Aggregation Functions. Means

Michel GRABISCH
University of Panthéon-Sorbonne
Paris, France
Jean-Luc MARICHAL
University of Luxembourg
Luxembourg
Radko MESIAR
Slovak University of Technology
Bratislava, Slovakia,
Endre PAP
University of Novi Sad
Novi Sad, Serbia

February 5, 2010

Abstract

This two-part state-of-the-art overview on aggregation theory summarizes the essential information concerning aggregation issues. An overview of aggregation properties is given, including the basic classification on aggregation functions. In this first part, the stress is put on means, i.e., averaging aggregation functions, both with fixed arity (n-ary means) and with multiple arities (extended means).

1 Introduction

Aggregation functions became in the last decade an independent field of mathematics and information sciences. The idea of aggregation functions is rather simple - they aim to summarize the information contained in an n-tuple of input values by means of a single representative value. Starting from the prototypical example - the arithmetic mean -, many other kinds of means were applied in numerous applications in various areas. Several other kinds of aggregation functions, such as the conjunctive and the disjunctive ones, are an indispensable mathematical model not only of logical operations in the area of many-valued logics, but also of many other theoretical and applied fields.

The basic feature of all aggregation functions is their nondecreasing monotonicity, expressing the idea that "an increase of any of the input values cannot decrease the output value". Another axiomatic constraint of aggregation functions concerns the boundary conditions, expressing the idea that "minimal (maximal) inputs are aggregated into minimal (maximal) output of the scale we work on".

Thus defined, the class of aggregation functions is huge, making the problem of choosing the right function (or family) for a given application a difficult one. Besides this practical consideration, the study of the main classes of aggregation functions, their properties and their relationships, is so complex and rich that it becomes a mathematical topic of its own.

A solid mathematical analysis of aggregation functions, able to answer both mathematical and practical concerns, was the main motivation for us to prepare a monograph [38]. From the recent related monographs, recall the handbook [7] and the application-oriented book [86]. The aim of this two-part invited state-of-the-art overview is to summarize the essential information about aggregation functions for Information Sciences readers, to open them the door to the rich world of important tools for information fusion. Recall that the field of aggregation theory and application was discussed in recent Information Sciences issues in many papers, for example [27, 28, 41, 48, 66, 67, 70, 87, 92, 93, 98, 97]. With the kind permission of the publisher, some parts of [38] were used in this manuscript. Moreover, to increase the transparency, proofs of several introduced results are omitted (or sketched only), however, for interested readers, always an indication where the full proofs can be found is given.

The paper is organized as follows. In the next section, basic properties of aggregation functions and several illustrative examples are given. Section 3 is devoted to means related to the arithmetic mean and means with some special properties. In Section 4, non-additive integral-based aggregation functions are discussed, stressing a prominent role of the Choquet and Sugeno integrals. The paper ends with some concluding remarks.

2 Basic definitions and examples

2.1 Aggregation functions

Aggregation functions are special real functions with inputs from a subdomain of the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$. We will deal with interval $\mathbb{I}$ domains, independently of their type (open, closed, ...). The framework of aggregation functions we deal with is constrained by the next definition, see also [7, 15, 47, 86] for the case of closed domain $\mathbb{I}$.

Definition 1. An *aggregation function* in $\mathbb{I}^n$ is a function $\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$ that

(i) is nondecreasing (in each variable)

(ii) fulfills the boundary conditions

$$\inf_{\mathbf{x}\in\mathbb{I}^n} \mathsf{A}^{(n)}(\mathbf{x}) = \inf \mathbb{I} \quad \text{and} \quad \sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{A}^{(n)}(\mathbf{x}) = \sup \mathbb{I}. \tag{1}$$

(iii) $\mathsf{A}^{(1)}(x) = x$ for all $x \in \mathbb{I}$.

An *extended aggregation function* in $\cup_{n\in\mathbb{N}}\mathbb{I}^n$ is a mapping $\mathsf{A} : \bigcup_{n\in\mathbb{N}} \mathbb{I}^n \to \mathbb{I}$ whose restriction $\mathsf{A}^{(n)} := \mathsf{A}|_{\mathbb{I}^n}$ to $\mathbb{I}^n$ is an aggregation function in $\mathbb{I}^n$ for any $n \in \mathbb{N}$.

The integer n represents the arity of the aggregation function, that is, the number of its variables. When no confusion can arise, the aggregation functions will simply be written A instead of $\mathsf{A}^{(n)}$. Observe that if $\mathbb{I} = [a, b]$, then the boundary conditions (1) can be rewritten into $\mathsf{A}^{(n)}(\mathbf{a}) = a$ and $\mathsf{A}^{(n)}(\mathbf{b}) = b$ for each $n \in \mathbb{N}$, where $\mathbf{a} = (a, \ldots, a)$ and $\mathbf{b} = (b, \ldots, b)$.

Now we introduce some basic aggregation functions.

(i) the arithmetic mean AM, defined by

$$\mathsf{AM}(\mathbf{x}) := \frac{1}{n}\sum_{i=1}^{n} x_i\,, \tag{2}$$

is an aggregation function in any domain $\mathbb{I}^n$ (if $\mathbb{I} = \overline{\mathbb{R}}$, the convention $+\infty + (-\infty) = -\infty$ is often considered).

(ii) the product $\Pi(\mathbf{x}) = \prod_{i=1}^{n} x_i$ $\quad(\mathbb{I} \in \{|0, 1|, |0, \infty|, |1, \infty|\}$, where $|a, b|$ means any of the four kinds of intervals, with boundary points a and b, and with the convention $0 \cdot \infty = 0)$;

(iii) the geometric mean $\mathsf{GM}(\mathbf{x}) = \left(\prod_{i=1}^{n} x_i\right)^{1/n}$ $\quad(\mathbb{I} \subset [0, \infty], 0 \cdot \infty = 0)$;

(iv) the minimum function $\mathsf{Min}(\mathbf{x}) = \min\{x_1, \ldots, x_n\}$ (arbitrary $\mathbb{I}$);

(v) the maximum function $\mathsf{Max}(\mathbf{x}) = \max\{x_1, \ldots, x_n\}$ (arbitrary $\mathbb{I}$);

(vi) the sum function $\Sigma(\mathbf{x}) = \sum_{i=1}^{n} x_i$ $(\mathbb{I} \in \{|0, \infty|, |-\infty, 0|, |-\infty, \infty|\}, +\infty + (-\infty) = -\infty)$;

(vii) k-order statistics $\mathsf{OS}_k : \mathbb{I}^n \to \mathbb{I}$ is defined for any interval $\mathbb{I}$ and $k \in \{1, \ldots, n\}$ as the k-th smallest input value, i.e., $\mathsf{OS}_k(\mathbf{x}) = x_j$ so that the cardinalities $|\{i \mid x_i \leqslant x_j\}| \geqslant k$ and $|\{i \mid x_i > x_j\}| < n - k$. Note that $\mathsf{OS}_1 = \mathsf{Min}$ and $\mathsf{OS}_n = \mathsf{Max}$;

(viii) k-th projection $\mathsf{P}_k : \mathbb{I}^n \to \mathbb{I}$ is defined for any interval $\mathbb{I}$ as the k-th argument, $\mathsf{P}_k(\mathbf{x}) = x_k$.

Based on many valued logics connectives [35, 40] we have the next classification of aggregation functions.

Definition 2. Consider an (extended) aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ ($\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$). Then

(i) A is called conjunctive whenever $\mathsf{A} \leqslant \mathsf{Min}$, i.e., $\mathsf{A}(\mathbf{x}) \leqslant \mathsf{Min}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{I}^n$ $(\mathbf{x} \in \cup_{n\in\mathbb{N}}\mathbb{I}^n)$.

(ii) A is called disjunctive whenever $\mathsf{A} \geqslant \mathsf{Max}$.

(iii) A is called internal whenever $\mathsf{Min} \leqslant \mathsf{A} \leqslant \mathsf{Max}$.

(iv) A is called mixed if it is neither conjunctive nor disjunctive nor internal.

In the particular case $\mathbb{I} = [0, 1]$ (or $\mathbb{I} = [a, b] \subset \mathbb{R}$), the standard duality of aggregation functions is introduced.

Definition 3. Let $\mathsf{A} : [0, 1]^n \to [0, 1]$ be an aggregation function. Then the function $\mathsf{A}^d : [0, 1]^n \to [0, 1]$ given by

$$\mathsf{A}^d(\mathbf{x}) = 1 - \mathsf{A}(1 - x_1, \ldots, 1 - x_n) \tag{3}$$

is called the dual aggregation function of A.

Evidently, A^d given by (3) is an aggregation function on $[0,1]$. Similarly, the dual extended aggregation function A^d to A acting on $[0,1]$ can be introduced. If $\mathbb{I} = [a,b] \subset \mathbb{R}$, then (3) should be modified into

$$\mathsf{A}^d(\mathbf{x}) = a + b - \mathsf{A}(a + b - x_1, \dots, a + b - x_n).$$

It is evident that the dual aggregation function for a conjunctive (respectively disjunctive, mixed) aggregation function is a disjunctive (respectively conjunctive, mixed) aggregation function.

Note that many properties defined for n-ary functions can be naturally adapted to extended functions. For example, the extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is said to be continuous if, for any $n \in \mathbb{N}$, the corresponding n-ary aggregation function $\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$ is continuous. These adaptations are implicitly assumed throughout, for example in Sections 2.2 and 2.3. Properties defined for extended aggregation functions will only be stressed explicitly (note that these properties make an important link between aggregation functions with fixed but different arities).

2.2 Monotonicity properties

Definition 4. The aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is *strictly increasing* (in each argument) if, for any $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$, if $\mathbf{x} \leqslant \mathbf{x}'$ and $\mathbf{x} \neq \mathbf{x}'$ (in short, if $\mathbf{x} < \mathbf{x}'$) then $\mathsf{A}(\mathbf{x}) < \mathsf{A}(\mathbf{x}')$.

Thus an aggregation function is strictly increasing if it presents a positive reaction to any increase of at least one input value.

If an aggregation function $\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$ presents a positive reaction in the case when all input values strictly increase, we have the next intermediate kind of monotonicity.

Definition 5. The aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is jointly strictly increasing (unanimously increasing) if, for any $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$,

$$x_i < x_i',\ \ \forall i \in \{1, \dots, n\} \quad \Rightarrow \quad \mathsf{A}(\mathbf{x}) < \mathsf{A}(\mathbf{x}').$$

For extended aggregation function $\mathsf{A} : \bigcup_{n\in\mathbb{N}} \mathbb{I}^n \to \mathbb{I}$, both introduced kinds of monotonicity mean that each n-ary aggregation function $\mathsf{A}^{(n)}, n \in \mathbb{N}$, possesses the relevant kind of monotonicity. Any strictly increasing aggregation function is necessarily also jointly strictly increasing, but not vice-versa. For example, the product Π on $]0,1]$ is strictly increasing and thus also jointly strictly increasing. However, the product Π on $[0,1]$ is only jointly strictly increasing but not strictly increasing. The *bounded sum* $\mathsf{S}_{\mathbf{L}}(\mathbf{x}) = \mathsf{Min}(\sum_{i=1}^n x_i, 1)$ on $[0,1]^n, n > 1$, is an example of an aggregation function which is not jointly strictly monotone.

2.3 Continuity properties

As already mentioned, the continuity of an extended aggregation function A means the classical continuity of n-ary functions $\mathsf{A}^{(n)}$. The same holds for the other kinds of continuity which are therefore introduced for n-ary functions only. We recall only a few of them, more details can be found in [38], Section 2.2.2.

The continuity property can be strengthened into the well-known Lipschitz condition [51]; see Zygmund [99].

Definition 6. Let $\|\cdot\| : \mathbb{R}^n \to [0,\infty[$ be a norm. If an aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ satisfies the inequality

$$|\mathsf{A}(\mathbf{x}) - \mathsf{A}(\mathbf{y})| \leqslant c\,\|\mathbf{x} - \mathbf{y}\| \qquad (\mathbf{x}, \mathbf{y} \in \mathbb{I}^n), \tag{4}$$

for some constant $c \in\,]0,\infty[$, then we say that A satisfies the *Lipschitz condition* or is *Lipschitzian* (with respect to $\|\cdot\|$). More precisely, any aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ satisfying (4) is said to be *c-Lipschitzian*. The greatest lower bound d of constants $c > 0$ in (4) is called the *best Lipschitz constant* (which means that A is d-Lipschitzian but, for any $u \in\,]0, d[$, F is not u-Lipschitzian).

Important examples of norms are given by the Minkowski norm of order $p \in [1,\infty[$, namely

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p},$$

also called the L_p-norm, and its limiting case $\|\mathbf{x}\|_\infty := \max_i |x_i|$, which is the Chebyshev norm. Evidently, each d-Chebyshev aggregation function (d-Lipschitz for the L_∞ norm) is also d-Lipschitz (for any $p \in [1,\infty[$), but not vice-versa. $\mathsf{Min}, \mathsf{Max}, \mathsf{AM}$ are 1-Chebyshev aggregation functions. The product Π on $[0,1]$ is 1-Lipschitz with respect to L_1-norm but n-Chebyshev (n is the best possible constant in this case). The geometric mean GM on $[0,1]$ is not Lipschitz (for any norm on $\mathbb{R}^n$).

We have also the next two weaker forms of continuity (using the lattice notation $\vee$ for supremum and $\wedge$ for infimum.

Definition 7. An aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is called *lower semi-continuous* or *left-continuous* if, for all $(\mathbf{x}^{(k)})_{k\in\mathbb{N}} \subset (\mathbb{I}^n)^{\mathbb{N}}$ such that $\vee_k \mathbf{x}^{(k)} \in \mathbb{I}^n$,

$$\bigvee_{k\in\mathbb{N}} \mathsf{A}(\mathbf{x}^{(k)}) = \mathsf{A}\Big(\bigvee_{k\in\mathbb{N}} \mathbf{x}^{(k)}\Big).$$

Definition 8. An aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is called *upper semi-continuous* or *right-continuous* if, for all $(\mathbf{x}^{(k)})_{k\in\mathbb{N}} \subset (\mathbb{I}^n)^{\mathbb{N}}$ such that $\wedge_k \mathbf{x}^{(k)} \in \mathbb{I}^n$,

$$\bigwedge_{k\in\mathbb{N}} \mathsf{A}(\mathbf{x}^{(k)}) = \mathsf{A}\Big(\bigwedge_{k\in\mathbb{N}} \mathbf{x}^{(k)}\Big).$$

Recall that an aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{R}$ is lower semi-continuous (respectively, upper semi-continuous) if and only if A is lower semi-continuous (respectively, upper semi-continuous) in each variable (for more details see, e.g., Klement et al. [45, Proposition 1.22]). Moreover, an aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{R}$ is continuous if and only if it is both lower and upper semi-continuous.

