UNIVERSITE DE LIEGE

FACULTE DES SCIENCES

AGGREGATION OPERATORS FOR MULTICRITERIA DECISION AID

JEAN-LUC MARICHAL

Dissertation submitted in fulfillment of the requirements
for the Doctorate Degree in Sciences

1998–1999

AGGREGATION OPERATORS FOR MULTICRITERIA DECISION AID

by

JEAN-LUC MARICHAL

University of Liège
Liège, Belgium

Dissertation submitted in fulfillment of the requirements
for the Doctorate Degree in Sciences

1998–1999

Jean-Luc MARICHAL

Department of Management, FEGSS
University of Liège
Boulevard du Rectorat 7 – B31
B-4000 Liège, Belgium

Email: jl.marichal@ulg.ac.be
URL: http://www.sig.egss.ulg.ac.be/marichal/

Jury members:

Y. CRAMA	Université de Liège Co-promoteur, membre du comité de thèse
D. DUBOIS	Université Paul Sabatier, Toulouse, France Rapporteur
M. GRABISCH	Thomson-CSF, LCR, Orsay, France Rapporteur, membre du comité de thèse
P. LAUBIN	Université de Liège Président de la Section de Mathématique
S. OVCHINNIKOV	San Francisco State University, U.S.A. Rapporteur
M. ROUBENS	Université de Liège Promoteur, membre du comité de thèse
J. SCHMETS	Université de Liège Président du jury

Contents

List of Tables

List of Figures

Introduction

In many domains we are faced with the problem of aggregating a collection of numerical readings to obtain a so-called mean or typical value. The main object of this dissertation deals with the aggregation procedures used in multicriteria decision making problems. In such problems, values to be aggregated are gathered in a score table and represent evaluations of alternatives according to various criteria. Aggregation operators are proposed to obtain a global score for each alternative taking into account the given criteria. These global scores are then exploited to establish a recommendation or prescription.

We do not intend to cover all the range of multicriteria decision problems, but merely to address the aggregation step. It is known that for most of the methodologies, an aggregation step exists, but the quantities to be aggregated may differ (mainly scores, degrees of satisfaction, preferences).

The main contribution of this thesis is the taking into consideration of interaction between criteria. Until recently, criteria were supposed to be independent and the aggregation operator which was often used was the weighted arithmetic mean, with all its well-known drawbacks. Such an operator is not suitable when interacting criteria are considered. However, this problem has been overcome by the contribution of fuzzy integrals, such as the Choquet and Sugeno integrals. We study in detail both of these families, as well as some others.

In Chapter 1 we give some formal definitions related to the problem of aggregation. In particular, the concept of aggregation operator (extended or not) is presented. As privileged examples, the averaging operators are studied.

Next, we introduce briefly the general background of multicriteria decision making, especially the aggregation phase with which we deal here.

In Chapter 2 we present a set of properties that could be required for an aggregation operator. Actually, in deciding on the form of the aggregation operator a number of properties should be associated with this operation. These properties are generally based on natural considerations corresponding to the idea of an aggregated value. Three categories of properties are presented. The first one contains some elementary properties such as the increasing monotonicity. The second one is devoted to stability properties related to the scales used to define the input values. The third one presents more technical conditions about the way of constructing the aggregated value.

Chapter 3 considers some particular families of aggregation operators, as well as axiomatic characterizations of these families on the basis of properties introduced in Chapter 2. We start by presenting the theory of quasi-arithmetic means, which is built from either the so-called bisymmetry equation or the decomposability condition. We extend this family by dropping the condition of strict monotonicity. We then present the theory of associative functions, in which we generalize some characterizations. We also study operators that are suitable for values defined on specific scale types, especially on interval scales.

In Chapter 4 we discuss the necessity to use the concept of fuzzy measure and integral to deal with aggregation of interactive criteria. Two main classes of fuzzy integrals are investigated and characterized: the Choquet and Sugeno integrals. Subfamilies of these integrals are also well studied, and the intersection of both families is described.

Chapter 5 is concerned with formal definitions of importance indices of criteria, as well as interaction indices between combinations of criteria. Such indices, like the Shapley interaction index, come from a connection with game theory and allow to have a best understanding of the interaction phenomena between criteria in multicriteria decision making problems. It is also shown that these interaction indices, as set functions, form equivalent representations of the fuzzy measure that weight the criteria.

In Chapter 6 we investigate the aggregation problem of interactive criteria when the input data (scores) are cardinal in nature. The use of the Choquet integral as an aggregation operator is justified in this context. In addition to the importance and interaction indices presented in Chapter 5, several concepts, such as veto and favor degrees and dispersion measure, are introduced to point out the behavioral properties of the Choquet integral and to facilitate the interpretation of the associated fuzzy measure. The inverse problem of identification of the fuzzy measure by means of such semantical considerations and learning data is also approached.

The difficult problem of aggregating values defined on ordinal scales is also discussed, especially through the use of the Sugeno integral, which can be viewed as the qualitative counterpart of the Choquet integral.

Chapter 7 deals with the problem of approximation of a set function by another one having a simpler form. Such an operation is also proposed for the particular case of fuzzy measures and Choquet integrals. The approximation of the Choquet integral by a weighted arithmetic mean is treated in detail.

Acknowledgements

In preparing this thesis, I have benefitted from the help of many people. I would especially like to express my gratitude to Professor Roubens for his guidance and counsel. His assistance was indispensable to the completion of this thesis.

I am also deeply indebted to Professor Crama for his valuable support as well as the priceless advice he gave me.

I also gratefully acknowledge Professor Dubois, Professor Grabisch, and Professor Ovchinnikov for proofreading the entire text and agreeing to report on it.

Finally, I must mention the invaluable help of my wife Pascale who gave up countless evenings and week-ends during the writing of this thesis.

Jean-Luc Marichal
November 1998
Liège, Belgium

Chapter 1

The aggregation problem

1.1 Basic definitions

Aggregation refers to the process of combining several numerical values into a single one, so that the final result of aggregation takes into account in a given manner all the individual values. Such an operation is used in many well-known disciplines such as statistical or economic measurement.

For instance, suppose that several individuals form quantifiable judgements either about a measure of an object (weight, length, area, height, volume, importance or other attributes) or about a ratio of two such measures (how much heavier, longer, larger, taller, more important, preferable, more meritorious etc. one object is than another). In order to reach a consensus on these judgements, classical synthesizing functions have been proposed: arithmetic mean, geometric mean, median and many others.

An issue of considerable interest in many areas is the aggregation of multiple criteria. In addition to the obvious case of decision making this problem arises in pattern recognition, information retrieval, expert systems, and neural networks.

In decision making, values to be aggregated are typically preference or satisfaction degrees. A preference degree tells to what extent an alternative a is preferred to an alternative b, and thus is a *relative* appraisal. By contrast, a satisfaction degree expresses to what extent a given alternative is satisfactory with respect to a given criterion. It is an *absolute* appraisal. We elaborate on this issue in Section 1.3.

We assume that the values to be aggregated belong to numerical scales, which can be of ordinal or cardinal type. On an ordinal scale, numbers have no other meaning that defining an order relation on the scale, and distances or differences between values cannot be interpreted. On a cardinal scale, distances between values are not quite arbitrary. Actually, there are several kinds of cardinal scales: on an interval scale, where the position of the zero is a matter of convention, values are defined up to a positive linear transformation, i.e. $\phi(x) = r\,x + s$, with $r > 0$ (e.g. temperatures expressed on the Celsius scale); on a ratio scale, where a true zero exists, values are defined up to a similarity transformation, i.e. $\phi(x) = r\,x$, with $r > 0$ (e.g. lengths expressed in inches). We will come back on these measurement aspects in Section 2.2.

Once values are defined we can aggregate them and obtain a new value. But this can be done in many different ways according to what is expected from the aggregation operation, what is the nature of the values to be aggregated, and what scale types have been used. Thus, for a given problem, any aggregation operator should not be used. In other terms, the use of a given aggregation operator should always be justified.

Now, let us introduce the concept of aggregation operator in a formal way. We make a distinction between aggregation operators having one definite number of arguments and extended aggregation operators defined for all number of arguments.

Let E, F be non-empty real intervals, finite or infinite. E denotes the definition set of the values to be aggregated, and F denotes the set of the possible results of the aggregation. Usually, E is a closed interval $[a, b]$, or an open interval $]a, b[$, or the real line $\mathbb{R}$ itself. We also denote by E° the *interior* of E, that is the corresponding open set.

Definition 1.1.1 An *aggregation operator* is a function $M^{(n)} : E^n \to F$, where $n \in \mathbb{N}_0$.

For example, the arithmetic mean as an aggregation operator is defined by

$$\mathrm{AM}^{(n)}(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

The integer n represents the number of values to be aggregated. When no confusion can arise, the aggregation operators will be written M instead of $M^{(n)}$.

Definition 1.1.2 An *extended aggregation operator* is a sequence $M = (M^{(n)})_{n \in \mathbb{N}_0}$, the n-th element of which is an aggregation operator $M^{(n)} : E^n \to F$.

For example, the arithmetic mean as an extended aggregation operator is the sequence $(\mathrm{AM}^{(n)})_{n \in \mathbb{N}_0}$.

Of course, an extended aggregation operator is a multi-dimensional operator, which can be viewed as a mapping

$$M : \bigcup_{n \geq 1} E^n \to F.$$

For any $n \in \mathbb{N}_0$ and any $x \in E^n$, we then have $M(x) = M^{(n)}(x)$.

We let $A_n(E, F)$ denote the set of all aggregation operators from E^n to F. Also, $A(E, F)$ denotes the set of all extended aggregation operators whose n-th element is in $A_n(E, F)$.

We will also use the notation N_n for the set $\{1, \ldots, n\}$. When no ambiguity can arise, we will write N instead of N_n.

In order to avoid heavy notations, we also introduce the following terminology. It will be used all along this dissertation.

- For all $k \in \mathbb{N}_0$ and all $x \in E$, we set $k \odot x := x, \ldots, x$ (k times). For instance,

  $$M(3 \odot x, 2 \odot y) = M(x, x, x, y, y).$$

- For all $S, T \subseteq N$, set difference of S and T is denoted by $S \setminus T$. Cardinality of sets $S, T, \ldots$ will be denoted whenever possible by corresponding lower cases $s, t, \ldots$, otherwise by the standard notation $|S|, |T|, \ldots$. Moreover, we will often omit braces for singletons, e.g. writing $v(i)$, $S \cup i$ instead of $v(\{i\})$, $S \cup \{i\}$. Also, for pairs, triples, we will write ij, ijk instead of $\{i, j\}$, $\{i, j, k\}$, as for example $S \cup ijk$.

- For any subset $S \subseteq N$, $e_S^{(n)}$ is the *characteristic vector* (or *incidence vector*) of S, i.e. the vector of $\{0, 1\}^n$ whose i-th component is 1 if and only if $i \in S$. We also introduce the complementary characteristic vector of $S \subseteq N$ by $\overline{e}_S^{(n)} = e_{N \setminus S}^{(n)}$.

  When there is no fear of ambiguity, the superscript (n) will be omitted.

- For any $n \in \mathbb{N}_0$, we let Π_n denote the set of all permutations of N_n. For any $\pi \in \Pi_n$ and any $S \subseteq N$, we set $\pi(S) := \{\pi(i) \,|\, i \in S\}$.
- $\wedge$, $\vee$ denote respectively the minimum and maximum operations.
- Given a vector $(x_1, \ldots, x_n)$ and a permutation $\pi \in \Pi_n$, the notation $[x_1, \ldots, x_n]_\pi$ means $x_{\pi(1)}, \ldots, x_{\pi(n)}$, that is, the permutation π of the indices. Moreover, we let $(\cdot)$ denote the particular permutation which arranges all the elements $x_1, \ldots, x_n$ by increasing values: that is, $x_{(1)} \leq \ldots \leq x_{(n)}$.

  The so-called *median* of an odd number of values $x_1, \ldots, x_{2k-1}$ is simply defined by
$$\mathrm{median}(x_1, \ldots, x_{2k-1}) := x_{(k)}.$$

We now give a small list of well-known aggregation operators. Special aggregation operators will be developed in subsequent chapters. In the definitions below, x will stand for $(x_1, \ldots, x_n)$.

- The *arithmetic mean* operator AM is defined by
$$\mathrm{AM}(x) = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{1.1}$$
- For any weight vector $\omega = (\omega_1, \ldots, \omega_n) \in [0,1]^n$ such that
$$\sum_{i=1}^{n} \omega_i = 1,$$
  the *weighted arithmetic mean* operator WAM_ω and the *ordered weighted averaging* operator OWA_ω associated to ω, are respectively defined by
$$\mathrm{WAM}_\omega(x) = \sum_{i=1}^{n} \omega_i\, x_i \tag{1.2}$$
$$\mathrm{OWA}_\omega(x) = \sum_{i=1}^{n} \omega_i\, x_{(i)}. \tag{1.3}$$
- For any $k \in N$, the *projection* operator P_k and the *order statistic* operator OS_k associated to the k-th argument, are respectively defined by
$$\mathrm{P}_k(x) = x_k \tag{1.4}$$
$$\mathrm{OS}_k(x) = x_{(k)}. \tag{1.5}$$
- For any non-empty subset $S \subseteq N$, the *partial minimum* operator $\min_S$ and the *partial maximum* operator $\max_S$ associated to S, are respectively defined by
$$\min{}_S(x) = \min_{i \in S} x_i \tag{1.6}$$
$$\max{}_S(x) = \max_{i \in S} x_i. \tag{1.7}$$

Since means are the most common aggregation operators, it is worth studying them in details. The next section deals with this issue.

1.2 The concept of mean

A considerable amount of literature about the concept of mean (or average) and the properties of several means (like the median, the arithmetic mean, the geometric mean, the root-power mean, the harmonic mean, etc.) has been already produced in the 19th century and has often treated the significance and the interpretation of these specific aggregation operators.

Cauchy [23] considered in 1821 the mean of n independent variables $x_1, \ldots, x_n$ as a function $M(x_1, \ldots, x_n)$ which should be internal to the set of x_i values:

$$\min\{x_1, \ldots, x_n\} \leq M(x_1, \ldots, x_n) \leq \max\{x_1, \ldots, x_n\}. \tag{1.8}$$

Such functions are also known in literature as *compensative* functions (see Section 2.1.5).

The concept of mean as a *numerical equalizer* is usually ascribed to Chisini [26], who gives in 1929 the following definition (p. 108):

> Let $y = g(x_1, \ldots, x_n)$ be a function of n independent variables $x_1, \ldots, x_n$ representing homogeneous quantities. A mean of $x_1, \ldots, x_n$ with respect to the function g 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,
> $$g(M, \ldots, M) = g(x_1, \ldots, x_n).$$

When g is considered as the sum, the product, the sum of squares, the sum of inverses, the sum of exponentials, or is proportional to $[(\sum_i x_i^2)/(\sum_i x_i)]^{1/2}$ as for the duration of oscillations of a composed pendulum of n elements of same weights, the solution of Chisini's equation corresponds respectively to the arithmetic mean, the geometric mean, the quadratic mean, the harmonic mean, the exponential mean and the antiharmonic mean, which is defined as

$$M(x_1, \ldots, x_n) = \Big(\sum_i x_i^2\Big)\Big/\Big(\sum_i x_i\Big).$$

Unfortunately, as noted by de Finetti [34, 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 the above equation) fulfils the Cauchy's internality property. The following quote from Ricci [150, p. 39] could be considered as another possible criticism to Chisini's view :

> ... when all values become equal, the mean equals any of them too. The inverse proposition is not true. If a function of several variables takes their common value when all variables coincide, this is not sufficient evidence for calling it a mean. For example, the function
> $$g(x_1, x_2, \ldots, x_n) = x_n + (x_n - x_1) + (x_n - x_2) + \cdots + (x_n - x_{n-1})$$
> equals x_n when $x_1 = \cdots = x_n$, but it is even greater than x_n as long as x_n is greater than every other variable.

In 1930, Kolmogoroff [107] and Nagumo [136] considered that the mean should be more than just a Cauchy mean or a numerical equalizer. They defined a *mean value* to be an infinite sequence of continuous, symmetric and strictly increasing (in each variable) real functions

$$M^{(1)}(x_1) = x_1, M^{(2)}(x_1, x_2), \ldots, M^{(n)}(x_1, \ldots, x_n), \ldots$$

satisfying the *idempotence* law: $M^{(n)}(x,\ldots,x) = x$ for all n and all x, and a certain kind of associative law:

$$M^{(k)}(x_1,\ldots,x_k) = x \;\Rightarrow\; M^{(n)}(x_1,\ldots,x_k,x_{k+1},\ldots,x_n) = M^{(n)}(x,\ldots,x,x_{k+1},\ldots,x_n) \quad (1.9)$$

for every natural integer $k \leq n$. They proved, independently of each other, that these conditions are necessary and sufficient for the quasi-arithmeticity of the mean, that is, for the existence of a continuous strictly monotonic function f such that $M^{(n)}$ may be written in the form

$$M^{(n)}(x_1,\ldots,x_n) = f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big] \qquad (n \in \mathbb{N}_0). \quad (1.10)$$

The *quasi-arithmetic means* (1.10) comprise most of the algebraic means of common use, and allow one to specify f in relation to operational conditioning (see Section 3.1.1). Some means however do not belong to this family: de Finetti [34, p. 380] observed that the antiharmonic mean is not increasing in each variable and that the median is not associative in the sense of (1.9).

The above properties defining a mean value seem to be natural enough. For instance, one can readily see that, for increasing means, the idempotence property is equivalent to Cauchy's internality (1.8), and both are accepted by all statisticians as requisites for means.

Note that Fodor and Marichal [67] generalized the Kolmogoroff-Nagumo's result by relaxing the condition that the means be strictly increasing, requiring only that they be increasing (see Section 3.2.2). The family obtained, which has a rather intricate structure, naturally includes the min and max operations.

Associativity of means (1.9) has been introduced first in 1926 by Bemporad [18, p. 87] in a characterization of the arithmetic mean. Under idempotence, this condition seems more natural, for it becomes equivalent to

$$\begin{gathered} M^{(k)}(x_1,\ldots,x_k) = M^{(k)}(x'_1,\ldots,x'_k) \\ \Downarrow \\ M^{(n)}(x_1,\ldots,x_k,x_{k+1},\ldots,x_n) = M^{(n)}(x'_1,\ldots,x'_k,x_{k+1},\ldots,x_n) \end{gathered}$$

which says that the mean does not change when altering some values without modifying their partial mean.

Notice that another definition of mean was given by Grabisch [79]: a mean operator (or averaging operator) is a symmetric, increasing and idempotent real function. As a consequence, such a function lies between min and max. Dubois and Prade [44, 46] requested also the continuity, and the fact that min and max are excluded from the family.

1.3 Aggregation in multicriteria decision making

As announced in the introduction, we study the aggregation problem in the framework of multicriteria decision making (MCDM). The purpose of this section is to present this setting in more details.

There are actually two main approaches of multicriteria decision making, namely multiattribute utility theory (see e.g. [105]), and the preference modelling approach (see e.g. [70]). In multiattribute utility theory, to each alternative is given an absolute score with respect to each criterion, and the global score, taking into account all the criteria, is obtained by aggregating all the partial scores. This is called the *cardinal approach*. By contrast, in preference modelling, a preference degree is assigned to every ordered pair of alternatives, with respect to each criterion. Then, a global preference degree is obtained by aggregating all the partial preference degrees. This is the *ordinal or relational approach*.

Whatever the approach to be taken, a necessary step is aggregation, and quantities to be aggregated are either scores or preference degrees. In this dissertation, we will often refer to the cardinal approach for the sake of simplicity, but this is not limitative.

1.3.1 Cardinal approach

Assume $A = \{a, b, c, \ldots\}$ is a non-empty set of *alternatives* or *acts*, among which the decision maker must choose. Such alternatives could represent possible solutions to a problem. Since our major concern here is the fomulation of aggregation operators rather than computational issues we shall assume that A is finite. Assume we have a collection of *goals* or *criteria* $N = \{1, \ldots, n\}$ we desire to satisfy. Each criterion i is represented by a mapping g_i from the set of alternatives A to a measurement scale $E_i \subseteq \mathbb{R}$. The value $g_i(a)$ is then called the *partial score* of a with respect to criterion i.

In most applications, it is assumed that each E_i is the unit interval[1] $[0, 1]$. In this case, $g_i(a)$ can be viewed as the degree to which the alternative a satisfies criterion i, or a degree of similarity between a and some ideal alternative according to criterion i. In the framework introduced by Bellman and Zadeh [17], the mapping g_i is viewed as the membership function of the fuzzy set [199] of alternatives that meet criterion i (see also Dubois and Prade [44] and Dubois and Koning [42]).

We do not address here the way of constructing the mappings g_i, and we suppose that the scores are given beforehand.

Of course the criteria have not always the same importance. It is then useful to define a weight ω_i associated to each criterion i. Such a weight represents the strength or importance of this criterion.

When criteria represent a panel of experts, the score $g_i(a)$ is then regarded as the rating of alternative a by expert i. Usually, all the experts have the same weight, but in certain applications each expert can have a coefficient of importance. Montero [127] then proposed to define the set of experts as a fuzzy set with membership function $\omega : N \to [0, 1]$. Given an individual $i \in N$, the value ω_i may be interpreted as the degree in which that individual is really a decision maker relative to the decision problem, or it can be viewed as the power (degree of importance, competence or ability) of his opinion.

Our central interest is the problem of constructing a single comprehensive criterion from the given criteria. The word *comprehensive* refers to the fact that the criterion resulting from combination is supposed to be representative of all the original criteria, and reflects the decision maker's attitude.

Formally, each alternative $a \in A$ can be assimilated with the vector of its partial scores, called *profile*[2]

$$a \in A \quad \longleftrightarrow \quad (g_1(a), \ldots, g_n(a)) \in E_1 \times \cdots \times E_n.$$

[1]This is not restrictive if we consider that scores are defined up to a positive linear transformation.

[2]Removing some alternatives, if necessary, we can assume that the profiles are all distinct.

From such a profile, one can compute a *global score* $M(g_1(a),\ldots,g_n(a))$ of the alternative a by means of an aggregation operator M which takes into account the weights associated to criteria.

Implicit in this formulation is the assumption that the global score is calculated in a pointwise manner, i.e. it only depends upon the evaluations of the g_i's at a. This assumption ensures us of satisfying the *independence of irrelevant alternatives condition* [14, 114]: modifying some profiles except that of alternative a does not alter the global score of a.

It is also assumed that the aggregated value only depends on individual scores and not on alternatives themselves. This means that the operator M is *neutral with respect to alternatives*:

$$g_i(a) = g_i(b),\ \forall i \quad \Rightarrow \quad M(g_1(a),\ldots,g_n(a)) = M(g_1(b),\ldots,g_n(b)).$$

Particularly, the same aggregation procedure should be used for each alternative.

For the sake of convenience, we denote the partial score $g_i(a)$ by x_i^a. Moreover, for any $S \subseteq N$, we set $E_S := \times_{i\in S} E_i$. Thus, an alternative a can be represented by a n-dimensional profile $x^a = (x_1^a,\ldots,x_n^a)$ in E_N. Table 1.1 shows a typical presentation of a multicriteria decision problem.

	criterion 1	$\cdots$	criterion n	global score
alternative a	x_1^a	$\cdots$	x_n^a	$M(x_1^a,\ldots,x_n^a)$
alternative b	x_1^b	$\cdots$	x_n^b	$M(x_1^b,\ldots,x_n^b)$
$\vdots$	$\vdots$		$\vdots$	$\vdots$

Table 1.1: Presentation of a multicriteria decison making problem

Once the global scores are computed, they can be used to rank the alternatives or select an alternative that best satisfies the given criteria. For instance the optimal alternative $a^* \in A$ could be selected such that

$$M(x^{a^*}) = \max_{a\in A} M(x^a).$$

To summarize, multicriteria decision making procedures consist of three main steps (phases) as follows.

1. *Modelling phase*

   In this phase we look for appropriate models for constructing the partial scores x_i^a and also for determining the importance of each criterion (i.e., the weights).

2. *Aggregation phase*

   In this step we try to find a unified (global) score for each alternative, on the basis of the partial scores and the weights.

3. *Exploitation phase*

   In this phase we transform the global information about the alternatives either into a complete ranking of the elements in A, or into a global choice of the best alternatives in A.

In the classical multiattribute utility (MAUT) model [62, 105], it is assumed that the preferences over A of the decision maker are expressed by a total preorder $\succeq$ (i.e., a strongly complete and transitive binary relation). Then the basic idea of MAUT consists in assuming the existence of a so-called utility function $u : A \to \mathbb{R}$ which represents $\succeq$, that is such that

$$a \succ b \iff u(a) > u(b). \tag{1.11}$$

Using such a preference model to establish a recommendation (choice of an alternative, ranking of all alternatives) in a decision aid study is straightforward and the main task of the analyst is to assess u. For this purpose, it is suggested to consider one-dimensional utility functions[3] $u_i : E_i \to \mathrm{I\!R}$ on each criterion/attribute, and then to aggregate them by a suitable operator M:

$$u(a) = M(u_1(x_i^a), \ldots, u_n(x_n^a)), \quad a \in A, \tag{1.12}$$

so that the function u so constructed verifies (1.11). The numbers $u_i(x_i^a)$ are real numbers, either positive or negative, but it is known that the u_i are defined up to a positive linear transformation. Thus we can consider that the u_i range in the unit interval $[0,1]$ without loss of generality. This is a necessary assumption when dealing with operators defined on $[0,1]^n$.

Under some well-known conditions (see e.g., Krantz *et al.* [108] or Wakker [186]), u can be obtained in an additive manner, that is

$$u(a) = \sum_{i=1}^{n} u_i(x_i^a), \quad a \in A. \tag{1.13}$$

In this case, modelling preferences amounts to assess the partial utilities u_i. Several techniques have been proposed to do so (see Keeney and Raiffa [105] or von Winterfeldt and Edwards [185]).

A very important notion in multicriteria decision making, closely related to the existence of an additive utility function, is that of preferential independence (see e.g. [186]). To introduce this concept, we note that through the natural identification of alternatives with their profiles in E_N, the preference relation $\succeq$ on A can be considered as a preference relation on E_N.

Definition 1.3.1 The subset S of criteria is said to be *preferentially independent* of $N \setminus S$ if, for all $x_S, y_S \in E_S$ and all $x_{N\setminus S}, z_{N\setminus S} \in E_{N\setminus S}$, we have

$$(x_S, x_{N\setminus S}) \succeq (y_S, x_{N\setminus S}) \iff (x_S, z_{N\setminus S}) \succeq (y_S, z_{N\setminus S}). \tag{1.14}$$

The whole set of criteria N is said to be *mutually preferentially independent* if S is preferentially independent of $N \setminus S$ for every $S \subseteq N$.

Roughly speaking, the preference of $(x_S, x_{N\setminus S})$ over $(y_S, x_{N\setminus S})$ is not influenced by the values $x_{N\setminus S}$. For some problems this principle might be violated as it can be seen in the following example.

Example 1.3.1 Let us consider a decision problem involving 4 cars, evaluated on 3 criteria: price, consumption and comfort.

	criterion 1 (price)	criterion 2 (consumption)	criterion 3 (comfort)
car 1	10.000 Euro	10 ℓ/100 km	very good
car 2	10.000 Euro	9 ℓ/100 km	good
car 3	30.000 Euro	10 ℓ/100 km	very good
car 4	30.000 Euro	9 ℓ/100 km	good

Suppose the consumer (decision maker) has the following preferences:

$$\text{car } 2 \succeq \text{car } 1 \qquad \text{and} \qquad \text{car } 3 \succeq \text{car } 4.$$

The reason may be that, as price increases, so does the importance of comfort. In this case, criteria 2 and 3 are not preferentially independent of criterion 1.

[3]These utility functions are used only to express the partial scores in a same measurement scale.

It is known [62, 165] that the mutual preferential independence among the criteria (1.14) is a necessary condition (but not sufficient) for a utility function to be additive, that is of the form (1.13). In other terms, in case of violation of this property, no additive utility function can model the preferences of the decision maker. We will see in Chapter 4 that the concept of fuzzy integral allows to overcome this problem.

1.3.2 Relational approach

The relational approach consists in *comparing* alternatives two by two, and expressing with a number the *degree of preference* of one alternative over the other, with respect to a criterion. These numbers are very often expressed by the help of fuzzy (valued) preference relations.

More formally, let A be a given set of alternatives and $R_1, R_2, \ldots, R_n$ be fuzzy binary relations on A representing n criteria. That is, for each criterion $i \in N$, R_i is a function from $A \times A$ to $[0,1]$ such that $R_i(a,b)$ reflects the degree to which a is declared to be not worse than b for criterion i. Thus $R_i(a,b)$ is a *relative* evaluation.

Such an approach has been developed essentially by Roy [157, 158] (ELECTRE methods) with ordinary crisp relations, and then by Blin [21], Saaty [159], Fodor and Roubens [69, 70] with fuzzy preference relations.

The *modelling phase* consists in looking for appropriate models for fuzzy monocriterion relation R_i and also for determining the weight ω_i of each criterion i.

All these preference relations are then aggregated to take into account all the criteria. We then look for an aggregation operator $M \in A_n([0,1], \mathbb{R})$ so that the global relation R, expressed by

$$R(a,b) = M(R_1(a,b), \ldots, R_n(a,b)), \quad a, b \in A,$$

reflects an overall opinion on pairs of alternatives[4] (*aggregation phase*).

This global relation R can be used to get a ranking of the alternatives, or to choose the set of the "best" alternatives (*exploitation phase*). Note that, in this approach, the transitivity property (in the usual sense, or in the max-min sense) is most often lost[5], so that the result is a partial ordering of the alternatives: some alternatives may be incomparable each other.

1.3.3 Equivalence classes of aggregation operators

When one is faced with the choice of an aggregation operator in a decision making problem, a fundamental question arises: what are the operators which lead to the same decision, i.e. to the same ranking of the alternatives? This question was addressed and solved by Grabisch [78].

Definition 1.3.2 Two operators $M_1 \in A_n(E, F_1)$ and $M_2 \in A_n(E, F_2)$ are said to be *strongly equivalent* if

$$M_1(x) < M_1(x') \;\Leftrightarrow\; M_2(x) < M_2(x'), \quad x, x' \in E^n.$$

Notation: $M_1 \sim M_2$.

[4]Of course we suppose that the *condition of independence of irrelevant preferences* is satisfied: if a profile of fuzzy relations $(R_1, \ldots, R_n)$ is modified in such a way that individual's paired comparisons among a pair of alternatives (a,b) are unchanged —$(R_1(x,y), \ldots, R_n(x,y))$ becomes $(R_1'(x,y), \ldots, R_n'(x,y))$ for all (x,y) belonging to $A \times A$ except for $(x,y) = (a,b)$— the aggregation resulting from the original and modified profiles should be unchanged for the pair (a,b). Hence, $R(a,b)$ depends only on $R_1(a,b), \ldots, R_n(a,b)$ and is a function of n arguments for every pair (a,b) of $A \times A$.

[5]A condition ensuring min-transitivity is presented in Section 2.2.4 (Proposition 2.2.7). Moreover, preservation of T-transitivity, where T is a t-norm, has been studied by Fodor and Ovchinnikov [66].

It is clear that $\sim$ is an equivalence relation. It ensures that the operators in the same equivalence class lead to exactly the same decisions, i.e. same ranking of the alternatives and same set of unordered (undecidable) alternatives. Grabisch [78] proved the following result.

Theorem 1.3.1 *Consider $M_1 \in A_n(E, F_1)$ and $M_2 \in A_n(E, F_2)$. Then M_1 and M_2 are strongly equivalent if and only if there exists a unique increasing bijection $g : F_1 \to F_2$, such that*

$$M_2(x) = g(M_1(x)), \quad x \in E^n.$$

This shows that any aggregation operator M that represents a utility function (1.12) is defined up to an increasing bijection.

Chapter 2

Aggregation properties

If we want to obtain a reasonable or satisfactory aggregation, any aggregation operator should not be used. To eliminate the "undesirable" operators, we can adopt an axiomatic approach and impose that these operators fulfil some selected properties. Such properties can be dictated by the nature of the values to be aggregated. For example, in some multicriteria evaluation methods, the aim is to assess a global absolute score to an alternative given a set of partial scores with respect to different criteria. Clearly, it would be unnatural to give as global score a value which is lower than the lowest partial score, or greater than the highest score, so that only *compensative* aggregation operators are allowed. Now, if preference degrees coming from transitive (in some sense) relations are combined, it may be requested that the result of combination remains transitive. Another example concerns the aggregation of opinions in voting procedures. If, as usual, the voters are anonymous, the aggregation operator must be symmetric.

Notice also that all the properties defined for aggregation operators can be naturally adapted to extended aggregation operators. For instance, $M \in A(E, \mathbb{R})$ is said to be symmetric if, for all $n \in \mathbb{N}_0$, the n-th aggregation operator $M^{(n)} \in A_n(E, \mathbb{R})$ in the sequence is symmetric.

In this chapter we present some properties that could be desirable for the aggregation of criteria. Of course, all these properties are not required with the same strength, and do not pertain to the same purpose. Some of them are imperative conditions whose violation leads to obviously counterintuitive aggregation modes. Others are technical conditions that just facilitate the representation or the calculation of the aggregation function. There are also facultative conditions that naturally apply in special circumstances but are not to be universally accepted. Note that analytic properties, such as the differentiability condition, which has been employed to characterize the weighted arithmetic mean [4, Sect. 5.3], will not be considered here since most of them are not interpretable in the multicriteria decision framework.

2.1 Elementary mathematical properties

2.1.1 Symmetry

Definition 2.1.1 (Sy) *Symmetry, commutativity, neutrality, anonymity:* $M \in A_n(E, \mathbb{R})$ is a symmetric operator if, for all $\pi \in \Pi_n$ and all $x \in E^n$, we have

$$M(x_1, \ldots, x_n) = M(x_{\pi(1)}, \ldots, x_{\pi(n)}).$$

The symmetry property (Sy) essentially implies that the indexing (ordering) of the arguments does not matter. This is required when combining criteria of equal importance or anonymous expert's opinions[1]; indeed, a symmetric operator is independent of the labels. Moreover, whatever the order in which the information is collected, the result will always be the same.

Notice that any symmetric operator is completely defined by means of compositions involving order statistics:

$$M(x_1, \ldots, x_n) = M(x_{(1)}, \ldots, x_{(n)}), \quad x \in E^n.$$

The following result, well known in the theory of groups, shows that the symmetry property can be checked with only two equalities, see e.g. Rotman [155, Exercise 2.9, p. 24].

Proposition 2.1.1 *$M \in A_n(E, \mathbb{R})$ fulfils (Sy) if and only if, for all $x \in E^n$, we have*
i) $M(x_2, x_1, x_3, \ldots, x_n) = M(x_1, x_2, x_3, \ldots, x_n)$
ii) $M(x_2, x_3, \ldots, x_n, x_1) = M(x_1, x_2, x_3, \ldots, x_n)$.

In situations when judges, criteria, or individual opinions are not equally important, the (Sy) property must be omitted. Previous research along this line was done by Cholewa [27] and Montero [126, 127]. They justified the weighted arithmetic mean as a general (not necessarily symmetric) aggregation rule.

2.1.2 Continuity

Definition 2.1.2 (Co) *Continuity:* $M \in A_n(E, \mathbb{R})$ is a continuous operator if it is a continuous function in the usual sense.

A continuous aggregation operator does not present any chaotic reaction to a small change of the arguments.

2.1.3 Monotonicity and such

Definition 2.1.3 (In) *Increasingness, monotonicity, non-decreasingness, non-negative responsiveness:* $M \in A_n(E, \mathbb{R})$ is increasing (in each argument) if, for all $x, x' \in E^n$, we have

$$x_i \leq x'_i \quad \forall i \in N \quad \Rightarrow \quad M(x) \leq M(x').$$

Definition 2.1.4 (SIn) *Strict increasingness, strict monotonicity, positive responsiveness:* $M \in A_n(E, \mathbb{R})$ is strictly increasing (in each argument) if, for all $x, x' \in E^n$, we have

$$x_i \leq x'_i \quad \forall i \in N, \text{ and } \exists j \in N \text{ such that } x_j < x'_j \quad \Rightarrow \quad M(x) < M(x').$$

An increasing aggregation operator presents a non-negative response to any increase of the arguments. In other terms, increasing a partial value cannot decrease the result. This operator is strictly increasing if, moreover, it presents a positive reaction to any increase of at least one partial value.

Definition 2.1.5 (UIn) *Unanimous increasingness:* $M \in A_n(E, \mathbb{R})$ is unanimously increasing if, for all $x, x' \in E^n$, we have
i) $x_i \leq x'_i \quad \forall i \in N \quad \Rightarrow \quad M(x) \leq M(x')$
ii) $x_i < x'_i \quad \forall i \in N \quad \Rightarrow \quad M(x) < M(x')$.

A unanimously increasing operator is increasing and presents a positive response whenever all the arguments strictly increase. For instance, we observe that, on $[0,1]^n$, the maximum operator $M(x) = \max x_i$ fulfils (UIn) whereas the bounded sum $M(x) = \min(\sum_{i=1}^n x_i, 1)$ does not.

[1] Of course, symmetry is more natural in voting procedures than in multicriteria decision making.

2.1.4 Idempotence

In a variety of applications, it is desirable that the aggregation operators satisfy the *unanimity* property, i.e. if all x_i are identical, $M(x_1, \dots, x_n)$ restitutes the common value. This is notably the case for means.

Definition 2.1.6 (Id) *Idempotence, agreement, unanimity, identity, reflexivity:* $M \in A_n(E, \mathbb{R})$ is idempotent if, for all $x \in E$, we have

$$M(x, \dots, x) = x.$$

Definition 2.1.7 (WId) *Weak idempotence, boundary conditions:* $M \in A_n([a,b], \mathbb{R})$ is weakly idempotent if

$$M(a, \dots, a) = a \quad \text{and} \quad M(b, \dots, b) = b.$$

2.1.5 Location in the real line

Aggregation operators can be roughly divided into three classes, each possessing very distinct behavior: *conjunctive* operators, *disjunctive* operators and *compensative* operators.

