = 3^m \cdot 2^{-mk_1} = 6^m \cdot 2^{-m(k_1+1)} \leqslant 12^m \cdot h^m \leqslant 12^m \cdot \alpha^{-m} \cdot |S|.$$

On the other hand, since $2^{-k_1} - h > 2^{-(k_1+1)}$ and $\alpha h \geqslant \alpha \cdot 2^{-(k_1+2)}$, the figure $J \ominus S$ can be subdivided into a finite number of non-overlapping intervals with edges greater than, or equal to, $\alpha \cdot 2^{-(k_1+2)}$, and therefore, as in the proof of Lemma 11.1, into a finite number of non-overlapping intervals whose edges have lengths between $\alpha \cdot 2^{-(k_1+2)}$ and $\alpha \cdot 2^{-(k_1+1)}$. Therefore, denoting by I any interval of this subdivision, we find $r(I) \geqslant 2^{-m}$ and $|I| \geqslant \alpha^m \cdot 2^{-m(k_1+2)} = 12^{-m} \alpha^m \cdot |J|$. Consequently, by supposing the interval S, and *a fortiori* the cube J, sufficiently small, we may assume that $\delta(J) < \eta$ and that each of the intervals of the subdivision in question contains points of E. It follows, by condition (ii) of our lemma, that $G(J \ominus S) > 0$, and so, by (11.10), that $G(J) > G(S) > b \cdot |S| \geqslant 12^{-m} \cdot \alpha^m \cdot b \cdot |J|$. Thus among the 3^m cubes of the net $\mathfrak{Q}_{k_1}$ which make up the cube J, there is one at least, P say, such that $G(P) > 12^{-m} \cdot \alpha^m \cdot b \cdot |P|$, and the cubes P and $J \supset P$, thus defined, clearly satisfy the conditions (a) and (b) of (11.9).

It now follows, on account of (11.9) and condition (i) of the lemma, that (with the help of Vitali's theorem in the form (3.8)) we can determine in Q a finite system of non-overlapping intervals $P_1, P_2, \ldots, P_n$ belonging to the nets $\mathfrak{Q}_k$, such that:

$$(11.11) \qquad G(P_i) > 12^{-m} \cdot \alpha^m b \cdot |P_i| \quad \textit{and} \quad P_i \cdot E \neq 0 \quad \textit{for} \quad i = 1, 2, \ldots, n,$$

$$(11.12) \qquad \sum_{i=1}^{n} |P_i| > (1 - \varepsilon) \cdot |Q|.$$

Among the cubes of the nets $\mathfrak{Q}_k$, we shall consider specially two classes of cubes. A cube of a net $\mathfrak{Q}_k$ contained in Q will be said to be of the first class if it is one of the cubes $P_1, P_2, \ldots, P_n$; and of the second class if it contains points of E and if further among the 2^m cubes $(\mathfrak{Q}_{k+1})$ composing it, there exists at least one which does not overlap with any cube P_i. Since the number of cubes P_i is finite, there exists a net $\mathfrak{Q}_K$ such that no cube of this net contains cubes of the first class. Let $\mathfrak{A}$ be the set of all the cubes of the first or second class contained in Q and belonging to the nets $\mathfrak{Q}_k$ for $k \leqslant K$.

The set of these cubes covers the whole cube Q. For if not, there would certainly exist in the net $\mathfrak{Q}_K$ a cube $I_0 \subset Q$ not contained in any cube $(\mathfrak{A})$. Now, since I_0 contains no cube of the first class, I_0 would not contain any point of the set E; and since, by hypothesis, I_0 is not contained in any cube of the first or second class, we could, starting with I_0, form in Q a finite ascending sequence of cubes without points in common with E and which belong respectively to the nets $\mathfrak{Q}_K, \mathfrak{Q}_{K-1}, \dots, \mathfrak{Q}_{k_0}$, where $\mathfrak{Q}_{k_0}$ is the net containing the cube Q. But the last term of this sequence of cubes is evidently the cube Q itself, and we arrive at a contradiction since $E \subset Q$.

Let us now remark that since all the cubes $(\mathfrak{A})$ belong to the nets of the regular sequence $\{\mathfrak{Q}_k\}$, it follows that, of any two overlapping cubes $(\mathfrak{A})$, one is always contained in the other. Hence, we can replace the system of cubes $\mathfrak{A}$ by another system $\mathfrak{A}_1 \subset \mathfrak{A}$ which also covers Q, and which, this time, consists of non-overlapping cubes. Let A be the sum of the cubes $(\mathfrak{A}_1)$ of the first class. On account of (11.11) we have

(11.13) $$G(A) \geqslant 12^{-m} \cdot \alpha^m \cdot b \cdot |A|.$$

Moreover, since the figure $Q \ominus A$ is formed of a finite number of cubes of the second class which do not overlap, it follows from condition (ii) of the lemma that

(11.14) $$G(Q \ominus A) \geqslant 0.$$

Finally, in each cube I of the second class, there is always a cube which is contained in $Q \ominus \sum_{i=1}^{n} P_i$ and whose volume is $2^{-m} \cdot |I|$. It therefore follows from (11.12) that $|Q \ominus A| < 2^m \cdot \varepsilon \cdot |Q|$, and in virtue of (11.14) and (11.13) we find

$$G(Q) \geqslant G(A) \geqslant 12^{-m} \cdot \alpha^m \cdot b \cdot |A| > 12^{-m} \cdot \alpha^m \cdot b \cdot (1 - 2^m \varepsilon) \cdot |Q|,$$

which completes the proof.

(11.15) ***Theorem.*** *Any additive function of an interval F is derivable at almost all the points x at which either $\underline{F}(x) > -\infty$, or $\overline{F}(x) < +\infty$.*

Proof. Consider the set of the points x at which $F(x) > -\infty$ and suppose, if possible, that the set A of the points x at which $\overline{F}(x) > \underline{F}(x) > -\infty$ is of positive measure. We could then determine a number $\alpha > 0$, and a set $B \subset A$ of positive outer measure, such that $\underline{F}_{(\alpha)}(x) \neq \infty$ and $\overline{F}_{(\alpha)}(x) - \underline{F}_{(\alpha)}(x) > \alpha$ at every point x of B. We may clearly assume that $\alpha \leqslant 2^{-m}$.

Let ε be any positive number. Let us denote for any integer p by B_p the set of the points x of B at which $p\varepsilon < \underline{F}_{(\alpha)}(x) \leqslant (p+1)\,\varepsilon$, and let p_0 be an integer such that $|B_{p_0}| > 0$. We can determine a number $\sigma > 0$ and a set $E \subset B_{p_0}$, whose measure is not zero, so as to have $F(I) > p_0\,\varepsilon \cdot |I|$ whenever the interval I is subject to the conditions $\delta(I) < \sigma$, $\mathrm{r}(I) \geqslant \alpha$ and $E \cdot I \neq 0$.

Now write $G(I) = F(I) - p_0\,\varepsilon \cdot |I|$ (where I denotes any interval). Thus defined, the function G clearly fulfils the conditions:

1° $0 < \underline{G}_{(\alpha)}(x) < 2\varepsilon$ and $\overline{G}_{(\alpha)}(x) > \alpha$ at any point $x \in E$,

2° $G(I) > 0$ for any interval I such that $\delta(I) < \sigma$, $\mathrm{r}(I) \geqslant \alpha$ and $E \cdot I \neq 0$.

By Lemma 11.1 we can therefore determine a cube Q, belonging to one of the nets $\mathfrak{Q}_k$, so as to have $\delta(Q) < \sigma$, $|E \cdot Q| > (1-\varepsilon) \cdot |Q|$ and $G(Q) \leqslant 8^m \alpha^{-m} \cdot 2\varepsilon \cdot |Q|$. From the first two of these relations and from conditions 1° and 2°, it follows, on account of Lemma 11.8, that $G(Q) > 12^{-m} \alpha^{m+1} \cdot (1 - 2^m \varepsilon) |Q|$. Thus $12^{-m} \alpha^{m+1} \cdot (1 - 2^m \varepsilon) \leqslant 8^m \alpha^{-m} \cdot 2\varepsilon$ for every $\varepsilon > 0$, and this is clearly impossible. We arrive at a contradiction and this shows that $|A| = 0$, i. e. that $\underline{F}(x) = \overline{F}(x)$ for almost all x for which $\underline{F}(x) > -\infty$.

It remains to be shown that the set of the points x at which the derivative $F'(x)$ is infinite, is of measure zero. Suppose then, if possible, that $F'(x) = +\infty$ at each point x of a set M of positive measure. We may clearly assume that there exists a number $\eta > 0$ such that $F(I) > 0$ whenever I is an interval containing points of M and subject to the conditions $\delta(I) < \eta$ and $\mathrm{r}(I) \geqslant 2^{-m}$. Therefore, denoting by R any cube which belongs to one of the nets $\mathfrak{Q}_k$ and which satisfies the relations $|M \cdot R| > (1 - 2^{-(m+1)}) \cdot |R|$ and $\delta(R) < \eta$, we find easily from Lemma 11.8 that $F(R) > 2^{-1} \cdot 12^{-m} \cdot b \cdot |R|$ for every finite number b. We thus again arrive at a contradiction and this completes the proof.

It should be remarked that for the validity of Lemma 11.1 it is enough to suppose merely that $\alpha < 1$ (instead of $\alpha \leqslant 2^{-m}$). Similarly in condition (ii) of Lemma 11.8, the inequality $\mathrm{r}(I) \geqslant 2^{-m}$ may be replaced by $\mathrm{r}(I) \geqslant \alpha$. The proofs of the lemmas remain essentially the same; we need only observe that if I is an interval whose parameter of regularity is greater than, or equal to 2^{-m}, and if $\alpha < 1$ is a positive number, the interval I can always be subdivided into a finite number of non-overlapping subintervals I_j, where $j = 1, 2, \ldots$, such that $\mathrm{r}(I_j) \geqslant \alpha$ and $|I_j| \geqslant k_\alpha \cdot |I|$, where k_α is a constant depending only on α.

We can now easily see that Theorem 11.15 may be stated in a slightly more general form as follows: *any additive function of an interval F is derivable at almost all the points x at which either $\underline{F}_{(\alpha)}(x) > -\infty$ or $\overline{F}_{(\alpha)}(x) < +\infty$, where α is any positive number less than* 1. The question whether the condition $\alpha < 1$ is necessary here, does not seem to have been solved yet completely. It may however easily be proved (by the method of nets used in the proof of Lemma 11.8) that *for any additive function F, the set of the points x for which either $\overline{F}_{(1)}(x) = -\infty$ or $\underline{F}_{(1)}(x) = +\infty$ is of measure zero.* For a discussion of these questions, *vide* the memoir of A. J. Ward [5].