Example 1. (i) An important example of a left-continuous (lower semi-continuous) but noncontinuous aggregation function is the nilpotent minimum $\mathsf{T}^{\mathbf{nM}} : [0,1]^2 \to [0,1]$,

$$\mathsf{T}^{\mathbf{nM}}(x_1, x_2) := \begin{cases} \mathsf{Min}(x_1, x_2) & \text{if } x_1 + x_2 > 1 \\ 0 & \text{otherwise.} \end{cases}$$

(ii) The drastic product $\mathsf{T_D} : [0,1]^n \to [0,1]$, given by

$$\mathsf{T_D}(\mathbf{x}) := \begin{cases} \mathsf{Min}(\mathbf{x}) & \text{if } |\{i \in \{1,\ldots,n\} \mid x_i < 1\}| < 2 \\ 0 & \text{otherwise,} \end{cases}$$

is a noncontinuous but upper semi-continuous aggregation function.

2.4 Symmetry

The next property we consider is *symmetry*, also called *commutativity*, *neutrality*, or *anonymity*. The standard commutativity of binary operations $x*y = y*x$, well known in algebra, can be easily generalized to n-ary functions, with $n \geqslant 2$, as follows.

Definition 9. $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is a *symmetric* aggregation function if

$$\mathsf{A}(\mathbf{x}) = \mathsf{A}(\sigma(\mathbf{x}))$$

for any $\mathbf{x} \in \mathbb{I}^n$ and for any permutation σ of $(1,\ldots,n)$, where $\sigma(\mathbf{x}) = (x_{\sigma(1)},\ldots,x_{\sigma(n)})$.

The symmetry property essentially means that the aggregated value does not depend on the order of the arguments. This is required when combining criteria of equal importance or anonymous experts' opinions.

Up to the projections P_k, all aggregation functions introduced so far are symmetric. For example, Σ, AM, GM, S_L, T_L, Π, Min, Max are symmetric functions. A prominent example of non-symmetric aggregation functions is the weighted arithmetic mean $\mathsf{WAM}_{\mathbf{w}}(x_1,\ldots,x_n) = \sum_{i=1}^n w_i x_i$, where the nonnegative weights w_i are constrained by $\sum_{i=1}^n w_i = 1$ (and at least one weight $w_i \neq \frac{1}{n}$, otherwise $\mathsf{WAM}_{\mathbf{w}} = \mathsf{AM}$ is symmetric).

The following result, well-known in group theory, shows that the symmetry property can be checked with only two equalities; see for instance Rotman [78, Exercise 2.9 p. 24].

Proposition 1. $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is a symmetric aggregation function if and only if, for all $\mathbf{x} \in \mathbb{I}^n$, we have

(i) $\mathsf{A}(x_2, x_1, x_3, \ldots, x_n) = \mathsf{A}(x_1, x_2, x_3, \ldots, x_n)$,

(ii) $\mathsf{A}(x_2, x_3, \ldots, x_n, x_1) = \mathsf{A}(x_1, x_2, x_3, \ldots, x_n)$.

This simple test is very efficient, especially when symmetry does not appear immediately, like in the 4-variable expression

$$(x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_4) \vee (x_1 \wedge x_3 \wedge x_4) \vee (x_2 \wedge x_3 \wedge x_4),$$

which is nothing other than the 4-ary order statistic $x_{(2)}$.

2.5 Idempotency

Definition 10. $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is an *idempotent* aggregation function if $\delta_{\mathsf{A}} = \mathrm{id}$ (see Definition 12), that is,

$$\delta_{\mathsf{A}}(x) = \mathsf{A}(n \cdot x) = x \qquad (x \in \mathbb{I}),$$

where $n \cdot x = \underbrace{x,\ldots,x}_{n}$.

Idempotency is in some areas supposed to be a natural property of aggregation functions, e.g., in multicriteria decision making (see for instance Fodor and Roubens [31]), where it is commonly accepted that if all criteria are satisfied at the same degree x, implicitly assuming the commensurateness of criteria, then also the overall score should be x.

It is evident that $\mathsf{OS}_k, \mathsf{P}_k,$, AM, GM, $\mathsf{WAM}_{\mathbf{w}}$, Min, Max, and Med are idempotent aggregation functions, where the *median* of an odd number of values $(x_1, \ldots, x_{2k-1})$ is defined by

$$\mathsf{Med}(x_1, \ldots, x_{2k-1}) := x_{(k)},$$

and for an even number of values $(x_1, \ldots, x_{2k})$, the median is defined by

$$\mathsf{Med}(x_1, \ldots, x_{2k}) := \mathsf{AM}(x_{(k)}, x_{(k+1)}) = \frac{x_{(k)} + x_{(k+1)}}{2}.$$

Σ and Π are not idempotent aggregation functions. Recall also that any nondecreasing and idempotent function $\mathsf{F} : \mathbb{I}^n \to \overline{\mathbb{R}}$ is an aggregation function.

2.5.1 Idempotent elements

Definition 11. An element $x \in \mathbb{I}$ is *idempotent* for $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ if $\delta_{\mathsf{A}}(x) = x$.

In $[0,1]^n$ the product Π has no idempotent elements other than the extreme elements 0 and 1. As an example of an aggregation function in $[0,1]^n$ which is not idempotent but has a non-extreme idempotent element, take an arbitrarily chosen element $c \in]0,1[$ and define the aggregation function $\mathsf{A}_{\{c\}} : [0,1]^n \to [0,1]$ as follows:

$$\mathsf{A}_{\{c\}}(\mathbf{x}) := \mathsf{Med}\Big(0, c + \sum_{i=1}^{n}(x_i - c), 1\Big), \tag{5}$$

where Med is the standard median function (i.e., in ternary case the "middle" input, between the smallest one and the greatest one). It is easy to see that the only idempotent elements for $\mathsf{A}_{\{c\}}$ are 0, 1, and c.

Under some constraints, also non-idempotent aggregation functions can be transformed into idempotent one.

Definition 12. Let $\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$ be an aggregation function such that the diagonal section $\delta_{\mathsf{A}^{(n)}} : \mathbb{I} \to \mathbb{I}$, $\delta_{\mathsf{A}^{(n)}}(\mathbf{x}) = \mathsf{A}^{(n)}(x, \ldots, x)$, is strictly increasing, and

$$\mathrm{ran}(\delta_{\mathsf{A}^{(n)}}) = \{\delta_{\mathsf{A}^{(n)}}(x) \mid x \in \mathbb{I}\} = \mathrm{ran}(\mathsf{A}^{(n)}) = \{\mathsf{A}^{(n)}(\mathbf{x}) \mid x \in \mathbb{I}^n\},$$

is called idempotizable, and the idempotent aggregation function $AI^{(n)} : \mathbb{I}^n \to \mathbb{I}$ given by $AI^{(n)}(x) = \delta^{-1}_{\mathsf{A}^{(n)}}(\mathsf{A}^{(n)}(x))$ is called idempotized $\mathsf{A}^{(n)}$.

Note that this idempotization process preserves some properties of the original aggregation function $\mathsf{A}^{(n)}$, such as the symmetry and continuity. A typical example of an idempotized aggregation function is the arithmetic mean AM (geometric mean GM) which is the idempotized sum Σ (idempotized product Π).

2.5.2 Strong idempotency

The idempotency property has been generalized to extended functions as follows; see Calvo et al. [16].

Definition 13. $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is *strongly idempotent* if, for any $m \in \mathbb{N}$,

$$\mathsf{A}(m \cdot \mathbf{x}) = \mathsf{A}(\mathbf{x}) \qquad (\mathbf{x} \in \cup_{n\in\mathbb{N}}\mathbb{I}^n).$$

For instance, if $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is strongly idempotent then we have

$$\mathsf{A}(x_1, x_2, x_1, x_2) = \mathsf{A}(x_1, x_2).$$

Proposition 2. Suppose $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is strongly idempotent. Then A is idempotent if and only if $\mathsf{A}(x) = x$ for all $x \in \mathbb{I}$.

According to our convention on unary aggregation functions, namely $\mathsf{A}(x) = x$ for all $x \in \mathbb{I}$, it follows immediately from the previous proposition that any strongly idempotent extended aggregation function is idempotent. Typical examples of symmetric strongly idempotent extended aggregation functions are $\mathsf{AM}, \mathsf{GM}, \mathsf{Min}$ and Max. P_1 is a strongly idempotent extended aggregation function which is not symmetric. As an example of an extended aggregation function A which is idempotent, symmetric but not strongly idempotent, define an extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ by

$$\mathsf{A}(\mathbf{x}) = \begin{cases} \mathsf{Min}(\mathbf{x}) & \text{if } n \text{ is odd,} \\ \mathsf{Max}(\mathbf{x}) & \text{if } n \text{ is even.} \end{cases} \tag{6}$$

2.6 Associativity

Definition 14. $\mathsf{A} : \mathbb{I}^2 \to \mathbb{I}$ is *associative* if, for all $(x_1, x_2, x_3) \in \mathbb{I}^3$, we have

$$\mathsf{A}\big(\mathsf{A}(x_1, x_2), x_3\big) = \mathsf{A}\big(x_1, \mathsf{A}(x_2, x_3)\big). \tag{7}$$

Basically, associativity concerns binary aggregation functions. Min, Max, Π, Σ, S_L, T_L, T^{nM}, T_D are examples of binary associative aggregation functions. On the other side, AM and GM are not associative. Nevertheless, the associativity can be extended to any finite number of arguments, both for n-ary aggregation functions and for extended aggregation functions. Recall that, for an extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$, and for two vectors $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{I}^n$ and $\mathbf{x}' = (x'_1, \ldots, x'_m) \in \mathbb{I}^m$, we use the convenient notation $\mathsf{A}(\mathbf{x}, \mathbf{x}')$ to represent $\mathsf{A}(x_1, \ldots, x_n, x'_1, \ldots, x'_m)$, and similarly for more than two vectors. Also, if $\mathbf{x} \in \mathbb{I}^0$ is an empty vector then it is simply dropped from the function. For instance, $\mathsf{A}(\mathbf{x}, \mathbf{x}') = \mathsf{A}(\mathbf{x}')$ and $\mathsf{A}(\mathsf{A}(\mathbf{x}), \mathsf{A}(\mathbf{x}')) = \mathsf{A}(\mathsf{A}(\mathbf{x}'))$.

Definition 15. $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is *associative* if $\mathsf{A}(x) = x$ for all $x \in \mathbb{I}$ and if

$$\mathsf{A}(\mathbf{x}, \mathbf{x}') = \mathsf{A}(\mathsf{A}(\mathbf{x}), \mathsf{A}(\mathbf{x}'))$$

for all $\mathbf{x}, \mathbf{x}' \in \cup_{n\in\mathbb{N}_0}\mathbb{I}^n$.

As the next proposition shows, associativity means that each subset of consecutive arguments can be replaced with their partial aggregation without changing the overall aggregation.

Proposition 3. An extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is associative if and only if $\mathsf{A}(x) = x$ for all $x \in \mathbb{I}$ and

$$\mathsf{A}(\mathbf{x}, \mathsf{A}(\mathbf{x}'), \mathbf{x}'') = \mathsf{A}(\mathbf{x}, \mathbf{x}', \mathbf{x}'')$$

for all $\mathbf{x}, \mathbf{x}', \mathbf{x}'' \in \cup_{n\in\mathbb{N}_0}\mathbb{I}^n$.

Associativity is also a well-known algebraic property which allows one to omit "parentheses" in an aggregation of at least three elements. Implicit in the assumption of associativity is a consistent way of going unambiguously from the aggregation of n elements to $n+1$ elements, which implies that any associative extended aggregation function A is completely determined by its binary aggregation function $\mathsf{A}^{(2)}$. Indeed, by associativity, we clearly have

$$\mathsf{A}(x_1, \ldots, x_{n+1}) = \mathsf{A}\big(\mathsf{A}(x_1, \ldots, x_n), x_{n+1}\big).$$

For practical purpose we can start with the aggregation procedure before knowing all inputs to be aggregated. Additional input data are then simply aggregated with the current aggregated output.

Each associative idempotent extended aggregation function is necessary strongly idempotent. For a fixed arity $n > 2$, we can introduce the associativity as follows.

Definition 16. Let $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ be an n-ary aggregation function. Then it is associative if, for all $x_1, \ldots, x_{2n-1} \in \mathbb{I}$, we have

$$\begin{aligned}\mathsf{A}(\mathsf{A}(x_1, \ldots, x_n), x_{n+1}, \ldots, x_{2n-1}) &= \mathsf{A}(x_1, \mathsf{A}(x_2, \ldots, x_{n+1}), x_{n+2}, \ldots, x_{2n-1}) \\ &= \mathsf{A}(x_1, \ldots, x_{n-1}, \mathsf{A}(x_n, \ldots, x_{2n-1})).\end{aligned}$$

2.7 Decomposability

Introduced first in Bemporad [8, p. 87] in a characterization of the arithmetic mean, associativity of means has been used by Kolmogoroff [49] and Nagumo [71] to characterize the so-called mean values. More recently, Marichal and Roubens [62] proposed to call this property "decomposability" in order not to confuse it with classical associativity. Alternative names, such as *associativity with repetitions* or *weighted associativity*, could be naturally considered as well.

When symmetry is not assumed, it is necessary to rewrite this property in such a way that the first variables are not privileged. To abbreviate notations, for nonnegative integers m, n, we write $\mathsf{A}(m{\cdot}x, n{\cdot}y)$ what means the repetition of arguments, i.e., $\mathsf{A}(\underbrace{x, \ldots, x}_{m}, \underbrace{y, \ldots, y}_{n})$. We then consider the following definition.

Definition 17. An extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is *decomposable* if $\mathsf{A}(x) = x$ for all $x \in \mathbb{I}$ and if

$$\mathsf{A}(\mathbf{x}, \mathbf{x}') = \mathsf{A}\big(k \cdot \mathsf{A}(\mathbf{x}), k' \cdot \mathsf{A}(\mathbf{x}')\big) \tag{8}$$

for all $k, k' \in \mathbb{N}_0$, all $\mathbf{x} \in \mathbb{I}^k$, and all $\mathbf{x}' \in \mathbb{I}^{k'}$.

