On the characterization of means defined on a linear space

By ZSOLT PÁLES (Debrecen)

Preliminaires

In 1931 KOLMOGOROV [6], NAGUMO [7] and DE FINETTI [1] gave a characterization of the quasiarithmetic means. KOLMOGOROV obtained the following.

Theorem A. A discrete symmetric mean (defined on an interval $I \subseteq \mathbb{R}$) is quasiarithmetic if and only if it is a continuous and a strictly monotonous function and associative

(Throughout this section we use the terminology of [8])

Several generalizations of the quasiarithmetic means were born in the last years. (See Bajraktarevic [2], Daróczy [3], Páles [8].) One of them is the concept of quasideviation means introduced by the author in [8]. In [8] the following result is proved:

Theorem B. A discrete symmetric mean (defined on a real open interval) is a quasideviation mean if and only if it is infinitesimal and strongly intern.

In the characterization of the quasiarithmetic means and quasideviation means the concept of associativity and strongly internity play the most important roles. There is a natural way to define the discrete means, the associative means and the strongly intern means on linear spaces. (See Definitions 1, 2 and 3, respectively.) (But our concept of mean is different from the terminology of HILLE [5].) This involves the following problem: What are the associative and the strongly intern discrete symmetric means on linear spaces?

Our discussion is restricted to the case when the domain of the given discrete symmetric mean is a convex set having at least two dimensions. (If the domain has only one dimension then the given mean may be regarded as a mean defined on a real interval.)

In this article completely solve the above-mentioned problem stating its solution in Theorems 1, 2 and 3.

1. Notations and basic concepts

Throughout this paper \mathbf{R} , \mathbf{R}_+ , \mathbf{N} , X, and D denote the set of real numbers, the set of positive real numbers, the set of natural numbers, a real linear space, and a fixed convex subset of X, respectively.

If $x_1, ..., x_n \in X$, then $]x_1, ..., x_n[$ denotes the set

$$\left\{ \sum_{i=1}^n \lambda_i x_i | \lambda_i > 0, \ \sum_{i=1}^n \lambda_i = 1 \right\}.$$

If $x=(x_1, ..., x_n)\in X^n$, $y=(y_1, ..., y_m)\in X^m$ then let

$$(x, y) := (x_1, ..., x_n, y_1, ..., y_m).$$

Let further

$$\mathscr{D} := \mathscr{D}(D) := \bigcup_{n=1}^{\infty} D^n.$$

Definition 1. The function $M: \mathcal{D} \to D$ is called a symmetric discrete mean (on D) if it has the following properties:

(i) For $x = (x_1, ..., x_n) \in D^n \subset D$,

$$M(x) \in]x_1, ..., x_n[;$$

(ii) For $n \in \mathbb{N}$, the function $M_n := M|_{D^n}$ is a symmetric function of its variables.

Definition 2. The discrete symmetric mean $M: \mathcal{D} \to D$ is associative if, for $x_1 \in D^{k_1}, \ldots, x_n \in D^{k_n}$,

$$M(x_1, ..., x_n) = M(\underbrace{M(x_1), ..., M(x_1), ..., \underbrace{M(x_n), ..., M(x_n)}_{k_n})}_{k_n}$$

Definition 3. The discrete symmetric mean $M: \mathcal{D} \to D$ is strongly intern if, for $x_1, ..., x_n \in \mathcal{D}$,

$$M(x_1, ..., x_n) \in]M(x_1), ..., M(x_n)[.$$

Remarks. (1) Using property (i) it is easy to check that an associative discrete symmetric mean is strongly intern.

(2) If $X=\mathbb{R}$ and D is a real interval then the Definitions 1, 2 and 3 give the corresponding concepts on the real line (introduced in [8]).

2. Strongly intern means

In this section and in the next section we shall assume that the set D has at least three points which are not on the same line.

Theorem 1. A discrete symmetric mean $M: \mathcal{D} \to D$ is strongly intern if and only if there exists a function $f: D \to \mathbb{R}_+$ such that

(2.1)
$$M(x_1, ..., x_n) = \frac{\sum_{i=1}^n f(x_i) x_i}{\sum_{i=1}^n f(x_i)} =: M_f(x_1, ..., x_n)$$

for $x_1, ..., x_n \in D$, $n \in \mathbb{N}$.

