

UNIFORMLY CONVEX SPACES*

BY

JAMES A. CLARKSON†

1. INTRODUCTION

The spaces with which we shall deal in this paper are Banach spaces, that is, linear, metric, complete, normed spaces, which possess in addition a certain property of *convexity* of the norm. Expressed in geometrical terms this property is simple: it is that the mid-point of a variable chord of the unit sphere of the space cannot approach the surface of the sphere unless the length of the chord goes to zero. Additional interest is given to the notion by the fact that, as we shall prove, the well known spaces L_p and l_p possess this property for p exceeding unity.

Several writers‡ have considered the problem of defining an integral of a function whose domain is in Euclidean space (or even a more general space) and whose range lies in a Banach space. Bochner§ has pointed out that such a function may be absolutely continuous without being an integral in his sense, or indeed without being differentiable at any point. We shall prove that if the range space is uniformly convex in our sense such phenomena do not occur, and that for these spaces the situation is quite analogous to the theory for ordinary complex functions.

2. UNIFORMLY CONVEX SPACES

Let B denote a Banach space, with elements $x, y, \dots$. We denote the norm of an element x by $\|x\|$.

DEFINITION 1. *A Banach space B will be said to be uniformly convex if to each ϵ , $0 < \epsilon \leq 2$, there corresponds a $\delta(\epsilon) > 0$ such that the conditions*

$$\|x\| = \|y\| = 1, \quad \|x - y\| \geq \epsilon$$

* Presented to the Society, February 29, 1936; received by the editors January 21, 1936.

† National Research Fellow.

‡ L. M. Graves, *Riemann integration and Taylor's theorem in general analysis*, these Transactions, vol. 29 (1927), pp. 163–172.

S. Bochner, *Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20 (1933), pp. 262–276.

N. Dunford, *Integration in general analysis*, these Transactions, vol. 37 (1935), pp. 441–453.

G. Birkhoff, *Integration of functions with values in a Banach space*, these Transactions, vol. 38 (1935), pp. 357–378.

§ See Bochner, *Absolut-additive abstracte Mengenfunktionen*, Fundamenta Mathematicae, vol. 21 (1933), pp. 211–213. We take this opportunity to acknowledge our indebtedness to Professor Bochner for his suggestions concerning this paper.

imply

$$\left\| \frac{x+y}{2} \right\| \leq 1 - \delta(\epsilon).$$

We remark that Euclidean spaces of all dimensions, Hilbert space, and hyper-Hilbert spaces,* are all uniformly convex. This follows, for example, from the identity

$$(1) \quad \|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

which is known† to be characteristic of such spaces.

We shall need to recall the definition of the *product*‡ of a finite number of Banach spaces. Let $B_1, B_2, \dots, B_k$ be k Banach spaces with elements $x^1, x^2, \dots, x^k$. The product space $B = B_1 \times B_2 \times \dots \times B_k$ is defined as the set of all ordered k -tuples $x = (x^1, x^2, \dots, x^k)$, where addition and complex multiplication are defined in obvious fashion, and where the norm of the elements x of B is required, in addition to the usual properties, to have the property that $\|x\| \rightarrow 0$ is equivalent to $\|x^i\| \rightarrow 0$ ($i = 1, 2, \dots, k$). B is also a Banach space.

It is clear that the product of uniformly convex Banach spaces is not in general uniformly convex unless we require something more of the norm in the product space. In our first theorem we lay down a condition sufficient for this.

Let $N(a_1, a_2, \dots, a_k)$ be a non-negative continuous function of the non-negative variables a_i . We say that N is

(a) *homogeneous*, if for $c \geq 0$,

$$N(ca_1, ca_2, \dots, ca_k) = cN(a_1, a_2, \dots, a_k);$$

(b) *strictly convex*, if

$$N(a_1 + b_1, a_2 + b_2, \dots, a_k + b_k) < N(a_1, a_2, \dots, a_k) + N(b_1, b_2, \dots, b_k)$$

unless $a_i = cb_i$ ($i = 1, 2, \dots, k$). In the latter case we have equality by condition (a);

(c) *strictly increasing*, if it is strictly increasing in each variable separately.

A familiar example of a function N satisfying these conditions is

* That is, those spaces which satisfy all the postulates of Hilbert space except that of separability. See J. v. Neumann, *Mathematische Grundlagen der Quantenmechanik*, Berlin, 1932, pp. 37–38, for discussion of an example.

† J. v. Neumann and Jordan, *On inner products in linear metric spaces*, Annals of Mathematics, vol. 36 (1935), pp. 719–724.

‡ Cf. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 181.

$N = (\sum_{i=1}^k a_i^p)^{1/p}$ ($p > 1$); here condition (b) becomes the inequality of Minkowski.

Suppose now that a finite number of Banach spaces, $B_1, B_2, \dots, B_k$ are given, and that B is their product. We shall call B a *uniformly convex product* of the B_i if the norm of an element $x = (x^1, x^2, \dots, x^k)$ of B is defined by

$$\|x\| = N(\|x^1\|, \|x^2\|, \dots, \|x^k\|),$$

where N is a continuous non-negative function satisfying the conditions (a)-(c).

We now prove

THEOREM 1. *The uniformly convex product of a finite number of uniformly convex Banach spaces is uniformly convex.*[†]

We see at once that the norm thus defined satisfies the usual rules, including the triangle inequality. Let $B = B_1 \times B_2 \times \dots \times B_k$, and let $\{x_i\} = \{(x_i^1, x_i^2, \dots, x_i^k)\}$, $\{y_i\} = \{(y_i^1, y_i^2, \dots, y_i^k)\}$ be two sequences of points of B such that $\|x_i\| = \|y_i\| = 1$, $\|x_i + y_i\| \rightarrow 2$ as $i \rightarrow \infty$. We must show that $\|x_i - y_i\| \rightarrow 0$.

