

INJECTIVITY IN BANACH SPACES AND THE MAZUR-ULAM THEOREM ON ISOMETRIES

BY

JULIAN GEVIRTZ

ABSTRACT. A mapping f of an open subset U of a Banach space X into another Banach space Y is said to be (m, M) -isometric if it is a local homeomorphism for which $M \geq D^+f(x)$ and $m \leq D^-f(x)$ for all $x \in U$, where $D^+f(x)$ and $D^-f(x)$ are, respectively, the upper and lower limits of $|f(y) - f(x)|/|y - x|$ as $y \rightarrow x$. For $0 < \rho \leq 1$ we find a number $\mu(\rho) > 1$ which has the following property: Let X and Y be Banach spaces and let U be an open convex subset of X containing a ball of radius r and contained in the concentric ball of radius R . Then all (m, M) -isometric mappings of U into Y are injective if $M/m < \mu(r/R)$. We also derive similar injectivity criteria for a more general class of connected open sets U . The basic tool used is an approximate version of the Mazur-Ulam theorem on the linearity of distance preserving transformations between normed linear spaces.

Throughout, X and Y denote real Banach spaces. The theorem mentioned in the title was proved in [9] and states that a distance preserving mapping f of X onto Y for which $f(0) = 0$ is a linear transformation. A proof may also be found in [1].

If $U \subset X$ and $f: U \rightarrow Y$, then at each point x of U one defines $D^+f(x)$ and $D^-f(x)$, respectively, as the upper and lower limits of $|f(y) - f(x)|/|y - x|$ as y tends to x . Following John [4], a mapping f of an open subset U of X into Y is said to be (m, M) -isometric if it is a local homeomorphism (i.e., continuous, open and locally one-to-one) and, furthermore, $0 < m \leq D^-f(x)$ and $D^+f(x) \leq M$ for all x in U . Less precisely, f is called μ -isometric if it is (m, M) -isometric for some m, M with $M/m = \mu$, or simply *quasi-isometric* if it is μ -isometric for some μ .

We are concerned here with injectivity criteria for quasi-isometric mappings. For a given connected open subset U of X we define $\mu_0(U)$ to be the infimum of all μ for which there exists a noninjective μ -isometric mapping of U into some Banach space Y . In other words, $\mu_0(U)$ is the largest number t for which $\mu < t$ implies that all μ -isometric mappings of U are injective. John [4, 7] established that there is a universal constant C with the property that if X is a Hilbert space and $U \subset X$ is an open convex set containing a ball of radius r and contained in the concentric ball of radius R , then $\mu_0(U) \geq 1 + Cr/R$. John's proof of this result relies in essential ways on the hypothesis that the norm of X arises from an inner product. We shall derive an analogous lower bound for $\mu_0(U)$ for convex U without the assumption that X is a Hilbert space (see Theorem 2 and its corollary). More generally, we shall show how

Received by the editors June 1, 1981 and, in revised form, July 28, 1981.

1980 *Mathematics Subject Classification*. Primary 46B20.

Key words and phrases. Quasi-isometric mapping, injectivity, uniform domain.

©1982 American Mathematical Society
0002-9947/82/0000-0174/\$03.75

one may derive lower bounds for $\mu_0(U)$ for a much wider class of domains. The domains that we deal with are the uniform domains introduced by Martio and Sarvas [8] in the context of Euclidean spaces. The bounds we obtain are independent of the Banach space in which U lies and only depend on two parameters whose values give rough limitations on the shape of U (see Theorem 4 and Remark 4). The main tool used in the proofs of these injectivity criteria is an approximate version of the Mazur-Ulam theorem given in Proposition 2.

If one applies John's result to the case in which U is a ball, the lower bound one obtains for $\mu_0(U)$ is close to 1, since the constant C is a small number. A similar situation prevails when our results are specialized to balls (see Remark 2). However, there is another kind of argument that can be used to derive injectivity criteria for quasi-isometric mappings when the domain is a ball. Using such an argument John [4] showed that if $\mu < ((1 + \sqrt{5})/2)^{1/2} = 1.272\dots$, then all μ -isometric mappings of a ball in a Hilbert space are injective and in [4, 6] he showed that the same conclusion follows under the additional assumption that the image space is also a Hilbert space if $\mu < \sqrt{2} = 1.414\dots$. Here again it is possible to get by without the hypothesis that X is a Hilbert space, as is shown in [3] where it is proved that if U is a ball in a Banach space, then $\mu_0(U) \geq 1.114\dots$.

Before proceeding we fix our notation and terminology. X and Y shall always denote real Banach spaces. The r -neighborhood of a set A is denoted by $B(A, r)$ and we abbreviate $B(\{x\}, r)$ and $B(\{0\}, r)$ by $B(x, r)$ and $B(r)$, respectively. The corresponding closed balls are denoted by $\bar{B}(x, r)$ and $\bar{B}(r)$. As usual, $[x, y]$ denotes the closed segment determined by x and y . A bounded subset A of X is said to be *symmetric* with respect to a point a if x in A implies that $2a - x$ is also in A . The center a of a bounded symmetric set is, of course, unique. If A is bounded and symmetric with respect to a , we define $\text{rad } A$ to be $\inf\{r > 0: A \subset B(a, r)\}$.