We shall now proceed to prove the second theorem of Ward, in which the ordinary extreme derivates $\underline{F}(x)$ and $\overline{F}(x)$ of Theorem 11.15 are replaced by the strong derivates $\underline{F}_s(x)$ and $\overline{F}_s(x)$. It should be remarked however, that we cannot at the same time replace, in the assertion of Theorem 11.15, derivability in the ordinary sense by derivability in the strong sense: in fact, in general, a non-negative function, even when it is absolutely continuous, may yet have a strong upper derivate which is everywhere infinite (see p. 133 above).

We shall begin by proving the following lemma which is similar to Lemma 11.1.

(11.16) ***Lemma.*** *If G is an additive function of an interval in $\boldsymbol{R}_m$ and if for some fixed number a we have $0 < \underline{G}_s(x) < a$ at every point of a set E of positive outer measure, then given any $\varepsilon > 0$ there exists an interval Q such that*

(11.17) $$\delta(Q) < \varepsilon, \quad \mathrm{r}(Q) > 2^{-m}, \quad E \cdot Q \neq 0 \quad \text{and} \quad G(Q) < 3^m \cdot a \cdot |Q|.$$

Proof. Let us write for brevity $\gamma = 1/3^m$. We may suppose (by replacing, if necessary, the set E by a subset of positive outer measure) that $G(I) > 0$ for every interval I containing points of E, which has diameter less than a positive number $\sigma < \varepsilon$. Let $x_0 \epsilon E$ be a point of outer density for E and let $J = [a_1, b_1; a_2, b_2; \ldots; a_m, b_m]$ be an interval containing x_0 such that

(11.18) $\delta(J) < \sigma$, $|E \cdot J| > (1-\gamma)\cdot|J|$ *and* $G(J) < a\cdot|J|$.

Let us denote by l the smallest edge of J and by n_1 the positive integer satisfying the inequality

$$n_1 l \leqslant b_1 - a_1 < (n_1+1)\, l. \tag{11.19}$$

Writing $d_1 = (b_1 - a_1)/n_1$, let us subdivide the interval J into n_1 equal non-overlapping subintervals

$$J_i = [a_1 + (i-1)\, d_1,\ a_1 + i\, d_1;\ a_2, b_2; \dots ;\ a_m, b_m]$$

where $i = 1, 2, \dots, n_1$. We shall call an interval J_i of the first kind if $|E \cdot J_i| > (1-3\gamma)\cdot|J_i|$, and of the second kind in the opposite case. Denoting by $\sum_i'$ a summation over the indices i corresponding to intervals of the second kind, we see easily, on account of the second of the relations (11.18), that $\sum_i' 3\gamma\cdot|J_i| < \gamma\cdot|J| = \sum_{i=1}^{n_1} \gamma\cdot|J_i|$. Hence, if p and q are the number of intervals J_i of the first and second kind respectively, we find $3q < n_1 = p + q$. Now let us subdivide the interval J into a finite number of non-overlapping subintervals, in such a manner that each of these is the sum of a certain number of intervals J_i among which exactly one is of the first kind. Since $2q < p$, the intervals of this subdivision include some which coincide with certain intervals J_i of the first kind, and their number is at least equal to $p - q > n_1/3$. Thus if we denote their sum by A, we find

$$|A| > |J|/3. \tag{11.20}$$

On the other hand, the figure $J \ominus A$ consists of a finite number of non-overlapping intervals each of which contains an interval J_i of the first kind, and therefore points of E. Consequently, $G(J \ominus A) > 0$, and, on account of (11.18) and (11.20)

$$G(A) < G(J) < a\cdot|J| < 3a\cdot|A|.$$

It follows that among the intervals J_i of the first kind of which the figure A is formed, there exists one at least, J_{i_0} say, such that $G(J_{i_0}) < 3a\cdot J_{i_0}$.

Let us write, for brevity, $a_1^0 = a_1 + (i_0 - 1)\, d_1$, $b_1^0 = a_1 + i_0\, d_1$ and $J^{(1)} = J_{i_0} = [a_1^0, b_1^0; a_2, b_2; \dots; a_m, b_m]$. By the above, $J^{(1)} \subset J$, $G(J^{(1)}) < 3a\cdot|J^{(1)}|$ and, since $J^{(1)}$ coincides with an interval J_i of the first kind, $|E\cdot J^{(1)}| > (1-3\gamma)\cdot|J^{(1)}|$; finally by (11.19), $l \leqslant b_1^0 - a_1^0 = d_1 \leqslant 2l$.

If we now operate on $J^{(1)}$ just as we formerly did on J (except that we replace γ by 3γ, α by 3α and the linear interval $[a_1, b_1]$ by $[a_2, b_2]$), we obtain an interval $J^{(2)}=[a_1^0, b_1^0; a_2^0, b_2^0; a_3, b_3; \dots; a_m, b_m] \subset J^{(1)}$ such that $G(J^{(2)}) < 3^2 \cdot \alpha \cdot |J^{(2)}|$, $|E \cdot J^{(2)}| > (1-3^2\gamma) \cdot |J^{(2)}|$ and $l \leqslant b_j^0 - a_j^0 \leqslant 2l$ for $j=1, 2$.

Proceeding in this way m times, we obtain after m operations an interval $J^{(m)}=[a_1^0, b_1^0; a_2^0, b_2^0; \dots; a_m^0, b_m^0] \subset J$ such that $G(J^{(m)}) < 3^m \alpha \cdot |J^{(m)}|$, $|E \cdot J^{(m)}| > (1-3^m\gamma) \cdot |J^{(m)}|$ and $l \leqslant b_j^0 - a_j^0 \leqslant 2l$ for $j=1, 2, \dots, m$. It follows that $\mathrm{r}(J^{(m)}) \geqslant 2^{-m}$, and if we write $Q=J^{(m)}$ and substitute $\gamma=3^{-m}$, we find at once that the interval Q fulfils the conditions (11.17).

(11.21) ***Theorem.*** *If F is an additive function of an interval, we have $F'(x)=\underline{F}_s(x) \neq \infty$ $[F'(x)=\overline{F}_s(x) \neq \infty]$ at almost all the points at which $\underline{F}_s(x) > -\infty$ $[\overline{F}_s(x) < +\infty]$.*

Thus, in particular, the function F is derivable in the strong sense at almost all the points at which both the extreme strong derivates $\underline{F}_s(x)$ and $\overline{F}_s(x)$ are finite.

Proof. Since $\underline{F}(x) \geqslant \underline{F}_s(x)$ holds for all x, the function F is, by Theorem 11.15, derivable (in the ordinary sense) at almost all the points x for which $\underline{F}_s(x) > -\infty$, and we have only to show further that at almost all these points $F'(x)=\underline{F}_s(x)$. Suppose therefore that the set of the points x for which $F'(x) > \underline{F}_s(x) > -\infty$ is of positive measure. We could then determine a number $\alpha > 0$ and a set B of positive measure such that $F'(x)-\underline{F}_s(x) > \alpha$ at every point $x \in B$.

For brevity, write $\varepsilon=\alpha \cdot 3^{-(m+1)}$, and let B_p denote the set of the points $x \in B$ for which $p\varepsilon < \underline{F}_s(x) \leqslant (p+1) \cdot \varepsilon$. Let p_0 be an integer such that $|B_{p_0}| > 0$, and write $G(I)=F(I)-p_0\varepsilon \cdot |I|$ (where I is any interval). Since $G'(x) > \underline{G}_s(x)+\alpha > \alpha$ at every point $x \in B_{p_0}$, we can determine a positive number σ and a set $E \subset B_{p_0}$ of positive measure, such that $G(Q) > \alpha \cdot |Q|$ whenever Q is an interval satisfying the conditions

$$(11.22) \qquad \delta(Q) < \sigma, \quad \mathrm{r}(Q) > 2^{-m} \quad \textit{and} \quad E \cdot Q \neq 0.$$

But since $0 < \underline{G}_s(x) < 2\varepsilon$ at every point $x \in E \subset B_{p_0}$, there exists by Lemma 11.16, an interval Q subject to the conditions (11.22) and such that $G(Q) < 3^m \cdot 2\varepsilon \cdot |Q| < \alpha \cdot |Q|$. We thus arrive at a contradiction and this proves the theorem.

*§ 12. **A theorem of Hardy-Littlewood.** The theorem of Jessen, Marcinkiewicz and Zygmund concerning strong derivation of indefinite integrals, which was mentioned in § 10, p. 133, is connected with an important inequality due to G. H. Hardy and J. E. Littlewood [2]. This inequality, which was established in connection with certain problems of the theory of trigonometrical series, thus obtains a new and interesting application.

We reproduce in this § the elegant proof given by F. Riesz [5] (cf. also A. Zygmund [I, pp. 241—245]) for this inequality. Although simpler than the other proofs, it requires nevertheless some rather delicate considerations. Certain parts of the argument have been touched up in accordance with suggestions communicated to the author by Zygmund.

The reasonings of this § concern functions of a real variable.

(12.1) ***F. Riesz's lemma.*** *Let $F(x)$ be a continuous function on an interval $[a, b]$ and k a finite number. Let E be the set of the points x, interior to the interval $[a, b]$, for each of which the inequality $F(x)-F(u)>k\cdot(x-u)$ is fulfilled by at least one point u subject to $a<u<x$.*

Then the set E is either empty, or else expressible as the sum of a sequence $\{(a_n, b_n)\}$ of open non-overlapping intervals such that $F(b_n)-F(a_n)\geqslant k\cdot(b_n-a_n)$.

Proof. By subtracting from $F(x)$ the linear function kx, we may suppose that $k=0$. Then E is the set of all the points x of the open interval (a, b), for each of which there exists a point u such that $F(u)<F(x)$ and $a<u<x$. Since the function F is continuous, the set E is clearly open, — and, unless empty, it is therefore expressible as the sum of a sequence $\{(a_n, b_n)\}$ of non-overlapping open intervals. We have to prove $F(a_n)\leqslant F(b_n)$ for each n.

To see this, let us fix an index n and suppose that $F(a_n)>F(b_n)$. Let h be any number such that

$$F(a_n)>h>F(b_n), \tag{12.2}$$

and let x_0 be the lower bound of the points x of the interval $[a_n, b_n]$ for which $F(x)=h$. By (12.2) the point x_0 belongs to the open interval (a_n, b_n), and so to the set E; thus there exists a point y such that $F(y)<F(x_0)=h<F(a_n)$ and $a<y<x_0$. This last relation implies $a<y<a_n$, since, by (12.2) and by the definition of the point

x_0, the inequality $F(y) < h$ cannot hold for any y of the interval $[a_n, x_0]$. Thus $F(y) < F(a_n)$ and $a < y < a_n$, and consequently $a_n \epsilon E$; but this is clearly contradictory, since a_n is an end-point of one of the non-overlapping open intervals which constitute the set E.