In the first place we assert that $\lim_{i \rightarrow \infty} (\|x_i^j\| - \|y_i^j\|) = 0$ ($j = 1, 2, \dots, k$). Indeed, we have, using conditions (a)-(c),

$$\begin{aligned} \|x_i + y_i\| &= N(\|x_i^1 + y_i^1\|, \dots, \|x_i^k + y_i^k\|) \\ &\leq N(\|x_i^1\| + \|y_i^1\|, \dots, \|x_i^k\| + \|y_i^k\|) \\ &\leq N(\|x_i^1\|, \dots, \|x_i^k\|) + N(\|y_i^1\|, \dots, \|y_i^k\|) \\ &= \|x_i\| + \|y_i\| = 2, \end{aligned}$$

so that

$$N(\|x_i^1\| + \|y_i^1\|, \dots, \|x_i^k\| + \|y_i^k\|) \rightarrow 2.$$

Now the last expression is < 2 unless $\|x_i^j\| = \|y_i^j\|$ ($j = 1, 2, \dots, k$) by (b) and our normalization; by the continuity of N our above assertion then follows.

Assume that $\|x_i - y_i\|$ does not $\rightarrow 0$. We may then by selecting an appropriate subsequence of i 's assume that

* See, for example, Hardy, Littlewood and Pólya, *Inequalities*, Cambridge, 1934, p. 31.

† In general, of course, $\|x\|$ will depend upon the x_i themselves, and not merely upon their norms. If this is the case, however, the question of the uniform convexity of the product space is entirely independent of the uniform convexity of the factor spaces, and no theorem of the type of Theorem 1 can be stated.

$$(2) \quad \begin{cases} \lim_{i \rightarrow \infty} \|x_i - y_i\| = \alpha > 0; \\ \lim_{i \rightarrow \infty} \|x_i^j\| = \lim_{i \rightarrow \infty} \|y_i^j\| = \beta_j & (j = 1, 2, \dots, k); \\ \lim_{i \rightarrow \infty} \|x_i^1 - y_i^1\| = \gamma > 0; \\ \|x_i^1\| > 0, \quad \|y_i^1\| > 0 & (i = 1, 2, \dots). \end{cases}$$

Set

$$w_i = \beta_1 x_i^1 / \|x_i^1\|,$$

$$z_i = \beta_1 y_i^1 / \|y_i^1\|.$$

Then

$$\lim_{i \rightarrow \infty} \|w_i - x_i^1\| = \lim_{i \rightarrow \infty} \|z_i - y_i^1\| = 0,$$

so that

$$\lim_{i \rightarrow \infty} \|w_i - z_i\| = \gamma.$$

Then since B_1 is uniformly convex,

$$\limsup_{i \rightarrow \infty} \|w_i + z_i\| < 2\beta_1,$$

and hence

$$\limsup_{i \rightarrow \infty} \|x_i^1 + y_i^1\| < 2\beta_1.$$

Since by (2)

$$\limsup_{i \rightarrow \infty} \|x_i^j + y_i^j\| \leq 2\beta_j \quad (j = 2, 3, \dots, k),$$

it follows that

$$\begin{aligned} \limsup_{i \rightarrow \infty} \|x_i + y_i\| &\leq N \left(\limsup_{i \rightarrow \infty} \|x_i^1 + y_i^1\|, \dots, \limsup_{i \rightarrow \infty} \|x_i^k + y_i^k\| \right) \\ &< N(2\beta_1, \dots, 2\beta_k) = 2N(\beta_1, \dots, \beta_k) = 2, \end{aligned}$$

from which contradiction of our assumption the theorem follows.

3. SPACES L_p AND l_p ($p > 1$)

We might now attempt to prove the uniform convexity of space l_p by extending the argument of Theorem 1 to an infinite number of factors. We prefer to prove this fact and the corresponding statement for L_p by exhibiting a set of inequalities for these spaces which are in close analogy with the identity (1). We collect these in

THEOREM 2. For space L_p or l_p , with $p \geq 2$, the following inequalities between the norms of two arbitrary elements x and y of the space are valid (here q is the conjugate index, $q = p/(p-1)$):

$$(3) \quad 2(\|x\|^p + \|y\|^p) \leq \|x+y\|^p + \|x-y\|^p \leq 2^{p-1}(\|x\|^p + \|y\|^p);$$

$$(4) \quad 2(\|x\|^p + \|y\|^p)^{q-1} \leq \|x+y\|^q + \|x-y\|^q;$$

$$(5) \quad \|x+y\|^p + \|x-y\|^p \leq 2(\|x\|^q + \|y\|^q)^{p-1}.$$

For $1 < p \leq 2$ these inequalities hold in the reverse sense.

We first point out that for all values of p the right-hand side of (3) is equivalent to the left-hand side, while (4) is equivalent to (5); to see this, set $x+y=\xi$, $x-y=\eta$ and reduce. We make full use of this fact in proving the theorem.

Proof of (4). First consider $1 < p \leq 2$. We commence by showing that for x, y , any two complex numbers, we have

$$(6) \quad |x+y|^q + |x-y|^q \leq 2(|x|^p + |y|^p)^{q-1}.$$

To prove this, assume that $|x| \geq |y|$, and divide (6) by $|x|^q$, reducing (6) to*

$$(7) \quad |1+c|^q + |1-c|^q \leq 2(1+|c|^p)^{q-1},$$

with $|c| \leq 1$. Setting $c = \rho e^{i\theta}$, we see by elementary calculus methods that it suffices to consider $\theta=0$; i.e., $0 \leq c \leq 1$; (7) is trivial for $c=0$ or $c=1$, so we need only consider $0 < c < 1$. Making the further transformation $c = (1-z)/(1+z)$ (so that $0 < z < 1$) we reduce (7) to the form

$$S \equiv \frac{1}{2}\{(1+z)^p + (1-z)^p\} - (1+z^q)^{p-1} \geq 0.$$

Expanding each term of S in its Taylor's series, we have

$$\begin{aligned} \frac{1}{2}\{(1+z)^p + (1-z)^p\} &= 1 + \frac{p(p-1)}{2!}z^2 + \frac{p(p-1)(2-p)(3-p)}{4!}z^4 + \dots \\ &\quad + \frac{p(p-1)(2-p) \cdots (2k-1-p)}{(2k)!}z^{2k} + \dots; \\ (1+z^q)^{p-1} &= 1 + (p-1)z^q - \frac{(p-1)(2-p)}{2!}z^{2q} + \dots \\ &\quad + \frac{(p-1)(2-p) \cdots (2k-1-p)}{(2k-1)!}z^{(2k-1)q} \\ &\quad - \frac{(p-1)(2-p) \cdots (2k-p)}{(2k)!}z^{2kq} + \dots. \end{aligned}$$