As the following proposition shows, decomposability means that each element of any subset of consecutive arguments can be replaced with their partial aggregation without changing the overall aggregation.

Proposition 4. An extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is decomposable if and only if

$$\mathsf{A}\big(\mathbf{x}, k' \cdot \mathsf{A}(\mathbf{x}'), \mathbf{x}''\big) = \mathsf{A}(\mathbf{x}, \mathbf{x}', \mathbf{x}'')$$

for all $k' \in \mathbb{N}_0$, all $\mathbf{x}' \in \mathbb{I}^{k'}$, and all $\mathbf{x}, \mathbf{x}'' \in \cup_{n\in\mathbb{N}_0}\mathbb{I}^n$.

Among till now introduced extended aggregation functions, AM, GM, Min, Max, P_1 are decomposable, and Σ, Π, S_L are not decomposable.

2.8 Bisymmetry

Another grouping property is the bisymmetry.

Definition 18. An aggregation function $\mathsf{A} : \mathbb{I}^2 \to \mathbb{I}$ is *bisymmetric* if for all $(x_1, x_2, x_3, x_4) \in \mathbb{I}^4$, we have

$$\mathsf{A}\big(\mathsf{A}(x_1, x_2), \mathsf{A}(x_3, x_4)\big) = \mathsf{A}\big(\mathsf{A}(x_1, x_3), \mathsf{A}(x_2, x_4)\big).$$

The bisymmetry property is very easy to handle and has been investigated from the algebraic point of view by using it mostly in structures without the property of associativity. For a list of references see Aczél [2, Section 6.4] (see also Aczél and Dhombres [3, Chapter 17], and Soublin [82]).

For n arguments, bisymmetry takes the following form (see Aczél [1]).

Definition 19. An aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is *bisymmetric* if

$$\begin{aligned} & \mathsf{A}\big(\mathsf{A}(x_{11}, \ldots, x_{1n}), \ldots, \mathsf{A}(x_{n1}, \ldots, x_{nn})\big) \\ = \; & \mathsf{A}\big(\mathsf{A}(x_{11}, \ldots, x_{n1}), \ldots, \mathsf{A}(x_{1n}, \ldots, x_{nn})\big) \end{aligned}$$

for all square matrices

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nn} \end{pmatrix} \in \mathbb{I}^{n\times n}.$$

Bisymmetry expresses the condition that aggregation of all the elements of any square matrix can be performed first on the rows, then on the columns, or conversely. However, since only square matrices are involved, this property seems not to have a good interpretation in terms of aggregation. Its usefulness remains theoretical. We then consider it for extended functions as follows; see Marichal et al. [61].

Definition 20. An extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is *strongly bisymmetric* if $\mathsf{A}(x) = x$ for all $x \in \mathbb{I}$, and if, for any $n, p \in \mathbb{N}$, we have

$$\begin{aligned} & \mathsf{A}\big(\mathsf{A}(x_{11}, \ldots, x_{1n}), \ldots, \mathsf{A}(x_{p1}, \ldots, x_{pn})\big) \\ = \; & \mathsf{A}\big(\mathsf{A}(x_{11}, \ldots, x_{p1}), \ldots, \mathsf{A}(x_{1n}, \ldots, x_{pn})\big) \end{aligned}$$

for all matrices

$$\begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix} \in \mathbb{I}^{p\times n}.$$

Remark 1. Contrary to bisymmetry, the strong bisymmetry property can be justified rather easily. Consider n judges (or criteria, attributes, etc.) giving a numerical score to each of p candidates. These scores, assumed to be defined on the same scale, can be put in a $p \times n$ matrix like

$$\begin{array}{c} \\ C_1 \\ \vdots \\ C_p \end{array} \begin{array}{c} \begin{array}{ccc} J_1 & \cdots & J_n \end{array} \\ \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix} \end{array}$$

Suppose now that we want to aggregate all the entries (scores) of the matrix in order to obtain an overall score of the p candidates. A reasonable way to proceed could be the following. First aggregate the scores of each candidate (aggregation over the rows of the matrix), and then aggregate these overall values. An alternative way to proceed would be to first aggregate the scores given by each judge (aggregation over the columns of the matrix), and then aggregate these values. The strong bisymmetry property means that these two ways to aggregate must lead to the same overall score, which is a natural property. Of course, we could as well consider only one candidate, n judges, and p criteria (assuming commensurateness of the scores along the criteria). In this latter setting, strong bisymmetry seems very natural as well.

Note that AM, GM, Min, Max, Π, Σ are strongly bisymmetric. The extended aggregation function A given by (6) is bisymmetric but not strongly bisymmetric. n-ary aggregation function A_c, see (5), is not bisymmetric whenever $c \in \,]0,1[\,$.

2.9 Some other properties

The neutral element is again a well-known notion coming from the area of binary operations. Recall that for a binary operation $*$ defined on a domain X, an element $e \in X$ is called a *neutral element* (of the operation $*$) if

$$x * e = e * x = x \qquad (x \in X).$$

Clearly, any binary operation $*$ can have at most one neutral element. From the previous equalities we can see that the action of the neutral element of a binary operation has the same effect as its omission. This idea is the background of the general definition.

Definition 21. Let $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ be an extended aggregation function. An element $e \in \mathbb{I}$ is called an *extended neutral element* of A if, for any $i \in \{1,\ldots,n\}$ and any $\mathbf{x} \in \mathbb{I}^n$ such that $x_i = e$, then

$$\mathsf{A}(x_1,\ldots,x_n) = \mathsf{A}(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n).$$

So the extended neutral element can be omitted from the input values without influencing the aggregated value. In multicriteria decision making, assigning a score equal to the extended neutral element (if it exists) to some criterion means that only the other criteria fulfillments are decisive for the overall evaluation.

For n-ary functions, there is an alternative approach, given in the following definition:

Definition 22. An element $e \in \mathbb{I}$ is called a *neutral element* of an aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ if, for any $i \in \{1, \ldots, n\}$ and any $x \in \mathbb{I}$, we have $\mathsf{A}(x_{\{i\}}e) = x$, where $x_{\{i\}}e$ means that all coordinates are e except the ith which is x.

Clearly, if $e \in \mathbb{I}$ is an extended neutral element of an extended aggregation function $\mathsf{A} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$, with $\mathsf{A}^{(1)}(x) = x$, then e is a neutral element of all $\mathsf{A}^{(n)}$, $n \in \mathbb{N}$. For instance, $e = 0$ is an extended neutral element for the extended sum function Σ.

Definition 23. An element $a \in \mathbb{I}$ is called an *annihilator element* of an aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ if, for any $\mathbf{x} \in \mathbb{I}^n$ such that $a \in \{x_1, \ldots, x_n\}$, we have $\mathsf{A}(\mathbf{x}) = a$.

Proposition 5. Consider an aggregation function $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$. If A is conjunctive and $a := \inf \mathbb{I} \in \mathbb{I}$ then a is an annihilator element. Dually, if A is disjunctive and $b := \sup \mathbb{I} \in \mathbb{I}$ then b is an annihilator element.

The converse of Proposition 5 is false. For instance, in $[0, 1]^n$, 0 is an annihilator of the geometric mean GM, which is not conjunctive.

The previously discussed properties of aggregation functions were mostly algebraic ones (up to the continuity). From several analytical properties of aggregation functions, we recall some properties related to additivity.

Definition 24. An aggregation function $\mathsf{A}^{(n)} : \mathbb{I}^n \to \mathbb{I}$ is called

(i) additive, if for any $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$ such that also $\mathbf{x} + \mathbf{y} \in \mathbb{I}^n$, we have

$$\mathsf{A}^{(n)}(\mathbf{x} + \mathbf{y}) = \mathsf{A}^{(n)}(\mathbf{x}) + \mathsf{A}^{(n)}(\mathbf{y});$$

(ii) modular, if for any $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$, we have

$$\mathsf{A}^{(n)}(\mathbf{x} \vee \mathbf{y}) + \mathsf{A}^{(n)}(\mathbf{x} \wedge \mathbf{y}) = \mathsf{A}^{(n)}(\mathbf{x}) + \mathsf{A}^{(n)}(\mathbf{y});$$

(iii) superadditive, if for any $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$ such that also $\mathbf{x} + \mathbf{y} \in \mathbb{I}^n$, we have

$$\mathsf{A}^{(n)}(\mathbf{x} + \mathbf{y}) \geqslant \mathsf{A}^{(n)}(\mathbf{x}) + \mathsf{A}^{(n)}(\mathbf{y});$$

(iv) supermodular, if for any $\mathbf{x}, \mathbf{y} \in \mathbb{I}^n$, we have

$$\mathsf{A}^{(n)}(\mathbf{x} \vee \mathbf{y}) + \mathsf{A}^{(n)}(\mathbf{x} \wedge \mathbf{y}) \geqslant \mathsf{A}^{(n)}(\mathbf{x}) + \mathsf{A}^{(n)}(\mathbf{y}).$$

Note that additivity implies modularity, and the last implies supermodularity and additivity implies superadditivity. Arithmetic mean AM is an example of an aggregation function satisfying all properties mentioned in Definition 24. The product Π on $[0, 1]$ is supermodular and superadditive but not modular (and thus neither additive).

For more specific properties of aggregation functions we recommend to consider [38], Chapter 2.

3 Means and averages

Means and averaging functions seem to be from the historical point of view the first aggregation functions (not bearing the name "aggregation function" in the time of their birth) exploited in many areas. Already discovered and studied by the ancient Greeks, see Antoine [5, Chapter 3] for a historical discussion of the various Greek notions of "mean", the concept of mean has given rise today to a very wide field of investigation with a huge variety of applications. Actually, a tremendous amount of literature on the properties of several means (such as the arithmetic mean, the geometric mean, etc.) has already been produced, especially since the 19th century, and is still developing today. For a good overview, see the expository paper by Frosini [32] and the remarkable monograph by Bullen [13].

The first modern definition of mean was probably due to Cauchy [17] who considered in 1821 a *mean* as an internal function $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$, i.e., $\mathsf{Min} \leqslant \mathsf{M} \leqslant \mathsf{Max}$. We adopt this approach and assume further that a mean should be a nondecreasing function.

As usual, $\mathbb{I}$ represents a nonempty real interval, bounded or not. The more general cases where $\mathbb{I}$ includes $-\infty$ and/or ∞ will always be mentioned explicitly.

Definition 25. An n-ary *mean* in $\mathbb{I}^n$ is an internal aggregation function $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$. An *extended mean* in $\cup_{n\in\mathbb{N}}\mathbb{I}^n$ is an extended function $\mathsf{M} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ whose restriction to each $\mathbb{I}^n$ is a mean.

It follows that a mean is nothing other than an idempotent aggregation function. Moreover, if $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is a mean in $\mathbb{I}^n$, then it is also a mean in $\mathbb{J}^n$, for any subinterval $\mathbb{J} \subseteq \mathbb{I}$.

The concept of mean as an *average* or *numerical equalizer* is usually ascribed to Chisini [20, p. 108], who gave in 1929 the following definition:

> Let $y = \mathsf{F}(x_1, \ldots, x_n)$ be a function of n independent variables $x_1, \ldots, x_n$. A mean of $x_1, \ldots, x_n$ with respect to the function F is a number M such that, if each of $x_1, \ldots, x_n$ is replaced by M, the function value is unchanged, that is,
>
> $$\mathsf{F}(M, \ldots, M) = \mathsf{F}(x_1, \ldots, x_n).$$

When F is considered as the sum, the product, the sum of squares, the sum of inverses, or the sum of exponentials, the solution of Chisini's equation corresponds respectively to the arithmetic mean, the geometric mean, the quadratic mean, the harmonic mean, and the exponential mean.

Unfortunately, as noted by de Finetti [25, p. 378] in 1931, Chisini's definition is so general that it does not even imply that the "mean" (provided there exists a real and unique solution to Chisini's equation) fulfills Cauchy's internality property.

To ensure existence, uniqueness, and internality of the solution of Chisini's equation, we assume that F is nondecreasing and idempotizable. Therefore we propose the following definition:

Definition 26. A function $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is an *average* in $\mathbb{I}^n$ if there exists a nondecreasing and idempotizable function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ such that $\mathsf{F} = \delta_{\mathsf{F}} \circ \mathsf{M}$. In this case, we say that M is an *average associated with* F *in* $\mathbb{I}^n$.

Averages are also known as *Chisini means* or *level surface means*. The average associated with F is also called the F-*level mean* (see Bullen [13, VI.4.1]). The following result shows that, thus defined, the concepts of mean and average coincide.

Proposition 6. The following assertions hold:

(i) Any average is a mean.

(ii) Any mean is the average associated with itself.

(iii) Let M be the average associated with a function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$. Then M is the average associated with a function $\mathsf{G} : \mathbb{I}^n \to \mathbb{R}$ if and only if there exists an increasing bijection $\varphi : \mathrm{ran}(\mathsf{F}) \to \mathrm{ran}(\mathsf{G})$ such that $\mathsf{G} = \varphi \circ \mathsf{F}$.

Proposition 6 shows that, thus defined, the concepts of mean and average are identical and, in a sense, rather general. Note that some authors (see for instance Bullen [13, p. xxvi], Sahoo and Riedel [79, Section 7.2], and Bhatia [12, Chapter 4]) define the concept of mean by adding conditions such as continuity, symmetry, and homogeneity, which is $\mathsf{M}(r\,\mathbf{x}) = r\,\mathsf{M}(\mathbf{x})$ for all admissible $r \in \mathbb{R}$.

3.1 Quasi-arithmetic means

A well-studied class of means is the class of *quasi-arithmetic means* (see for instance Bullen [13, Chapter IV]), introduced as extended aggregation functions as early as 1930 by Kolmogoroff [49], Nagumo [71], and then as n-ary functions in 1948 by Aczél [1]. In this section we introduce the quasi-arithmetic means and describe some of their properties and axiomatizations.

Definition 27. Let $f : \mathbb{I} \to \mathbb{R}$ be a continuous and strictly monotonic function. The n-ary *quasi-arithmetic mean generated by f* is the function $\mathsf{M}_f : \mathbb{I}^n \to \mathbb{I}$ defined as

$$\mathsf{M}_f(\mathbf{x}) := f^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big). \tag{9}$$

The *extended quasi-arithmetic mean generated by f* is the function $\mathsf{M}_f : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ whose restriction to $\mathbb{I}^n$ is the n-ary quasi-arithmetic mean generated by f.

