

INTEGRATION OF FUNCTIONS WITH VALUES IN A BANACH SPACE†

BY

GARRETT BIRKHOFF‡

1. **Introduction.** The central concern of this paper is the integration of functions with values in a complete normed vector space, or “Banach” space \mathfrak{B} . This question has already been studied by Graves and Bochner,§ but we shall approach it from an entirely independent angle, most easily understood as an extension of Fréchet’s elegant interpretation|| of the Lebesgue integral.

Fréchet considers a function $f(p)$ from an abstract domain \mathfrak{S} with a σ -field Σ of measurable sets, to the real number system R . To each partition Δ of \mathfrak{S} into finite or enumerable sets σ_i (of measures $m(\sigma_i)$) of Σ he assigns a “relative upper integral”

$$J^*(f, \Delta) \equiv \sum_i m(\sigma_i) \cdot \sup_{p \in \sigma_i} f(p)$$

and a dual “relative lower integral”

$$J_*(f, \Delta) \equiv \sum_i m(\sigma_i) \cdot \inf_{p \in \sigma_i} f(p),$$

assuming that both series are unconditionally convergent.

It is evident that $J_*(f, \Delta) \leq J^*(f, \Delta')$ for any Δ, Δ' . Therefore the intersection of the “relative integral ranges”

$$J_*(f, \Delta) \leq x \leq J^*(f, \Delta),$$

for fixed f and variable Δ , is not empty. If it consists of a single point $J(f)$ of R , then $f(p)$ is called “integrable,” and $J(f)$ is called the “integral” of $f(p)$.

Our integral may be obtained from Fréchet’s by making two alterations. R must be replaced by an arbitrary Banach space \mathfrak{B} , and the “relative integral range” must be redefined as the least closed convex set containing all sums $\sum_i m(\sigma_i) \cdot f(p_i)$ [$p_i \in \sigma_i$], assuming again the unconditional convergence of all such series.

† Presented to the Society, September 4, 1934; received by the editors August 16, 1934, and, in revised form, February 11, 1935.

‡ Society of Fellows, Harvard University.

§ L. M. Graves, *Riemann integration and Taylor’s theorem in general analysis*, these Transactions, vol. 29 (1927), pp. 163–77. S. Bochner, *Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20 (1933), pp. 262–76.

|| Fréchet, *Sur l’intégrale d’une fonctionnelle étendue à un ensemble abstrait*, Bulletin de la Société Mathématique de France, vol. 43 (1915), pp. 248–65.

The essential task is to prove that any two integral ranges of the same function overlap, leading immediately to the recognition of integrability as the property of having relative integral ranges of arbitrarily small diameter.

2. **Outline.** With this in mind, the outline of the paper is very easy to remember.

The essential technical facts are established in §§3–9 by a study of convexity and unconditional convergence.† Interesting incidental results are obtained, but the emphasis is on the large number of ways in which the sets of a given unconditionally convergent series of sets can be replaced, without destroying unconditional convergence or enlarging the closure of the “convex hull” of the vector sum.

The definition of the integral sketched in the introduction is then stated in full, together with some remarks on “completely additive set functions.” These occupy §§10–14.

They are naturally followed by a discussion of the properties of the integral, a few of which may be stated here. The integral of any integrable function T is a completely additive set function depending linearly on T [§§15, 18]. Finite-valued functions are everywhere dense in the “space” of these set functions [§17]. If \mathfrak{B} is separable, then two integrands define the same set function if and only if they are “equivalent” as functions [§20]. And any rectifiable curve in Hilbert space has a tangent at almost every point [§21].

The paper concludes in §§22–25 with counterexamples (such as of nowhere differentiable integrals), with a demonstration that our integral genuinely includes those of Graves and Bochner, and with the enumeration of some unsolved problems.

3. **Calculus of complexes.** The object of this section is to familiarize the reader with the formal properties of two natural operations on non-vacuous sets, or “complexes” of vectors.

Accordingly, let \mathfrak{B} be any vector space,‡ whose elements we shall denote by Greek, and whose (real) coefficients by italic letters. Let further B_1, B_2, B_3, \dots denote complexes of elements of \mathfrak{B} , in the sense just defined.

We introduce the notation

$$b_1 \cdot B_1 + \dots + b_r \cdot B_r = \sum_{i=1}^r b_i \cdot B_i$$

† First studied by M. W. Orlicz, *Beiträge zur Theorie der Orthogonalentwicklungen*, *Studia Mathematica*, vol. 1 (1929), pp. 1–39 and 249–55.

‡ As defined for instance in S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 26. We shall refer to this volume in later footnotes as Banach.

for the set of all sums $b_1 \cdot \beta_1 + \cdots + b_r \cdot \beta_r$, $[\beta_i \in B_i]$. This is to be distinguished from $b_1 \cdot B_1 + \cdots + b_r \cdot B_r$, which will be used to denote the point-set sum of the $b_i \cdot B_i$, that is, the set of all elements of the form $b_i \cdot \beta_i$.

It is quite evident that if we denote the origin by Θ , then the following six properties of vector spaces hold for the vector sums of complexes:

- V1: $B_1 + B_2 = B_2 + B_1$,
 V2: $B_1 + (B_2 + B_3) = (B_1 + B_2) + B_3$,
 V3: $b \cdot (B_1 + B_2) = b \cdot B_1 + b \cdot B_2$,
 V4: $b_1 \cdot (b_2 \cdot B) = b_1 b_2 \cdot B$,
 V5: $1 \cdot B = B$,
 V6: $B + \Theta = B$ and $0 \cdot B = \Theta$.

Therefore if we define a "vectoroid" space to be any system satisfying conditions V1–V6, we can assert

THEOREM 1. *The non-vacuous subsets of \mathfrak{B} are the elements of a vectoroid space.*

Calculations based on V1–V6 will be regarded as evident in the remainder of the paper, and performed without explanation.

4. **Convex hulls.** Let again B be any complex of \mathfrak{B} . By the "convex hull" of B [in symbols, $\text{Co}(B)$] we mean the set of all elements of the form $b_1 \beta_1 + \cdots + b_r \beta_r$, where $b_i \geq 0$, $\beta_i \in B$, and $b_1 + \cdots + b_r = 1$. A convex complex is of course one which is its own convex hull, and any convex hull is convex, i.e., $\text{Co}(\text{Co}(B)) = \text{Co}(B)$. We observe in passing without proof

THEOREM 2. *$B \subset \mathfrak{B}$ is convex if and only if it satisfies $(m_1 + m_2)B = m_1 B + m_2 B$ for all $m_1, m_2 \geq 0$.*

Now by V4, $b_1 m \beta_1 + \cdots + b_r m \beta_r = m(b_1 \beta_1 + \cdots + b_r \beta_r)$, which shows that $\text{Co}(m \cdot B) = m \cdot \text{Co}(B)$. Evidently also if $\alpha_i \in A$, $\beta_i \in B$, $c_i \geq 0$, $c_i + \cdots + c_r = 1$, and $\xi = c_1(\alpha_1 + \beta_1) + \cdots + c_r(\alpha_r + \beta_r)$, then

$$\xi = (c_1 \alpha_1 + \cdots + c_r \alpha_r) + (c_1 \beta_1 + \cdots + c_r \beta_r) \in \text{Co}(A) + \text{Co}(B)$$

proving $\text{Co}(A + B) \subset \text{Co}(A) + \text{Co}(B)$.

But it is geometrically obvious† that given $a_i, b_i \geq 0$ such that $a_1 + \cdots + a_r = b_1 + \cdots + b_r = 1$, $c_{i,j} \geq 0$ exists satisfying $\sum_j c_{i,j} = a_i$ and $\sum_i c_{i,j} = b_j$, whence $\sum_{i,j} c_{i,j} = 1$. Therefore if $\eta = \sum_i a_i \alpha_i + \sum_j b_j \beta_j \in \text{Co}(A) + \text{Co}(B)$, writing $\eta = \sum_{i,j} c_{i,j}(\alpha_i + \beta_j)$, we see that $\eta \in \text{Co}(A + B)$, proving $\text{Co}(A) + \text{Co}(B) \subset \text{Co}(A + B)$.

This completes the proof of

† For both the a_i and the b_j can be regarded as dividing a unit line segment into disjoint intervals, the intersections of which are subintervals whose lengths $c_{i,j}$ have the desired properties.