Definition 2.1.8 (Conj) *Conjunctiveness:* $M \in A_n(E, \mathbb{R})$ is conjunctive if, for all $x \in E^n$, we have

$$M(x) \leq \min x_i.$$

Conjunctive operators combine values as if they were related by a logical "and" operator. That is, the result of combination can be high only if all the values are high. t-norms are the suitable functions (defined on $[0,1]^n$) for doing conjunctive aggregation (see Section 3.3.5). However, they generally do not satisfy properties which are often requested for multicriteria aggregation, as idempotence, compensativeness, scale invariance, etc.

Definition 2.1.9 (Disj) *Disjunctiveness:* $M \in A_n(E, \mathbb{R})$ is disjunctive if, for all $x \in E^n$, we have

$$\max x_i \leq M(x).$$

Disjunctive operators combine values as an "or" operator, so that the result of combination is high if at least one value is high. Such operators are in this sense dual of conjunctive operators. The most common disjunctive operators are t-conorms (defined on $[0,1]^n$). As t-norms, t-conorms do not possess suitable properties for criteria aggregation.

Definition 2.1.10 (Comp) *Compensativeness:* $M \in A_n(E, \mathbb{R})$ is compensative if, for all $x \in E^n$, we have

$$\min x_i \leq M(x) \leq \max x_i.$$

Between conjunctive and disjunctive operators, there is room for a third category, namely *compensative* or *compromise* aggregation operators. They are located between min and max, which are the bounds of the t-norm and t-conorm families. In this kind of operators, a bad (resp. good) score on one criterion can be compensated by a good (resp. bad) one on another criterion, so that the result of combination will be medium, see Figure 2.1. If we add the properties (Sy, In), then we get the family of *averaging operators*, where border functions min and max are usually excluded.

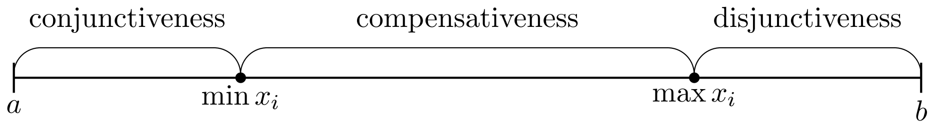


Figure 2.1: Location of M in $[a,b]$

Since E is connected, any compensative aggregation operator defined on E^n necessarily takes its values in E. Moreover, the following result presents an immediate link between (Comp) and (Id) that makes the latter all the more natural (see also de Finetti [34, p. 379]).

Proposition 2.1.2 *For every $M \in A_n(E, \mathbb{R})$, we have*

$$\text{(Comp)} \Rightarrow \text{(Id)} \tag{2.1}$$

$$\text{(In, Id)} \Rightarrow \text{(Comp)} \tag{2.2}$$

Regarding conjunctive and disjunctive two-place operators, we have the following.

Proposition 2.1.3 *Let $M \in A_2(E, \mathbb{R})$ fulfilling (In).*
i) If $b \in E$ is the right endpoint of E then M fulfils (Conj) if and only if

$$M(b,x) \leq x \quad \text{and} \quad M(x,b) \leq x \quad \forall x \in E.$$

ii) If $a \in E$ is the left endpoint of E then M fulfils (Disj) if and only if

$$M(a,x) \geq x \quad \text{and} \quad M(x,a) \geq x \quad \forall x \in E.$$

Proof. *i*) Let $x, y \in E$ with $x \leq y$. We simply have $M(x,y) \leq M(x,b) \leq x = \min(x,y)$ and $M(y,x) \leq M(b,x) \leq x = \min(x,y)$.
ii) Similar to *i*). ■

Conjunctive, disjunctive and compensative operators form a large disjoint covering of operators on $\mathbb{R}$ or $[a,b]$, but there are some operators which do not belong to one of these categories.

In an experimental study on the evaluation of tiles, Zimmermann and Zysno [202] pointed out the fact that human aggregation procedure is compensatory. Moreover, they showed that the arithmetic mean leads to a biased evaluation, because this operator does not take into account interaction between criteria. Zimmermann and Zysno have therefore proposed to mix the product and the probabilistic sum to a proportion $\gamma \in [0,1]$ in order to produce a kind of compensation between criteria:

$$M(x) = \Big(\prod_{i=1}^n x_i\Big)^{1-\gamma}\Big(1 - \prod_{i=1}^n (1-x_i)\Big)^{\gamma}, \quad x \in [0,1]^n.$$

This aggregation operator fulfils (Sy, Co, In) but not (Id). It covers a range from the product ($\gamma = 0$, conjunctive attitude) to the probabilistic sum ($\gamma = 1$, disjunctive attitude).

It is known that t-norms are very often used to model the conjunction in multiple-valued logic [12, 164]. But there are many situations in real life in which other functions, such as means, are taken. In this sense, a more general class than those of t-norms might be considered to model the conjunction.

On this issue, Trillas *et al.* [182] introduced a definition of the conjunction other than (Conj). According to them, $M \in A_2([0,1],[0,1])$ is said to be a *conjunction* if, for all $x, y \in [0,1]$,

$$x \leq y \ \Rightarrow \ \exists z \in [0,1] \text{ such that } M(y,z) = x.$$

In this sense, any continuous t-norm is a conjunction. Furthermore, the geometric mean is a conjunction whereas the arithmetic mean is not.

Similarly, $M \in A_2([0,1],[0,1])$ is said to be a *disjunction* if, for all $x, y \in [0,1]$,

$$x \leq y \quad \Rightarrow \quad \exists z \in [0,1] \text{ such that } M(x,z) = y.$$

2.2 Stability properties

As explained in Chapter 1, depending on the kind of scale which is used, allowed operations on values are restricted. For example, aggregation on ordinal scales should be limited to operations involving comparisons only, such as medians, and order statistics, while linear operations are allowed on interval scales.

To be precise, a *scale of measurement* is a mapping which assigns real numbers to objects being measured. Stevens [175, 176] defined the *scale type* of a scale by giving a class of *admissible transformations*, transformations that lead from one acceptable scale to another. For instance, we call a scale a *ratio scale* if the class of admissible transformations consists of the similarities $\phi(x) = r\,x$, $r > 0$. In this case, the scale value is determined up to choice of a unit. Mass is an example of a ratio scale. The transformation from Kilograms into pounds, for example, involves the admissible transformation $\phi(x) = 2.2\,x$. Length (inches, centimeters) and time intervals (years, seconds) are two other examples of ratio scales. We call a scale an *interval scale* if the class of admissible transformations consists of the positive linear transformations $\phi(x) = r\,x + s$, $r > 0$. The scale value is then determined up to choices of unit and zero point. Temperature (except where there is an absolute zero) defines an interval scale. Thus, transformation from Centigrade into Fahrenheit involves the admissible transformation $\phi(x) = (9/5)x + 32$. We call a scale an *ordinal scale* if the class of admissible transformations consists of the strictly increasing functions ϕ. Here the scale value is determined only up to order. For example, the scale of air quality being used in a number of cities is an ordinal scale. It assigns a number 1 to unhealthy air, 2 to unsatisfactory air, 3 to acceptable air, 4 to good air, and 5 to excellent air. We could just as well use the numbers 1, 7, 8, 15, 23, or the numbers 1.2, 6.5, 8.7, 205.6, 750, or any numbers that preserve the order. Definitions of other scale types can be found in the book by Roberts [151] on measurement theory, see also Roberts [152, 153].

It is clear that certain numerical statements involving measurements are meaningless in the sense that their truth value depends on which scale is employed. A classical exemple is the statement "The ratio of today's maximum and minimum temperatures is 1.14" which is "meaningless unless a particular representation, e.g., °C, is specified" [113, 179]. To give a second example, suppose that $x_1, \ldots, x_n$ are measured according to an ordinal scale, then the arithmetic mean comparison is meaningless. As illustration of this statement, let us consider the pairs of scores $(3,5)$ and $(1,8)$:

$$\frac{3+5}{2} < \frac{1+8}{2}$$

but according to the following admissible transformation ($\phi(1) = 1$, $\phi(3) = 4$, $\phi(5) = 7$, $\phi(8) = 8$),

$$\frac{4+7}{2} > \frac{1+8}{2} \;!!$$

A statement using scales of measurement is said to be *meaningful* if the truth or falsity of the statement is invariant when every scale is replaced by another acceptable version of it [151, p. 59]. For example, an aggregation operator is meaningful if the ranking of alternatives induced by the aggregation does not depend on scale transformation. This means that, when scores are defined according to an interval scale, using scores defined on a $[0, 100]$ scale or on $[-2, 3]$ scale has no influence on the ranking of alternatives.

In 1959, Luce [112] observed that the general form of a functional relationship between variables is greatly restricted if we know the scale type of the variables. These restrictions are discovered by formulating a functional equation from knowledge of the admissible transformations.

Luce's method is based on the principle, called the *principle of theory construction*, that an admissible transformation of the independent variables should lead to an admissible transformation of the dependent variable. For example, suppose that $f(a) = M(f_1(a), \ldots, f_n(a))$, where f is a ratio scale and $f_1, \ldots, f_n$ are all ratio scales, with the units chosen independently. Then, by the principle of theory construction, we get the functional equation

$$M(r_1\, x_1, \ldots, r_n\, x_n) = R(r_1, \ldots, r_n)\, M(x_1, \ldots, x_n),$$
$$r_i > 0, \quad R(r_1, \ldots, r_n) > 0.$$

Aczél, Roberts, and Rosenbaum [9] showed that the solutions of this equation are given by (see Theorem 3.4.2):

$$M(x_1, \ldots, x_n) = a \prod_{i=1}^{n} g_i(x_i), \quad \text{with} \quad a > 0, g_i > 0,$$

and

$$g_i(x_i\, y_i) = g_i(x_i)\, g_i(y_i).$$

In this section we present some functional equations related to certain scale types. The interested reader can find more details in [8, 9] and a good survey in [153].

2.2.1 Ratio, difference and interval scales

Definition 2.2.1 (SSi) *Stability for the admissible similarity transformations, positive homogeneousness, homogeneity of degree one with respect to multiplication:* $M \in A_n(E, \mathbb{R})$ is stable for the admissible similarity transformations if

$$M(r\, x_1, \ldots, r\, x_n) = r\, M(x_1, \ldots, x_n)$$

for all $x \in E^n$ and all $r > 0$ such that $r\, x_i \in E$ for all $i \in N$.

Definition 2.2.2 (STr) *Stability for the admissible translations, homogeneousness with respect to addition, translativity:* $M \in A_n(E, \mathbb{R})$ is stable for the admissible translations if

$$M(x_1 + s, \ldots, x_n + s) = M(x_1, \ldots, x_n) + s \tag{2.3}$$

for all $x \in E^n$ and all $s \in \mathbb{R}$ such that $x_i + s \in E$ for all $i \in N$.

Definition 2.2.3 (SPL) *Stability for the admissible positive linear transformations:* $M \in A_n(E, \mathbb{R})$ is stable for the admissible positive linear transformations if

$$M(r\, x_1 + s, \ldots, r\, x_n + s) = r\, M(x_1, \ldots, x_n) + s$$

for all $x \in E^n$ and all $r > 0, s \in \mathbb{R}$ such that $r\, x_i + s \in E$ for all $i \in N$.

The choice of the interval $[0,1]$ is not restrictive if we consider that scores are defined up to a positive linear transformation, as it is the case for example in multiattribute utility theory.

Proposition 2.2.1 *For every $M \in A_n(E, \mathbb{R})$, the following assertions hold:*

$$i) \quad \text{(SSi, STr)} \Leftrightarrow \text{(SPL)} \tag{2.4}$$

$$ii) \quad \text{If } E \ni 0 \text{ then (SSi)} \Rightarrow M(e_\emptyset) = 0 \tag{2.5}$$

$$iii) \quad \text{If } E \ni 0 \text{ then (SPL)} \Rightarrow \text{(Id)} \tag{2.6}$$

Proof. $i)$ and $ii)$ Trivial.
$iii)$ It suffices to use $i)$ and $ii)$. ■

It clearly turns out, by the previous proposition, that if $E \ni 0$ then the condition "$r > 0$" in the statement of (SSi) or (SPL) can be replaced by "$r \geq 0$" without any effect.

More general definitions related to stability have been studied, even when the variables x_i correspond to independent scales. We now present some of them, see Aczél *et al.* [9] (see also Aczél and Roberts [8]).

Definition 2.2.4 Consider $M \in A_n(E, \mathbb{R})$. Then the property

(SRR) corresponds to same ratio scales for the independent variables and a ratio scale for the dependent variable. The corresponding functional equation is

$$M(r\,x_1, \ldots, r\,x_n) = R(r)M(x_1, \ldots, x_n)$$

where $x \in E^n$, $R > 0$, $r > 0$ and $r\,x_i \in E$ for all $i \in N$.

(IRR) corresponds to independent ratio scales for the independent variables and a ratio scale for the dependent variable. The corresponding functional equation is

$$M(r_1\,x_1, \ldots, r_n\,x_n) = R(r_1, \ldots, r_n)M(x_1, \ldots, x_n)$$

where $x \in E^n$, $R > 0$, $r_i > 0$ and $r_i\,x_i \in E$ for all $i \in N$.

(SII) corresponds to same interval scales for the independent variables and an interval scale for the dependent variable. The corresponding functional equation is

$$M(r\,x_1 + s, \ldots, r\,x_n + s) = R(r,s)M(x_1, \ldots, x_n) + S(r,s)$$

where $x \in E^n$, $R > 0$, $r > 0$ and $r\,x_i + s \in E$ for all $i \in N$.

(ISZII) corresponds to independent interval scales with same zero for the independent variables and an interval scale for the dependent variable. The corresponding functional equation is

$$M(r_1\,x_1 + s, \ldots, r_n\,x_n + s) = R(r_1, \ldots, r_n, s)M(x_1, \ldots, x_n) + S(r_1, \ldots, r_n, s)$$

where $x \in E^n$, $R > 0$, $r > 0$ and $r_i\,x_i + s \in E$ for all $i \in N$.

(ISUII) corresponds to independent interval scales with same unit for the independent variables and an interval scale for the dependent variable. The corresponding functional equation is

$$M(r\,x_1 + s_1, \ldots, r\,x_n + s_n) = R(r, s_1, \ldots, s_n)M(x_1, \ldots, x_n) + S(r, s_1, \ldots, s_n)$$

where $x \in E^n$, $R > 0$, $r > 0$ and $r\,x_i + s_i \in E$ for all $i \in N$.

(III) corresponds to independent interval scales for the independent variables and an interval scale for the dependent variable. The corresponding functional equation is

$$\begin{aligned} M(r_1 x_1 + s_1, \dots, r_n x_n + s_n) &= R(r_1, \dots, r_n, s_1, \dots, s_n) M(x_1, \dots, x_n) \\ &\quad + S(r_1, \dots, r_n, s_1, \dots, s_n) \end{aligned}$$

where $x \in E^n$, $R > 0$, $r_i > 0$ and $r_i x_i + s_i \in E$ for all $i \in N$.

2.2.2 Inversion scales

Definition 2.2.5 (SSN) *Stability for the standard negation:* $M \in A_n([0,1], \mathbb{R})$ is stable for the standard negation if

$$M(1 - x_1, \dots, 1 - x_n) = 1 - M(x_1, \dots, x_n), \quad x \in [0,1]^n.$$

The (SSN) property means that a reversal of the scale has no effect on the evaluation. As observed by Dubois and Koning [42], if we assume that the alternatives are rated in terms of *distaste* intensities instead of preference intensities then the global distaste should be built from individual distastes with the same aggregation operator as preferences. Indeed distaste and preference are just a matter of naming the assessment criterion (choosing the *good* alternatives or choosing the *bad* ones) and the aggregation operator should not depend on this name.

For a two-place function M, (SSN) expresses self-duality of M (compare with De Morgan laws in fuzzy sets theory, see e.g. [70]). This condition can be extended by using any strong negation $\varphi^{-1}(1-\varphi(x))$ instead of $1-x$, where $\varphi : [0,1] \to [0,1]$ is a continuous strictly increasing function fulfilling $\varphi(0) = 0$ and $\varphi(1) = 1$, see Trillas [181].

Proposition 2.2.2 *For every $M \in A_n([0,1], \mathbb{R})$, we have (SSi, SSN) $\Rightarrow$ (SPL).*

Proof. By (2.4), it suffices to show that M fulfils (STr). Let $x \in [0,1]^n$ and $s \in [-1,1]$ such that $x_i + s \in [0,1]$ for all $i \in N$. We have to prove that (2.3) holds. By (2.5) and (SSN) we have $M(e_N) = 1 - M(e_\emptyset) = 1$. So we can assume $s \in]-1,1[$. For all $i \in N$, set $y_i = x_i/(1-s)$. We then have $y_i \in [0,1]$ for all $i \in N$ and

$$\begin{aligned} M(x_1, \dots, x_n) &= (1-s)\ M(y_1, \dots, y_n) \quad \text{(SSi)} \\ &= (1-s) - (1-s)\ M(1-y_1, \dots, 1-y_n) \quad \text{(SSN)} \\ &= (1-s) - M[(1-s) - x_1, \dots, (1-s) - x_n] \quad \text{(SSi)} \\ &= -s + M(x_1 + s, \dots, x_n + s) \quad \text{(SSN)} \end{aligned}$$

■

In many situations, particularly concerning ratio judgements, it is reasonable to assume the following *reciprocal* property [5, 6].

Definition 2.2.6 (Rec) *reciprocal property:* Assume that E is an interval of positive numbers which with every element x contains also its reciprocal $1/x$. $M \in A_n(E,E)$ fulfils the reciprocal property if

$$M(1/x_1, \dots, 1/x_n) = 1/M(x_1, \dots, x_n), \quad x \in E^n.$$

Let a and b be two objects about which the ratio judgements are made (for instance, how much heavier a is than b). If we interchange a and b, then reasonably the judgements change into their reciprocals (if a is judged to be twice as heavy as b, then b should be judged half as heavy as a). The assumption (Rec) is that in this case also the aggregated judgement turns into its reciprocal.

2.2.3 Ordinal scales

The automorphism group of E, that is the group of all strictly increasing bijections $\phi : E \to E$ is denoted by $\Phi(E)$ and the set of all strictly increasing functions $\phi : E \to E$ by $\Phi'(E)$. Of course, for any E, we have $\Phi(E) \subset \Phi'(E)$. We also denote the vector $(\phi(x_1), \ldots, \phi(x_n))$ by $\phi(x)$, for all $x \in E^n$.

Definition 2.2.7 (OS, OS') *Ordinal stability:* $M \in A_n(E, E)$ is ordinally stable if, for all $\phi \in \Phi(E)$ (resp. $\Phi'(E)$), we have

$$M(\phi(x)) = \phi(M(x)), \quad x \in E^n. \tag{2.7}$$

Ovchinnikov [141, 142] showed that ordinal stability is a special form of the invariance property in measurement theory. Indeed, let us define an $(n+1)$-ary relation R on E^n by

$$R(y, x_1, \ldots, x_n) \Leftrightarrow y = M(x_1, \ldots, x_n), \qquad \forall y, x_1, \ldots, x_n \in E.$$

This relation is Φ-invariant if

$$R(y, x_1, \ldots, x_n) = R(\phi(y), \phi(x_1), \ldots, \phi(x_n)),$$

wich is equivalent to (2.7).

The following proposition has been proved by Ovchinnikov [141, Theorem 4.1] in the compensative case (see also [72, 121]). It shows that an operator which is insensitive to any scale change is forced to be ordinal in nature.

Proposition 2.2.3 *Let $M \in A_n(E, E)$ fulfilling (OS). Then*

$$M(x) \in \{x_1, \ldots, x_n\} \cup \{\inf E, \sup E\}, \quad x \in E^n.$$

Furthermore, if E is open or if M fulfils (OS') then

$$M(x) \in \{x_1, \ldots, x_n\}, \quad x \in E^n.$$

Proof. Consider $x = (x_1, \ldots, x_n) \in E^n$ reordered as $x_{(1)} \leq \ldots \leq x_{(n)}$ and set $x_0 := M(x)$. Suppose the result is false. We then have three exclusive cases:

- If $x_{(i)} < x_0 < x_{(i+1)}$ for one $i \in \{1, \ldots, n-1\}$ then there are elements $u, v \in E$ and a function $\phi \in \Phi(E)$ such that $x_{(i)} < u < x_0 < v < x_{(i+1)}$, $\phi(t) = t$ on $E \setminus [x_{(i)}, x_{(i+1)}]$, and $\phi(u) = v$. This implies $\phi(x_0) > x_0$, which is impossible because

$$\phi(x_0) = \phi(M(x)) = M(\phi(x)) = M(x) = x_0.$$

- If $x_0 < x_{(1)}$ then there are $v \in E$ and a function $\phi \in \Phi(E)$ such that $x_0 < v < x_{(1)}$, $\phi(t) = t$ for all $t \geq x_{(1)}$, and $\phi(x_0) = v$. This implies $\phi(x_0) > x_0$, a contradiction.
- The case $x_{(n)} < x_0$ can be treated as the previous one.

■

In the second part of Proposition 2.2.3, the assumption that E is open is necessary when not considering (OS'); indeed, if $a := \inf E \in E$ for example, then any $\phi \in \Phi(E)$ is such that $\phi(a) = a$ and thus the constant function $M(x) = a$ fulfils (OS).

As for ratio and interval scales (see Definition 2.2.4), a more general definition of the ordinal stability has been proposed: the comparison meaningfulness for ordinal values, see Orlov [140].

Definition 2.2.8 (CM, CM') *Comparison meaningfulness for ordinal values:* $M \in A_n(E, \mathbb{R})$ is comparison meaningful for ordinal values if, for all $\phi \in \Phi(E)$ (resp. $\Phi'(E)$) and all $x, x' \in E^n$, we have
i) $M(x) = M(x') \Rightarrow M(\phi(x)) = M(\phi(x'))$,
ii) $M(x) < M(x') \Rightarrow M(\phi(x)) < M(\phi(x'))$.

(CM) (resp. (CM')) is a weaker requirement than (OS) (resp. (OS')), but which is sufficient in decision making, as far as only the ranking of alternatives has importance. Moreover, for any $M \in A_n(E, E)$, we have (OS') $\Rightarrow$ (OS) and, for any $M \in A_n(E, \mathbb{R})$, we have (CM') $\Rightarrow$ (CM).

The following result is an adaptation of Lemma 2.2 in [141].

Proposition 2.2.4 *i) For every $M \in A_n(E, E)$, we have (Id, CM) $\Rightarrow$ (OS) $\Rightarrow$ (CM). If E is open, we have (Id, CM) $\Leftrightarrow$ (OS).*
ii) For every $M \in A_n(E, E)$, we have (Id, CM') $\Leftrightarrow$ (OS').

Proof. Let us prove *i*). The proof is identical for *ii*).
Let $x \in E^n$ and set $x_0 := M(x)$. By (Id), we have $M(x) = M(x_0, \ldots, x_0)$ and thus, for all $\phi \in \Phi(E)$,

$$M(\phi(x)) \stackrel{\text{(CM)}}{=} M(\phi(x_0), \ldots, \phi(x_0)) \stackrel{\text{(Id)}}{=} \phi(x_0) = \phi(M(x))$$

and M fulfils (OS). Conversely, it is clear that any $M \in A_n(E, E)$ fulfilling (OS) satisfies (CM) and if E is open then, by Proposition 2.2.3, M fulfils (Id). ■

When independent ordinal scales are considered, comparison meaningfulness takes the following form.

Definition 2.2.9 (CMIS) *Comparison meaningfulness for ordinal values with independent scales:* $M \in A_n(E, \mathbb{R})$ is comparison meaningful for ordinal values with independent scales if, for all $\phi_1, \ldots, \phi_n \in \Phi(E)$ and all $x, x' \in E^n$, we have
i) $M(x) = M(x') \Rightarrow M(\phi_1(x_1), \ldots, \phi_n(x_n)) = M(\phi_1(x'_1), \ldots, \phi_n(x'_n))$,
ii) $M(x) < M(x') \Rightarrow M(\phi_1(x_1), \ldots, \phi_n(x_n)) < M(\phi_1(x'_1), \ldots, \phi_n(x'_n))$.

The following two properties also concern ordinal values. They express a kind of stability for minitive and maxitive translations.

Definition 2.2.10 (SMin) *Stability for minimum with a constant vector:* $M \in A_n(E, \mathbb{R})$ is stable for minimum with a constant vector if

$$M(x_1 \wedge r, \ldots, x_n \wedge r) = M(x_1, \ldots, x_n) \wedge r$$

for all $x \in E^n$ and all $r \in E$.

Definition 2.2.11 (SMax) *Stability for maximum with a constant vector:* $M \in A_n(E, \mathbb{R})$ is stable for maximum with a constant vector if

$$M(x_1 \vee r, \ldots, x_n \vee r) = M(x_1, \ldots, x_n) \vee r$$

for all $x \in E^n$ and all $r \in E$.

(SMin) and (SMax) were introduced by Fodor and Roubens [71]. Clearly, they are related to an algebra which uses min and max operations instead of classical sum and product operations. In this sense, they look very much like (SSi) and (STr) respectively. Unfortunately, we do not know any practical interpretation of these properties. Their investigation here is purely theoretical.

Proposition 2.2.5 *For every $M \in A_n(E, \mathbb{R})$, we have (SMin, SMax) $\Rightarrow$ (Id).*

Proof. For all $x \in E$, we have, by (SMin, SMax),

$$M(x, \ldots, x) = M(x, \ldots, x) \wedge x \leq M(x, \ldots, x) \vee x = M(x, \ldots, x),$$

and thus $M(x, \ldots, x) = x$. ∎

We also introduce the following properties.

Definition 2.2.12 (SMinB) *Stability for minimum between Boolean and constant vectors:* $M \in A_n([0,1], \mathbb{R})$ is stable for minimum between Boolean and constant vectors if

$$M(r\, e_T) \in \{M(e_T), r\}$$

for all $T \subseteq N$ and all $r \in [0,1]$.

Definition 2.2.13 (SMaxB) *Stability for maximum between Boolean and constant vectors:* $M \in A_n([0,1], \mathbb{R})$ is stable for maximum between Boolean and constant vectors if

$$M(e_T + r\, \overline{e}_T) \in \{M(e_T), r\}$$

for all $T \subseteq N$ and all $r \in [0,1]$.

The following proposition justifies the names given to (SMinB) and (SMaxB) properties.

Proposition 2.2.6 *Consider $M \in A_n([0,1], \mathbb{R})$ fulfilling (In, Id). Then M fulfils (SMinB) if and only if*

$$M(x_1 \wedge r, \ldots, x_n \wedge r) = M(x_1, \ldots, x_n) \wedge r, \quad x \in \{0,1\}^n,\ r \in [0,1]. \tag{2.8}$$

Likewise, M fulfils (SMaxB) if and only if

$$M(x_1 \vee r, \ldots, x_n \vee r) = M(x_1, \ldots, x_n) \vee r, \quad x \in \{0,1\}^n,\ r \in [0,1]. \tag{2.9}$$

Proof. Let us consider the case of (SMinB). The other one can be treated similarly.
(Sufficiency) Trivial.
(Necessity) If $x \in \{0,1\}^n$ then there exists $T \subseteq N$ such that $x = e_T$. Next, let $r \in [0,1]$.
If $M(e_T) \leq r$ then

$$M(r\, e_T) \overset{\text{(In)}}{\leq} M(e_T) \leq r,$$

otherwise, if $M(e_T) \geq r$ then

$$M(r\, e_T) \overset{\text{(In)}}{\leq} M(r\, e_N) \overset{\text{(Id)}}{=} r \leq M(e_T).$$

Therefore, by (SMinB), we simply have,

$$M(x_1 \wedge r, \ldots, x_n \wedge r) = M(r\, e_T) = M(e_T) \wedge r = M(x_1, \ldots, x_n) \wedge r.$$

∎

2.2.4 Additivity and related properties

Definition 2.2.14 (Add) *Additivity:* $M \in A_n(E, \mathbb{R})$ is additive if, for all $x, x' \in E$, we have

$$M(x_1 + x'_1, \ldots, x_n + x'_n) = M(x_1, \ldots, x_n) + M(x'_1, \ldots, x'_n).$$

Definition 2.2.15 (Min) *Minitivity:* $M \in A_n(E, \mathbb{R})$ is minitive if, for all $x, x' \in E$, we have

$$M(x_1 \wedge x'_1, \ldots, x_n \wedge x'_n) = M(x_1, \ldots, x_n) \wedge M(x'_1, \ldots, x'_n).$$

Definition 2.2.16 (Max) *Maxitivity:* $M \in A_n(E, \mathbb{R})$ is maxitive if, for all $x, x' \in E$, we have

$$M(x_1 \vee x'_1, \ldots, x_n \vee x'_n) = M(x_1, \ldots, x_n) \vee M(x'_1, \ldots, x'_n).$$

The (Add) property is very classical. The (Min) and (Max) properties are less common but can be very useful when aggregating fuzzy preference relations.

A fuzzy binary relation R on a set A of alternatives is min-transitive (resp. negatively max-transitive) if, for all $a, b, c \in A$,

$$R(a, c) \wedge R(c, b) \leq R(a, b) \quad \text{(resp. } R(a, b) \leq R(a, c) \vee R(c, b)\text{)}.$$

The next proposition shows that it is useful to assume the (Min) and (Max) properties when we consider aggregation of min-transitive (or negatively max-transitive) fuzzy binary relations (see also [70, Sect. 7.3.1]).

Proposition 2.2.7 *Let $M \in A_n([0,1], \mathbb{R})$ fulfilling (In). Let A be a set of alternatives and $R_1, \ldots, R_n$ be min-transitive (resp. negatively max-transitive) fuzzy binary relations on A. Then the aggregated fuzzy relation R defined as*

$$R(a, b) = M(R_1(a, b), \ldots, R_n(a, b)), \quad \forall a, b \in A,$$

is a min-transitive (resp. negatively max-transitive) fuzzy binary relation if and only if M fulfils (Min) (resp. (Max)).

Proof. Consider the case of min-transitivity. The other one can be treated similarly.

(Necessity). Set $x_i^{ab} := R_i(a, b)$ for all $a, b \in A$ and all $i \in N$. By hypothesis, whenever $x_i^{ac} \wedge x_i^{cb} \leq x_i^{ab}$ for all $a, b, c \in A$ and all $i \in N$, we have

$$M(x_1^{ac}, \ldots, x_n^{ac}) \wedge M(x_1^{cb}, \ldots, x_n^{cb}) \leq M(x_1^{ab}, \ldots, x_n^{ab})$$

for all $a, b, c \in A$. In the particular case where $x_i^{ac} \wedge x_i^{cb} = x_i^{ab}$ for all $a, b, c \in A$ and all $i \in N$, since M fulfils (In), we obtain that:

$$\begin{aligned} M(x_1^{ab}, \ldots, x_n^{ab}) &= M(x_1^{ac} \wedge x_1^{cb}, \ldots, x_n^{ac} \wedge x_n^{cb}) \\ &\leq M(x_1^{ac}, \ldots, x_n^{ac}) \wedge M(x_1^{cb}, \ldots, x_n^{cb}) \end{aligned}$$

for all $a, b, c \in A$. Finally, we have that:

$$M(x_1^{ac} \wedge x_1^{cb}, \ldots, x_n^{ac} \wedge x_n^{cb}) = M(x_1^{ac}, \ldots, x_n^{ac}) \wedge M(x_1^{cb}, \ldots, x_n^{cb})$$

for all $a, b, c \in A$. Therefore, M fulfils (Min).

(Sufficiency). Suppose that $R_i(a,c) \wedge R_i(c,b) \leq R_i(a,b)$ for all $a,b,c \in A$ and all $i \in N$. We have, using (Min) and (In) successively,

$$\begin{aligned} & M(R_1(a,c),\dots,R_n(a,c)) \wedge M(R_1(c,b),\dots,R_n(c,b)) \\ = \; & M(R_1(a,c) \wedge R_1(c,b),\dots,R_n(a,c) \wedge R_n(c,b)) \\ \leq \; & M(R_1(a,b),\dots,R_n(a,b)) \end{aligned}$$

for all $a,b,c \in A$. Therefore, R is min-transitive. ■

We now present the concept of comonotonicity, which appeared as early as 1952 in Hardy *et al.* [97]. In the context we are interested in it is defined as follows.

Definition 2.2.17 Two vectors $x, x' \in E^n$ are said to be *comonotonic* if there exists a permutation $\pi \in \Pi_n$ such that

$$x_{\pi(1)} \leq \cdots \leq x_{\pi(n)} \quad \text{and} \quad x'_{\pi(1)} \leq \cdots \leq x'_{\pi(n)}.$$

Thus π orders the components of x and x' simultaneously. Another way to say that x and x' are comonotonic is that $(x_i - x_j)(x'_i - x'_j) \geq 0$ for every $i,j \in N$. Thus if $x_i < x_j$ for some i,j then $x'_i \leq x'_j$.

The following example has been given in [124, 125] to understand intuitively the notion of comonotonicity in multicriteria decision making.

Example 2.2.1 Let us assume that a consumer wants to buy a new car. His/her criteria are cost and performance. The set of possible choices are a Ferrari, 5 Renault cars, 5 Peugeot cars. We can now assume that the set of Renault cars constitutes a comonotonic set of his/her possible choices. This means that the consumer can say that if a criteria is better satisfied than another for one Renault car, this will be true for the other Renault cars. To do this, the consumer must put "mentally" the criteria on the same scale of satisfaction in order to compare them, since obviously, he/she cannot a priori compare the cost to the performance. We also see here the interest of the commensurability assumption (see Section 6.1.1.)

Definition 2.2.18 (CoAdd) *Comonotonic additivity:* $M \in A_n(E, \mathbb{R})$ is comonotonic additive if

$$M(x_1 + x'_1, \dots, x_n + x'_n) = M(x_1, \dots, x_n) + M(x'_1, \dots, x'_n)$$

for any two comonotonic vectors $x, x' \in E$.

Definition 2.2.19 (CoMin) *Comonotonic minitivity:* $M \in A_n(E, \mathbb{R})$ is comonotonic minitive if

$$M(x_1 \wedge x'_1, \dots, x_n \wedge x'_n) = M(x_1, \dots, x_n) \wedge M(x'_1, \dots, x'_n)$$

for any two comonotonic vectors $x, x' \in E$.

Definition 2.2.20 (CoMax) *Comonotonic maxitivity:* $M \in A_n(E, \mathbb{R})$ is comonotonic maxitive if

$$M(x_1 \vee x'_1, \dots, x_n \vee x'_n) = M(x_1, \dots, x_n) \vee M(x'_1, \dots, x'_n)$$

for any two comonotonic vectors $x, x' \in E$.

The concept of comonotonic additivity has appeared first in [35] and more recently in [160]. Comonotonic minitivity and maxitivity were introduced for the first time (in the context of fuzzy integrals) in [33]. Note that a justification of these two latter properties has been given by Ralescu and Ralescu [148] in the framework of aggregation of fuzzy subsets.

2.3 More elaborate mathematical properties

The following properties concern the "decomposability" of the aggregation procedure. It means that it is possible to partition the set of criteria (voters, etc.) into disjoint subgroups, build the partial aggregation for each subgroup and then combine these partial results to get the global value. This condition may take several forms. The strongest one we will present is associativity. Other weaker formulations will also be presented: decomposability, autodistributivity, bisymmetry, self-identity.

2.3.1 Associativity

We consider first the associativity functional equation. Associativity of addition means that $(a+b)+c = a+(b+c)$, so we can write $a+b+c$ unambiguously. If we write the addition operation as a two-place function $f(a,b) = a+b$, then associativity says that $f(f(a,b),c) = f(a, f(b,c))$. For general f, this is the *associativity functional equation.*

Definition 2.3.1 (A) *Associativity (for two arguments):* $M \in A_2(E,E)$ is associative if, for all $x \in E^3$, we have

$$M(M(x_1,x_2),x_3) = M(x_1, M(x_2,x_3)). \tag{2.10}$$

A large number of papers deal with the associativity functional equation (2.10) even in the field of real numbers. In complete generality its investigation naturally constitutes a principal subject of algebra. For a list of references see Aczél [4, Sect. 6.2].

With the use of a graphical representation linked to clustering procedures, we obtain Figure 2.2. From (2.10), it is clear that we have to suppose that the range of $M^{(2)}$ is a subset of E.

Figure 2.2: Associativity for two-place operators

Basically, associativity concerns aggregation of only two arguments. But it can be extended to any finite number of arguments.

Definition 2.3.2 (A) *Associativity:* $M \in A(E,E)$ is associative if $M(x) = x$ for all $x \in E$, and

$$\begin{aligned} & M(x_1,\ldots,x_j,x_{j+1},\ldots,x_k,x_{k+1},\ldots,x_n) \\ = \; & M(x_1,\ldots,x_j,M(x_{j+1},\ldots,x_k),x_{k+1},\ldots,x_n) \end{aligned} \tag{2.11}$$

for all $n \in \mathbb{N}_0$, all $x \in E^n$ and all integers j,k such that $0 \le j < k \le n$.

Associativity means that each subset of consecutive elements from $x \in E^n$ can be replaced by their partial aggregation without changing the global aggregation. On a graphical basis, we obtain Figure 2.3.

Associativity is also a well-known algebraic property which allows to omit "parentheses" in an aggregation of at least three elements. Observe that, if the extended operator M is associative,

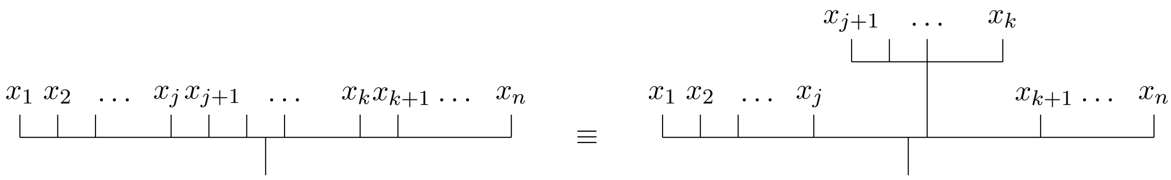


Figure 2.3: The associativity property

then the function $M^{(2)}$ is associative (just set $n = 3$ in (2.11)). Implicit in the assumption of associativity is a consistent way of going unambiguously from the aggregation of n elements to $n+1$ elements, i.e. if M is associative:

$$M(x_1, \ldots, x_{n+1}) = M(M(x_1, \ldots, x_n), x_{n+1}), \quad n \in \mathbb{N}_0.$$

Thus associativity mandates how we define the extended aggregation operator and does so in a way that keeps some consistency between the aggregation of n and $n+1$ elements.

