RETURN ON THE 1932 PROOF OF THE MAZUR-ULAM THEOREM

LUCAS BORBOLETA

Abstract. This work revisits the proof given in 1932 by Mazur and Ulam for their theorem.

1. Introduction

In 1932, Mazur and Ulam proved that any isometry f of a normed space E, with $f(0) = 0$, is necessarily a linear transformation [1]. This article revisits that theorem and its proof. First, the conditions of the theorem are precisely rephrased. Second, the auxiliary definitions, lemmas, and proof plan of [1] are reproduced . Third, the proof and its steps are analyzed and commented in details.

2. Definition and Theorem

The vector spaces are considered on the real field $\mathbb{R}$. Their norms are general, so they can be non-euclidean.

Definition 1 (Isometry)**.** An application $f : E \to F$, from a vector normed space E onto another vector normed space F, is an isometry if $\forall (x, y) \in E^2, \|f(x) - f(y)\| = \|x - y\|$.

Theorem 2. *An isometry $f : E \to F$, from a vector normed space E onto another vector normed space F, with $f(0) = 0$, is always a linear application.*

3. Auxiliary definitions, lemmas, and proof plan

Let us rephrase the proof of [1], focusing on accuracy, with a few adaptation of notations, but keeping its essential steps and concepts, like "metric center" and "symmetric center" of a bounded subset.

Definition 3 (Diameter)**.** For any subset A of a vector normed space G, its diameter, noted $Diam(A)$, is defined as the supremun of distance between any two points of A:

$$Diam(A) \triangleq Sup\left\{\|x - y\| : (x, y) \in A^2\right\}$$

By convention, $Diam(\emptyset) \triangleq 0$. A subset A is said "bounded" if $Diam(A) < \infty$.

Definition 4 (Metric Center)**.** For any bounded subset A of a vector normed space G, its metric center of order zero, noted $Metc_0(A)$, is a set defined as:

$$Metc_0(A) \triangleq \left\{z \in A : \forall x \in A, \|x - z\| \leq \frac{1}{2} Diam(A)\right\}$$

A metric center of order $n \in \mathbb{N}^*$, for any bounded subset $A \subset G$, noted $Metc_n(A)$, is defined recursively as:

$$Metc_n(A) \triangleq Metc_0(Metc_{n-1}(A))$$

Date: 2012-10-14 (typo fixing 2012-12-23).
Key words and phrases. Norms, Euclidean Spaces, Inner Product Spaces, Linear Isometries.

The metric center, for any bounded subset $A \subset G$, noted $Metc(A)$, is defined as the intersection of the metric centers of all orders:

$$Metc(A) \triangleq \bigcap_{n\in\mathbb{N}} Metc_n(A)$$

Lemma 5. *If $Diam(A)$ is finite then the sequence of diameters from metric centers of order n converges to zero:*

$$Diam(Metc_n(A)) \underset{n\to+\infty}{\longrightarrow} 0$$

Corollary 6. *The metric center of any bounded subset $A \subset G$ of a normed vector space is either empty or a singleton.*

Lemma 7. *If $f : E \to F$ is an isometry, and if $\{x_0\}$ is the metric center of a bounded subset $A \subset E$, then $\{f(x_0)\}$ is the metric center of the (bounded) subset $f(A) \subset F$:*

$$Metc(A) = \{x_0\} \Rightarrow Metc(f(A)) = \{f(x_0)\}$$

Definition 8 (Symmetric Center)**.** For any subset A of vector space G, its symmetric center, noted $Symc(A)$, is a set defined as:

$$Symc(A) \triangleq \{z \in A : \forall x \in A, (2z - x) \in A\}$$

Lemma 9. *For any bounded subset $A \subset G$ of a normed vector space, if its symmetric center is not empty, then it is also its metric center.*

$$\forall A \subset G, Diam(A) < \infty \text{ and } Symc(A) \neq \emptyset \Rightarrow Metc(A) = Symc(A)$$

Corollary 10. *For any bounded subset $A \subset G$ of a normed vector space, its symmetric center is either empty or a singleton.*