We denote the Blaschke distance function on sets by D ; that is, for $A_1, A_2 \subset X$, $D(A_1, A_2) = \inf\{r > 0: A_1 \subset B(A_2, r) \text{ and } A_2 \subset B(A_1, r)\}$. We point out that if f satisfies $|f(x) - f(y)| \leq M|x - y|$ for all $x, y \in A_1 \cup A_2$, then $D(f(A_1), f(A_2)) \leq MD(A_1, A_2)$. For $A \subset X$ and $\alpha > 0$ we define $T(A, \alpha) = \{x \in X: A \subset \bar{B}(x, \alpha)\} = \cap \{\bar{B}(y, \alpha): y \in A\}$. For $x, y \in X$ we define $S(x, y, \alpha) = T([x, y], \alpha) = T(\{x, y\}, \alpha) = \bar{B}(x, \alpha) \cap \bar{B}(y, \alpha)$. We also define $C(A, \alpha) = A \cap T(A, \alpha)$. The following simple properties are used in the sequel: $S(x, y, \alpha) \neq \emptyset$ if $\alpha \geq |x - y|/2$. If A is convex, then so is $C(A, \alpha)$ and if A is symmetric with respect to a , then so is $C(A, \alpha)$.

Let $U \subset X$ and $f: U \rightarrow Y$. Then f is said to be (m, M) -rigid on U if

$$m|x - y| \leq |f(x) - f(y)| \leq M|x - y|$$

for all $x, y \in U$. Although as a global condition this is much stronger than (m, M) -isometry, a fundamental result of John [4, 5] (see Lemma 9 below) says that if f is (m, M) -isometric in $B(a, r)$, then f is (m, M) -rigid in $B(a, (m/M)r)$.

We now motivate what is to follow. The basis of the proof of the Mazur-Ulam theorem is the fact that the affine structure of a Banach space can be defined exclusively in terms of the metric. More precisely, it is possible to characterize the midpoint $a = (x + y)/2$ of $[x, y]$ without reference to the underlying algebraic

structure. This is done as follows: One defines sets S_n recursively by

$$S_1 = S(x, y, |x - y|/2) \quad \text{and} \quad S_{n+1} = C(S_n, |x - y|/2^n)$$

for $n \geq 1$. It is not difficult to show that for all n , $a \in S_n$ and that $\text{rad } S_n \rightarrow 0$. This then yields the purely metric characterization of a as the unique point belonging to all of the S_n . Using the relationship between (m, M) -isometric and (m, M) -rigid mappings mentioned in the preceding paragraph together with the indicated proof of the Mazur-Ulam theorem, John [4, 5] showed that if f is $(1, 1)$ -isometric in a suitable neighborhood of $[x, y]$, then $f((x + y)/2) = (f(x) + f(y))/2$. (He then used this to conclude that a $(1, 1)$ -isometric mapping of a connected open subset U of X coincides on U with an affine distance preserving mapping of all of X .) This leads one to believe that if f is (m, M) -isometric in a neighborhood of $[x, y]$, then $f((x + y)/2)$ will be very close to $(f(x) + f(y))/2$ for M/m sufficiently close to 1 (see Proposition 2). This approximate linearity is the basis of the proofs of our injectivity criteria. To actually put this idea into effect we need certain intuitively clear continuity properties of $S(x, y, \alpha)$ and $C(A, \alpha)$ which are stated in Lemmas 2, 3 and 4. The important point is that the moduli of continuity do not depend on the Banach space containing the sets in question. We now give the details.

LEMMA 1. *Let $a \in X$, $\rho > 0$. Let $W \subset X$ be star-shaped with respect to a . For every subset A of $\bar{B}(a, \rho)$ for which $a \in A$ and for all $\beta, \gamma > \rho$ there holds*

$$(1) \quad D(W \cap T(A, \beta), W \cap T(A, \gamma)) \leq |\beta - \gamma| \frac{\beta}{\beta - \rho}.$$

PROOF. It obviously suffices to prove the assertion for $a = 0$. Assume first that $\beta \geq \gamma$. Then if $x \in T(A, \beta)$ and $y \in A$, we have

$$\left| \frac{\gamma - \rho}{\beta - \rho} x - y \right| = \left| \frac{\gamma - \rho}{\beta - \rho} (x - y) - \frac{\beta - \gamma}{\beta - \rho} y \right| \leq \frac{\gamma - \rho}{\beta - \rho} \beta + \frac{\beta - \gamma}{\beta - \rho} \rho = \gamma,$$

which says that $((\gamma - \rho)/(\beta - \rho))T(A, \beta) \subset T(A, \gamma)$. Since $0 \in A$, we have $T(A, \beta) \subset \bar{B}(\beta)$ and since W is star-shaped with respect to a , we have

$$\begin{aligned} W \cap T(A, \gamma) &\subset W \cap T(A, \beta) \subset \frac{\gamma - \rho}{\beta - \rho} (W \cap T(A, \beta)) \\ &\quad + \frac{\beta - \gamma}{\beta - \rho} (W \cap T(A, \beta)) \\ &\subset \left(W \cap \frac{\gamma - \rho}{\beta - \rho} T(A, \beta) \right) + \frac{\beta - \gamma}{\beta - \rho} \bar{B}(\beta) \\ &\subset (W \cap T(A, \gamma)) + \bar{B}\left(\frac{\beta - \gamma}{\beta - \rho} \beta\right). \end{aligned}$$

This gives (1) in the case that $\beta \geq \gamma$. If $\gamma \geq \beta$, then the same argument gives

$$D(W \cap T(A, \beta), W \cap T(A, \gamma)) \leq |\beta - \gamma| \gamma / (\gamma - \rho) \leq |\beta - \gamma| \beta / (\beta - \rho).$$

LEMMA 2. *Let $x, y \in X$ and let $\beta, \gamma > |x - y|/2$. Then*

$$D(S(x, y, \beta), S(x, y, \gamma)) \leq \frac{|\beta - \gamma|}{1 - |x - y|/2\beta}.$$

PROOF. This follows immediately upon application of Lemma 1 with $A = [x, y]$, $W = X$ and $\rho = |x - y|/2$, since $S(x, y, \alpha) = T([x, y], \alpha)$.