Besides the results treated in this §, many other applications of Lemma 12.1 are given by F. Riesz [6; 7; 8], particularly in the theory of derivation of functions of a real variable. Cf. also S. Izumi [1]. The lemma might also have been used in the considerations of § 9 (instead of appealing to the theorems of § 8 on rectifiable curves).

To shorten our notations we shall restrict ourselves in the rest of this § to functions defined in the open interval $(0, 1)$; and we shall agree to write $\mathrm{E}[f > a]$ for $\underset{x}{\mathrm{E}}[f(x) > a;\ 0 < x < 1]$. The symbols $\mathrm{E}[f \geqslant a]$, $\mathrm{E}[b > f > a]$ and so on, will have similar meanings.

Two measurable functions g and h in $(0, 1)$ will be called (in accordance with the terminology of F. Riesz) *equi-measurable* if $|\mathrm{E}[g > a]| = |\mathrm{E}[h > a]|$ for every finite number a. We see at once that we then also have

$$|\mathrm{E}[g \geqslant a]| = |\mathrm{E}[h \geqslant a]|, \quad \mathrm{E}|[b > g \geqslant a]| = |\mathrm{E}[b > h \geqslant a]|, \quad \text{etc.}$$

(12.3) *If two non-negative measurable functions g and h in the interval* $(0, 1)$ *are equi-measurable, their definite integrals over this interval are equal.*

To see this, let us associate with the function g a non-decreasing sequence $\{g_n\}$ of simple functions by writing $g_n(x) = (k-1)/2^n$ when $(k-1)/2^n \leqslant g(x) < k/2^n$ and $k = 1, 2, \ldots, 2^n \cdot n$, and $g_n(x) = n$ when $g(x) \geqslant n$. Similarly with g replaced by h, we define the sequence $\{h_n\}$ converging to the function h. If we calculate directly the integrals of the functions g_n and h_n over $(0, 1)$ by formula 10.1 of Chap. I, p. 20, we see at once from the fact that the given functions g and h are equi-measurable that $\int_0^1 g_n(x)\,dx = \int_0^1 h_n(x)\,dx$. Making $n \to \infty$, this gives $\int_0^1 g(x)\,dx = \int_0^1 h(x)\,dx$ as asserted.

If f is a continuous function in $(0, 1)$ which is not constant on any set of positive measure, and if m and M denote the lower and upper bound of f respectively, the function $g(y) = |\mathrm{E}[f > y]|$ is evidently continuous and decreases from 1 to 0 in the open interval

(m, M). Its inverse function is therefore continuous and decreasing in $(0, 1)$ and, as we easily verify, equi-measurable with the given function.

We shall extend this process with suitable modifications, to arbitrary measurable functions finite almost everywhere in $(0, 1)$.

With any such a function $f(x)$, we associate the function $f^{\alpha}(x)$ defined for each x of $(0, 1)$ as the upper bound of the numbers y for which $|E[f>y]|>x$. The function $f^{\alpha}(x)$ is clearly finite and non-increasing in $(0, 1)$. To show that this function is equi-measurable with $f(x)$, let y_0 be any finite number, and let x_0 denote the upper bound of the set $E[f^{\alpha}>y_0]$, or else $x_0=0$ if this set is empty. Then since $|E[f^{\alpha}>y_0]|=x_0$, it has to be proved that $|E[f>y_0]|=x_0$.

We have, in the first place, $f^{\alpha}(x_0+\varepsilon)\leqslant y_0$ for every $\varepsilon>0$ (provided, of course, that $x_0+\varepsilon<1$), so that $|E[f>y_0+\varepsilon]|\leqslant x_0+\varepsilon$, and therefore $|E[f>y_0]|\leqslant x_0$. On the other hand, $f^{\alpha}(x_0-\varepsilon)>y_0$ for every $\varepsilon>0$ (provided that $x_0-\varepsilon>0$), so that $|E[f>y_0]|>x_0-\varepsilon$, whence $|E[f>y_0]|\geqslant x_0$, and finally $|E[f>y_0]|=x_0=E[f^{\alpha}>y_0]$.

We shall further define, in connection with any summable function $f(x)$, three functions $f^{\beta_1}(x)$, $f^{\beta_2}(x)$ and $f^{\beta}(x)$. At any point x of $(0, 1)$ we shall denote by $f^{\beta_1}(x)$ the upper bound of the mean values of f on the intervals (u, x) contained in $(0, x)$, i.e. the upper bound of the numbers $\dfrac{1}{x-u}\displaystyle\int_u^x f(t)\,dt$ for $0<u<x$. Similarly, $f^{\beta_2}(x)$ will denote the upper bound of the means $\dfrac{1}{v-x}\displaystyle\int_x^v f(t)\,dt$ for $x<v<1$. Finally, $f^{\beta}(x)$ will denote the larger of the two numbers $f^{\beta_1}(x)$ and $f^{\beta_2}(x)$, or what comes to the same, the upper bound of the means $\dfrac{1}{v-u}\displaystyle\int_u^v f(t)\,dt$ where u and v are subject to the condition $0<u<x<v<1$.

(12.4) ***Lemma.*** *If $f(x)$ is a non-negative measurable function in the open interval $(0, 1)$ and if E is a set contained in this interval, then*
$$\int_E f(x)\,dx \leqslant \int_0^{|E|} f^{\alpha}(x)\,dx.$$

Proof. Let f_1 be the function equal to f on the set E and to 0 elsewhere. We evidently have $f_1^\alpha(x) \leqslant f^\alpha(x)$ at each point x of the interval (0,1). Furthermore $f_1^\alpha(x)=0$ as soon as $x>|E|$. Therefore on account of (12.3) we find $\int_E f(x)\,dx=\int_0^1 f_1(x)\,dx=\int_0^1 f_1^\alpha(x)\,dx \leqslant \int_0^{|E|} f^\alpha(x)\,dx$.

(12.5) ***Lemma.*** *If $f(x)$ is a non-negative summable function in the interval (0,1), then for each point x of the interval, we have* $f^{\beta_1\alpha}(x) \leqslant \frac{1}{x}\int_0^x f^\alpha(t)\,dt$.

Proof. Let x_0 be any point in (0,1), let $y_0=f^{\beta_1\alpha}(x_0)$, and let ε denote an arbitrary positive number. We write $A=\mathrm{E}[f^{\beta_1\alpha}>y_0-\varepsilon]$ and $B=\mathrm{E}[f^{\beta_1}>y_0-\varepsilon]$. Then since the function $f^{\beta_1\alpha}$ is non-increasing, we have $|A| \geqslant x_0$ and therefore, remembering that the functions $f^{\beta_1\alpha}$ and f^{β_1} are equi-measurable, $|B|=|A| \geqslant x_0$.

Now B is the set of the points x for each of which there exists a point u subject to the conditions $\int_u^x f(t)\,dt>(y_0-\varepsilon)\cdot(x-u)$ and $0<u<x$. Therefore, applying F. Riesz's Lemma 12.1 to the indefinite integral of f, we find easily that B is an open set and that $\int_B f(t)\,dt \geqslant (y_0-\varepsilon)\cdot|B|$. It follows by Lemma 12.4 that

$$y_0-\varepsilon \leqslant \frac{1}{|B|}\int_B f(t)\,dt \leqslant \frac{1}{|B|}\int_0^{|B|} f^\alpha(t)\,dt. \tag{12.6}$$

Now since $|B| \geqslant x_0$ and since the function f^α is non-increasing, the last term of (12.6) cannot exceed $\frac{1}{x_0}\int_0^{x_0} f^\alpha(t)\,dt$; and since ε is an arbitrary positive number, we must have $f^{\beta_1\alpha}(x_0)=y_0 \leqslant \frac{1}{x_0}\int_0^{x_0} f^\alpha(t)dt$. This completes the proof.

(12.7) ***Theorem of Hardy-Littlewood.*** *If $f(x)$ is a non-negative summable function in* (0, 1) *and ε is a positive number, then*

$$\int_0^1 f^\beta(x)\,dx \leqslant A\cdot\int_0^1 f(x)\cdot\log^+ f(x)\,dx+B\cdot\int_0^1 f(x)\,dx+\varepsilon, \tag{12.8}$$

where A and B are constants depending on ε, but not on f.

Proof. We first evaluate the integral of f^{β_1} over $(0, 1)$. According to (12.3) and Lemma 12.5 we have

$$\int_0^1 f^{\beta_1}(x)\,dx = \int_0^1 f^{\beta_1 \alpha}(x)\,dx \leqslant \int_0^1 \left[\int_0^x \frac{f^\alpha(y)}{x}\,dy\right] dx.$$

In virtue of Fubini's theorem the last member of this inequality is the surface integral of the function $f^\alpha(y)/x$ over the triangle $0 \leqslant x \leqslant 1$, $0 \leqslant y \leqslant x$. Therefore inverting the order of integration in this member, we find

$$(12.9) \qquad \int_0^1 f^{\beta_1}(x)\,dx \leqslant \int_0^1 \left[\int_y^1 \frac{dx}{x}\right] f^\alpha(y)\,dy = \int_0^1 f^\alpha(y)\cdot|\log y|\,dy.$$

Let now $\eta < 1$ be a positive number such that $\int_0^\eta |\log y| / \sqrt{y}\,dy < \varepsilon/2$.

Let us denote by E_1 the set of the points y of the interval $(0, \eta)$ at which $f^\alpha(y) \leqslant 1/\sqrt{y}$, and by E_2, the set of the remaining points of this interval. We find

$$(12.10) \qquad \int_0^1 f^\alpha(y)\cdot|\log y|\,dy \leqslant \int_{E_1} \frac{|\log y|}{\sqrt{y}}\,dy + 2\int_{E_2} f^\alpha(y)\cdot\log f^\alpha(y)\,dy +$$

$$+ |\log \eta| \cdot \int_\eta^1 f^\alpha(y)\,dy \leqslant 2\int_0^1 f^\alpha(y)\cdot\log^+ f^\alpha(y)\,dy + |\log \eta|\cdot\int_0^1 f^\alpha(y)\,dy + \varepsilon/2.$$

Further, since the functions f and f^α are equi-measurable, so are the functions $f\cdot\log^+ f$ and $f^\alpha\cdot\log^+ f^\alpha$, and it therefore follows from (12.9) and (12.10) that

$$(12.11) \qquad \int_0^1 f^{\beta_1}(x)\,dx \leqslant 2\int_0^1 f(x)\cdot\log^+ f(x)\,dx + |\log\eta|\cdot\int_0^1 f(x)\,dx + \varepsilon/2.$$

A similar inequality clearly holds when on the left-hand side of (12.11) f^{β_1} is replaced by f^{β_2}, and on adding the two inequalities we find $\int_0^1 f^\beta(x)\,dx \leqslant \int_0^1 f^{\beta_1}(x)\,dx + \int_0^1 f^{\beta_2}(x)\,dx \leqslant 4\int_0^1 f(x)\cdot\log^+ f(x)\,dx +$

$+|2\log\eta|\cdot\int_0^1 f(x)\,dx + \varepsilon$; this gives (12.8) with $A = 4$ and with $B = 2|\log\eta|$.