2.3.2 Decomposability

It can be easily verified that the arithmetic mean as an extended operator does not solve the associativity equation (2.11). So, it seems interesting to know whether there exists a functional equation, similar to associativity, which can be solved by the arithmetic mean, or even by other means such as the geometric mean, the quadratic mean, etc.

On this subject, an acceptable equation, called *associativity of means*, has been proposed for symmetric extended operators (Sy), and can be formulated as follows:

$$M(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = M(k \odot M(x_1, \ldots, x_k), x_{k+1}, \ldots, x_n)$$

for all $k, n \in \mathbb{N}_0$ such that $k \leq n$. It was already mentioned in Section 1.2 that, under idempotence (Id), this condition says that the global aggregation does not change when altering some values without modifying their partial aggregation:

$$M(x_1, \ldots, x_k) = M(x'_1, \ldots, x'_k) \;\Rightarrow\; M(x_1, \ldots, x_k, x_{k+1}, \ldots, x_n) = M(x'_1, \ldots, x'_k, x_{k+1}, \ldots, x_n)$$

for all $k, n \in \mathbb{N}_0$ such that $k \leq n$. Introduced first in Bemporad [18, p. 87] in a characterization of the arithmetic mean, associativity of means has been used by Kolmogoroff [107] and Nagumo [136] to characterize the so-called mean values. More recently, Marichal and Roubens [121] proposed to call this property "decomposability" in order not to confuse it with classical associativity (A).

When symmetry is not assumed, it is necessary to rewrite this property in such a way that the first variables are not privileged. We then propose to generalize this concept in two ways: decomposability (D) and strong decomposability (SD).

Definition 2.3.3 (D) *Decomposability:* $M \in A(E, E)$ is decomposable if $M(x) = x$ for all $x \in E$, and

$$\begin{aligned} & M(x_1, \ldots, x_j, x_{j+1}, \ldots, x_k, x_{k+1}, \ldots, x_n) \\ = \; & M(x_1, \ldots, x_j, (k-j) \odot M(x_{j+1}, \ldots, x_k), x_{k+1}, \ldots, x_n) \end{aligned}$$

for all $n \in \mathbb{N}_0$, all $x \in E^n$ and all integers j, k such that $0 \leq j < k \leq n$.

The decomposability property means that each element of any subset of consecutive elements from $x \in E^n$ can be replaced by their partial aggregation without changing the global aggregation (see Figure 2.4).

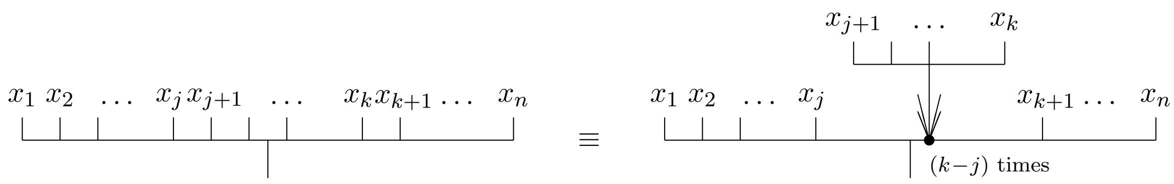


Figure 2.4: The decomposability property

Proposition 2.3.1 *Let $M \in A(E,E)$ fulfilling (D). If $M^{(2)}$ fulfils (Sy) then so does M.*

Proof. Let us proceed by induction on $n \geq 2$. Assume that $M^{(n)}$ fulfils (Sy) for a fixed $n \geq 2$. Let $(x_1, \ldots, x_{n+1}) \in E^{n+1}$. By (D), we have

$$\begin{aligned} M^{(n+1)}(x_1, \ldots, x_{n+1}) &= M^{(n+1)}(x_1, n \odot M^{(n)}(x_2, \ldots, x_{n+1})) \\ &= M^{(n+1)}(n \odot M^{(n)}(x_1, \ldots, x_n), x_{n+1}). \end{aligned}$$

Since $M^{(n)}$ fulfils (Sy), we have

$$M^{(n+1)}(x_2, x_1, x_3, \ldots, x_{n+1}) = M^{(n+1)}(x_1, x_2, x_3, \ldots, x_n) = M^{(n+1)}(x_2, x_3, \ldots, x_{n+1}, x_1).$$

By Proposition 2.1.1, $M^{(n+1)}$ fulfils (Sy). ∎

Definition 2.3.4 (SD) *Strong decomposability:* $M \in A(E,E)$ is strongly decomposable if $M(x) = x$ for all $x \in E$, and

$$M\Big(\sum_{i \in K} x_i\, e_i + \sum_{i \notin K} x_i\, e_i\Big) = M\Big(M^{(k)}(x_{i_1}, \ldots, x_{i_k})\, e_K + \sum_{i \notin K} x_i\, e_i\Big)$$

for all $n \in \mathbb{N}_0$, all $x \in E^n$ and all $K = \{i_1, \ldots, i_k\} \subseteq N$ with $i_1 < \cdots < i_k$.

Strong decomposability means that each element of any subset of elements (which are not necessarily consecutive) from $x \in E^n$ can be replaced by their partial aggregation without changing the global aggregation. Under idempotence (Id), this property is equivalent to:

$$M(x_{i_1}, \ldots, x_{i_k}) = M(x'_{i_1}, \ldots, x'_{i_k}) \quad \Rightarrow \quad M\Big(\sum_{i \in K} x_i\, e_i + \sum_{i \notin K} x_i\, e_i\Big) = M\Big(\sum_{i \in K} x'_i\, e_i + \sum_{i \notin K} x_i\, e_i\Big),$$

for all $n \in \mathbb{N}_0$, all $x, x' \in E^n$ and all $K = \{i_1, \ldots, i_k\} \subseteq N$ with $i_1 < \cdots < i_k$.

Proposition 2.3.2 *For every $M \in A(E,E)$, the following assertions hold:*

$$\begin{aligned} &i) && \text{(SD)} \Rightarrow \text{(D)} && (2.12) \\ &ii) && \text{(Sy, SD)} \Leftrightarrow \text{(Sy, D)} && (2.13) \\ &iii) && \text{(Id, A)} \Rightarrow \text{(D)} && (2.14) \end{aligned}$$

Proof. *i*) and *ii*) Trivial.

iii) Let $j, k, n \in \mathbb{N}$ such that $0 \leq j < k \leq n$. For all $x \in E^n$, we have

$$\begin{aligned} & M(x_1, \ldots, x_j, x_{j+1}, \ldots, x_k, x_{k+1}, \ldots, x_n) \\ = \;& M(x_1, \ldots, x_j, M(x_{j+1}, \ldots, x_k), x_{k+1}, \ldots, x_n) \quad \text{(A)} \\ = \;& M(x_1, \ldots, x_j, M[(k-j) \odot M(x_{j+1}, \ldots, x_k)], x_{k+1}, \ldots, x_n) \quad \text{(Id)} \\ = \;& M(x_1, \ldots, x_j, (k-j) \odot M(x_{j+1}, \ldots, x_k), x_{k+1}, \ldots, x_n) \quad \text{(A)}. \end{aligned}$$

■

According to (2.14), it seems more useful to consider decomposable idempotent extended operators rather than associative idempotent extended operators. As example, the arithmetic mean, which seems to be a quite acceptable idempotent extended operator, is decomposable but not associative.

Proposition 2.3.3 *Let $M \in A(E, E)$ fulfilling (Id, SD). Then we have*

i) $M(k \odot x_1, \ldots, k \odot x_n) = M(x_1, \ldots, x_n)$ *for all $k, n \in \mathbb{N}_0$ and all $x \in E^n$.* (2.15)

ii) $M(x_{11}, \ldots, x_{1n}; \ldots; x_{p1}, \ldots, x_{pn}) = M[M(x_{11}, \ldots, x_{1n}), \ldots, M(x_{p1}, \ldots, x_{pn})]$ (2.16)
for all matrices $X = (x_{ij}) \in E^{p \times n}$, where $n, p \in \mathbb{N}_0$.

iii) $M(x_1, \ldots, x_n) = M(x'_n, \ldots, x'_1)$ *for all $n \in \mathbb{N}, n \geq 2$ and all $x \in E^n$,* (2.17)
where $x'_j = M(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$ for all $j \in N$.

Proof. The proof is an adaptation of that of Propositions 1 and 2 in Nagumo [136, Sect. 1].

i) For all $k, n \in \mathbb{N}_0$ and all $x \in E^n$, we have

$$\begin{aligned} M(k \odot x_1, \ldots, k \odot x_n) &= M[k\, n \odot M(x_1, \ldots, x_n)] \quad \text{(SD)} \\ &= M(x_1, \ldots, x_n) \quad \text{(Id)}. \end{aligned}$$

ii) For all matrices $X = (x_{ij}) \in E^{p \times n}$, where $n, p \in \mathbb{N}_0$, we have

$$\begin{aligned} M(x_{11}, \ldots, x_{1n}; \ldots; x_{p1}, \ldots, x_{pn}) &= M[n \odot M(x_{11}, \ldots, x_{1n}), \ldots, n \odot M(x_{p1}, \ldots, x_{pn})] \quad \text{(SD)} \\ &= M[M(x_{11}, \ldots, x_{1n}), \ldots, M(x_{p1}, \ldots, x_{pn})] \quad \text{(by (2.15))}. \end{aligned}$$

iii) For all $n \in \mathbb{N}, n \geq 2$ and all $x \in E^n$, we have, by (2.15),

$$M(x_1, \ldots, x_n) = M[(n-1) \odot x_1, \ldots, (n-1) \odot x_n]$$

and by using (SD) with subset $K_j = \{j, n+j, 2n+j, \ldots, (n-2)n+j\}$ for $j = 1, \ldots, n$, we obtain

$$M[(n-1) \odot x_1, \ldots, (n-1) \odot x_n] = M(x'_n, \ldots, x'_1; \ldots; x'_n, \ldots, x'_1).$$

Therefore, we have

$$\begin{aligned} M(x_1, \ldots, x_n) &= M(x'_n, \ldots, x'_1; \ldots; x'_n, \ldots, x'_1) \\ &= M((n-1) \odot M(x'_n, \ldots, x'_1)) \quad \text{(by (2.16))} \\ &= M(x'_n, \ldots, x'_1) \quad \text{(Id)}, \end{aligned}$$

which was to be proved. ■

2.3.3 Autodistributivity

Definition 2.3.5 (AD) *Autodistributivity, self-distributivity (for two arguments):* $M \in A_2(E,E)$ is autodistributive if, for all $x \in E^3$, we have

$$\begin{aligned} M(x_1, M(x_2,x_3)) &= M(M(x_1,x_2), M(x_1,x_3)) \\ \text{and} \quad M(M(x_1,x_2),x_3) &= M(M(x_1,x_3), M(x_2,x_3)). \end{aligned} \tag{2.18}$$

The autodistributivity equations (2.18) were investigated both in general algebraic structures and for real numbers in particular. A list of references can be found in [4, Sect. 6.5] (see also [7, Chap. 17]).

2.3.4 Bisymmetry and related properties

Definition 2.3.6 (B) *Bisymmetry for two arguments, mediality:* $M \in A_2(E,E)$ is bisymmetric if for all $x \in E^4$, we have

$$M(M(x_1,x_2), M(x_3,x_4)) = M(M(x_1,x_3), M(x_2,x_4)).$$

The bisymmetry property (also called mediality) 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 — in a certain respect, it has been used as a substitute for associativity (A) and also for symmetry (Sy). For a list of references see [4, Sect. 6.4] (see also [7, Chap. 17]).

Proposition 2.3.4 *For every $M \in A_2(E,E)$, we have*
i) (Sy, A) $\Rightarrow$ (B),
ii) (Id, B) $\Rightarrow$ (AD),
iii) (SIn, AD) $\Rightarrow$ (Id).

Proof. See [46, Sect. 1.2.1]. ∎

Definition 2.3.7 (B) *Bisymmetry (for $n \geq 2$ arguments):* $M \in A_n(E,E)$ is bisymmetric if

$$\begin{aligned} & M(M(x_{11},\ldots,x_{1n}),\ldots,M(x_{n1},\ldots,x_{nn})) \\ = \; & M(M(x_{11},\ldots,x_{n1}),\ldots,M(x_{1n},\ldots,x_{nn})) \end{aligned} \tag{2.19}$$

for all square matrices

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

Bisymmetry expresses 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 in MCDM. Its usefulness remains theoretical.

Definition 2.3.8 (GB) *General bisymmetry:* $M \in A(E,E)$ fulfils the general bisymmetry property if $M(x) = x$ for all $x \in E$, and

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

for all matrices

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix} \in E^{p\times n},$$

where $n, p \in \mathbb{N}_0$.

Contrary to (B), the (GB) property can be justified rather easily. Consider n judges (or criteria, attributes, etc.) giving a score to each of p candidates. These scores, assumed to be defined on a same scale, can be put in a $p \times n$ matrix as follows:

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

Suppose now that we want to aggregate all the scores in the matrix in order to obtain a global score of the p candidates. A reasonable way to proceed could be the following: aggregate the scores of each candidate (aggregation on the rows of the matrix), and then aggregate these global values. A dual way to proceed would be: aggregate the scores given by each judge (aggregation on the columns of the matrix), and then aggregate these values.

The general bisymmetry property for an aggregation operator means that these two ways to aggregate must lead to the same global score. This is a rather natural property.

Proposition 2.3.5 *If $M \in A(E,E)$ fulfils (GB) then, for all $n \in \mathbb{N}$ with $n \geq 2$, $M^{(n)}$ fulfils (B).*

Proposition 2.3.6 *For every $M \in A(E,E)$, we have*
i) (Sy, A) $\Rightarrow$ (GB),
ii) (Id, SD) $\Rightarrow$ (GB).

Proof. *i)* Trivial.
ii) For all matrices $X = (x_{ij}) \in E^{p\times n}$, where $n, p \in \mathbb{N}_0$, we have

$$\begin{aligned} & M(M(x_{11},\ldots,x_{1n}),\ldots,M(x_{p1},\ldots,x_{pn})) \\ = \;& M(x_{11},\ldots,x_{1n};\ldots;x_{p1},\ldots,x_{pn}) \quad \text{(by (2.16))} \\ = \;& M[M(x_{11},\ldots,x_{p1}),\ldots,M(x_{1n},\ldots,x_{pn}),\ldots,M(x_{11},\ldots,x_{p1}),\ldots,M(x_{1n},\ldots,x_{pn})] \quad \text{(SD)} \\ = \;& M[p\odot M(M(x_{11},\ldots,x_{p1}),\ldots,M(x_{1n},\ldots,x_{pn}))] \quad \text{(by (2.16))} \\ = \;& M(M(x_{11},\ldots,x_{p1}),\ldots,M(x_{1n},\ldots,x_{pn})) \quad \text{(Id)} \end{aligned}$$

which proves the result. ∎

A matrix $X \in E^{p\times n}$ is *ordered* if its elements satisfy $x_{ij} \leq x_{kl}$ for all $i \leq k$ and $j \leq l$. It is said to be *orderable* if it is possible to make it ordered by permuting some rows and/or some columns. Notice that not all matrices are orderable as the following example shows:

$$X = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}.$$

Definition 2.3.9 (BOM) *Bisymmetry for orderable matrices:* $M \in A_n(E, E)$ is bisymmetric for orderable matrices if

$$\begin{aligned} & M([M([x_{11}, \ldots, x_{1n}]_{\pi'}), \ldots, M([x_{n1}, \ldots, x_{nn}]_{\pi'})]_{\pi}) \\ = \; & M([M([x_{11}, \ldots, x_{n1}]_{\pi}), \ldots, M([x_{1n}, \ldots, x_{nn}]_{\pi})]_{\pi'}) \end{aligned}$$

for all $\pi, \pi' \in \Pi_n$ and all ordered square matrices

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

Definition 2.3.10 (GBOM) *General bisymmetry for orderable matrices:* $M \in A(E, E)$ fulfils the general bisymmetry for orderable matrices if

$$\begin{aligned} & M([M([x_{11}, \ldots, x_{1n}]_{\pi'}), \ldots, M([x_{p1}, \ldots, x_{pn}]_{\pi'})]_{\pi}) \\ = \; & M([M([x_{11}, \ldots, x_{p1}]_{\pi}), \ldots, M([x_{1n}, \ldots, x_{pn}]_{\pi})]_{\pi'}) \end{aligned}$$

for all $\pi \in \Pi_p$, all $\pi' \in \Pi_n$ and all ordered matrices

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

Of course, (B) implies (BOM) and (GB) implies (GBOM).

We now give a justification of (GBOM). Consider the same situation as above: n judges give a score to each of p candidates. Now we start by removing some values—for example the lowest score given by each judge, or the lowest score obtained by each candidate. In general, it does not make sense anymore to aggregate as before.

However, there are situations where it still make sense: if the worst candidate is the same for each judge then, when removing this candidate, we get a score matrix for $(p-1)$ candidates and n judges, and we can aggregate as before. Likewise, if the most intolerant judge is the same for each candidate then, when removing the judge, we get a score matrix for p candidates and $(n-1)$ judges, and we can aggregate. Clearly, if we wish to take into account all the possibilities to remove judges and candidates, we have to consider orderable score matrices.

2.3.5 Self-identity

The property of *self-identity* for extended aggregation operators was introduced by Yager [196, 197] and formulated as follows:

Definition 2.3.11 (SId) *Self-identity:* $M \in A(E, E)$ is a self-identity extended aggregation operator if $M(x) = x$ for all $x \in E$, and

$$M(x_1, \ldots, x_n, M(x_1, \ldots, x_n)) = M(x_1, \ldots, x_n).$$

for all $n \in \mathbb{N}_0$ and $x \in E^n$.

Thus we see that in the case of self-identity extended operators adding an element equal to the already established value does not change the aggregation value.

In this definition, the last argument is privileged. This can make sense in some situations [198], even when (Sy) is not assumed. For instance, consider a situation in which the arguments are temporal in nature, in this case x_i indicates the i-th observed reading. In situations in which we feel that the basic underlying process generating the readings is changing we may desire to give more emphasis to the later readings rather than to the former ones.

Yager and Rybalov [198] established the following result.

Proposition 2.3.7 *Let $M \in A(E, E)$.*
i) (SId) $\Rightarrow$ (Id).
ii) Under (In), we have

$$x \begin{Bmatrix} < \\ > \end{Bmatrix} M(x_1, \ldots, x_n) \quad \Rightarrow \quad M(x_1, \ldots, x_n, x) \begin{Bmatrix} \leq \\ \geq \end{Bmatrix} M(x_1, \ldots, x_n).$$

The second part of the result essentially implies that values greater than the already established value tend to increase the score while those below it tend to decrease this score.

2.3.6 Separability

The property of *separability*, suggested by Aczél and Saaty [10], means that the influences of the individual judgements can be separated.

Definition 2.3.12 (Sep) *Separability: $M \in A_n(E, E)$* is separable if

$$M(x_1, \ldots, x_n) = g(x_1) \circ \cdots \circ g(x_n), \quad x_1, \ldots, x_n \in E, \tag{2.20}$$

where $g : E \to E$ is continuous and non-constant and $\circ$ is a continuous *associative* and *cancellative* operation mapping $E \times E$ into E, i.e.

$$(u \circ v) \circ w = u \circ (v \circ w), \quad \forall u, v, w \in E,$$

and

$$\begin{cases} u \circ t = v \circ t, \text{ for any } t \in E, & \Rightarrow \quad u = v, \\ t \circ u = t \circ v, \text{ for any } t \in E, & \Rightarrow \quad u = v. \end{cases}$$

Note that if not all judging individuals have the same weight when the judgements are aggregated then these different influences should be reflected in different functions $g_1, \ldots, g_n$ and (2.20) must be replaced by

$$M(x_1, \ldots, x_n) = g_1(x_1) \circ \cdots \circ g_n(x_n), \quad x_1, \ldots, x_n \in E.$$

An interesting study of such non-symmetric functions can be found in Aczél [5].

Chapter 3

Some classes of aggregation operators

Given any aggregation operator, we can ask for a motivation of its use, i.e. for reasonable conditions (or properties) which lead to this operator. Conversely, we can specify *a priori* some conditions and determine all the aggregation operators satisfying these.

Proposing an interesting axiomatic characterization of an operator (or a family of operators) is not an easy task. Mostly, aggregation operators can be characterized by different sets of conditions. Nevertheless the various possible characterizations are not equally important. Some of them involve purely technical conditions with no clear interpretation and the result becomes useless. Some other involve conditions that contain explicitly the result and the characterization becomes trivial. On the contrary, there are characterizations involving only natural conditions, easily interpretable. In fact, this is the only case where the result should be seen as a significant contribution. It improves our understanding of the operator considered and provides strong arguments to justify (or reject) its use in the context of decision making.

The aim of this chapter is to present characterizations of some families of aggregation operators (or extended aggregation operators). Most of the characterizations presented here, like those of the quasi-arithmetic means, involve rather natural properties.

In Section 3.1 the family of bisymmetric aggregation operators is studied in the presence of some properties such as continuity and idempotence. Some characterizations are also presented. Section 3.2 deals with the extended aggregation operators that fulfil the decomposability property. In particular, a description of the quasi-arithmetic means that are not necessarily strictly increasing is given. In Section 3.3 we present the theory of associative functions and generalize some well-known characterizations. Sections 3.4 and 3.5 consider some operators that are suitable for aggregation of values defined on specific scale types, especially the ordinal and interval scales.

Whenever the form of E is not specified in a statement, it is understood as an arbitrary real interval, finite or infinite.

3.1 Bisymmetric operators

3.1.1 Quasi-arithmetic means

Let E be a real interval, finite or infinite. It has been proved by Aczél [2] (see also [4, Sect. 6.4] and [7, Chap. 17]) that the quasi-arithmetic means are the only symmetric, continuous, strictly increasing, idempotent, real functions $M \in A_n(E, E)$ which satisfy the bisymmetry condition (2.19). The statement of this result is formulated as follows.

Theorem 3.1.1 *$M \in A_n(E,E)$ fulfils (Sy, Co, SIn, Id, B) if and only if there exists a continuous strictly monotonic function $f : E \to \mathbb{R}$ such that*

$$M(x) = f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big], \quad x \in E^n. \tag{3.1}$$

The *quasi-arithmetic means* (3.1) are compensative aggregation operators and cover a wide spectrum of means including arithmetic, quadratic, geometric, harmonic, root-power and exponential means as it can be seen in Table 3.1.

$f(x)$	$M(x)$	name
x	$\frac{1}{n}\sum x_i$	arithmetic
x^2	$\sqrt{\frac{1}{n}\sum x_i^2}$	quadratic
$\log x$	$\sqrt[n]{\prod x_i}$	geometric
x^{-1}	$\frac{1}{\frac{1}{n}\sum \frac{1}{x_i}}$	harmonic
x^α $(\alpha \in \mathbb{R}_0)$	$\Big(\frac{1}{n}\sum x_i^\alpha\Big)^{\frac{1}{\alpha}}$	root-power
$e^{\alpha x}$ $(\alpha \in \mathbb{R}_0)$	$\frac{1}{\alpha}\ln\Big[\frac{1}{n}\sum e^{\alpha x_i}\Big]$	exponential

Table 3.1: Examples of quasi-arithmetic means

The function f occuring in (3.1) is called a *generator* of M. It was also proved that f is determined up to a linear transformation: with $f(x)$, every function $g(x) = r\,f(x) + s$ $(r, s \in \mathbb{R}, r \neq 0)$ belongs to the same M, but no other function.

Note that Aczél and Alsina [6] proved that the quasi-arithmetic means can be characterized only by two property: idempotence (Id) and separability (Sep) (with continuous non-constant g and continuous, cancellative and associative $\circ$).

Theorem 3.1.2 *$M \in A_n(E,E)$ fulfils (Id, Sep) if and only if there exists a continuous strictly monotonic function $f : E \to \mathbb{R}$ such that M is of the form (3.1).*

Nagumo [136] investigated some subfamilies of the class of quasi-arithmetic means. He proved the following result (see also [6, Sect. 4] and [7, Chap. 15]).

Theorem 3.1.3 *i) $M \in A_n(E,E)$ is a quasi-arithmetic mean fulfilling (STr) if and only if*

- *either M is the arithmetic mean:*

$$M(x) = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad x \in E^n,$$

- *or M is the exponential mean: there exists $\alpha \in \mathbb{R}_0$ such that*

$$M(x) = \frac{1}{\alpha}\ln\Big[\frac{1}{n}\sum_{i=1}^{n} e^{\alpha\, x_i}\Big], \quad x \in E^n.$$

ii) Let $E = \mathbb{R}_0^+$ or a subinterval. $M \in A_n(E,E)$ is a quasi-arithmetic mean fulfilling (SSi) if and only if

- *either M is the geometric mean:*

$$M(x) = \Big(\prod_{i=1}^{n} x_i\Big)^{\frac{1}{n}}, \quad x \in E^n,$$

- *or M is the root-power mean: there exists $\alpha \in \mathbb{R}_0$ such that*

$$M(x) = \Big(\frac{1}{n}\sum_{i=1}^{n} x_i^{\alpha}\Big)^{\frac{1}{\alpha}}, \quad x \in E^n. \tag{3.2}$$

Thus the arithmetic mean is the only quasi-arithmetic mean fulfilling (SPL) when $E = \mathbb{R}_0^+$ or a subinterval. This result was already reached in 1926 by Bemporad [18, p. 87]. We will see in Section 4.2.4 that this remains true with any $E \supseteq [0,1]$ (see Corollary 4.2.4).

Let us denote by $M_{(\alpha)}$ the root-power mean (3.2) generated by $\alpha \in \mathbb{R}_0$. It is well known that, if $\alpha_1 < \alpha_2$ then $M_{(\alpha_1)}(x) \leq M_{(\alpha_2)}(x)$ for all $x \in]0,+\infty[^n$ (equality if and only if all x_i are equal).

The family of root-power means was studied by Dujmovic [54, 55, 56] and Dyckhoff and Pedrycz [57]. It encompasses most of traditionally known means: the arithmetic mean $M_{(1)}$, the harmonic mean $M_{(-1)}$, the quadratic mean $M_{(2)}$, and three limit cases: the geometric mean $M_{(0)}$, the minimum $M_{(-\infty)}$ and the maximum $M_{(+\infty)}$ (see e.g. [1]).

In addition to the Nagumo's result (Theorem 3.1.3), we have the following characterizations.

Theorem 3.1.4 *i) Let E be an interval of positive numbers which with every element x contains also its reciprocal $1/x$. $\ln(E)$ denotes the set of all $\ln x$, where x is in E. Then $M \in A_n(E,E)$ is a quasi-arithmetic mean fulfilling (Rec) if and only if there exists a continuous, strictly monotonic, odd function $\omega : \ln(E) \to \mathbb{R}$ such that*

$$M(x) = \exp\Big[\omega^{-1}\Big(\frac{1}{n}\sum_{i=1}^{n}\omega(\ln x_i)\Big)\Big], \quad x \in E^n.$$

ii) $M \in A_n([0,1],[0,1])$ is a quasi-arithmetic mean fulfilling (SSN) if and only if there exists a continuous, strictly monotonic, odd function $\omega : [-\frac{1}{2},\frac{1}{2}] \to \mathbb{R}$ such that

$$M(x) = \frac{1}{2} + \omega^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n}\omega(x_i - \frac{1}{2})\Big], \quad x \in [0,1]^n.$$

Proof. *i*) see Aczél and Alsina [6].

ii) $M \in A_n([0,1],[0,1])$ is a quasi-arithmetic mean fulfilling (SSN) if and only if $F \in A_n([\frac{1}{\sqrt{e}}, \sqrt{e}\,],[\frac{1}{\sqrt{e}}, \sqrt{e}\,])$, defined by

$$F(z) = \frac{1}{\sqrt{e}} \exp[M(\ln(\sqrt{e}\, z_1), \ldots, \ln(\sqrt{e}\, z_n))], \quad z \in [\frac{1}{\sqrt{e}}, \sqrt{e}\,]^n,$$

is a quasi-arithmetic mean fulfilling (Rec). We then can conclude by *i*). ∎

Back to Theorem 3.1.1, note that Aczél [2] also investigated the case where (Sy) and (Id) are dropped (see also [4, Sect. 6.4] and [7, Chap. 17]). He showed the power of the concept of bisymmetry by the fact that the theory of quasi-arithmetic means does not loose much of its force if both (Sy) and (Id) fail to hold. He obtained the following result.

Theorem 3.1.5 *i) $M \in A_n(E,E)$ fulfils (Co, SIn, Id, B) if and only if there exist a continuous strictly monotonic function $f : E \to \mathbb{R}$ and real numbers $\omega_1, \ldots, \omega_n > 0$ fulfilling $\sum_i \omega_i = 1$ such that*

$$M(x) = f^{-1}\Big[\sum_{i=1}^{n} \omega_i\, f(x_i)\Big], \quad x \in E^n. \tag{3.3}$$

ii) $M \in A_n(E,E)$ fulfils (Co, SIn, B) if and only if there exist a continuous strictly monotonic function $f : E \to \mathbb{R}$ and real numbers $p_1, \ldots, p_n > 0, q \in \mathbb{R}$ such that

$$M(x) = f^{-1}\Big[\sum_{i=1}^{n} p_i\, f(x_i) + q\Big], \quad x \in E^n. \tag{3.4}$$

The *quasi-linear means* (3.3) and the *quasi-linear functions* (3.4) are weighted aggregation operators. In the set of properties given here for these operators, the weights are not *a priori* given. The question of uniqueness with respect to f is dealt in details in [4, Sect. 6.4].

Table 3.2 provides some particular cases of quasi-linear means.

$f(x)$	$M(x)$	name of weighted mean
x	$\sum \omega_i\, x_i$	arithmetic
x^2	$\sqrt{\sum \omega_i\, x_i^2}$	quadratic
$\log x$	$\prod x_i^{\omega_i}$	geometric
x^α ($\alpha \in \mathbb{R}_0$)	$\left(\sum \omega_i\, x_i^\alpha\right)^{\frac{1}{\alpha}}$	root-power

Table 3.2: Examples of quasi-linear means

Note that Aczél [4, Sect. 6.5] showed that the two-place quasi-linear means are the general continuous strictly increasing solutions of the autodistributivity equations (2.18) (see also Aczél and Dhombres [7, Chap. 17]).

Theorem 3.1.6 *$M \in A_2(E,E)$ fulfils (Co, SIn, AD) if and only if there exists a continuous strictly monotonic function $f : E \to \mathbb{R}$ and an arbitrary constant $\omega \in]0,1[$ such that*

$$M(x,y) = f^{-1}[(1-\omega)\,f(x) + \omega\,f(y)], \quad (x,y) \in E^2.$$

Of course, by adding (Sy) to the previous theorem, we obtain the two-place quasi-arithmetic means. This indicates the equivalence between (B) and (AD) for strictly increasing means.

As Fuchs [73] has shown, the results and proofs of Theorems 3.1.1 and 3.1.5 can be applied *mutatis mutandis* to arbitrary completely ordered sets. In order to avoid arguments concerning metric, the continuity property has been replaced by a property of pure algebraic character.

3.1.2 Non-strict quasi-arithmetic means

We now generalize Theorem 3.1.1 by relaxing (SIn) into (In). Thus, we describe the class of operators fulfilling (Sy, Co, In, Id, B). These operators will be called *non-strict quasi-arithmetic means*. Contrary to the class of quasi-linear functions, this family of operators has a rather intricate structure. It is very similar to that of *ordinal sums* (see Sect. 3.3.3) well-known in the theory of semigroups, see e.g. [110, 128]. Note that all the results we present in this section have been published in Fodor and Marichal [67].

We shall confine ourselves to functions with *two* variables. The case of n variables remains an open problem. Moreover, we will assume that E is a closed real interval $[a,b]$.

Lemma 3.1.1 *If $M \in A_2([a,b],[a,b])$ fulfils (Sy, Co, In, Id, B) then the following conditions are equivalent:*

$$\begin{array}{rl} i) & M(a,x) < x < M(x,b) \quad \forall x \in]a,b[\\ ii) & x < M(x,y) < y \quad \forall x,y \in]a,b[, x<y \\ iii) & M \text{ fulfils (SIn) on }]a,b[. \end{array}$$

Proof. *ii)* $\Rightarrow$ *i)*. For all $x \in]a,b[$, there exist $u,v \in]a,b[$ such that $a<u<x<v<b$. From *ii)* we have $M(a,x) \le M(u,x) < x < M(x,v) \le M(x,b)$.

i) $\Rightarrow$ *ii)*. Assume first that there exist $x_0, y_0 \in]a,b[, x_0 < y_0$ such that $M(x_0,y_0) = y_0$. Define

$$X := \{x \in [a,b] \,|\, x \le y_0 \text{ and } M(x,y_0) = y_0\}.$$

On the one hand, it is clear that $X \neq \emptyset$ since $x_0 \in X$. On the other hand, by (Co), X is closed. Introducing $x^* := \inf X$, we have $a < x^* \le x_0 < y_0$ since, from *i)*, we have $a \notin X$. Moreover, by (In), we have $[x^*, y_0] \subseteq X$. We should have $x^* > M(a,y_0)$. Indeed, if $x^* \le M(a,y_0)$, then, since $M(a,y_0) < y_0$ by hypothesis, we have $M(a,y_0) \in X$, that is $M(M(a,y_0),y_0) = y_0$ and

$$\begin{aligned} M(M(a,x^*),y_0) &\overset{\text{(Id)}}{=} M(M(a,x^*),M(y_0,y_0)) \overset{\text{(B)}}{=} M(M(a,y_0),M(x^*,y_0)) \\ &= M(M(a,y_0),y_0) = y_0. \end{aligned}$$

Since $M(a,x^*) \le x^* < y_0$, we have $M(a,x^*) \in X$ and, by the definition of x^*, we have $M(a,x^*) = x^*$, which contradicts *i)*.

It follows that $M(a,y_0) < x^* < y_0 = M(x^*,y_0)$ and, by (Co), there exists $z \in (a,x^*)$ such that $x^* = M(z,y_0)$. Consequently, we have, using (Id) and (B),

$$M(M(z,x^*),y_0) = M(M(z,x^*),M(y_0,y_0)) = M(M(z,y_0),M(x^*,y_0)) = M(x^*,y_0) = y_0.$$

Since $M(z,x^*) \le x^* < y_0$, we have $M(z,x^*) \in X$ and, by the definition of x^*, we have $M(z,x^*) = x^*$. Finally, we have, using (Id) and (B),

$$\begin{aligned} x^* &= M(x^*,x^*) = M(x^*,M(z,y_0)) = M(M(x^*,x^*),M(z,y_0)) \\ &= M(M(x^*,z),M(x^*,y_0)) = M(x^*,y_0) = y_0, \end{aligned}$$

a contradiction. Consequently, we have $M(x,y) < y$ for all $x,y \in]a,b[, x<y$. One can prove in a similar way that $x < M(x,y)$.

ii) $\Leftrightarrow$ *iii)*. Aczél has proved that, under the assumptions of this lemma, the condition *ii)* is equivalent to

$$M(x,y) = f^{-1}\Big[\frac{f(x)+f(y)}{2}\Big] \quad \forall x,y \in]a,b[$$

where f is any continuous strictly monotonic function on $]a,b[$ (see [4, pages 281–284]), which is sufficient. ∎

Before stating the following important result we need to introduce some subfamilies. For all $\theta \in [a,b]$, we define $\mathcal{B}_{a,b,\theta}$ as the set of operators $M \in A_2([a,b],[a,b])$ which fulfil (Sy, Co, In, Id, B) and such that $M(a,b)=\theta$. The extreme cases $\mathcal{B}_{a,b,a}$ and $\mathcal{B}_{a,b,b}$ will play an important role in the sequel. We can notice that $\min \in \mathcal{B}_{a,b,a}$ and $\max \in \mathcal{B}_{a,b,b}$.

Theorem 3.1.7 *$M \in A_2([a,b],[a,b])$ fulfils (Sy, Co, In, Id, B) if and only if there exist two numbers α and β fulfilling $a \le \alpha \le \beta \le b$, two operators $M_{a,\alpha,\alpha} \in \mathcal{B}_{a,\alpha,\alpha}$ and $M_{\beta,b,\beta} \in \mathcal{B}_{\beta,b,\beta}$, and a continuous strictly monotonic function $f : [\alpha,\beta] \to \mathbb{R}$ such that, for all $x,y \in E$,*

$$M(x,y) = \begin{cases} M_{a,\alpha,\alpha}(x,y) & \text{if } x,y \in [a,\alpha], \\ M_{\beta,b,\beta}(x,y) & \text{if } x,y \in [\beta,b], \\ f^{-1}\Big[\dfrac{f[\text{median}(\alpha,x,\beta)] + f[\text{median}(\alpha,y,\beta)]}{2}\Big] & \text{otherwise.} \end{cases}$$

Proof. (Sufficiency) Indeed, we can easily show that the operators M defined in the statement fulfil (Sy, Co, In, Id, B).

(Necessity) Assume that $M \in A_2([a,b],[a,b])$ fulfils (Sy, Co, In, Id, B). Define

$$X_a := \{x \in [a,b] \mid M(a,x) = x\} \quad \text{and} \quad X_b := \{x \in [a,b] \mid M(x,b) = x\}.$$

On the one hand, it is clear that $X_a \neq \emptyset$ and $X_b \neq \emptyset$ since $a \in X_a$ and $b \in X_b$. On the other hand, by (Co), X_a and X_b are closed. Introducing $\alpha := \sup X_a$ and $\beta := \inf X_b$, we have $\alpha \le \beta$, otherwise we would have

$$M(a,b) \ge M(a,\alpha) = \alpha > \beta = M(\beta,b) \ge M(a,b),$$

a contradiction.