THEOREM 3.† $\text{Co}(m \cdot B) = m \cdot \text{Co}(B)$ and $\text{Co}(A + B) = \text{Co}(A) + \text{Co}(B)$.

That is, abstractly speaking, the correspondence $B \rightarrow \text{Co}(B)$ is a homeomorphism carrying the vectoroid space of Theorem 1 into the vectoroid subspace of convex complexes.

5. **The norm and diameter of convex hulls.** We now add the permanent assumption that \mathfrak{B} is “normed,” that is, that there is associated with \mathfrak{B} a rule assigning to every $\xi \in \mathfrak{B}$ a number $\|\xi\|$ called the “norm” of ξ , and satisfying

$$N1: \quad \|\Theta\| = 0 \text{ and } \|\xi\| > 0 \text{ for } \xi \neq \Theta,$$

$$N2: \quad \|\xi + \eta\| \leq \|\xi\| + \|\eta\|,$$

$$N3: \quad \|c \cdot \xi\| = |c| \cdot \|\xi\|.$$

Such a rule automatically associates with every bounded complex B of \mathfrak{B} the “norm” $\|B\| \equiv \sup_{\beta \in B} \|\beta\|$. It also associates with B a “diameter” $\rho(B) = \|B - B\| \leq 2\|B\|$. Moreover the norms of bounded complexes clearly satisfy $N1-N3$.

Because of the convexity of the norm function, we can prove

THEOREM 4. $\|\text{Co}(B)\| = \|B\|$ and $\rho(\text{Co}(B)) = \rho(B)$.

For if $b_i \geq 0$, $\sum_i b_i = 1$, and $\beta_i \in B$, then

$$\left\| \sum_{i=1}^r b_i \cdot \beta_i \right\| \leq \sum_{i=1}^r \|b_i \cdot \beta_i\| = \sum_{i=1}^r b_i \cdot \|\beta_i\| \leq \sum_{i=1}^r b_i \cdot \|B\| = \|B\|$$

proving that $\|\text{Co}(B)\| \leq \|B\|$. But obviously $\|\text{Co}(B)\| \geq \|B\|$, and so $\|\text{Co}(B)\| = \|B\|$. The second half of Theorem 4 follows since

$$\rho(\text{Co}(B)) = \|\text{Co}(B) - \text{Co}(B)\| = \|\text{Co}(B - B)\| = \|B - B\| = \rho(B).$$

6. **Limits and closure.** Hereafter we shall assume that \mathfrak{B} is not only normed but “complete,” that is, that every Cauchy sequence of elements of \mathfrak{B} tends to a limit element of \mathfrak{B} . Further, we shall denote by \overline{B} the closure of any complex B in \mathfrak{B} . The reader will find no difficulty in proving that $\text{Co}(B)$ is the least closed convex set containing B . The truth of the formula $\overline{A + B} \subset \overline{A} + \overline{B}$ is equally obvious.

THEOREM 5. $\left\| \sum_{i=1}^r B_i \right\|$ is not altered if we replace the B_i by the closures of their convex hulls. Moreover if $0 \leq c_i \leq 1$, then $\left\| \sum_{i=1}^r c_i \cdot B_i \right\|$ is bounded by the norm $\left\| \sum_{k=1}^s B_{i(k)} \right\|$ of the vector sum of some set of the B_i .

† Theorem 3 has long been known. Cf. for instance T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Berlin, 1934, p. 29.

The first statement follows directly from Theorems 3-4 and the equality $\|\overline{A}\| = \|A\|$. To prove the second, it is sufficient to note that $c_i \cdot B_i \subset \text{Co}(B_i + \Theta)$. Hence by Theorem 3,

$$\sum_{i=1}^r c_i \cdot B_i \subset \text{Co} \left(\sum_{i=1}^r (B_i + \Theta) \right),$$

and

$$\begin{aligned} \left\| \sum_{i=1}^r c_i \cdot B_i \right\| &\leq \left\| \text{Co} \left(\sum_{i=1}^r (B_i + \Theta) \right) \right\| = \left\| \sum_{i=1}^r (B_i + \Theta) \right\| \\ &= \sup \left\| \sum_{k=1}^s B_{i(k)} \right\| \end{aligned}$$

by what we have just proved, Theorem 4, and definition.

7. Unconditional summation of elements. The primary object of §§7-9 is to translate unconditional convergence into terms of the calculus of complexes and limitations on norm, in order to be able to handle the properties of relative integral ranges. For this purpose Theorem 6 is not strictly necessary.

An enumerable aggregate Ξ of elements $\xi_1, \xi_2, \xi_3, \dots$ (which need not be distinct) of \mathfrak{B} is called "unconditionally summable to ξ ," if and only if every arrangement α of all the elements of Ξ gives a series $\Xi^{(\alpha)}: \xi_{\alpha(1)} + \xi_{\alpha(2)} + \xi_{\alpha(3)} + \dots$ convergent to ξ . Under these conditions, the series $\Xi^{(\alpha)}$ are called "unconditionally convergent to ξ ."

It is clear that Ξ ($\Xi^{(\alpha)}$) is unconditionally summable (convergent) to ξ if and only if to every $\epsilon > 0$ corresponds a number N so large that the sum η of any finite set of terms of Ξ ($\Xi^{(\alpha)}$) including ξ_1, \dots, ξ_N satisfies $\|\eta - \xi\| < \epsilon$.

Now let $\Xi: \xi_1 + \xi_2 + \xi_3 + \dots$ and $\Xi': \xi'_1 + \xi'_2 + \xi'_3 + \dots$ be any two unconditionally convergent series. By $c \cdot \Xi$ we mean the series $c \cdot \xi_1 + c \cdot \xi_2 + c \cdot \xi_3 + \dots$, and by $\Xi + \Xi'$ the series $\Xi'': \xi''_1 + \xi''_2 + \xi''_3 + \dots$, where $\xi''_i = \xi_i + \xi'_i$. The reader can easily see that these series are unconditionally convergent, and that the operations of addition and of multiplication by a scalar possess all of the usual vector properties. Therefore the unconditionally convergent series of \mathfrak{B} are the elements of a vector space \mathfrak{C} .

Let $B(\Xi)$ denote the (bounded) set of the finite partial sums of the elements of Ξ . By the "norm" $\|\Xi\|$ of Ξ we mean $\|B(\Xi)\|$. Since $B(\Xi + \Xi') \subset B(\Xi) + B(\Xi')$ and $B(c \cdot \Xi) = c \cdot B(\Xi)$, we see that \mathfrak{C} may be regarded as normed in the sense of §5.

We shall now prove that \mathfrak{C} is complete relative to this norm, which amounts to asserting

THEOREM 6. *The unconditionally convergent series of \mathfrak{B} are the elements of a second Banach space.*

Let $\Xi_1, \Xi_2, \Xi_3, \dots$ be any sequence of unconditionally convergent series of elements of \mathfrak{B} , such that to any $\epsilon > 0$ corresponds N so large that $m \geq N$ and $n \geq N$ imply $\|\Xi_m - \Xi_n\| < \epsilon$. Clearly the i th terms ξ_i^k of the Ξ_k are uniformly convergent Cauchy sequences, with limits ξ_i . Let Ξ denote the formal series $\xi_1 + \xi_2 + \xi_3 + \dots$. The proof is complete if Ξ is unconditionally convergent and

$$\lim_{n \rightarrow \infty} \|\Xi - \Xi_n\| = 0.$$

But to any $\epsilon > 0$ corresponds N so large that if $m, n \geq N$, then $\|\Xi_m - \Xi_n\| < \epsilon$. And we can find M so large that if $M < k(1) < \dots < k(r)$, then $\|\sum_{i=1}^r \xi_{k(i)}^N\| < \epsilon$. It follows that under the same hypotheses,

$$\begin{aligned} \left\| \sum_{i=1}^r \xi_{k(i)} \right\| &= \left\| \lim_{n \rightarrow \infty} \sum_{i=1}^r \xi_{k(i)}^n \right\| \\ &\leq \left\| \sum_{i=1}^r \xi_{k(i)}^N \right\| + \epsilon < 2\epsilon \end{aligned}$$

so that Ξ must be unconditionally convergent. But now if we pick *any* $j(1) < \dots < j(s)$, then for $n \geq N$,

$$\begin{aligned} \left\| \sum_{i=1}^s (\xi_{j(i)} - \xi_{j(i)}^n) \right\| &= \left\| \lim_{m \rightarrow \infty} \sum_{i=1}^s (\xi_{j(i)}^m - \xi_{j(i)}^n) \right\| \\ &= \lim_{m \rightarrow \infty} \left\| \sum_{i=1}^s (\xi_{j(i)}^m - \xi_{j(i)}^n) \right\| \leq \epsilon \end{aligned}$$

by hypothesis. Hence $\|\Xi - \Xi_n\| = \epsilon$, completing the proof.