***§ 13. Strong derivation of the indefinite integral.** We proceed to prove the theorem of Jessen, Marcinkiewicz and Zygmund. We shall give the proof for the case of the plane; its extension to spaces $\boldsymbol{R}_m$ of any number of dimensions (cf. §10, p. 133) presents no fresh difficulties and is effected by means of the well-known inequality of Jensen.

We shall begin with some auxiliary remarks. Suppose given a non-negative function $f(x, y)$ summable over the open square $J_0=(0,1;0,1)$. By Fubini's theorem, the function $f(x, y)$ is summable in x over $(0,1)$ for almost all y of $(0,1)$. Denote by H the set of these values of y. For any $y \in H$ and for any x of the interval $(0,1)$, we shall denote (cf. §12, p. 144) by $f^{\beta}(x, y)$ the upper bound of the mean $\dfrac{1}{v-u}\displaystyle\int_u^v f(t, y)\,dt$ for $0<u<x<v<1$; and whenever $y \in CH$, we shall write, for definiteness, $f^{\beta}(x, y)=0$ identically in x. We shall prove that the function $f^{\beta}(x, y)$ thus associated with any function $f(x, y)$ which is summable over the open square $J_0=(0,1;0,1)$, is measurable.

For this purpose, let a and b denote two positive numbers, and write $g_{a,b}(x, y)=\displaystyle\int_{x-a}^{x+b} f(t, y)\,dt$ when $y \in H$ and $0 \leqslant x-a<x+b \leqslant 1$, and $g_{a,b}(x, y)=0$ elsewhere in J_0. We shall begin by showing that each of the functions $g_{a,b}(x, y)$ is measurable. By Lusin's theorem, or more directly by the theorem of Vitali-Carathéodory (Chap. III, § 7), the function f is equal almost everywhere to the limit of a non-decreasing sequence $\{f^{(n)}\}$ of non-negative, bounded, upper semi-continuous functions. Now, let us put $g_{a,b}^{(n)}(x, y)=\displaystyle\int_{x-a}^{x+b} f^{(n)}(t, y)\,dt$ when $0 \leqslant x-a<x+b \leqslant 1$, and $g_{a,b}^{(n)}(x,y)=0$ elsewhere. As is easy to show (e. g. by means of Theorem 12.11, Chap. I), each of the functions $g_{a,b}^{(n)}(x, y)$ is then also upper semi-continuous, and since, as we readily see, $g_{a,b}(x, y)=\lim\limits_n g_{a,b}^{(n)}(x,y)$ almost everywhere, the function $g_{a,b}(x,y)$ is measurable.

Finally, with the same notation as above, if $\{u_p\}$ is the sequence of rational numbers of the interval $(0,1)$ we have

$$f^{\beta}(x, y)=\underset{p,q}{\text{upper bound}}\; g_{u_p,u_q}(x, y)/(u_p-u_q)$$

at any point (x, y) of J_0. Thus the function $f^\beta(x, y)$ is also measurable, and this proves our assertion.

(13.1) ***Theorem.*** *If $f(x, y)$ is a measurable function in the plane $\boldsymbol{R}_2$ and if the function $f \log^+ |f|$ is summable, then the indefinite integral of f is almost everywhere derivable in the strong sense.*

Proof. Clearly we need only consider the function f in the open square $J_0 = (0, 1; 0, 1)$; and we may also suppose that this function is non-negative.

We write $g_n(x, y) = f(x, y)$ wherever $f(x, y) \leqslant n$, and $g_n(x, y) = n$ wherever $f(x, y) > n$; we write further $h_n(x, y) = f(x, y) - g_n(x, y)$ and we denote by σ an arbitrary positive number. The functions $h_n^\beta(x, y)$ are measurable and non-negative; so that by Theorem 12.7 of Hardy-Littlewood,

$$\begin{aligned}(13.2)\qquad &\iint_{J_0} h_n^\beta(x, y)\, dxdy \leqslant \\ &\leqslant A \iint_{J_0} h_n(x, y) \cdot \log^+ h_n(x, y)\, dxdy + B \iint_{J_0} h_n(x, y)\, dxdy + \tfrac{1}{2}\sigma^2,\end{aligned}$$

where A and B are finite constants depending only on σ. And since the integrals on the right-hand side of (13.2) tend to 0 as $n \to \infty$, there exists a positive integer N such that the left-hand side of (13.2) becomes less than σ^2 for $n = N$. Therefore, writing for brevity $h(x, y) = h_N(x, y)$ and $g(x, y) = g_N(x, y)$, we have

$$(13.3)\qquad \iint_{J_0} h^\beta(x, y)\, dxdy < \sigma^2,$$

so that in particular the function $h^\beta(x, y)$, besides being measurable and non-negative, is summable on J_0.

Now denote by E the set of the points (x_0, y_0) of J_0 such that 1^0 $\int_0^1 h^\beta(x_0, t)\, dt < +\infty$, and 2^0 the indefinite integral $\int_0^y h^\beta(x_0, t)\, dt$ has at the point $y = y_0$ the derivative $h^\beta(x_0, y_0)$ with respect to y. Since, by Theorem 6.3, condition 2^0 is fulfilled for almost all y_0 of $(0, 1)$ provided that condition 1^0 is satisfied, it follows at once from Fubini's theorem in the form (8.6), Chap. III (cf. also Theorem 6.7, Chap. III) that $|E| = |J_0| = 1$.

Let us write F, H and G for the indefinite integrals of the functions f, h and g, respectively, in J_0. Let (x_0, y_0) be a point of the set E and $I = [x_0 - u_1, x_0 + u_2; y_0 - v_1, y_0 + v_2]$ any interval containing (x_0, y_0) and contained in J_0. We have

$$\frac{H(I)}{|I|}=\frac{1}{v_2+v_1}\int_{-v_1}^{v_2}\left[\frac{1}{u_2+u_1}\int_{-u_1}^{u_2}h(x_0+u,\, y_0+v)\, du\right]dv \leqslant$$
$$\leqslant \frac{1}{v_2+v_1}\int_{-v_1}^{v_2}h^{\beta}(x_0,\, y_0+v)\, dv,$$

whence making $\delta(I)\to 0$ we obtain $\overline{H}_s(x_0, y_0)\leqslant h^{\beta}(x_0, y_0)$. Thus, since (x_0, y_0) is an arbitrary point of the set $E\subset J_0$ of outer measure 1, and since the extreme derivate $\overline{H}_s(x_0, y_0)$ is measurable (cf. Theorem 4.2), it follows from (13.3) that $0\leqslant \underline{H}_s(x, y)\leqslant \overline{H}_s(x, y)\leqslant\sigma$ at every point (x, y) of J_0, except at most a set of measure less than σ. On the other hand, since the function $g=g_N$ is bounded, its indefinite integral G is, by Theorem 10.7, derivable in the strong sense almost everywhere. Therefore $\overline{F}_s(x, y)-\underline{F}_s(x, y)\leqslant\sigma$ at all but a subset of measure σ of the points of J_0; and so finally, since σ is an arbitrary positive number, $\overline{F}_s(x, y)=\underline{F}_s(x, y)$ almost everywhere in J_0, which completes the proof.

By Ward's Theorem 11.21, to prove that the non-negative function F is almost everywhere derivable in the strong sense, it is enough to show that $\overline{F}_s(x)<+\infty$ almost everywhere. Hence by using Theorem 11.21, the proof of Theorem 13.1 might be slightly shortened.

*§ 14. Symmetrical derivates.

If Φ is an additive function of a set in a space $\boldsymbol{R}_m$, we shall denote by $\overline{\mathrm{D}}_{\mathrm{sym}}\,\Phi(x)$ the *upper*, and by $\underline{\mathrm{D}}_{\mathrm{sym}}\,\Phi(x)$ the *lower, symmetrical derivate* of Φ at a point x, these being defined respectively as the upper, and as the lower, limit of the ratio $\Phi(S)/|S|$ where S represents a closed sphere of centre x and of radius tending to zero. It is obvious that, for any point x whatsoever $\overline{\mathrm{D}}\Phi(x)\geqslant\overline{\mathrm{D}}_{\mathrm{sym}}\,\Phi(x)\geqslant\underline{\mathrm{D}}_{\mathrm{sym}}\,\Phi(x)\geqslant\underline{\mathrm{D}}\Phi(x)$.

Following A. J. Ward [5], we shall establish a decomposition theorem in terms of symmetrical derivates, which is similar to Theorem 9.6. We shall begin by the following "covering theorem":

(14.1) ***Theorem.*** *If Φ is an additive function of a set in $\boldsymbol{R}_m$ and E a bounded set measurable* ($\mathfrak{B}$), *contained in an open set G, then for any $\varepsilon>0$ there exists in G an enumerable sequence of closed spheres $\{S_k\}$ such that* (i) *the centre of each S_k belongs to E and the radius is less than ε,* (ii) *$S_i\cdot S_j=0$ whenever $i\neq j$, and* (iii) *the spheres S_k cover together the whole of the set E, with the possible exception of a subset on which the function Φ vanishes identically.*

Proof. We can clearly assume (by replacing, if necessary, the function Φ by its absolute variation) that the function Φ is monotone non-negative.

a) We shall first prove that, with the hypotheses of the theorem, there always exists in G a finite system of equal spheres $\{S_k\}$ which sasisfy the conditions (i) and (ii) and cover the set E except perhaps for a set $T \subset E$ such that

$$\Phi(T) \leqslant (1-1/4^{m+1} m^m) \cdot \Phi(E). \tag{14.2}$$

To see this, let A be a subset of E, measurable $(\mathfrak{B})$, such that $\Phi(A) \geqslant \frac{1}{2}\Phi(E)$ and $\varrho(A, CG) > 0$. Let n_0 be a positive integer such that $m/n_0 < \varrho(A, CG)$ and $m/n_0 < \varepsilon$.