Definition 11 (Middle)**.** For any two distinct elements a and b of vector space G, its middle, noted $Mid(a, b)$, is a set defined as:

$$Mid(a, b) \triangleq \left\{x \in G : \|x - a\| = \|x - b\| = \frac{1}{2}\|a - b\|\right\}$$

Claim 12. The element $\frac{a+b}{2}$ is in the symmetric center of the subset $Mid(a, b)$:

$$\frac{a+b}{2} \in Symc(Mid(a, b))$$

Fact 13. *From the lemma 9, since the symmetric center $Symc(Mid(a, b))$ is not empty, it is also the metric center of $Mid(a, b)$, and by the way reduced to a singleton:*

$$Symc(Mid(a, b)) = Metc(Mid(a, b)) = \left\{\frac{a+b}{2}\right\}$$

From this point, let us consider an isometry $f : E \to F$, with $f(0) = 0$.

Fact 14. *From the lemma 7, the singleton $\left\{f\left(\frac{a+b}{2}\right)\right\}$ is also metric center of $f(Mid(a, b))$:*

$$Metc(f(Mid(a, b))) = \left\{f\left(\frac{a+b}{2}\right)\right\}$$

Claim 15. The singleton $\left\{\frac{f(a)+f(b)}{2}\right\}$ is the symmetric center of the subset $f\left(Mid\left(a,b\right)\right)$:

$$Symc\left(f\left(Mid\left(a,b\right)\right)\right)=\left\{\frac{f\left(a\right)+f\left(b\right)}{2}\right\}$$

Fact 16. *From the lemma 9,* $Metc\left(f\left(Mid\left(a,b\right)\right)\right)=Symc\left(f\left(Mid\left(a,b\right)\right)\right)$*, so* $\left\{f\left(\frac{a+b}{2}\right)\right\}=\left\{\frac{f(a)+f(b)}{2}\right\}$*, which implies:*

$$f\left(\frac{a+b}{2}\right)=\frac{f\left(a\right)+f\left(b\right)}{2}$$

Fact 17. *Considering new variables x and y, such that $a=2x$ and $b=2y$, one gets* $f\left(x+y\right)=\frac{f(2x)+f(2y)}{2}$*.*

Considering the case $y=0$, and remembering the theorem hypothesis $f\left(0\right)=0$, one gets $f\left(x\right)=\frac{f(2x)}{2}\Leftrightarrow f\left(2x\right)=2f\left(x\right)$*.*

So one concludes $f\left(x+y\right)=f\left(x\right)+f\left(y\right)$*.*

Claim 18. The transformation f , being additive and continuous (as an isometry), is linear.

4. Analysis and comments of the proof

4.1. **About definition 1.** Being an isometry, the application f, with $f\left(0\right)=0$, verifies $\forall x\in E,\left\|f\left(x\right)\right\|=\left\|x\right\|$. Conversely, if f *would* be *just linear* and such that $\forall x\in E,\left\|f\left(x\right)\right\|=\left\|x\right\|$, then it would imply $f\left(0\right)=0$ and $\forall\left(x,y\right)\in E^{2},\left\|f\left(x-y\right)\right\|=\left\|f\left(x\right)-f\left(y\right)\right\|=\left\|x-y\right\|$. So defining an isometry by the sole requirement $\forall x\in E,\left\|f\left(x\right)\right\|=\left\|x\right\|$ only works (equivalently to 1) for a linear application. So the definition 1 cannot be "simplified".

4.2. **About definition 4.** The definition of [1] is ambiguous regarding "metric center of order zero". Should one understand $Metc_{0}\left(A\right)\subset A$ or just $Metc_{0}\left(A\right)\subset G$? However, for the recursive definition of "metric center of order $n\geq 1$", [1] clearly requires $Metc_{n}\left(A\right)\subset A$. So by extension, one also requires $Metc_{0}\left(A\right)\subset A$.

4.3. **About lemma 5.** Here is a proof of this lemma. Let us consider $Metc_{0}\left(A\right)=\left\{z\in A:\forall x\in A,\left\|x-z\right\|\leq\frac{1}{2}Diam\left(A\right)\right\}$.