PROOF. (i) First we prove that the mean M_f defined in (2.1) is strongly intern for each function $f: D \to \mathbb{R}_+$.

Let

$$x_1 = (x_{11}, ..., x_{1k_1}) \in D^{k_1},$$

 \vdots
 $x_n = (x_{n1}, ..., x_{nk_n}) \in D^{k_n}$

 $(k_1, ..., k_n, n \in \mathbb{N})$. Then

(2.2)
$$M_{f}(x_{1}, ..., x_{n}) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{k_{i}} f(x_{ij}) x_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{k_{i}} f(x_{ij})} = \frac{\sum_{i=1}^{n} \left(\sum_{j=1}^{k_{i}} f(x_{ij})\right) M_{f}(x_{i})}{\sum_{i=1}^{n} \sum_{j=1}^{k_{i}} f(x_{ij})} = \sum_{i=0}^{n} \lambda_{i} M_{f}(x_{i})$$

where $\lambda_i = \frac{\sum\limits_{i=1}^{k_i} f(x_{ij})}{\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{k_i} f(x_{ij})}$. It is easy to see that $\lambda_1, ..., \lambda_n > 0$ and $\sum\limits_{i=1}^{n} \lambda_i = 1$. Thus

(2.2) implies

$$M_f(x_1, ..., x_n) \in]M_f(x_1), ..., M_f(x_n)[.$$

Hence M_f is strongly intern.

(ii) To prove that every strongly intern mean M has the representation (2.1) let $x_0 \in D$ be an arbitrary but fixed point. If $x_0 \neq x \in D$ then $M(x_0, x)$ lies on the open segment $]x_0, x[$ i.e. there exists a unique value $0 < \lambda_x < 1$ such that

(2.3)
$$M(x_0, x) = \lambda_x x_0 + (1 - \lambda_x) x.$$

Let $f: D \rightarrow \mathbb{R}_+$ be defined as

(2.4)
$$f(x) := \begin{cases} 1 & \text{if } x = x_0, \\ \frac{1}{\lambda_x} - 1 & \text{if } x \neq x_0. \end{cases}$$

Then, by (2.3), we have

(2.5)
$$M(x_0, x) = \frac{f(x_0)}{f(x) + f(x_0)} x_0 + \frac{f(x)}{f(x) + f(x_0)} x = M_f(x_0, x).$$

Let, for $n \in \mathbb{N}$, S_n be the statement that

$$(2.6) M(x_1, ..., x_n) = M_f(x_1, ..., x_n)$$

for all $x_1, ..., x_n \in D$.

First we show that S_2 is valid. Let $x_1, x_2 \in D$ be arbitrary points. If $x_1 = x_2$, $x_1 = x_0$ or $x_2 = x_0$ then there is nothing to prove. Thus we may assume that x_0, x_1, x_2 are pairwise distinct points. We shall distinguish two cases:

Case I. x_0, x_1, x_2 are noncollinear points. Then $x_1, x_2, M(x_0, x_1), M(x_0, x_2)$ are pairwise distinct and noncollinear points too. Thus the segments

$$M(x_0, x_1), x_2$$
 and $M(x_0, x_2), x_1$

have at most one common point. The strong internity of M and M_f and (2.5) imply

$$M(x_0, x_1, x_2) \in]M(x_0, x_1), x_2[\cap]M(x_0, x_2), x_1[=$$

$$=]M_f(x_0, x_1), x_2[\cap]M_f(x_0, x_2), x_1[\ni M_f(x_0, x_1, x_2).$$

Therefore

$$(2.7) M(x_0, x_1, x_2) = M_f(x_0, x_1, x_2).$$

If $M(x_1, x_2) \neq M_f(x_1, x_2)$ then $]M(x_1, x_2), x_0[$ and $]M_f(x_1, x_2), x_0[$ are disjoint open segments. But, using (2.7), we have

$$]M(x_1, x_2), x_0[\in M(x_0, x_1, x_2) = M_f(x_0, x_1, x_2)\in]M_f(x_1, x_2), x_0[.$$

This contradiction shows that

$$M(x_1, x_2) = M_f(x_1, x_2)$$

in Case I.