8. Unconditional summation of complexes. Suppose similarly Φ is an aggregate of enumerable complexes B_1, B_2, B_3, \dots of \mathfrak{B} . Φ will be called "unconditionally summable" to a given complex B if and only if every series $\beta_1 + \beta_2 + \beta_3 + \dots$ [$\beta_i \in B_i$] is unconditionally convergent, and B is the locus of the sums of such series. We shall abbreviate this by writing $\sum_i B_i = B$.

In order that Φ be unconditionally summable it is necessary as well as sufficient that to any $\epsilon > 0$ correspond N so large that $N < k(1) < \dots < k(r)$ implies $\|B_{k(1)} + \dots + B_{k(r)}\| < \epsilon$. For otherwise we could form an infinite series of elements from a sequence of such sets of complexes which was not unconditionally convergent no matter how the gaps between the different terms were filled in by elements from the remaining complexes of Φ .

Keeping this in mind, we can prove without difficulty

THEOREM 7. If $\sum_i B_i = B$, then $\sum_i \overline{\text{Co}(B_i)} \subset \overline{\text{Co}(B)}$ and

$$\overline{\text{Co}(B)} = \overline{\sum_i \text{Co}(B_i)} = \overline{\sum_i \overline{\text{Co}(B_i)}}.$$

In the first place (using the above notation), by Theorem 5

$$\left\| \sum_{i=1}^r \overline{\text{Co}(B_{k(i)})} \right\| = \left\| \sum_{i=1}^r B_{k(i)} \right\| < \epsilon,$$

and consequently the aggregate of the $\overline{\text{Co}(B_i)}$ is unconditionally summable. Secondly, by the theory of limits, for sufficiently large N it is true that every point of

$$\sum_{i=1}^N \overline{\text{Co}(B_i)} \subset \overline{\text{Co}\left(\sum_{i=1}^N B_i\right)}$$

is within ϵ of some point of $\overline{\text{Co}(B)}$, and hence in the limit $\sum_i \overline{\text{Co}(B_i)} \subset \overline{\text{Co}(B)}$.

But conversely every point b of $\overline{\text{Co}(B)}$ can be approximated by a sum $b^* = \sum_{i=1}^s m_i \beta_i^*$, where $m_i \geq 0$, $\sum_{i=1}^s m_i = 1$, and $\beta_i^* \in B$. And $\beta_i^* = \sum_{k=1}^\infty \beta_i^k$ [$\beta_i^k \in B_k$] whence $b^* = \sum_{i=1}^s m_i \sum_{k=1}^\infty \beta_i^k$. And since the sum of s unconditionally summable aggregates is itself unconditionally summable,

$$b^* = \sum_{k=1}^\infty \sum_{i=1}^s m_i \beta_i^k = \sum_{k=1}^\infty \tilde{\beta}_k \quad [\tilde{\beta}_k \in \text{Co}(B_k)].$$

That is, $\overline{\text{Co}(B)} \subset \overline{\sum_i \text{Co}(B_i)}$.

Now consider the following triple inequality,

$$\overline{\sum_i \text{Co}(B_i)} \subset \overline{\text{Co}(B)} \subset \overline{\sum_i \text{Co}(B_i)} \subset \overline{\sum_i \overline{\text{Co}(B_i)}}.$$

The first relation follows from $\sum_i \overline{\text{Co}(B_i)} \subset \overline{\text{Co}(B)}$, which was proved above. The second we have just proved, and the third is utterly obvious. This completes the demonstration of Theorem 7.

THEOREM 8. If the aggregate Φ of complexes B_i is unconditionally summable, and the $|m_i|$ are uniformly bounded by a finite constant K , then the aggregate of complexes $m_i \cdot B_i$ is unconditionally summable, and

$$\left\| \overline{\text{Co}\left(\sum_i m_i \cdot B_i\right)} \right\| \leq 2K \cdot \sup \left\| \sum_{k=1}^r B_{i(k)} \right\|.$$

Theorem 8 follows directly if we can prove unconditional summability from the relations

$$\begin{aligned}
\left\| \overline{\text{Co} \left(\sum_i m_i \cdot B_i \right)} \right\| &= \left\| \sum_i m_i \cdot B_i \right\| \\
&\leq \left\| \sum_i |m_i| \cdot (B_i + \Theta) + \sum_i (-|m_i|) \cdot (B_i + \Theta) \right\| \\
&\leq 2K \cdot \left\| \sum_i \text{Co} (B_i + \Theta) \right\| = 2K \cdot \left\| \sum_i (B_i + \Theta) \right\| \\
&\equiv 2K \cdot \sup \left\| \sum_{k=1}^r B_{i(k)} \right\|
\end{aligned}$$

in which the first relation results from Theorem 4, the second from separating positive and negative coefficients, the third and fifth from inclusion of the left-hand complex in the right, and the fourth from Theorem 7.

But unconditional summability results from the three facts (1) $m_i \cdot B_i \subset K \cdot \text{Co}(B_i + \Theta) - K \cdot \text{Co}(B_i + \Theta)$, (2) Theorem 7, (3) the set of the finite partial sums of the $(B_i + \Theta)$ not involving the first N terms is the same as that of the B_i . For these show, taken in reverse order, that if $\sum_i B_i$ exists, then so do (1) $\sum_i (B_i + \Theta)$, (2) $\sum_i \text{Co}(B_i + \Theta)$, and hence (by Theorem 6) $\sum_i [K \cdot \text{Co}(B_i + \Theta) - K \cdot \text{Co}(B_i + \Theta)]$, (3) $\sum_i m_i \cdot B_i$.

9. Replacement of single series by double series. We are now in a position to prove the essential

THEOREM 9. *If $m_j^i \geq 0$, $\sum_i m_j^i = m_j$ for every j , $B_j^i \subset B_j$, every B_i is bounded, and $\sum_i m_i B_i = B$, then $\sum_{i,j} m_j^i \cdot B_j^i \subset \overline{\text{Co}(B)}$, whence*

$$\overline{\text{Co} \left(\sum_{i,j} m_j^i \cdot B_j^i \right)} \subset \overline{\text{Co}(B)}.$$

For to any $\epsilon > 0$ corresponds M so large that the norm of any finite sum of complexes B_k [$k > M$] is less than $\frac{1}{2}\epsilon$. And since B_1, \dots, B_M are bounded, we can choose N satisfying

$$(A) \quad \sum_{i=1}^M \sum_{j=N+1}^{\infty} m_j^i < \epsilon/2 [\|B_1\| + \dots + \|B_M\|].$$

If therefore we exclude the MN complexes $m_j^i \cdot B_j^i$ for which $i \leq M, j \leq N$, we see that any finite sum of the remaining $m_j^i \cdot B_j^i$ is composed of a set for which $i \leq M, j > N$, plus a set for which $i > M$. But by the triangle inequality on norm and (A), the norm of the first sum $< \frac{1}{2}\epsilon$, while by Theorem 5 and construction the norm of the second sum $< \frac{1}{2}\epsilon$. Therefore the norm of the whole sum $< \epsilon$, and the $m_j^i \cdot B_j^i$ are unconditionally summable.

Again, every element of $\sum_{i=1}^M \sum_{j=1}^N m_j^i \cdot B_j^i$ lies within $\frac{1}{2}\epsilon$ of some element of $\overline{\text{Co}(\sum_{i=1}^M m_i \cdot B_i)}$, and hence within ϵ of some element of $\overline{\text{Co}(\sum_i m_i \cdot B_i)}$. Therefore in the limit, by Theorem 7, we obtain

$$\sum_{i,j} m_j^i \cdot B_j^i \subset \overline{\text{Co}(B)}.$$

10. **Admissible domains.** Hitherto we have confined our attention to properties of the range of the functions which we shall try to integrate. We shall in the present section consider the domain.