The condition $B(a, \delta) \subset C(A, \beta)$ appearing in the hypotheses of the next two lemmas is easily shown to be equivalent to $B(a, \delta) \subset A \subset \bar{B}(a, \beta - \delta)$.

LEMMA 3. Let $\delta > 0$. For every bounded convex symmetric subset $A \neq \emptyset$ of X with center a and all β, γ for which $B(a, \delta) \subset C(A, \beta)$ and $\gamma > \text{rad } A$ there holds

$$D(C(A, \beta), C(A, \gamma)) \leq \left(1 + \frac{\text{rad } A}{\delta}\right) |\beta - \gamma|.$$

PROOF. We apply Lemma 1 with $\rho = \text{rad } A$ and $W = A$. Since $B(a, \delta) \subset C(A, \beta)$ implies that $\beta - \delta \geq \text{rad } A$, we conclude that

$$D(C(A, \beta), C(A, \gamma)) \leq |\beta - \gamma| \frac{\beta}{\beta - \rho} \leq |\beta - \gamma| \frac{\rho + \delta}{\delta} = \left(1 + \frac{\rho}{\delta}\right) |\beta - \gamma|.$$

LEMMA 4. Let $\delta > 0$. For every bounded convex symmetric subset $A \neq \emptyset$ of X with center a and all $F \subset X$ and β for which $C(F, \beta) \neq \emptyset$ and $B(a, \delta) \subset C(A, \beta)$ there holds

$$D(C(A, \beta), C(F, \beta)) \leq \left(1 + \frac{4 \text{rad } A}{\delta}\right) D(A, F).$$

PROOF. For any sets R and S and any numbers $\gamma, \eta > 0$ the following relations are easily verified:

$$(2) \quad T(R, \gamma) \subset T(B(R, \eta), \gamma + \eta),$$

$$(3) \quad B(S, \eta) \cap T(S, \gamma + \eta) \subset B(C(S, \gamma + 2\eta), \eta).$$

Let $D(R, S) < \eta$. By (2),

$$\begin{aligned} C(R, \gamma) &= R \cap T(R, \gamma) \subset B(S, \eta) \cap T(B(R, \eta), \gamma + \eta) \\ &\subset B(S, \eta) \cap T(S, \gamma + \eta). \end{aligned}$$

Applying (3) to this we conclude that

$$(4) \quad C(R, \gamma) \subset B(C(S, \gamma + 2\eta), \eta).$$

Let $\varepsilon = D(A, F) < \eta < \delta/2$ and $r = \text{rad } A$. Then $\beta - 2\eta > r$ and thus

$$C(A, \beta - 2\eta) \neq \emptyset.$$

Applying (4) with $R = A$, $S = F$ and $\gamma = \beta - 2\eta$ we obtain

$$(5) \quad C(A, \beta - 2\eta) \subset B(C(F, \beta), \eta).$$

Another application of (4) with $R = F$, $S = A$ and $\gamma = \beta$ yields

$$(6) \quad C(F, \beta) \subset B(C(A, \beta + 2\eta), \eta).$$

Since by Lemma 3 we have that $D(C(A, \beta), C(A, \beta \pm 2\eta)) \leq 2\eta(1 + (r/\delta))$, (5) and (6) give $D(C(A, \beta), C(F, \beta)) \leq 2\eta(1 + (r/\delta)) + \eta$. Since η is arbitrarily close to ε , we conclude that

$$D(C(A, \beta), C(F, \beta)) \leq (3 + 2(r/\delta))\varepsilon \leq (1 + 4(r/\delta))\varepsilon.$$

If, on the contrary, $\varepsilon = D(A, F) \geq \delta/2$ and $C(F, \beta) \neq \emptyset$, we also have

$$\begin{aligned} D(C(A, \beta), C(F, \beta)) &\leq D(C(A, \beta), \{a\}) + D(\{a\}, C(F, \beta)) \\ &\leq 2r + \varepsilon \leq \varepsilon \left(1 + \frac{4r}{\delta}\right). \end{aligned}$$

LEMMA 5. Let $A \subset U \subset X$ and let $f: U \rightarrow Y$ be (m, M) -rigid. Then for all $\alpha > 0$

$$f(U \cap T(A, \alpha)) \subset f(U) \cap T(f(A), M\alpha) \subset f(U \cap T(A, (M/m)\alpha)).$$

PROOF. We have

$$\begin{aligned} f(U \cap T(A, \alpha)) &= f\left(\bigcap \{U \cap \bar{B}(x, \alpha) : x \in A\}\right) \\ &\subset \bigcap \{f(U) \cap \bar{B}(f(x), M\alpha) : x \in A\} \\ &= f(U) \cap T(f(A), M\alpha). \end{aligned}$$

Since f^{-1} is $(1/M, 1/m)$ -rigid on $f(U)$ we have

$$f^{-1}(f(U) \cap T(f(A), M\alpha)) \subset U \cap T(A, (M/m)\alpha),$$

from which the second inclusion follows.

We now introduce a sequence of sets which generalizes the one used in the proof of the Mazur-Ulam theorem. Let $x, y \in X$ and let $\delta \geq |x - y|$. For $\alpha \geq 1$ we define $S_n = S_n(x, y, \delta, \alpha)$ recursively as follows: $S_1 = S(x, y, \alpha\delta/2)$ and for $n \geq 1$, $S_{n+1} = C(S_n, \alpha^{n+1}\delta/2^n)$. We have

LEMMA 6. Let $x \neq y$ be points in X , $a = (x + y)/2$ and $1 < \alpha < 2$. Then for $n \geq 1$ S_n is convex and symmetric with respect to a , $\text{rad } S_n \leq (\alpha/2)^n \delta$ and $S_n \supset B(a, (\alpha - 1)\alpha^{n-1}\delta/2^n)$.