* The proof of (7) given here is due to J. S. Frame.

Hence

$$\begin{aligned}
 S &= \sum_{k=1}^{\infty} \left[\frac{p(p-1)(2-p) \cdots (2k-1-p)}{(2k)!} z^{2k} \right. \\
 &\quad \left. - \frac{(p-1)(2-p) \cdots (2k-1-p)}{(2k-1)!} z^{(2k-1)q} \right. \\
 &\quad \left. + \frac{(p-1)(2-p) \cdots (2k-p)}{(2k)!} z^{2kq} \right] \\
 &= \sum_{k=1}^{\infty} \frac{(2-p)(3-p) \cdots (2k-p)}{(2k-1)!} z^{2k} \left[\frac{1 - z^{(2k-p)/(p-1)}}{(2k-p)/(p-1)} - \frac{1 - z^{2k/(p-1)}}{2k/(p-1)} \right].
 \end{aligned}$$

But $(1-z^t)/t$, for $t > 0$ and $0 < z < 1$, is a decreasing function of t ; hence the series for S has non-negative terms, and (7) is established.

Turning now to the proof of (4), we consider first space l_p . Let the two elements considered be $x = (x_1, x_2, \cdots)$, $y = (y_1, y_2, \cdots)$; (4) states that

$$\begin{aligned}
 (8) \quad &\left[\sum_{i=1}^{\infty} |x_i + y_i|^p \right]^{q/p} + \left[\sum_{i=1}^{\infty} |x_i - y_i|^p \right]^{q/p} \\
 &\leq 2 \left[\sum_{i=1}^{\infty} (|x_i|^p + |y_i|^p) \right]^{q-1}.
 \end{aligned}$$

Now one form of Minkowski's Inequality states that if A_i, B_i are any two sets of non-negative numbers, finite or infinite in number, and $0 < s \leq 1$, then

$$(\sum A_i^s)^{1/s} + (\sum B_i^s)^{1/s} \leq (\sum (A_i + B_i)^s)^{1/s},$$

proper account being taken of convergence. Setting

$$\frac{p}{q} = s, \quad |x_i + y_i|^q = A_i, \quad |x_i - y_i|^q = B_i,$$

we infer that the left side of (8) is

$$\leq \left[\sum_{i=1}^{\infty} (|x_i + y_i|^q + |x_i - y_i|^q)^{p/q} \right]^{q/p}$$

which by (6) is

$$\leq \left[\sum_{i=1}^{\infty} (2\{|x_i|^p + |y_i|^p\}^{q/p})^{p/q} \right]^{q/p} = \left[\sum_{i=1}^{\infty} 2^{p/q} \{|x_i|^p + |y_i|^p\} \right]^{q/p}.$$

Since $q/p = q-1$, this is our result; (4), then, stands proved for l_p ($1 < p \leq 2$).

To extend this result to space L_p ($1 < p \leq 2$), let $[0, 1]$ be the interval over which the functions of our space are to be defined. We consider first two func-

tions $x(t)$, $y(t)$ which are step functions on a division of $[0, 1]$ into equal parts. It is readily verified that for such functions the relation (4) reduces to the l_p case already treated, and as these functions form a dense set in L_p the result follows by continuity of the norm.

Consider now (4) for $p > 2$. Again let x , y be any two elements of l_p ; the relation which we must prove is (8) with the sense reversed. Letting A_i , B_i , and s have the same values as above, and again applying Minkowski's Inequality, which is now reversed in sense since s exceeds 1, we conclude that the left side of (8) is

$$(9) \quad \geq \left[\sum_{i=1}^{\infty} (|x_i + y_i|^q + |x_i - y_i|^q)^{p/q} \right]^{q/p}.$$

Now (5) (or its equivalent, (4)) has already been proved for $p < 2$; hence employing it for complex numbers we have that for $p > 2$,

$$|x_i + y_i|^q + |x_i - y_i|^q \geq 2(|x_i|^p + |y_i|^p)^{q-1}.$$

From this we have that (9) is

$$\geq \left[\sum_{i=1}^{\infty} (2\{|x_i|^p + |y_i|^p\}^{q-1})^{p/q} \right]^{q/p} = 2 \left[\sum_{i=1}^{\infty} (|x_i|^p + |y_i|^p) \right]^{q-1},$$

our result for l_p . The passage to space L_p is carried out exactly as in the case for $1 < p \leq 2$, with which the demonstration of relations (4) and (5) is complete.

Proof of (3). We take $p \geq 2$, and consider the right-hand inequality. We show that this is implied by (5): that for a , $b \geq 0$ we have

$$(10) \quad 2(a^q + b^q)^{p-1} \leq 2^{p-1}(a^p + b^p).$$

For suppose $a \leq b > 0$, which entails no real loss of generality; dividing (10) by $b^{q(p-1)} = b^p$ we obtain

$$2(c^q + 1)^{p-1} \leq 2^{p-1}(c^p + 1) \quad (0 \leq c \leq 1),$$

or

$$2^{p-2} \frac{c^p + 1}{(c^q + 1)^{p-1}} \geq 1,$$

which, being raised to the power $1/p$, gives

$$H(c) \equiv 2^{(p-2)/p} \frac{(c^p + 1)^{1/p}}{(c^q + 1)^{1/q}} \geq 1.$$

$H(1) = 1$, and the result then follows by noting that dH/dc is < 0 for the in-

terval considered. For $1 < p \leq 2$ the inequality (10) holds in the reverse sense, the proof being identical. This completes the proof of Theorem 2.