Remark 2. (i) Each quasi-arithmetic mean M_f is a mean in the sense of Definition 25. It is also the average associated with $n(f\circ\mathsf{M}_f)$. For instance, the arithmetic mean $\mathsf{M}_f = \mathsf{AM}$ (with $f = \mathrm{id}$) is the average associated with the sum Σ.

(ii) In certain applications it may be convenient to extend the range of f to the extended real line $\overline{\mathbb{R}} = [-\infty, \infty]$. Evidently, in this case it is necessary to define the expression $\infty - \infty$, which will often be considered as $-\infty$.

The class of quasi-arithmetic means comprises most of the algebraic means of common use such as the arithmetic mean and the geometric mean. Table 1 provides some well-known instances of quasi-arithmetic means.

The function f occurring in (9) is called a *generator* of M_f. Aczél [1] showed that f is determined up to a linear transformation. More generally, we have the following result (see Bullen et al. [14, p. 226]):

$f(x)$	$\mathsf{M}_f(\mathbf{x})$	name	notation
x	$\frac{1}{n}\sum_{i=1}^{n} x_i$	arithmetic mean	AM
x^2	$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^{1/2}$	quadratic mean	QM
$\log x$	$\left(\prod_{i=1}^{n} x_i\right)^{1/n}$	geometric mean	GM
x^{-1}	$\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}}$	harmonic mean	HM
x^α $(\alpha \in \mathbb{R}\setminus\{0\})$	$\left(\frac{1}{n}\sum_{i=1}^{n} x_i^\alpha\right)^{1/\alpha}$	root-mean-power	$\mathsf{M}_{\mathrm{id}^\alpha}$
$e^{\alpha x}$ $(\alpha \in \mathbb{R}\setminus\{0\})$	$\frac{1}{\alpha}\ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{\alpha x_i}\right)$	exponential mean	EM_α

Table 1: Examples of quasi-arithmetic means

Proposition 7. Let $f, g : \mathbb{I} \to \mathbb{R}$ be continuous and strictly monotonic functions. Assume also that g is increasing (respectively, decreasing). Then

(i) $\mathsf{M}_f \leqslant \mathsf{M}_g$ if and only if $g \circ f^{-1}$ is convex (respectively, concave);

(ii) $\mathsf{M}_f = \mathsf{M}_g$ if and only if $g \circ f^{-1}$ is linear, that is,

$$g(x) = r f(x) + s \qquad (r, s \in \mathbb{R},\ r \neq 0).$$

We now present an axiomatization of the class of quasi-arithmetic means as extended aggregation functions, originally called *mean values*. The next theorem brings an axiomatization of quasi-arithmetic means as extended aggregation functions. This axiomatization was obtained independently by Kolmogoroff [49] and Nagumo [71] in 1930.

Theorem 1. $\mathsf{F} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{R}$ is symmetric, continuous, strictly increasing, idempotent, and decomposable if and only if there is a continuous and strictly monotonic function $f : \mathbb{I} \to \mathbb{R}$ such that $\mathsf{F} = \mathsf{M}_f$ is the extended quasi-arithmetic mean generated by f.

Another axiomatization of n-ary quasi-arithmetic means is due to Aczél [1].

Theorem 2. The function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ is symmetric, continuous, strictly increasing, idempotent, and bisymmetric if and only if there is a continuous and strictly monotonic function $f : \mathbb{I} \to \mathbb{R}$ such that $\mathsf{F} = \mathsf{M}_f$ is the quasi-arithmetic mean generated by f.

Remark 3. Note that the results given in Theorem 1, and Theorem 2 can be extended to subintervals $\mathbb{I}$ of the extended real line containing ∞ or $-\infty$ with a slight modification of the requirements. Namely, the codomain of F should be $[-\infty, \infty]$, and the strict monotonicity and continuity are required for bounded input vectors only. Observe also that if $\mathbb{I} = [-\infty, \infty]$ then the corresponding quasi-arithmetic means are no more continuous due to the non-continuity of the standard summation on $[-\infty, \infty]$.

Adding some particular property, a special subfamily of quasi-arithmetic means is obtained. Due to Nagumo [71] we have the next result.

Theorem 3. (i) The quasi-arithmetic mean $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is difference scale invariant, i.e., $\mathsf{M}(\mathbf{x} + \mathbf{c}) = \mathsf{M}(\mathbf{x}) + c$ for all $\mathbf{x} \in \mathbb{I}^n$ and $\mathbf{c} = (c, \ldots, c) \in \mathbb{R}^n$ such that $\mathbf{x} + \mathbf{c} \in \mathbb{I}^n$, if and only if either M is the arithmetic mean AM or M is the exponential mean, i.e., there exists $\alpha \in \mathbb{R} \setminus \{0\}$ such that

$$\mathsf{M}(\mathbf{x}) = \frac{1}{\alpha} \ln\Big(\frac{1}{n}\sum_{i=1}^{n} e^{\alpha x_i}\Big).$$

(ii) Assume $\mathbb{I} \subseteq \,]0, \infty[$. The quasi-arithmetic mean $\mathsf{M} : \mathbb{I}^n \to \mathbb{I}$ is ratio scale invariant, i.e., $\mathsf{M}(c\mathbf{x}) = c\mathsf{M}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{I}^n$ and $c > 0$ such that $c\mathbf{x} \in \mathbb{I}^n$, if and only if either M is the geometric mean GM or M is the root-mean-power, i.e., there exists $\alpha \in \mathbb{R} \setminus \{0\}$ such that

$$\mathsf{M}(\mathbf{x}) = \Big(\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha}\Big)^{\frac{1}{\alpha}}.$$

A modification of Theorem 2, where the symmetry requirement is omitted, yields an axiomatic characterization of weighted quasi-arithmetic means, see Aczél [1] (these are called also quasi-linear means in some sources).

Theorem 4. The function $\mathsf{F} : \mathbb{I}^n \to \mathbb{R}$ is continuous, strictly increasing, idempotent, and bisymmetric if and only if there is a continuous and strictly monotonic function $f : \mathbb{I} \to \mathbb{R}$ and real numbers $w_1, \ldots, w_n > 0$ satisfying $\sum_i w_i = 1$ such that

$$\mathsf{F}(\mathbf{x}) = f^{-1}\Big(\sum_{i=1}^{n} w_i\, f(x_i)\Big) \qquad (\mathbf{x} \in \mathbb{I}^n). \tag{10}$$

Weighted quasi-arithmetic means can be seen as transformed weighted arithmetic means. The latter means are trivially characterized by the additivity property, i.e., $\mathsf{F}(\mathbf{x} + \mathbf{y}) = \mathsf{F}(\mathbf{x}) + \mathsf{F}(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y}, \mathbf{x} + \mathbf{y} \in \mathrm{Dom}(\mathsf{F})$ (domain of F).

Proposition 8. $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$ is additive, nondecreasing, and idempotent if and only if there exists a weight vector $\mathbf{w} \in [0, 1]^n$ satisfying $\sum_i w_i = 1$ such that $\mathsf{F} = \mathsf{WAM}_{\mathbf{w}}$.

Corollary 1. $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$ is additive, nondecreasing, symmetric, and idempotent if and only if $\mathsf{F} = \mathsf{AM}$ is the arithmetic mean.

Remark 4. As will be discussed in Section 4, the weighted arithmetic means $\mathsf{WAM}_{\mathbf{w}}$ are exactly the Choquet integrals with respect to additive normalized capacities; see Proposition 10 (v). If we further assume the symmetry property, we obtain the arithmetic mean AM.

$f(x)$	$\mathsf{M}(\mathbf{x})$	name	notation
x	$\sum_{i=1}^{n} w_i\, x_i$	weighted arithmetic mean	$\mathsf{WAM}_{\mathbf{w}}$
x^2	$\Big(\sum_{i=1}^{n} w_i\, x_i^2\Big)^{1/2}$	weighted quadratic mean	$\mathsf{WQM}_{\mathbf{w}}$
$\log x$	$\prod_{i=1}^{n} x_i^{w_i}$	weighted geometric mean	$\mathsf{WGM}_{\mathbf{w}}$
x^α $(\alpha \in \mathbb{R}\setminus\{0\})$	$\Big(\sum_{i=1}^{n} w_i\, x_i^\alpha\Big)^{1/\alpha}$	weighted root-mean-power	$\mathsf{WM}_{\mathrm{id}^\alpha,\mathbf{w}}$

Table 2: Examples of quasi-linear means

A natural way to generalize the quasi-arithmetic mean consists in incorporating weights as in the quasi-linear mean (10). To generalize a step further, we could assume that the weights are not constant. On this issue, Losonczi [53, 54] considered and investigated in 1971 nonsymmetric functions of the form

$$\mathsf{M}(\mathbf{x}) = f^{-1}\Bigg(\frac{\sum_{i=1}^n p_i(x_i)f(x_i)}{\sum_{i=1}^n p_i(x_i)}\Bigg),$$

where $f : \mathbb{I} \to \mathbb{R}$ is a continuous and strictly monotonic function and $p_1, \dots, p_n : \mathbb{I} \to]0, \infty[$ are positive valued functions. The special case where $p_1 = \cdots = p_n$ was previously introduced in 1958 by Bajraktarević [6] who defined the concept of *quasi-arithmetic mean with weight function* as follows (see also Páles [74]). The subcase where $f = \mathrm{id}$, called *Beckenbach-Gini means* or *mixture operators*, has been investigated by Marques Pereira and Ribeiro [63], Matkowski [65], and Yager [96].

Definition 28. Let $f : \mathbb{I} \to \mathbb{R}$ be a continuous and strictly monotonic function and let $p : \mathbb{I} \to]0, \infty[$ be a positive valued function. The n-ary *quasi-arithmetic mean generated by f with weight function p* is the function $\mathsf{M}_{f,p} : \mathbb{I}^n \to \mathbb{I}$ defined as

$$\mathsf{M}_{f,p}(\mathbf{x}) = f^{-1}\Bigg(\frac{\sum_{i=1}^n p(x_i)f(x_i)}{\sum_{i=1}^n p(x_i)}\Bigg).$$

The *extended quasi-arithmetic mean generated by f with weight function p* is the function $\mathsf{M}_{f,p} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ whose restriction to $\mathbb{I}^n$ is the n-ary quasi-arithmetic mean generated by f with weight function p.

It is very important to note that, even though quasi-arithmetic means with weight function are (clearly) idempotent, they need not be nondecreasing, which implies that they need not be means or even aggregation functions. To give an example, consider the case where $n = 2$, $f(x) = x$, and $p(x) = 2x + 1$, that is,

$$\mathsf{M}_{f,p}(x_1, x_2) = \frac{2x_1^2 + 2x_2^2 + x_1 + x_2}{2x_1 + 2x_2 + 2}.$$

We can readily see that the section $x \mapsto \mathsf{M}_{f,p}(x, 1)$ of this binary function is not nondecreasing.

Assuming that the weight function p is nondecreasing and differentiable, Marques Pereira and Ribeiro [63] and Mesiar and Špirková [68] found sufficient conditions on p to ensure nondecreasing monotonicity of $\mathsf{M}_{\mathrm{id},p}$. Here we assume $\mathbb{I} = [0, 1]$ as in [68] but the conditions easily extend to arbitrary intervals. For extended arithmetic means $\mathsf{M}_{\mathrm{id},p}$ with weight function p, the simplest sufficient condition to ensure nondecreasing monotonicity is $p(x) \geqslant p'(x) \geqslant 0$. A more general one is

$$p(x) \geqslant (1 - x)p'(x) \geqslant 0 \qquad (x \in [0, 1]).$$

Remark 5. For extended quasi-arithmetic means $\mathsf{M}_{f,p}$ with weight function p, assuming that both f and p are increasing and differentiable and that $\mathrm{ran}(f) = [0, 1]$, the above sufficient condition generalizes into

$$f'(x)\, p(x) \geqslant (1 - f(x))\, p'(x) \geqslant 0 \qquad (x \in [0, 1]).$$

For n-ary arithmetic means $\mathsf{M}_{\mathrm{id},p}$ with weight function p, we also have the sufficient condition

$$\frac{p^2(x)}{(n-1)\, p(1)} + p(x) \geqslant (1 - x)p'(x) \qquad (x \in [0, 1]).$$

Remark 6. It is worth mentioning that $\mathsf{M}_{f,p}$ can also be obtained by the minimization problem

$$\mathsf{M}_{f,p}(\mathbf{x}) = \arg\min_{r \in \mathbb{R}} \sum_{i=1}^{n} p(x_i)\, (f(x_i) - f(r))^2.$$

Evidently, in the same way, classical quasi-arithmetic means (p is constant) and weighted quasi-arithmetic means (replace $p(x_i)$ with w_i) are obtained. For more details, see Calvo et al. [16] and Mesiar and Špirková [69].

3.2 Constructions of means

There are several methods on how to construct means (binary, n-ary, extended). Integral-based methods are discussed in Section 4.

Means can also be constructed by minimization of functions. This construction method will be thoroughly discussed in our subsequent paper.

To give here a simple example, consider weights $w_1, w_2 \in]0, \infty[$ and minimize (in r) the expression

$$f(r) = w_1|x_1 - r| + w_2(x_2 - r)^2.$$

This minimization problem leads to the unique solution

$$r = \mathsf{M}(x_1, x_2) = \mathsf{Med}\big(x_1, x_2 - \frac{w_1}{2w_2}, x_2 + \frac{w_1}{2w_2}\big),$$

which defines a mean $\mathsf{M}:\mathbb{R}^2\to\mathbb{R}$.

Any nonsymmetric function F can be symmetrized by replacing it variables $x_1,\ldots,x_n$ with corresponding order statistics functions $x_{(1)}$ (minimal input), $x_{(2)},\ldots,x_{(n)}$ (maximal input).

One of the simplest examples is given by the *ordered weighted averaging function*

$$\mathsf{OWA}_{\mathbf{w}}(\mathbf{x})=\sum_{i=1}^{n} w_i\, x_{(i)}, \tag{11}$$

which merely results from the symmetrization of the corresponding weighted arithmetic mean $\mathsf{WAM}_{\mathbf{w}}$.

Remark 7. The concept of ordered weighted averaging function was introduced by Yager in 1988; see Yager [95], and also the book [94] edited by Yager and Kacprzyk. Note however that linear (not necessarily convex) combinations of ordered statistics were already studied previously in statistics; see for instance Weisberg [89] (and David and Nagaraja [23, Section 6.5] for a more recent overview). Since then, the family of these functions has been axiomatized in various ways; see for instance Fodor et al. [30] and Marichal and Mathonet [60]. Also, these functions are exactly the Choquet integrals with respect to symmetric normalized capacities; see Proposition 10 (vi).