Case II. x_0, x_1, x_2 are collinear points. Then, by our assumptions on D, there exists $x_3 \in D$ such that x_0, x_1, x_2, x_3 are noncollinear points. It is easy to check that the points

$$M(x_0, x_1), M(x_2, x_3), M(x_0, x_2), M(x_1, x_3)$$

are also noncollinear. Consequently the segments

$$M(x_0, x_1), M(x_2, x_3)$$
 and $M(x_0, x_2), M(x_1; x_3)$

have at most one common point. Using the strong internity of M and M_f , (2.5), and Case I we obtain

$$M(x_0, x_1, x_2, x_3) \in$$

$$\in]M(x_0, x_1), M(x_2, x_3)[\cap]M(x_0, x_2), M(x_1, x_3)[=$$

= $]M_f(x_0, x_1), M_f(x_2, x_3)[\cap]M_f(x_0, x_2), M_f(x_1, x_3)[\in]$

$$\in M_f(x_0, x_1, x_2, x_3).$$

Thus

$$(2.8) M(x_0, x_1, x_2, x_3) = M_f(x_0, x_1, x_2, x_3).$$

If $M(x_1, x_2) \neq M_f(x_1, x_2)$, then

$$[M(x_1, x_2), M(x_0, x_3)[$$
 and $[M_f(x_1, x_2), M_f(x_0, x_3)[$

are disjoint open segments since $M(x_0, x_3) = M_f(x_0, x_3)$.

On the other hand, (2.8) implies

$$]M(x_1, x_2), M(x_0, x_3)[\in M(x_0, x_1, x_2, x_3) =$$

= $M_f(x_0, x_1, x_2, x_3)\in]M_f(x_1, x_2), M_f(x_0, x_3)[$

This contradiction completes the proof of S_2 .

To show that S_n (n>2) is also valid we use induction. Let n>2 and assume that S_k is valid for $n>k\in\mathbb{N}$. Let $x_1,\ldots,x_n\in D$. If $x_1=\ldots=x_n$ then (2.6) is obviously satisfied. Thus, without loss of the generality, we may assume that $x_{n-1}\neq x_n$.

If $M(x_{n-1}, x_n) = M(x_1, ..., x_{n-2})$ then, by S_{n-2} and S_2 , we have

$${M(x_1, ..., x_n)} =]M(x_1, ..., x_{n-2}), M(x_{n-1}, x_n)[=]M_f(x_1, ..., x_{n-2}), M_f(x_{n-1}, x_n)[= {M_f(x_1, ..., x_n)}]$$

i.e. (2.6) is satisfied.

If $M(x_{n-1}, x_n) \neq M(x_1, ..., x_{n-2})$ then choose $x_{n+1} \in D$ such that the point systems

$$x_{n-1}, x_n, x_{n+1}$$

and

$$(2.9) M(x_1, ..., x_{n-2}), M(x_{n-1}, x_n), x_{n+1}$$

form real triangles.

Then

$$M(x_{n-1}, x_n), M(x_{n-1}, x_{n+1}), M(x_n, x_{n+1})$$

are noncollinear points. Thus the intersection

$$H :=]M(x_1, ..., x_{n-2}, x_{n-1}), M(x_n, x_{n+1})[\cap]M(x_1, ..., x_{n-2}, x_n), M(x_{n-1}, x_{n+1})[\cap]M(x_1, ..., x_{n-2}, x_{n+1}), M(x_{n-1}, x_n)[$$

consists of a single element which is, by the strong internity of M,

$$M(x_1, ..., x_{n+1}).$$

Using S_{n-1} and S_2 and the strong internity of M_f it can easily be verified that $M_f(x_1, ..., x_{n+1})$ also belongs to H. Therefore

$$M(x_1,...,x_{n+1})=M_f(x_1,...,x_{n+1}).$$

Then the intersection of the open segments

$$M(x_1, ..., x_n), x_{n+1}[$$
 and $M_f(x_1, ..., x_n), x_{n+1}[$

is nonvoid. But their endpoint x_{n+1} is common, consequently

$$M_f(x_1, ..., x_n), M(x_1, ..., x_n), x_{n+1}$$

are collinear points. Applying S_{n-2} and S_2 it is easy to check that $M(x_1, ..., x_n)$

and $M_f(x_1, ..., x_n)$ belong to the segment

$$]M(x_1,...,x_{n-2}), M(x_{n-1},x_n)[.$$

The assumption $M(x_1, ..., x_n) \neq M_f(x_1, ..., x_n)$ would imply that (2.9) is a collinear point system. Hence (2.6) is satisfied.