COROLLARY. For $p > 1$ the spaces L_p and l_p are uniformly convex.

It is of some interest to compute the function $\delta(\epsilon)$ of Definition 1 explicitly in this case. Consider $p \geq 2$, and set $\|x\| = \|y\| = 1$ in (3);

$$\|x + y\|^p + \|x - y\|^p \leq 2^p.$$

Then if $\|x - y\| \geq \epsilon$ ($0 < \epsilon \leq 2$), we see that

$$\left\| \frac{x + y}{2} \right\| \leq \left[1 - \left(\frac{\epsilon}{2} \right)^p \right]^{1/p},$$

so that an admissible value for $\delta(\epsilon)$ is $1 - [1 - (\epsilon/2)^p]^{1/p}$. For $1 < p \leq 2$, we have $\delta(\epsilon) = 1 - [1 - (\epsilon/2)^q]^{1/q}$, by (4).

4. A "STRONG" TRIANGLE RELATION

For uniformly convex spaces the ordinary "triangular inequality," which states that if $y = \sum_{i=1}^k x_i$, then $\|y\| \leq \sum_{i=1}^k \|x_i\|$, can be replaced by a considerably stronger inequality.

For two non-vanishing vectors, x and y , of the space B , we define a generalized "angle" between the vectors, which we denote by $\alpha[x, y]$, as follows:

$$\alpha[x, y] = \left\| \frac{x}{\|x\|} - \frac{y}{\|y\|} \right\|.$$

We note the following properties of α :

- (i) $\alpha[x, y] = \alpha[y, x]$; $0 \leq \alpha[x, y] \leq 2$.
- (ii) $\alpha[x, y] \leq \alpha[x, z] + \alpha[z, y]$.
- (iii) $\alpha[x, y] = 0$ is equivalent to the condition $y = cx$ ($c > 0$).

We now state our inequality in

THEOREM 3. Let $x_1, x_2, \dots, x_k, y$ be non-vanishing elements of B , a uniformly convex Banach space, with

$$y = \sum_{i=1}^k x_i.$$

Let $\alpha_i = \alpha[x_i, y]$ ($i = 1, 2, \dots, k$). Then

$$(11) \quad \|y\| \leq \sum_{i=1}^k [1 - 2\delta(\alpha_i)] \|x_i\|,$$

where δ is the function of Definition 1, and $\delta(0) = 0$.

Since for each i we have

$$\left\| \frac{x_i}{\|x_i\|} - \frac{y}{\|y\|} \right\| = \alpha_i,$$

then by Definition 1, for each i ,

$$\left\| \frac{x_i}{\|x_i\|} + \frac{y}{\|y\|} \right\| \leq 2 - 2\delta(\alpha_i),$$

or

$$\left\| \|y\|x_i + \|x_i\|y \right\| \leq (2 - 2\delta(\alpha_i))\|x_i\|\|y\|;$$

summing these inequalities over i , we obtain

$$\sum_{i=1}^k \left\| \|y\|x_i + \|x_i\|y \right\| \leq \sum_{i=1}^k (2 - 2\delta(\alpha_i))\|x_i\|\|y\|.$$

If we replace the sum of norms on the left by the norm of the sum, this becomes

$$\left\| \|y\|y + y \sum_{i=1}^k \|x_i\| \right\| \leq \|y\| \sum_{i=1}^k (2 - 2\delta(\alpha_i))\|x_i\|,$$

or, since the left side equals $\|y\|(\|y\| + \sum_{i=1}^k \|x_i\|)$,

$$\|y\| + \sum_{i=1}^k \|x_i\| \leq 2 \sum_{i=1}^k \|x_i\| - 2 \sum_{i=1}^k \delta(\alpha_i)\|x_i\|.$$

Rearranging the terms we have the inequality of the theorem.

It will be noticed that in a uniformly convex space, if the sum of the lengths of two sides of a triangle is equal to the length of the third side, the triangle is degenerate, and conversely;* that is, if x and y are $\neq 0$, and $\|x+y\| = \|x\| + \|y\|$, then $x = cy (c > 0)$: this statement is a consequence of (11) and (iii). We can show from this that in Definition 1 the value of $\delta(2)$ may be assumed to be 1. Indeed, suppose that $\|x\| = \|y\| = 1$, and $\|x-y\| = 2$. Then

$$\|x + (-y)\| = \|x\| + \|-y\|;$$

hence by the property just noted, $x = -y$, and the value $\delta(2) = 1$ is admissible. We can now state as a further property of the function $\alpha[x, y]$ that

(iv) $\alpha[x, y] = 2$ is equivalent to the condition $y = cx (c < 0)$.

Using the particular value of $\delta(\epsilon)$ obtained above for $L_p[l_p]$ we have the following strengthening of Minkowski's Inequality:

* We say that a space in which this condition is satisfied is "strictly convex."

COROLLARY. Let $x_1, x_2, \dots, x_k, y$ be non-vanishing elements of L_p or l_p ($p \geq 2$), $y = \sum_{i=1}^k x_i$. If $\alpha_i = \alpha[x_i, y]$, then

$$(12) \quad \|y\| \leq \sum_{i=1}^k [2(1 - (\alpha_i/2)^p)^{1/p} - 1] \|x_i\|.$$

For $1 < p \leq 2$, p must be replaced by q in this inequality.