PROOF. It is easy to verify that the assertion is true for $n = 1$. Assume inductively that it is true for a given $n \geq 1$. By definition $S_{n+1} = C(S_n, \alpha^{n+1}\delta/2^n)$. The inductive hypothesis implies that S_{n+1} is convex and symmetric with respect to a . If $u \in S_{n+1}$, then $|u - z| \leq \alpha^{n+1}\delta/2^n$ for all $z \in S_n$. In particular, for the point $z = 2a - u$ of S_n this implies that $|u - a| \leq (\alpha/2)^{n+1}\delta$, so that $\text{rad } S_{n+1} \leq (\alpha/2)^{n+1}\delta$. Finally, let

$$u \in B(a, (\alpha - 1)\alpha^n\delta/2^{n+1}) \subset B(a, (\alpha - 1)\alpha^{n-1}\delta/2^n) \subset S_n.$$

If $z \in S_n$, then $|u - z| \leq |u - a| + |a - z| \leq (\alpha - 1)\alpha^n\delta/2^{n+1} + (\alpha/2)^n\delta < \alpha^{n+1}\delta/2^n$, which means that $u \in S_{n+1}$.

We state as lemmas three basic facts proved by John.

LEMMA 7 [5, FUNDAMENTAL LEMMA]. Let A be a convex subset of X and let $f: A \rightarrow Y$ satisfy $D^+f(x) \leq M$ for all x in A . Then $|f(x) - f(y)| \leq M|x - y|$ for all x, y in A .

LEMMA 8 [5, THEOREM II]. Let $B(a, r) \subset X$ and let $f: B(a, r) \rightarrow Y$ be a local homeomorphism which satisfies $D^-f(x) \geq m$ for all $x \in B(a, r)$. Then $f(B(a, r)) \supset B(f(a), mr)$.

LEMMA 9 [5, THEOREM III]. Let $B(a, r) \subset X$ and let $f: B(a, r) \rightarrow Y$ be (m, M) -isometric. Then f is (m, M) -rigid in $B(a, (m/M)r)$.

The following two propositions together represent an approximate form of the Mazur-Ulam theorem, or to be more precise, of its proof.

PROPOSITION 1. Let $x, y \in X$ and $|x - y| = d$. Let $1 < \alpha < 2$, $0 < m \leq M$ and $\mu = M/m$. Let $f: B((x + y)/2, \rho) \rightarrow Y$ be (m, M) -isometric, where $\rho > \mu(1 + \alpha\mu)d/2$. Writing S_n and S'_n for $S_n(x, y, d, \alpha)$ and $S_n(f(x), f(y), Md, \alpha)$, respectively, there holds

$$D(f(S_n), S'_n) \leq M(\mu - 1) \frac{\alpha}{8} \left(\frac{\alpha^2 + 4\alpha - 1}{\alpha + 7} \right) \left(\frac{\alpha + 7}{\alpha - 1} \right)^n |x - y|.$$

PROOF. Lemma 9 implies that f is (m, M) -rigid in $B((x + y)/2, \rho/\mu)$. Since $\rho/\mu > (1 + \alpha\mu)d/2$, this ball contains $\bar{B}(x, \alpha\mu d/2) \cup \bar{B}(y, \alpha\mu d/2)$, and so by Lemma 8 its image contains $\bar{B}(f(x), \alpha Md/2) \cup \bar{B}(f(y), \alpha Md/2)$. Upon application of Lemma 5 with $U = B((x + y)/2, \rho/\mu)$ and $A = \{x, y\}$ we obtain

$$(7) \quad f(S_1) \subset S'_1 \subset f(S(x, y, \alpha\mu d/2)).$$

Since $S_{n+1} = S_n \cap T(S_n, \alpha^{n+1}d/2^n)$ and $S_n \subset S_1 \subset B((x + y)/2, \rho/\mu)$, Lemma 5 with $U = A = S_n$ implies that

$$(8) \quad f(S_{n+1}) \subset C(f(S_n), \alpha^{n+1}Md/2^n) \subset f(C(S_n, \alpha^{n+1}\mu d/2^n)).$$

We abbreviate $D_n = D(f(S_n), S'_n)$. Formula (7) implies

$$\begin{aligned} D_1 &\leq D(f(S(x, y, \alpha d/2)), f(S(x, y, \alpha\mu d/2))) \\ &\leq MD(S(x, y, \alpha d/2), S(x, y, \alpha\mu d/2)), \end{aligned}$$

and upon application of Lemma 2 we conclude that $D_1 \leq M\alpha^2 d(\mu - 1)/(2(\alpha - 1))$.

Now let $n \geq 1$. Writing $W = C(f(S_n), \alpha^{n+1}Md/2^n) \neq \emptyset$, we have

$$(9) \quad D_{n+1} \leq D(f(S_{n+1}), W) + D(S'_{n+1}, W).$$

Formula (8) gives

$$\begin{aligned} D(f(S_{n+1}), W) &\leq D(f(S_{n+1}), f(C(S_n, \alpha^{n+1}\mu d/2^n))) \\ &= D(f(C(S_n, \alpha^{n+1}d/2^n)), f(C(S_n, \alpha^{n+1}\mu d/2^n))) \\ &\leq MD(C(S_n, \alpha^{n+1}d/2^n), C(S_n, \alpha^{n+1}\mu d/2^n)). \end{aligned}$$