We shall define as an "admissible domain" any space \mathfrak{S} of points in which is defined a so-called " σ -ring" Σ of "measurable" point sets satisfying

D1: The complement $\mathfrak{S} - \sigma$ of any one set σ , and the product $\sigma_1 \cdot \sigma_2$ and the sum $\sigma_1 + \sigma_2$ of any two sets σ_1 and σ_2 of Σ , are in Σ .

D2: To every set σ of Σ corresponds a number $m(\sigma)$ called the "measure" of σ .

D3: $m(\sigma)$ is zero, finite and positive, or $+\infty$.

D4: If $\sigma = \sigma_1 + \sigma_2 + \sigma_3 + \dots$ is the sum of finite or enumerable disjoint sets σ_i of Σ , then σ is in Σ , and $m(\sigma) = m(\sigma_1) + m(\sigma_2) + m(\sigma_3) + \dots$.

By a "decomposition" of \mathfrak{S} we mean any choice Δ_k of finite or enumerable disjoint non-vacuous measurable sets of finite measure, whose point-set sum is \mathfrak{S} . We shall adopt the fixed notation $\sigma_1^k, \sigma_2^k, \dots$ for the sets of composition of Δ_k .

By the "product" $\Delta_1 \cdot \Delta_2$ of two decompositions Δ_1 and Δ_2 we mean the decomposition of \mathfrak{S} into those sets $\sigma_i^1 \cdot \sigma_j^2$ which are non-vacuous.

11. **Completely additive set functions.** Since integration will be defined relative to the σ -ring Σ , it is only natural that we should define a (single-valued) "set function" as a function J assigning to each set σ of Σ a single "value" $J(\sigma)$ in \mathfrak{B} .

The "sum" $K = J_1 + J_2$ of two set functions J_1 and J_2 , and the "product" $K^* = c \cdot J_1$ of J_1 by a real scalar c , are of course defined by the identities $K(\sigma) = J_1(\sigma) + J_2(\sigma)$ and $K^*(\sigma) = c \cdot J_1(\sigma)$. And J is called "completely additive" if and only if the hypothesis that σ is the sum of finite or enumerable disjoint sets σ_i of Σ implies the conclusion that the values $J(\sigma_i)$ are unconditionally summable to $J(\sigma)$.

LEMMA. *If J is completely additive, then the set of the $J(\sigma)$ [$\sigma \in \Sigma$] has a finite upper bound.*

Otherwise we could choose $\sigma_1, \sigma_2, \sigma_3, \dots$ by induction so as to satisfy $\|J(\sigma_1)\| > 1$ and $\|J(\sigma_{i+1})\| > 3\|J(\sigma_i)\|$. And the series of the $(J(\sigma_i - \sigma_i \cdot (\sigma_1 + \dots + \sigma_{i-1})))$ could not be unconditionally summable.

The (finite) least upper bound to the $\|J(\sigma_i)\|$ will be called the "norm" of J , denoted by $\|J\|$.

THEOREM 10. *The completely additive set functions of \mathfrak{S} to \mathfrak{B} are a Banach space $\mathfrak{F}(\mathfrak{S}; \mathfrak{B})$.*

Every property of Banach space is obvious except completeness. But since the $J_n(\sigma)$ are a uniformly convergent Cauchy sequence, it is obvious that they tend uniformly to a limit set function J .

It remains to prove that J is completely additive. But for each choice of $\sigma = \sigma_1 + \sigma_2 + \sigma_3 + \cdots$, this is a corollary of Theorem 6. This completes the proof of Theorem 10.

There are three superficial remarks, which, although apparently inconsequential, should perhaps be made. In the first place, the proof of Theorem 10 can be duplicated to show that $\mathfrak{F}(\mathfrak{S}; \mathfrak{B})$ is imbedded in the Banach "super-space" of bounded set functions. Secondly, every permutation $\sigma \rightarrow \pi(\sigma)$ of the sets of Σ induces an isometric linear transformation $\Pi: J \rightarrow J_\pi$ of this super-space into itself, defined by the equation $J_\pi(\sigma) = J(\pi^{-1}(\sigma))$. And thirdly, if this permutation preserves inclusion relations (i.e., is itself induced by a measure-preserving permutation of the points of \mathfrak{S}), then it carries $\mathfrak{F}(\mathfrak{S}; \mathfrak{B})$ into itself, and so defines an isometric linear transformation on it.

12. Admissible point functions. By a "function" (more precisely, point function) T of an admissible domain \mathfrak{S} to a Banach space \mathfrak{B} we shall mean from now on a rule assigning to each point p of \mathfrak{S} one or more† "images" in \mathfrak{B} . More generally, if σ is any complex in \mathfrak{S} , we shall use $T(\sigma)$ to denote the complex of the images of the points of σ .

The "sum" $V = T + U$ of two such functions T and U , and the "product" $W = k \cdot T$ of a function T by a real scalar k are naturally defined by setting $V(p) = T(p) + U(p)$ and $W(p) = k \cdot T(p)$. This defines the admissible functions as elements of a "vectoroid" space, which becomes a vector space if we restrict ourselves to single-valued functions.

13. Summability and integral ranges. We now lay down

DEFINITION 1. *A function T is called "summable" under the decomposition Δ of \mathfrak{S} if and only if each $T(\sigma_i)$ is bounded, and the aggregate of the $m(\sigma_i) \cdot T(\sigma_i)$ is unconditionally summable.*

† The idea that since an integral range is a multiple-valued set function, we lose nothing by allowing T to be multiple-valued, is due to A. Kolmogoroff, *Untersuchungen über den Integralbegriff*, Mathematische Annalen, vol. 103 (1930), pp. 654–96.

DEFINITION 2. If T is summable under Δ , then the set

$$J_{\Delta}(T) \equiv \overline{\text{Co} \left(\sum_i m(\sigma_i) \cdot T(\sigma_i) \right)}$$

is called the "integral range" of T relative to Δ .

THEOREM 11. If the function T is summable under two decompositions Δ and Δ_1 , then T is summable under the product decomposition $\Delta \cdot \Delta_1$, and

$$J_{\Delta \cdot \Delta_1}(T) \subset J_{\Delta}(T) \cdot J_{\Delta_1}(T).$$

Therefore any two integral ranges of T overlap.

Suppose the sets of decomposition of \mathfrak{S} under Δ and Δ_1 are σ_i and σ'_i , respectively. Then if we denote $m(\sigma_i)$ by m_i , $T(\sigma_i)$ by B_i , $m(\sigma_i \cdot \sigma'_j)$ by m_j^i , and $T(\sigma_i \cdot \sigma'_j)$ by B_j^i , the hypotheses of Theorem 9 are clearly fulfilled. It follows that T is summable under the product $\Delta \cdot \Delta_1$ of Δ and Δ_1 , and that $J_{\Delta \cdot \Delta_1}(T) \subset J_{\Delta}(T)$. The rest of the conclusion follows by symmetry.

14. **The integrable functions and their integrals.** We are now ready for

DEFINITION 3. A function T will be called integrable if and only if the inferior limit of the diameters of its integral ranges is zero.

THEOREM 12. If T is integrable, then the intersection of the integral ranges of T is a single element $J(T)$ of \mathfrak{B} .

We can choose a set of integral ranges $J_{\Delta_1}(T)$, $J_{\Delta_2}(T)$, $J_{\Delta_3}(T)$, \dots of diameters <1 , $<\frac{1}{2}$, $<\frac{1}{4}$, \dots . Since these are closed and overlap, their intersection is a point. But since every integral range of T is closed and overlaps every $J_{\Delta_k}(T)$, this point is contained in every integral range of T .

DEFINITION 4. The $J(T)$ of Theorem 12 is called the integral of T over \mathfrak{S} .

THEOREM 13. T is integrable if and only if to every $\epsilon > 0$ corresponds a decomposition Δ under which the aggregate $m(\sigma_i) \cdot T(\sigma_i)$ is unconditionally summable and has a diameter $< \epsilon$.

For since the diameter is bounded, so is each $T(\sigma_i)$. And by Theorem 4, $\rho(J_{\Delta}(T)) < \epsilon$. And these are the only facts about integrability not assumed.