We point out that in inequalities (11) and (12) the factors associated with the $\|x_i\|$ may assume any values between -1 and 1 inclusive. These relations are obviously reminiscent of the corresponding relation between vectors in the Euclidean plane, $\|y\| = \sum \cos a_i \|x_i\|$, $a_i =$ angle between the vectors x_i and y .

For use in a later argument we state the following corollary to Theorem 3.

COROLLARY. Let $y, x_1, x_2, \dots, x_k, w_1, \dots, w_l$ be non-vanishing elements of a uniformly convex space, with $y = \sum_{i=1}^k x_i + \sum_{i=1}^l w_i$, and with $\alpha[x_i, y] > \epsilon > 0$ ($i = 1, 2, \dots, k$). Then

$$\sum_{i=1}^k \|x_i\| \leq \frac{1}{2\delta(\epsilon)} \left(\sum_{i=1}^k \|x_i\| + \sum_{i=1}^l \|w_i\| - \|y\| \right).$$

This follows immediately from

$$\|y\| \leq \sum_{i=1}^k (1 - 2\delta(\epsilon)) \|x_i\| + \sum_{i=1}^l \|w_i\|;$$

which itself is implied by Theorem 3.

5. FUNCTIONS FROM A EUCLIDEAN TO A UNIFORMLY CONVEX SPACE

A function of bounded variation from a Euclidean space into a uniformly convex Banach space is differentiable at almost all points; we shall formulate and prove this statement in precise fashion as our next theorem. As we have noted above, Bochner has shown that this theorem is not true for Banach spaces in general. Bochner's example is a function from one-dimensional Euclidean space into M , the space of bounded functions. The following example, in which the range space is taken to be space L , the space of all summable functions $f(s)$ on the interval $0 \leq s \leq 1$, shows that this phenomenon may still occur in separable spaces.

To each point t , $0 \leq t \leq 1$, let correspond the element $\phi_t \in L$ defined as follows:

$$\phi_t = \phi_t(s) = \begin{cases} 1 & (s \leq t), \\ 0 & (s > t). \end{cases}$$

The reader will easily verify that the norm in L of the difference quotient

$$\Delta(h, t) = \frac{\phi_{t+h} - \phi_t}{h} \quad (h \neq 0)$$

is identically equal to unity, and hence that the point function ϕ_t , since it satisfies a Lipschitz condition, is of bounded variation and indeed is absolutely continuous. However, as $h \rightarrow 0$, $\Delta(h, t)$ approaches no element of L even in the weak sense,* and hence ϕ_t has no (strong) derivative at any point; that is, $\lim_{h \rightarrow 0} \Delta(h, t)$ does not exist.

In our proof of the abstract form of the Lebesgue theorem we employ the terminology and definitions of Saks.† The basic method of the proof, however, is contained in a proof of Tonelli‡ that rectifiable plane curves admit tangents almost everywhere.

We recall a few definitions.

By an *interval* of the n -dimensional Euclidean space $\mathcal{E}_n$ is meant the set of points $(x_1, x_2, \dots, x_n)$ satisfying the inequalities $a_i \leq x_i \leq b_i$, where (a_i, b_i) are n number pairs with $a_i < b_i$. In case all $b_i - a_i$ are equal, the interval is called a *cube*, and will always be denoted by I . By an *elementary figure*, R , is meant the sum of a finite number of intervals (or the null set). In order that the set of elementary figures may be closed under our manipulations, the notions of difference and intersection of two figures, denoted symbolically by $\ominus$ and $\odot$, are defined as

$$\begin{aligned} R_1 \ominus R_2 &= \overline{(R_1 - R_2)}^0, \\ R_1 \odot R_2 &= \overline{(R_1 \cdot R_2)}^0, \end{aligned}$$

where E^0 , $\bar{E}$ denote the interior and closure respectively of the set E . We denote the measure of the figure by $|R|$.

A function $F(R)$ which makes correspond to each elementary figure, contained in a fixed figure R_0 , an element in a linear space, is called *additive* if

$$F(R_1 + R_2) = F(R_1) + F(R_2)$$

whenever R_1 and R_2 are non-overlapping; that is, the set $R_1 \cdot R_2$ is of measure zero. $F(R)$ is said to be of *bounded variation* in R_0 if for any finite set of non-overlapping figures in R_0 : $R_1, R_2, \dots, R_k$, the sum $\sum_{i=1}^k \|F(R_i)\|$ is less than a fixed bound. $F(R)$ is *absolutely continuous*§ in R_0 if given $\epsilon > 0$ there corresponds a $\delta_\epsilon > 0$ such that if $|R| < \delta_\epsilon$, then $\|F(R)\| < \epsilon$.

* See Banach, loc. cit., p. 136.

† S. Saks, *Théorie de l'Intégrale*, Chapters I-III, esp. p. 47.

‡ Tonelli, *Fondamenti di Calcolo della Variazione*, pp. 48-56.

§ These last two definitions, because of the setting, differ slightly in form from those of Saks, but are equivalent to his in case the range space is the set of real numbers.

Finally, $F(R)$ is said to be *differentiable* at a point p if the limit

$$F'(p) = \lim_{|I| \rightarrow 0} \frac{F(I)}{|I|}$$

exists, where I is an arbitrary cube containing p .

THEOREM 4. *Let $F(R)$ be an additive function of elementary figures, defined for the figures within a fixed figure R_0 , and assuming values in a uniformly convex Banach space B . Let F be of bounded variation in R_0 . Then F is differentiable almost everywhere in R_0 .*

We first note that in proving the existence of $F'(p)$ we may without loss of generality consider only those functions $F(R)$ such that

$$(13) \quad \|F(R)\| \geq |R|$$

for all R . Indeed, we have to prove that

$$(14) \quad \lim_{|I| \rightarrow 0} \frac{F(I)}{|I|}$$

exists, where I is an arbitrary cube containing the point p . Let $B_1 = B \times \mathcal{E}_1$, where $\|(x, t)\| = [\|x\|^2 + t^2]^{1/2}$, and consider the function of figures $F_1(R) = (F(R), |R|)$, whose range is in B_1 . The function F_1 satisfies our conditions, since B_1 is uniformly convex by Theorem 1, and also (13); hence, assuming the theorem proved in this case, we have almost everywhere that

$$F'_1(p) = \lim_{|I| \rightarrow 0} \frac{(F(I), |I|)}{|I|} \quad (p \in I) \quad (15)$$

exists. This, however, implies the existence of (14), which justifies our preliminary remark.