The symmetrization process can naturally be applied to the quasi-linear mean (i.e., to the weighted quasi-arithmetic mean) (10) to produce the *quasi-ordered weighted averaging function* $\mathsf{OWA}_{\mathbf{w},f}:\mathbb{I}^n\to\mathbb{R}$, which is defined as

$$\mathsf{OWA}_{\mathbf{w},f}(\mathbf{x}):=f^{-1}\Big(\sum_{i=1}^{n} w_i f(x_{(i)})\Big),$$

where the generator $f:\mathbb{I}\to\mathbb{R}$ is a continuous and strictly monotonic function; see Fodor et al. [30].

The classical mean value formulas (Lagrange, Cauchy) lead to the concept of Lagrange (Cauchy) means, see [14] and [10, 11].

Definition 29. Let $f:\mathbb{I}\to\mathbb{R}$ be a continuous and strictly monotonic function. The *Lagrangian mean* $\mathsf{M}_{[f]}:\mathbb{I}^2\to\mathbb{I}$ associated with f is a mean defined as

$$\mathsf{M}_{[f]}(x,y):=\begin{cases} f^{-1}\left(\dfrac{1}{y-x}\displaystyle\int_x^y f(t)\,\mathrm{d}t\right) & \text{if } x\neq y\\ x & \text{if } x=y.\end{cases} \tag{12}$$

The uniqueness of the generator is the same as for quasi-arithmetic means: Let f and g be two generators of the same Lagrangian mean. Then, there exist $r,s\in\mathbb{R}$, $r\neq 0$ such that $g(x)=rf(x)+s$; see [10, Corollary 7], [14, p. 344], and [64, Theorem 1].

Many classical means are Lagrangian. The arithmetic mean, the geometric mean, and the so-called *Stolarsky means* [83], defined by

$$\mathsf{M}_S(x,y):=\begin{cases}\left(\dfrac{x^r-y^r}{r(x-y)}\right)^{\frac{1}{r-1}} & \text{if } x\neq y\\ x & \text{if } x=y,\end{cases}$$

correspond to taking $f(x) = x$, $f(x) = 1/x^2$, and $f(x) = x^{r-1}$, respectively, in (12). The harmonic mean, however, is not Lagrangian.

In general, some of the most common means are both quasi-arithmetic and Lagrangian, but there are quasi-arithmetic means, like the harmonic one, which are not Lagrangian. Conversely, the *logarithmic mean*

$$\mathsf{M}(x,y) := \begin{cases} \dfrac{x-y}{\log x - \log y} & \text{for } x,y>0,\ x\neq y \\ x & \text{for } x=y>0, \end{cases}$$

is an example of a Lagrangian mean (actually a Stolarsky mean, $f(x) = 1/x$), that is not quasi-arithmetic. A characterization of the class of Lagrangian means and a study of its connections with the class of quasi-arithmetic means can be found in Berrone and Moro [10]. Further properties of Lagrangian means and other extensions are investigated for instance in Aczél and Kuczma [4], Berrone [9], Głazowska [34], Horwitz [42, 43], Kuczma [50], Sándor [80], and Wimp [90].

Definition 30. Let $f,g : \mathbb{I} \to \mathbb{R}$ be continuous and strictly monotonic functions. The *Cauchy mean* $\mathsf{M}_{[f,g]} : \mathbb{I}^2 \to \mathbb{I}$ associated with the pair (f,g) is a mean defined as

$$\mathsf{M}_{[f,g]}(x,y) := \begin{cases} f^{-1}\left(\dfrac{1}{g(y)-g(x)} \displaystyle\int_x^y f(t)\,\mathrm{d}g(t)\right) & \text{if } x\neq y \\ x & \text{if } x=y. \end{cases}$$

We note that any Cauchy mean is continuous, idempotent, symmetric, and strictly increasing.

When $g = f$ (respectively, g is the identity function), we retrieve the quasi-arithmetic (respectively, the Lagrangian) mean generated by f. The *anti-Lagrangian mean* [11] is obtained when f is the identity function. For example, the harmonic mean is an anti-Lagrangian mean generated by the function $g = 1/x^2$. We also note that the generator of an anti-Lagrangian mean is defined up to a non-zero affine transformation.

Further studies on Cauchy means can be found for instance in Berrone [9], Lorenzen [52], and Losonczi [55, 56]. Extensions of Lagrangian and Cauchy means, called *generalized weighted mean values*, including discussions on their monotonicity properties, can be found in Chen and Qi [18, 19], Qi et al. [75, 76, 77], and Witkowski [91].

3.3 Associative means

The class of continuous, nondecreasing, idempotent, and associative binary functions is described in the next theorem. The result is due to Fodor [29] who obtained this description in a more general framework, where the domain of variables is any connected order topological space. Alternative proofs were obtained independently in Marichal [57, Section 3.4] and [58, Section 5].

Theorem 5. $\mathsf{M} : \mathbb{I}^2 \to \mathbb{I}$ is continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha, \beta \in \mathbb{I}$ such that

$$\mathsf{M}(x,y) = (\alpha \wedge x) \vee (\beta \wedge y) \vee (x \wedge y). \tag{13}$$

Notice that, by distributivity of $\wedge$ and $\vee$, M can be written also in the equivalent form:

$$\mathsf{M}(x,y) = (\beta \vee x) \wedge (\alpha \vee y) \wedge (x \vee y).$$

The graphical representation of M is given in Figure 1.

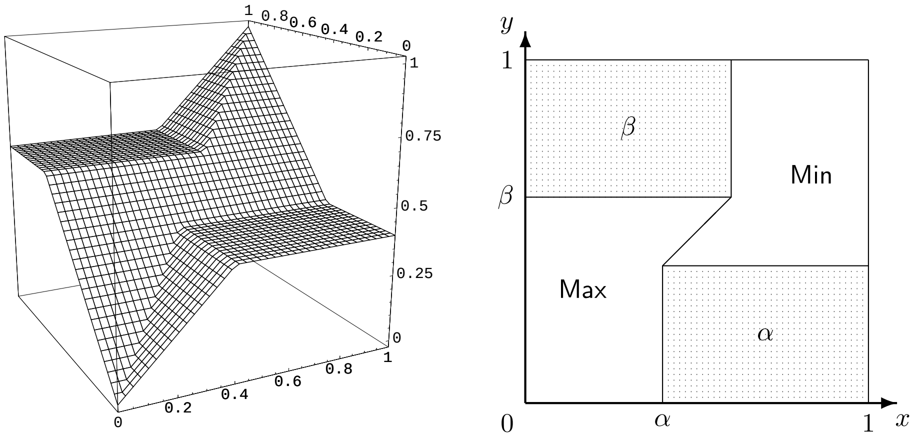

Figure 1: Representation on $[0,1]^2$ of function (13) when $\alpha \leqslant \beta$

Theorem 5 can be generalized straightforwardly to extended means as follows.

Theorem 6. $\mathsf{M} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha, \beta \in \mathbb{I}$ such that, for any $n \in \mathbb{N}$,

$$\mathsf{M}^{(n)}(\mathbf{x}) = (\alpha \wedge x_1) \vee \Big(\bigvee_{i=2}^{n-1} (\alpha \wedge \beta \wedge x_i)\Big) \vee (\beta \wedge x_n) \vee \Big(\bigwedge_{i=1}^{n} x_i\Big). \tag{14}$$

Remark 8. Means described in Theorem 5 are nothing other than idempotent binary lattice polynomial functions, that is, binary Sugeno integrals; see Section 4. We also observe that the n-ary lattice polynomial function given in (14) is an n-ary Sugeno integral defined from a particular normalized capacity.

The special case of symmetric associative means was already discussed by Fung and Fu [33] and revisited in Dubois and Prade [26]. It turns out that these functions are the α-medians, i.e., for $\alpha \in \mathbb{I}$ α-median , $\mathsf{Med}_\alpha : \mathbb{I}^n \to \mathbb{I}$ is given by

$$\mathsf{Med}_\alpha(\mathbf{x}) = \mathsf{Med}(x_1, \ldots, x_n, \underbrace{\alpha, \ldots, \alpha}_{n-1}) = \mathsf{Med}(\mathsf{Min}(\mathbf{x}), \alpha, \mathsf{Max}(\mathbf{x})).$$

The description is the following.

Theorem 7. $\mathsf{M} : \mathbb{I}^2 \to \mathbb{I}$ is symmetric, continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha \in \mathbb{I}$ such that $\mathsf{M} = \mathsf{Med}_\alpha$. Similarly, $\mathsf{M} : \cup_{n\in\mathbb{N}}\mathbb{I}^n \to \mathbb{I}$ is symmetric, continuous, nondecreasing, idempotent, and associative if and only if there exist $\alpha \in \mathbb{I}$ such that, for any $n \in \mathbb{N}$, $\mathsf{M}^{(n)} = \mathsf{Med}_\alpha^{(n)}$.

Remark 9. Since the conjunction of symmetry and associativity implies bisymmetry, we immediately see that the α-medians Med_α are particular nonstrict arithmetic means.

Czogała and Drewniak [22] have examined the case when M has a neutral element $e \in \mathbb{I}$. They obtained the following result.

Theorem 8. If $\mathsf{M} : \mathbb{I}^2 \to \mathbb{I}$ is nondecreasing, idempotent, associative, and has a neutral element $e \in \mathbb{I}$, then there is a nonincreasing function $g : \mathbb{I} \to \mathbb{I}$ with $g(e) = e$ such that, for all $x, y \in \mathbb{I}$,

$$\mathsf{M}(x,y) = \begin{cases} x \wedge y & \text{if } y < g(x) \\ x \vee y & \text{if } y > g(x) \\ x \wedge y \text{ or } x \vee y & \text{if } y = g(x). \end{cases}$$

Furthermore, if M is continuous, then $\mathsf{M} = \mathsf{Min}$ or $\mathsf{M} = \mathsf{Max}$.

Remark 10. (i) Fodor [29] showed that Theorem 8 still holds in the more general framework of connected order topological spaces.

(ii) The restriction of Theorem 8 to symmetric functions corresponds to idempotent uninorms.

4 Aggregation functions based on nonadditive integrals

The preceding section has developed the notion of means, which can be viewed as a variation of the idea of finite sum. Another generalization is the notion of integral, where the sum becomes infinite. Beside the classical Riemann integral, many types of integral exist, but there is one which is of particular interest to us, namely the Lebesgue integral, which is defined with respect to a measure. Indeed, the classical notion of measure extends the notion of weight to infinite universes, and the Lebesgue integral on a finite universe coincides with the weighted arithmetic mean. Therefore, the existence of more general notions of measure than the classical additive one, together with the appropriate integrals, offer a new realm of aggregation functions when these integrals are limited to a finite universe. These general measures may be called *nonadditive measures*, and the corresponding integrals *nonadditive integrals*, however a more precise term would be *monotonic measures* since additivity is replaced by monotonicity, although the most common name —which we will use—, is *capacity*, as coined by Choquet [21]. The term *fuzzy measure* introduced by Sugeno [85] is often used in the fuzzy set community, although this term is mathematically misleading since no fuzziness is involved there.

There are basically two types of integrals defined with respect to a capacity, namely the Choquet integral and the Sugeno integral, leading to two interesting classes of aggregation functions, developed in this section.

We start by introducing the notion of capacity.

Definition 31. Let $X = \{1, \ldots, n\}$ be a given universe. A set function $\mu : 2^X \to [0, \infty[$ is called a *capacity* if $\mu(\varnothing) = 0$ and $\mu(A) \leqslant \mu(B)$ whenever $A \subset B \subset X$ (monotonicity). If in addition $\mu(X) = 1$, the capacity is *normalized*.

If μ does not satisfy monotonicity, it is called a *game*. If μ takes only values 0 and 1, then μ is called a *0-1 capacity*. For any capacity μ, the *dual* capacity is defined by $\mu^d(A) := \mu(X) - \mu(X \setminus A)$, for any $A \subseteq X$.

Let μ be a capacity on X and $A, B \subseteq X$. We say that μ is *additive* if $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever A and B are disjoint; it is *symmetric* if $\mu(A) = \mu(B)$ whenever $|A| = |B|$; it is *maxitive* if $\mu(A \cup B) = \mu(A) \vee \mu(B)$, and it is *minitive* if $\mu(A \cap B) = \mu(A) \wedge \mu(B)$.

We give several fundamental examples of capacities.

(i) The smallest normalized capacity is $\mu_{\min}(A) := 0$, $\forall A \subsetneq X$, while the greatest one is $\mu_{\max}(A) := 1$, $\forall A \subseteq X, A \neq \varnothing$;

(ii) For any $i \in X$, the *Dirac measure centered on* i is defined by, for any $A \subseteq X$

$$\delta_i(A) := \begin{cases} 1 & \text{if } i \in A \\ 0 & \text{otherwise.} \end{cases}$$

(iii) For any integer k, $1 \leqslant k \leqslant n$, the *threshold measure* τ_k is defined by

$$\tau_k(A) := \begin{cases} 1 & \text{if } |A| \geqslant k \\ 0 & \text{otherwise.} \end{cases}$$

The class of all capacities on X can be partitionned into the so-called k-additive capacities, $k = 1, \ldots, n$. We need some additional definitions to introduce this new concept (details can be found in [36, 39, 37]). For a given capacity μ on X, their *Möbius transform* $m^\mu : X \to \mathbb{R}$ and *interaction transform* $I^\mu : X \to \mathbb{R}$ are defined by

$$m^u(A) := \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B)$$
$$I^\mu(A) := \sum_{B \supseteq A} \frac{1}{b - a + 1} m^\mu(B)$$

for all $A \subseteq X$, with $a = |A|, b = |B|$. A remarkable particular case is $A = \{i\}$, $i \in X$, for the interaction transform, since we recover the well-known Shapley value, whose more familiar expression is

$$I^\mu(\{i\}) =: \phi_i(\mu) = \sum_{A \subseteq X \setminus i} \frac{(n-a-1)!a!}{n!} (\mu(A \cup i) - \mu(A)).$$

A capacity μ is *k-additive* for some $2 \leq k \leq n$ if $m^\mu(A) = 0$ for all A such that $|A| > k$, and there exists some $A \subseteq X$, $|A| = k$, such that $m^\mu(A) \neq 0$. Note that due to the definition of the interaction transform, the above definition can be equivalently written with I^μ instead of m^μ.