Let $x,y \in [a,b]$. If $x,y \in [a,\alpha]$, then we have $M(x,y) = M_{a,\alpha,\alpha}(x,y)$, where $M_{a,\alpha,\alpha} \in \mathcal{B}_{a,\alpha,\alpha}$. Likewise, if $x,y \in [\beta,b]$, then we have $M(x,y) = M_{\beta,b,\beta}(x,y)$, where $M_{\beta,b,\beta} \in \mathcal{B}_{\beta,b,\beta}$. Otherwise, we have two mutually exclusive cases:

- If $\alpha = \beta$, then we have

  $$\alpha = M(a,\alpha) \le M(x,y) \le M(\alpha,b) = \alpha,$$

  that is $M(x,y) = \alpha$.

- If $\alpha < \beta$, then we have

$$M(a,y) = M(\alpha,y) \quad \forall y \in [\alpha, M(\alpha,b)], \tag{3.5}$$
$$M(x,b) = M(x,\beta) \quad \forall x \in [M(a,\beta),\beta]. \tag{3.6}$$

Indeed, if $y \in [\alpha, M(\alpha,b)]$ then, by (Co), there exists $z \in]\alpha, b[$ such that $y = M(\alpha,z)$. So, we have

$$\begin{aligned} M(a,y) &= M(M(a,a),M(\alpha,z)) = M(M(a,\alpha),M(a,z)) \\ &= M(M(\alpha,\alpha),M(a,z)) = M(M(\alpha,a),M(\alpha,z)) \\ &= M(\alpha,y), \end{aligned}$$

which proves (3.5). We can show that (3.6) is true by using the same argument.

Moreover, we have

$$M(\alpha,\beta) = M(\alpha,b) = M(a,\beta) = M(a,b). \tag{3.7}$$

Indeed, setting $\theta := M(a,b)$, we have

$$\alpha = M(a,\alpha) \le M(a,\beta) \le \theta \le M(\alpha,b) \le M(\beta,b) = \beta$$

and we can apply (3.5) and (3.6). Therefore, we have

$$\begin{aligned} \theta &= M(M(a,b),\theta) = M(M(a,\theta),M(\theta,b)) = M(M(\alpha,\theta),M(\theta,\beta)) \\ &= M(M(\alpha,\beta),\theta) = M(M(a,\alpha),M(\beta,b)) = M(\alpha,\beta), \end{aligned}$$

and

$$M(\alpha,b) = M(M(a,\alpha),b) = M(M(\alpha,b),\theta) = M(M(\alpha,b),M(\alpha,\beta)) = M(\alpha,\beta),$$

and

$$M(a,\beta) = M(a,M(\beta,b)) = M(\theta,M(a,\beta)) = M(M(\alpha,\beta),M(a,\beta)) = M(\alpha,\beta),$$

which proves (3.7).

We also have

$$M(a,x) = M(\alpha,x) \quad \forall x \in [\alpha,\beta], \tag{3.8}$$
$$M(x,b) = M(x,\beta) \quad \forall x \in [\alpha,\beta]. \tag{3.9}$$

By (3.5)–(3.7), it suffices to prove that $M(a,x) = M(\alpha,x)$ for all $x \in [\theta,\beta]$, and $M(x,b) = M(x,\beta)$ for all $x \in [\alpha,\theta]$.

M is continuous, thus for any $x \in [\theta,\beta]$ there exists $z \in [a,b]$ such that $x = M(\beta,z)$. Thus we have

$$\begin{aligned} M(a,x) &= M(a,M(\beta,z)) = M(M(a,\beta),M(a,z)) \\ &= M(M(\alpha,\beta),M(a,z)) = M(M(\beta,z),\alpha) \\ &= M(\alpha,x) \end{aligned}$$

which proves (3.8). We can prove (3.9) similarly.

For any $x \le \alpha$, $y \ge \beta$ we do have $M(x,y) = \theta$. Indeed, from (3.7), we have $\theta = M(a,\beta) \le M(x,y) \le M(\alpha,b) = \theta$.

Finally, by Theorem 3.1.1 and Lemma 3.1.1, it suffices to show that $M(\alpha,x) < x < M(x,\beta)$ for all $x \in]\alpha,\beta[$. Suppose the first inequality is not true. Then, from (3.8), there exists $x \in]\alpha,\beta[$ such that $M(a,x) = M(\alpha,x) = x$, which contradict the definition of α. We can prove the second inequality in a similar way. ■

As we can see, the previous characterization partitions the definition set $[a, b]^2$ into at most nine cells. On each one of them, M takes a well-defined form. Figure 3.1 presents graphics showing the particular case of $f(x) = x$ (non-strict arithmetic mean). Out of the two extreme cells, which correspond to the families $\mathcal{B}_{a,\alpha,\alpha}$ and $\mathcal{B}_{\beta,b,\beta}$, we find the arithmetic mean with arguments depending on each cell itself. This comes from the presence of the *median* function in the expression of M. For example, if $x \in [\beta, b]$ and $y \in [\alpha, \beta]$ then we get the arithmetic mean for the arguments β and y. In order to obtain a clear and readable three-dimensional representation, we have chosen the *min* and *max* operators in the extreme cells.

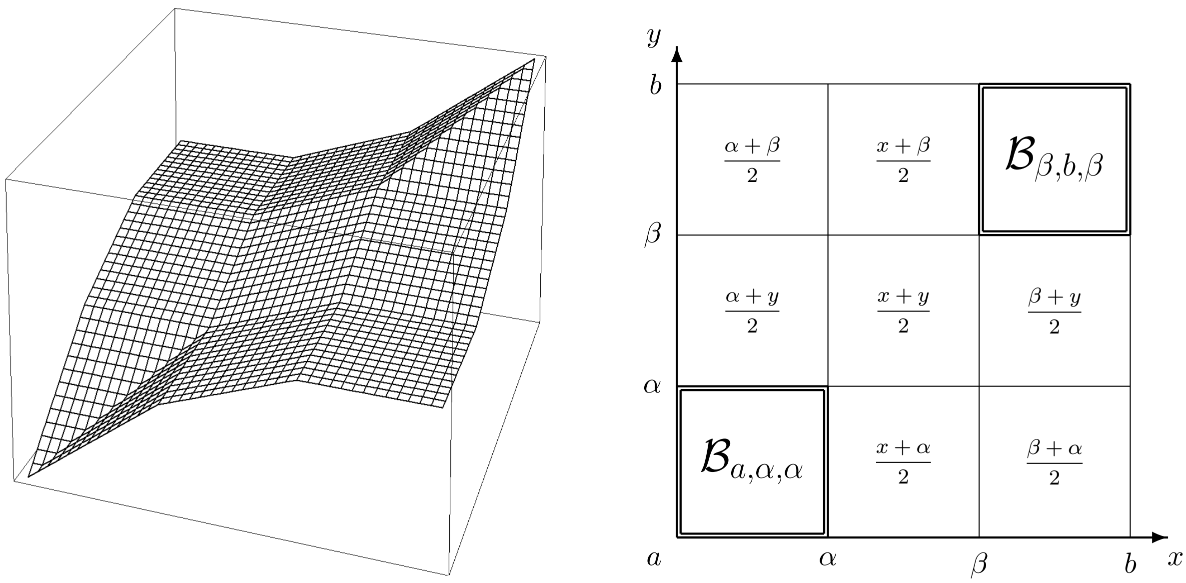


Figure 3.1: Example of non-strict arithmetic mean

Now, our task consists in describing the two families $\mathcal{B}_{a,\alpha,\alpha}$ and $\mathcal{B}_{\beta,b,\beta}$. Because of (Id), they can be assimilated with $\mathcal{B}_{a,b,b}$ and $\mathcal{B}_{a,b,a}$ respectively, simply by the help of a redefinition of the bounds of the intervals $[a, \alpha]$ and $[\beta, b]$.

Before going on, let us consider a lemma.

Lemma 3.1.2 *$M \in \mathcal{B}_{a,b,a}$ (resp. $\mathcal{B}_{a,b,b}$) is strictly increasing on $]a, b[^2$ if and only if there exists a continuous strictly increasing (resp. decreasing) function $g : [a, b] \to \mathbb{R}$, with $g(a) = 0$ (resp. $g(b) = 0$), such that*

$$M(x, y) = g^{-1}\Big[\sqrt{g(x)\, g(y)}\,\Big], \quad (x, y) \in [a, b]^2. \tag{3.10}$$

Proof. Let us consider the case $\mathcal{B}_{a,b,a}$, the other one is symmetric.

(Sufficiency) Easy.

(Necessity) Let $M \in \mathcal{B}_{a,b,a}$ be strictly increasing on $]a, b[^2$. From Theorem 3.1.1, there exists a function f which is continuous and strictly monotonic on $]a, b[$, such that

$$2f(M(x, y)) = f(x) + f(y) \quad \forall x, y \in]a, b[. \tag{3.11}$$

Replacing f by $-f$, if necessary, we can assume that f is strictly increasing on $]a,b[$. By the continuity of M, we have

$$\lim_{x\to a^+} M(x,y) = M(a,y) = a \quad \forall y \in]a,b[.$$

Then assume that $\lim_{x\to a^+} f(x) = r \in \mathbb{R}$. From (3.11), we have $f(y) = r$ for all $y \in]a,b[$, which is impossible since f is strictly increasing on $]a,b[$. Therefore, we have $\lim_{x\to a^+} f(x) = -\infty$.

From (3.11), we also have $\lim_{y\to b^-} f(y) \in \mathbb{R}$. Then let $g(x)$ be the continuous extension of the function $\exp f(x)$ on $[a,b]$, that is, $g(a) = 0$ and $g(x) = \exp f(x)$ on $]a,b]$. The function g thus defined is continuous and strictly increasing on $[a,b]$ and (3.11) becomes

$$\ln g[M(x,y)] = \frac{\ln g(x) + \ln g(y)}{2} \quad \forall x,y \in]a,b]$$

and so we have

$$M(x,y) = g^{-1}\Big[\sqrt{g(x)\,g(y)}\,\Big]$$

on $]a,b]^2$ and even on $[a,b]^2$ since M is continuous. ■

Now, we can present a description of the two families $\mathcal{B}_{a,b,a}$ and $\mathcal{B}_{a,b,b}$. Let us start with the first one. It corresponds to the cell in the upper right-hand corner in the partition of the definition set $[a,b]^2$ (see Figure 3.1), for which we have the boundary condition $M(a,b) = a$. The *quasi-geometric means* (3.10) play a central role in this description.

Theorem 3.1.8 *We have $M \in \mathcal{B}_{a,b,a}$ if and only if*

- *either*
$$M(x,y) = \min(x,y) \quad \forall x,y \in [a,b];$$
- *or there exists a continuous strictly increasing function $g : [a,b] \to \mathbb{R}$, with $g(a) = 0$, such that*
$$M(x,y) = g^{-1}\Big[\sqrt{g(x)\,g(y)}\,\Big] \quad \forall x,y \in [a,b];$$
- *or there exist a countable index set $K \subseteq \mathbb{N}$, a family of disjoint open subintervals $\{]a_k,b_k[\,|\,k \in K\}$ of $[a,b]$ and a family $\{g_k\,|\,k \in K\}$ of continuous strictly increasing functions $g_k : [a_k,b_k] \to \mathbb{R}$, with $g_k(a_k) = 0$, such that, for all $x,y \in [a,b]$,*

$$M(x,y) = \begin{cases} g_k^{-1}\Big[\sqrt{g_k[\min(x,b_k)]\,g_k[\min(y,b_k)]}\,\Big] & \textit{if } \exists k \in K \textit{ such that } \min(x,y) \in]a_k,b_k[; \\ \min(x,y) & \textit{otherwise.} \end{cases}$$

Proof. (Sufficiency) One can easily check that the operators M defined in the statement belong to $\mathcal{B}_{a,b,a}$.

(Necessity) Let $x,y \in [a,b]$ and $M \in \mathcal{B}_{a,b,a}$. Define a set $X \subseteq [a,b]$ by

$$X := \{x \in [a,b]\,|\,M(x,b) = x\}.$$

It is clear that X is closed and non-empty. Thus $Y := [a,b] \setminus X$ is open and bounded. In fact $Y = \emptyset$ if and only if $M(x,b) = x$ for all $x \in [a,b]$, that is

$$M(x,y) = \min(x,y)$$

since assuming $x \leq y$, with $x, y \in [a, b]$, we have $M(x, y) \leq M(x, b) = x = M(x, x) \leq M(x, y)$ and hence $M(x, y) = x$.

In the other extreme case we have $Y =]a, b[$, that is $X = \{a, b\}$, if and only if $x < M(x, b)$ for all $x \in]a, b[$. However $M(a, a) = a$ and $M(a, b) = a$ imply $M(a, x) = a < x$ for all $x \in]a, b[$. It follows, from Lemma 3.1.1 that $M(x, y)$ is strict on $]a, b[^2$ and from Lemma 3.1.2 that

$$M(x, y) = g^{-1}\Big[\sqrt{g(x)\, g(y)}\,\Big],$$

where g is any continuous strictly increasing function on $[a, b]$, with $g(a) = 0$.

Consider the remaining case, that is $\emptyset \subsetneq Y \subsetneq]a, b[$. Then there exists a countable index set $K \subseteq \mathbb{N}$ and a class of pairwise disjoint open intervals $\{]a_k, b_k[\,|\, k \in K\}$ of $[a, b]$ such that

$$Y = \bigcup_{k \in K}]a_k, b_k[.$$

For all $k \in K$, we obviously have $M(a_k, b) = a_k$ and $M(b_k, b) = b_k$ since $a_k, b_k \in X$, but also

$$M(x, b) > x \quad \forall x \in (a_k, b_k), \tag{3.12}$$
$$M(a_k, x) = a_k \quad \forall x \in [a_k, b], \tag{3.13}$$
$$M(b_k, x) = b_k \quad \forall x \in [b_k, b]. \tag{3.14}$$

To establish (3.12), we can notice that $x \in]a_k, b_k[$ implies $x \notin X$. For (3.13) and (3.14), we obviously have

$$a_k = M(a_k, a_k) \leq M(a_k, x) \leq M(a_k, b) = a_k, \quad \forall x \in [a_k, b],$$
$$b_k = M(b_k, b_k) \leq M(b_k, x) \leq M(b_k, b) = b_k, \quad \forall x \in [b_k, b].$$

Then we can see that, if $\min(x, y) \in X$, then

$$M(x, y) = \min(x, y).$$

If $\min(x, y) \in Y$, that is $\min(x, y) \in]a_k, b_k[$ for one $k \in K$, then, assuming that $x \in]a_k, b_k[$ and $y \in [b_k, b]$, we have

$$M(x, y) = M(x, b_k). \tag{3.15}$$

Indeed, since from (3.13), we have $M(a_k, b_k) = a_k$ and since $M(b_k, b_k) = b_k$, then, by continuity of M, there exists $z \in (a_k, b_k)$ such that $x = M(z, b_k)$. Then, from (3.14) we have

$$\begin{aligned} M(x, y) &= M(M(z, b_k), M(y, y)) = M(M(z, y), M(b_k, y)) = M(M(z, y), b_k) \\ &= M(M(z, y), M(b_k, b_k)) = M(M(z, b_k), M(y, b_k)) = M(x, b_k). \end{aligned}$$

Now, we can show that if $x, y \in]a_k, b_k[$, then

$$M(x, y) = g_k^{-1}\Big[\sqrt{g_k(x)\, g_k(y)}\,\Big]$$

where g_k is any continuous strictly increasing function on $[a_k, b_k]$, with $g_k(a_k) = 0$. It is sufficient, from Lemmas 3.1.1 and 3.1.2, to show that

$$M(a_k, x) < x < M(x, b_k), \quad \forall x \in]a_k, b_k[.$$

The first inequality comes from (3.13). For the second one, we notice that if $x = M(x, b_k)$ for one $x \in]a_k, b_k[$ then, from (3.15), we would have $x = M(x, b_k) = M(x, b)$, which contradicts (3.12). ■

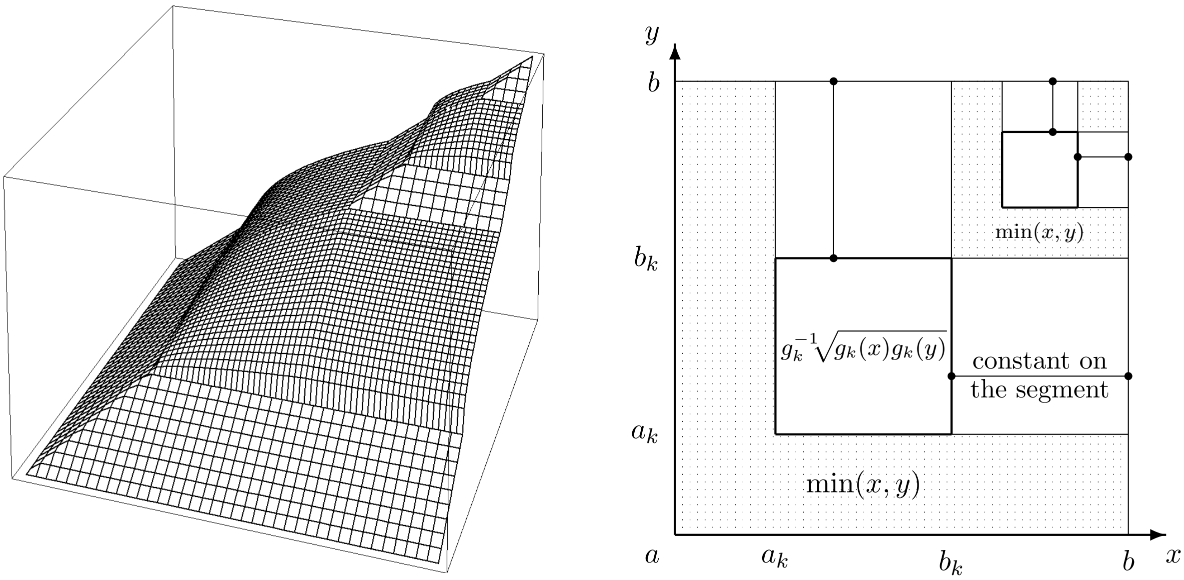


Figure 3.2: Description of $\mathcal{B}_{a,b,a}$

We can note that, in the previous result, the third possibility includes the first two. Indeed, we get the first case if no subinterval $]a_k, b_k[$ is considered and we get the second one if only one subinterval is considered and if it corresponds to $]a, b[$. Consequently, only the third case could have been presented, the first two cases being simply degenerations of the third.

Figure 3.2 represents an example from the third case. As we can see, there exists a partition of $[a, b]$ in disjoint open subintervals $]a_k, b_k[$ that divide the definition set $[a, b]^2$ into several pieces. If, given $(x, y) \in [a, b]^2$, there exists an index $k \in K$ for which $\min(x, y)$ lies in $]a_k, b_k[$ then (x, y) is in one of the unshaded regions. A quasi-geometric mean $g_k^{-1}\sqrt{g_k(x)g_k(y)}$ is defined in the central square of this region. Moreover, due to the presence of the *min* function in the expression of M, there are constant values on each horizontal or vertical segment going from the edge of the square to the extremity of the definition set. Finally, the *min* function is defined in all the shaded regions.

Now, let us turn to the second family $\mathcal{B}_{a,b,b}$ that corresponds to the boundary condition $M(a, b) = b$. The next theorem presents a description very similar to the previous one.

Theorem 3.1.9 *We have $M \in \mathcal{B}_{a,b,b}$ if and only if*

- *either*
$$M(x, y) = \max(x, y) \quad \forall x, y \in [a, b];$$
- *or there exists a continuous strictly decreasing function $g : [a, b] \to \mathbb{R}$, with $g(b) = 0$, such that*
$$M(x, y) = g^{-1}\Big[\sqrt{g(x)\, g(y)}\Big] \quad \forall x, y \in [a, b];$$
- *or there exist a countable index set $K \subseteq \mathbb{N}$, a family of disjoint open subintervals $\{]a_k, b_k[\,|\, k \in K\}$ of $[a, b]$ and a family $\{g_k \,|\, k \in K\}$ of continuous strictly decreasing func-*

tions $g_k : [a_k, b_k] \to \mathbb{R}$*, with* $g_k(b_k) = 0$*, such that, for all* $x, y \in [a, b]$*,*

$$M(x,y) = \begin{cases} g_k^{-1}\Big[\sqrt{g_k[\max(a_k,x)]\, g_k[\max(a_k,y)]}\,\Big] & \textit{if } \exists k \in K \textit{ such that } \max(x,y) \in]a_k, b_k[; \\ \max(x,y) & \textit{otherwise.} \end{cases}$$

Figure 3.3 represents an example from the third case. We will not stress on this representation, which is very similar to the one of Figure 3.2.

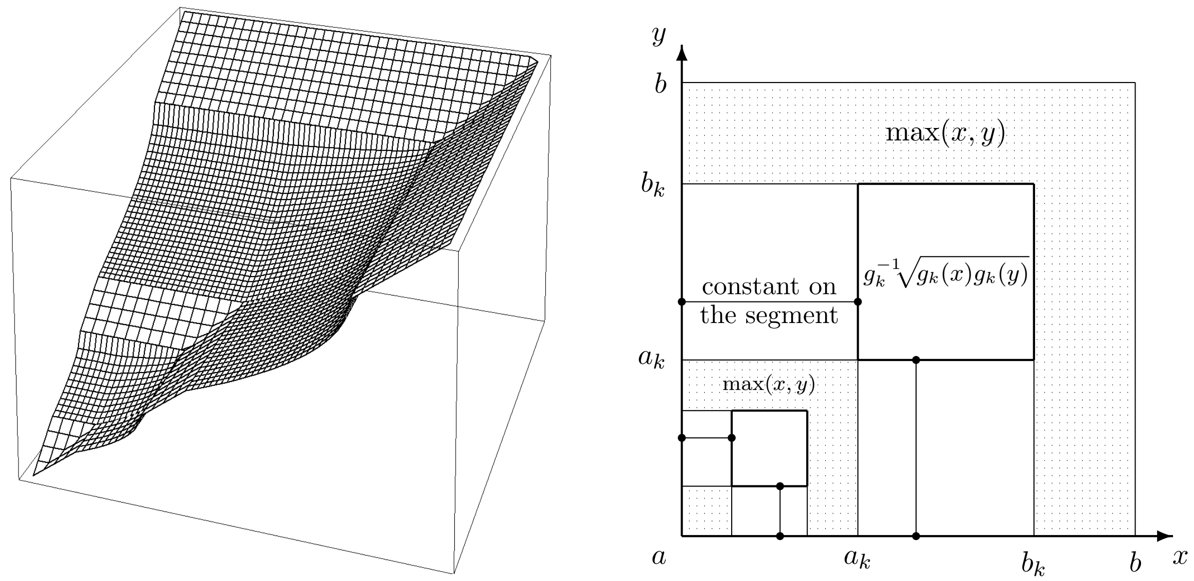


Figure 3.3: Description of $\mathcal{B}_{a,b,b}$

To conclude, we note that Theorems 3.1.7, 3.1.8 and 3.1.9, taken together, give a complete description of the non-strict quasi-arithmetic means with two variables, defined from the bisymmetry property.

3.2 Decomposable extended operators

3.2.1 Quasi-arithmetic means

Let E be a real interval, finite or infinite. Kolmogoroff [107] and Nagumo [136], in their pioneering work, considered the class of extended aggregation operators fulfilling (Sy, Co, SIn, Id, D), also called *mean values*. They established, independently of each other, the following result.

Theorem 3.2.1 *$M \in A(E, E)$ fulfils (Sy, Co, SIn, Id, D) if and only if there exists a continuous strictly monotonic function $f : E \to \mathbb{R}$ such that, for all $n \in \mathbb{N}_0$,*

$$M^{(n)}(x) = f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big], \quad x \in E^n.$$

Theorem 3.2.1 gives a characterization of the class of quasi-arithmetic means (see Section 3.1.1) defined for *all* and not for *any fixed* number of variables. It turns out, according to

Theorems 3.1.1 and 3.2.1, that the (B) property is an adequate substitute of (D) for defining quasi-arithmetic means with a definite number of variables. Note that the connection between (B) and (D) for quasi-arithmetic means was discussed by Horváth [102].

Now, suppose $n \in \mathbb{N}_0$ fixed and consider the situation where each value x_i, $i \in N_n$, is *weighted* by a non-negative rational number ω_i (also called degree of significance of i). If these weights $(\omega_1, \ldots, \omega_n)$ are given according to a ratio scale, they are not univoqually estimated but are such that any other system of acceptable weights $(\omega'_1, \ldots, \omega'_n)$ corresponds to

$$\omega'_i = C\,\omega_i \quad \forall i \in N_n, \quad C \text{ is a strictly positive rational number.}$$

$(\omega_1, \ldots, \omega_n)$ can be modified using a similarity transformation into

$$\omega'_i = \frac{\omega_i}{\sum_i \omega_i} \quad (\sum_i \omega'_i = 1)$$

or

$$p_i = q\,\omega'_i,$$

where $p_i, q \in \mathbb{N}$, $q \neq 0$, $\sum_i p_i = q$.

It is possible to simulate the effect of a quasi-linear mean (3.3) with the weights $(\omega_1, \ldots, \omega_n)$ only by using a quasi-arithmetic mean M in the following way:

$$M(p_1 \odot x_1, \ldots, p_n \odot x_n) = f^{-1}\Big[\frac{1}{q}\sum_i p_i\, f(x_i)\Big] = f^{-1}\Big[\sum_i \omega'_i\, f(x_i)\Big].$$

Now, we show that (Sy) is unnecessary in Theorem 3.2.1, provided that decomposability is considered in its general form (SD). Thus we prove that the class of extended operators satisfying (Co, SIn, Id, SD) coincides with that of extended quasi-arithmetic mean operators. This result can also be found in Marichal [117]. From (2.12) and (2.13), we can see that this is a generalization of Theorem 3.2.1.

One could think that (SD) alone implies (Sy). Nevertheless, the non-symmetric extended operator $M = (\mathrm{P}_1^{(n)})_{n \in \mathbb{N}_0}$ fulfils (SD)[1].

Lemma 3.2.1 *If A corresponds to the matrix*

$$A = \begin{pmatrix} \theta & \theta & 0 \\ 1-\theta & 0 & \theta \\ 0 & 1-\theta & 1-\theta \end{pmatrix}, \qquad \theta \in]0,1[,$$

then

$$\lim_{i \to +\infty} A^i = \frac{1}{D}\begin{pmatrix} \theta^2 & \theta^2 & \theta^2 \\ \theta(1-\theta) & \theta(1-\theta) & \theta(1-\theta) \\ (1-\theta)^2 & (1-\theta)^2 & (1-\theta)^2 \end{pmatrix}$$

with $D = \theta^2 + \theta(1-\theta) + (1-\theta)^2$.

Proof. The eigenvalues of A correspond to the solutions of $\det(A - zI) = 0$ or

$$(z-1)[\theta(1-\theta) - z^2] = 0.$$

[1]Remember that $\mathrm{P}_1^{(n)}$ is the n-place projection associated to the first argument, see (1.4).

Three distinct eigenvalues are obtained: $z_1 = 1, z_2 = \sqrt{\theta(1-\theta)}, z_3 = -\sqrt{\theta(1-\theta)}$ and A can be diagonalized:

$$\Delta = S^{-1}AS = \mathrm{diag}\Big(1, \sqrt{\theta(1-\theta)}, -\sqrt{\theta(1-\theta)}\,\Big).$$

We also have the following eigenvectors:

$$S_1 = \begin{pmatrix} s_{11} \\ s_{21} \\ s_{31} \end{pmatrix} = \begin{pmatrix} \theta^2 \\ \theta(1-\theta) \\ (1-\theta)^2 \end{pmatrix}, \qquad S_2 = \begin{pmatrix} s_{12} \\ s_{22} \\ s_{32} \end{pmatrix} = \begin{pmatrix} -\sqrt{\theta} \\ \sqrt{\theta} - \sqrt{1-\theta} \\ \sqrt{1-\theta} \end{pmatrix},$$

$$S_3 = \begin{pmatrix} s_{13} \\ s_{23} \\ s_{33} \end{pmatrix} = \begin{pmatrix} \sqrt{\theta} \\ -\sqrt{\theta} - \sqrt{1-\theta} \\ \sqrt{1-\theta} \end{pmatrix}.$$

A can be expressed in the form: $A = S\Delta S^{-1}$ and

$$A^i = S\Delta^i S^{-1}, \quad \forall i \in \mathbb{N}_0.$$

Finally, setting $s'_{ij} := (S^{-1})_{ij}$, we have

$$\lim_{i\to+\infty} A^i = S\Big(\lim_{i\to+\infty} \Delta^i\Big)S^{-1} = (\, s'_{11}S_1 \quad s'_{12}S_1 \quad s'_{13}S_1 \,).$$

We only have to determine $s'_{11}, s'_{12}, s'_{13}$ such that

$$(\, s'_{11} \quad s'_{12} \quad s'_{13} \,)\, S = (\, 1 \quad 0 \quad 0 \,)$$

and we can see that

$$s'_{11} = s'_{12} = s'_{13} = \frac{1}{D}.$$

■

Theorem 3.2.2 *$M \in A(E,E)$ fulfils (Co, SIn, Id, SD) if and only if there exists a continuous strictly monotonic function $f : E \to \mathbb{R}$ such that, for all $n \in \mathbb{N}_0$,*

$$M^{(n)}(x) = f^{-1}\Big[\frac{1}{n}\sum_{i=1}^{n} f(x_i)\Big], \quad x \in E^n.$$

Proof. (Sufficiency) Trivial (see Theorem 3.2.1).

(Necessity) Let $M \in A(E,E)$ fulfilling (Co, SIn, Id, SD). By Propositions 2.3.6 and 2.3.5, $M^{(2)}$ fulfils (Co, SIn, Id, B). Next, by Theorem 3.1.5, there exists a continuous strictly monotonic function $f : E \to \mathbb{R}$ and a real number $\theta \in]0,1[$ such that

$$M(x_1,x_2) = f^{-1}\Big[\theta\, f(x_1) + (1-\theta)\, f(x_2)\Big], \quad \forall x_1,x_2 \in E.$$

Define $\Omega := f(E) = \{f(x) \,|\, x \in E\}$. The extended operator $F \in A(\Omega,\Omega)$ defined by

$$F(z_1,\ldots,z_n) := f\Big[M(f^{-1}(z_1),\ldots,f^{-1}(z_n))\Big], \quad \forall z \in \Omega^n,\ \forall n \in \mathbb{N}_0,$$

also fulfils (Co, SIn, Id, SD) and is such that

$$F(z_1,z_2) = \theta\, z_1 + (1-\theta)\, z_2, \quad \forall z_1,z_2 \in \Omega. \tag{3.16}$$

Now, let us show that

$$F(z_1, z_2, z_3) = \frac{1}{D}\Big[\theta^2\, z_1 + \theta(1-\theta)\, z_2 + (1-\theta)^2\, z_3\Big], \quad \forall z_1, z_2, z_3 \in \Omega, \tag{3.17}$$

with $D = \theta^2 + \theta(1-\theta) + (1-\theta)^2$. We have successively

$$\begin{aligned} F(z_1, z_2, z_3) &= F(F(z_1, z_2), F(z_1, z_3), F(z_2, z_3)) \quad \text{(by (2.17))} \\ &= F(\theta\, z_1 + (1-\theta)\, z_2, \theta\, z_1 + (1-\theta)\, z_3, \theta\, z_2 + (1-\theta)\, z_3) \quad \text{(by (3.16))} \\ &= F((z_1, z_2, z_3)\, A) \end{aligned}$$

where A is the matrix defined in Lemma 3.2.1. By iteration, we obtain

$$\begin{aligned} F(z_1, z_2, z_3) &= F((z_1, z_2, z_3)\, A) = F((z_1, z_2, z_3)\, A^2) \\ &= F((z_1, z_2, z_3)\, A^i) \quad \forall i \in \mathbb{N}_0. \end{aligned}$$

We then have

$$\begin{aligned} F(z_1, z_2, z_3) &= \lim_{i \to +\infty} F((z_1, z_2, z_3)\, A^i) \quad \text{(constant numerical sequence)} \\ &= F((z_1, z_2, z_3) \lim_{i \to +\infty} A^i) \quad \text{(by (Co))} \\ &= F\Big(3 \odot \frac{1}{D}\Big[\theta^2\, z_1 + \theta(1-\theta)\, z_2 + (1-\theta)^2\, z_3\Big]\Big) \quad \text{(Lemma 3.2.1)} \\ &= \frac{1}{D}\Big[\theta^2\, z_1 + \theta(1-\theta)\, z_2 + (1-\theta)^2\, z_3\Big] \quad \text{(by (Id))} \end{aligned}$$

which proves (3.17).

Now we show that θ must be $1/2$. (SD) implies

$$F(z_1, z_2, z_3) = F(F(z_1, z_3), z_2, F(z_1, z_3)).$$

By (3.16) and (3.17), this identity becomes

$$\theta(1-\theta)(1-2\,\theta)(z_3 - z_1) = 0,$$

that is $\theta = 1/2$.

Consequently, $M^{(2)}$ fulfils (Sy). Moreover, by Proposition 2.3.1, M also fulfils (Sy). We then conclude by Theorem 3.2.1. ■

3.2.2 Non-strict quasi-arithmetic means

We now generalize Theorem 3.2.1 by relaxing (SIn) into (In). Thus, we investigate the class of extended operators fulfilling (Sy, Co, In, Id, D). Since its description is very similar to that given in Section 3.1.2, we will call those extended operators the *non-strict quasi-arithmetic means*. Here again, we assume that E is a closed real interval $[a, b]$. The results of this section can also be found in Fodor and Marichal [67].

Before presenting the result, we need a lemma. Its statement can be extracted from the Kolmogoroff's main proof (see [107]).

Lemma 3.2.2 *If $M \in A([a, b], [a, b])$ fulfils (Sy, Co, In, Id, D) then there exists a function $\psi : [0, 1] \to [a, b]$ which is continuous on $]0, 1[$ and increasing on $[0, 1]$, with $\psi(0) = a$ and $\psi(1) = b$, such that, for all $n \in \mathbb{N}_0$,*

$$M^{(n)}[\psi(t_1), \ldots, \psi(t_n)] = \psi\Big(\frac{1}{n}\sum_{i=1}^{n} t_i\Big), \quad t \in [0, 1]^n. \tag{3.18}$$

For all $\theta \in [a, b]$, we define $\mathcal{D}_{a,b,\theta}$ as the set of extended operators $M \in A([a, b], [a, b])$ which fulfil (Sy, Co, In, Id, D) and such that $M(a, b) = \theta$. The extreme cases $\mathcal{D}_{a,b,a}$ and $\mathcal{D}_{a,b,b}$ will play an important role in the sequel.

Theorem 3.2.3 *$M \in A([a, b], [a, b])$ fulfils (Sy, Co, In, Id, D) if and only if there exist two numbers α and β fulfilling $a \leq \alpha \leq \beta \leq b$, two extended operators $M_{a,\alpha,\alpha} \in \mathcal{D}_{a,\alpha,\alpha}$ and $M_{\beta,b,\beta} \in \mathcal{D}_{\beta,b,\beta}$, and a continuous strictly monotonic function $f : [\alpha, \beta] \to \mathbb{R}$ such that, for all $n \in \mathbb{N}_0$ and all $x \in E^n$,*

$$M(x) = \begin{cases} M_{a,\alpha,\alpha}(x) & \text{if } \max_i x_i \in [a, \alpha], \\ M_{\beta,b,\beta}(x) & \text{if } \min_i x_i \in [\beta, b], \\ f^{-1}\Big[\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} f[\text{median}(\alpha, x_i, \beta)]\Big] & \text{otherwise.} \end{cases}$$

Proof. (Sufficiency). We can easily show that M satisfies the announced properties.

(Necessity). According to Lemma 3.2.2, there exists a function $\psi : [0, 1] \to [a, b]$ which is continuous on $]0, 1[$ and increasing on $[0, 1]$, with $\psi(0) = a$ and $\psi(1) = b$, such that (3.18) holds for all $n \in \mathbb{N}_0$.

Define α and β in the following way:

$$a \leq \alpha = \lim_{t \to 0^+} \phi(t) \leq \lim_{t \to 1^-} \phi(t) = \beta \leq b.$$

Then, for all $k \in \mathbb{N}_0$, we have

$$M(k \odot a, \alpha) = \alpha \tag{3.19}$$
$$M(\beta, k \odot b) = \beta. \tag{3.20}$$

Indeed, according to Lemma 3.2.2 and by continuity of ψ and of M, we have

$$M(k \odot a, \alpha) = \lim_{t \to 0^+} M(k \odot \psi(0), \psi(t)) = \lim_{t \to 0^+} \psi\Big(\frac{t}{k+1}\Big) = \alpha$$

and

$$M(\beta, k \odot b) = \lim_{t \to 1^-} M(\psi(t), k \odot \psi(1)) = \lim_{t \to 1^-} \psi\Big(\frac{k+t}{k+1}\Big) = \beta.$$

Then let $n \in \mathbb{N}_0$ and $x \in [a, b]^n$. If $\max_i x_i \in [a, \alpha]$ then, from (3.19), we have $M(x) = M_{a,\alpha,\alpha}(x)$, where $M_{a,\alpha,\alpha} \in \mathcal{D}_{a,\alpha,\alpha}$. Likewise, if $\min_i x_i \in [\beta, b]$ then, from (3.20), we have $M(x) = M_{\beta,b,\beta}(x)$, where $M_{\beta,b,\beta} \in \mathcal{D}_{\beta,b,\beta}$. Otherwise, we have two mutually exclusive cases:

i) If $\alpha = \beta$ then, from (3.19) and (3.20), we have

$$\alpha = M((n-1) \odot a, \alpha) \leq M(x) \leq M(\alpha, (n-1) \odot b) = \alpha,$$

that is $M(x) = \alpha$.

ii) If $\alpha < \beta$ then ψ is strictly increasing on $[0, 1]$. Indeed, suppose it is not true and there exist $t_1, t_2 \in]0, 1[$, $t_1 < t_2$, such that $\psi(t_1) = \psi(t_2)$.