We may apply Lemma 3 with $a = (x + y)/2$, $A = S_n$, $\beta = \alpha^{n+1}d/2^n$, $\gamma = \alpha^{n+1}\mu d/2^n$ and $\delta = (\alpha - 1)\alpha^n d/2^{n+1}$, since by Lemma 6 $\text{rad } A \leq (\alpha/2)^n d < \gamma$ and $C(A, \beta) = S_{n+1} \supset B(a, \delta)$, and conclude that

$$(10) \quad D(f(S_{n+1}), W) \leq M(\mu - 1) \frac{\alpha^{n+1}d}{2^n} \left(1 + \frac{\text{rad } A}{\delta} \right) \leq K(\mu - 1),$$

where $K = Mad(\alpha + 1)/(\alpha - 1)$. Similarly, we may apply Lemma 4 with $a = (f(x) + f(y))/2$, $A = S'_n$, $F = f(S_n)$, $\beta = \alpha^{n+1}Md/2^n$ and $\delta = (\alpha - 1)\alpha^n Md/2^{n+1}$ to obtain

$$(11) \quad D(S'_{n+1}, W) = D(C(A, \beta), C(F, \beta)) \leq \left(1 + \frac{4 \text{rad } A}{\delta} \right) D_n \leq LD_n,$$

where $L = (\alpha + 7)/(\alpha - 1)$, since in this case Lemma 6 implies that $\text{rad } A \leq (\alpha/2)^n M d$ and $C(A, \beta) = S'_{n+1} \supset B(a, \delta)$. Formulas (9), (10) and (11) imply that $D_{n+1} \leq K(\mu - 1) + LD_n$. By induction we then get

$$\begin{aligned} D_n &\leq K \left(\frac{L^{n-1} - 1}{L - 1} \right) (\mu - 1) + D_1 L^{n-1} \leq \left(\frac{K}{L - 1} (\mu - 1) + D_1 \right) L^{n-1} \\ &\leq \frac{Md\alpha}{8} \left(\frac{\alpha^2 + 4\alpha - 1}{\alpha + 7} \right) \left(\frac{\alpha + 7}{\alpha - 1} \right)^n (\mu - 1), \end{aligned}$$

as desired.

PROPOSITION 2. Let $x, y \in X$, $0 < m \leq M$, $\mu = M/m < 2$ and $1 < \alpha < 2$. Let $f: B((x + y)/2, \rho) \rightarrow Y$ be (m, M) -isometric, where $\rho > \mu(1 + \alpha\mu)|x - y|/2$. Then

$$\left| \frac{f(x) + f(y)}{2} - f\left(\frac{x + y}{2}\right) \right| \leq c_1(\mu - 1)^{c_2} M |x - y|,$$

where

$$c_1 = \alpha(\alpha^2 + 4\alpha - 1)/(8(\alpha - 1)) + 1$$

and

$$c_2 = (\log(2/\alpha))(\log((\alpha + 7)/(\alpha - 1)) + \log(2/\alpha))^{-1}.$$

PROOF. By Proposition 1 and Lemma 6 we have for all $n \geq 1$

$$\begin{aligned} (12) \quad \left| \frac{f(x) + f(y)}{2} - f\left(\frac{x + y}{2}\right) \right| &\leq D(f(S_n), S'_n) + \text{rad } S'_n \\ &\leq \left(\frac{\alpha(\alpha^2 + 4\alpha - 1)}{8(\alpha + 7)} (\mu - 1)C^n + E^n \right) M |x - y|, \end{aligned}$$

where $C = (\alpha + 7)/(\alpha - 1)$ and $E = \alpha/2$. For given α and μ we use an integer n which is chosen in such a way that this last expression takes the form $c_1(\mu - 1)^{c_2} M |x - y|$, apart from negligible differences. Explicitly, we write n in the form $n = -\beta(\log(\mu - 1))/\log C + \xi$, where $0 < \beta < 1$ and $0 \leq \xi < 1$. Since $\mu < 2$, we have that $n \geq 1$. Also, $(\mu - 1)C^n = (\mu - 1)^{1-\beta} C^\xi \leq (\mu - 1)^{1-\beta} C$ and $E^n = (\mu - 1)^{-\beta(\log E)/(\log C)} E^\xi \leq (\mu - 1)^{-\beta(\log E)/(\log C)}$. If β is determined so that $1 - \beta = -\beta(\log E)/(\log C)$, then

$$1 - \beta = (\log 1/E)/(\log C + \log 1/E) = c_2$$

and the desired conclusion follows from (12).

REMARK 1. It is a simple matter to verify numerically that c_2 attains its maximum for α in $(1, 2)$ at $\alpha = \alpha_0 = 1.1572\dots$ and that for this value of α , c_1 and c_2 have values $K_1 = 5.5704\dots$ and $K_2 = 0.1216\dots$, respectively. Henceforth α_0 , K_1 and K_2 will denote these numbers.

LEMMA 10. Let $x, y \in X$ and let $f: B([x, y], \epsilon) \rightarrow Y$ be (m, M) -isometric, where $\mu = M/m < 2$. Let $N > \mu(1 + \alpha_0\mu)|x - y|/\epsilon$ be an integer. Then

$$\begin{aligned} &|f(y) - f(x) - N(f(x + (y - x)/N) - f(x))| \\ &\leq 2(N - 1)K_1(\mu - 1)^{K_2} M |y - x|. \end{aligned}$$

PROOF. Let $x_i = x + (i/N)(y - x)$ for $0 \leq i \leq N$. We have $x_i = (x_{i+1} + x_{i-1})/2$ and $|x_{i+1} - x_{i-1}| = 2|y - x|/N$. Since f is (m, M) -isometric in $B(x_i, \epsilon)$ and $\epsilon > \mu(1 + \alpha_0\mu)|x_{i+1} - x_{i-1}|/2$, we have by Proposition 2 that

$$|(f(x_{i+1}) + f(x_{i-1}))/2 - f(x_i)| \leq 2K_1(\mu - 1)^{K_2}M|y - x|/N.$$

Summing this from $i = 1$ to $i = j - 1$ we deduce that

$$|f(x_j) - f(x_{j-1}) - (f(x_1) - f(x_0))| \leq 4(j - 1)K_1(\mu - 1)^{K_2}M|y - x|/N.$$

Finally, summing from $j = 1$ to $j = N$ we obtain the desired bound.