15. **Integrals are completely additive set functions.** In this section it will be shown that the integral of any integrable function T is a completely additive set function depending linearly on T . To this end we prove

THEOREM 14. If the function T is integrable over \mathfrak{S} , then it is integrable over every set σ of Σ to an element of \mathfrak{B} which will be denoted by $J(T, \sigma)$, and the set function $J(T, \sigma)$ is completely additive.

Let Δ be any decomposition of \mathfrak{S} into sets σ_i under which T is summable. Then writing $\sigma_i^1 = \sigma \cdot \sigma_i$ and $\sigma_i^2 = (\mathfrak{S} - \sigma) \cdot \sigma_i$, we see $T(\sigma_i^k) \subset T(\sigma_i)$ [$k=1, 2$] and $\sum_k m(\sigma_i^k) = m(\sigma_i)$. Therefore by Theorem 9

$$\sum_{i,k} m(\sigma_i^k) \cdot T(\sigma_i^k) \subset \sum_i m(\sigma_i) \cdot T(\sigma_i),$$

whence obviously

$$\sum_i m(\sigma_i^1) \cdot T(\sigma_i^1) + \sum_i m(\sigma_i^2) \cdot T(\sigma_i^2) \subset J_\Delta(T).$$

It is a corollary[†] that $\sum_i m(\sigma_i^1) \cdot T(\sigma_i^1)$ is of diameter at most $\rho(J_\Delta(T))$, and hence by Theorem 13 T is integrable over σ .

Similarly, if Δ and Δ_1 are any two decompositions of \mathfrak{S} into subsets σ_i and σ_i^1 respectively, then by Theorem 9

$$\sum_i J(T, \sigma_i) \subset \sum_i J_{\Delta \cdot \Delta_1}(T, \sigma_i) \subset J_{\Delta \cdot \Delta_1}(T) \subset J_\Delta(T),$$

whence, in the limit, $\sum_i J(T, \sigma_i) = J(T)$. Now replacing \mathfrak{S} by an arbitrary set σ of Σ , we complete the proof.

DEFINITION 5. By the "norm" of an integrable function T , is meant the real number $\|T\| \equiv \sup_{\sigma \in \Sigma} \|J(T, \sigma)\|$.

THEOREM 15. If Δ is any decomposition of \mathfrak{S} , if the function T is integrable over every set σ_i of composition of Δ , and the aggregate of the $J(T, \sigma_i)$ is unconditionally summable, then T is integrable over \mathfrak{S} and $J(T) = \sum_i J(T, \sigma_i)$.

Decompose each σ_i by a decomposition Δ_i under which $\|J_{\Delta_i}(T, \sigma_i) - J(T, \sigma_i)\| < \epsilon/2^i$. Then the corresponding decomposition of \mathfrak{S} will be summable, and its integrated range will be within a sphere of radius ϵ of $\sum_i J(T, \sigma_i)$.

THEOREM 16. If T and U are integrable functions, and m is a real number, then $m \cdot T$ and $T + U$ are integrable, $J(m \cdot T) = m \cdot J(T)$, and $J(T + U) = J(T) + J(U)$.

The conclusions about $m \cdot T$ are evident, since if $\rho(J_\Delta(T)) < \epsilon$, then $J_\Delta(m \cdot T) = m \cdot J_\Delta(T)$ is of diameter $< m\epsilon$. Those about $T + U = V$ follow since if $\rho(J_\Delta(T)) < \epsilon$ and $\rho(J_{\Delta_1}(U)) < \epsilon$, then

$$J_{\Delta \cdot \Delta_1}(V) \subset J_\Delta(T) + J_{\Delta_1}(U)$$

which is of diameter less than 2ϵ .

COROLLARY. $\|m \cdot T\| = |m| \cdot \|T\|$ and $\|T + U\| \leq \|T\| + \|U\|$.

16. Multiplication by a scalar function. We shall now prove a very powerful result,

[†] Since $\rho(A) \leq \rho(A+B) \leq \rho(A) + \rho(B)$, for any A and B .

THEOREM 17. *If T is integrable over \mathfrak{S} , and $f(p)$ is any real-valued bounded Lebesgue integrable function over \mathfrak{S} , then the function $U(p) = f(p) \cdot T(p)$ is integrable, and*

$$\|U\| \leq 2F \cdot \|T\| \quad \text{where} \quad F \equiv \sup_{p \in \mathfrak{S}} |f(p)|.$$

Let $\epsilon > 0$ be given. By taking the product of a finite number of suitably chosen decompositions, we arrive at a decomposition Δ into subsets σ_i which satisfies the following conditions.

- (1) On any σ_i , $-F \leq c_i \leq f(p) \leq c_i + \epsilon \leq F + \epsilon$.
- (2) The aggregate $m(\sigma_i) \cdot T(\sigma_i)$ is unconditionally summable.
- (3) $\|\sum_i m(\sigma_i) \cdot T(\sigma_i) - J(T)\| < \epsilon$.

Now by Theorem 8, the aggregates $m(\sigma_i) \cdot U(\sigma_i)$, $c_i \cdot J(T, \sigma_i)$, and $\epsilon_i \cdot J(T, \sigma_i)$ [$0 \leq \epsilon_i \leq \epsilon$] are unconditionally summable. Moreover for any choice of $p_i \in \sigma_i$,

$$\begin{aligned} & \left\| \sum_i m(\sigma_i) \cdot U(p_i) - \sum_i c_i \cdot J(T, \sigma_i) \right\| \\ (B) \quad & \leq \left\| \sum_i c_i \cdot [m(\sigma_i)T(p_i) - J(T, \sigma_i)] \right\| + \left\| \sum_i \epsilon_i m(\sigma_i) \cdot T(p_i) \right\| \\ & \leq 2F \cdot \epsilon + \epsilon[\|T\| + \epsilon] = \epsilon[2F + \|T\| + \epsilon]. \end{aligned}$$

(The first inequality comes from writing $U(p_i) = c_i T(p_i) + \epsilon_i T(p_i)$; the first half of the second inequality from Theorem 8, and the second half from assumption (3) and Theorem 8.)

But $\epsilon[2F + \|T\| + \epsilon]$ tends to zero with ϵ ; consequently, by Theorem 13, $U(p)$ is integrable. The relation $\|J(U)\| \leq 2F \cdot \|T\|$ now follows by Theorem 8, since $\|T\|$ bounds the norms of the partial sums of the $J(T, \sigma_i)$ and $|c_i| \leq F$. Replacing \mathfrak{S} by σ , we have similarly $\|J(U, \sigma)\| \leq 2F \cdot \sup_{\sigma} \|J(T, \sigma^*)\| \leq \|T\|$ (by definition of $\|T\|$). This completes the proof of the theorem.

17. Denseness of finite-valued functions. We shall next show that in a certain sense every integrable function can be approximated arbitrarily closely by finite-valued functions.

THEOREM 18. *Functions assuming only a finite number of distinct values are everywhere dense in the vector space of integrable functions normed by Definition 5.*

To any integrable function T and number $\epsilon > 0$ corresponds a decomposition Δ of \mathfrak{S} into subsets σ_i satisfying $\|J_\Delta(T) - J(T)\| < \frac{1}{2}\epsilon$. Moreover by Theorem 8 we can choose n so large that

$$\left\| \sum_{n+1}^{\infty} m_i^* \cdot T(\sigma_i) \right\| < \frac{1}{2}\epsilon$$

for all m_i^* satisfying $0 \leq m_i^* \leq m(\sigma_i)$.

Define the function U by the equations

$$U(p) = \begin{cases} J(T, \sigma_i)/m(\sigma_i) & \text{if } p \in \sigma_i, \quad m(\sigma_i) > 0, \quad \text{and } k \leq n, \\ \Theta & \text{otherwise.} \end{cases}$$

U is clearly finite-valued and integrable.