Denote by $\mathfrak{P}$ the net in the space $\boldsymbol{R}_m$, which consists of the cubes of the form $[p_1/n_0, (p_1+1)/n_0; p_2/n_0, (p_2+1)/n_0; \ldots, p_m/n_0, (p_m+1) n_0]$ where $p_1, p_2, \ldots, p_m$ are arbitrary integers. We can clearly subdivide the net $\mathfrak{P}$ into $(4m)^m$ families of cubes, $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_{(4m)^m}$ say, such that the distance between any two cubes belonging to the same family is not less than $(4m-1)/n_0$. Denote, for each $k=1, 2, \ldots, (4m)^m$, by A_k the part of the set A covered by the cubes of the family $\mathfrak{P}_k$. Then there exists a positive integer $k_0 \leqslant (4m)^m$ such that

$$\Phi(A_{k_0}) \geqslant \Phi(A)/(4m)^m \geqslant \Phi(E)/4^{m+1} m^m. \tag{14.3}$$

Now let $P_1, P_2, \ldots, P_r$ be those cubes of $\mathfrak{P}_{k_0}$ which contain points of A_{k_0}. With each P_i we can associate a closed sphere S_i, of radius m/n_0, whose centre belongs to $A_{k_0} \cdot P_i$. The system of spheres $S_1, S_2, \ldots, S_r$ thus defined is contained in G and clearly satisfies the conditions (i) and (ii) of the theorem. Again, since $P_i \subset S_i$ for every $i=1, 2, \ldots, r$, the spheres S_i cover the whole of the set E with the possible exception of the points of the set $T=E-A_{k_0}$, which, in virtue of (14.3), fulfils the condition (14.2)

b) We now pass on to the proof of the assertion of the theorem. By what has already been proved, we can define by induction a sequence $\{\mathfrak{S}_n\}_{n=1,2,\ldots}$ of finite systems of closed spheres with centres in E and radii less than ε, subject to the following two conditions: 1^0 If B_0 denotes the empty set and B_n, for $n \geqslant 1$, the sum of the spheres belonging to $\mathfrak{S}_1 + \mathfrak{S}_2 + \ldots + \mathfrak{S}_n$, then the system $\mathfrak{S}_{n+1}$, where $n \geqslant 0$, consists of a finite number of closed spheres, contained in the open set $G-B_n$, no two of which have common points; 2^0 $\Phi(E-B_{n+1}) \leqslant (1-h_m) \cdot \Phi(E-B_n)$ where $h_m = 1/4^{m+1} m^m$ and $n=0, 1, \ldots$ Now, arranging the spheres belonging to the family

$\mathfrak{S}_1+\mathfrak{S}_2+...+\mathfrak{S}_n+...$ in a sequence $\{S_i\}$, we see at once that the latter fulfils conditions (i) and (ii) of the theorem. On the other hand, by 2^0 we have $\Phi(E-B_n)\leqslant(1-h_m)^n\cdot\Phi(E)$ for each n; whence, denoting by B the sum of all the spheres S_i, it follows that $\Phi(E-B)=0$, which establishes condition (iii) and completes the proof.

Theorem 14.1 may be established in a slightly more general form:

Given a bounded set E measurable ($\mathfrak{B}$), a sequence of positive numbers $\{r_n\}$ converging to 0 and a family of closed sets $\mathfrak{A}$, suppose that with each point x of E there are associated two finite numbers $\alpha=\alpha(x)$, $N=N(x)$, and a sequence $\{A_n(x)\}$ of sets ($\mathfrak{A}$) such that $\mathrm{S}(x;r_n)\subset A_n(x)\subset\mathrm{S}(x;\alpha r_n)$ for $n>N(x)$.

Then, for any sequence $\{\Phi_n\}$ of additive functions of a set, we can extract from $\mathfrak{A}$ a sequence of sets $\{A_i=A_{n_i}(x_i)\}$ such that (i) $x_i\in E$ for $i=1,2,...$, (ii) $A_i\cdot A_j=0$ whenever $i\neq j$, and (iii) the sets A_i cover the whole of the set E, with the exception at most of a set of measure zero on which all the functions Φ_n vanish identically.

(14.4) ***Lemma.*** *If Φ is an additive function of a set in $\boldsymbol{R}_m$, and if $\underline{\mathrm{D}}_{\mathrm{sym}}\Phi(x)>0$ at each point x of a bounded set X measurable ($\mathfrak{B}$), then $\Phi(X)\geqslant 0$.*

Proof. Let us denote, for every positive integer n, by X_n the set of the points $x\in X$ such that $\Phi(S)\geqslant 0$ whenever S is a closed sphere of centre x and radius less than $1/n$. Each set X_n is evidently measurable ($\mathfrak{B}$), in fact closed in X. Hence, for any $\varepsilon>0$, we can associate with each X_n an open set $G_n\supset X_n$ such that $\mathrm{W}(\Phi;G_n-X_n)\leqslant\varepsilon$ (cf. Theorems 6.9 and 6.10, Chap. III). Next, keeping n fixed for the moment, we can (on account of Theorem 14.1) define in G_n a sequence $\{S_k\}$ of closed spheres with centres in X_n and radii less than $1/n$, such that (i) $S_i\cdot S_j=0$ whenever $i\neq j$, and (ii) the spheres S_k cover the whole of the set X_n with the exception at most of a set T on which the function Φ vanishes identically. Since $\Phi(S_k)\geqslant 0$ for every k, we find by (i) and (ii) that $\Phi(X_n)\geqslant -[\mathrm{W}(\Phi;T)+\mathrm{W}(\Phi;G_n-X_n)]\geqslant-\varepsilon$. Hence, as $X=\lim_n X_n$ and ε is an arbitrary positive number, it follows that $\Phi(X)\geqslant 0$, which completes the proof.

(14.5). ***Theorem.*** *If Φ is an additive function of a set in $\boldsymbol{R}_m$, and if A_∞ denotes the set of x at which one at least of the derivates $\overline{\mathrm{D}}_{\mathrm{sym}}\Phi(x)$ and $\underline{\mathrm{D}}_{\mathrm{sym}}\Phi(x)$ is infinite, then for any bounded set X measurable ($\mathfrak{B}$), we have*

$$\Phi(X)=\Phi(X\cdot A_\infty)+\int_X \mathrm{D}\Phi(x)\,dx. \tag{14.6}$$

Consequently, if $\underline{\mathrm{D}}_{\mathrm{sym}}\Phi(x)>-\infty$ at every point x and $\mathrm{D}\Phi(x)\geqslant 0$ at almost every point x of a bounded set X measurable ($\mathfrak{B}$), then $\Phi(X)\geqslant 0$.

Proof. We firstly remark that if $-\infty<\underline{\mathrm{D}}_{\mathrm{sym}}\Phi(x)$ at each point x of a bounded set Q measurable ($\mathfrak{B}$) and of measure zero, then $\Phi(Q)\geqslant 0$. In fact, denoting for each positive integer n by Q_n the set of the points x of Q at which $-n<\underline{\mathrm{D}}_{\mathrm{sym}}\Phi(x)$, and writing $\Phi_n(X)=\Phi(X)+n\cdot|X|$, we obtain $\underline{\mathrm{D}}_{\mathrm{sym}}\Phi_n(x)>0$ at every point $x\in Q_n$. Hence, by Lemma 14.4, we must have $\Phi(Q_n)=\Phi_n(Q_n)\geqslant 0$, and making $n\to\infty$ we find $\Phi(Q)\geqslant 0$. By symmetry we also have $\Phi(Q)\leqslant 0$ whenever Q is a bounded set measurable ($\mathfrak{B}$) of measure zero, such that $\overline{\mathrm{D}}\Phi(x)<+\infty$ at each point x of Q.

We pass on to the proof of formula (14.6). By Theorem 7.3, there exists a set A, measurable ($\mathfrak{B}$) and of measure zero, such that the relation

$$\Phi(X)=\Phi(X\cdot A)+\int_X \mathrm{D}\Phi(x)\,dx \tag{14.7}$$

holds whenever X is a bounded set measurable ($\mathfrak{B}$). Since the set A_∞ is of measure zero, we see at once from the equation (14.7) that the function Φ must vanish identically for all the subsets of $A_\infty - A$, which are bounded and measurable ($\mathfrak{B}$). On the other hand, by what has just been proved, the function vanishes also for all subsets of $A-A_\infty$. Hence the set A_∞ may be taken in place of the set A in (14.7) and this gives (14.6). Finally, if $\underline{\mathrm{D}}_{\mathrm{sym}}\Phi(x)>-\infty$ at every point x of a bounded set X measurable ($\mathfrak{B}$), then $\Phi(X\cdot A_\infty)\geqslant 0$ and the second part of the theorem follows at once from the first.

Let us mention the following consequence of Theorem 14.5: *If at each point x both the symmetrical derivates of a given additive function of a set are finite, the latter is absolutely continuous.* For ordinary derivates the corresponding proposition has long been known (cf. H. Lebesgue [5, p. 423]) and is moreover included in Theorem 15.7 of this chapter, as well as in Theorem 2.1 of Chap. VI.

***§ 15. Derivation in abstract spaces.** With certain hypotheses, a process of derivation may be defined for additive functions of a set in any separable metrical space, and for such a process, theorems similar to those of §§ 7 and 9 may be established.

(15.1) ***Lemma.*** *If Φ is an additive function of a set ($\mathfrak{B}$) on a metrical space $\boldsymbol{M}$, then given any set X measurable ($\mathfrak{B}$) and any $\varepsilon>0$, there exists an open set G such that*

$$\mathrm{W}(\Phi;\, G-X)<\varepsilon \quad \textit{and} \quad \mathrm{W}(\Phi;\, X-G)<\varepsilon. \tag{15.2}$$

Proof. Let $\mathfrak{B}_0$ denote the class of the sets X measurable ($\mathfrak{B}$) for each of which there exists, however we choose $\varepsilon > 0$, an open set G satisfying the relations (15.2). Since any closed set F is the limit of a descending sequence of open sets, we observe easily (cf. Theorems 5.1 and 6.4, Chap. I) that there exists for each $\varepsilon > 0$ an open set $G \supset F$ such that $W(\Phi; G - F) < \varepsilon$. The class $\mathfrak{B}_0$ thus includes all closed sets; to prove that $\mathfrak{B} = \mathfrak{B}_0$, it suffices, therefore, to show that the class $\mathfrak{B}_0$ is additive.

To do this, we choose $\varepsilon > 0$ and denote by X the sum of a sequence $\{X_n\}_{n=1,2,\ldots}$ of sets ($\mathfrak{B}_0$). To each set X_n there corresponds an open set G_n such that $W(\Phi; G_n - X_n) < \varepsilon/2^n$ and $W(\Phi; X_n - G_n) < \varepsilon/2^n$. Writing $G = \sum_n G_n$, we clearly find that the inequalities (15.2) are satisfied. Therefore $X \in \mathfrak{B}_0$.