THEOREM 1. Let $x, y \in X$ and let $f: B([x, y], \epsilon) \rightarrow Y$ be (m, M) -isometric. If $\mu = M/m < 2$, then

$$|f(y) - f(x)| \geq m(1 - 2K_1(\mu - 1)^{K_2}\mu^2(1 + \alpha_0\mu)|y - x|/\epsilon)|y - x|.$$

PROOF. Using the smallest N that satisfies the requirement of the preceding lemma, we have $N - 1 \leq \mu(1 + \alpha_0\mu)|y - x|/\epsilon$. Also, f is (m, M) -isometric in $B(x, \epsilon) \supset B(x, \mu(1 + \alpha_0\mu)|y - x|/N)$ and so by Lemma 9 f is (m, M) -rigid in $B(x, (1 + \alpha_0\mu)|y - x|/N)$ which contains x and $x + (y - x)/N$. Thus,

$$|f(x + (y - x)/N) - f(x)| \geq m|y - x|/N.$$

The desired conclusion now follows from Lemma 10.

We now use this theorem to obtain lower bounds for $\mu_0(U)$ for convex U . For short we say that $U \subset X$ is (r, R) -convex if it is open and convex and $B(a, r) \subset U \subset B(a, R)$ for some $a \in X$.

THEOREM 2. Let $U \subset X$ be an (r, R) -convex set and let $f: U \rightarrow Y$ be an (m, M) -isometric mapping. Then f is injective if $\mu = M/m$ satisfies

$$(13) \quad (\mu - 1)^{K_2}\mu^3(1 + \alpha_0\mu) < \frac{1}{16K_1} \frac{r}{R}.$$

PROOF. Let U and f be as in the statement and assume that $\mu < 2$. Obviously we may assume that $B(r) \subset U \subset B(R)$. Let $x, y \in U$ and let $d = |y - x| > 0$. Let $x' = (1 - t)x$ and $y' = (1 - t)y$, where $0 < t < 1$. We have $|y' - x'| = (1 - t)d$, $|x - x'|, |y - y'| \leq tR$. By the convexity of U we have that $(1 - t)U + B(tr) \subset U$, so that $B([x', y'], tr) \subset U$. Let G stand for the expression $2K_1(\mu - 1)^{K_2}\mu^2(1 + \alpha_0\mu)$ appearing in Theorem 1. Suppose that there is a $t \in (0, 1)$ for which $1 - G(1 - t)d/(tr) \geq \frac{1}{2}$, or equivalently,

$$(14) \quad t \geq \frac{2Gd}{r + 2Gd}.$$

We would then have by Theorem 1 and Lemma 7 that

$$\begin{aligned} |f(y) - f(x)| &\geq |f(y') - f(x')| - |f(x) - f(x')| - |f(y) - f(y')| \\ &\geq md(1 - t)/2 - 2MtR. \end{aligned}$$

Thus, $f(y) \neq f(x)$ provided that this last expression is positive; that is, provided that

$$(15) \quad t < \frac{md}{md + 4MR}.$$

Hence f will be one-to-one on U provided that for any $x, y \in U$ there is a t in $(0, 1)$ which satisfies (14) and (15). Since this is true if $\mu G < \frac{1}{8}(r/R)$, which is equivalent to (13), and (13) implies that $\mu < 2$, we are done.

REMARK 2. In the special case that U is a ball we may take $r = R$. In this case we have that $\mu_0(U) \geq t_0$, where t_0 is the solution of the equation $(t - 1)^{K_2} t^3 (1 + \alpha_0 t) = (16K_1)^{-1}$. Simple calculations show that t_0 is approximately $1 + 1.7(10)^{-19}$.

Theorem 2 immediately gives

COROLLARY. If U is (r, R) -convex, then $\mu_0(U) \geq 1 + \gamma_1(r/R)^{\gamma_2}$, where $\gamma_1 = (16K_1 t_0^3 (1 + \alpha_0 t_0))^{-1/K_2} = t_0 - 1$ and $\gamma_2 = 1/K_2 \sim 8.22$.

REMARK 3. John [4, 7] proved that there is a universal constant γ with the property that if X and Y are Hilbert spaces, then any μ -isometric mapping of an (r, R) -convex subset of X into Y is one-to-one provided that $\mu < 1 + \gamma(r/R)$. The hypothesis that X is a Hilbert space is essential to the proof given there, although the assumption that Y is also a Hilbert space can be dispensed with, basically because the existence of a μ -isometric mapping of an open subset of a Hilbert space into Y implies that Y cannot be too different from a Hilbert space. (Formally: If $f: U \rightarrow Y$ is μ -isometric, then f^{-1} is well defined and μ -isometric in some ball of Y . Since mappings of the form $g(y) = f^{-1}(ty + b)$ are also μ -isometric, we see that if U is (r, R) -convex and $f: U \rightarrow Y$ is noninjective, then for some g of the indicated form $g \circ f$ is a noninjective μ^2 -isometric mapping of U into X . Thus, if X is a Hilbert space, $\mu \geq (1 + \gamma(r/R))^{1/2} \geq 1 + \gamma'(r/R)$ for a suitable γ' .) Simple examples in the Euclidean plane show that 1 is the smallest exponent of r/R that will work in such a theorem. Our only purpose here has been to show how injectivity criteria similar to John's can be deduced for general Banach spaces and we have made no attempt to derive the smallest value of γ_2 obtainable by arguments based on the proof of the Mazur-Ulam theorem, preferring simplicity instead. Indeed, it is fairly clear that a more careful treatment would result in a value of γ_2 smaller than 8.22. It would be of interest to determine if the corollary is even true for $\gamma_2 = 1$ with a suitable value of γ_1 .