- If $Metc_{0}\left(A\right)$ has less than two elements then $Diam\left(Metc_{0}\left(A\right)\right)=0$.
- On the opposite, if $Metc_{0}\left(A\right)$ has at least two distinct elements z_1 and z_2, then they satisfy the property $\left\|z_1-z_2\right\|\leq\frac{1}{2}Diam\left(A\right)$, so $Diam\left(Metc_{0}\left(A\right)\right)\leq\frac{1}{2}Diam\left(A\right)$.

In both cases, $Diam\left(Metc_{0}\left(A\right)\right)\leq\frac{1}{2}Diam\left(A\right)$. By recurrence, one gets for any $n\in\mathbb{N}$: $Diam\left(Metc_{n}\left(A\right)\right)\leq\frac{1}{2^{n+1}}Diam\left(A\right)$. And one concludes $Diam\left(Metc_{n}\left(A\right)\right)\underset{n\to+\infty}{\longrightarrow}0$.

4.4. **About corollary 6.** Here is a proof of this corollary. Let us make the hypothesis that two distinct elements z_1 and z_2 are members of $Metc\left(A\right)$. This means two distinct elements z_1 and z_2 are members of $Metc_{n}\left(A\right)$ for any $n\in\mathbb{N}$. This implies $\left\|z_1-z_2\right\|\leq\frac{1}{2^{n+1}}Diam\left(A\right)$, for any $n\in\mathbb{N}$. However, z_1 and z_2 are distinct, so $\left\|z_1-z_2\right\|=\delta>0$. Combining these properties yields to $\exists\delta>0,\forall n\in\mathbb{N},\delta\leq\frac{1}{2^{n+1}}Diam\left(A\right)$. But this is incompatible with the lemma 5: $Diam\left(Metc_{n}\left(A\right)\right)\underset{n\to+\infty}{\longrightarrow}0$. This proofs that $Metc\left(A\right)$ cannot have two distinct elements z_1 and z_2.

4.5. **About lemma 7.** Here is a proof.

- First, let us remark that $\|f(x)-f(y)\| = \|x-y\|$ implies $Diam(f(A)) = Diam(A)$. By the way, if A is bounded then be also $f(A)$.
- Second, for any $z \in Metc_0(A)$, and for any $x \in A$, $\|x-z\| \leq \frac{1}{2}Diam(A)$, so equivalently $\|f(x)-f(z)\| \leq \frac{1}{2}Diam(f(A))$, and then $f(z) \in Metc_0(f(A))$. This means $Metc_0(f(A)) = f(Metc_0(A))$.
- Third, by definition, $Metc_n(A) = Metc_0(Metc_{n-1}(A))$, so by recurrence one obtains $Metc_n(f(A)) = f(Metc_n(A))$. Then by taking the intersection over the family $n \in \mathbb{N}$, one gets $Metc(f(A)) = \bigcap_{n\in\mathbb{N}} Metc_n(f(A)) = \bigcap_{n\in\mathbb{N}} f(Metc_n(A)) = f(Metc(A))$. Indeed, since f is an isometry, it is injective: $f(x_1) = f(x_2) \Leftrightarrow \|f(x_1)-f(x_2)\| = 0$, $\|f(x_1)-f(x_2)\| = \|x_1-x_2\|$, $\|x_1-x_2\| = 0 \Leftrightarrow x_1 = x_2$. So if for any $n \in \mathbb{N}$, $y \in f(Metc_n(A))$, it implies $\exists x \in E$, with $y = f(x)$, such that $\forall n \in \mathbb{N}, x \in Metc_n(A)$, so $x \in Metc(A)$.

4.6. **About lemma 9.** Here is a proof.

- Let us consider $z \in Symc(A)$ and $x \in A$. This implies $z \in A$ and $(2z-x) \in A$. Since $x \in A$ and $(2z-x) \in A$ then $\|(2z-x)-x\| \leq Diam(A)$. This inequality can be rewritten $\|z-x\| \leq \frac{1}{2}Diam(A)$. This implies $z \in Metc_0(A)$.
- If one checks that $Symc(A)$ is a singleton then necessarily $z \in Metc(A)$. One can understand that[1] makes the hypothesis that $Symc(A)$ is a singleton. Here, one obtains the same result by requiring $Diam(A) < \infty$ in the lemma.