3. Associative means

Theorem 2. A discrete symmetric mean $M: \mathcal{D} \to D$ is associative if and only if there exists a function $f: D \to \mathbb{R}_+$ satisfying the functional equation

(3.1)
$$2f(M_f(x, y)) = f(x) + f(y), \quad x, y \in D$$
 and

 $(3.2) M = M_f$

We prove Theorem 2 in three steps.

Lemma 1. If $M: \mathcal{D} \to D$ is an associative discrete symmetric mean then there exists a function $f: D \to \mathbb{R}_+$ such that (3.1) and (3.2) are fulfilled.

PROOF. If M is associative then it is strongly intern. Applying Theorem 1 we see that there exists a function $f: D \to \mathbb{R}_+$ such that (3.2) is valid. To prove that f satisfies (3.1) let $x, y \in D$ be arbitrary. Choose $z \in D$ such that $z \neq M_f(x, y) = :\mu$. Using the associativity of $M = M_f$ we easily get

(3.3)
$$M_f(x, y, z) = M_f(\mu, \mu, z) =: v.$$

(3.3) implies the following equations:

(3.4)
$$f(x)x + f(y)y + f(z)z = (f(x) + f(y) + f(z))v$$

and

(3.5)
$$2f(\mu)\mu + f(z)z = (2f(\mu) + f(z))\nu.$$

From (3.4) and (3.5) it follows that

$$f(x)x+f(y)y-2f(\mu)\mu = (f(x)+f(y)-2f(\mu))\nu.$$

Since

$$\mu = \frac{f(x)x + f(y)y}{f(x) + fy},$$

thus we obtain

(3.6)
$$(f(x)+f(y)-2f(\mu))(\mu-\nu) = 0.$$

The strong internity of M_f and $z \neq M_f(x, y)$ imply that $\mu \neq v$. Consequently it follows from (3.6) that f satisfies (3.1).

Lemma 2. If the function $f: D \rightarrow \mathbb{R}_{+}$ satisfies the functional equation (3.1) then

(3.7)
$$nf(M_f(x_1, ..., x_n)) = f(x_1) + ... + f(x_n)$$

for all $n \in \mathbb{N}$ and $x_1, ..., x_n \in D$.

PROOF. We prove (3.7) using a Cauchy-type induction i.e. first we show that (3.7) is valid for $n=2^k$, $k \in \mathbb{N}$.

For $n=2^0$ and $n=2^1$ (3.7) is obvious. Assume that (3.7) is valid for $n=2, ..., 2^k$ and let $x_1, ..., x_{2^{k+1}} \in D$ be arbitrary points. Denote, for the sake of brevity,

$$x := (x_1, ..., x_{2^{k+1}}), \quad x' := (x_1, ..., x_{2^k}), \quad x'' := (x_{2^k+1}, ..., x_{2^{k+1}}).$$

Then

$$\begin{split} M_f(x) &= \frac{\sum\limits_{i=1}^{2^k} f(x_i) x_i + \sum\limits_{i=2^k+1}^{2^{k+1}} f(x_i) x_i}{\sum\limits_{i=1}^{2^k} f(x_i) + \sum\limits_{i=2^k+1}^{2^{k+1}} f(x_i)} = \frac{\sum\limits_{i=1}^{2^k} f(x_i) M_f(x') + \sum\limits_{i=2^k+1}^{2^{k+1}} f(x_i) M_f(x'')}{\sum\limits_{i=1}^{2^k} f(x_i) + \sum\limits_{i=2^k+1}^{2^{k+1}} f(x_i)} = \\ &= \frac{2^k f(M_f(x')) M_f(x') + 2^k f(M_f(x'')) M_f(x'')}{2^k f(M_f(x')) + 2^k f(M_f(x''))} = M_f(M_f(x'), M_f(x'')). \end{split}$$