We now show how arguments of the same kind may be used to derive lower bounds for $\mu_0(U)$ for a more general class of connected open sets U . We need

LEMMA 11. Let $u, a, b \in X$ and let $f: B(u, \rho) \rightarrow Y$ be (m, M) -isometric, where $\mu = M/m < 2$ and $\rho > (1 + \mu + \alpha_0 \mu^2) \max\{|a|, |b|\}$. Then

$$|f(u + a + b) + f(u) - f(u + a) - f(u + b)| \leq 4K_1(\mu - 1)^{K_2} M(|a| + |b|).$$

PROOF. We may apply Proposition 2 with $x = u + a + b$ and $y = u$ since in this case we have

$$B((x + y)/2, \rho - |a + b|/2) \subset B(u, \rho)$$

and

$$\rho - |a + b|/2 > \mu(1 + \alpha_0 \mu) |a + b|/2 = \mu(1 + \alpha_0 \mu) |x - y|/2.$$

Thus we conclude that $|f(u + a + b) + f(u) - 2f(u + (a + b)/2)|$ is bounded above by $2K_1(\mu - 1)^{K_2}M(|a| + |b|)$. Similarly, we may apply Proposition 2 with $x = u + a$ and $y = u + b$ to conclude that $|f(u + a) + f(u + b) - 2f(u + (a + b)/2)|$ is also bounded above by $2K_1(\mu - 1)^{K_2}M(|a| + |b|)$. From these two bounds the desired result follows immediately.

If C is a curve in X parametrized by $\varphi: [\alpha, \beta] \rightarrow X$, the length of C is, as usual, the supremum of $\sum |\varphi(t_i) - \varphi(t_{i-1})|$ over all partitions $\alpha = t_0 < t_1 < \dots < t_k = \beta$ of $[\alpha, \beta]$. In what follows we use the same symbol to denote a curve and the set of points lying on it.

THEOREM 3. *Let $x, y \in X$ and let C be a curve of length L joining x to y . Let $f: B(C, \epsilon) \rightarrow Y$ be (m, M) -isometric. If $\mu = M/m < 2$, then*

$$|f(y) - f(x)| \geq m(|y - x| - 12K_1\mu^2(\mu - 1)^{K_2}(2 + \alpha_0\mu)L^2/\epsilon).$$

PROOF. Let $\varphi: [0, L] \rightarrow X$ be the arc length parametrization of C , where $\varphi(0) = x$ and $\varphi(L) = y$. Let N be the integer defined by $N - 1 \leq (1 + \mu + \alpha_0\mu^2)L/\epsilon < N$ and let $x_i = \varphi(i/N)$ for $0 \leq i \leq N$. Let $z_i = x_i - x_{i-1}$. Obviously, $|z_i| \leq L/N$. We have $(1 + \mu + \alpha_0\mu^2)L/N < \epsilon$. Since $B(x_{j-1}, \epsilon) \subset B(C, \epsilon)$ we may apply Lemma 11 with $u = x_{j-1}$, $a = z_i$ and $b = z_j$ to conclude

$$|f(x_j + z_i) - f(x_j) - (f(x_{j-1} + z_i) - f(x_{j-1}))| \leq 8K_1(\mu - 1)^{K_2}ML/N.$$

Summing this from $j = 1$ to $j = i - 1$ we obtain

$$|f(x_i) - f(x_{i-1}) - (f(x + z_i) - f(x))| \leq 8(i - 1)K_1(\mu - 1)^{K_2}ML/N$$

for $1 \leq i \leq N$, since $x_0 = x$. If we now sum this from $i = 1$ to $i = N$ we obtain

$$(16) \quad \left| f(y) - f(x) - \sum_{i=1}^N (f(x + z_i) - f(x)) \right| \leq 4(N - 1)K_1(\mu - 1)^{K_2}ML.$$

In order to apply Lemma 10 to $B([x, x + z_i], \epsilon - L/N)$ we need

$$N > \mu(1 + \alpha_0\mu)|z_i|/(\epsilon - L/N).$$

But since $\epsilon - L/N > \mu(1 + \alpha_0\mu)L/N$, we have that $\mu(1 + \alpha_0\mu)|z_i|/(\epsilon - L/N) < 1$, so that the only condition on N is that it be greater than 0. Thus, since $B([x, x + z_i], \epsilon - L/N) \subset B(C, \epsilon)$ we may apply Lemma 10 to conclude

(17)

$$|f(x + z_i) - f(x) - N(f(x + z_i/N) - f(x))| \leq 2(N - 1)K_1(\mu - 1)^{K_2}M|z_i|.$$

Now let $w_i = (z_1 + z_2 + \dots + z_i)/N$. Obviously, $|w_i| \leq iL/N^2$. We may apply Lemma 11 with $u = x$, $a = w_i$ and $b = z_{i+1}/N$ to conclude

$$\begin{aligned} |f(x + w_{i+1}) - f(x + w_i) - (f(x + z_{i+1}/N) - f(x))| \\ \leq 4K_1(\mu - 1)^{K_2}M(i + 1)L/N^2. \end{aligned}$$

Summing this from $i = 1$ to $i = N - 1$ and adding and subtracting $f(x + w_1) - f(x) = f(x + z_1/N) - f(x)$ we obtain

$$(18) \quad \left| f(x + w_N) - f(x) - \sum_{i=1}^N (f(x + z_i/N) - f(x)) \right| \leq 2K_1(\mu - 1)^{K_2} M(N - 1)(N + 2)L/N^2.$$

The bounds (16), (17) and (18) together imply

$$|f(y) - f(x) - N(f(x + w_N) - f(x))| \leq 12(N - 1)K_1(\mu - 1)^{K_2} ML.$$

But $|w_N| \leq L/N$ and $\mu L/N < \varepsilon$. Thus, by Lemma 9, $|f(x + w_N) - f(x)| \geq m|w_N| = m|y - x|/N$. Taking into account that $\mu \geq 1$ and the definition of N we arrive at the desired conclusion.