But if σ is any set of Σ , and $\sigma \cdot \sigma_i$ and $\sigma_i - \sigma \cdot \sigma_i$ are denoted by σ_i^1 and σ_i^2 respectively, then $0 \leq m(\sigma_i^1) \leq m(\sigma_i)$, and so

$$\left\| \sum_{n+1}^{\infty} J(T - U, \sigma_i^1) \right\| = \left\| \text{Co} \left(\sum_{n+1}^{\infty} m(\sigma_i^1) \cdot T(\sigma_i^1) \right) \right\| < \frac{1}{2}\epsilon.$$

Again, $\Theta \epsilon [J_{\Delta}(T, \sigma_i^k) - J(T, \sigma_i^k)]$ for all i, k . Hence

$$\sum_{i=1}^n [J_{\Delta}(T, \sigma_i^1) - J(T, \sigma_i^1)] \subset \sum_{i,k} [J_{\Delta}(T, \sigma_i^k) - J(T, \sigma_i^k)],$$

and so

$$\left\| \sum_{i=1}^n J(T - U, \sigma_i^1) \right\| \leq \|J_{\Delta}(T) - J(T)\| < \frac{1}{2}\epsilon.$$

By the complete additivity of $J(T, \sigma)$, we now get $\|J(T - U)\| < \epsilon$.

Theorem 18 can be extended even further to a result whose proof, although not difficult, is so long that we shall omit it. Accordingly, we state without proof

THEOREM 19. *The integrable functions of euclidean space to any separable Banach space \mathfrak{S} are a separable space under the norm of Definition 5.*

The only construction involved is that of replacing each of the first n σ_i by an approximating point-set sum of a finite number of intervals with rational coordinates, and $J(T, \sigma_i)/m(\sigma_i)$ by a nearby element of \mathfrak{B} . The everywhere dense functions are then defined as Θ except on a finite number of intervals with rational coordinates, and constant on each such interval.

18. Effect of linear transformation of the range. It is almost self-evident that all of the arguments so far are preserved under linear transformations, since these preserve both sums and limits. Accordingly, we prove

THEOREM 20. *If T is any integrable function of \mathfrak{S} to \mathfrak{B} , and $\alpha: \beta \rightarrow \alpha(\beta)$ is any linear transformation of \mathfrak{B} into the Banach space \mathfrak{A} , then (i) the function*

$U(p) \equiv \alpha(T(p))$ of \mathfrak{S} to \mathfrak{A} is integrable, (ii) $J(U) = \alpha(J(T))$, (iii) if we denote the modulus[†] of α by a , then $\|U\| \leq a \cdot \|T\|$.

If T is summable under a decomposition Δ of \mathfrak{S} , then so is U , and $J_\Delta(U) = \alpha(J_\Delta(T))$. This is true by definition of U for single terms $m(\sigma_i) \cdot U(\sigma_i)$; it remains true for finite sums since α is additive, and under passage to the limit since $\|\alpha(\beta) - \alpha(\beta')\|/\|\beta - \beta'\|$ is bounded. The proof is completed by letting $\rho(J_\Delta(T))$ tend to zero.

Theorem 20, and the fact that for real-valued functions our integral reduces to Fréchet's interpretation of the Lebesgue integral, leads us to formulate

THEOREM 21. *Let T be any integrable function of \mathfrak{S} to the Banach space \mathfrak{B} , and $f(\xi)$ a variable linear functional with domain \mathfrak{B} . For each f , $f(J(T))$ is the Lebesgue integral $\int_{\mathfrak{S}} f(T(p)) dm(\mathfrak{S})$ of the real function $f(T(p))$ with domain \mathfrak{S} , and so $J(T)$ is the intersection (as f varies) of the hyperplanes of elements η of \mathfrak{B} satisfying $f(\eta) = \int_{\mathfrak{S}} f(T(p)) dm(\mathfrak{S})$.*

The first statement is a corollary of Theorem 19, and the second follows from the fact[‡] that to every $\xi \neq J(T)$ there corresponds a linear functional f such that $f(\xi) \neq f(J(T))$.

We can deduce from Theorem 21 and known§ results

COROLLARY 1. *If $\sum_{k=1}^{\infty} \|T_k\| < +\infty$, $T(p) = \sum_{k=1}^{\infty} T_k(p)$ is defined almost everywhere and integrable, then $J(T) = \sum_{k=1}^{\infty} J(T_k)$.*

COROLLARY 2. *A function of a square to a Banach space \mathfrak{B} can only fail to satisfy the Theorem of Fubini because of non-integrability.*

COROLLARY 3. *If $T_0(t)$, $T_1(t)$, \dots , $T_n(t)$ are functions of the line interval $[0, x]$ to \mathfrak{B} such that*

$$T_k(t) - T_k(0) = J(T_{k+1}, [0, t])$$

for $k=0, 1, \dots, n-1$ and $0 \leq t \leq x$, and if B_x denotes the range of $T_n(t)$ on $[0, x]$, then

$$T_0(x) \subset \sum_{k=0}^{n-1} (x^k/k!) \cdot T_k(0) + (x^n/n!) \cdot \overline{\text{Co}(B_x)}.$$

(This is Taylor's formula with the remainder.)

[†] That is, $\sup_{\beta \neq 0} \|\alpha(\beta)\|/\|\beta\|$. Banach calls this the norm of α ; we have not conformed to his usage for fear of confusion with the other norms which we have defined.

[‡] Banach, p. 55.

[§] C. Carathéodory, *Vorlesungen über reelle Funktionen*, Leipzig, 1927. We refer to Theorem 13, p. 441, and Theorem 1, p. 627. I am stating Corollary 3 without reference.

19. **Uniform convergence and the method of iteration.** We can prove by a very superficial argument

THEOREM 22. *If T_1, T_2, T_3, \dots are integrable functions of \mathfrak{S} to \mathfrak{B} , $m(\mathfrak{S})$ is finite, and the T_n converge uniformly to T , then T is integrable and $J(T) = \lim_{n \rightarrow \infty} J(T_n)$.*

For n exists so large that $\|T(p) - T_n(p)\| < \epsilon/(2m(\mathfrak{S}))$ for all p , and Δ such that every element of $J_\Delta(T_n)$ lies within $\frac{1}{2}\epsilon$ of $J(T_n)$. Under these circumstances every element of $J_\Delta(T)$ lies within $\frac{1}{2}\epsilon$ of some element of $J_\Delta(T_n)$, and hence within ϵ of $J(T_n)$. The theorem is now obvious.

THEOREM 23. *If $K(x, y)$ is any function of the square $0 \leq x, y \leq 1$ to \mathfrak{B} which is uniformly continuous in x , Lebesgue integrable in y , and satisfies $\|K(x, y)\| \leq M < +\infty$, then the integral equation*

$$(1) \quad S(x) = T(x) + c \int_0^1 K(x, y) S(y) dy$$

has a unique solution $S(x)$ provided $2cM < 1$.

First we construct a solution by iteration. We set $S_k(x) = \Theta$, and $S_{k+1}(x) = T(x) + c \int_0^1 K(x, y) \cdot S_k(y) dy$. By Theorem 17 these integrals exist, and satisfy

$$\|S_{k+1}(x) - S_k(x)\| = [2cM]^k \cdot \|T\|$$

by induction and Theorem 17. Hence the series of differences converges uniformly to a limit satisfying (1).

Theorem 17 may also be used to show that as the homogeneous equation corresponding has no solution, the solution obtained in this way by iteration is unique.

As a matter of fact, inspection shows that the entire Fredholm theory of integral equations of the form (1) carries over to the case where $S(x)$ and $T(x)$ are permitted to be vector functions. The only change is that when the determinant of $K(x, y)$ vanishes, the solutions to the equation

$$(2) \quad S(x) = c \int_0^1 K(x, y) S(y) dy$$

are of the form $f_1(x)\xi_1 + \dots + f_r(x)\xi_r$, where the $f_i(x)$ are linearly independent *real* solutions of (2), and the ξ_i are arbitrary vector constants.

20. Consequences of separability. Since the sum of enumerable sets of measure zero is again of measure zero, it is possible to extend certain theorems concerning the Lebesgue integral to our integral in the case that \mathfrak{B} is *separable*, that is, contains an enumerable everywhere dense set of elements.

In the first place, if we define two functions T and U of \mathfrak{S} to \mathfrak{B} as “equivalent” if and only if $T(p) = U(p)$ except on a set of measure zero, then

THEOREM 24. *If \mathfrak{B} is separable, then two point functions T and U of \mathfrak{S} to \mathfrak{B} give rise to the same set function $J(T, \sigma) = J(U, \sigma)$ if and only if T and U are equivalent.*

The sufficiency of equivalence is obvious. But conversely, if $J(T, \sigma) = J(U, \sigma)$, then by Theorem 21 and the theory of the Lebesgue integral, $f(T(p)) = f(U(p))$ except on a set of measure zero, for any functional $f(\xi)$. But we can define functionals $f_i(\xi)$ of modulus unity† such that $f_i(\xi_i) = \|\xi_i\|$ for each of an enumerable everywhere dense set of elements ξ_i of \mathfrak{B} .