Thus

$$\begin{split} 2^{k+1}f\big(M_f(x)\big) &= 2^{k+1}f\big(M_f(M_f(x'),M_f(x''))\big) = \\ &= 2^kf\big(M_f(x')\big) + 2^kf\big(M_f(x'')\big) = \sum_{i=1}^{2^k}f(x_i) + \sum_{i=2^k+1}^{2^{k+1}}f(x_i) = \sum_{i=1}^{2^{k+1}}f(x_i). \end{split}$$

To prove (3.7) in the general case let $n \in \mathbb{N}$ and $x_1, ..., x_n \in D$ be arbitrary. Choose $k \in \mathbb{N}$ such that $2^k \ge n$ and let

$$y := x_{n+1} := \dots := x_{2^k} := M_f(x_1, \dots, x_n).$$

Then the strong internity of M_f gives

$$M_f(x_1,...,x_n)=M_f(x_1,...,x_{n-1})=y.$$

Thus

$$nf(M_f(x_1, ..., x_n)) = 2^k f(M_f(x_1, ..., x_{2^k})) + (n-2^k)f(y) =$$

$$= \sum_{i=1}^{2^k} f(x_i) + (n-2^k)f(y) = \sum_{i=1}^n f(x_i).$$

which was to be proved.

Lemma 3. If $f: D \to \mathbb{R}_+$ satisfies the functional equation (3.7) for all $n \in \mathbb{N}$ and $x_1, ..., x_n \in D$ then M_f is an associative mean.

PROOF. Let $x_i = (x_{i1}, ..., x_{ik_i}) \in D^{k_i}$, i = 1, ..., n be arbitrary. Then

$$M_f(x_1, ..., x_n) = \frac{\sum_{i=1}^n \sum_{j=1}^{k_i} f(x_{ij}) x_{ij}}{\sum_{i=1}^n \sum_{j=1}^{k_i} f(x_{ij})} = \frac{\sum_{i=1}^n \left(\sum_{j=1}^{k_i} f(x_{ij})\right) M_f(x_i)}{\sum_{i=1}^n \sum_{j=1}^{k_i} f(x_{ij})} =$$

$$= \frac{\sum_{i=1}^{n} k_i f(M_f(x_i)) M_f(x_i)}{\sum_{i=1}^{n} k_i f(M_f(x_i))} = M_f(\underbrace{M_f(x_1), ..., M_f(x_1)}_{k_1}, ..., \underbrace{M_f(x_n), ..., M_f(x_n)}_{k_n}).$$

This completes the proof of Theorem 2.

Remarks. (1) The functional equation (3.1) was studied in the real case (when D is a real interval) in ACZÉL [1]. ACZÉL proved that every continuous solution of (3.1) has the form

$$f(x) = 1/(\alpha x + \beta), \quad x \in D \subseteq \mathbb{R}$$

where α , $\beta \in \mathbb{R}$.

(2) We have proved a bit more since we used only that M is strongly intern and $M = M_f$ satisfies (3.3) for $x, y, z \in D$.

Theorem 3. Let X be an (at least two dimensional) topological vector space and let D be a convex subset of X such that there exists an interior point of D. Let $M: \mathcal{D} \to D$ be a discrete symmetric mean. Suppose that

$$(x, y) \mapsto M(x, y), (x, y) \in D^2$$

is a continuous function. Then M is associative if and only if there exist a continuous linear functional l and a real constant c such that $M = M_f$ where, for $x \in D$,

(3.8)
$$f(x) = 1/(l(x)+c)$$

and

$$(3.9) l(x) + c > 0.$$

PROOF. By Theorem 2 it is enough to prove that

(i) the continuity of

(3.10)
$$(x, y) \mapsto M_f(x, y), (x, y) \in D^2$$

implies that f is continuous;

(ii) every continuous solution of (3.1) has the form (3.8) where l is a continuous linear functional on X and $c \in \mathbb{R}$ with (3.9).