The following concept was introduced by Martio and Sarvas [8]. The formulation given here is essentially taken from Gehring and Osgood [2]. We say that an open subset U of X is an (a, b) -uniform domain if any two points x and y of U may be joined by a curve $E \subset U$ with the following properties:

- (i) E has finite length $L \leq a|x - y|$.
- (ii) If $\varphi: [0, L] \rightarrow X$ is the arc length parametrization of E , then $B(\varphi(t), b \min\{t, L - t\}) \subset U$ for all $t \in [0, L]$.

THEOREM 4. *Let $U \subset X$ be an (a, b) -uniform domain. If*

$$(19) \quad \mu < 2(1 - 108K_1\mu^2(\mu - 1)^{K_2}(2 + \alpha_0\mu)a^2/b),$$

then all μ -isometric mappings of U are injective.

PROOF. Let U be an (a, b) -uniform domain and let $f: U \rightarrow Y$ be (m, M) -isometric, where $\mu = M/m < 2$. Let $x, y \in U$ be distinct points and let E be the curve of the preceding definition. Let $\varphi: [0, L] \rightarrow X$ be the arc length parametrization of E with $\varphi(0) = x$ and $\varphi(L) = y$. Let $x' = \varphi(L/(6a))$ and $y' = \varphi(5L/(6a))$. Lemma 7 together with simple properties of arc length implies that $|f(x) - f(x')|$ and $|f(y) - f(y')|$ bounded above by $ML/(6a)$. Since $|x - x'|, |y - y'| \leq L/(6a)$ and $L \leq a|y - x|$, we have that $|y' - x'| \geq |y - x| - 2L/(6a) \geq 2L/(3a) > 0$.

Let C denote the portion of E between x' and y' ; that is $C = \varphi([L/(6a), 5L/(6a)])$. The second condition of the preceding definition implies that $B(C, bL/(6a)) \subset U$. Taking into account that the length of C is less than L which is in turn at most $3a|x' - y'|/2$, we conclude from Theorem 3 that

$$|f(y') - f(x')| \geq m|y' - x'| \left((1 - 108K_1\mu^2(\mu - 1)^{K_2}(2 + \alpha_0\mu)a^2/b) \right).$$

Since $|f(x) - f(x')| + |f(y) - f(y')| \leq ML/(3a) \leq M|y' - x'|/2$, we see that

$$|f(y) - f(x)| \geq |y' - x'| \left(m(1 - 108K_1\mu^2(\mu - 1)^{K_2}(2 + \alpha_0\mu)a^2/b) - M/2 \right).$$

Since $|y' - x'| > 0$, we have that $f(y) \neq f(x)$ provided that (19) holds and since (19) obviously implies that $\mu < 2$, we are done.

REMARK 4. Theorem 4 says that if U is (a, b) -uniform, then $\mu_0(U) \geq t_{a,b}$, where $t_{a,b} > 1$ satisfies the equation $t + 216K_1t^2(t - 1)^{K_2}(2 + \alpha_0t)a^2/b = 2$. For the case

that $X = R^n$, the Euclidean space of dimension n , Martio and Sarvas [8, Theorem 3.8] obtained a lower bound for $\mu_0(U)$ for such domains. However, their bound depends on the dimension n in addition to a and b .

Reasoning similar to that of the proof of Theorem 2 can be used to show that an (r, R) -convex set is $(2, r/(2R))$ -uniform so that Theorem 4 yields a result similar to the corollary to Theorem 2 but with a much smaller value of γ_1 .

ACKNOWLEDGMENT. The author wishes to thank the referee for his valuable criticisms of the original version of this paper. The statements and proofs of Lemmas 1, 2 and 3 are also due to him. The author is also indebted to Professor F. W. Gehring for telling him about the paper of O. Martio and J. Sarvas to which reference has been made.

REFERENCES

1. S. Banach, *Théorie des opérations linéaires*, PWN, Warsaw, 1932.
2. F. W. Gehring and B. G. Osgood, *Uniform domains and the quasi-hyperbolic metric*, J. Analyse Math. **36** (1979), 50–74.
3. J. Gevirtz, *Injectivity of quasi-isometric mappings of balls*, Proc. Amer. Math. Soc. **85** (1982), 345–349.
4. F. John, *Quasi-isometric mappings in Hilbert space*, New York Univ., Courant Inst. Math. Sci., Res. Rep. No. IMM-NYU 336, 1965.
5. ———, *On quasi-isometric mappings. I*, Comm. Pure Appl. Math. **21** (1968), 77–110.
6. ———, *On quasi-isometric mappings. II*, Comm. Pure Appl. Math. **22** (1969), 265–278.
7. ———, *Distance changes in deformations with small strain*, (Studies and essays presented to Yu-Why Chen), Math. Research Center, National Taiwan Univ., Taipei, 1970, pp. 1–15.
8. O. Martio and J. Sarvas, *Injectivity theorems in plane and space*, Ann. Acad. Sci. Fenn. Ser. AI Math. **4** (1978/1979), 383–401.
9. S. Mazur and S. Ulam, *Sur les transformations isométriques d'espaces vectoriels normés*, C. R. Acad. Sci. Paris Sér. **194** (1932), 946–948.

INSTITUTO DE MATEMÁTICA, PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE, CASILLA 114-D, SANTIAGO, CHILE