We first show that the limit

$$(15) \quad \lim_{|I| \rightarrow 0} \frac{F(I)}{\|F(I)\|} \quad (p \in I)$$

exists for almost all p .

Given a point p , and $\epsilon > 0$, we define

$$A(p, \epsilon) = \sup \alpha[F(I), F(I')],$$

where I, I' are arbitrary cubes with measure less than ϵ containing p . Let $A(p) = \lim_{\epsilon \rightarrow 0} A(p, \epsilon)$. A necessary and sufficient condition that (15) exist is that $A(p) = 0$.

Let $\epsilon, \epsilon' > 0$ be fixed. Let $R_0 = \sum_{i=1}^k R_i$, R_i non-overlapping elementary figures. By assumption the sums $\sum_{i=1}^k \|F(R_i)\|$ have a least upper bound M for all choices of the R_i ; we fix this choice so that

$$(16) \quad M - \sum_{i=1}^k \|F(R_i)\| < \epsilon'.$$

Let $E = E_p \{A(p) > 2\epsilon\}$; then $m^*E = \sum_{i=1}^k m^*ER_i^0$. Each set ER_i^0 is covered (in the sense of Vitali)[†] by a set of cubes I_i^j , $\subset R_i^0$, with

$$(17) \quad \alpha[F(I_i^j), F(R_i)] > \epsilon.$$

Then by Vitali's covering theorem there is a finite set of these, $\{I_i^j\}$ ($j = 1, 2, \dots, n_i$), disjunct, with

$$(18) \quad \sum_{j=1}^{n_i} |I_i^j| > \frac{1}{2} m^*ER_i^0 = \frac{1}{2} m^*ER_i \quad (i = 1, 2, \dots, k).$$

Now by additivity

$$F(R_i) = \sum_j F(I_i^j) + F\left(R_i \ominus \sum_j I_i^j\right),$$

so that by (17) and the second corollary to Theorem 3 we infer

$$\sum_j \|F(I_i^j)\| \leq \frac{1}{2\delta(\epsilon)} \left[\sum_j \|F(I_i^j)\| + \left\| F\left(R_i \ominus \sum_j I_i^j\right) \right\| - \|F(R_i)\| \right],$$

or, adding these inequalities over i ,

$$\begin{aligned} & \sum_i \sum_j \|F(I_i^j)\| \\ & \leq \frac{1}{2\delta(\epsilon)} \left[\sum_i \sum_j \|F(I_i^j)\| + \sum_i \left\| F\left(R_i \ominus \sum_j I_i^j\right) \right\| - \sum_i \|F(R_i)\| \right]. \end{aligned}$$

The sum of the first two terms in the bracket is $\leq M$; hence by (13) and (16) we can conclude that

$$\sum_i \sum_j |I_i^j| \leq \frac{\epsilon'}{2\delta(\epsilon)},$$

which in conjunction with (18) gives us

$$m^*E = \sum_i m^*ER_i < 2 \sum_i \sum_j |I_i^j| \leq \frac{\epsilon'}{\delta(\epsilon)}.$$

[†] For a definition of a Vitali covering, and a proof of Vitali's theorem, see, for example, Saks, loc. cit., pp. 33-36.

Since E is independent of ϵ' , this implies $mE=0$, which in turn implies $A(p)=0$ almost everywhere. The existence of the limit (15) almost everywhere has now been established.

Now to arrive at the conclusion that the limit (14) exists, we again form the function $F_1(R) = (F(R), |R|)$ with range in space $B_1 = B \times \mathcal{E}_1$. Since (15) holds almost everywhere, we have the existence of

$$\lim_{|I| \rightarrow 0} \frac{F_1(I)}{\|F_1(I)\|} = \lim_{|I| \rightarrow 0} \frac{(F(I), |I|)}{(\|F(I)\|^2 + |I|^2)^{1/2}},$$

and hence that of both limits

$$\lim_{|I| \rightarrow 0} \frac{F(I)}{(\|F(I)\|^2 + |I|^2)^{1/2}}, \quad \lim_{|I| \rightarrow 0} \frac{|I|}{(\|F(I)\|^2 + |I|^2)^{1/2}}$$

except for a null set. If the second limit does not vanish except for a null set, then (14) will follow. Now where the second limit vanishes, we have

$$(19) \quad \lim_{|I| \rightarrow 0} \frac{\|F(I)\|}{|I|} = +\infty.$$

But given any $N > 0$ let E_N be the set on which

$$\limsup_{|I| \rightarrow 0} \frac{\|F(I)\|}{|I|} > N.$$

Then corresponding to every point $p \in E_N$ there is a set of cubes $I_i \ni p$ with $|I_i| \rightarrow 0$ and $\|F(I_i)\| > N|I_i|$; by another application of Vitali's theorem there is a finite set of these, non-overlapping, with

$$\sum |I_i| \geq \frac{1}{2} m^* E_N,$$

which yields the estimate

$$m^* E_N \leq 2 \sum |I_i| < \frac{2}{N} \sum \|F(I_i)\| \leq \frac{2M}{N}.$$

Thus (19) can only occur on a null set, and (14) exists almost everywhere. This completes the proof of Theorem 4.