4.1 Choquet integral based aggregation functions

The Choquet integral was introduced in 1953 by Choquet [21]. It is a generalization of the Lebesgue integral where the (classical) measure is replaced by a capacity.

Definition 32. Let μ be a capacity on $X = \{1, \dots, n\}$ and $\mathbf{x} \in [0, \infty[^n$. The *Choquet integral* of $\mathbf{x}$ with respect to μ is defined by

$$\mathcal{C}_\mu(\mathbf{x}) := \sum_{i=1}^{n} (x_{\sigma(i)} - x_{\sigma(i-1)})\mu(A_{\sigma(i)})$$

with σ a permutation on $\{1, \dots, n\}$ such that $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant \cdots \leqslant x_{\sigma(n)}$, with the convention $x_{\sigma(0)} := 0$, and $A_{\sigma(i)} := \{\sigma(i), \dots, \sigma(n)\}$.

We will often use the more compact notation $x_{(i)}$ instead of $x_{\sigma(i)}$, which was already introduced in Section 3.

It is straightforward to see that an equivalent formula is

$$\mathcal{C}_\mu(\mathbf{x}) = \sum_{i=1}^{n} x_{(i)} \big(\mu(A_{(i)}) - \mu(A_{(i+1)})\big),$$

with $A_{(n+1)} := \varnothing$. It can be shown that $\mathcal{C}_\mu$ is a continuous aggregation function, for any capacity μ. However, if $\mathcal{C}_\mu$ is defined on $\mathbb{I}$ which is a bounded interval $[0, a]$, it is necessary that the capacity is normalized, otherwise the bounds of interval will not be preserved.

If $\mathbf{x} \in \mathbb{R}^n$, there exist two ways of defining $\mathcal{C}_\mu(\mathbf{x})$, according to how the symmetry is done. The usual one is the following:

$$\mathcal{C}_\mu(\mathbf{x}) = \mathcal{C}_\mu(\mathbf{x}^+) - \mathcal{C}_{\mu^d}(\mathbf{x}^-)$$

for all $\mathbf{x} \in \mathbb{R}^n$, where $x_i^+ := \mathbf{x}_i \vee 0$ for all $i \in X$, and $\mathbf{x}^- := (-\mathbf{x})^+$. For the sake of concision, we do not detail further this topic here and refer the reader to [38]. The Choquet integral with $\mathbb{I} = \mathbb{R}$ has a very simple form in terms of the Möbius transform:

$$\mathcal{C}_\mu(\mathbf{x}) = \sum_{A \subseteq X} \Big(m^\mu(A) \cdot \bigwedge_{i \in A} x_i\Big).$$

A last remark before studying properties of the Choquet integral is that this aggregation function when $\mathbb{I} = [0, 1]$ can be obtained as the unique linear parcimonious interpolation on the vertices of the hypercube $[0, 1]^n$. This remarkable property comes from the fact that for any capacity μ, we have $\mathcal{C}_\mu(\mathbf{1}_A) = \mu(A)$ for any subset $A \subseteq X$, where $\mathbf{1}_A$ is the characteristic vector of A. Indeed, vertices of $[0, 1]^n$ correspond to the vectors $\mathbf{1}_A$, $A \subseteq X$, and for a given $\mathbf{x} \in [0, 1]^n$ which is not a vertex, the interpolation is done with the vertices of the canonical simplex

$$[0, 1]^n_\sigma := \{\mathbf{x} \in [0, 1]^n \mid x_{\sigma(1)} \leqslant \cdots \leqslant x_{\sigma(n)}\}$$

with σ any permutation such that $\mathbf{x} \in [0, 1]^n_\sigma$.

Definition 33. Let $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$. We say that $\mathbf{x}, \mathbf{x}'$ are *comonotonic* if there exists a permutation σ on $\{1, \dots, n\}$ such that $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant \cdots \leqslant x_{\sigma(n)}$ and $x'_{\sigma(1)} \leqslant x'_{\sigma(2)} \leqslant \cdots \leqslant x'_{\sigma(n)}$ (equivalently if there is no $i, j \in \{1, \dots, n\}$ such that $x_i > x_j$ and $x'_i < x'_j$).

Proposition 9. The Choquet integral satisfies the following properties:

(i) The Choquet integral is linear with respect to the capacity: for any capacities μ_1, μ_2 on X, any $\lambda_1, \lambda_2 \geqslant 0$,

$$\mathcal{C}_{\lambda_1\mu_1+\lambda_2\mu_2} = \lambda_1\mathcal{C}_{\mu_1} + \lambda_2\mathcal{C}_{\mu_2}.$$

If one restricts to normalized capacities, then the condition $\lambda_1 + \lambda_2 = 1$ is needed.

(ii) The Choquet integral satisfies *comonotonic additivity*, i.e., for any comonotonic vectors $\mathbf{x}, \mathbf{x}' \in [0, \infty[^n$, and any capacity μ,

$$\mathcal{C}_\mu(\mathbf{x} + \mathbf{x}') = \mathcal{C}_\mu(\mathbf{x}) + \mathcal{C}_\mu(\mathbf{x}').$$

(iii) Let μ, μ' be two capacities on X. Then $\mu \leqslant \mu'$ if and only if $\mathcal{C}_\mu \leqslant \mathcal{C}_{\mu'}$.

(iv) If μ is a 0-1 capacity, then

$$\mathcal{C}_\mu(\mathbf{x}) = \bigvee_{\substack{A\subseteq\{1,\dots,n\}\\ \mu(A)=1}} \bigwedge_{i\in A} x_i, \quad \forall \mathbf{x} \in [0,1]^n,$$

.

(v) The Choquet integral $\mathcal{C}_\mu$ is symmetric if and only if μ is symmetric.

(vi) The Choquet integral on $\mathbb{R}^n$ is invariant to positive affine transformation (interval scale change), that is

$$\mathcal{C}_\mu(c\mathbf{x} + a\mathbf{1}_X) = c \cdot \mathcal{C}_\mu(\mathbf{x}) + a$$

for any $c > 0$ and $a \in \mathbb{R}$.

(vii) For any capacity μ on X, we have $(\mathcal{C}_\mu)^d = \mathcal{C}_{\mu^d}$, i.e., the dual of the Choquet integral is the Choquet integral. w.r.t. its dual capacity.

The next theorem gives a characterization of the Choquet integral.

Theorem 9. Let $\mathsf{F} : \mathbb{R}^n \to \mathbb{R}$ be a given function. Then there exists a unique normalized capacity μ such that $\mathsf{F} = \mathcal{C}_\mu$ if and only if F fulfills the following properties:

(i) Comonotonic additivity;

(ii) Nondecreasing monotonicity;

(iii) $\mathsf{F}(\mathbf{1}_{\{1,\dots,n\}}) = 1$, $\mathsf{F}(\mathbf{0}) = 0$.

Moreover, μ is defined by $\mu(A) := \mathsf{F}(\mathbf{1}_A)$.

This result was shown by De Campos and Bolaños [24] in the case of $\mathbb{I} = [0, \infty[$, assuming in addition positive homogeneity, which can be deduced from (i) and (ii). The proof in the continuous case is due to Schmeidler [81].

We give the relation of the Choquet integral with other aggregation functions.

Proposition 10. Let μ be a normalized capacity and consider $\mathbb{I} = \mathbb{R}$. The following holds

(i) $\mathcal{C}_\mu = \mathsf{Min}$ if and only if $\mu = \mu_{\min}$.

(ii) $\mathcal{C}_\mu = \mathsf{Max}$ if and only if $\mu = \mu_{\max}$.

(iii) $\mathcal{C}_\mu = \mathsf{OS}_k$ (k-order statistics) if and only if μ is the threshold measure τ_{n-k+1}.

(iv) $\mathcal{C}_\mu = \mathsf{P}_k$ (k-th projection) if and only if μ is the Dirac measure δ_k.

(v) $\mathcal{C}_\mu = \mathsf{WAM}_\mathbf{w} = \check{\mathcal{C}}_\mu$ if and only if μ is additive, with $\mu(\{i\}) = w_i$, $\forall i \in \{1, \dots, n\}$.

(vi) $\mathcal{C}_\mu = \mathsf{OWA}_\mathbf{w}$ if and only if μ is symmetric, with $w_i = \mu(A_{n-i+1}) - \mu(A_{n-i})$, $i = 2, \dots, n$, and $w_1 = 1 - \sum_{i=2}^n w_i$, where A_i is any subset of X with $|A_i| = i$ (equivalently, $\mu(A) = \sum_{j=0}^{i-1} w_{n-j}$, $\forall A, |A| = i$).

Let us come back to k-additive capacities. The Choquet integral has an interesting expression in terms of the interaction transform when the capacity is 2-additive:

$$\mathcal{C}_\mu(\mathbf{x}) = \sum_{\substack{i,j \in X \\ I_{ij} > 0}} (x_i \wedge x_j) I_{ij} + \sum_{\substack{i,j \in X \\ I_{ij} < 0}} (x_i \vee x_j) |I_{ij}| + \sum_{i \in X} \Big(\phi_i x_i - \frac{1}{2} \sum_{j \neq i} |I_{ij}| \Big),$$

with $I_{ij} := I^\mu(\{i, j, \})$ and $\phi_i := I^\mu(\{i\})$. It is important to note that the above expression is a convex sum of at most $\frac{n(n+1)}{2}$ terms, grouped in three parts: a conjunctive one, a disjunctive one and an additive one.

We end this section by considering multilevel Choquet integrals. An interesting question is the following: what new aggregation function family do we obtain by combining the output of several different Choquet integrals by another Choquet integral, possibly iterating this process on several levels? For example, we may consider a 3-level aggregation of $\mathbf{x} \in \mathbb{R}^3$:

$$\mathsf{A}(\mathbf{x}) := \mathcal{C}_{\mu_3}(\mathcal{C}_{\mu_{11}}(x_1, x_2), \mathcal{C}_{\mu_2}(\mathcal{C}_{\mu_{11}}(x_1, x_2), \mathcal{C}_{\mu_{12}}(x_1, x_2, x_3)), x_3).$$

This question has been solved by Murofushi and Narukawa [72], and it is based on a result by Ovchinnikov on piecewise linear functions [73]. The answer is simply that one gets nothing new after the second level.

First we give a formal definition of a multilevel Choquet integral.

Definition 34. Let $\Gamma \subseteq \mathbb{R}^n$. For any $i \in [n]$, the projection P_i is a *0-level Choquet integral*. Let us consider $\mathsf{F}_i : \Gamma \to \mathbb{R}$, $i \in M := \{1, \dots, m\}$, being k_i-level Choquet integrals, and a capacity μ on M. Then

$$\mathsf{F}(\mathbf{x}) := \mathcal{C}_\mu(\mathsf{F}_1(\mathbf{x}), \dots, \mathsf{F}_m(\mathbf{x}))$$

is a *k-level Choquet integral*, with $k := \mathsf{Max}(k_1, \dots, k_m) + 1$. A *multilevel Choquet integral* is a function that is a k-level Choquet integral for some integer $k > 1$.

The result is the following.

Theorem 10. Let $\Gamma \subseteq \mathbb{R}^n$ be a convex closed n-dimensional set, and $\mathsf{F} : \Gamma \to \mathbb{R}$. The following are equivalent.

(i) F is a multilevel Choquet integral.

(ii) F is a 2-level Choquet integral, with all capacities of the first level being additive, and the capacity of the second level being 0-1 valued.

(iii) F is nondecreasing, positively homogeneous, and continuous piecewise linear.

4.2 Sugeno integral based aggregation functions

The Sugeno integral was introduced by Sugeno in 1972 [84], independently of the work of Choquet. Yet, there are striking similarities between the two definitions, up to the fact that the usual arithmetic operations of the Choquet integral are replaced by the lattice operations $\vee, \wedge$. However, the introduction of these two lattices operations implies many fundamental differences in their properties, and makes the Sugeno integral very close to the lattice polynomial functions.

We will see that the definition of the Sugeno integral makes sense only if $\mathbb{I} = [0, \mu(X)]$.

Definition 35. Let μ be a capacity on $X = \{1, \ldots, n\}$, and $\mathbf{x} \in [0, \mu(X)]^n$. The *Sugeno integral* of $\mathbf{x}$ with respect to μ is defined by

$$\mathcal{S}_\mu(\mathbf{x}) := \bigvee_{i=1}^{n} \big(x_{\sigma(i)} \wedge \mu(A_{\sigma(i)})\big)$$

with the same notations as in Definition 32. It can be shown that the Sugeno integral is a continuous aggregation function. The condition $\mathbb{I} = [0, \mu(X)]$ is necessary to fulfill the boundary conditions. As for the Choquet integral, we have $\mathcal{S}_\mu(\mathbf{1}_A) = \mu(A)$ for all $A \subseteq X$.

It can be shown that two other equivalent expressions are:

$$\mathcal{S}_\mu(\mathbf{x}) = \bigwedge_{i=1}^{n} \big(x_{\sigma(i)} \vee \mu(A_{\sigma(i+1)})\big) \tag{15}$$

$$= \mathsf{Med}(x_1, \ldots, x_n, \mu(A_{\sigma(2)}), \ldots, \mu(A_{\sigma(n)})), \tag{16}$$

with $A_{\sigma(n+1)} := \varnothing$, and Med is the classical median function.

The Sugeno integral has a close relation with weighted lattice polynomial functions. They are inductively defined as follows: (i) for any $k \in X$ and any $c \in \mathbb{I}$, the projection P_k and the constant function c are weighted lattice polynomial functions; (ii) if p, q are weighted lattice polynomial functions, then $p \vee q$ and $p \wedge q$ are weighted lattice polynomial functions; every weighted lattice polynomial function is formed by finitely many applications of rules (i) and (ii).

Each weighted lattice polynomial function $p : \mathbb{I}^n \to \mathbb{I}$ can be written both in conjunctive and disjunctive normal forms [59]:

$$p(\mathbf{x}) = \bigvee_{A \subseteq X} \Big(\alpha(A) \wedge \bigwedge_{i \in A} x_i\Big) = \bigwedge_{A \subseteq X} \Big(\beta(A) \vee \bigvee_{i \in A} x_i\Big), \tag{17}$$

where $\alpha, \beta : 2^X \to \mathbb{I}$ are some set functions. Based on (17), we have the next representation of the Sugeno integral.