Then we have, for all $p, q \in \mathbb{N}, p \leq q, q \neq 0$,

$$M(p \odot \psi(t_1), (q-p) \odot \psi(0)) = M(p \odot \psi(t_2), (q-p) \odot \psi(0)),$$

that is, from Lemma 3.2.2,

$$\psi\Big(\frac{p}{q}\,t_1\Big) = \psi\Big(\frac{p}{q}\,t_2\Big).$$

Therefore, for any rational number $r \in [0,1]$, we have

$$\psi(r\,t_1) = \psi(r\,t_2),$$

which still holds, by continuity of ψ, for all real number $r \in [0,1]$. Choosing $r = t_1/t_2 \in]0,1[$, the previous equality becomes

$$\psi(r\,t_1) = \psi(t_1) = \psi(t_2).$$

By iteration, we get

$$\psi(r^m\,t_1) = \psi(t_2) \quad \forall m \in \mathbb{N}_0$$

and by continuity of ψ,

$$\alpha = \lim_{m\to+\infty} \psi(r^m\,t_1) = \psi(t_2).$$

One can show, in a similar way, that $\psi(t_1) = \beta$. Indeed we have, for all $p, q \in \mathbb{N}, p \leq q, q \neq 0$,

$$M(p \odot \psi(t_1), (q-p) \odot \psi(1)) = M(p \odot \psi(t_2), (q-p) \odot \psi(1)),$$

that is, from Lemma 3.2.2,

$$\psi(1 - r\,(1-t_1)) = \psi(1 - r\,(1-t_2))$$

for all $r \in [0,1]$. Choosing $r = (1-t_2)/(1-t_1) \in]0,1[$, the previous equality implies

$$\psi(1-(1-t_1)) = \psi(t_1) = \psi(t_2) = \psi(1 - r\,(1-t_1)).$$

By iteration and then by continuity of ψ, we get

$$\psi(t_1) = \lim_{m\to+\infty} \psi(1 - r^m\,(1-t_1)) = \beta.$$

Finally, we have $\alpha = \beta$, a contradiction. Consequently, ψ is strictly increasing on $]0,1[$ and thus on $[0,1]$. Since ψ is continuous on $]0,1[$, its inverse ψ^{-1} is defined on $]\alpha,\beta[\cup\{a,b\}$ and is continuous on $]\alpha,\beta[$.

Set $n_1, n_2, n_3 \in \mathbb{N}_0$ such that $n_1, n_3 < n$ and $n_1 + n_2 + n_3 = n$. Let us investigate the expression

$$M(x_1,\ldots,x_{n_1},y_1,\ldots,y_{n_2},z_1,\ldots,z_{n_3})$$

with $x_1,\ldots,x_{n_1} \in [a,\alpha]$, $y_1,\ldots,y_{n_2} \in]\alpha,\beta[$ and $z_1,\ldots,z_{m_3} \in [\beta,b]$.

Using (Co) and Lemma 3.2.2 successively, we have

$$\begin{aligned}
M(n_1 \odot a, y_1,\ldots,y_{n_2}, n_3 \odot \beta) &= \lim_{t\to1^-} M\Big[n_1 \odot \psi(0), \psi\psi^{-1}(y_1),\ldots,\psi\psi^{-1}(y_{n_2}), n_3 \odot \psi(t)\Big] \\
&= \lim_{t\to1^-} \psi\Big[\,\frac{n_1}{n}\,0 + \frac{1}{n}\sum_{i=1}^{n_2}\psi^{-1}(y_i) + \frac{n_3}{n}\,t\,\Big] \\
&= \psi\Big[\,\frac{n_1}{n}\,0 + \frac{1}{n}\sum_{i=1}^{n_2}\psi^{-1}(y_i) + \frac{n_3}{n}\,1\,\Big].
\end{aligned}$$

Since $n_1 < n$, this latter expression is also equal to

$$M(n_1 \odot \alpha, y_1,\ldots,y_{n_2}, n_3 \odot b)$$

and thus finally to

$$M(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}, z_1, \ldots, z_{n_3})$$

since, by (In), we have

$$\begin{aligned} M(m_1 \odot a, y_1, \ldots, y_{n_2}, n_3 \odot \beta) &\leq M(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}, z_1, \ldots, z_{n_3}) \\ &\leq M(n_1 \odot \alpha, y_1, \ldots, y_{n_2}, n_3 \odot b). \end{aligned}$$

Then, let $f(x)$ be the continuous extension on $[\alpha, \beta]$ of the function $\psi^{-1}(x)$, that is, $f(\alpha) = 0$, $f(\beta) = 1$ and $f(x) = \psi^{-1}(x)$ on $]\alpha, \beta[$. The function f is thus continuous and strictly monotonic on $[\alpha, \beta]$ and we have

$$M(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}, z_1, \ldots, z_{n_3}) = f^{-1}\Big[\frac{n_1}{n} f(\alpha) + \frac{1}{n} \sum_{i=1}^{n_2} f(y_i) + \frac{n_3}{n} f(\beta)\Big].$$

■

We now intend to describe the two families $\mathcal{D}_{a,\alpha,\alpha}$ and $\mathcal{D}_{\beta,b,\beta}$. By (Id), they can be assimilated with $\mathcal{D}_{a,b,b}$ and $\mathcal{D}_{a,b,a}$ respectively, simply by the help of a redefinition of the bounds of the intervals $[a, \alpha]$ and $[\beta, b]$.

Lemma 3.2.3 *Let E be any real interval, finite or infinite. Let $M \in A(E, E)$ fulfilling (Id, D). If $M^{(2)} = \min$ (resp. $\max$) then $M^{(n)} = \min$ (resp. $\max$) for all $n \in \mathbb{N}_0$.*

Proof. Let us proceed by induction. Suppose that $M^{(n)} = \min$ for a fixed $n \geq 2$. By (2.13) and Proposition 2.3.1, M fulfils (SD). Let $x \in E^{n+1}$ with $x_1 \leq \ldots \leq x_{n+1}$. Using twice (2.17), we simply have

$$M^{(n+1)}(x_1, \ldots, x_{n+1}) = M^{(n+1)}(x_1, \ldots, x_1, x_2) = M^{(n+1)}(x_1, \ldots, x_1) = x_1 = \min x_i.$$

The same can be done for the max operation. ■

Lemma 3.2.4 *Let $M \in A(E, E)$ fulfilling (Co, Id, D) and let $f : E \to \mathbb{R}$ be a continuous strictly monotonic function. If*

$$M^{(2)}(x, y) = f^{-1}\Big[\frac{f(x) + f(y)}{2}\Big], \quad (x, y) \in E^2$$

then, for all $n \in \mathbb{N}_0$, we have

$$M^{(n)}(x) = f^{-1}\Big[\frac{1}{n} \sum_{i=1}^{n} f(x_i)\Big], \quad x \in E^n.$$

Proof. On the one hand, by Proposition 2.3.1, M fulfils (Sy).

On the other hand, define $\Omega := f(E) = \{f(x) \,|\, x \in E\}$. The extended operator $F \in A(\Omega, \Omega)$ defined by

$$F(z_1, \ldots, z_n) := f[M(f^{-1}(z_1), \ldots, f^{-1}(z_n))], \quad \forall z \in \Omega^n, \quad \forall n \in \mathbb{N}_0,$$

also fulfils (Sy, Co, Id, D) and we have $F(z_1, z_2) = (z_1 + z_2)/2$ for all $z_1, z_2 \in \Omega$.

By using (2.17), we can prove by induction that, for all $n \in \mathbb{N}_0$,

$$F^{(n)}(z) = \frac{1}{n} \sum_{i=1}^{n} z_i, \quad z \in \Omega^n$$

(see Nagumo [136, Sect. 4]). This allows to end the proof. ■

Now, turn to the description of $\mathcal{D}_{a,b,a}$ and $\mathcal{D}_{a,b,b}$. These descriptions are very similar to that of Theorems 3.1.8 and 3.1.9.

Theorem 3.2.4 *We have $M \in \mathcal{D}_{a,b,a}$ if and only if*

- *either we have, for all $n \in \mathbb{N}_0$,*
$$M(x) = \min_i x_i \quad \forall x \in [a,b]^n,$$

- *or there exists a continuous strictly increasing function $g : [a,b] \to \mathbb{R}$, with $g(a) = 0$, such that, for all $n \in \mathbb{N}_0$,*
$$M(x) = g^{-1}\Big[\sqrt[n]{\prod_i g(x_i)} \Big] \quad \forall x \in [a,b]^n,$$

- *or there exist a countable index set $K \subseteq \mathbb{N}$, a family of disjoint open subintervals $\{]a_k, b_k[\,|\, k \in K\}$ of $[a,b]$ and a family $\{g_k \,|\, k \in K\}$ of continuous strictly increasing functions $g_k : [a_k, b_k] \to \mathbb{R}$, with $g_k(a_k) = 0$, such that, for all $n \in \mathbb{N}_0$ and all $x \in [a,b]^n$,*
$$M(x) = \begin{cases} g_k^{-1}\Big[\sqrt[n]{\prod_i g_k[\min(x_i, b_k)]} \Big], & \textit{if } \exists k \in K \textit{ such that } \min_i x_i \in]a_k, b_k[\\ \min_i x_i, & \textit{otherwise.} \end{cases}$$

Proof. (Sufficiency) One can easily check that the extended operators M defined in the statement belong to $\mathcal{D}_{a,b,a}$.

(Necessity) From (2.13) and Propositions 2.3.6 and 2.3.5, $M^{(2)} \in \mathcal{B}_{a,b,a}$ and we can use Theorem 3.1.8. Let $n \in \mathbb{N}_0$ and $x \in [a,b]^n$. We have three mutually exclusive cases.

- $M^{(2)} = \min$, and so, by Lemma 3.2.3, $M^{(n)} = \min$.

- There exists a continuous strictly increasing function $g : [a,b] \to \mathbb{R}$, with $g(a) = 0$, such that $M(x,y) = g^{-1}\sqrt{g(x)\, g(y)}$ for all $x, y \in [a,b]$. In that case, defining $f(x) := \ln g(x)$ on $]a,b]$, we have, by Lemma 3.2.4,
$$M(x) = g^{-1}\Big[\sqrt[n]{\prod_i g(x_i)} \Big]$$
for all $x \in]a,b]^n$ and even for all $x \in [a,b]^n$ since M fulfils (Co).

- There exist a countable index set $K \subseteq \mathbb{N}$, a family of disjoint subintervals $\{]a_k, b_k[\,|\, k \in K\}$ of $[a,b]$ and a family $\{g_k \,|\, k \in K\}$ of continuous strictly increasing functions $g_k : [a,b] \to \mathbb{R}$, with $g_k(a_k) = 0$, such that, for all $x, y \in [a,b]$,
$$M(x,y) = \begin{cases} g_k^{-1}\Big[\sqrt{g_k[\min(x, b_k)]\, g_k[\min(y, b_k)]}\Big] & \text{if } \exists k \in K \text{ such that } \min(x,y) \in]a_k, b_k[; \\ \min(x,y) & \text{otherwise.} \end{cases}$$
Suppose that there exists $k \in K$ such that $\min_i x_i \in]a_k, b_k[$. Then for all $j \in N_n$, we have
$$M^{(n)}(x_1, \ldots, x_j, x_{j+1}, \ldots, x_n) = M^{(n)}(x_1, \ldots, x_j, b_k, \ldots, b_k)$$
whenever $x_1, \ldots, x_j \in]a_k, b_k[$ and $x_{j+1}, \ldots, x_n \in [b_k, b]$. Indeed, if $n = 2$ then if $x \in]a_k, b_k[$ and $y \in [b_k, b]$, we have $M(x,y) = M(x, b_k)$. Suppose the result is true for n ($n \geq 2$) and

also $x_1, \ldots, x_j \in]a_k, b_k[$ and $x_{j+1}, \ldots, x_{n+1} \in [b_k, b]$, $j \in N_{n+1}$. So, using twice (2.17), we have

$$\begin{aligned} & M^{(n+1)}(x_1, \ldots, x_j, x_{j+1}, \ldots, x_{n+1}) \\ = \; & M^{(n+1)}(M^{(n)}(x_1, \ldots, x_j, b_k, \ldots, b_k), \ldots, M^{(n)}(x_2, \ldots, x_j, b_k, \ldots, b_k)) \\ = \; & M^{(n+1)}(x_1, \ldots, x_j, b_k, \ldots, b_k) \end{aligned}$$

Thus, the result is still true for $n+1$.

To end, we can use Lemma 3.2.4 to show that if $x \in]a_k, b_k[$ then

$$M(x) = g^{-1}\Big[\sqrt[n]{\prod_i g(x_i)} \Big].$$

Hence the result. ∎

Theorem 3.2.5 *We have $M \in \mathcal{D}_{a,b,b}$ if and only if*

- *either we have, for all $n \in \mathbb{N}_0$,*

$$M(x) = \max_i x_i \quad \forall x \in [a,b]^n,$$

- *or there exists a continuous strictly decreasing function $g : [a,b] \to \mathbb{R}$, with $g(b) = 0$, such that, for all $n \in \mathbb{N}_0$,*

$$M(x) = g^{-1}\Big[\sqrt[n]{\prod_i g(x_i)} \Big] \quad \forall x \in [a,b]^n,$$

- *or there exist a countable index set $K \subseteq \mathbb{N}$, a family of disjoint open subintervals $\{]a_k, b_k[\mid k \in K\}$ of $[a,b]$ and a family $\{g_k \mid k \in K\}$ of continuous strictly decreasing functions $g_k : [a_k, b_k] \to \mathbb{R}$, with $g_k(b_k) = 0$, such that, for all $n \in \mathbb{N}_0$ and all $x \in [a,b]^n$,*

$$M(x) = \begin{cases} g_k^{-1}\Big[\sqrt[n]{\prod_i g_k[\max(a_k, x_i)]} \Big], & \text{if } \exists k \in K \text{ such that } \max_i x_i \in]a_k, b_k[\\ \max_i x_i, & \text{otherwise.} \end{cases}$$

3.3 Associative operators and extended operators

Before dealing with associative operators, we will need to introduce some useful concepts: A *semigroup* (E, M) is a set E with an associative internal operation M defined on it. As usual, we will assume that E is a real interval, finite or infinite. An element $e \in E$ is

a) an *identity* for M if $M(e,x) = M(x,e) = x$ for all $x \in E$,

b) an *zero* (or *annihilator*) for M if $M(e,x) = M(x,e) = e$ for all $x \in E$,

c) an *idempotent* for M if $M(e,e) = e$.

For any semigroup (E, M), it is clear that there is at most one identity and at most one zero for M in E, and both are idempotents.

We also need to introduce the concept of *ordinal sum*, well-known in the theory of semigroups (see e.g. [29, 110]).

Definition 3.3.1 Let K be a totally ordered set and $\{(E_k, M_k) \mid k \in K\}$ be a collection of disjoint semigroups indexed by K. Then the *ordinal sum* of $\{(E_k, M_k) \mid k \in K\}$ is the set-theoretic union $\cup_{k\in K} E_k$ under the following binary operation:

$$M(x,y) = \begin{cases} M_k(x,y), & \text{if } \exists\, k \in K \text{ such that } x, y \in E_k \\ \min(x,y), & \text{if } \exists\, k_1, k_2 \in K, k_1 \neq k_2 \text{ such that } x \in E_{k_1} \text{ and } y \in E_{k_2}. \end{cases}$$

The ordinal sum is a semigroup under the above defined operation.

3.3.1 Case of strictly increasing operators

Aczél [3] investigated the general continuous, strictly increasing, real solution on E^2 of the associativity functional equation (2.10). He proved the following (see also [4, Sect. 6.2]).

Theorem 3.3.1 *Let E be a real interval, finite or infinite, which is open on one side. $M \in A_2(E, E)$ fulfils (Co, SIn, A) if and only if there exists a continuous and strictly monotonic function $f : E \to \mathbb{R}$ such that*

$$M(x,y) = f^{-1}[f(x) + f(y)], \quad (x,y) \in E^2. \tag{3.21}$$

It was also proved that the function f occuring in (3.21) is determined up to a multiplicative constant, that is, with $f(x)$ all functions $g(x) = r\, f(x)$ ($r \in \mathbb{R}_0$) belongs to the same M, and only these.

Moreover, the function f is such that, if $e \in E$ then

$$M(e,e) = e \Leftrightarrow f(e) = 0. \tag{3.22}$$

Indeed, if $M(e,e) = e$ then, by (3.21), we have $2f(e) = f(e)$, hence $f(e) = 0$. Conversely, suppose $f(e) = 0$. By (3.21), we have $0 = 2f(e) = f(M(e,e))$. Since f is strictly monotonic, we have $M(e,e) = e$.

By (3.22) and because of strict monotonicity of f, there is at most one idempotent for M (which is, actually, the identity) and hence M cannot be idempotent (Id). Therefore, there is no operator fulfilling (Co, SIn, Id, A). However, we can notice that every operator fulfilling (Co, SIn, A) satisfies (Sy). The sum ($f(x) = x$) and the product ($f(x) = \log x$) are well-known examples of continuous, strictly increasing, associative operators.

According to Ling [110], any semigroup (E, M) satisfying the hypotheses of Theorem 3.3.1 is called *Aczélian*.

Recall that each associative extended aggregation operator $M \in A(E, E)$ is uniquely determined by its two-place function. Thus, we have immediately the following result.

Corollary 3.3.1 *Let E be a real interval, finite or infinite, which is open on one side. $M \in A(E, E)$ fulfils (Co, SIn, A) if and only if there exists a continuous and strictly monotonic function $f : E \to \mathbb{R}$ such that, for all $n \in \mathbb{N}_0$,*

$$M(x) = f^{-1}\Big[\sum_{i=1}^{n} f(x_i)\Big], \quad x \in E^n.$$

3.3.2 Archimedean semigroups

Some authors tempted to generalize Theorem 3.3.1 by relaxing (SIn) into (In). But it seems that the class of operators fulfilling (Co, In, A) has not been described yet. However, under some additional conditions, results have been obtained.

First, we state a representation theorem attributed very often to Ling [110]. In fact, her main theorem can be deduced from previously known results on topological semigroups, see Faucett [59] and Mostert and Shields [128]. Nevertheless, the advantage of Ling's approach is twofold: treating two different cases in a unified manner and establishing elementary proofs.

Theorem 3.3.2 *Let $E = [a, b]$. $M \in A_2(E, E)$ fulfils (Co, In, A) and*

$$M(b, x) = x \quad \forall x \in E \tag{3.23}$$

$$M(x, x) < x \quad \forall x \in E^\circ \tag{3.24}$$

if and only if there exists a continuous strictly decreasing function $f : E \to [0, +\infty]$, with $f(b) = 0$, such that

$$M(x, y) = f^{-1}[\min(f(x) + f(y), f(a))] \quad \forall x, y \in E. \tag{3.25}$$

The requirement that E be closed is not really a restriction. If E is any real interval, finite or infinite, with right endpoint b (b can be $+\infty$), then we can replace condition (3.23) by

$$\lim_{t \to b^-} M(t, t) = b, \quad \lim_{t \to b^-} M(t, x) = x \quad \forall x \in E.$$

Any function f solving equation (3.25) is called an *additive generator* (or simply *generator*) of M. Moreover, we can easily see that any function M of the form (3.25) is symmetric (Sy) and, by Proposition 2.1.3, it is also conjunctive (Conj).

Condition (3.23) expresses that b is a *left identity* for M. It turns out, from (3.25), that b acts as an identity, and a as a zero. Condition (3.24) simply expresses that there are no idempotents for M in $]a, b[$: indeed, by (In) and (3.23), we always have $M(x, x) \leq M(b, x) = x$ for all $x \in [a, b]$.

Depending on whether $f(a)$ is finite or infinite (recall that $f(a) \in [0, +\infty]$), M takes a well-defined form (see Fodor and Roubens [70, Sect. 1.3] and Schweizer and Sklar [164]):

- $f(a) < +\infty$ if and only if M has *zero divisors* (i.e. $\exists\, x, y \in]a, b[$ such that $M(x, y) = a$). In this case, there exists a continuous strictly increasing function $g : [a, b] \to [0, 1]$, with $g(a) = 0$ and $g(b) = 1$ such that

  $$M(x, y) = g^{-1}[\max(g(x) + g(y) - 1, 0)] \quad \forall x, y \in [a, b]. \tag{3.26}$$

  To see this, it suffices to set $g(x) := 1 - f(x)/f(a)$. For associative extended aggregation operators $M \in A([a, b], [a, b])$, (3.26) becomes

  $$M(x) = g^{-1}\Big[\max\Big(\sum_{i=1}^{n} g(x_i) - n + 1, 0\Big)\Big] \quad \forall x \in [a, b]^n, \quad \forall n \in \mathbb{N}_0.$$

- $\lim_{t \to a^+} f(x) = +\infty$ if and only if M is strictly increasing on $]a, b[$. In this case, there exists a continuous strictly increasing function $g : [a, b] \to [0, 1]$, with $g(a) = 0$ and $g(b) = 1$ such that

  $$M(x, y) = g^{-1}[g(x)\, g(y)] \quad \forall x, y \in [a, b], \tag{3.27}$$

To see this, it suffices to set $g(x) := \exp(-f(x))$. For associative extended aggregation operators $M \in A([a,b],[a,b])$, (3.27) becomes

$$M(x) = g^{-1}\Big[\prod_{i=1}^{n} g(x_i)\Big] \quad \forall x \in [a,b]^n, \quad \forall n \in \mathbb{N}_0.$$

Of course, Theorem 3.3.2 can also be written under a dual form as follows.

Theorem 3.3.3 *Let $E = [a,b]$. $M \in A_2(E,E)$ fulfils (Co, In, A) and*

$$M(a,x) = x \quad \forall x \in E \tag{3.28}$$
$$M(x,x) > x \quad \forall x \in E^{\circ} \tag{3.29}$$

if and only if there exists a continuous strictly increasing function $f : E \to [0,+\infty]$, with $f(a) = 0$, such that

$$M(x,y) = f^{-1}[\min(f(x)+f(y), f(b))] \quad \forall x,y \in E. \tag{3.30}$$

Here again, E can be any real interval, even infinite. The functions M of the form (3.30) are symmetric (Sy) and disjunctive (Disj). There are no interior idempotents. The left endpoint a acts as an identity and the right endpoint b acts as a zero.

Once more, two mutually exclusive cases can be examined:

- $f(b) < +\infty$ if and only if M has zero divisors (i.e. $\exists\, x,y \in]a,b[$ such that $M(x,y) = b$). In this case, there exists a continuous strictly increasing function $g : [a,b] \to [0,1]$, with $g(a) = 0$ and $g(b) = 1$ such that

$$M(x,y) = g^{-1}[\min(g(x)+g(y),1)] \quad \forall x,y \in [a,b]. \tag{3.31}$$

  To see this, it suffices to set $g(x) := f(x)/f(b)$. For associative extended aggregation operators $M \in A([a,b],[a,b])$, (3.31) becomes

$$M(x) = g^{-1}\Big[\min\Big(\sum_{i=1}^{n} g(x_i), 1\Big)\Big] \quad \forall x \in [a,b]^n, \quad \forall n \in \mathbb{N}_0.$$

- $\lim_{t\to b^-} f(x) = +\infty$ if and only if M is strictly increasing on $]a,b[$. In this case, there exists a continuous strictly increasing function $g : [a,b] \to [0,1]$, with $g(a) = 0$ and $g(b) = 1$ such that

$$M(x,y) = g^{-1}[1-(1-g(x))\,(1-g(y))] \quad \forall x,y \in [a,b], \tag{3.32}$$

  To see this, it suffices to set $g(x) := 1-\exp(-f(x))$. For associative extended aggregation operators $M \in A([a,b],[a,b])$, (3.32) becomes

$$M(x) = g^{-1}\Big[1-\prod_{i=1}^{n}(1-g(x_i))\Big] \quad \forall x \in [a,b]^n, \quad \forall n \in \mathbb{N}_0.$$

Any semigroup fulfilling the assumptions of Theorem 3.3.2 or 3.3.3 is called *Archimedean*, see Ling [110]. In other words, any semigroup (E,M) is said to be *Archimedean* if M fulfils (Co, In, A), one endpoint of E is an identity for M, and there are no idempotents for M in E°. We can make a distinction between conjunctive and disjunctive Archimedean semigroups depending on whether the identity is the right endpoint of E or the left endpoint of E respectively. An Archimedean semigroup is called *properly Archimedean* or *Aczélian* if every additive generator f is unbounded; otherwise it is *improperly Archimedean.*

Ling [110, Sect. 6] proved that every Archimedean semigroup is obtainable as a limit of Aczélian semigroups.

3.3.3 A class of non-decreasing operators

We now intend to describe the class of operators $M \in A_2([a,b],[a,b])$ fulfilling (Co, In, WId, A). For all $\theta \in [a,b]$, we define $\mathcal{A}_{a,b,\theta}$ as the set of operators $M \in A_2([a,b],[a,b])$ which fulfil (Co, In, WId, A) and such that $M(a,b) = M(b,a) = \theta$. The extreme cases $\mathcal{A}_{a,b,a}$ and $\mathcal{A}_{a,b,b}$ will play an important role in the sequel. The results proved by the author can be found in Marichal [116].

Theorem 3.3.4 *$M \in A_2([a,b],[a,b])$ fulfils (Co, In, WId, A) if and only if there exist $\alpha, \beta \in [a,b]$ and two operators $M_{a,\alpha\wedge\beta,\alpha\wedge\beta} \in \mathcal{A}_{a,\alpha\wedge\beta,\alpha\wedge\beta}$ and $M_{\alpha\vee\beta,b,\alpha\vee\beta} \in \mathcal{A}_{\alpha\vee\beta,b,\alpha\vee\beta}$ such that, for all $x,y \in [a,b]$,*

$$M(x,y) = \begin{cases} M_{a,\alpha\wedge\beta,\alpha\wedge\beta}(x,y), & \text{if } x,y \in [a,\alpha\wedge\beta] \\ M_{\alpha\vee\beta,b,\alpha\vee\beta}(x,y), & \text{if } x,y \in [\alpha\vee\beta,b] \\ (\alpha\wedge x)\vee(\beta\wedge y)\vee(x\wedge y), & \text{otherwise.} \end{cases}$$

Proof. (Sufficiency) We can easily see that the operators M defined in the statement fulfil (Co, In, WId). The only property we have to prove is associativity.

Assume $\alpha \le \beta$ (the other case can be treated similarly) and let $x,y,z \in [a,b]$.

1. If $y,z \le \alpha$ then
   - if $x \le \alpha$ then $M = M_{a,\alpha,\alpha}$ and $M(M(x,y),z) = M(x,M(y,z))$;
   - if $x > \alpha$ then $M(M(x,y),z) = M(\alpha,z) = \alpha = M(x,M(y,z))$.
2. The case $y,z \ge \beta$ can be treated similarly.
3. In the remaining cases,
   - if $z \le \alpha$ and $y > \alpha$ then $M(x,y) \ge \alpha$ and $M(M(x,y),z) = \alpha = M(x,\alpha) = M(x,M(y,z))$;
   - if $z \ge \beta$ and $y < \beta$ then $M(x,y) \le \beta$ and $M(M(x,y),z) = \beta = M(x,\beta) = M(x,M(y,z))$;
   - if $\alpha \le z \le \beta$ then $M(M(x,y),z) = z = M(x,z) = M(x,M(y,z))$.

(Necessity) Set $\alpha := M(b,a)$, $\beta := M(a,b)$ and suppose $\alpha \le \beta$. The other case can be treated similarly.

We have

$$M(\alpha,a) = M(b,\alpha) = \alpha, \tag{3.33}$$
$$M(a,\beta) = M(\beta,b) = \beta. \tag{3.34}$$

Indeed, we have for instance, $M(\alpha,a) = M(M(b,a),a) = M(b,M(a,a)) = M(b,a) = \alpha$.

We have

$$M(x,y) = \alpha \quad \forall x,y \in [a,b], y \le \alpha \le x, \tag{3.35}$$
$$M(x,y) = \beta \quad \forall x,y \in [a,b], x \le \beta \le y, \tag{3.36}$$

Indeed, we have for instance, by (In) and (3.33), $\alpha = M(b,\alpha) \ge M(x,y) \ge M(\alpha,a) = \alpha$.

We have

$$M(a,y) = y \quad \forall y \in [a,\beta] \tag{3.37}$$
$$M(b,y) = y \quad \forall y \in [\alpha,b] \tag{3.38}$$

Indeed, for instance, if z increases from a to β, $M(a,z)$ increases continuously from a to β. Using the intermediate-value theorem, this implies that: $\forall y \in [a,\beta]$, $\exists z \in [a,\beta]$ such that $y = M(a,z)$ and

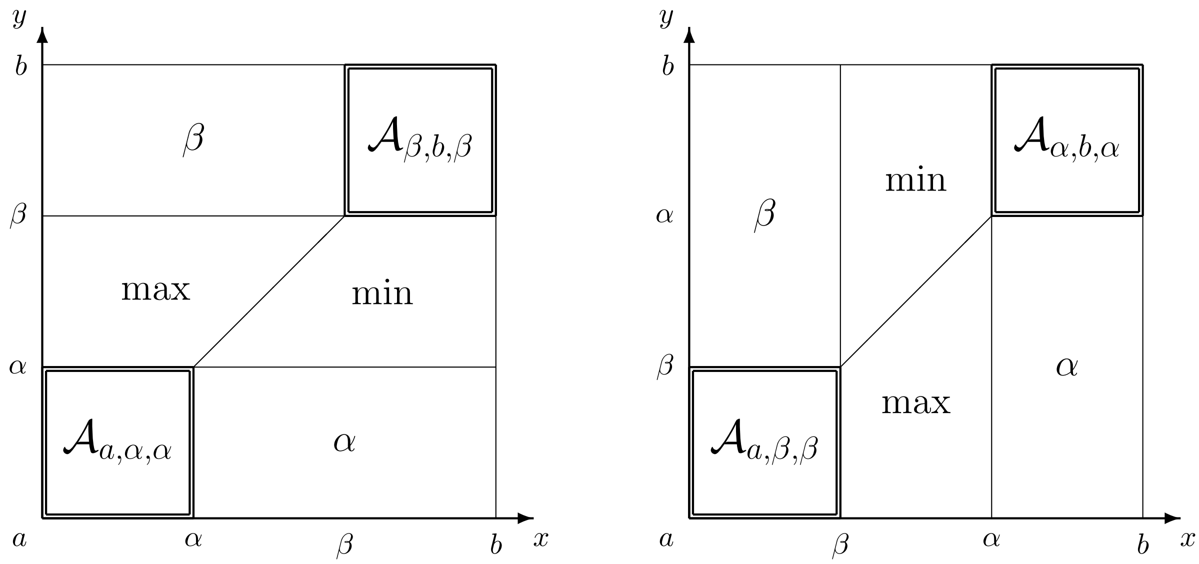
$$M(a,y) = M(a,M(a,z)) = M(M(a,a),z) = M(a,z) = y.$$

To end the proof, we have to prove that

$$M(x,y) = y \quad \forall x \in [a,b]\ \forall y \in [\alpha,\beta].$$

Indeed, by (In) and (3.37)–(3.38), we simply have $y = M(a,y) \leq M(x,y) \leq M(b,y) = y$. ∎

As we can note, the previous characterization partitions the definition set $[a,b]^2$ into several pieces. On each one of them, M takes a well defined form. Figure 3.4 presents graphics showing this partition and the corresponding values of the function.



Case of $\alpha \leq \beta$ Case of $\alpha \geq \beta$

Figure 3.4: Operators fulfilling (Co, In, WId, A) on $[a,b]^2$

Now, our task consists in describing the two families $\mathcal{A}_{a,b,a}$ and $\mathcal{A}_{a,b,b}$. For this purpose, consider a proposition.

Proposition 3.3.1 *If $M \in A_2([a,b],[a,b])$ fulfils (Co, In, A) then the following assertions are equivalent:*

i) *b is an identity for M.*
ii) *a is a zero, and b is an idempotent for M.*
iii) *$M(a,b) = M(b,a) = a$, and b is an idempotent for M.*

The assertions remain equivalent if the endpoints a and b are exchanged.

Proof. *i)* or *ii)* $\Rightarrow$ *iii)* Trivial.

iii) $\Rightarrow$ *ii)* For all $x \in [a,b]$, we have $M(a,x) \leq M(a,b) = a$, so that $M(a,x) = a$.

iii) $\Rightarrow$ *i)* If z increases from a to b, $M(b,z)$ increases continuously from a to b. Using the intermediate-value theorem, this implies that: $\forall x \in [a,b]$, $\exists z \in [a,b]$ such that $x = M(b,z)$ and

$$M(b,x) = M(b, M(b,z)) = M(M(b,b),z) = M(b,z) = x.$$

We can prove similarly that $M(x,b) = x$ for all $x \in [a,b]$. ∎

Now, let us turn to the description of $\mathcal{A}_{a,b,a}$. Mostert and Shields [128, p. 130, Theorem B] proved the following.

Theorem 3.3.5 *$M \in A_2([a,b],[a,b])$ fulfils (Co, A) and is such that a acts as a zero and b as an identity if and only if*

- *either*
$$M(x,y) = \min(x,y) \quad \forall x,y \in [a,b],$$
- *or there exists a continuous strictly decreasing function $f : [a,b] \to [0,+\infty]$, with $f(b) = 0$, such that*
$$M(x,y) = f^{-1}[\,\min(f(x) + f(y), f(a))\,] \quad \forall x,y \in [a,b].$$
(conjunctive Archimedean semigroup)
- *or there exist a countable index set $K \subseteq \mathbb{N}$, a family of disjoint open subintervals $\{]a_k,b_k[\,|\,k \in K\}$ of $[a,b]$ and a family $\{f_k\,|\,k \in K\}$ of continuous strictly decreasing function $f_k : [a_k,b_k] \to [0,+\infty]$, with $f_k(b_k) = 0$, such that, for all $x,y \in [a,b]$,*
$$M(x,y) = \begin{cases} f_k^{-1}[\,\min(f_k(x) + f_k(y), f_k(a_k))\,], & \text{if } \exists\, k \in K \text{ such that } x,y \in [a_k,b_k] \\ \min(x,y), & \text{otherwise.} \end{cases}$$
(ordinal sum of conjunctive Archimedean semigroups and one-point semigroups)

By Proposition 3.3.1, $\mathcal{A}_{a,b,a}$ is the family of operators $M \in A_2([a,b],[a,b])$ fulfilling (Co, In, A) and such that a acts as a zero and b as an identity. Consequently, the description of the family $\mathcal{A}_{a,b,a}$ is also given by Theorem 3.3.5 (see also Figure 3.5). Moreover, it turns out that all operators fulfilling the assumptions of this result satisfy (Sy, In, Conj).

Theorem 3.3.5 can also be written under a dual form as follows.

Theorem 3.3.6 *$M \in A_2([a,b],[a,b])$ fulfils (Co, A) and is such that a acts as an identity and b as a zero if and only if*

- *either*
$$M(x,y) = \max(x,y) \quad \forall x,y \in [a,b],$$
- *or there exists a continuous strictly increasing function $f : [a,b] \to [0,+\infty]$, with $f(a) = 0$, such that*
$$M(x,y) = f^{-1}[\,\min(f(x) + f(y), f(b))\,] \quad \forall x,y \in [a,b].$$
(disjunctive Archimedean semigroup)
- *or there exist a countable index set $K \subseteq \mathbb{N}$, a family of disjoint open subintervals $\{]a_k,b_k[\,|\,k \in K\}$ of $[a,b]$ and a family $\{f_k\,|\,k \in K\}$ of continuous strictly increasing function $f_k : [a_k,b_k] \to [0,+\infty]$, with $f_k(a_k) = 0$, such that, for all $x,y \in [a,b]$,*
$$M(x,y) = \begin{cases} f_k^{-1}[\,\min(f_k(x) + f_k(y), f_k(b_k))\,], & \text{if } \exists\, k \in K \text{ such that } x,y \in [a_k,b_k] \\ \max(x,y), & \text{otherwise.} \end{cases}$$
(ordinal sum of disjunctive Archimedean semigroups and one-point semigroups)

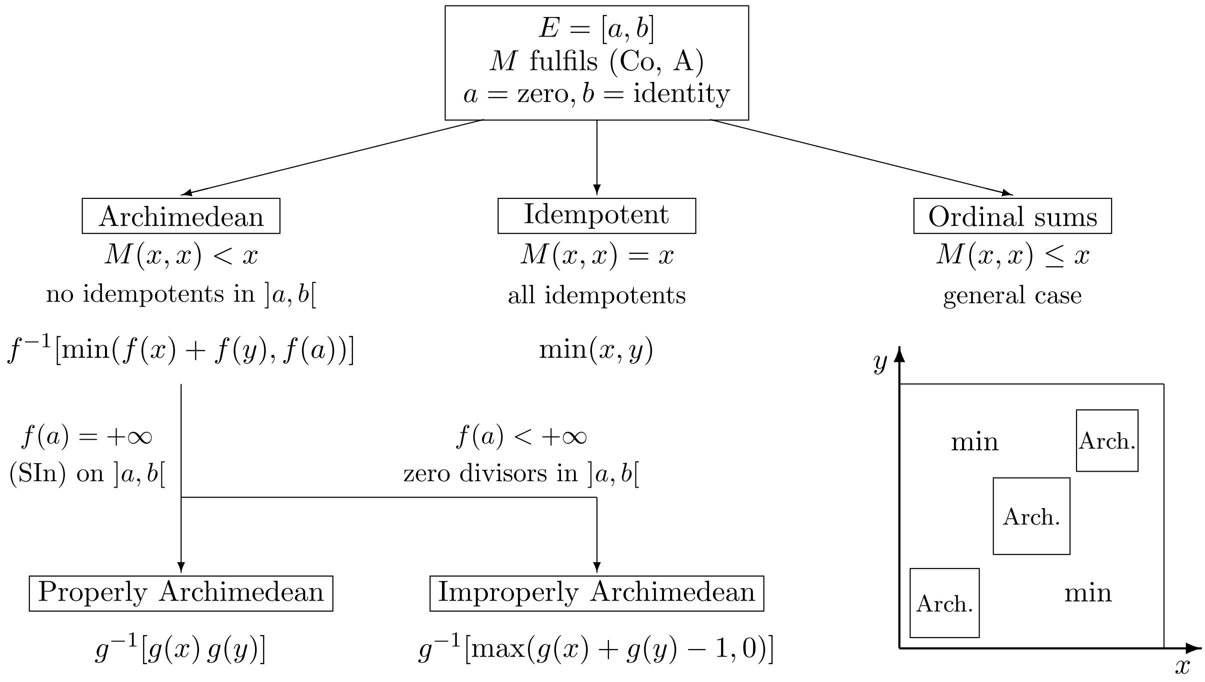


Figure 3.5: Description of $\mathcal{A}_{a,b,a}$

As above, we can see that $\mathcal{A}_{a,b,b}$ is the family of operators $M \in A_2([a, b], [a, b])$ fulfilling (Co, In, A) and such that a acts as an identity and b as a zero. The description of the family $\mathcal{A}_{a,b,b}$ is thus given by Theorem 3.3.6 (see also Figure 3.6). Moreover, all operators fulfilling the assumptions of this result satisfy (Sy, In, Disj).