(i) Let $x_0 \in D$ and suppose that f is not continuous at x_0 . Then there exists a sequence $\{x_n\}$ such that $x_n \to x_0$ and $f(x_n)$ converges to a limit point different from $f(x_0)$ in the interval $[0, \infty]$. Let $x_0 \neq x \in D$. Then, by the continuity of (3.10),

(3.11)
$$\lim_{n \to \infty} M(x_n, x) = M(x_0, x).$$

On the other hand, by the properties of $\{f(x_n)\}\$, (3.11) cannot be valid.

(ii) We show that the function g:=1/f satisfies the Jensen functional equation on D. This gives the representation (3.8) for f. Let $x, y \in D$. Applying Lemma 2 we have that

$$(k+l)f(M_f(\underbrace{x,\ldots,x}_k,\underbrace{y,\ldots,y}_l)) = kf(x)+lf(y)$$

i.e.

$$f\left(\frac{kf(x)x+lf(y)y}{kf(x)+lf(y)}\right) = \frac{k}{k+l}f(x) + \frac{l}{k+l}f(y)$$

for each $k, l \in \mathbb{N}$. It follows from the continuity of f that

$$f\left(\frac{\lambda f(x)x + \mu f(y)y}{\lambda f(x) + \mu f(y)}\right) = \frac{\lambda}{\lambda + \mu} f(x) + \frac{\mu}{\lambda + \mu} f(y)$$

for every λ , $\mu > 0$.

Let $\lambda = \lambda^*/f(x)$, $\mu = \mu^*/f(y)$. Then

$$f\left(\frac{\lambda^{*}x + \mu^{*}y}{\lambda^{*} + \mu^{*}}\right) = \frac{\lambda^{*}}{\frac{\lambda^{*}}{f(x)} + \frac{\mu^{*}}{f(y)}} + \frac{\mu^{*}}{\frac{\lambda^{*}}{f(x)} + \frac{\mu^{*}}{f(y)}}$$

i.e.

(3.12)
$$\frac{1}{f\left(\frac{\lambda^* x + \mu^* y}{\lambda^* + \mu^*}\right)} = \frac{\lambda^*}{\lambda^* + \mu^*} \frac{1}{f(x)} + \frac{\mu^*}{\lambda^* + \mu^*} \frac{1}{f(y)}$$

for λ^* , $\mu^* > 0$.

(3.12) means that g=1/f satisfies the Jensen functional equation on D. Thus the proof is complete.

Corollary. Let X be an (at least two dimensional) topological vector space. Let, further, $M: \mathcal{D}(X) \to X$ be an associative discrete mean. Suppose that

$$(x, y) \mapsto M(x, y), (x, y) \in D^2$$

is a continuous function Then M is the arithmetic mean on X.

PROOF. By Theorem 3, $M = M_f$ where f has the form (3.8). Since D = X, (3.9) can be satisfied if and only if $l(x) \equiv 0$. Therefore $f \equiv 1/c$ and M is the arithmetic mean on X.

References

- [1] J. Aczél, Vorlesungen über Funktionalgleichungen und ihre Anwendungen, Berlin 1961.
- [2] M. BAJRAKTAREVIC, Sur une equation fonctionelle aux valeurs moyennes, Glasnik Mat. Fiz. Astr. 13 (1958), 243—248.
- [3] Z. DARÓCZY, Über eine Klasse von Mittelwerten, Publ. Math. (Debrecen) 19 (1972), 211-217.
- [4] B. DE FINETTI, Sul concetto di media, Giornale dell' Ist. Ital. d. Attarii 2 (1931), 369-396.
- [5] E. HILLE, On a class of adjoint functional equations, Acta Sci. Math. Szeged, 34 (1973), 141-161.
- [6] A. N. Kolmogorov, Sur la notion de la moyenne, Atta Accad. dei Lincei, (6) 12 (1930), 388-391.
- [7] M. NAGUMO, Über eine Klasse der Mittelwerte, Japanese J. of Math., 7 (1930), 235-257.
- [8] Zs. Páles, On characterization of quasideviation means. Acta Math. Acad. Sci. Hungar., 40 (1982), 243—260.

(Received December 20, 1981.)