THEOREM 5. *Under the hypotheses of Theorem 4, the derivative of $F(R)$ is integrable in the sense of Bochner.*

Let I_0 be a cube containing R_0 , and for each positive integer k let I_0 be subdivided in symmetrical fashion into 2^{nk} sub-cubes, I_j^k . Now in I_0 define the functions $F_k(p)$ as follows: if p is interior to a cube $I_j^k \subset R_0$, define $F_k(p) = F(I_j^k)/|I_j^k|$; for all other points of I_0 let $F_k(p)$ vanish. Then almost

everywhere in R_0 we have $\lim_{k \rightarrow \infty} F_k(p) = F'(p)$, and, since the F_k are "finite valued functions," $F'(p)$ is measurable as a consequence. Now to show that it is integrable we need only demonstrate* that the real function $\|F'(p)\|$, which is also measurable, is integrable. Since $\|F_k(p)\| \rightarrow \|F'(p)\|$ almost everywhere, this will follow from the Fatou lemma† if the integrals $\int_{R_0} \|F_k(p)\|$ are bounded. But for each k ,

$$\int_{R_0} \|F_k(p)\| \leq \int_{I_0} \|F_k(p)\| = \sum_{I_j^k \subset R_0} \frac{\|F(I_j^k)\|}{|I_j^k|} |I_j^k| \leq M,$$

which completes the proof.

THEOREM 6. *Let the hypotheses of Theorem 4 be satisfied, and in addition let $F(R)$ be absolutely continuous in R_0 . Then F is the integral of its derivative; for every elementary figure $R \subset R_0$,*

$$F(R) = \int_R F'(p).$$

That $F'(p)$ exists almost everywhere, and is a summable function, we have already seen; set $G(R) = \int_R F'(p)$ for each figure $R \subset R_0$. Bochner has shown‡ that $G'(p)$ exists and equals $F'(p)$ almost everywhere; moreover G is additive and absolutely continuous. Then the function $H = F - G$ is an additive absolutely continuous function of figures whose derivative vanishes almost everywhere. We assert that from this it follows that H vanishes identically.§ Indeed, let R be any elementary figure, and let Ξ be the set on which $H'(p) = 0$. Let $\epsilon > 0$ be chosen, and let $\delta(\epsilon)$ be the function whose existence is asserted in the definition of absolute continuity. The set $R \cdot \Xi$ is covered (Vitali) by a set of cubes I for each of which we have

$$(20) \quad H(I) < \delta(\epsilon) |I|.$$

Then by Vitali's theorem there is a finite set of these cubes, $\{I_i\}$, disjunct, such that

$$mR \cdot \Xi - \frac{\delta(\epsilon)}{2} < m(R \cdot \Xi \cdot \sum I_i) \leq m \sum I_i < mR \cdot \Xi + \frac{\delta(\epsilon)}{2},$$

* Since Bochner's integral is absolutely convergent.

† Cf. Schlesinger and Plessner, *Lebesguesche Integrale und Fouriersche Reihen*, Berlin, 1926, p. 91.

‡ See the first cited paper of Bochner, p. 269.

§ As the proof shows, this assertion is true if B is any Banach space. In fact, since Theorem 5 employs the assumption of uniform convexity only to assert the existence of $F'(R)$, we can say that if $F(R)$ is an additive absolutely continuous function from $\mathcal{E}_n$ into any Banach space, which is differentiable almost everywhere, then $F'(p)$ is integrable (Bochner), and F is its indefinite integral.

and as Ξ includes almost all points of R_0 , we may rewrite this suppressing the factor Ξ , as

$$(21) \quad mR - \frac{\delta(\epsilon)}{2} < mR \odot \sum I_i \leq m \sum I_i < mR + \frac{\delta(\epsilon)}{2}.$$

Now

$$(\sum I_i \ominus [R \odot \sum I_i]) + R = \sum I_i + (R \ominus [R \odot \sum I_i]),$$

so that by additivity and the triangle relation one has

$$(22) \quad \|H(R)\| \leq \sum \|H(I_i)\| + \|H(R \ominus [R \odot \sum I_i])\| \\ + \|H(\sum I_i \ominus [R \odot \sum I_i])\|.$$

From (21) we deduce the inequalities

$$(23) \quad \begin{cases} m(R \ominus [R \odot \sum I_i]) = mR - mR \odot \sum I_i < \delta(\epsilon), \\ m(\sum I_i \ominus [R \odot \sum I_i]) = m \sum I_i - mR \odot \sum I_i < \delta(\epsilon); \end{cases}$$

combining the relations (20), (21), (22), and (23), and remembering that $mR' < \delta(\epsilon)$ implies $\|H(R')\| < \epsilon$, we infer that

$$\|H(R)\| < \delta(\epsilon) \left[mR + \frac{\delta(\epsilon)}{2} \right] + 2\epsilon.$$

As we may assume that $\delta(\epsilon) \leq \epsilon$, this is arbitrarily small with ϵ , $\|H(R)\|$ must vanish, and the proof is complete.

COROLLARY. *An additive function of elementary figures, with values in a uniformly convex space, which is absolutely continuous and of bounded variation, can be extended to all measurable sets while preserving these properties, and in only one way.*

By a standard argument we can now prove, for point functions in the one-dimensional case,

THEOREM 7. *A function $\phi(t)$ from an interval $a \leq t \leq b$ into B , a uniformly convex space, which is of bounded variation, is (strongly) differentiable almost everywhere. The derivative, $\phi'(t)$, is integrable (Bochner). If in addition $\phi(t)$ is absolutely continuous, then**

$$\phi(t) = \phi(a) + \int_a^t \phi'(t) dt.$$

We omit the proof, which presents no difficulty.

* Compare G. Birkhoff, loc. cit., Theorem 26, which is the special case of Theorem 7 in which B is Hilbert space. The last assertion of this theorem appears to be in error.