Theorems 3.3.4, 3.3.5 and 3.3.6, taken together, give a complete description of the family of operators $M \in A_2([a, b], [a, b])$ fulfilling (Co, In, WId, A). Imposing some additional conditions leads to the following immediate corollaries.

Corollary 3.3.2 *$M \in A_2([a, b], [a, b])$ fulfils (Co, SIn, WId, A) if and only if there exists a continuous strictly increasing function $g : [a, b] \to [0, 1]$, with $g(a) = 0$ and $g(b) = 1$ such that*

- *either*
$$M(x, y) = g^{-1}[g(x)\, g(y)] \quad \forall x, y \in [a, b],$$
- *or*
$$M(x, y) = g^{-1}[g(x) + g(y) - g(x)\, g(y)] \quad \forall x, y \in [a, b].$$

Corollary 3.3.3 *$M \in A_2([a, b], [a, b])$ fulfils (Sy, Co, In, WId, A) if and only if there exist $\alpha \in [a, b]$ and two functions $M_{a,\alpha,\alpha} \in \mathcal{A}_{a,\alpha,\alpha}$ and $M_{\alpha,b,\alpha} \in \mathcal{A}_{\alpha,b,\alpha}$ such that, for all $x, y \in [a, b]$,*

$$M(x, y) = \begin{cases} M_{a,\alpha,\alpha}(x, y), & \text{if } x, y \in [a, \alpha] \\ M_{\alpha,b,\alpha}(x, y), & \text{if } x, y \in [\alpha, b] \\ \alpha, & \text{otherwise.} \end{cases}$$

Corollary 3.3.4 *$M \in A_2([a, b], [a, b])$ fulfils (Co, In, WId, A) and has exactly one identity element in $[a, b]$ if and only if $M \in \mathcal{A}_{a,b,a} \cup \mathcal{A}_{a,b,b}$.*

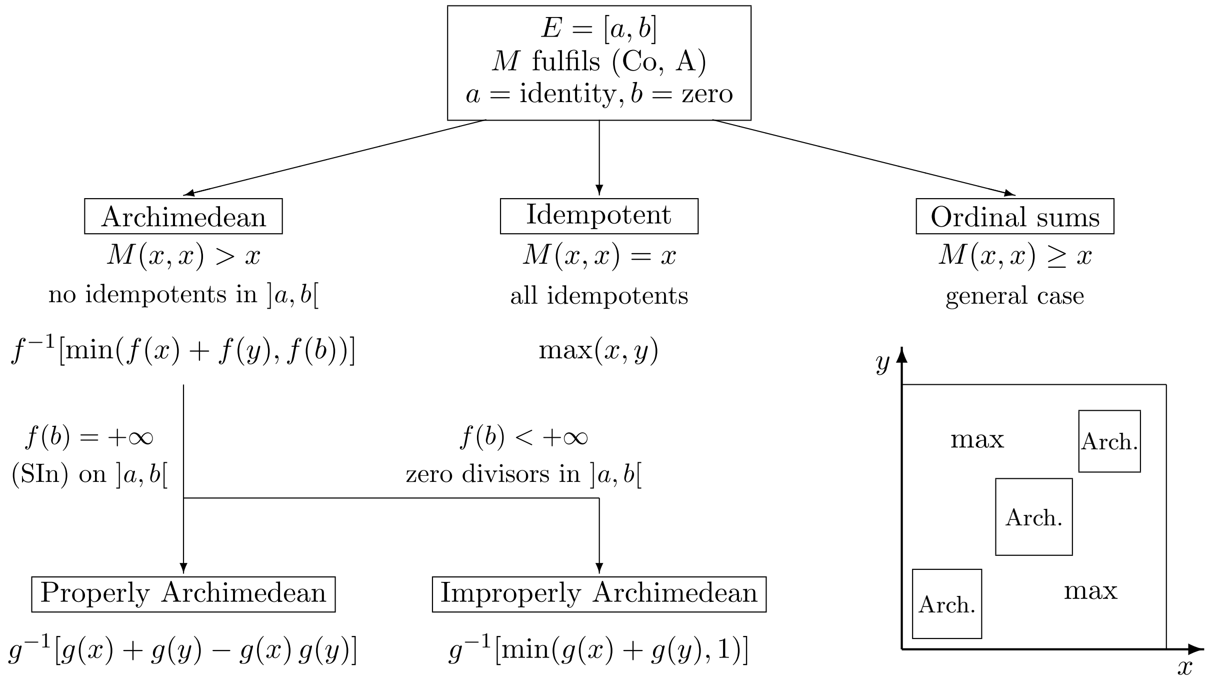


Figure 3.6: Description of $\mathcal{A}_{a,b,b}$

3.3.4 Compensative operators

Now, we investigate the case of compensative operators (Comp). By (2.1), it is equivalent to consider idempotent operators (Id). Although we have seen in Section 3.3.1 that there is no operator fulfilling (Co, SIn, Id, A), the class of operators fulfilling (Co, In, Id, A) is not empty and its description can be deduced from Theorem 3.3.4. However, Fodor [65] had already obtained this description in a more general framework. In his result, E can be any connected order topological space. In particular, E can be an arbitrary real interval, even infinite.

Theorem 3.3.7 *Let E be a real interval, finite or infinite. $M \in A_2(E, E)$ fulfils (Co, In, Id, A) if and only if there exist $\alpha, \beta \in E$ such that*

$$M(x, y) = (\alpha \wedge x) \vee (\beta \wedge y) \vee (x \wedge y), \quad (x, y) \in E^2. \tag{3.39}$$

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

$$M(x, y) = (\beta \vee x) \wedge (\alpha \vee y) \wedge (x \vee y), \quad (x, y) \in E^2.$$

On the basis of (3.39), the graphical representation of M can be illustrated (see Figure 3.7). For associative extended aggregation operators $M \in A(E, E)$, the statement can be formulated as follows.

Theorem 3.3.8 *Let E be a real interval, finite or infinite. $M \in A(E, E)$ fulfils (Co, In, Id, A) if and only if there exist $\alpha, \beta \in E$ such that*

$$M(x) = (\alpha \wedge x_1) \vee \Big(\bigvee_{i=2}^{n-1} (\alpha \wedge \beta \wedge x_i)\Big) \vee (\beta \wedge x_n) \vee \Big(\bigwedge_{i=1}^{n} x_i\Big) \quad \forall x \in E^n, \quad \forall n \in \mathbb{N}_0 \tag{3.40}$$

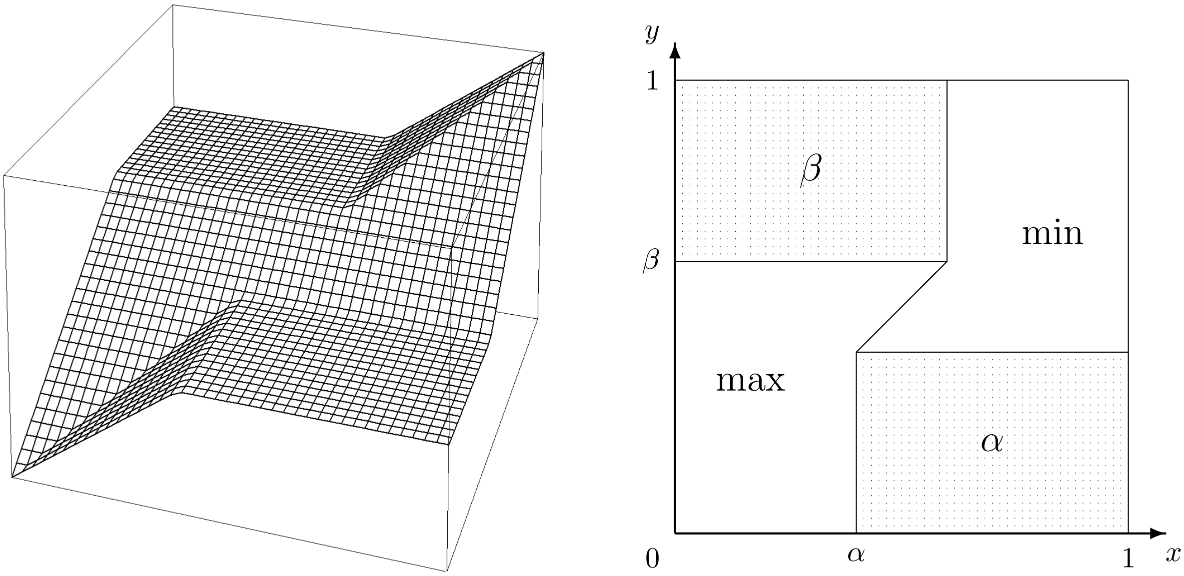


Figure 3.7: Representation on $[0,1]^2$ of $M(x,y) = (\alpha \wedge x) \vee (\beta \wedge y) \vee (x \wedge y)$ in case of $\alpha \leq \beta$

It is worth noting that there exists no weighted associative extended operators allowing to assign to each partial value x_i a positive weight ω_i. This comes from the fact that any associative extended operator is completely determined by its two-place operator. Moreover, when (Id) is assumed, no simulation similar to that used for quasi-arithmetic mean operators (see Section 3.2.1) can be used, for we have, under (Id, A),

$$M(p_1 \odot x_1, \ldots, p_n \odot x_n) = M(x_1, \ldots, x_n) \quad \forall p_1, \ldots, p_n \in \mathbb{N}_0.$$

Before Fodor [65], the symmetric case (Sy) was obtained by Fung and Fu [75] and in a revisited way by Dubois and Prade [44]. Now, the result can be formulated as follows.

Theorem 3.3.9 *Let E be a real interval, finite or infinite.*
i) $M \in A_2(E,E)$ fulfils (Sy, Co, In, Id, A) if and only if there exists $\alpha \in E$ such that

$$M(x,y) = \text{median}(x,y,\alpha) \quad \forall x,y \in E \tag{3.41}$$

ii) $M \in A(E,E)$ fulfils (Sy, Co, In, Id, A) if and only if there exists $\alpha \in E$ such that

$$M(x) = \text{median}\Big(\bigwedge_{i=1}^{n} x_i, \bigvee_{i=1}^{n} x_i, \alpha\Big) \quad \forall x \in E^n, \quad \forall n \in \mathbb{N}_0. \tag{3.42}$$

Since (Sy, A) implies (B), we immediately see that (3.41) is a particular case of non-strict arithmetic mean (see Theorem 3.1.7).

The previous three theorems show that the idempotence property is seldom consistent with associativity. For instance, the associative mean (3.42) is not very decisive since it leads to the predefined value α as soon as there exist $x_i \leq \alpha$ and $x_j \geq \alpha$.

Operators (3.39)–(3.42) will be investigated in more details in Section 4.3.

Czogała and Drewniak [32] have examined the case when M has an identity element $e \in E$. They obtained the following.

Theorem 3.3.10 *Let E be a real interval, finite or infinite.*
i) If $M \in A_2(E,E)$ fulfils (In, Id, A) and has an identity element $e \in E$, then there is a decreasing function $g : E \to E$ with $g(e) = e$ such that, for all $x, y \in E$,

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

ii) If $M \in A_2(E,E)$ fulfils (Co, In, Id, A) and has an identity element $e \in E$, then $M = \min$ or $\max$.

Fodor [65] showed that the previous result still holds in the more general framework of connected order topological spaces.

3.3.5 Triangular norms and conorms

In fuzzy set theory, one of the main topics consists in defining fuzzy logical connectives which are appropriate extensions of logical connectives **AND**, **OR** and **NOT** in the case when the valuation set is the unit interval $[0,1]$ rather than $\{0,1\}$.

Fuzzy connectives modelling **AND** and **OR** are called *triangular norms* (t-norms for short) and *triangular conorms* (t-conorms) respectively, see [12, 164].

Definition 3.3.2 *i)* A t-norm is a function $T : [0,1]^2 \to [0,1]$ fulfilling (Sy, In, A) and having 1 as identity.
ii) A t-conorm is a function $S : [0,1]^2 \to [0,1]$ fulfilling (Sy, In, A) and having 0 as identity.

The investigation of these functions has been made by Schweizer and Sklar [162, 163] and Ling [110]. See also Dubois and Prade [46] and the references mentioned there.

Of course, the family of continuous t-norms is nothing less than the class $\mathcal{A}_{0,1,0}$, and the family of continuous t-conorms is the class $\mathcal{A}_{0,1,1}$. These families have been described in Section 3.3.3. Moreover, in this context, Corollary 3.3.4 gives a characterization of their union:

Corollary 3.3.5 *$M \in A_2([0,1],[0,1])$ fulfils (Co, In, WId, A) and has exactly one identity in $[0,1]$ if and only if M is a continuous t-norm or a continuous t-conorm.*

It is well known from literature that t-norms and t-conorms are extensively used in fuzzy set theory, especially in modelling fuzzy connectives and implications (see [188]). Applications to practical problems require the use of, in a sense, the most appropriate t-norm or t-conorm. On this issue, Fodor [63] presented a method to construct new t-norms from t-norms.

It is worth noting that some properties of t-norms, such as associativity, do not play any essential role in preference modelling and choice theory. Recently, some authors [11, 57, 201] have investigated non-associative binary operation on $[0,1]$ in different contexts. These operators can be viewed as a generalization of t-norms and t-conorms in the sense that both are contained in this kind of operations. Moreover, Fodor [64] defined and investigated the concept of weak t-norms. His results were usefully applied to the framework of fuzzy strict preference relations.

We will not stress on this topic of t-norms and t-conorms. The interested reader can consult the book of Fodor and Roubens [70].

3.4 Operators that are stable for some scale transformations

3.4.1 Ratio, interval and inversion scales

A foundational paper of Aczél *et al.* [9] gives the general solutions of the functional equations related to (SRR), (IRR), etc. We present these solutions as well as some related results.

Theorem 3.4.1 *$M \in A_n(\mathbb{R}_0^+, \mathbb{R}_0^+)$ fulfils (SRR) if and only if*

$$M(x) = g(x_1)\, F(\frac{x_2}{x_1}, \dots, \frac{x_n}{x_1}), \quad x \in \mathbb{R}_0^+,$$

with $F \in A_{n-1}(\mathbb{R}_0^+, \mathbb{R}_0^+)$ and $g : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ such that $g(x\,y) = g(x)g(y)$ for all $x, y \in \mathbb{R}_0^+$. (If $n = 1$ then $F =$ constant.)

Proof. See Aczél *et al.* [9, case #2] and Aczél and Dhombres [7, Chap. 20]. ∎

Theorem 3.4.2 *$M \in A_n(\mathbb{R}_0^+, \mathbb{R}_0^+)$ fulfils (IRR) if and only if*

$$M(x) = a \prod_{i=1}^{n} g_i(x_i), \quad x \in \mathbb{R}_0^+,$$

with $a > 0$ and $g_i : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ such that $g_i(x_i\, y_i) = g_i(x_i)g_i(y_i)$ for all $x_i, y_i \in \mathbb{R}_0^+$.

Moreover, $M \in A_n(\mathbb{R}_0^+, \mathbb{R}_0^+)$ fulfils (Co, IRR) if and only if

$$M(x) = a \prod_{i=1}^{n} x_i^{c_i}, \quad x \in \mathbb{R}_0^+,$$

with $a > 0$ and $c_i \in \mathbb{R}$ for all i.

Proof. See Aczél *et al.* [9, case #4]. ∎

Theorem 3.4.3 *Non-constant operators $M \in A_n(]0,1],]0,1])$ that fulfil (IRR) are characterized by*

$$M(x) = a \prod_{i=1}^{n} x_i^{c_i}, \quad x \in \mathbb{R}_0^+,$$

with $a \in]0,1]$ and $c_i \geq 0$ for all i.

Proof. See Fodor and Roubens [70, Theorem 5.9]. ∎

Theorem 3.4.4 *$M \in A_n(\mathbb{R}, \mathbb{R})$ fulfils (SII) if and only if*

$$M(x) = \begin{cases} \mathrm{S}(x)\, F\Big(\frac{x_1 - \mathrm{AM}(x)}{\mathrm{S}(x)}, \dots, \frac{x_n - \mathrm{AM}(x)}{\mathrm{S}(x)}\Big) + a\, \mathrm{AM}(x) + b, & \text{if } \mathrm{S}(x) \neq 0, \\ a\, x_1 + b, & \text{if } \mathrm{S}(x) = 0 \ (\text{i.e. } x_1 = \cdots = x_n), \end{cases}$$

or

$$M(x) = \begin{cases} g(\mathrm{S}(x))\, F\Big(\frac{x_1 - \mathrm{AM}(x)}{\mathrm{S}(x)}, \dots, \frac{x_n - \mathrm{AM}(x)}{\mathrm{S}(x)}\Big) + b, & \text{if } \mathrm{S}(x) \neq 0, \\ b, & \text{if } \mathrm{S}(x) = 0, \end{cases}$$

where $a, b \in \mathbb{R}$, $\mathrm{S}(x) = \sqrt{\sum_{i=1}^{n} (x_i - \mathrm{AM}(x))^2}$, $F \in A_n(\mathbb{R}, \mathbb{R})$ arbitrary, and $g : \mathbb{R} \to \mathbb{R}_0^+$ such that $g(x\,y) = g(x)g(y)$ for all $x, y \in \mathbb{R}$. (If $n = 1$ then $M(x) = a\,x + b$).

Proof. See Aczél *et al.* [9, case #5]. ■

From the previous theorem and the immediate equivalence (Id, SII) $\Leftrightarrow$ (SPL) in $\mathbb{R}^n$, we deduce a description of the operators that fulfil (SPL).

Theorem 3.4.5 *$M \in A_n(\mathbb{R},\mathbb{R})$ fulfils (SPL) if and only if*

$$M(x) = \begin{cases} \mathrm{S}(x)\, F\Big(\frac{x_1-\mathrm{AM}(x)}{\mathrm{S}(x)}, \ldots, \frac{x_n-\mathrm{AM}(x)}{\mathrm{S}(x)}\Big) + \mathrm{AM}(x), & \text{if } \mathrm{S}(x) \neq 0, \\ x_1, & \text{if } \mathrm{S}(x) = 0 \ (i.e.\ x_1 = \cdots = x_n), \end{cases}$$

where $\mathrm{S}(x) = \sqrt{\sum_{i=1}^n (x_i - \mathrm{AM}(x))^2}$, *and* $F \in A_n(\mathbb{R},\mathbb{R})$ *arbitrary. (If $n = 1$ then $M(x) = x$).*

An alternative form of the previous theorem is the following (see also Proposition 3.5.1.)

Theorem 3.4.6 *$M \in A_n(\mathbb{R},\mathbb{R})$ fulfils (SPL) if and only if*

$$M(x) = \begin{cases} (x_{(n)} - x_{(1)})\, F\left(\frac{x_1-x_{(1)}}{x_{(n)}-x_{(1)}}, \ldots, \frac{x_n-x_{(1)}}{x_{(n)}-x_{(1)}}\right) + x_{(1)}, & \text{if } x_{(n)} > x_{(1)} \\ x_{(1)}, & \text{if } x_{(n)} = x_{(1)}. \end{cases}$$

where $F \in A_n(\mathbb{R},\mathbb{R})$ is arbitrary. (If $n = 1$ then $M(x) = x$).

Theorem 3.4.7 *$M \in A_n(\mathbb{R},\mathbb{R})$ fulfils (ISUII) if and only if*

$$M(x) = \sum_{i=1}^n a_i\, x_i + b, \quad x \in \mathbb{R}^n,$$

where $a_i, b \in \mathbb{R}$ are arbitrary constants.

Proof. See Aczél *et al.* [9, case #9]. ■

Theorem 3.4.8 *Let $M \in A_n(\mathbb{R},\mathbb{R})$. Then the following assertions are equivalent.*

i) *M fulfils (ISZII).*

ii) *M fulfils (III).*

iii) *There exists $j \in N$ such that*

$$M(x) = a\, x_j + b, \quad x \in \mathbb{R}^n,$$

where $a, b \in \mathbb{R}$ are arbitrary constants.

Proof. See Aczél *et al.* [9, cases #7 and #11]. ■

The following result, due to Silvert [172], gives the form of *symmetric sums*, that is the aggregation operators satisfying (Sy, Co, In, WId, SSN).

Theorem 3.4.9 *If $M \in A_n([0,1],[0,1])$ fulfils (Sy, Co, In, WId, SSN) then there exists an increasing continuous function $g : [0,1]^n \to [0,1]$ fulfilling $g(0,\ldots,0) = 0$ such that*

$$M(x) = \frac{g(x_1,\ldots,x_n)}{g(x_1,\ldots,x_n) + g(1-x_1,\ldots,1-x_n)}, \quad x \in [0,1]^n. \tag{3.43}$$

The representation (3.43) is not unique. For instance $\lambda\, g$, $\lambda \neq 0$, and $g = M$ all generate M.
Note also that Dombi [38] investigated the family of strictly increasing associative symmetric sums (see also [46]).

3.4.2 Ordinal scales

Let us introduce the concept of Boolean max-min functions (see Marichal [115]). The word 'Boolean' refers to the fact that these functions are generated by set functions that range in $\{0,1\}$. As mean operators are clearly cardinal in nature (i.e. only cardinal information can be aggregated), Boolean max-min functions are suitable for ordinal information.

Definition 3.4.1 For any set function $c : 2^N \to \{0,1\}$ such that $c_\emptyset = 0$ and

$$\bigvee_{T\subseteq N} c_T = 1,$$

the *Boolean max-min* function $\mathrm{B}_c^{\vee\wedge} : \mathbb{R}^n \to \mathbb{R}$ associated to c is defined by

$$\mathrm{B}_c^{\vee\wedge}(x) = \bigvee_{T\subseteq N} \Big[c_T \wedge (\bigwedge_{i\in T} x_i)\Big] = \bigvee_{\substack{T\subseteq N\\ c_T=1}} \bigwedge_{i\in T} x_i.$$

The best known form of a Boolean max-min function is the following: M is a Boolean max-min function if and only if there exists a finite family $\{T_k\}_{k=1}^m$ of non-empty subsets of N such that

$$M(x) = \bigvee_{k=1}^{m} \bigwedge_{i\in T_k} x_i.$$

Let us introduce the following sets which partition $\mathbb{R}^n$:

$$\mathcal{O}_\pi := \{x \in \mathbb{R}^n \,|\, x_{\pi(1)} \le \cdots \le x_{\pi(n)}\}, \quad \pi \in \Pi_n. \tag{3.44}$$

It is clear that any Boolean max-min function $M = \mathrm{B}_c^{\vee\wedge}$ is such that

$$\forall\, \pi \in \Pi_n,\ \exists\, k \in N : M(x) = x_k,\ \forall x \in \mathcal{O}_\pi. \tag{3.45}$$

The converse is not true: there exist many non-continuous functions M that satisfy (3.45), see [143, Lemma 4.3].

Proposition 3.4.1 *For any $\pi \in \Pi_n$, we have*

$$\mathrm{B}_c^{\vee\wedge}(x) = x_{\pi(j)}, \quad x \in \mathcal{O}_\pi,$$

with

$$j = \bigvee_{\substack{T\subseteq N\\ c_T=1}} \bigwedge_{i\in\pi^{-1}(T)} i.$$

Proof. If $x \in \mathcal{O}_\pi$, we have

$$\begin{aligned}\bigvee_{\substack{T\subseteq N\\ c_T=1}} \bigwedge_{i\in T} x_i &= \bigvee_{\substack{T\subseteq N\\ c_T=1}} \bigwedge_{\pi(i)\in T} x_{\pi(i)}\\ &= \bigvee_{\substack{T\subseteq N\\ c_T=1}} \bigwedge_{i\in\pi^{-1}(T)} x_{\pi(i)},\end{aligned}$$

which leads to the result. ■

Using classical distributivity of min and max operations, we can see that any Boolean max-min function can be put in a "conjunctive" form

$$\bigwedge_{T\subseteq N}\Big[d_T \vee (\bigvee_{i\in T} x_i)\Big] = \bigwedge_{\substack{T\subseteq N\\ d_T=0}} \bigvee_{i\in T} x_i$$

with an appropriate set function $d : 2^N \to \{0,1\}$ such that $d_\emptyset = 1$ and $\bigwedge_{T\subseteq N} d_T = 0$ (see Section 4.3 for more details).

The following result is due to Marichal and Mathonet [118].

Theorem 3.4.10 *Let $M \in A_n([a,b],[a,b])$. Then the following three assertions are equivalent:*

i) M fulfils (Co, OS').

ii) M fulfils (In, OS').

iii) There exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$.

Proof. *i)* $\Rightarrow$ *ii)* Consider $x = (x_1,\ldots,x_n) \in [a,b]^n$ reordered as $x_{(1)} \le \ldots \le x_{(n)}$ and set $x_{(0)} := a$, $x_{(n+1)} := b$. Let $k \in N$ and consider the function $f_k : [a,b] \to \mathbb{R}$ defined by

$$f_k(t) := M(x_1,\ldots,x_{k-1},t,x_{k+1},\ldots,x_n), \quad t\in[a,b].$$

We shall show that f_k is increasing on $[a,b]$. Let $i \in \{0,1,\ldots,n\}$ such that $x_{(i)} < x_{(i+1)}$. Then f_k is increasing on $]x_{(i)},x_{(i+1)}[$. Indeed, if $t_1,t_2 \in]x_{(i)},x_{(i+1)}[$, $t_2 > t_1$, then there exists $\phi \in \Phi'([a,b])$ such that $\phi(t_1) = t_2$, $\phi(x_{(j)}) = x_{(j)}$ for all $j \in \{0,1,\ldots,n+1\}$ and $\phi(t) \ge t$ for all $t\in[a,b]$. By (OS'), we then have

$$f_k(t_2) = f_k(\phi(t_1)) = \phi(f_k(t_1)) \ge f_k(t_1).$$

Finally, by (Co), f_k is increasing on $[a,b]$.

ii) $\Rightarrow$ *iii)* Set $M_T := M(a\,\overline{e}_T + b\,e_T)$ for all $T\subseteq N$. By Proposition 2.2.3, we have $M_T \in \{a,b\}$ for all $T\subseteq N$. Moreover, for all $x\in[a,b]$ and all $T\subseteq N$, we have

$$M(a\,\overline{e}_T + x\,e_T) = M_T \wedge x \tag{3.46}$$
$$M(x\,e_T + b\,\overline{e}_T) = M_{N\setminus T} \vee x \tag{3.47}$$

Indeed, taking $\phi(t) = \frac{x-a}{b-a}(t-a)+a$, we have, by (OS'),

$$M(a\,\overline{e}_T + x\,e_T) = M(\phi(a)\overline{e}_T + \phi(b)e_T) = \phi(M_T)$$

which proves (3.46). Similarly, taking $\phi(t) = \frac{b-x}{b-a}(t-a)+x$, we have

$$M(x\,e_T + b\,\overline{e}_T) = M(\phi(a)e_T + \phi(b)\overline{e}_T) = \phi(M_{N\setminus T})$$

which proves (3.47).

Let $x\in[a,b]^n$. On the one hand, for all $T\subseteq N$, we have

$$M(x) \overset{\text{(In)}}{\ge} M[a\,\overline{e}_T + (\bigwedge_{i\in T} x_i)e_T] \overset{(3.46)}{=} M_T \wedge (\bigwedge_{i\in T} x_i)$$

and thus

$$M(x) \ge \bigvee_{T\subseteq N}[M_T \wedge (\bigwedge_{i\in T} x_i)].$$

On the other hand, let $T^* \subseteq N$ such that $M_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$ is maximum and set

$$J := \{j \in N \mid x_j \leq M_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)\}.$$

We should have $J \neq \emptyset$: indeed, if $x_j > M_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$ for all $j \in N$, we have, since $M_N = b$,

$$M_N \wedge (\bigwedge_{i\in N} x_i) > M_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$$

which contradicts the definition of T^*. Moreover, we should have $M_{N\setminus J} = a$: indeed, if $M_{N\setminus J} = b$ with $N \setminus J \neq \emptyset$, we have, by definition of J,

$$M_{N\setminus J} \wedge (\bigwedge_{i\in N\setminus J} x_i) > M_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$$

a contradiction. Finally, we have,

$$M(x) \overset{\text{(In)}}{\leq} M[(M_{T^*} \wedge (\bigwedge_{i\in T^*} x_i))e_J + b\,\overline{e}_J] \overset{(3.47)}{=} M_{T^*} \wedge (\bigwedge_{i\in T^*} x_i) = \bigvee_{T\subseteq N} [M_T \wedge (\bigwedge_{i\in T} x_i)].$$

Therefore, setting $c_T := (M_T - a)/(b-a)$ for all $T \subseteq N$, we have $c_\emptyset = 0, c_N = 1$ and

$$M(x) = \bigvee_{T\subseteq N} [M_T \wedge (\bigwedge_{i\in T} x_i)] = \bigvee_{\substack{T\subseteq N\\ M_T=b}} \bigwedge_{i\in T} x_i = \bigvee_{\substack{T\subseteq N\\ c_T=1}} \bigwedge_{i\in T} x_i.$$

iii) $\Rightarrow$ *i)* Clearly M fulfils (Co). It also fulfils (OS') since we have $\phi(x \vee y) = \phi(x) \vee \phi(y)$ and $\phi(x \wedge y) = \phi(x) \wedge \phi(y)$ for all $\phi \in \Phi'([a,b])$ and all $x, y \in [a,b]$. ■

It is worth noting that when (OS') is replaced by (OS) in Theorem 3.4.10, the equivalence between the assertions does not hold anymore. Indeed, since $\Phi([a,b])$ is the set of all continuous strictly increasing functions $\phi : [a,b] \to [a,b]$ with boundary conditions $\phi(a) = a$ and $\phi(b) = b$, we immediately see that the non-continuous operator

$$M(x) = \begin{cases} b, & \text{if } \max_i x_i = b, \\ \min_i x_i, & \text{else,} \end{cases} \tag{3.48}$$

fulfils (In, OS), which is sufficient.

In the more general case when E is any doubly homogeneous linear order[2] [141, 143], the equivalence between *i)* and *iii)* in Theorem 3.4.10 remains true and was independently established by Ovchinnikov [143, Theorem 5.3] using a rather different approach. Since any connected open set in $\mathbb{R}$ is a doubly homogeneous linear order, the result can be stated as follows.

Theorem 3.4.11 *Assume that E is open. Then $M \in A_n(E,E)$ fulfils (Co, OS or OS') if and only if there exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$.*

We have already observed in the remark regarding Proposition 2.2.3 that, when E is not open, there exist operators $M \in A_n(E,E)$ fulfilling (Co, OS) other than $\mathrm{B}_c^{\vee\wedge}$. However, as the following result shows, those operators do not fulfil (Id). In other words, the operators that fulfil (Co, Id, OS) are the Boolean max-min functions. By Proposition 2.2.4, we can even replace (Co, Id, OS) by (Co, Id, CM). We will also assume that M ranges in $\mathbb{R}$, thus making our result all the more general.

[2]A linear order E is said to be doubly homogeneous if, for any $x_1, x_2, y_1, y_2 \in E$ such that $x_1 < x_2$ and $y_1 < y_2$, there is an automorphism $\phi : E \to E$ such that $\phi(x_1) = y_1$ and $\phi(x_2) = y_2$.

Theorem 3.4.12 *Let E be any real interval, finite or infinite. $M \in A_n(E, \mathbb{R})$ fulfils (Co, Id, CM or CM') if and only if there exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$.*

Proof. (Sufficiency) Trivial.

(Necessity) Since (CM') implies (CM), we can assume that M fulfils (Co, Id, CM). Let us show that we have

$$M(x) \in E^\circ, \quad \forall x \in (E^\circ)^n. \tag{3.49}$$

Suppose for example that there exists $x \in (E^\circ)^n$ such that $M(x) \leq \inf E$. By (Id), we have $x_{(1)} < x_{(n)}$. Moreover, there exists $J \in \mathbb{N}_0$ such that $\inf E + 1/j < x_{(1)}$ for all $j \geq J$. By (Id), we have

$$M(x) \leq M(\inf E + 1/j, \ldots, \inf E + 1/j), \quad j \geq J. \tag{3.50}$$

Now, let us consider a sequence $(\phi_i)_{i\in\mathbb{N}_0}$ such that, for all $i \in \mathbb{N}_0$ we have $\phi_i \in \Phi(E)$ and $\phi_i(t) = x_{(n)} + (t - x_{(n)})/i$ on $[x_{(1)}, x_{(n)}]$. By (3.50), we have, for all $j \geq J$,

$$\begin{aligned} x_{(n)} &\overset{\text{(Id)}}{=} M(x_{(n)}, \ldots, x_{(n)}) = M(\lim_{i\to\infty} \phi_i(x)) \overset{\text{(Co)}}{=} \lim_{i\to\infty} M(\phi_i(x)) \\ &\overset{\text{(CM)}}{\leq} \lim_{i\to\infty} M(\phi_i(\inf E + 1/j)) \overset{\text{(Id)}}{=} \lim_{i\to\infty} \phi_i(\inf E + 1/j). \end{aligned}$$

Letting $j \to \infty$, we obtain

$$x_{(n)} \leq \inf E < x_{(1)} < x_{(n)},$$

a contradiction. Of course, a similar argument can be used for $\sup E$. Hence (3.49) is proved.

By Proposition 2.2.4, it is clear that the restriction of M to $(E^\circ)^n$ fulfils the assumptions of Theorem 3.4.11. Hence, there exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$ on $(E^\circ)^n$ and even on E^n since M is continuous. ∎

To study the particular case of symmetric operators, we need to investigate the *order statistics* (1.5) (cf. van der Waerden [184, Sect. 17]). Ovchinnikov [141, Sect. 7] has proved that any order statistic can be put in the following two forms:

$$x_{(k)} = \begin{cases} \displaystyle\bigvee_{1\leq i_1<\cdots<i_{n-k+1}\leq n} (x_{i_1} \wedge \cdots \wedge x_{i_{n-k+1}}) & \text{(disjunctive normal form)} \\ \displaystyle\bigwedge_{1\leq i_1<\cdots<i_k\leq n} (x_{i_1} \vee \cdots \vee x_{i_k}) & \text{(conjunctive normal form)} \end{cases} \tag{3.51}$$

Therefore, the order statistic OS_k is a Boolean max-min function such that $c_T = 1$ if and only if $|T| = n - k + 1$. See Section 4.4.3 for more details.

The following result gives a characterization of the class of order statistics. Proved first in 1981 by Orlov [140] on $\mathbb{R}^n$, and then in 1993 by Marichal and Roubens [121, Theorem 1] on E^n, it was finally shown in the more general framework of ordered sets in 1996 by Ovchinnikov [141, Theorem 4.3].

Theorem 3.4.13 *$M \in A_n(E, \mathbb{R})$ fulfils (Sy, Co, Comp, CM) if and only if there exists $k \in N$ such that $M = \mathrm{OS}_k$.*

Proof. We give here the proof obtained by Marichal and Roubens [121].

(Sufficiency) Trivial.

(Necessity) By (Comp), $M \in A_n(E, E)$. By (2.1), M fulfils (Id), and by Proposition 2.2.4, M fulfils (OS). Then consider $z \in (E^\circ)^n$ such that $z_1 < \cdots < z_n$. By (Comp) and Proposition 2.2.3, there exists $k \in N$ such that $M(z) = z_k$.

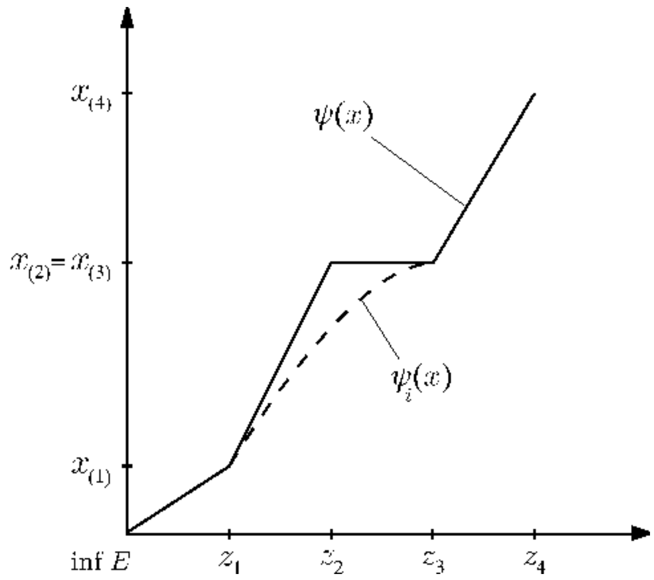


Figure 3.8: Surjection $\psi : E \to E$ such that $\psi(z_j) = x_{(j)}$ for all $j \in N$

Let $x \in E^n$ and consider a continuous non-decreasing surjection $\psi : E \to E$ such that $\psi(z_j) = x_{(j)}$ for all $j \in N$ (see Figure 3.8).