But if $T(p) - U(p) = \xi \neq \Theta$, then choosing ξ_i satisfying $\|\xi_i - \xi\| < \frac{1}{2}\|\xi\|$, clearly $f_i(\xi) > 0$. Hence except on the set of measure zero on which some $f_i(T(p) - U(p)) \neq 0$, $T(p) = U(p)$. This proves Theorem 24.

COROLLARY. *If \mathfrak{B} is separable, then every integrable function is single-valued except on a set of measure zero.*

Again, suppose \mathfrak{S} is the line interval $[a, b]$. A function $T(x)$ of \mathfrak{S} to \mathfrak{B} is said to have the “strong” derivative $T'(x)$ at the point x if and only if $(1/h) \cdot [T(x+h) - T(x)]$ tends to $T'(x)$ as $|h|$ tends to zero; it is said to have the “weak” derivative $T'(x)$ if and only if for every linear functional $f(\xi)$,

$$\lim_{|h| \rightarrow 0} [f(T(x+h)) - f(T(x))]/h = f(T'(x)).$$

THEOREM 25. *If the space of the functionals of \mathfrak{B} is separable, then the integral $J(T, [a, x])$ of any integrable function T of a line interval to \mathfrak{B} is weakly differentiable except on a set of measure zero.*

Theorem 25 is a corollary of Theorem 21 and the truth of the theorem for the real functions $f(J(T, [a, x]))$.

21. Rectifiable curves in Hilbert space. It is well known that any real function of bounded variation is differentiable almost everywhere. We shall extend this proposition.

THEOREM 26. *Every function $h(x)$ of bounded variation‡ of an interval $[a, b]$ to Hilbert space \mathfrak{H} is (strongly) differentiable except on a set of measure zero. Moreover its derivative is integrable to $h(b) - h(a)$.*

In the first place, if \mathfrak{F} and \mathfrak{G} are orthogonal complements§ in \mathfrak{H} , then the

† Banach, p. 55. Of course the f_i are not in general everywhere dense in the space of the functionals of \mathfrak{B} .

‡ I.e., such that $\sup \sum_i \|h(x_i) - h(x_{i-1})\| < +\infty$ ($x_0 < \dots < x_n$). This upper limit we shall call the “total variation” of $h(x)$.

§ We have borrowed without explanation standard terms such as orthogonal complement and orthonormal vector from M. H. Stone's *Linear Transformations in Hilbert Space*, New York, 1932.

projections $f(x)$ of $h(x)$ onto \mathfrak{F} and $g(x)$ of $h(x)$ onto \mathfrak{G} are both of bounded variation. While if we denote by $|f|$, $|g|$, and $|h|$ the total variations of $f(x)$, $g(x)$, and $h(x)$ respectively, then

$$(1) \quad (|f|^2 + |g|^2)^{1/2} \leq |h| \leq |f| + |g|.$$

Now divide $[a, b]$ so finely by points $a = x_0 < \dots < x_n = b$ that

$$(2) \quad |h| - \epsilon = \sum_1^n \|h(x_i) - h(x_{i-1})\| \leq |h|.$$

The $h(x_i)$ will all lie on a linear manifold \mathfrak{F} of dimensions at most $n+1$. And if \mathfrak{G} is the orthogonal complement of \mathfrak{F} , then by elementary algebra and conditions (1)–(2),

$$(3) \quad |g| \leq (|h|^2 - |f|^2)^{1/2} \leq (2\epsilon \cdot |h|)^{1/2}.$$

Further, construct the non-decreasing real variation function

$$v_g(x) = \int_a^x \|dg(t)\| \leq |g| \leq (2\epsilon \cdot |h|)^{1/2}.$$

By the theory of real functions, $v_g'(x)$ exists and does not exceed $(2\epsilon \cdot |h|)^{1/4}$ except on a set S of measure at most $(2\epsilon \cdot |h|)^{1/4}$. But clearly

$$(4) \quad \left\| \frac{g(x + \Delta x) - g(x)}{\Delta x} \right\| \leq \frac{1}{\Delta x} \int_x^{x+\Delta x} \|dg(t)\| = \frac{1}{\Delta x} \int_x^{x+\Delta x} dv_g.$$

Therefore if x is not on S , then, for small enough Δx ,

$$(5) \quad \left\| \frac{h(x + \Delta x) - h(x)}{\Delta x} - \frac{f(x + \Delta x) - f(x)}{\Delta x} \right\| = \left\| \frac{g(x + \Delta x) - g(x)}{\Delta x} \right\| \leq (2\epsilon \cdot |h|)^{1/4}.$$

Consequently except on a set S of measure $= (2\epsilon \cdot |h|)^{1/4}$ of $[a, b]$, for each x and sufficiently small Δx , the difference quotients $\|\{g(x + \Delta x) - g(x)\}/\Delta x\|$ are bounded by $(2\epsilon \cdot |h|)^{1/4}$.

But since \mathfrak{F} has only a finite number of dimensions, all the components of $f(x)$, and hence $f(x)$ itself, have finite derivatives except on a set S' of measure zero. Therefore if $|\Delta x| < M_x$, a sufficiently small positive function of x , then the diameter of the set of the $\{h(x + \Delta x) - h(x)\}/\Delta x$ is at most $2 \cdot (2\epsilon \cdot |h|)^{1/4}$ for any x not on $S \cup S'$ [of measure $\leq (2\epsilon \cdot |h|)^{1/4}$]. The proof of the first assertion is completed by letting ϵ tend to zero.

To prove the second assertion, decompose $[a, b]$ so that the integral ranges of $f'(x)$ and $v_g'(x)$ are both of diameters $< \frac{1}{2}\epsilon$. It will follow that the

integral range of $h(x)$ lies within $\epsilon + |g|$ of $f(b) - f(a)$. The remainder of the proof is obvious.

If we use the variation-function $s(x) = \int_a^x \|dh(t)\|$ as the parameter for the domain, then we get the geometrical

COROLLARY 1. *Any rectifiable curve in Hilbert space has a tangent at almost every point.*

We are more interested in the analytical

COROLLARY 2. *If $h(t)$ is a bounded function of a line interval to Hilbert space, and $g(x) = J(h, [a, x])$, then $g(x)$ is (strongly) differentiable almost everywhere to $h(x)$.*

Corollary 2 depends on Theorem 24.

22. Examples of non-differentiable integrals. In this and the following section we shall give seven examples illustrating various theoretical points. In all seven we shall understand the symbol \mathfrak{H} to denote Hilbert space, and \mathfrak{B} to denote the space of real bounded functions $y = y(x)$ on the interval $[0, 1]$, with the norm $\|y\| = \sup_x y(x)$. We shall also use $\xi_{i,j}$ to denote a doubly infinite set of orthonormal vectors in \mathfrak{H} .

We shall first show by two examples that the hypotheses of Corollary 2 of Theorem 26 are indispensable.

EXAMPLE 1. *There exists a totally discontinuous bounded integrable function of the interval $[0, 1]$ to \mathfrak{B} , whose integral, although of bounded variation, is nowhere even weakly differentiable.*

Let y_r in \mathfrak{B} correspond to the function $y_r(x) = 0$ on $[0, r]$, $y_r(x) = 1$ on $[r, 1]$. Graves observed[†] that the function $T(r) = y_r$ of $[0, 1]$ to \mathfrak{B} was bounded, totally discontinuous, and integrable. Since $y(a)$ is a linear functional $f_a(y)$ of the elements $y: y(x)$ of \mathfrak{B} , and $U(s) = J(T, [0, s])$ is not differentiable at a with respect to f_a , $U(s)$ is nowhere even weakly differentiable, although there is a bound on the norm of its differential quotients $\Delta U / \Delta s$.