The notion of a "continuous rectifiable curve" in a uniformly convex space may be defined in a manner quite analogous to the procedure for ordinary plane curves, and by using the "angle" α defined above, the notion of a "tangent" to such a curve may be made precise. We omit the details, but state without proof the following theorem, the demonstration of which may be made to depend directly on the preceding theory:

THEOREM 8. *A continuous rectifiable curve in a uniformly convex space possesses a tangent at almost all of its points.*

6. A GENERALIZATION. CONCLUSION

It has been called to the attention of the author by J. von Neumann that the "differentiability properties" of the space expressed in the last five theorems remain true under a somewhat weaker assumption. It suffices to assume, not that the space B is uniformly convex, but merely that there can be defined a norm in the space, equivalent to the given norm (i.e., such that the notion of limit is the same with both), with respect to which the space is uniformly convex. The only point that is not immediately obvious in the verification of this statement is that the definitions of bounded variation and absolute continuity are equivalent with the different norms, and this is clear if we notice that when $\|x\|_1, \|x\|_2$ are two equivalent norms, the ratio $\|x\|_1 / \|x\|_2$ is bounded for $\|x\|_2 \neq 0$.

It follows from this that these properties are possessed by all finite dimensional Euclidean spaces with an arbitrary "Minkowski norm"; that is, a norm determined by any convex region centrally symmetric about the origin; since any such metric is clearly equivalent to the Euclidean metric. For example, all spaces $l_k^{(*)}$ with a finite number, k , of coordinates, possess these properties.

It may be proved directly without a great deal of difficulty that the space l_1 also possesses these properties. The proof rests essentially on the denumerable coordinate system in space l_1 . If a function $\phi(t)$, say from a real segment into (real) l_1 , is of bounded variation, then each of its components $x_i(t)$ is a real function of bounded variation, and the sum of their variations is finite. Then $\phi(t)$ may be decomposed into the difference of two "monotone" functions by decomposing each of the $x_i(t)$; by means of a theorem of Fubini, that a convergent series of monotonically increasing functions is term-wise differentiable, the result follows.

A question which will naturally arise in the reader's mind is whether "differentiability theorems" of the above nature can be proved under the still weaker assumption that the space B in question is strictly convex,* or

* See the definition above.

that a norm equivalent to the given norm exists with respect to which B is strictly convex. The answer to this question is negative: indeed, the essential role played by the uniformity assumption is clearly exhibited by the following theorem.

THEOREM 9. *Any separable Banach space may be given a new norm, equivalent to the original norm, with respect to which the space is strictly convex.*

We first demonstrate that the theorem is true of space C (the space of continuous functions $f(t)$ in the interval $0 \leq t \leq 1$, where $\|f\| = \max |f(t)|$). Let $\{t_n\}$ be a dense sequence of points in $(0, 1)$; then as F. J. Murray has remarked, the sequence of bounded linear functionals $F_n(f) = f(t_n)$ forms a "total" set in the space $\bar{C}$ conjugate to C ; that is, the only element of space C for which all of these functionals vanish is the zero element. Now let the space C be renormed with the following definition:

$$\|f\|_1 = \left[\|f\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(f)|^2 \right]^{1/2};$$

it is readily verified that $\|f\|_1$ is a norm equivalent to $\|f\|$. But suppose that for two non-vanishing elements f and g we have $\|f+g\|_1 = \|f\|_1 + \|g\|_1$. This implies

$$\begin{aligned} & \left[(\|f\| + \|g\|)^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(f) + F_n(g)|^2 \right]^{1/2} \\ &= \left[\|f\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(f)|^2 \right]^{1/2} + \left[\|g\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(g)|^2 \right]^{1/2}, \end{aligned}$$

and hence there exists a constant $c > 0$ with $cF_n(f) = F_n(g)$ ($n = 1, 2, \dots$). As the sequence $\{F_n\}$ is total, this implies that $cf = g$; i.e., the renormed space is strictly convex.

Now let B be any separable Banach space. Banach has shown* that there is a closed linear manifold of space C which is in one-to-one linear isometric correspondence with B . Then the above defined renorming of C furnishes a means of renorming the space B in the manner required, which finishes the demonstration of the theorem.

In the opposite direction we have

THEOREM 10. *None of the following spaces can be renormed so as to be uniformly convex:† space L ; space M (bounded functions); space C (continuous functions); space m (bounded sequences); space c (convergent sequences).*

* Banach, loc. cit., p. 185.

† For the precise definition of these spaces see, for example, Banach, loc. cit., pp. 9–12.

It need only be verified that Theorems 4–8 do not hold in the spaces enumerated. Our above example, and the example given by Bochner, dispose of spaces L and M , and the property of space C used in the last theorem, that it is a “universal space” for all separable Banach spaces, together with our example in space L , furnishes an immediate proof for this space. The following example, of an absolutely continuous (even Lipschitzean) function, from the interval $0 \leq t \leq 1$ into space c , which is not differentiable, even in the weak sense, at any point, settles the matter for space c ; as this space forms a closed linear manifold of space m , the same example will do for the latter space.

To construct the example, we first form the sequence $f_n(t)$ of real functions on $(0, 1)$ as follows:

$$f_1(t) = \begin{cases} t & (0 \leq t \leq \frac{1}{2}), \\ 1 - t & (\frac{1}{2} \leq t \leq 1); \end{cases}$$

$$f_n(t) = \begin{cases} \frac{1}{2}f_{n-1}(2t) & (0 \leq t \leq \frac{1}{2}), \\ f_n(t - \frac{1}{2}) & (\frac{1}{2} \leq t \leq 1). \end{cases}$$

The function $\phi(t)$ from $(0, 1)$ into space c is now defined by setting $\phi(t) = \{f_1(t), f_2(t), \dots\}$. We leave to the reader the simple verification that $\phi(t)$ possesses the required properties.

The conjugate space of c is the separable space l_1 . We see, then, that this property, separability of the conjugate space, is neither necessary nor sufficient for differentiability properties of the type discussed above, thus answering in the negative a question posed by G. Birkhoff* in his paper referred to.

* Loc. cit. p. 378.