It is easy to build a sequence $(\psi_i)_{i\in \mathbb{N}_0}$ with $\psi_i \in \Phi(E)$, such that

$$\lim_{i\to\infty} \psi_i(t) = \psi(t), \quad t \in E.$$

We then have

$$\begin{aligned} M(x) &\overset{\text{(Sy)}}{=} M(x_{(1)}, \ldots, x_{(n)}) = M(\psi(z)) = M(\lim_{i\to\infty} \psi_i(z)) \overset{\text{(Co)}}{=} \lim_{i\to\infty} M(\psi_i(z)) \\ &\overset{\text{(OS)}}{=} \lim_{i\to\infty} \psi_i(M(z)) = \lim_{i\to\infty} \psi_i(z_k) = \psi(z_k) = x_{(k)}. \end{aligned}$$

Hence the result. ∎

We see that the order statistics are the symmetric Boolean max-min functions. Moreover, combining (2.1) and Theorems 3.4.12 and 3.4.13 provides the following result.

Theorem 3.4.14 *Let $M \in A_n(E, \mathbb{R})$. Then the following assertions are equivalent.*

i) M fulfils (Sy, Co, Id, CM).

ii) M fulfils (Sy) and there exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$.

iii) There exists $k \in N$ such that $M = \mathrm{OS}_k$.

Adding (Sy) in Theorems 3.4.10 and 3.4.11 allows to point out other characterizations of the class of order statistics.

We now describe the class of operators $M \in A_n([a, b], \mathbb{R})$ fulfilling (Co, CM). The result can be found in Marichal and Mathonet [118]. Before presenting it we shall go through a number of lemmas. Moreover, for any partition (R, S, T) of N, we define the function $M_{(R,S,T)} \in A_n([a, b], \mathbb{R})$ by

$$M_{(R,S,T)}(t) := M(a\, e_R + t\, e_S + b\, e_T), \quad t \in [a, b].$$

We also set $M_T := M(a\, \overline{e}_T + b\, e_T)$ for all $T \subseteq N$.

Lemma 3.4.1 *Let $M \in A_n([a,b],\mathbb{R})$ fulfilling (Co, CM). Then, for all partition (R,S,T) of N, the function $M_{(R,S,T)}$ is constant or strictly monotonic.*

Proof. Let $x_0, y_0 \in]a,b[$ such that $x_0 < y_0$ and suppose $M_{(R,S,T)}(x_0) < M_{(R,S,T)}(y_0)$ (resp. $>,=$). Let $x,y \in]a,b[$ such that $x < y$. There exists $\phi \in \Phi([a,b])$ such that $\phi(x_0) = x$ and $\phi(y_0) = y$. By (CM), we have $M_{(R,S,T)}(x) < M_{(R,S,T)}(y)$ (resp. $>,=$). Finally, by (Co), $M_{(R,S,T)}$ is strictly increasing (resp. strictly decreasing, constant) on $[a,b]$. ■

Lemma 3.4.2 *Let $M \in A_n([a,b],\mathbb{R})$ fulfilling (Co, CM). Then, there exist $T,T' \subseteq N$ such that $M_T \leq M(x) \leq M_{T'}$ for all $x \in [a,b]^n$.*

Proof. Let us consider the case of the lower bound. The other one can be treated similarly. By (Co), there exists $x^* \in [a,b]^n$ such that $M(x^*) \leq M(x)$ for all $x \in [a,b]^n$. Let

$$C := \{x_i^* \,|\, x_i^* \in]a,b[\}.$$

If $C = \emptyset$ then we can conclude immediately. Else, let

$$K := \{k \in N \,|\, x_k^* = \min C\} \neq \emptyset.$$

Choosing $k \in K$, we have

$$M(x^*) = M(x^* + (t - x_k^*)e_K), \quad t \in]a, x_k^*[;$$

indeed, suppose that there exists $t \in]a, x_k^*[$ such that $M(x^*) < M(x^* + (t - x_k^*)e_K)$ and consider $\phi \in \Phi([a,b])$ such that $\phi(t) = x_k^*$ and $\phi(x_i^*) = x_i^*$ for all $i \in N \setminus K$. By (CM), we have $M(\phi(x^*)) < M(x^*)$, a contradiction.

By (Co), we have

$$M(x^*) = M(x^* + (a - x_k^*)e_K).$$

We can iterate this process until obtaining $M(x^*) = M_T$ with $T = \{i \in N \,|\, x_i^* = b\}$. ■

Lemma 3.4.3 *Let $M \in A_n([a,b],\mathbb{R})$ fulfilling (Co, CM). Then, we have $M_\emptyset \leq M(x) \leq M_N$ or $M_N \leq M(x) \leq M_\emptyset$ for all $x \in [a,b]^n$.*

Proof. Let us consider the case of the lower bound. The other one can be treated similarly. By Lemma 3.4.2, there exists $T \subseteq N$ such that $M_T \leq M(x)$ for all $x \in [a,b]$. By Lemma 3.4.1, we have three mutually exclusive cases:

- If $M_{(\emptyset,N,\emptyset)}$ is strictly increasing then $M_\emptyset < M_N$. Let us show that $M_\emptyset \leq M_T$. Suppose it is not true. By Lemma 3.4.1, $M_{(\emptyset,N\setminus T,T)}$ is strictly increasing since

  $$M_{(\emptyset,N\setminus T,T)}(a) = M_T < M_N = M_{(\emptyset,N\setminus T,T)}(b).$$

  By (Co), there exists $r \in]a,b[$ such that $M_\emptyset = M_{(\emptyset,N\setminus T,T)}(r)$. Then, there exists $\phi_i \in \Phi([a,b])$, such that $\phi_i(r) = b - (b-r)/i$ for all $i \in \mathbb{N}_0$. By (CM), we have

  $$M_\emptyset = M_{(\emptyset,N\setminus T,T)}(\phi_i(r)), \quad i \in \mathbb{N}_0,$$

  and by (Co), $M_\emptyset = M_N$, a contradiction.

- If $M_{(\emptyset,N,\emptyset)}$ is strictly decreasing then $M_N < M_\emptyset$. Suppose $M_T < M_N$. By Lemma 3.4.1, $M_{(N\setminus T,T,\emptyset)}$ is strictly decreasing since

$$M_{(N\setminus T,T,\emptyset)}(a) = M_\emptyset > M_T = M_{(N\setminus T,T,\emptyset)}(b).$$

We can conclude as in the previous case.

- If $M_{(\emptyset,N,\emptyset)}$ is constant then $M_\emptyset = M_N$. Suppose $M_T < M_\emptyset$. By Lemma 3.4.1, $M_{(N\setminus T,T,\emptyset)}$ is strictly decreasing and $M_{(\emptyset,N\setminus T,T)}$ is strictly increasing. Taking $r \in]a,b[$, we have

$$M_T = M_{(N\setminus T,T,\emptyset)}(b) < M_{(N\setminus T,T,\emptyset)}(r) < M_{(N\setminus T,T,\emptyset)}(a) = M_\emptyset = M_N.$$

Hence, by (Co), there exists $s \in]a,b[$ such that

$$M_{(N\setminus T,T,\emptyset)}(r) = M_{(\emptyset,N\setminus T,T)}(s).$$

Consider $\phi_i \in \Phi([a,b])$, such that $\phi_i(r) = b-(b-r)/i$ and $\phi_i(s) = b-(b-s)/i$ for all $i \in \mathbb{N}_0$. By (CM) and (Co), we have $M_T = M_N$, a contradiction.

∎

Theorem 3.4.15 *$M \in A_n([a,b],\mathbb{R})$ fulfils (Co, CM) if and only if*
- *either M is constant,*
- *or there exist a set function c and a continuous strictly monotonic function $g : [a,b] \to \mathbb{R}$ such that $M = g \circ \mathrm{B}_c^{\vee\wedge}$.*

Proof. (Sufficiency) Easy.
(Necessity) By Lemma 3.4.1, $M_{(\emptyset,N,\emptyset)}$ is constant or strictly monotonic. If $M_{(\emptyset,N,\emptyset)}$ is constant then, by Lemma 3.4.3, so is M. In the other case, the function $g := M_{(\emptyset,N,\emptyset)}$ is a continuous bijection from $[a,b]$ onto $[M_\emptyset \wedge M_N, M_\emptyset \vee M_N]$. By Lemma 3.4.3, $M' := g^{-1} \circ M$ is well defined. Moreover, we can readily see that M' is an aggregation operator defined on $[a,b]^n$ and fulfilling (Co, Id, CM). Theorem 3.4.12 then allows to conclude. ∎

Corollary 3.4.1 *$M \in A_n([a,b],\mathbb{R})$ fulfils (Sy, Co, CM) if and only if*
- *either M is constant,*
- *or there exist $k \in N$ and a continuous strictly monotonic function $g : [a,b] \to \mathbb{R}$ such that $M = g \circ \mathrm{OS}_k$.*

We also have the following result.

Theorem 3.4.16 *$M \in A_n([a,b],[a,b])$ fulfils (Co, OS) if and only if*
- *either $M = a$ or $M = b$ (constant operators),*
- *or there exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$.*

Proof. (Sufficiency) Easy.
(Necessity) By Proposition 2.2.4, M fulfils (CM) and, by Theorem 3.4.15, we have two exclusive cases:

- M is constant and, in this case, we must have $\phi(M) = M$ for all $\phi \in \Phi([a,b])$, that is $M = a$ or b.

- There exists a set function c and a continuous strictly monotonic function $g : [a,b] \to \mathbb{R}$ such that $M = g \circ \mathrm{B}_c^{\vee\wedge}$. In this case, we have

$$M(x) \in]a,b[, \quad \forall x \in]a,b[^n,$$

and the restriction of M to $]a,b[^n$ fulfils the assumptions of Theorem 3.4.11, which is sufficient.

■

Kim [106, Corollary 1.2] showed that considering independent ordinal scales in multicriteria decision making leads to the presence of a dictator criterion.

Theorem 3.4.17 *$M \in A_n(\mathbb{R}, \mathbb{R})$ fulfils (Co, CMIS) if and only if*

- *either M is constant,*
- *or there exist $k \in N$ and a continuous strictly monotonic function $g : \mathbb{R} \to \mathbb{R}$ such that $M = g \circ \mathrm{P}_k$.*

The classical definition of continuity uses a distance between aggregated values and makes use of the cardinal properties of the scores x_i. Though (OS) and (Co) are not contradictory, coupling these two axioms is somewhat awkward since (OS) implies that the cardinal properties of the scores x_i should not be used. In the following three results, we drop the continuity property.

Let R be a total preorder on N. A *cell* $\mathcal{O}_R$ in E^n is defined by

$$\mathcal{O}_R := \{x \in E^n \,|\, x_i < x_j \Leftrightarrow i\,P\,j \text{ and } x_i = x_j \Leftrightarrow i\,I\,j\},$$

where P and I are the asymmetric and symmetric parts of R, respectively. The set of all cells forms a partition of E^n. Ovchinnikov [143, Theorem 5.1] described the class of functions fulfilling (OS) when E is any doubly homogeneous linear order. For connected real sets, the result can be stated as follows.

Theorem 3.4.18 *Assume that E is open. $M \in A_n(E,E)$ fulfils (OS) if and only if for all total preorder R on N, there exists $k \in N$ such that $M(x) = x_k$ for all $x \in \mathcal{O}_R$.*

Ovchinnikov [141, Theorem 4.2] also proved the following result.

Theorem 3.4.19 *Let $M \in A_n(E, \mathbb{R})$ fulfilling (Sy, Comp, CM). Then the restriction of M to*

$$\mathcal{A} = \bigcup_{\pi \in \Pi_n} \{x \in (E^\circ)^n \,|\, x_{\pi(1)} < \cdots < x_{\pi(n)}\} = \{x \in (E^\circ)^n \,|\, x_{(1)} < \cdots < x_{(n)}\}$$

is an order statistic OS_k.

According to the previous theorem, we see that the 'pathological functions' such as (3.48) occur only on the 'borders' of regions like $\{x \in (E^\circ)^n \,|\, x_{\pi(1)} < \cdots < x_{\pi(n)}\}$. In a sense, an operator fulfilling (Sy, Comp, CM) is an order statistic 'almost everywhere' on E^n.

It is interesting to observe that, since E^n is the closure of $\mathcal{A}$ (see Theorem 3.4 in [141]), adding (Co) in Theorem 3.4.19 allows to retrieve Theorem 3.4.13.

Theorem 3.4.20 *$M \in A_n([a,b], \mathbb{R})$ fulfils (In, Id, CM') if and only if there exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$.*

Proof. (Sufficiency) Trivial.

(Necessity) By (2.2), we have $M \in A_n([a,b],[a,b])$. By Proposition 2.2.4, M fulfils (OS'). We then can conclude by Theorem 3.4.10. ■

3.4.3 Additive, minitive and maxitive operators

In [4, Sect. 5.1], Aczél proved the following result.

Theorem 3.4.21 *i) $M \in A_n(E, \mathbb{R})$ fulfils (Add) if and only if*

$$M(x) = \sum_{i=1}^{n} g_i(x_i), \quad x \in E^n,$$

with $g_i : E \to \mathbb{R}$ such that $g_i(x_i + y_i) = g_i(x_i) + g_i(y_i)$ for all $x_i, y_i \in E$.
ii) $M \in A_n(E, \mathbb{R})$ fulfils (Co, Add) if and only if

$$M(x) = \sum_{i=1}^{n} a_i\, x_i, \quad x \in E^n,$$

with $a_i \in \mathbb{R}$ for all $i \in N$.

The following theorem gives a description of the aggregation operators fulfilling (Min) or (Max) (see also [49]).

Theorem 3.4.22 *i) $M \in A_n([a,b], \mathbb{R})$ fulfils (Min) if and only if there exist increasing functions $g_i : [a,b] \to \mathbb{R}$, $i \in N$, such that*

$$M(x) = \bigwedge_{i=1}^{n} g_i(x_i), \quad x \in [a,b]^n.$$

ii) $M \in A_n([a,b], \mathbb{R})$ fulfils (Max) if and only if there exist increasing functions $h_i : [a,b] \to \mathbb{R}$, $i \in N$, such that

$$M(x) = \bigvee_{i=1}^{n} h_i(x_i), \quad x \in [a,b]^n.$$

Proof. *i)* (Necessity) Let $x \in [a,b]^n$. By (Min), we have

$$M(x) = \bigwedge_{i=1}^{n} M(x_i\, e_i + b\, \overline{e}_i) = \bigwedge_{i=1}^{n} g_i(x_i)$$

where $g_i(t) = M(t\, e_i + b\, \overline{e}_i)$, $i \in N$. Moreover, for all $i \in N$, g_i is increasing; indeed, if $t, t' \in [a,b]$, $t \le t'$, we have that

$$g_i(t) = g_i(t \wedge t') = g_i(t) \wedge g_i(t')$$

implies $g_i(t) \le g_i(t')$.

(Sufficiency) We have

$$g_i(t \wedge t') = g_i(t) \wedge g_i(t')$$

for all $t, t' \in [a,b]$ and all $i \in N$; indeed, if $t \le t'$, we have $g_i(t) \le g_i(t')$ and

$$g_i(t \wedge t') = g_i(t) = g_i(t) \wedge g_i(t').$$

We then can conclude.

ii) Similar to *i)*. ■

3.5 Operators that are stable for positive linear transformations

In this section, E will be any real interval (finite or infinite) containing $[0,1]$. Moreover, we set $\theta_S^{(n)} := M^{(n)}(e_S)$ and $\overline{\theta}_S^{(n)} := M^{(n)}(\overline{e}_S)$ for all $S \subseteq N$, and the superscript (n) will often be omitted.

This section aims at describing the family of all aggregation operators fulfilling three specific properties. The first two are (In, SPL). The third property is chosen among well-known algebraic properties such as associativity, decomposability and bisymmetry (see Section 2.3 for details). All the results of this section can be found in Marichal *et al.* [120].

3.5.1 Preliminary characterizations

The following proposition shows that all the operators fulfilling (SPL) can be considered on $[0,1]^n$ without loss of generality (see also Theorem 3.4.6).

Proposition 3.5.1 *Assume $E \supseteq [0,1]$. Any $M \in A_n(E, \mathbb{R})$ fulfilling (SPL) is completely defined by its restriction to $[0,1]^n$.*

Proof. Let $M' \in A_n([0,1], \mathbb{R})$ denote the restriction to $[0,1]^n$ of M, that is $M' = M$ on $[0,1]^n$. By (SPL), we have, for all $x \in E^n$,

$$M(x) = \begin{cases} x_{(1)}, & \text{if } x_{(n)} = x_{(1)} \\ (x_{(n)} - x_{(1)})\, M'\left(\frac{x_1 - x_{(1)}}{x_{(n)} - x_{(1)}}, \ldots, \frac{x_n - x_{(1)}}{x_{(n)} - x_{(1)}}\right) + x_{(1)}, & \text{otherwise.} \end{cases}$$

We then can conclude. ■

The next proposition is very interesting and easy to prove. A quite similar statement can be found in [7, Chap. 15] (see remark that follows Proposition 9).

Proposition 3.5.2 *Assume $E \supseteq [0,1]$. $M \in A_2(E, \mathbb{R})$ fulfils (In, SPL) if and only if there exist $\alpha, \beta \in [0,1]$ such that*

$$M(x,y) = \alpha\, x + \beta\, y + (1 - \alpha - \beta)(x \wedge y) \quad \forall x, y \in E. \tag{3.52}$$

Moreover, we have $\alpha = M(1,0)$ and $\beta = M(0,1)$.

Proof. (Sufficiency) Easy.

(Necessity) Let $x, y \in E$. If $x \leq y$ then we have

$$M(x,y) \overset{\text{(SPL)}}{=} (y - x)\, M(0,1) + x = (1 - \beta)\, x + \beta\, y$$

with $\beta = M(0,1)$. Moreover, $\beta \in [0,1]$ since, by (2.2) and (2.6), M is compensative.

One proceeds similarly if $x \geq y$. ■

Corollary 3.5.1 *Assume $E \supseteq [0,1]$. $M \in A_2(E, \mathbb{R})$ fulfils (Sy, In, SPL) if and only if there exists $\omega \in [0,1]^2$ such that $M = \mathrm{OWA}_\omega$.*

Figure 3.9 shows the graphical representation on $[0,1]^2$ of the function defined in (3.52). It is worth comparing it with Figure 3.7.

Some particular examples according to the values of α and β can be found in Table 3.3.

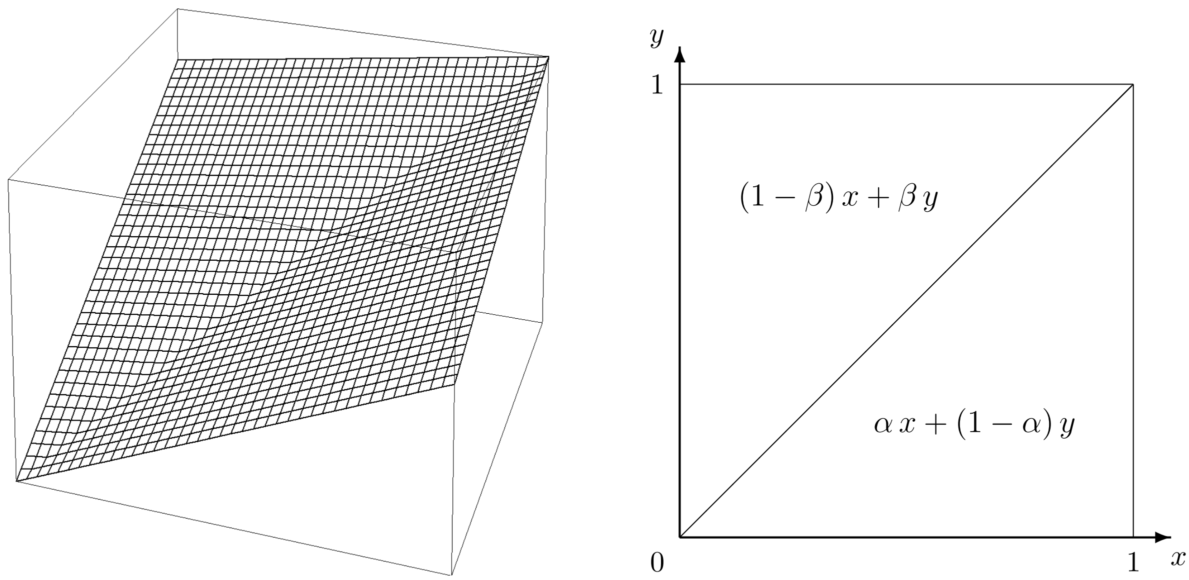


Figure 3.9: Representation on $[0,1]^2$ of $M(x,y) = \alpha\,x + \beta\,y + (1-\alpha-\beta)(x \wedge y)$

values of (α,β)	M
$(\alpha,\beta) = (0,0)$	min
$(\alpha,\beta) = (1,1)$	max
$(\alpha,\beta) = (1,0)$	P_1
$(\alpha,\beta) = (0,1)$	P_2
$\alpha+\beta = 1$	$\mathrm{WAM}_{(\alpha,\beta)}$
$\alpha = \beta$	$\mathrm{OWA}_{(1-\beta,\beta)}$

Table 3.3: Some examples of two-place operators fulfilling (In, SPL)

3.5.2 Bisymmetric and autodistributive operators

Theorem 3.5.1 *Assume $E \supseteq [0,1]$ and let $M \in A_2(E,\mathrm{I\!R})$. Then the following three assertions are equivalent:*

i) *M fulfils (In, SPL, B)*
ii) *M fulfils (In, SPL, AD)*
iii) $M \in \{\min,\max\} \cup \{\mathrm{WAM}_\omega \,|\, \omega \in [0,1]^2\}$.

Proof. *iii)* $\Rightarrow$ *i)* Trivial.

i) $\Rightarrow$ *ii)* It is a straightforward consequence of (2.6) and Proposition 2.3.4.

ii) $\Rightarrow$ *iii)* Set $\alpha := M(1,0)$ and $\beta := M(0,1)$. By Proposition 3.5.2, we only have to prove that $(\alpha,\beta) \in \{(0,0),(1,1)\}$ or $\alpha+\beta = 1$. Since M fulfils (AD), we must have

$$M(x_1, M(x_2,x_3)) = M(M(x_1,x_2), M(x_1,x_3)), \quad x \in E^3.$$

Substituting $x = (\alpha,1,0)$ into this identity, we obtain, by (3.52), $\alpha = 0$ or $\alpha = 1$ or $\alpha+\beta = 1$. Similarly, for $x = (\beta,0,1)$, we obtain, by (3.52), $\beta = 0$ or $\beta = 1$ or $\alpha+\beta = 1$. ■

Corollary 3.5.2 *Assume $E \supseteq [0,1]$ and let $M \in A_2(E, \mathbb{R})$. Then the following three assertions are equivalent:*
i) *M fulfils (Sy, In, SPL, B)*
ii) *M fulfils (Sy, In, SPL, AD)*
iii) $M \in \{\min, \max, \mathrm{AM}\}$.

We now intend to describe the family of operators $M \in A_n(E, E)$ $(n \geq 2)$ fulfilling (In, SPL, B). For this purpose, we need three technical lemmas.

Lemma 3.5.1 *Let $n \in \mathbb{N}_0, n \geq 2$. If $M \in A_n([0,1],[0,1])$ fulfils (In, SPL, B) and if there exists $S \subseteq N$ such that $\theta_S \in]0,1[$ then M fulfils (Add).*

Proof. Let $x, y \in [0,1]^n$ with $x_i + y_i \in [0,1]$ for all $i \in N$, and set $\lambda := \inf\{\theta_S, 1-\theta_S\} \in]0,1[$. Let us show that

$$M(x+y) = M(x) + M(y). \tag{3.53}$$

(i) Assume first that $x_i, y_i \leq \lambda$ for all $i \in N$. Consider the square matrix X of n rows r_i and n columns c_j $(i, j \in N)$, where r_i is defined as follows (for $i \in N$):

$$r_i = \frac{1}{2}\Big(\frac{x_i}{\theta_S} + 1\Big)e_S + \frac{1}{2}\,\frac{y_i}{1-\theta_S}\,\overline{e}_S.$$

On the one hand, by (SPL), we have, for all $i \in N$:

$$M(r_i) = \frac{1}{2}\,\frac{y_i}{1-\theta_S} + \frac{1}{2}\Big(\frac{x_i}{\theta_S} + 1 - \frac{y_i}{1-\theta_S}\Big)\,\theta_S = \frac{x_i+y_i}{2} + \frac{\theta_S}{2},$$

and thus

$$M(M(r_1), \ldots, M(r_n)) = \frac{1}{2}\,M(x+y) + \frac{\theta_S}{2}.$$

On the other hand, for all $j \in N$, we have

$$M(c_j) = \begin{cases} \frac{1}{2}(\frac{M(x)}{\theta_S} + 1), & \text{if } j \in S, \\ \frac{1}{2}\frac{M(y)}{1-\theta_S}, & \text{otherwise.} \end{cases}$$

However, by (2.6) and (2.2), M fulfils (Comp). Hence, by (SPL), we have,

$$M(M(c_1), \ldots, M(c_n)) = M\Big[\frac{1}{2}\Big(\frac{M(x)}{\theta_S} + 1\Big)e_S + \frac{1}{2}\frac{M(y)}{1-\theta_S}\,\overline{e}_S\Big] = \frac{1}{2}\,[M(x)+M(y)] + \frac{\theta_S}{2}.$$

Since M fulfils (B), we have (3.53).
(ii) In the general case, we have,

$$M(x+y) \overset{\text{(SPL)}}{=} \frac{1}{\lambda}\,M(\lambda\,x + \lambda\,y) \overset{(i)}{=} \frac{1}{\lambda}\,M(\lambda\,x) + \frac{1}{\lambda}\,M(\lambda\,y) \overset{\text{(SPL)}}{=} M(x) + M(y).$$

■

Lemma 3.5.2 *Let $n \in \mathbb{N}_0, n \geq 2$. If $M \in A_n([0,1],[0,1])$ fulfils (In, SPL, B), and if $\theta_S \in \{0,1\}$ for all $S \subseteq N$, then, setting $S_{\max} := \{i \in N \,|\, \theta_i = 1\}$ and assuming $S_{\max} \neq \emptyset$, we have, for all $S \subseteq N$:*

$$S \cap S_{\max} = \emptyset \Rightarrow \theta_S = 0.$$

Proof. The result is trivial if $n = 2$. Otherwise, use induction over $|S|$.

The result holds for $|S| \in \{0, 1\}$. Assume that it holds for $|S| = s \geq 1$ and show that it holds for $|S| = s + 1$. Assume that $S \cap S_{\max} = \emptyset$ and consider the square matrix X of n rows r_i and n columns c_j $(i, j \in N)$, where r_i is defined as follows (for $i \in N$):

$$r_i = \begin{cases} e_\emptyset, & \text{if } i \notin S, \\ e_S, & \text{if } i = q, \\ e_N, & \text{if } i \in S \setminus q, \end{cases}$$

where $q \in S$. On the one hand, for all $i \in S \setminus q$, we have $M(r_i) = 1$. On the other hand, for all $j \in N$, we have

$$M(c_j) = \begin{cases} 0, & \text{if } j \notin S \text{ (by induction)} \\ \theta_S, & \text{if } j \in S. \end{cases}$$

Indeed, if $j \notin S$, we have $c_j = e_{S \setminus q}$ and $(S \setminus q) \cap S_{\max} = \emptyset$.

Now, let $j_0 \in S \setminus q$. In particular, we have $j_0 \notin S_{\max}$. Consider the square matrix X' of n rows r'_i and n columns c'_j $(i, j \in N)$, where $r'_i = r_i$ for all $i \in N \setminus j_0$ and $r'_{j_0} = \overline{e}_{j_0}$. We then have

$$M(\overline{e}_{j_0}) = \overline{\theta}_{j_0} = 1$$

since, by (In), $\overline{\theta}_{j_0} \geq \theta_i$ for all $i \in S_{\max}$. So we have

$$M(M(r'_1), \ldots, M(r'_n)) = M(M(r_1), \ldots, M(r_n)).$$

However, since $c'_{j_0} = e_{S \setminus j_0}$, we have $M(c'_{j_0}) = 0$ (by induction) and

$$M(c'_j) = \begin{cases} 0, & \text{if } j \notin S, \\ 0, & \text{if } j = j_0, \\ \theta_S, & \text{if } j \in S \setminus j_0. \end{cases}$$

Consequently, since M fulfils (SPL) and (B), we have

$$\begin{aligned} \theta_S^2 &= M(M(c_1), \ldots, M(c_n)) = M(M(r_1), \ldots, M(r_n)) \\ &= M(M(r'_1), \ldots, M(r'_n)) = M(M(c'_1), \ldots, M(c'_n)) = \theta_S \, M(e_{S \setminus j_0}) = 0. \end{aligned}$$

■

Lemma 3.5.3 *Let $n \in \mathbb{N}_0, n \geq 2$. If $M \in A_n([0, 1], [0, 1])$ fulfils (In, SPL, B), if $\theta_S \in \{0, 1\}$ for all $S \subseteq N$, and $\theta_i = 0$ for all $i \in N$, then, setting $S_{\min} = \{i \in N \mid \overline{\theta}_i = 0\}$, we have $S_{\min} \neq \emptyset$ and $\theta_{S_{\min}} = 1$.*

Proof. If $n = 2$ then $\overline{\theta}_1 = \theta_2 = 0$ and $\overline{\theta}_2 = \theta_1 = 0$. Otherwise, let $S^*_{\min} \subseteq N$ with a minimal cardinality such that $\theta_{S^*_{\min}} = 1$. The existence of such a $S^*_{\min}$ is trivial since $\theta_N = 1$, and we clearly have $|S^*_{\min}| \geq 2$. Let us prove that $S^*_{\min} \subseteq S_{\min}$. Assume that there exists $k \in S^*_{\min}$ such that $\overline{\theta}_k = 1$. Consider the square matrix X of n rows r_i and n columns c_j $(i, j \in N)$, where r_i is defined as follows (for $i \in N$):

$$r_i = \begin{cases} e_\emptyset, & \text{if } i \notin S^*_{\min}, \\ e_{S^*_{\min}}, & \text{if } i = k, \\ \overline{e}_k & \text{if } i \in S^*_{\min} \setminus k. \end{cases}$$

On the one hand, we have, for all $i \in N$:

$$M(r_i) = \begin{cases} 0, & \text{if } i \notin S^*_{\min} \\ 1, & \text{if } i \in S^*_{\min} \end{cases}$$

and thus
$$M(M(r_1),\dots,M(r_n)) = M(e_{S^*_{\min}}) = \theta_{S^*_{\min}} = 1.$$
On the other hand, for all $j \in N$, we have
$$M(c_j) = \begin{cases} 0, & \text{if } j \notin S^*_{\min} \quad \text{since } |S^*_{\min}| \text{ is minimal,} \\ 0, & \text{if } j = k \quad \text{since } \theta_i = 0 \; \forall i \in N, \\ 1, & \text{if } j \in S^*_{\min} \setminus k \quad \text{since } c_j = e_{S^*_{\min}}. \end{cases}$$
Since M fulfils (B), we have
$$1 = M(M(c_1),\dots,M(c_n)) = M(e_{S^*_{\min}\setminus k}),$$
a contradiction since $|S^*_{\min}|$ is minimal.

Now, let us prove that $S_{\min} \subseteq S^*_{\min}$. For all $i \notin S^*_{\min}$, we have, by (In), $\overline{\theta}_i \geq \theta_{S^*_{\min}} = 1$, that is $i \notin S_{\min}$. ■

Theorem 3.5.2 *Assume $E \supseteq [0,1]$. $M \in A_n(E,E)$ fulfils (In, SPL, B) if and only if*
$$M \in \{\min_S, \max_S \,|\, S \subseteq N\} \cup \{\mathrm{WAM}_\omega \,|\, \omega \in [0,1]^n\}.$$

Proof. (Sufficiency) Trivial.

(Necessity) By Proposition 3.5.1, we can assume that $M \in A_n([0,1],[0,1])$. Let $x \in [0,1]^n$. The values θ_S ($S \subseteq N$) fulfil the assumptions of exactly one of Lemmas 3.5.1, 3.5.2 or 3.5.3. We then have three exclusive cases:

(i) Under the assumptions of Lemma 3.5.1, we have, by (SPL),
$$M(x) = \sum_{i=1}^n \theta_i \, x_i,$$
with, by (Id), $\sum_i \theta_i = M(n \odot 1) = 1$.

(ii) Under the assumptions of Lemma 3.5.2, there exists $p \in S_{\max}$ such that $x_p = \max_{i \in S_{\max}} x_i$ and we have
$$M(x) \overset{\text{(In)}}{\leq} M(x_p \, e_{S_{\max}} + \overline{e}_{S_{\max}}) \overset{\text{(SPL)}}{=} x_p + (1-x_p)\, \overline{\theta}_{S_{\max}} = x_p$$
and
$$M(x) \overset{\text{(In)}}{\geq} M(x_p \, e_p) \overset{\text{(SPL)}}{=} x_p \, \theta_p = x_p.$$
Therefore, we have
$$M(x) = \max_{i \in S_{\max}} x_i.$$

(iii) Under the assumptions of Lemma 3.5.3, there exists $p \in S_{\min}$ such that $x_p = \min_{i \in S_{\min}} x_i$ and we have, since $p \in S_{\min}$:
$$M(x) \overset{\text{(In)}}{\leq} M(x_p \, e_p + \overline{e}_p) \overset{\text{(SPL)}}{=} x_p + (1-x_p)\, \overline{\theta}_p = x_p$$
and
$$M(x) \overset{\text{(In)}}{\geq} M(x_p \, e_{S_{\min}}) \overset{\text{(SPL)}}{=} x_p \, \theta_{S_{\min}} = x_p.$$
Therefore, we have
$$M(x) = \min_{i \in S_{\min}} x_i.$$
■

Corollary 3.5.3 *Assume $E \supseteq [0,1]$. $M \in A_n(E,E)$ fulfils (Sy, In, SPL, B) if and only if*
$$M \in \{\min, \max, \mathrm{AM}\}.$$

3.5.3 Extended operators fulfilling general bisymmetry

The (GB) property is stronger than simply fulfilling (B) for each n. Thus, for example, we can satisfy (In, SPL, B) and have radically different operations for each degree of cardinality of aggregation. That is we could have an extended aggregation operator such that

$$\begin{aligned}
M(x_1, x_2) &= \min(x_1, x_2), \\
M(x_1, x_2, x_3) &= \frac{1}{3}\sum_{i=1}^{3} x_i, \\
M(x_1, x_2, x_3, x_4) &= x_3, \\
M(x_1, x_2, x_3, x_4, x_5) &= \max(x_1, x_4).
\end{aligned}$$

Thus in this situation we have no restriction on the process of extending the functions to greater cardinalities in a consistent way.

As Theorem 3.5.3 below shows, (GB) brings some requirements for consistency in extending an operator's cardinality. Let us begin with two lemmas.

Lemma 3.5.4 *If $M \in A([0,1],[0,1])$ fulfils (In, SPL, GB) and if there exists $p \in \mathbb{N}_0$, $p \geq 2$ such that*

$$M^{(p)} \in \{\mathrm{WAM}_\omega^{(p)} \mid \omega \in [0,1]^p\} \setminus \{\mathrm{P}_i^{(p)} \mid i \in N_p\}$$

then, for all $n \in \mathbb{N}_0$, there exists $\omega \in [0,1]^n$ such that $M^{(n)} = \mathrm{WAM}_\omega$.

Proof. There exist $\omega \in [0,1]^p$ and $\{i_1, i_2\} \subseteq N_p$ with $i_1 \neq i_2$, such that $M^{(p)} = \mathrm{WAM}_\omega^{(p)}$ with $\omega_{i_1}, \omega_{i_2} \neq 0$.

Let $n \in \mathbb{N}_0, n \geq 2$, and let us show that $M^{(n)}$ fulfils (Add). Let $x, y \in [0,1]^n$ with $x_i + y_i \in [0,1]$ for all $i \in N_n$. As in Lemma 3.5.1, we can assume $x_i, y_i \leq \inf\{\omega_{i_1}, \omega_{i_2}\} \in]0,1[$ for all $i \in N_n$. Consider the matrix X of p rows r_i $(i \in N_p)$ and n columns c_j $(j \in N_n)$, where r_i is defined as follows (for $i \in N_p$):

$$r_i = \begin{cases} \frac{1}{\omega_{i_1}}\, x & \text{if } i = i_1, \\ \frac{1}{\omega_{i_2}}\, y & \text{if } i = i_2, \\ e_\emptyset & \text{if } i \in N_p \setminus i_1 i_2. \end{cases}$$

Since M fulfils (SPL, GB), we have $M^{(n)}(x+y) = M^{(n)}(x) + M^{(n)}(y)$. We then can conclude as in the proof of Theorem 3.5.2. ∎

Lemma 3.5.5 *If $M \in A([0,1],[0,1])$ fulfils (In, SPL, GB) and if there exists $p \in \mathbb{N}_0$, $p \geq 2$ such that*

$$M^{(p)} \in \{\min_S^{(p)} \mid S \subseteq N_p\} \setminus \{\mathrm{P}_i^{(p)} \mid i \in N_p\}$$
$$\textit{(resp. } M^{(p)} \in \{\max_S^{(p)} \mid S \subseteq N_p\} \setminus \{\mathrm{P}_i^{(p)} \mid i \in N_p\}\textit{)}$$

then, for all $n \in \mathbb{N}_0$, there exists $S \subseteq N_n$ such that $M^{(n)} = \min_S$ (resp. $M^{(n)} = \max_S$).

Proof. Assume that there exist $S \subseteq N_p$ and $\{i_1, i_2\} \subseteq N_p$ with $i_1 \neq i_2$, such that $M^{(p)} = \min_S^{(p)}$. The case $\max_S^{(p)}$ can be treated similarly.