Proposition 11. For any $\mathbf{x} \in [0, \mu(X)]^n$ and any capacity μ on X, the Sugeno integral of $\mathbf{x}$ with respect to μ can be written as

$$\mathcal{S}_\mu(\mathbf{x}) = \bigvee_{A \subseteq X} \left(\bigwedge_{i \in A} x_i \wedge \mu(A) \right)$$

$$\mathcal{S}_\mu(\mathbf{x}) = \bigwedge_{A \subseteq X} \left(\bigvee_{i \in A} x_i \vee \mu(X \setminus A) \right).$$

The following result gives the exact relation between weighted lattice polynomial functions and the Sugeno integral.

Theorem 11. Let $\mathsf{F} : [0,1]^n \to [0,1]$ be a function. The following assertions are equivalent.

(i) There exists a unique normalized capacity μ such that $\mathsf{F} = \mathcal{S}_\mu$;

(ii) F is an idempotent weighted lattice polynomial function;

(iii) F is an endpoint-preserving weighted lattice polynomial function.

A consequence of this theorem is that the multilevel Sugeno integral is the Sugeno integral: one does not get anything new by combining Sugeno integrals.

The Sugeno integral possesses several particular properties given in the next propositions.

Proposition 12. The Sugeno integral satisfies the following properties:

(i) The Sugeno integral commutes with max-min combinations of capacities: for any nonnegative games μ_1, μ_2 on X, any $\lambda_1, \lambda_2 \in [0, \infty[$,

$$\mathcal{S}_{(\lambda_1 \wedge \mu_1) \vee (\lambda_2 \wedge \mu_2)} = (\lambda_1 \wedge \mathcal{S}_{\mu_1}) \vee (\lambda_2 \wedge \mathcal{S}_{\mu_2})$$

$$\mathcal{S}_{(\lambda_1 \vee \mu_1) \wedge (\lambda_2 \vee \mu_2)} = (\lambda_1 \vee \mathcal{S}_{\mu_1}) \wedge (\lambda_2 \vee \mathcal{S}_{\mu_2}),$$

with the convention $((\lambda_1 \vee \mu_1) \wedge (\lambda_2 \vee \mu_2))(\varnothing) := 0$.

(ii) The Sugeno integral satisfies *comonotonic maxitivity* and *comonotonic minitivity*: for any comonotonic vectors $\mathbf{x}, \mathbf{x}' \in [0, \mu(X)]^n$, and any capacity μ,

$$\mathcal{S}_\mu(\mathbf{x} \vee \mathbf{x}') = \mathcal{S}_\mu(\mathbf{x}) \vee \mathcal{S}_\mu(\mathbf{x}')$$

$$\mathcal{S}_\mu(\mathbf{x} \wedge \mathbf{x}') = \mathcal{S}_\mu(\mathbf{x}) \wedge \mathcal{S}_\mu(\mathbf{x}').$$

(iii) Let μ, μ' be two nonnegative games on X. Then $\mu \leqslant \mu'$ if and only if $\mathcal{S}_\mu \leqslant \mathcal{S}_{\mu'}$.

(iv) $\mathcal{C}_\mu = \mathcal{S}_\mu$ if and only if μ is a $0-1$ capacity.

(v) For any normalized capacity μ and any $\mathbf{x} \in [0,1]^n$, $|\mathcal{C}_\mu(\mathbf{x}) - \mathcal{S}_\mu(\mathbf{x})| \leqslant \frac{1}{4}$.

(vi) For any capacity μ on X, we have $(\mathcal{S}_\mu)^d = \mathcal{S}_{\mu^d}$.

Proposition 13. The Sugeno integral $\mathcal{S}_\mu$ satisfies the following properties:

(i) Symmetry (or neutrality, commutativity) if and only if μ is symmetric;

(ii) Additivity if and only if μ is a 0-1 additive capacity (Dirac measure);

(iii) Maxitivity if and only if μ is a maxitive capacity (possibility measure);

(iv) Minitivity if and only if μ is a minitive capacity (necessity measure).

The next characterization is due to Marichal [57]. Still others can be found in this reference.

Theorem 12. Let $\mathsf{F} : [0,1]^n \to [0,1]$. Then there exists a capacity μ on X such that $\mathsf{F} = \mathcal{S}_\mu$ if and only if F satisfies (i) nondecreasingness, (ii) $\mathcal{S}_\mu(\alpha \vee \mathbf{x}) = \alpha \vee \mathcal{S}_\mu(\mathbf{x})$ ($\vee$-homogeneity), and (iii) $\mathcal{S}_\mu(\alpha \wedge \mathbf{x}) = \alpha \wedge \mathcal{S}_\mu(\mathbf{x})$ ($\wedge$-homogeneity).

Now we show the relation of the Sugeno integral with other aggregation functions.

We begin by introducing several aggregation functions.

Definition 36. Let $\mathbf{w} \in [0,1]^n$ satisfying $\bigvee_{i=1}^n w_i = 1$. Then, for any $\mathbf{x} \in [0,1]^n$:

(i) The *weighted maximum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{WMax}_{\mathbf{w}}(\mathbf{x}) := \bigvee_{i=1}^{n} (w_i \wedge x_i), \quad \forall \mathbf{x} \in \mathbb{I}^n.$$

(ii) The *weighted minimum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{WMin}_{\mathbf{w}}(\mathbf{x}) := \bigwedge_{i=1}^{n} ((1 - w_i) \vee x_i), \quad \forall \mathbf{x} \in \mathbb{I}^n.$$

(iii) The *ordered weighted maximum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{OWMax}_{\mathbf{w}}(\mathbf{x}) := \bigvee_{i=1}^{n} (w_i \wedge x_{(i)}), \quad \forall \mathbf{x} \in \mathbb{I}^n,$$

with $x_{(1)} \leqslant \cdots \leqslant x_{(n)}$, and $w_1 \geqslant w_2 \geqslant \cdots \geqslant w_n$.

(iv) The *ordered weighted minimum* with respect to $\mathbf{w}$ is the aggregation function defined by

$$\mathsf{OWMin}_{\mathbf{w}}(\mathbf{x}) := \bigwedge_{i=1}^{n} ((1 - w_i) \vee x_{(i)}), \quad \forall \mathbf{x} \in \mathbb{I}^n.$$

with $x_{(1)} \leqslant \cdots \leqslant x_{(n)}$, and $w_1 \leqslant w_2 \leqslant \cdots \leqslant w_n$.

Proposition 14. Let μ be a capacity. The following holds.

(i) $\mathcal{S}_\mu = \mathsf{Min}$ if and only if $\mu = \mu_{\min}$.

(ii) $\mathcal{S}_\mu = \mathsf{Max}$ if and only if $\mu = \mu_{\max}$.

(iii) $\mathcal{S}_\mu = \mathsf{OS}_k$ if and only if μ is the threshold capacity τ_{n-k+1}.

(iv) $\mathcal{S}_\mu = \mathsf{P}_k$ if and only if μ is the Dirac measure δ_k.

(v) $\mathcal{S}_\mu = \mathsf{WMax}_{\mathbf{w}}$ if and only if μ is a normalized maxitive capacity, with $\mu(\{i\}) = w_i$, for all $i \in \{1, \ldots, n\}$.

(vi) $\mathcal{S}_\mu = \mathsf{WMin}_{\mathbf{w}}$ if and only if μ is a normalized minitive capacity, with $\mu(X \setminus \{i\}) = \mu(X) - w_i$, for all $i \in X$.

(vii) $\mathcal{S}_\mu = \mathsf{OWMax}_{\mathbf{w}}$ if and only if μ is a normalized symmetric capacity such that $\mu(A) = w_{n-|A|+1}$, for any $A \subseteq X$, $A \neq \varnothing$.

(viii) $\mathcal{S}_\mu = \mathsf{OWMin}_{\mathbf{w}}$ if and only if μ is a normalized symmetric capacity such that $\mu(A) = 1 - w_{n-|A|}$, for any $A \subsetneq X$.

(ix) The set of Sugeno integrals with respect to 0-1 capacities coincides with the set of lattice polynomial functions.

5 Concluding remarks

In this paper we have focused on internal aggregation functions. We have discussed some of their properties, some construction methods (nonadditive integrals, for example) and some representation theorems. Much more details can be found in monographs of Bullen [13], and in our monograph [38]. Some of internal, i.e., idempotent aggregation functions will be discussed in our subsequent paper due to construction methods discussed there. As a typical example recall the weighted median, which for integer weights can be defined straightforwardly as

$$\mathsf{Med}_{\mathbf{w}}(x_1, \ldots, x_n) = \mathsf{Med}(w_1 \cdot x_1, \ldots, w_n \cdot x_n),$$

while for real (nonnegative) weights it is linked to the minimization problem of expression $\sum_{i=1}^{n} w_1 |x_i - r|$.

In the nonadditive integral domain we have discussed only the Choquet and the Sugeno integrals, though there are several other types of nonadditive integrals which may be interested (see, e.g., section 5.6 in [38]). Without going deeper into details, we recall briefly two of them, based on t-conorms, uninorms and copulas (these particular aggregation functions on $[0,1]$ are discussed in our subsequent paper). The (S, U)-integral based on continuous t-conorm S and a uninorm U satisfying the restricted distributivity relation

$$\mathsf{U}(x, \mathsf{S}(y, z)) = \mathsf{S}(\mathsf{U}(x, y), \mathsf{U}(x, z))$$

for all $x, y, z \in [0,1]$ such that $\mathsf{S}(y, z) < 1$ is given, for any $\mathbf{x} \in [0,1]^n$ and S-additive fuzzy measure μ on X, $\mu(A \cup B) = \mathsf{S}(\mu(A), \mu(B))$ whenever A and B are disjoint, by

$$(\mathsf{S}, \mathsf{U})_\mu(\mathbf{x}) = \mathsf{S}_{j=1}^{k} \left(\mathsf{S}_{i=1}^{m} (\mathsf{U}(x_i, \mu(\{i\}))) \right).$$

As an example take $\mathsf{S} = \mathsf{S}_L$ the bounded sum and $\mathsf{U} = \Pi$ the product. Then $(\mathsf{S}, \mathsf{U})_\mu$ coincide with the Weber integral [88]. Any S_L-additive fuzzy measure on X is determined by the values of μ on singletons, $w_i = \mu(\{i\})$, and $\mu(A) = \min\left(\sum_{i \in A} w_i, 1\right)$. Note that necessarily $\sum_{i=1}^{n} w_i \geqslant 1$, and

$$(\mathsf{S}, \mathsf{U})_\mu(\mathbf{x}) = \min\left(\sum_{i=1}^{n} w_i x_i, 1\right).$$

Obviously, if $\sum_{i=1}^{n} w_i = 1$, the weighted arithmetic mean $\mathsf{WAM}_{\mathbf{w}}$ is recovered.

Copulas C are linked to the probability measures P_{C} on $([0,1]^2, \mathcal{B}([0,1]^2)$ with uniform marginals, $\mathsf{C}(x,y) = P_{\mathsf{C}}([0,x] \times [0,y])$. A copula based integral (see Imaoka [44] for special copulas and Klement at al. [46] for general copulas) is given for any $\mathbf{x} \in [0,1]^n$ and any fuzzy measure μ on X, by two equivalent formulas

$$\begin{aligned} I_{\mathsf{C}}(\mathbf{x},\mu) &= \sum_{i=1}^{n} \big(\mathsf{C}(x_{\sigma(i)}, \mu(A_{\sigma(i)})) - \mathsf{C}(x_{\sigma(i-1)}, \mu(A_{\sigma(i)}))\big) \\ &= \sum_{i=1}^{n} \big(\mathsf{C}(x_{\sigma(i)}, \mu(A_{\sigma(i)})) - \mathsf{C}(x_{\sigma(i)}, \mu(A_{\sigma(i+1)}))\big), \end{aligned}$$

with the convention $A_{\sigma(n+1)} = \varnothing$, and $x_{\sigma(0)} = 0$. Observe that $I_{\Pi}(\cdot,\mu) = \mathcal{C}_\mu$ is the Choquet integral, while $I_{\mathsf{Min}}(\cdot,\mu) = \mathcal{S}_\mu$ is the Sugeno integral.

In a consequent paper we discuss conjunctive and disjunctive aggregation functions, and also some mixed aggregation functions related to both conjunctive and disjunctive aggregation functions. Moreover, several construction methods for aggregation functions will be introduced.

Acknowledgment

The authors gratefully acknowledge the support of their respective institutions during the long period of writing of the manuscript. During this period, we benefitted from the support of various projects which facilitated communication and collaboration among us, in particular the bilateral project between France and Serbia *Pavle Savić* N° 11092SF supported by EGIDE, Paris, and bilateral project between Slovakia and Serbia SK-SRB-19, the internal research project supported by the University of Luxembourg "Mathematics Research in Decision Making and Operations Research", F1R-MTH-PUL-09MRDO,, project APVV-0012-07 supported by the Slovak grant agency, the national grants MNTRS (Serbia, Project 144012), Provincial Secretariat for Science and Technological Development of Vojvodina. The article has drawn heavily upon material from Michel Grabisch, Jean-Luc Marichal, Radko Mesiar, Endre Pap, Aggregation Functions, ©M. Grabisch, J.-L. Marichal, R. Mesiar and E. Pap 2009, published by Cambridge University Press, extracts reproduced with permission.

References

[1] J. Aczél. On mean values. *Bull. Amer. Math. Soc.*, 54:392–400, 1948.

[2] J. Aczél. *Lectures on functional equations and their applications.* Mathematics in Science and Engineering, Vol. 19. Academic Press, New York, 1966. Translated by Scripta Technica, Inc. Supplemented by the author. Edited by Hansjorg Oser.

[3] J. Aczél and J. Dhombres. *Functional equations in several variables*, volume 31 of *Encyclopedia of Mathematics and its Applications.* Cambridge

University Press, Cambridge, 1989. With applications to mathematics, information theory and to the natural and social sciences.

[4] J. Aczél and M. Kuczma. On two mean value properties and functional equations associated with them. *Aequationes Math.*, 38(2-3):216–235, 1989.

[5] C. Antoine. *Les moyennes*, volume 3383 of *Que Sais-Je? [What Do I Know?]*. Presses Universitaires de France, Paris, 1998.

[6] M. Bajraktarević. Sur une équation fonctionnelle aux valeurs moyennes. *Glasnik Mat.-Fiz. Astronom. Društvo Mat. Fiz. Hrvatske. Ser. II*, 13:243–248, 1958.

[7] G. Beliakov, A. Pradera, and T. Calvo. *Aggregation Functions: A Guide for Practitioners*. Studies in Fuziness and Soft Computing. Springer, Berlin, 2007.