Again, suppose that $\varepsilon > 0$ and that $X = CY$, where $Y \in \mathfrak{B}_0$. There will then exist an open set H such that $W(\Phi; Y - H) < \varepsilon/2$ and $W(\Phi; H - Y) < \varepsilon$. Consequently writing $P = CH$, we find that

$$(15.3) \qquad W(\Phi; P - X) < \varepsilon/2 \qquad \textit{and} \qquad W(\Phi; X - P) < \varepsilon.$$

But since the set P is closed, there exists an open set G such that $G \supset P$ and $W(\Phi; G - P) < \varepsilon/2$; and this implies, on account of (15.3), the inequalities (15.2) and so completes the proof.

We shall call *net* in a metrical space $\boldsymbol{M}$ any finite or enumerable family of sets measurable ($\mathfrak{B}$) no two of which have common points and which together cover the space $\boldsymbol{M}$. The sets constituting a net will be called its *meshes*. A sequence $\{\mathfrak{M}_n\}$ of nets will be termed *regular*, if each mesh of $\mathfrak{M}_{n+1}$ (where $n > 0$) is contained in a mesh of $\mathfrak{M}_n$ and if further $\Delta(\mathfrak{M}_n) \to 0$ as $n \to \infty$ (where $\Delta(\mathfrak{M}_n)$ denotes the characteristic number of $\mathfrak{M}_n$; cf. Chap. II, p. 40). It is easy to see that in order that there exist a regular sequence of nets in a metrical space, it is necessary and sufficient that this space be separable.

In the rest of this § we shall keep fixed a separable metrical space $\boldsymbol{M}$ and we shall suppose given in $\boldsymbol{M}$ a regular sequence $\mathfrak{M} = \{\mathfrak{M}_n\}$ of nets and a measure μ which is defined for the sets measurable ($\mathfrak{B}$) and which is subject to the condition $\mu(\boldsymbol{M}) < +\infty$. Let Φ be an additive function of a set ($\mathfrak{B}$) on $\boldsymbol{M}$. For $x \in M$, where M is any mesh of a net $\mathfrak{M}_n$, let us write

$$d_n(x) = \begin{cases} \Phi(M)/\mu(M) & \textit{when} \quad \mu(M) \neq 0, \\ +\infty & \textit{when} \quad \mu(M) = 0 \quad \textit{and} \quad \Phi(M) \geqslant 0, \\ -\infty & \textit{when} \quad \mu(M) = 0 \quad \textit{and} \quad \Phi(M) < 0. \end{cases}$$

The functions $d_n(x)$ are thus defined on the whole space $\boldsymbol{M}$ and are measurable $(\mathfrak{B})$. Let us write $(\mu, \mathfrak{M})\overline{\mathrm{D}}\Phi(x) = \limsup_n d_n(x)$. The number $(\mu, \mathfrak{M})\overline{\mathrm{D}}\Phi(x)$ thus defined will be called *upper derivate of the function Φ at the point x with respect to the measure μ and the regular sequence of nets $\mathfrak{M}$*. Considered as a function of x, this upper derivate is clearly measurable $(\mathfrak{B})$. Similarly we define the lower derivate $(\mu, \mathfrak{M})\underline{\mathrm{D}}\Phi(x)$. If at a point x the two numbers $(\mu, \mathfrak{M})\overline{\mathrm{D}}\Phi(x)$ and $(\mu, \mathfrak{M})\underline{\mathrm{D}}\Phi(x)$ are equal, their common value will be written $(\mu, \mathfrak{M})\mathrm{D}\Phi(x)$ and called *derivative of the function Φ at x with respect to the measure μ and the regular sequence of nets $\mathfrak{M}$*. For the rest of this §, a measure μ and a regular sequence of nets $\mathfrak{M}$ will be kept fixed in the space $\boldsymbol{M}$.

(15.4) ***Lemma.*** *Let Φ be an additive function of a set $(\mathfrak{B})$ on the space $\boldsymbol{M}$. Then*

(i) *if the inequality $(\mu,\mathfrak{M})\overline{\mathrm{D}}\Phi(x)\geqslant k$, where k is a finite number, holds at every point x of a set A measurable $(\mathfrak{B})$, we have $\Phi(A)\geqslant k\cdot\mu(A)$;*

(ii) *if at each point x of a set B measurable $(\mathfrak{B})$ and of measure (μ) zero, the derivative $(\mu,\mathfrak{M})\mathrm{D}\Phi(x)$ either does not exist or else exists and is finite, $\Phi(B)=0$.*

Proof. *re* (i). By subtracting from the function Φ the function $k\cdot\mu$, we may assume that $k=0$. Let ε be any positive number. By Lemma 15.1 there exists an open set G such that

$$(15.5) \qquad \mathrm{W}(\Phi; G-A)<\varepsilon \qquad \textit{and} \qquad \mathrm{W}(\Phi; A-G)<\varepsilon.$$

Let $\widetilde{\mathfrak{M}}_1$ be the set of the meshes M of the net $\mathfrak{M}_1$ such that

$$(15.6) \qquad M\subset G \qquad \textit{and} \qquad \Phi(M)\geqslant-\varepsilon\cdot\mu(M),$$

and generally, for $n\geqslant 1$, let $\widetilde{\mathfrak{M}}_{n+1}$ be the set of the meshes M of the net $\mathfrak{M}_{n+1}$ which fulfil the conditions (15.6) and are not contained in any of the meshes of $\widetilde{\mathfrak{M}}_1+\widetilde{\mathfrak{M}}_2+\ldots+\widetilde{\mathfrak{M}}_n$. By arranging the sets belonging to $\widetilde{\mathfrak{M}}_1+\widetilde{\mathfrak{M}}_2+\ldots+\widetilde{\mathfrak{M}}_n+\ldots$ in a sequence $\{M_k\}$, we have $\Phi(M_k)>-\varepsilon\cdot\mu(M_k)$ for $k=1, 2, \ldots$, and $A\cdot G\subset\sum_k M_k$. Since the sets M_k are measurable $(\mathfrak{B})$ and no two of them have common points, it therefore follows from (15.5) that $\Phi(A)=\Phi(A\cdot G)+\Phi(A-G)\geqslant\Phi(\sum_k M_k)-2\varepsilon\geqslant \geqslant-\varepsilon\cdot\sum_k\mu(M_k)-2\varepsilon\geqslant-\varepsilon\cdot[\mu(G)+2]$, and so that $\Phi(A)\geqslant 0$.

re (ii). Denote, for any positive integer n, by B_n the set of the points x of B at which $\overline{\mathrm{D}}\Phi(x)\geqslant -n$. On account of (i) we have $\Phi(B_n)\geqslant -n\cdot\mu(B_n)=0$ for each n, and, since $B=\lim\limits_n B_n$, this gives $\Phi(B)\geqslant 0$. By symmetry $\Phi(B)\leqslant 0$, and so finally $\Phi(B)=0$.

(15.7) ***Theorem.*** *If Φ is a function of a set ($\mathfrak{B}$), which is additive on the space $\boldsymbol{M}$, the derivative $(\mu, \mathfrak{M})\mathrm{D}\Phi(x)$ exists almost everywhere and is integrable ($\mathfrak{B}$, μ) on $\boldsymbol{M}$; moreover, if $E_{+\infty}$ and $E_{-\infty}$ denote the sets of the points at which $(\mu, \mathfrak{M})\mathrm{D}\Phi(x)=+\infty$ and $(\mu, \mathfrak{M})\mathrm{D}\Phi(x)=-\infty$ respectively, we have*

$$(15.8)\qquad \Phi(X)=\Phi(X\cdot E_{+\infty})+\Phi(X\cdot E_{-\infty})+\int_X(\mu, \mathfrak{M})\mathrm{D}\Phi(x)\,d\mu(x)$$

and

$$(15.9)\qquad \mathrm{W}(\Phi;\,X)=\Phi(X\cdot E_{+\infty})+|\Phi(X\cdot E_{-\infty})|+\int_X|(\mu, \mathfrak{M})\mathrm{D}\Phi(x)|\,d\mu(x)$$

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

Proof. By Theorem 14.6, Chap. I, there exist a function of a point f integrable ($\mathfrak{B}$, μ) on $\boldsymbol{M}$ and an additive function of a set Θ singular ($\mathfrak{B}$, μ) on $\boldsymbol{M}$ such that

$$(15.10)\qquad \Phi(X)=\Theta(X)+\int_X f(x)\,d\mu(x)\quad \textit{for every set } X\in\mathfrak{B}.$$

Let E be a set measurable ($\mathfrak{B}$) such that $\mu(E)=0$ and that the function Θ vanishes identically on $\mathrm{C}E$. Writing, for brevity, D, $\overline{D}$ and $\underline{D}$ in place of $(\mu,\mathfrak{M})\mathrm{D}$, $(\mu, \mathfrak{M})\overline{\mathrm{D}}$ and $(\mu,\mathfrak{M})\underline{\mathrm{D}}$ respectively, let us denote for any pair of integers $n>0$ and k, by $P_{n,k}$ the set of the points x at which $\overline{D}\Phi(x)\geqslant(k+1)/n>k/n\geqslant f(x)$. If we substitute $P_{n,k}\cdot\mathrm{C}E$ for X in (15.10), we find on account of Lemma 15.4 (i) that

$$\Phi(P_{n,k}\cdot\mathrm{C}E)\geqslant\frac{k+1}{n}\mu(P_{n,k}\cdot\mathrm{C}E)\geqslant\frac{k}{n}\cdot\mu(P_{n,k}\cdot\mathrm{C}E)\geqslant\int_{P_{n,k}\cdot\mathrm{C}E}f(x)d\mu(x)=\Phi(P_{n,k}\cdot\mathrm{C}E),$$

and so that $\mu(P_{n,k})=\mu(P_{n,k}\cdot\mathrm{C}E)=0$. Therefore $\overline{D}\Phi(x)\leqslant f(x)$ at almost all points x. By symmetry $\underline{D}\Phi(x)\geqslant f(x)$ must also hold almost everywhere in $\boldsymbol{M}$. Therefore the derivative $D\Phi(x)$ exists and equals $f(x)$ at almost all the points x of $\boldsymbol{M}$, and the identity (15.10) takes the form

$$(15.11)\ \Phi(X)=\Theta(X)+\int_X D\Phi d\mu=\Phi(E\cdot X)+\int_X D\Phi d\mu\quad \textit{for every set } X\in\mathfrak{B}.$$

Moreover, since $D\Phi(x)=f(x)\neq\pm\infty$ almost everywhere, the set $E_{+\infty}+E_{-\infty}$ is of measure (μ) zero, and it follows directly from (15.11) that the function Φ vanishes identically on the set $(E_{+\infty}+E_{-\infty})-E$.