Let us proof that if two distinct elements z_1 and z_2 are members of $Symc(A)$ then $Diam(A) = \infty$. Indeed, let us construct a sequence: $\left(z_1^{(0)}, z_2^{(0)}\right) \triangleq (z_1, z_2)$, $\forall n \in \mathbb{N}^*$, $\left(z_1^{(n)}, z_2^{(n)}\right) \triangleq \left(2z_1 - z_2^{(n-1)}, 2z_2 - z_1^{(n-1)}\right)$. By definition of $Symc(A)$, if $z_1^{(n-1)}$ and $z_2^{(n-1)}$ are in the symmetric center, so be also $z_1^{(n)}$ and $z_2^{(n)}$. So all the pairs $\left(z_1^{(n)}, z_2^{(n)}\right)$ are in $Symc(A)$. By recurrence, one checks that $\left(z_1^{(n)}, z_2^{(n)}\right) = (z_1 + n(z_1-z_2), z_2 + n(z_2-z_1))$. This implies $\left\|z_1^{(n)} - z_2^{(n)}\right\| = (2n+1)\cdot\|z_1-z_2\|$ and $\left\|z_1^{(n)} - z_2^{(n)}\right\| \underset{n\to\infty}{\longrightarrow} \infty$.

4.7. **About corollary 10.** This corollary appears as a combination of lemmas 6 and 9. However, the proof of lemma 9 already discovered it as a by product: For any bounded subset A, $Symc(A)$ is either empty or singleton.

4.8. **About definition 11.** Let us remark that $Mid(a,b)$ is a bounded subset. Indeed, let us consider two of its elements x and y. The distance $\|x-y\|$ can be bounded as follows: $\|x-y\| = \|(x-a)+(a-y)\| \leq \|x-a\| + \|y-a\| \leq 2\frac{\|a-b\|}{2}$.

4.9. **About claim 12.** Here is a proof. The symmetric center of subset $Mid(a,b)$ is defined as $Symc(Mid(a,b)) = \{z \in Mid(a,b) : \forall x \in Mid(a,b), (2z-x) \in Mid(a,b)\}$. Let us consider $z \in Mid(a,b)$ and $y \triangleq 2\left(\frac{a+b}{2}\right) - z = a+b-z$. One checks $y-a = b-z$ and $y-b = a-z$. Since $z \in Mid(a,b)$, it implies $\|z-a\| = \|z-b\| = \frac{1}{2}\|a-b\|$. So one checks $\|y-a\| = \|y-b\| = \frac{1}{2}\|a-b\|$. One concludes $z \in Mid(a,b) \Rightarrow \left(2\left(\frac{a+b}{2}\right) - z\right) \in Mid(a,b)$, and so $\frac{a+b}{2} \in Symc(Mid(a,b))$.

4.10. **About claim 15.** Here is a proof. One makes the auxiliary hypothesis that $f(Mid(a,b)) = Mid(f(a), f(b))$, then being in transposed conditions of the claim 12, one concludes that $\frac{f(a)+f(b)}{2} \in Symc(Mid(f(a), f(b))) = Symc(f(Mid(a,b)))$. So let us proof the auxiliary hypothesis.

Let us consider the condition $x \in Mid(a,b)$. It is equivalent to $\|x-a\| = \|x-b\| = \frac{1}{2}\|a-b\|$. But f being an isometry, it is also equivalent to $\|f(x)-f(a)\| = \|f(x)-f(b)\| = \frac{1}{2}\|f(a)-f(b)\|$. Which is in turn equivalent to $f(x) \in Mid(f(a), f(b))$. This means $f(Mid(a,b)) = Mid(f(a), f(b))$.