[8] G. Bemporad. Sul principio della media aritmetica. (Italian). *Atti Accad. Naz. Lincei*, 3(6):87–91, 1926.

[9] L. R. Berrone. Decreasing sequences of means appearing from non-decreasing functions. *Publ. Math. Debrecen*, 55(1-2):53–72, 1999.

[10] L. R. Berrone and J. Moro. Lagrangian means. *Aequationes Math.*, 55(3):217–226, 1998.

[11] L. R. Berrone and J. Moro. On means generated through the Cauchy mean value theorem. *Aequationes Math.*, 60(1-2):1–14, 2000.

[12] R. Bhatia. *Positive definite matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2007.

[13] P. S. Bullen. *Handbook of means and their inequalities*, volume 560 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 2003.

[14] P. S. Bullen, D. S. Mitrinović, and P. M. Vasić. *Means and their inequalities*, volume 31 of *Mathematics and its Applications (East European Series)*. D. Reidel Publishing Co., Dordrecht, 1988. Translated and revised from the Serbo-Croatian.

[15] T. Calvo, A. Kolesárová, M. Komorníková, and R. Mesiar. Aggregation operators: properties, classes and construction methods. In *Aggregation operators: new trends and applications*, pages 3–104. Physica, Heidelberg, 2002.

[16] T. Calvo, R. Mesiar, and R. R. Yager. Quantitative weights and aggregation. *IEEE Trans. Fuzzy Syst.*, 12(1):62–69, 2004.

[17] A. L. Cauchy. *Cours d'analyse de l'Ecole Royale Polytechnique, Vol. I. Analyse algébrique*. Debure, Paris, 1821.

[18] C.-P. Chen and F. Qi. A new proof for monotonicity of the generalized weighted mean values. *Adv. Stud. Contemp. Math. (Kyungshang)*, 6(1):13–16, 2003.

[19] C.-P. Chen and F. Qi. New proofs of monotonicities of generalized weighted mean values. *Tamkang J. Math.*, 35(4):301–304, 2004.

[20] O. Chisini. Sul concetto di media. (Italian). *Periodico di matematiche*, 9(4):106–116, 1929.

[21] G. Choquet. Theory of capacities. *Ann. Inst. Fourier, Grenoble*, 5:131–295 (1955), 1953–1954.

[22] E. Czogała and J. Drewniak. Associative monotonic operations in fuzzy set theory. *Fuzzy Sets and Systems*, 12(3):249–269, 1984.

[23] H. A. David and H. N. Nagaraja. *Order statistics*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2003.

[24] L. M. de Campos and M. J. Bolaños. Characterization and comparison of Sugeno and Choquet integrals. *Fuzzy Sets and Systems*, 52(1):61–67, 1992.

[25] B. de Finetti. Sul concetto di media. (Italian). *Giorn. Ist. Ital. Attuari*, 2(3):369–396, 1931.

[26] D. Dubois and H. Prade. Criteria aggregation and ranking of alternatives in the framework of fuzzy set theory. In *Fuzzy sets and decision analysis*, volume 20 of *Stud. Management Sci.*, pages 209–240. North-Holland, Amsterdam, 1984.

[27] F. Durante, S. Saminger-Platz, and P. Sarkoci. On representations of 2-increasing binary aggregation functions. *Information Sciences*, 178:23:4534–4541, 2008.

[28] A. G. Emrouznejad, A. Improving minimax disparity model to determine the OWA operator weights. *Information Sciences*, in Press,.

[29] J. C. Fodor. An extension of Fung-Fu's theorem. *Internat. J. Uncertain. Fuzziness Knowledge-Based Systems*, 4(3):235–243, 1996.

[30] J. C. Fodor, J.-L. Marichal, and M. Roubens. Characterization of some aggregation functions arising from MCDM problems. In *Fuzzy logic and soft computing*, pages 194–201. World Sci. Publishing, River Edge, NJ, 1995.

[31] J. C. Fodor and M. Roubens. *Fuzzy preference modelling and multicriteria decision support*. Kluwer, Dordrecht, 1994.

[32] V. Frosini. Averages. In *Italian Contributions to the Methodology of Statistics*, pages 1–17. Cleup, Padova, 1987.

[33] L. W. Fung and K. S. Fu. An axiomatic approach to rational decision making in a fuzzy environment. In *Fuzzy sets and their applications to cognitive and decision processes (Proc. U. S.-Japan Sem., Univ. Calif., Berkeley, Calif., 1974)*, pages 227–256. Academic Press, New York, 1975.

[34] D. Głazowska and J. Matkowski. An invariance of geometric mean with respect to Lagrangian means. *J. Math. Anal. Appl.*, 331(2):1187–1199, 2007.

[35] S. Gottwald. *A treatise on many-valued logics*, volume 9 of *Studies in Logic and Computation*. Research Studies Press Ltd., Baldock, 2001.

[36] M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92(2):167–189, 1997.

[37] M. Grabisch. The interaction and Möbius representations of fuzzy measures on finite spaces, k-additive measures: a survey. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals — Theory and Applications*, pages 70–93. Physica Verlag, 2000.

[38] M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation functions*. Encyclopedia of Mathematics and its Applications 127. Cambridge University Press, 2009.

[39] M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. *Mathematics of Operations Research*, 25(2):157–178, 2000.

[40] P. Hájek. *Metamathematics of Fuzzy Logic*. Kluwer Academic Publishers, Dordrecht, 1998.

[41] M. S. P. J. V. R. Hidders, J. On the expressibility of functions in xquery fragments. *Information Systems*, 33:4-5:435–455, 2008.

[42] A. Horwitz. Means and divided differences. *J. Math. Anal. Appl.*, 191(3):618–632, 1995.

[43] A. Horwitz. Invariant means. *J. Math. Anal. Appl.*, 270(2):499–518, 2002.

[44] H. Imaoka. A proposal of opposite-Sugeno integral and a uniform expression of fuzzy integrals. In *Int. Joint Conf. 4th IEEE Int. Conf. on Fuzzy Systems and 2nd Int. Fuzzy Engineering Symp.*, pages 583–590, Yokohama, March 1995.

[45] E. P. Klement, R. Mesiar, and E. Pap. *Triangular norms*, volume 8 of *Trends in Logic—Studia Logica Library*. Kluwer Academic Publishers, Dordrecht, 2000.

[46] E. P. Klement, R. Mesiar, and E. Pap. Measure-based aggregation operators. *Fuzzy Sets and Systems*, 142(1):3–14, 2004. Aggregation techniques.

[47] A. Kolesárová and M. Komorníková. Triangular norm-based iterative compensatory operators. *Fuzzy Sets and Systems*, 104(1):109–120, 1999.

[48] A. M. Kolesárová and R. G., Mesiar. Weighted ordinal means. *Information Sciences*, 177:18:3822–3830, 2007.

[49] A. N. Kolmogoroff. Sur la notion de la moyenne. (French). *Atti Accad. Naz. Lincei*, 12(6):388–391, 1930.

[50] M. Kuczma. On the quasiarithmetic mean in a mean value property and the associated functional equation. *Aequationes Math.*, 41(1):33–54, 1991.

[51] R. Lipschitz. De explicatione per series trigonometricas instuenda functionum unius variables arbitrariarum et praecipue earum, quae per variablis spatium finitum valorum maximorum et minimorum numerum habent infintum disquisitio. *J. Reine Angew. Math.*, 63:296–308, 1864.

[52] G. Lorenzen. Why means in two arguments are special. *Elem. Math.*, 49(1):32–37, 1994.

[53] L. Losonczi. Über eine neue Klasse von Mittelwerten. *Acta Sci. Math. (Szeged)*, 32:71–81, 1971.

[54] L. Losonczi. General inequalities for nonsymmetric means. *Aequationes Math.*, 9:221–235, 1973.

[55] L. Losonczi. Homogeneous Cauchy mean values. In *Functional equations—results and advances*, volume 3 of *Adv. Math. (Dordr.)*, pages 209–218. Kluwer Acad. Publ., Dordrecht, 2002.

[56] L. Losonczi. Equality of two variable means revisited. *Aequationes Math.*, 71(3):228–245, 2006.

[57] J.-L. Marichal. On Sugeno integral as an aggregation function. *Fuzzy Sets and Systems*, 114(3):347–365, 2000.

[58] J.-L. Marichal. On the associativity functional equation. *Fuzzy Sets and Systems*, 114(3):381–389, 2000.

[59] J.-L. Marichal. Weighted lattice polynomials. *Discrete Mathematics*, to appear.

[60] J.-L. Marichal and P. Mathonet. A characterization of the ordered weighted averaging functions based on the ordered bisymmetry property. *IEEE Trans. Fuzzy Syst.*, 7(1):93–96, 1999.

[61] J.-L. Marichal, P. Mathonet, and E. Tousset. Characterization of some aggregation functions stable for positive linear transformations. *Fuzzy Sets and Systems*, 102(2):293–314, 1999.

[62] J.-L. Marichal and M. Roubens. Characterization of some stable aggregation functions. In *Proc. 1st Int. Conf. on Industrial Engineering and Production Management (IEPM'93)*, pages 187–196, Mons, Belgium, June 1993.

[63] R. A. Marques Pereira and R. A. Ribeiro. Aggregation with generalized mixture operators using weighting functions. *Fuzzy Sets and Systems*, 137(1):43–58, 2003. Preference modelling and applications (Granada, 2001).

[64] J. Matkowski. Mean value property and associated functional equations. *Aequationes Math.*, 58(1-2):46–59, 1999. Dedicated to János Aczél on the occasion of his 75th birthday.

[65] J. Matkowski. On invariant generalized Beckenbach-Gini means. In *Functional equations—results and advances*, volume 3 of *Adv. Math. (Dordr.)*, pages 219–230. Kluwer Acad. Publ., Dordrecht, 2002.

[66] O. Mayor, G.and Valero. Aggregation of asymmetric distances in computer science. *Information Sciences*, 180:6:803–812, 2010.

[67] R. Mesiar and E. Pap. Aggregation of infinite sequences. *Information Sciences*, 178:18:3557–3564, 2008.

[68] R. Mesiar and J. Špirková. Weighted means and weighting functions. *Kybernetika (Prague)*, 42(2):151–160, 2006.

[69] R. Mesiar and J. Špirková. Weighted aggregation operators based on minimization. In *Proc. Int. Summer School on Aggregation Operators and their Applications (AGOP'2007)*, pages 203–206, Ghent, Belgium, July 2007.

[70] R. Mesiar, J. Špirkova, and L. Vavrikova. Weighted aggregation operators based on minimization. *Information Sciences*, 178:4:1133–1140, 2008.

[71] M. Nagumo. Über eine klasse der mittelwerte. (German). *Japanese Journ. of Math.*, 7:71–79, 1930.

[72] Y. Narukawa and T. Murofushi. The n-step Choquet integral on finite spaces. In *Proc. of the 9th Int. Conf. on Information Proc. and Management of Uncertainty in Knowledge-Based Systems*, pages 539–543, Annecy, France, July 2002.

[73] S. Ovchinnikov. Max-min representation of piecewise linear functions. *Beiträge Algebra Geom.*, 43(1):297–302, 2002.

[74] Z. Páles. On the characterization of quasi-arithmetic means with weight function. *Aequationes Math.*, 32(2-3):171–194, 1987.

[75] F. Qi. Generalized weighted mean values with two parameters. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 454(1978):2723–2732, 1998.

[76] F. Qi, J.-Q. Mei, D.-F. Xia, and S.-L. Xu. New proofs of weighted power mean inequalities and monotonicity for generalized weighted mean values. *Math. Inequal. Appl.*, 3(3):377–383, 2000.

[77] F. Qi and S.-Q. Zhang. Note on monotonicity of generalized weighted mean values. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 455(1989):3259–3260, 1999.

[78] J. J. Rotman. *An introduction to the theory of groups*, volume 148 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, fourth edition, 1995.

[79] P. K. Sahoo and T. Riedel. *Mean value theorems and functional equations*. World Scientific Publishing Co. Inc., River Edge, NJ, 1998.

[80] J. Sándor. On the identric and logarithmic means. *Aequationes Math.*, 40(2-3):261–270, 1990.

[81] D. Schmeidler. Integral representation without additivity. *Proc. Amer. Math. Soc.*, 97(2):255–261, 1986.

[82] J.-P. Soublin. Étude algébrique de la notion de moyenne. *J. Math. Pures Appl. (9)*, 50:53–264, 1971.

[83] K. B. Stolarsky. Generalizations of the logarithmic mean. *Math. Mag.*, 48:87–92, 1975.

[84] M. Sugeno. Fuzzy measures and fuzzy integrals. *Trans. S.I.C.E. (Keisoku Jidōseigyō Gakkai)*, 8(2):218–226, 1972. In Japanese.

[85] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, 1974.

[86] V. Torra and Y. Narukawa. *Modeling decisions: Information Fusion and Aggregation Operators*. Cognitive Technologies. Springer, 2007.

[87] A. H. Truck, I. A tool for aggregation with words. *Information Sciences*, 179:14:2317–2324, 2009.

[88] S. Weber. $\perp$-decomposable measures and integrals for Archimedean t-conorms $\perp$. *J. Math. Anal. Appl.*, 101(1):114–138, 1984.

[89] H. Weisberg. The distribution of linear combinations of order statistics from the uniform distributions. *Ann. Math. Statist.*, 42:704–709, 1971.

[90] J. Wimp. Quadrature with generalized means. *Amer. Math. Monthly*, 93(6):466–468, 1986.

[91] A. Witkowski. Monotonicity of generalized weighted mean values. *Colloq. Math.*, 99(2):203–206, 2004.

[92] Z. Xu. Choquet integrals of weighted intuitionistic fuzzy information. *Information Sciences*, 180:5:726–736, 2010.

[93] R. Yager. On the dispersion measure of owa operators. *Information Sciences*, 179:22:3908–3919, 2009.

[94] R. Yager and J. Kacprzyk, editors. *The Ordered Weighted Averaging operators*. Kluwer Academic Publishers, USA, 1997. Theory and Applications.

[95] R. R. Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Systems Man Cybernet.*, 18(1):183–190, 1988.

[96] R. R. Yager. The power average operator. *IEEE Trans. Systems Man Cybernet.*, 31(6):724–731, 2001.

[97] L. Zadeh. Is there a need for fuzzy logic? *Information Sciences: an International Journal archive*.

[98] L. Zadeh. Toward a generalized theory of uncertainty (GTU); An outline. *Information Sciences*, 172:1–40, 2005.

[99] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1979 edition.