On the other hand, by Lemma 15.4 (ii), Φ vanishes identically on $E-(E_{+\infty}+E_{-\infty})$. Therefore in (15.11) the set E may be replaced by the set $E_{+\infty}+E_{-\infty}$, and the relation (15.11) becomes the required formula (15.8). Finally, since by Lemma 15.4 (i) the function Φ is non-negative for the subsets ($\mathfrak{B}$) of $E_{+\infty}$ and non-positive for the subsets ($\mathfrak{B}$) of $E_{-\infty}$, formula (15.9) follows at once from formula (15.8).

Let us mention specially the following corollary of Theorem 15.7:

(15.12) ***Theorem.*** *Suppose given in the space $\boldsymbol{M}$ two regular sequences of nets $\mathfrak{N}$ and $\mathfrak{P}$, and, as before, a measure μ defined for the sets ($\mathfrak{B}$) and subject to the condition $\mu(\boldsymbol{M})<+\infty$. Then for every function Φ of a set ($\mathfrak{B}$), which is additive on $\boldsymbol{M}$, we have almost everywhere $(\mu, \mathfrak{N})\mathrm{D}\Phi(x)=(\mu, \mathfrak{P})\mathrm{D}\Phi(x)$; moreover, if E denotes the set of the points x at which either one at least of the derivatives $(\mu, \mathfrak{N})\mathrm{D}\Phi(x)$ and $(\mu, \mathfrak{P})\mathrm{D}\Phi(x)$ does not exist, or else both exist but have different values, then the function Φ vanishes identically on E, i. e.* $\mathrm{W}(\Phi; E)=0$.

In fact, if we write, for brevity, D_1 and D_2 in place of $(\mu, \mathfrak{N})\mathrm{D}$ and $(\mu, \mathfrak{P})\mathrm{D}$ respectively, and if we denote by Θ the function of singularities of Φ, we have by the previous theorem

$$\Phi(X)=\Theta(X)+\int_X D_1\Phi(x)\,d\mu(x)=\Theta(X)+\int_X D_2\Phi(x)\,d\mu(x)$$

for every set X measurable ($\mathfrak{B}$). Equating the two integrals which occur in this relation, we obtain almost everywhere $D_1\Phi(x)=D_2\Phi(x)$.

Now the set E of the points at which this relation does not hold, may be expressed as the sum of three sets A_1, A_2 and A_3, where A_1 is the set of the points $x \in E$ at which one at least of the derivatives $D_1\Phi(x)$ and $D_2\Phi(x)$ does not exist, or else exists and is finite, A_2 the set of the points x at which $D_1\Phi(x)=+\infty$ and $D_2\Phi(x)=-\infty$, and A_3 the set of the points x at which $D_1\Phi(x)=-\infty$ and $D_2\Phi(x)=+\infty$. It follows directly from Lemma 15.4 (ii) that the function Φ vanishes identically on A_1. In the same way, it follows from part (i) of this lemma that we have simultaneously $\Phi(X)\geqslant 0$ and $\Phi(X)\leqslant 0$, and so $\Phi(X)=0$, for every subset X measurable ($\mathfrak{B}$) of A_2 or of A_3. Consequently $\mathrm{W}(\Phi; E)=0$, and this completes the proof.

Theorem 15.7, which corresponds, to a certain extent, to Theorem 9.6, was first proved by Ch. J. de la Vallée Poussin [1; cf. also I, p. 103] for derivation with respect to the Lebesgue measure, and with respect to the regular sequences of nets of half open intervals in Euclidean spaces. Strictly, the Lebesgue measure does not fulfil the condition which we laid down for the measure μ, since Euclidean space has infinite Lebesgue measure. Nevertheless it is easy to see that for the validity of Theorem 15.7 (as well as for that of the other propositions of this §) it suffices to suppose only that the meshes of the nets considered have finite measure.

For the derivation of additive functions of a set in abstract spaces, see also R. de Possel [1].

*** § 16. Torus space.** As an example and an application of the results of the preceding §, we shall discuss in this § a metrical space which, from the point of view of the theory of measure and integration, may be considered as one of the nearest generalizations of Euclidean spaces. This space, called torus space of an infinite number of dimensions, occurs in a more or less explicit form in the important researches of H. Steinhaus [2], of P. J. Daniell [2; 3], and of other authors, in connection with certain problems of probability; but the first systematic study of this space is due to B. Jessen [2].

Following Jessen, we shall call *torus space* $\boldsymbol{Q}_\omega$ the metrical space whose elements are the infinite sequences of real numbers $\xi=(x_1, x_2, \dots, x_n, \dots)$ where $0 \leqslant x_n < 1$ for $n=1, 2, \dots,$ the distance $\varrho(\xi, \eta)$ of two points $\xi=(x_1, x_2, \dots, x_n, \dots)$ and $\eta=(y_1, y_2, \dots, y_n, \dots)$ in $\boldsymbol{Q}_\omega$ being defined by the formula $\varrho(\xi, \eta)=\sum_n |y_n-x_n|/2^n$. By $\boldsymbol{Q}_m$, where m is any positive integer, we shall denote the half open cube $[0, 1; 0, 1; \dots; 0, 1)$ in the Euclidean space $\boldsymbol{R}_m$. If $\xi=(x_1, x_2, \dots, x_n, \dots)$ is a point of $\boldsymbol{Q}_\omega$, we shall denote, for any positive integer m, by ξ'_m the point $(x_1, x_2, \dots, x_m)$ of $\boldsymbol{Q}_m$, and by ξ''_m the point $(x_{m+1}, x_{m+2}, \dots)$ of $\boldsymbol{Q}_\omega$, and we shall write $\xi=(\xi'_m, \xi''_m)$. According to this notation, (ξ, η) is a point of $\boldsymbol{Q}_\omega$ whenever $\xi \epsilon \boldsymbol{Q}_m$ (where m is any positive integer) and $\eta \epsilon \boldsymbol{Q}_\omega$. So that, if $A \subset \boldsymbol{Q}_m$ and $B \subset \boldsymbol{Q}_\omega$, the set $A \times B$ (cf. Chap. III, §§ 8, 9) lies in the space $\boldsymbol{Q}_\omega$; and in particular $\boldsymbol{Q}_m \times \boldsymbol{Q}_\omega = \boldsymbol{Q}_\omega$.

We shall call *closed interval*, or simply *interval*, in the space $\boldsymbol{Q}_\omega$, any set of the form $I \times \boldsymbol{Q}_\omega$, where I is a closed subinterval of $\boldsymbol{Q}_m$ for some value of $m=1, 2, \dots$. Similarly, taking I to be an interval which is half open (on the left or on the right) in $\boldsymbol{Q}_m$, we define in the space $\boldsymbol{Q}_\omega$ the *half open intervals (on the left or on the right).*

Every (closed) interval J in $\boldsymbol{Q}_\omega$ has only one expression of the form $I \times \boldsymbol{Q}_\omega$ where I is an interval in a space $\boldsymbol{Q}_m$. (It is to be remarked that the space $\boldsymbol{Q}_\omega$ itself is not a closed interval in the sense of the definitions given above.) By the *volume* of the interval $J = I \times \boldsymbol{Q}_\omega$ we shall mean the volume of the interval I in $\boldsymbol{Q}_m \subset \boldsymbol{R}_m$ (cf. Chap. III, § 2). Just as in Euclidean spaces, the volume of an interval J in $\boldsymbol{Q}_\omega$ will be denoted by $|J|$ or $\mathrm{L}_\omega(J)$. Again, as in Euclidean spaces (cf. Chap. III, § 5), we shall extend the notion of volume in the space $\boldsymbol{Q}_\omega$ by defining for every set E in this space the *outer measure* $\mathrm{L}_\omega^*(E)$ of the set E as the lower bound of the sums $\sum_k |J_k|$ where $\{J_k\}$ is any sequence of intervals such that $E \subset \sum_k J_k^\circ$. Thus defined, the outer measure evidently fulfils the three conditions of Carathéodory (cf. Chap. II, § 4) and determines, first the class of sets measurable ($\mathfrak{L}_{L_\omega^*}$), and then the class of functions measurable ($\mathfrak{L}_{L_\omega^*}$). For brevity, the sets and the functions belonging respectively to these classes, will simply be termed *measurable.* Also by the *measure* of a set E in the space Q_ω we shall always mean its measure′(L_ω^*).

It is easily shown, with the help of Borel's Covering Theorem, that the measure of any closed interval coincides with its volume (cf. Chap. III, § 5, p. 65), so that we can, without ambiguity, write $|E|$ or $L_\omega(E)$ (omitting the asterisk) to denote the outer measure of any set E in Q_ω. We also see that the boundary of any closed interval is of measure zero. Finally, we remark that the whole space Q_ω is of measure 1.

We shall now define in Q_ω a regular sequence of nets (cf. § 15, p. 153) of intervals half open on the right. We shall, in fact, denote for any positive integer m, by $\mathfrak{Q}^{(m)}$ the finite system of 2^{m^2} intervals half open on the right

$$[k_1/2^m,\ (k_1+1)/2^m;\ k_2/2^m,\ (k_2+1)/2^m;\ \dots;\ k_m/2^m,\ (k_m+1)/2^m)\times Q_\omega$$

where the k_i are arbitrary non-negative integers less than 2^m. We see at once that each system $\mathfrak{Q}^{(m)}$ is a net in Q_ω. To see that the sequence of these nets is regular, we observe in the first place that each interval of $\mathfrak{Q}^{(m+1)}$ is contained in one of the intervals of $\mathfrak{Q}^{(m)}$. On the other hand, no interval of the net $\mathfrak{Q}^{(m)}$ can have a diameter exceeding the number $\sum_{k=1}^{m} 1/2^{m+k}+\sum_{k=m+1}^{\infty} 1/2^k \leqslant 1/2^{m-1}$, so that the characteristic number $\Delta(\mathfrak{Q}^{(m)})$ of the net $\mathfrak{Q}^{(m)}$ tends to zero as $m\to\infty$.

If x and y are two real numbers, $x\dot{+}y$ will denote the number $x+y-[x+y]$, where, as usual, $[x+y]$ stands for the largest integer not exceeding $x+y$. If $\xi=(x_1, x_2, \dots, x_n, \dots)$ and $\eta=(y_1, y_2, \dots, y_n, \dots)$ are two points of Q_ω, we shall write $\xi\dot{+}\eta$ for $(x_1\dot{+}y_1, x_2\dot{+}y_2, \dots, x_n\dot{+}y_n, \dots)$. The point $\xi\dot{+}\eta$ clearly belongs to Q_ω.