4.11. **About fact 17.** Let us remember that fact 17 is derived from considering two distinct elements a and b, or equivalently, two distinct elements x and y. Without breaking the hypothesis $x \neq y$, the fact 17 establishes, as an intermediate result, that $f(2x) = 2f(x)$. So finally, the properties $f(x+y) = f(x) + f(y)$, is proved both for $x \neq y$ and $x = y$. Let us recall that, by hypothesis of theorem 2, $f(0) = 0$.

4.12. **About linearity claim 18.** The original article [1] focused on additivity of f, that is $\forall (x,y) \in E^2, f(x+y) = f(x)+f(y)$. The linearity property is mentioned in a footnote of[1]: "this transformation, being additive and continuous (as isometry), is linear". Here are detailed arguments for checking the property $\forall x \in E, \forall \lambda \in \mathbb{R}, f(\lambda x) = \lambda f(x)$. It proceeds by successive enlargement of scalar sets.

(1) $\forall x \in E, f(-x) = -f(x)$. Proof: $f(x-x) = 0 = f(x) + f(-x)$, so $f(-x) = -f(x)$.
(2) $\forall x \in E, \forall n \in \mathbb{N}, f(nx) = nf(x)$. Proof: $f(0 \cdot x) = 0$, $f(1 \cdot x) = f(1 \cdot x)$, then the additivity of f, $f((n+1) \cdot x) = f(n \cdot x) + f(x)$, and a recurrence over n gets it.
(3) $\forall x \in E, \forall n \in \mathbb{Z}, f(nx) = nf(x)$. Proof: by combining the previous results.
(4) $\forall x \in E, \forall n \in \mathbb{N}^*, f\left(\frac{1}{n}x\right) = \frac{1}{n}f(x)$. Proof: $f\left(n\left(\frac{1}{n}x\right)\right) = f(x) = nf\left(\frac{1}{n}x\right)$, so $f\left(\frac{1}{n}x\right) = \frac{1}{n}f(x)$.
(5) $\forall x \in E, \forall r \in \mathbb{Q}, f(rx) = rf(x)$. Proof: by combining the previous results.
(6) $\forall x \in E, \forall \lambda \in \mathbb{R}, f(\lambda x) = \lambda f(x)$. Proof: since $\mathbb{Q}$ is dense in $\mathbb{R}$, it is enough to argue than the application $g : \lambda \in \mathbb{R} \longmapsto f(\lambda x) \in F$, defined for a given $x \in E$, is continuous. $\forall (x,y) \in \mathbb{R}^2, \|g(\lambda) - g(\mu)\| = \|f(\lambda x) - f(\mu x)\| = \|\lambda x - \mu x\| = |\lambda - \mu| \cdot \|x\|$, so it proofs $g(\lambda) \underset{\lambda \to \mu}{\longrightarrow} g(\mu)$.

4.13. **About the role of the theorem hypothesis** $f(0) = 0$. If one removes the hypothesis $f(0) = 0$ then one can check that $g(x) \triangleq f(x) - f(0)$ is linear. Indeed, first $g(x) - g(y) = f(x) - f(y)$ implies that the application g is an isometry. Second, since $g(0) = 0$, the application of the theorem 2 yields to the linearity of g. So removing the hypothesis $f(0) = 0$ yields to the conclusion that f could be expressed as $f(x) = f(0) + g(x)$, with g being linear. Such application f is said "affine". So the discussed hypothesis is not essential in order to obtain an interesting conclusion.

4.14. **About the role of the triangular inequality.** The triangular inequality of the norm is used once: for checking that $Mid(a,b)$ is a bounded subset. Is it essential? Or does it mean that the theorem could be extended to a domain that is larger than the vector normed spaces?

4.15. **About metric center.** Lemma 9 indicates that the metric center of a bounded subset A does not depend of the kind of norm $\|\cdot\|$ if such center is also symmetric center of A. Is it still true when $Symc(A) = \emptyset$?

5. License

References

[1] S. Mazur and Stanisław M. Ulam. Sur les transformations isometriques d'espaces vectoriels normés. *Comptes Rendus de l'Académie des Sciences*, 194:946–948, 1932.

E-mail address: `lucas.borboleta@free.fr`