EXAMPLE 2. *There exists an integrable function of the interval $[0, 1]$ to \mathfrak{H} , whose integral is nowhere strongly differentiable.*

Set $T_i(t) = 2^i \cdot \xi_{i,j}$ on $[j/2^i, 2^{-i} + j/2^i]$ [$j = 0, \dots, 2^i - 1$]. The function $V(t) = \sum_1^\infty T_i(t)$ exists almost everywhere, and is integrable. The proofs of these facts, and of the nowhere strong differentiability of $W(x) = J(V, [0, x])$, are left to the reader.

[†] Loc. cit., p. 164. Cf. also S. Bochner, *Absolut-additive Mengenfunktionen*, *Fundamenta Mathematicae*, vol. 21 (1933), pp. 211-13, in which a function essentially like $U(s)$ is shown to be nowhere strongly differentiable.

Example 1 is not integrable in the sense of Bochner, and Example 2 is not integrable in the sense of Graves. This shows that our integral is included in neither [cf. §24].

23. Other counterexamples. We shall list below examples showing that the hypotheses of various other theorems proved above cannot be eliminated.

EXAMPLE 3. *There exists a function T of the interval $[0, 1]$ to \mathfrak{B} such that $\|T(x)\| \equiv 1$ and yet $\|T\| = 0$.*

Let T assign to each point x of $[0, 1]$ the function $f_x: f_x(x) = 1, f_x(t) = 0$ for $t \neq x$. The proofs that $T(x)$ is integrable and that $J(T, \sigma) = 0$ for any σ of Σ are left to the reader. This example shows that we cannot drop the hypothesis of separability in Theorem 24.

EXAMPLE 4. *To any $\epsilon > 0$ corresponds a function T of the interval $[0, 1]$ to \mathfrak{S} such that $\|T(x)\| \equiv 1$ and yet $\|T\| < \epsilon$.*

Let us choose n so large that $n\epsilon > 1$. Let $T(x) = \xi_{k,1}$ on $[(k-1)/n^2, k/n^2)$ [$k = 1, \dots, n^2$]. The proof that $\|T\| < \epsilon$ is left to the reader.

EXAMPLE 5. *There exists a Cauchy sequence of integrable functions of the interval $[0, 1]$ to \mathfrak{S} converging (relative to the norm of Definition 5) to no limit function.*

Let $T_i(x) = \xi_{i,j}$ on $[(j-1)/2^i, j/2^i)$ [$j = 1, \dots, 2^i$]. Evidently $\|T_i\| = 2^{-i/2}$; hence if $U_n = T_1 + \dots + T_n$, then the sequence $\{U_n\}$ is a Cauchy sequence (under Definition 5).

Suppose $U(x)$ existed, satisfying $\lim_{n \rightarrow \infty} \|U_n - U\| = 0$. By Theorem 21 the scalar product (which is a linear functional) of $U(x)$ with each $\xi_{i,j}$ would have to be that of $T_i(x)$ with $\xi_{i,j}$, except on a set of measure zero. But there can be no $U(x)$ in \mathfrak{S} having the necessary $\xi_{i,j}$ -components; the sum of the squares tends to infinity.

Example 5 shows that although integrable functions constitute a normed vector space under Definition 5 (provided we consider two integrands as the "same" if they give rise to the same set function), this space is not in general complete.

EXAMPLE 6. *There is a function of the unit square to \mathfrak{S} not satisfying the Theorem of Fubini.*

We shall use the functions $T_i(x)$ of Example 5, and set $U(x, y) = 2^{-i} \cdot T_i(x)$ on $2^{-i} \leq y < 2^{-i+1}$, and Θ where not otherwise defined. The function $U(x, y)$ cannot be integrated with respect to y on a single line $x = \text{constant}$, and yet it is integrable over the square to

$$\sum_{i=1}^{\infty} \sum_{j=1}^{2^i} (2^{-i}) \cdot \xi_{i,j}.$$

EXAMPLE 7. *There exists an integrable function of the interval $[0, 1)$ to \mathfrak{S} , such that $\|T(x)\|$ is not integrable.*

Set $T(x) = (2^i/i) \cdot \xi_{i,1}$ on $[2^{-i}, 2^{-(i-1)})$. The proofs are left to the reader.

24. **Relation to integrals of Graves and Bochner.** In §22 it was noted that the integral used in this paper was included in neither Graves' nor Bochner's. It will now be shown that on the contrary it includes both.

Without repeating Graves' definition, it may be said that if $T(x)$ is integrable in his sense, then the integral of T is arbitrarily near every Riemann sum $\sum_i T(x_i) \Delta x_i$ under a suitable decomposition of the domain of T into a finite number of intervals. But if this is so, then by Theorems 12–13 proved above, the function is integrable to the same integral in the sense of Definitions 1–4.

Again, let $T(x)$ be any function integrable in the sense of Bochner over an abstract domain \mathfrak{S} of finite measure M . That is, suppose that

- (1) the real function $\|T(p)\|$ is summable;
- (2) $T(p) = \lim_{n \rightarrow \infty} T_n(p)$ for almost every p in \mathfrak{S} ;
- (3) each $T_n(p)$ is a "finite-valued" function.

Let $\epsilon > 0$ be given. On a set \mathfrak{S}_ϵ of measure $> M - \epsilon$, the $T_n(p)$ converge uniformly to $T(p)$.

Again, let $\delta > 0$ be given. Choose n so large that $\|T(p) - T_n(p)\| < \delta/M$ on \mathfrak{S}_ϵ . Then decompose \mathfrak{S}_ϵ into sets σ_i [$i = 1, \dots, s$] of measures m_i , on which $T_n(p)$ is constant and has the values ξ_i .

The integral range $J_\Delta(T, \mathfrak{S}_\epsilon)$ under this decomposition clearly lies within $M \cdot (\delta/M) = \delta$ of $\sum_{i=1}^s m_i \cdot \xi_i$. Hence T is integrable in my sense over \mathfrak{S}_ϵ , and the integral $J(T)$ of T is $\lim_{n \rightarrow \infty} \sum_{i=1}^n m_i \cdot \xi_i$, i.e., equal to Bochner's integral.

Since both Bochner's integral and mine are completely additive set functions, this can be extended to \mathfrak{S} , and thence to domains which are the sums of enumerable sets of finite measure.

25. **Open questions.** There are several open questions of a technical nature concerning the subject we have just treated.

One is as to the replaceability of the closure of the convex hull by the convex hull alone, without affecting the program. For instance, is $\sum_i \text{Co}(B_i) \subset \text{Co}(\sum_i B_i)$? The accomplishment of this might simplify the definitions and argument.

Another is that of generalizing the integral to the case that the range is merely a *topological* vector space, in which $\alpha_n \rightarrow \alpha$ and $\beta_n \rightarrow \beta$ imply $\alpha_n + \beta_n \rightarrow \alpha + \beta$, and $b_n \rightarrow b$ and $\beta_n \rightarrow \beta$ imply $b_n \beta_n \rightarrow b \cdot \beta$. Both these problems are trivial if we restrict ourselves to Graves' Riemann integral, but with unconditional convergence, they become very deep.

Two questions which I have been unable to answer are, do there exist (non-integrable) functions, the intersection of whose integral ranges is vacuous? or such that when we apply Theorem 21, the intersection of the hyperplanes

$$f(\eta) = \int_{\mathfrak{E}} f(T(p)) dm(\mathfrak{E})$$

turns out to be vacuous? [Added in proof: the answer to the second question is yes.]

Is it true that if the "conjugate" space of the linear functionals of \mathfrak{B} is separable, then the integral of any bounded integrable function is strongly differentiable almost everywhere? Or that every rectifiable curve in \mathfrak{B} has a tangent at almost every point?

And is it true that if unconditional convergence implies absolute convergence in \mathfrak{B} , then the vector space of integrable functions is complete under the norm of Definition 5?

It may be proved that any vectoroid space satisfying

$$V7: (b_1 + b_2) \cdot B = b_1 \cdot B + b_2 \cdot B$$

is a vector space. Is it true that if we merely require

$$V7': B + (-1) \cdot B = \Theta$$

we can still deduce all the properties of vector space? How, if at all, can $V1-V6$ be replaced by properties of the calculus of vector complexes in such a way that the addition of $V7'$ does give a complete set of axioms for vector spaces?

SOCIETY OF FELLOWS,
HARVARD UNIVERSITY,
CAMBRIDGE, MASS.