We shall call *translation by the vector* a, where a is a point of Q_ω, the transformation which makes correspond to each point ξ of Q_ω the point $\xi\dot{+}a$. The translation by the vector a will be termed of order m, if all, except at most the first m, coordinates of a vanish.

A function f in Q_ω will be termed *cylindrical of order* m, if $f(\xi)$ does not depend on the first m coordinates of the point ξ, i. e. if

$f(\xi)=f(\xi \dot{+} a)$ identically in ξ, for every point a whose coordinates, except perhaps the first m, all vanish. A set E in Q_ω will be termed *cylindrical of order m*, if its characteristic function is so, or, what amounts to the same, if $E=Q_m \times A$ where A is a set in Q_ω.

(16.1) ***Theorem.*** *A function which is measurable on Q_ω and cylindrical of every finite order, is constant almost everywhere, i. e. $f(\xi)=c$ for almost all points ξ of Q_ω, where c is a constant.*

Proof. Suppose first that the function f is bounded, and therefore integrable, on Q_ω. Denoting by Φ the indefinite integral of f, let us define, for each value of m and for each mesh Q of the net $\mathfrak{Q}^{(m)}$, $f^{(m)}(\xi)=\Phi(Q)/\mathrm{L}_\omega(Q)$ whenever $\xi \epsilon Q$. For every pair of meshes Q_1 and Q_2 of the same net $\mathfrak{Q}^{(m)}$, there always exists a translation of order m which transforms Q_1 into Q_2; therefore, since the function f is cylindrical of order m, it follows that $\Phi(Q_1)=\Phi(Q_2)$; and since further $\mathrm{L}_\omega(Q_1)=\mathrm{L}_\omega(Q_2)$, each of the functions $f^{(m)}(\xi)$ is constant on Q_ω. On the other hand, we deduce from Theorem 15.7 that $f(\xi)=\lim_m f^{(m)}(\xi)$ almost everywhere in Q_ω, i. e. that the function is almost everywhere identical with a constant.

Now let f be any measurable function which is cylindrical of every finite order. Let us write $f_n(\xi)=f(\xi)$ when $|f(\xi)|\leqslant n$ and $f(\xi)=n$ when $|f(\xi)|>n$. Each of the functions $f_n(\xi)$ is bounded and cylindrical of every finite order, so that by what has just been proved, each of these functions is constant almost everywhere. Therefore the same is true of the function $f(\xi)=\lim_n f_n(\xi)$ and this completes the proof.

The fundamental properties of our measure in the space Q_ω may be established by methods similar to those used in Euclidean spaces. To illustrate this, let us enumerate some of these properties.

Given any measurable set E and any $\varepsilon>0$, there exists a closed set F and an open set G such that $F\subset E\subset G$ and such that $|G-E|<\varepsilon$ and $|E-F|<\varepsilon$ (cf. Theorem 6.6, Chap. III).

From this we may deduce next Lusin's theorem (cf. Theorem 7.1, Chap. III): *If f is a finite function measurable on a set E, there exists for each $\varepsilon>0$, a closed set $F\subset E$ such that the function f is continuous on F and that* $|E-F|<\varepsilon$; and its immediate corollary: *any function which is measurable in Q_ω, is equal almost everywhere in Q_ω to a function measurable* ($\mathfrak{B}$). Finally Fubini's theorem (cf. Chap. III, § 8) may be stated as follows for the space Q_ω:

(16.2) ***Theorem.*** *If f is a non-negative measurable function in the space Q_ω, then for any positive integer m,*

(i) *the definite integral $\int\limits_{Q_\omega} f(\xi, \eta) d\mathrm{L}_\omega(\eta)$ exists for every $\xi \in Q_m$, except at most for those of a set of measure (L_m) zero,*

(ii) *the definite integral $\int\limits_{Q_m} f(\xi, \eta) d\mathrm{L}_m(\xi)$ exists for every $\eta \in Q_\omega$, except at most for those of a set of measure (L_ω) zero,*

(iii) $$\int\limits_{Q_\omega} f(\zeta) d\mathrm{L}_\omega(\zeta) = \int\limits_{Q_m}\Big[\int\limits_{Q_\omega} f(\xi,\eta) d\mathrm{L}_\omega(\eta)\Big] d\mathrm{L}_m(\xi) = \int\limits_{Q_\omega}\Big[\int\limits_{Q_m} f(\xi,\eta) d\mathrm{L}_m(\xi)\Big] d\mathrm{L}_\omega(\eta).$$

Proof. We begin by verifying this directly when f is the characteristic function of a closed interval, or of a half open interval, and then successively when f is the characteristic function of an open set, of a set $(\mathfrak{G}_\delta)$, of a set of measure zero, and finally of any measurable set. It follows at once that the theorem is valid in the case where f is a simple function, and then, by passage to the limit, in the general case where f is any non-negative measurable function.

The line of argument that we have sketched, does not differ substantially in any way from the proof of Fubini's theorem for Euclidean spaces, and is even in a sense simpler than the latter, since in proving Theorem 8.1 of Chap. III we had to allow for the possibility of there being hyperplanes of discontinuity of the functions U and V.

In the space Q_ω there is, however, as shown by B. Jessen [2, p. 273], another theorem of the Fubini type, whose proof requires new methods. This theorem allows integration over the space Q_ω to be, so to speak, reduced to integrations over the cubes Q_m in Euclidean spaces, whereas each of the three members of the relation (iii) of Theorem 16.2 contains an integration extended over the space Q_ω.

(16.3) ***Jessen's theorem.*** *If f is a non-negative measurable function in the space Q_ω, the integral*

(16.4) $$f_m(\zeta) = \int\limits_{Q_m} f(\xi, \zeta_m'') d\mathrm{L}_m(\xi), \quad \textit{where } m = 1, 2, \ldots,$$

exists, and we have

(16.5) $$\lim_m f_m(\zeta) = \int\limits_{Q_\omega} f(\xi) d\mathrm{L}_\omega(\xi),$$

for almost all ζ in Q_ω.

Proof. Let us first remark that if Q is a set of measure zero in Q_ω, it follows from Theorem 16.2, applied to the characteristic function of Q, that for any m whatever, the set $\underset{\xi}{\mathrm{E}}[(\xi,\eta)\,\epsilon\, Q;\ \xi\,\epsilon\, Q_m]$ is of measure (L_m) zero for almost all η of Q_ω. Hence (with the notation adopted p. 157) we also have $\mathrm{L}_m\{\underset{\xi}{\mathrm{E}}[(\xi,\zeta''_m)\epsilon Q]\}=0$ for almost all ζ of Q_ω. It follows that if g and h are two non-negative measurable functions which are almost everywhere equal in Q_ω, the integrals $\int_{Q_m} g(\xi,\zeta''_m)\,d\mathrm{L}_m(\xi)$ and $\int_{Q_m} h(\xi,\zeta''_m)\,d\mathrm{L}_m(\xi)$ are equal for almost all the ζ of Q_ω, whatever m may be. We may therefore, without loss of generality, assume in the proof of Theorem 16.3 that the given function f is measurable $(\mathfrak{B})$; for any measurable function is almost everywhere equal to a function measurable $(\mathfrak{B})$.

The integral in the formula (16.4) then clearly exists for every ζ, and moreover it follows directly from this formula that the function $f_m(\zeta)$ is cylindrical of order m. The upper and lower limits of the sequence $\{f_m(\zeta)\}$ are thus cylindrical of every finite order and by Theorem 16.1 we may write almost everywhere in Q_ω

$$\liminf_m f_m(\zeta)=A \qquad \textit{and} \qquad \limsup_m f_m(\zeta)=B$$

where A and B are constants. It remains to be proved that $A=M=B$, where M denotes the integral on the right-hand side of (16.5).

We shall prove in the first place that $A\geqslant M$. For this purpose, let A' be any number exceeding A (if $A=+\infty$ our assertion is obvious), and write

$$(16.6)\quad P_k=\underset{\zeta}{\mathrm{E}}[f_k(\zeta)\leqslant A'],\qquad S_m=\sum_{k=1}^{m}P_k \qquad \textit{and} \qquad S=\lim_m S_m.$$

The set S coincides, except for a set of measure zero, with the whole space Q_ω. Keeping an index m fixed, let us evaluate the integral of f_m over S_m. Writing $R_m=P_m$, $R_{m-1}=P_{m-1}\cdot\mathrm{C}P_m,\ldots,R_1=P_1\cdot\mathrm{C}P_2\cdot\ldots\cdot\mathrm{C}P_m$, we have

$$(16.7)\quad S_m=\sum_{k=1}^{m}R_k \qquad \text{and} \qquad (16.8)\quad R_i\cdot R_j=0 \quad \textit{whenever}\ 1\leqslant i<j\leqslant m.$$

On the other hand, since every function f_k is cylindrical of order k, so are the sets P_k and $\mathrm{C}P_k$ and therefore the sets R_k for $k=1,2,\ldots,m$. We may thus write (cf. above p. 159) $R_k=Q_k\times\widetilde{R}_k$ where $\widetilde{R}_k\subset Q_\omega$. According to (16.6), we have $f_k(\zeta)\leqslant A'$ for every $\zeta\,\epsilon\, R_k\subset P_k$ where $k=1,2,\ldots,m$, or, what amounts to the same by formula (16.4),

$\int_{Q_k} f(\xi, \eta)\, d\mathrm{L}_k(\xi) \leqslant A'$ for every $\eta \in \widetilde{R}_k$. Therefore, on account of Theorem 16.2, we obtain for $k=1, 2, \ldots, m$,

$$\int_{R_k} f(\zeta)\, d\mathrm{L}_\omega(\zeta) = \int_{\widetilde{R}_k} \Big[\int_{Q_k} f(\xi, \eta)\, d\mathrm{L}_k(\xi) \Big]\, d\mathrm{L}_\omega(\eta) \leqslant A' \cdot \mathrm{L}_\omega(\widetilde{R}_k) = A' \cdot \mathrm{L}_\omega(R_k),$$

whence it follows by (16.7) and (16.8) that $\int_{S_m} f(\zeta) d\mathrm{L}_\omega(\zeta) \leqslant A' \cdot \mathrm{L}_\omega(S_m) \leqslant A'$. Making $m \to \infty$, we obtain in the limit $M = \int_{Q_\omega} f(\zeta)\, d\mathrm{L}_\omega(\zeta) \leqslant A'$ and so $M \leqslant A$.

By symmetry $M \geqslant B$ and, since it is clear that $A \leqslant B$, this requires $A = M = B$ and completes the proof.