Let $n \in \mathbb{N}_0, n \geq 2$, and $x, y \in [0,1]^n$. Next, consider the matrix X of p rows r_i $(i \in N_p)$ and n columns c_j $(j \in N_n)$, where r_i is defined as follows (for $i \in N_p$):

$$r_i = \begin{cases} x, & \text{if } i = i_1, \\ y, & \text{if } i = i_2, \\ e_{N_n}, & \text{if } i \in N_p \setminus i_1 i_2. \end{cases}$$

Since M fulfils (GB), we see that it also fulfils (Min). Therefore, if $x \in [0,1]^n$ then

$$M^{(n)}(x) = \min_{i \in N_n} M^{(n)}(x_i\, e_i + \overline{e}_i) \overset{\text{(SPL)}}{=} \min_{i \in N_n} [(1 - x_i)\overline{\theta}_i^{(n)} + x_i].$$

Let us show that $\overline{\theta}_i^{(n)} \in \{0,1\}$ for all $i \in N_n$. Suppose it is not true. Since $M^{(n)}$ fulfils (In, SPL, B), by Lemma 3.5.1, $M^{(n)}$ fulfils (Add) and thus

$$M^{(n)} \in \{\mathrm{WAM}_\omega^{(n)} \,|\, \omega \in [0,1]^n\} \setminus \{\mathrm{P}_i^{(n)} \,|\, i \in N_n\}$$

which is impossible by Lemma 3.5.4. Indeed, we have

$$\{\mathrm{WAM}_\omega^{(p)} \,|\, \omega \in [0,1]^p\} \cap \{\min_S^{(p)} \,|\, S \subseteq N_p\} = \{\mathrm{P}_i^{(p)} \,|\, i \in N_p\} \tag{3.54}$$

since if $M^{(p)}$ belongs to the left-hand set of (3.54) then we necessarily have $\omega_i = \theta_i^{(p)} \in \{0,1\}$.

Finally, we have $M^{(n)} = \min_S^{(n)}$ with $S = \{i \in N_n \,|\, \overline{\theta}_i^{(n)} = 0\}$. ■

Theorem 3.5.3 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (In, SPL, GB) if and only if*
- *either: for all $n \in \mathbb{N}_0$, there exists $S \subseteq N_n$ such that $M^{(n)} = \min_S$,*
- *or: for all $n \in \mathbb{N}_0$, there exists $S \subseteq N_n$ such that $M^{(n)} = \max_S$,*
- *or: for all $n \in \mathbb{N}_0$, there exists $\omega \in [0,1]^n$ such that $M^{(n)} = \mathrm{WAM}_\omega$.*

Proof. (Sufficiency) We can easily check that all the extended operators mentioned in the statement fulfil (In, SPL, GB).

(Necessity) By Proposition 3.5.1, we can assume that $M \in A([0,1],[0,1])$. If, for all $n \in \mathbb{N}_0$, $M^{(n)} \in \{\mathrm{P}_i^{(n)} \,|\, i \in N_n\}$ then we can conclude immediately. Otherwise, there exists $p \in \mathbb{N}_0, p \geq 2$ such that $M^{(p)} \notin \{\mathrm{P}_i^{(p)} \,|\, i \in N_p\}$. By Proposition 2.3.5, $M^{(p)}$ fulfils (In, SPL, B). Applying Theorem 3.5.2, we can see that $M^{(p)}$ fulfils the assumptions of Lemma 3.5.4 or 3.5.5, and we can conclude again. ■

Corollary 3.5.4 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (Sy, In, SPL, GB) if and only if*

$$M = (\min{}^{(n)})_{n \in \mathbb{N}_0} \;\; or \;\; (\max{}^{(n)})_{n \in \mathbb{N}_0} \;\; or \;\; (\mathrm{AM}^{(n)})_{n \in \mathbb{N}_0}.$$

3.5.4 Decomposable and strongly decomposable extended operators

Lemma 3.5.6 *If $M \in A([0,1],[0,1])$ fulfils (In, SPL, D) then*

$$M^{(2)} \in \{\mathrm{WAM}_\omega, \mathrm{OWA}_\omega \,|\, \omega \in [0,1]^2\}.$$

Proof. Set $\alpha := M(1,0)$ and $\beta := M(0,1)$. By Proposition 3.5.2, we only have to prove that $\alpha = \beta$ or $\alpha + \beta = 1$. Let us proceed in two steps:

(i) We have successively

$$M(0,0,1) \overset{\text{(D)}}{=} M(0,\beta,\beta) \overset{\text{(SPL)}}{=} \beta\, M(0,1,1)$$

and

$$M(0,1,1) \overset{\text{(D)}}{=} M(\beta,\beta,1) \overset{\text{(SPL)}}{=} \beta + (1-\beta)\, M(0,0,1).$$

It follows that

$$M(0,0,1) = \frac{\beta^2}{\beta^2 - \beta + 1}$$

and, similarly,

$$M(1,0,0) = \frac{\alpha^2}{\alpha^2 - \alpha + 1}.$$

(ii) We have, using (D) and (SPL),

$$M(0,1,0) = \beta\, M(1,1,0) = \alpha\, M(0,1,1).$$

By (i), the latter equality becomes

$$\beta\Big(\alpha + (1-\alpha)\,\frac{\alpha^2}{\alpha^2-\alpha+1}\Big) = \alpha\Big(\beta + (1-\beta)\frac{\beta^2}{\beta^2-\beta+1}\Big)$$

or, after reduction, $\alpha = 0$ or $\beta = 0$ or $\alpha = \beta$ or $\alpha + \beta = 1$. If $\beta = 0$, i.e. $M(0,0,1) = 0$, we have successively

$$\alpha \overset{\text{(SPL)}}{=} M(\alpha,\alpha,1) \overset{\text{(D)}}{=} M(1,0,1) \overset{\text{(D)}}{=} M(1,0,0) = \frac{\alpha^2}{\alpha^2-\alpha+1}$$

and thus $\alpha = 0$ or $\alpha = 1$. We proceed similarly if $\alpha = 0$. Hence the result. ∎

Theorem 3.5.4 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (In, SPL, D) if and only if*
- *either: $M = (\min^{(n)})_{n\in\mathbb{N}_0}$,*
- *or: $M = (\max^{(n)})_{n\in\mathbb{N}_0}$,*
- *or: there exists $\theta \in [0,1]$ such that, for all $n \in \mathbb{N}_0$, we have $M^{(n)} = \mathrm{WAM}_\omega$ with*

$$\omega_i = \frac{(1-\theta)^{n-i}\,\theta^{i-1}}{\sum_{j=1}^n (1-\theta)^{n-j}\,\theta^{j-1}}, \quad i \in N_n.$$

Proof. (Sufficiency) We can easily check that all the extended operators mentioned in the statement fulfil the corresponding properties.

(Necessity) By Proposition 3.5.1, we can assume that $M \in A([0,1],[0,1])$. By Lemma 3.5.6, there exists $\theta \in [0,1]$ such that

$$M^{(2)} = \mathrm{WAM}_{(1-\theta,\theta)} \quad \text{or} \quad \mathrm{OWA}_{(1-\theta,\theta)}.$$

(i) Assume first that $M^{(2)} = \mathrm{WAM}_{(1-\theta,\theta)}$. Let us prove by induction over $n \geq 2$ that

$$M(x) = \frac{1}{D_n}\sum_{i=1}^n (1-\theta)^{n-i}\,\theta^{i-1} x_i, \quad x \in [0,1]^n,$$

where $D_n = \sum_{j=1}^n (1-\theta)^{n-j}\,\theta^{j-1}$. The result holds true for $n = 2$. Suppose it holds for a fixed $n \geq 2$ and show it still holds for $n+1$. Let $(x_1,\ldots,x_{n+1}) \in [0,1]^{n+1}$ and set $x := (x_1,\ldots,x_n)$. We then have

$$\begin{aligned} M^{(n+1)}(x_1,\ldots,x_{n+1}) &\overset{\text{(D)}}{=} M^{(n+1)}(n \odot M^{(n)}(x), x_{n+1}) \\ &\overset{\text{(SPL)}}{=} \begin{cases} (M^{(n)}(x) - x_{n+1})\,\overline{\theta}_{n+1}^{(n+1)} + x_{n+1} & \text{if } x_{n+1} \leq M^{(n)}(x), \\ (x_{n+1} - M^{(n)}(x))\,\theta_{n+1}^{(n+1)} + M^{(n)}(x) & \text{if } x_{n+1} \geq M^{(n)}(x). \end{cases} \end{aligned}$$

Let us show that $\overline{\theta}_{n+1}^{(n+1)}$ is uniquely determined. The same can be done for $\theta_{n+1}^{(n+1)}$. Using induction, we have

$$\overline{\theta}_{n+1}^{(n+1)} \overset{\text{(D)}}{=} M^{(n+1)}(1, n \odot \overline{\theta}_n^{(n)}) \overset{\text{(ind.)}}{=} M^{(n+1)}(1, n \odot (1 - \frac{\theta^{n-1}}{D_n}))$$

and

$$\theta_1^{(n+1)} \overset{\text{(D)}}{=} M^{(n+1)}(n \odot \theta_1^{(n)}, 0) \overset{\text{(ind.)}}{=} M^{(n+1)}(n \odot \frac{(1-\theta)^{n-1}}{D_n}, 0).$$

Hence, using (SPL), we have the linear system

$$\begin{cases} \overline{\theta}_{n+1}^{(n+1)} = \dfrac{\theta^{n-1}}{D_n}\, \theta_1^{(n+1)} + (1 - \dfrac{\theta^{n-1}}{D_n}) \\ \theta_1^{(n+1)} = \dfrac{(1-\theta)^{n-1}}{D_n}\, \overline{\theta}_{n+1}^{(n+1)} \end{cases}$$

whose determinant

$$\det \begin{pmatrix} 1 & -\frac{\theta^{n-1}}{D_n} \\ -\frac{(1-\theta)^{n-1}}{D_n} & 1 \end{pmatrix} = 1 - \frac{(1-\theta)^{n-1}\, \theta^{n-1}}{D_n^2}$$

is strictly positive. Indeed, we have $D_n \geq (1-\theta)^{n-1}$ and $D_n \geq \theta^{n-1}$, with at least one strict inequality. Consequently, $M^{(n+1)}(x_1, \ldots, x_{n+1})$ is uniquely determined and M corresponds to the third case in the statement of the theorem.

(*ii*) Now assume that $M^{(2)} = \mathrm{OWA}_{(1-\theta,\theta)}$. By Proposition 2.3.1, M fulfils (Sy). By (2.13), M fulfils (SD). By Propositions 2.3.6 and 2.3.5, $M^{(2)}$ fulfils (B). Finally, using Corollary 3.5.2, we have $M^{(2)} \in \{\min, \max, \mathrm{AM}\}$.

The operator AM is a particular case of (*i*) for which $\theta = 1/2$. For the other two cases, we can conclude by Lemma 3.2.3. ■

Theorem 3.5.5 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (In, SPL, SD) if and only if*

$$M = (\min{}^{(n)})_{n\in\mathbb{N}_0} \;\; or \;\; (\max{}^{(n)})_{n\in\mathbb{N}_0} \;\; or \;\; (\mathrm{P}_1^{(n)})_{n\in\mathbb{N}_0} \;\; or \;\; (\mathrm{P}_n^{(n)})_{n\in\mathbb{N}_0} \;\; or \;\; (\mathrm{AM}^{(n)})_{n\in\mathbb{N}_0}.$$

Proof. (Sufficiency) Trivial.

(Necessity) Suppose $M \neq (\min^{(n)})_{n\in\mathbb{N}_0}$ and $M \neq (\max^{(n)})_{n\in\mathbb{N}_0}$. Since (SD) implies (D), by Theorem 3.5.4, there exists $\theta \in [0,1]$ such that, for all $n \in \mathbb{N}_0$, we have $M^{(n)} = \mathrm{WAM}_\omega$ with

$$\omega_i = \frac{(1-\theta)^{n-i}\, \theta^{i-1}}{\sum_{j=1}^n (1-\theta)^{n-j}\, \theta^{j-1}}, \quad i \in N_n.$$

Since M fulfils (SD), we must have

$$M(M(x_1,x_3), x_2, M(x_1,x_3)) = M(x_1,x_2,x_3), \quad x \in E^3.$$

In particular, for $x = (0,0,1)$, the previous equality becomes $M(\theta,0,\theta) = M(0,0,1)$, that is, $(1-\theta)^2\,\theta + \theta^3 = \theta^2$. Hence, we have $\theta \in \{0,1,1/2\}$ that allows to complete the proof. ■

Corollary 3.5.5 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (Sy, In, SPL) and (D or SD) if and only if*

$$M = (\min{}^{(n)})_{n\in\mathbb{N}_0} \;\; or \;\; (\max{}^{(n)})_{n\in\mathbb{N}_0} \;\; or \;\; (\mathrm{AM}^{(n)})_{n\in\mathbb{N}_0}.$$

3.5.5 Associative operators

Theorem 3.5.6 *Assume $E \supseteq [0,1]$. $M \in A_2(E,E)$ fulfils (In, SPL, A) if and only if*

$$M \in \{\min, \max, \mathrm{P}_1, \mathrm{P}_2\}.$$

Proof. (Sufficiency) Trivial.

(Necessity) Set $\alpha := M(1,0)$ and $\beta := M(0,1)$. By Proposition 3.5.2, we only have to prove that $\alpha, \beta \in \{0,1\}$. Since M fulfils (A), we must have

$$M(M(x_1,x_2),x_3) = M(x_1, M(x_2,x_3)), \quad x \in E^3.$$

In particular, for $x = (1,0,0)$, we obtain, by (SPL), $\alpha \in \{0,1\}$. Also, for $x = (0,0,1)$, we obtain $\beta \in \{0,1\}$. ∎

Theorem 3.5.7 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (In, SPL, A) if and only if*

$$M = (\min{}^{(n)})_{n\in\mathbb{N}_0} \ \ or \ \ (\max{}^{(n)})_{n\in\mathbb{N}_0} \ \ or \ \ (\mathrm{P}_1^{(n)})_{n\in\mathbb{N}_0} \ \ or \ \ (\mathrm{P}_n^{(n)})_{n\in\mathbb{N}_0}.$$

Proof. (Sufficiency) Trivial.

(Necessity) We construct the sequence $(M^{(n)})_{n\in\mathbb{N}_0}$ by induction over n. The functions $M^{(2)}$ are given by Theorem 3.5.6. Thus assume that $M^{(k)} = \min$ for all $k \leq n$ for a fixed $n \geq 2$. By (A), we simply have

$$M(x_1,\ldots,x_{n+1}) = M(M(x_1,\ldots,x_n),x_{n+1}) = \min_{i\in N_{n+1}} x_i.$$

The other cases can be treated similarly. ∎

Corollary 3.5.6 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (Sy, In, SPL, A) if and only if*

$$M = (\min{}^{(n)})_{n\in\mathbb{N}_0} \ \ or \ \ (\max{}^{(n)})_{n\in\mathbb{N}_0}.$$

Chapter 4

Fuzzy measures and integrals

4.1 Definitions and motivations

A significant aspect of aggregation in multicriteria decision making is the difference in the importance of criteria, which is usually modelled by using different weights. Since these weights must be taken into account during the aggregation, it is necessary to use weighted operators, thus giving up the symmetry property (Sy). Until recently, the most often used weighted aggregation operators were averaging operators, such as the quasi-linear means (3.3).

On this topic, Cholewa [27] and Montero [126, 127] described a serie of axioms that weighted aggregation operators should follow, and proposed the weighted arithmetic mean (WAM) as a typical aggregation operator that satisfies these axioms.

However, the weighted arithmetic means and, more generally, the quasi-linear means present some drawbacks. None of these operators is able to model in some understandable way an interaction between criteria. Indeed, it is well known in multiattribute utility theory (MAUT) that these operators lead to *mutual preferential independence* among the criteria, which expresses in some sense the independence of the criteria (see Section 1.3). Since these aggregation operators are not appropriate when interactive criteria are considered, people usually tend to construct independent criteria, or criteria that are supposed to be so, causing some bias effect in evaluation.

In order to have a flexible representation of complex interaction phenomena between criteria (e.g. positive or negative synergy between some criteria), it is useful to substitute to the weight vector a non-additive set function allowing to define a weight not only on each criterion, but also on each subset of criteria.

For this purpose, the use of fuzzy measures have been proposed by Sugeno in 1974 [177] to generalize additive measures. It seems widely accepted that additivity is not suitable as a required property of set functions in many real situations, due to the lack of additivity in many facets of human reasoning. To be able to express human subjectivity, Sugeno proposed to replace the additivity property by a weaker one: *monotonicity*, and he called these non-additive monotonic measures *fuzzy measures*. It is important to note however that fuzzy measures have nothing to do with fuzzy sets.

The purpose of this chapter is to show the usefulness of fuzzy measures and integrals in multicriteria decision making. Two main classes of fuzzy integrals are investigated and characterized, namely the Choquet and Sugeno integrals. Some subclasses are also studied in detail.

In the present section, we introduce the concept of fuzzy measure, also called capacity. We will see that such a measure can be defined from a so-called pseudo-Boolean function. The concept of fuzzy integral is also introduced. In Section 4.2 the Choquet integral is studied and characterized axiomatically. The link with the so-called Lovász extension of pseudo-Boolean functions is pointed out as well. As particular Choquet integrals, the weighted arithmetic mean and the ordered weighted averaging are also investigated. Section 4.3 deals in detail with the Sugeno integral. We show that this integral can be written under several equivalent forms. We also characterize the class of all the Sugeno integrals as well as some well-known subclasses. Section 4.4 is devoted to the aggregation operators that are simultaneously Choquet and Sugeno integrals. This family corresponds to the Boolean max-min functions already encountered in Chapter 3.

4.1.1 The concept of fuzzy measure

We consider a discrete set of n elements $N = \{1, \ldots, n\}$. Depending on the application, these elements could be players of a cooperative game, criteria in a multicriteria decision problem, attributes, experts or voters in an opinion pooling problem, etc. 2^N indicates the power set of N, i.e. the set of all subsets of N.

Definition 4.1.1 A (discrete) *fuzzy measure* on N is a set function $\mu : 2^N \to [0,1]$ satisfying the following conditions[1]:
i) $\mu(\emptyset) = 0, \mu(N) = 1$,
ii) $S \subseteq T \;\Rightarrow\; \mu(S) \leq \mu(T)$.

Note that the monotonicity condition *ii)* is obviously equivalent to:

$$\mu(S \cup i) \geq \mu(S), \quad \forall i \in N,\;\; \forall S \subseteq N \setminus i. \tag{4.1}$$

Moreover, the value of $\mu(N)$ is not very important; it is set equal to one for normalization reasons only.

Thus, a fuzzy measure is a set of 2^n real values obeying certain boundary and monotonicity conditions, which can be put into a lattice form. For any $S \subseteq N$, $\mu(S)$ can be viewed as the weight of importance or strength of the combination S for the particular decision problem under consideration. Thus, in addition to the usual weights on criteria taken separately, weights on any combination of criteria are also defined. Monotonicity then means that adding a new element to a combination cannot decrease its importance.

We will always assume that the weights are numerical values and can add up. In other terms, expressions like $\mu(S) + \mu(T)$ or $\mu(S \cup i) - \mu(S)$ can be interpreted.
Throughout this dissertation we will often write μ_S instead of $\mu(S)$.

We now give some usual definitions about fuzzy measures. We say that a fuzzy measure is

- *additive* if $\mu_{S \cup T} = \mu_S + \mu_T$ whenever $S \cap T = \emptyset$,
- *superadditive* if $\mu_{S \cup T} \geq \mu_S + \mu_T$ whenever $S \cap T = \emptyset$,
- *supermodular* if $\mu_{S \cup T} + \mu_{S \cap T} \geq \mu_S + \mu_T$ for all $S, T \subseteq N$.

Subadditivity and submodularity can be defined as well by reversing the inequalities. Remark that supermodularity, which is sometimes called convexity[2] (see [170]), implies superadditivity.

[1] Definitions on infinite spaces usually require algebras and σ-algebras, but this is not necessary in the discrete case (see full details in this respect in Denneberg [36] and Grabisch *et al.* [89]).
[2] Although convexity is often associated to supermodularity, it is also closely related to submodularity, see e.g. Lovász [111, Proposition 4.1].

If a fuzzy measure is additive then it suffices to define the n coefficients (weights) $\mu_1, \ldots, \mu_n$ to define the measure entirely. In general, one needs to define the $2^n - 2$ coefficients corresponding to the 2^n subsets of N, except $\emptyset$ and N.

Another particular class of fuzzy measures is that of *0-1 fuzzy measures*, whose values are either 0 or 1.

It was very early noticed that fuzzy measures can model some kind of interaction between criteria, but this issue was not formalized until the proposal by Murofushi and Soneda [130] of an interaction index for a pair of criteria. Later, Grabisch proposed a generalization of this index [81] to any subset of criteria, and Grabisch and Roubens proposed an axiomatic basis for the interaction index [92], giving a consistent basis for dealing with the notion of interaction. We elaborate on this subject in Chapter 5.

It should be mentioned that the concept of fuzzy measure predates its use by Sugeno. Historically, this concept has been first introduced in 1953 by Choquet [28] as a *capacity*. Later, it was encountered under many different names, such as 'confidence measure' [43], 'non-additive probability' [161], or, as in [183], 'weighting function'.

In the sequel, we follow the tradition of Sugeno and use the name 'fuzzy measure'.

In some domains, the monotonicity condition is not a requisite for set functions. In cooperative game theory, any real valued set function $v : 2^N \to \mathbb{R}$, with $v(\emptyset) = 0$, is called the *characteristic function of a game* (see e.g. Shapley [169]). Such a set function assigns to each coalition S of players a real number $v(S)$ representing the *worth* (i.e. the amount of money the coalition will earn if the game is played) or the *power* of S. When a player i sows discord between members of a coalition S, then the power of this coalition could decrease: $v(S \cup i) \leq v(S)$.

One also defines the *unanimity game* for $T \subseteq N$, as the game v_T such that $v_T(S) = 1$ if and only if $S \supseteq T$, and 0 otherwise.

4.1.2 Pseudo-Boolean functions

Recall that, for any subset $S \subseteq N$, e_S represents the *characteristic vector* of S, i.e. the vector of $\{0,1\}^n$ whose i-th component is 1 if and only if $i \in S$. Geometrically, the characteristic vectors are the 2^n vertices of the hypercube $[0,1]^n$.

Definition 4.1.2 A pseudo-Boolean function is a function $f : \{0,1\}^n \to \mathbb{R}$.

Any real valued set function $v : 2^N \to \mathbb{R}$ can be assimilated unambiguously with a pseudo-Boolean function. The correspondence is straightforward: we have

$$f(x) = \sum_{T \subseteq N} v(T) \prod_{i \in T} x_i \prod_{i \notin T} (1 - x_i), \quad x \in \{0,1\}^n,$$

and $v(S) = f(e_S)$ for all $S \subseteq N$. We shall henceforth make this identification.

In particular, the pseudo-Boolean function that corresponds to a fuzzy measure is increasing in each variable and fulfils the boundary conditions: $f(0, \ldots, 0) = 0$ and $f(1, \ldots, 1) = 1$.

Hammer and Rudeanu [96] showed that any pseudo-Boolean function has a unique expression as a multilinear polynomial in n variables:

$$f(x) = \sum_{T \subseteq N} a(T) \prod_{i \in T} x_i, \quad x \in \{0,1\}^n, \tag{4.2}$$

with $a(T) \in \mathbb{R}$.

In game theory, the real coefficients $\{a(T)\}_{T\subseteq N}$ are called the *dividends* of the coalitions in game v, see [98, 145]. Moreover, equation (4.2) is the decomposition of the set function v into unanimity games: indeed, $\prod_{i\in T} x_i$ corresponds to the unanimity game v_T and we have, for all $S \subseteq N$,

$$v(S) = f(e_S) = \sum_{T\subseteq N} a(T) \prod_{i\in T} (e_S)_i = \sum_{T\subseteq N} a(T)\, v_T(S). \tag{4.3}$$

Thus, any game v has a canonical representation in terms of unanimity games that determine a linear basis for v. Note that Gilboa and Schmeidler [76] and Pap [146] extended this unanimity-basis representation to general (infinite) spaces of players.

In combinatorics, a viewed as a set function on N is called the *Möbius transform* of v (see e.g. Rota [154]), which is given by

$$a(S) = \sum_{T\subseteq S} (-1)^{s-t} v(T), \quad S \subseteq N, \tag{4.4}$$

where $s = |S|$ and $t = |T|$.

When a is given, it is possible to recover the original v by the so-called *zeta* transform[3]:

$$v(S) = \sum_{T\subseteq S} a(T), \quad S \subseteq N. \tag{4.5}$$

The existence of an inverse transformation shows clearly that the correspondance between a and v is one-to-one, and a is a *representation* of v. In the sequel we will often write a_T instead of $a(T)$.

Of course, any set of 2^n coefficients $\{a_T \,|\, T \subseteq N\}$ could not be the Möbius representation of a fuzzy measure: the boundary and monotonicity conditions must be ensured. In terms of the Möbius representation, those conditions are very easy to prove, see e.g. Chateauneuf and Jaffray [25]:

Proposition 4.1.1 *A set of 2^n coefficients a_T, $T \subseteq N$, corresponds to the Möbius representation of a fuzzy measure if and only if*

$$\begin{cases} a_\emptyset = 0, \quad \displaystyle\sum_{T\subseteq N} a_T = 1, \\ \displaystyle\sum_{T: i\in T\subseteq S} a_T \geq 0, \quad \forall S \subseteq N,\ \forall i \in S. \end{cases} \tag{4.6}$$

Conditions (4.6) represent 2 equalities and

$$\sum_{s=1}^{n} s \binom{n}{s} = n2^{n-1} \tag{4.7}$$

inequalities.

[3]The link between the zeta transform and the classical Riemann zeta function is discussed in [37].

4.1.3 Fuzzy integrals as a new aggregation tool

When a fuzzy measure is available on N, it is interesting to have tools capable of summarizing all the values of a function to a single point, in terms of the underlying fuzzy measure. These tools are the fuzzy integrals, a concept proposed by Sugeno [177, 178].

Fuzzy integrals are integrals of a real function with respect to a fuzzy measure, by analogy with Lebesgue integral which is defined with respect to an ordinary (i.e. additive) measure. As the integral of a function in a sense represents its average value, a fuzzy integral can be viewed as a particular case of averaging aggregation operator.

Contrary to the weighted arithmetic means, fuzzy integrals are able to represent a certain kind of interaction between criteria, ranging from redundancy (negative interaction) to synergy (positive interaction). For this reason they have been thoroughly studied in the context of multicriteria decision problems [79, 80, 84, 134].

There are several classes of fuzzy integrals, among which the most representative are those of Sugeno and Choquet[4].

The concept of Choquet integral was proposed by Schmeidler [160] and Murofushi and Sugeno [131, 132], using a concept introduced by Choquet in capacity theory [28]. Since this integral is viewed here as an aggregation operator, we will adopt a connective-like notation instead of the usual integral form, and the integrand will be a set of n values $x_1, \ldots, x_n$ of $\mathbb{R}$.

Definition 4.1.3 Let μ be a fuzzy measure on N. The *(discrete) Choquet integral of a function $x : N \to \mathbb{R}$ with respect to μ* is defined by

$$\mathcal{C}_\mu(x) := \sum_{i=1}^{n} x_{(i)} \, [\mu_{\{(i),\ldots,(n)\}} - \mu_{\{(i+1),\ldots,(n)\}}],$$

with the usual convention that $x_{(1)} \leq \cdots \leq x_{(n)}$.

For instance, if $x_3 \leq x_1 \leq x_2$, we have

$$\mathcal{C}_\mu(x_1, x_2, x_3) = x_3 \, [\mu_{\{3,1,2\}} - \mu_{\{1,2\}}] + x_1 \, [\mu_{\{1,2\}} - \mu_{\{2\}}] + x_2 \, \mu_{\{2\}}.$$

The Choquet integral is closely related to the Lebesgue integral, since both coincide when the measure is additive:

$$\mathcal{C}_\mu(x) = \sum_{i=1}^{n} \mu_i \, x_i, \quad x \in \mathbb{R}.$$

In this sense, the Choquet integral is a generalization of the Lebesgue integral.

We introduce now the concept of discrete Sugeno integral, viewed as an aggregation operator. For theorical developments, see [88, 177, 178].

Definition 4.1.4 Let μ be a fuzzy measure on N. The *(discrete) Sugeno integral of a function $x : N \to [0,1]$ with respect to μ* is defined by

$$\mathcal{S}_\mu(x) := \bigvee_{i=1}^{n} [x_{(i)} \wedge \mu_{\{(i),\ldots,(n)\}}].$$

[4]Definitions of fuzzy integrals are presented here in the restrictive case of finite spaces, for we deal with spaces of criteria which are finite. We refer the reader to [131, 132, 133, 177] for more complete definitions.

For instance, if $x_3 \leq x_1 \leq x_2$, we have

$$\mathcal{S}_\mu(x_1, x_2, x_3) = (x_3 \wedge \mu_{\{3,1,2\}}) \vee (x_1 \wedge \mu_{\{1,2\}}) \vee (x_2 \wedge \mu_{\{2\}}).$$

Of course, given a fuzzy measure μ on N, the Choquet and Sugeno integrals can be regarded as aggregation operators defined on $\mathbb{R}^n$ and $[0,1]^n$, respectively. But they are essentially different in nature, since the latter is based on non-linear operators (min and max), and the former on usual linear operators. Both compute a kind of distorded average of $x_1, \ldots, x_n$. More general definitions exist but will not be considered here (see [79, 89]).

In the following sections, we will investigate some properties of these fuzzy integrals as well as some axiomatic characterizations. For instance, the Choquet integral fulfils (SPL) and the Sugeno integral fulfils (SMin, SMax), which represent the counterpart of (SPL) for ordinal values. In this sense, it can be said that the Choquet integral is suitable for cardinal aggregation (where numbers have a real meaning), while the Sugeno integral seems to be more suitable for ordinal aggregation (where only order makes sense). See Sections 6.1 and 6.5.

4.2 The Choquet integral

Before studying the Choquet integral, we have to mention that it is also known in the framework of combinatorial optimization as Lovász extension. We now present this concept.

4.2.1 Lovász extension

Lovász [111, Sect. 3] has observed that any $x \in (\mathbb{R}^+)^n \setminus \{0\}$ can be written uniquely in the form

$$x = \sum_{i=1}^{k} \lambda_i \, e_{S_i} \tag{4.8}$$

where $\lambda_1, \ldots, \lambda_k > 0$ and $\emptyset \neq S_1 \subsetneq \cdots \subsetneq S_k \subseteq N$. For example, we have

$$\begin{aligned} (1,5,3) &= 2\,(0,1,0) + 2\,(0,1,1) + 1\,(1,1,1), \\ (0,5,3) &= 2\,(0,1,0) + 3\,(0,1,1). \end{aligned}$$

Hence any function $f : \{0,1\}^n \to \mathbb{R}$ with $f(0) = 0$ can be extended to $\hat{f} : (\mathbb{R}^+)^n \to \mathbb{R}$, by $\hat{f}(0) = 0$ and

$$\hat{f}(x) = \sum_{i=1}^{k} \lambda_i \, f(e_{S_i}) \quad \Big(x = \sum_{i=1}^{k} \lambda_i \, e_{S_i} \in (\mathbb{R}^+)^n \setminus \{0\}\Big);$$

indeed, $\hat{f}$ is well defined (due to the uniqueness of (4.8)) and $\hat{f}(x) = f(x)$ for all $x \in \{0,1\}^n$. The function $\hat{f}$ is called [74] the *Lovász extension* of f.

Now, the Lovász extension of an arbitrary function $f : \{0,1\}^n \to \mathbb{R}$ is defined by

$$\hat{f}(x) = f(0) + \hat{f}_0(x), \quad x \in (\mathbb{R}^+)^n,$$

where $\hat{f}_0$ is the Lovász extension of $f_0 = f - f(0)$.

As has been stated by Lovász ([111, Proposition 4.1]), $\hat{f}$ is convex (resp. concave, linear) if and only if f is submodular (resp. supermodular, modular). Conversely, the restriction of a linear function to $\{0,1\}^n$ is modular, but the restriction of a convex function to $\{0,1\}^n$ need not be submodular [111].

The hypercube $[0,1]^n$ can be subdivided into $n!$ simplices of the form

$$\mathcal{B}_\pi := \{x \in [0,1]^n \,|\, x_{\pi(1)} \leq \cdots \leq x_{\pi(n)}\}, \quad \pi \in \Pi_n.$$

Of course, for each $\pi \in \Pi_n$, we have $\mathcal{B}_\pi = \mathcal{O}_\pi \cap [0,1]^n$, where $\mathcal{O}_\pi$ is defined by (3.44). Moreover, $\mathcal{B}_\pi$ is the convex hull of

$$\{e_{\{\pi(i),\dots,\pi(n)\}}\}_{i=1}^{n+1}.$$

Singer [173, Sect. 2] has shown that $\hat{f}$ is defined on each cone $\mathcal{K}_\pi = \{\lambda\,\mathcal{B}_\pi \,|\, \lambda \geq 0\}$ as the unique affine function that coincides with f at the $n+1$ vertices of $\mathcal{B}_\pi$. More formally, $\hat{f}$ can be written as[5]:

$$\hat{f}(x) = f(0) + \sum_{i=1}^{n} x_{\pi(i)}\,[f(e_{\{\pi(i),\dots,\pi(n)\}}) - f(e_{\{\pi(i+1),\dots,\pi(n)\}})], \quad x \in \mathcal{K}_\pi. \tag{4.9}$$

It is well known [173] that $\mathcal{B}_\pi$ is a polytope with vertices $\varepsilon_i^\pi = e_{\{\pi(i),\dots,\pi(n)\}}$ $(i = 1,\dots,n+1)$. Thus, on each polytope $\mathcal{B}_\pi$, the graph of $\hat{f}$ is the portion of the unique hyperplane passing through $(\varepsilon_1^\pi, f(\varepsilon_1^\pi)),\dots,(\varepsilon_{n+1}^\pi, f(\varepsilon_{n+1}^\pi))$ (which is, actually, their convex hull).

Concerning the convexity of $\hat{f}$, Singer [173, Theorem 3.3] proved the following result.

Theorem 4.2.1 *Let $f : \{0,1\}^n \to \mathbb{R}$, with Lovász extension $\hat{f} : (\mathbb{R}^+)^n \to \mathbb{R}$. For any $\pi \in \Pi_n$, we set*

$$\Phi_\pi(x) := \sum_{i=1}^{n} x_{\pi(i)}\,[f(e_{\{\pi(i),\dots,\pi(n)\}}) - f(e_{\{\pi(i+1),\dots,\pi(n)\}})], \quad x \in \mathbb{R}^n.$$

Then the following statements are equivalent:

1. *f is submodular.*
2. *$\hat{f}$ is convex.*
3. *$\hat{f}$ is polyhedral convex.*
4. *We have*
$$\hat{f}(x) = f(0) + \max_{\pi \in \Pi_n} \Phi_\pi(x), \quad x \in (\mathbb{R}^+)^n.$$
5. *We have*
$$f(x) = f(0) + \max_{\pi \in \Pi_n} \Phi_\pi(x), \quad x \in \{0,1\}^n.$$
6. *We have*
$$f(0) + \Phi_\pi(x) \leq f(x), \quad x \in \{0,1\}^n,\, \pi \in \Pi_n.$$

According to Singer [173], the restriction of $\hat{f}$ to $[0,1]^n$ is called the *tight extension* of f associated to the standard triangulation $\{\mathcal{B}_\pi \,|\, \pi \in \Pi_n\}$ of $[0,1]^n$.

A practical form of $\hat{f}$ is the following.

Proposition 4.2.1 *Let f be a pseudo-Boolean function. Then the Lovász extension of f is given by*

$$\hat{f}(x) = \sum_{T \subseteq N} a_T \bigwedge_{i \in T} x_i, \quad x \in (\mathbb{R}^+)^n, \tag{4.10}$$

where a is the Möbius representation of f.

Proof. The function (4.10) agrees with f at all the vertices of $[0,1]^n$, and identifies with an affine function on each cone $\mathcal{K}_\pi$, which is sufficient. ∎

[5] Of course, continuity is ensured when passing from a cone to another.

4.2.2 Properties of the Choquet integral and equivalent forms

Let μ be a fuzzy measure on N. By (4.9), we immediately see that the Choquet integral $\mathcal{C}_\mu$, defined on $(\mathbb{R}^+)^n$, is nothing else than the Lovász extension of the pseudo-Boolean function f_μ which represents μ:

$$\mathcal{C}_\mu = \hat{f}_\mu \quad \text{on } (\mathbb{R}^+)^n.$$

Moreover, since the Choquet integral fulfils (SPL), by Proposition 3.5.1, we can define it on E^n, where $E \supseteq [0,1]$.

Thus, the Choquet integral is a piecewise affine function on E^n and we have

$$\mathcal{C}_\mu(e_S) = \mu_S, \quad S \subseteq N.$$

Moreover, we clearly see that $\mathcal{C}_\mu$ is an increasing function if and only if μ is as well.

Proposition 4.2.1 can be rewritten as follows, see also Chateauneuf and Jaffray [25].

Proposition 4.2.2 *Assume $E \supseteq [0,1]$. Any Choquet integral $\mathcal{C}_\mu : E^n \to \mathbb{R}$ can be written as*

$$\mathcal{C}_\mu(x) = \sum_{T\subseteq N} a_T \bigwedge_{i\in T} x_i, \quad x \in E^n, \tag{4.11}$$

where a is the Möbius representation of μ.

Of course, the set function a occuring in (4.11) is uniquely determined, that is

$$\sum_{T\subseteq N} a_T \bigwedge_{i\in T} x_i = \sum_{T\subseteq N} a'_T \bigwedge_{i\in T} x_i, \quad x \in E^n \quad \Leftrightarrow \quad a = a'.$$

Moreover, any Choquet integral can also be put in the form

$$\sum_{T\subseteq N} \alpha_T \bigvee_{i\in T} x_i.$$

To see this, we need a lemma.

Lemma 4.2.1 *For all $x \in \mathbb{R}^n$, we have*

$$\bigvee_{i=1}^n x_i = \sum_{\substack{K\subseteq N\\K\neq\emptyset}} (-1)^{k+1} \bigwedge_{i\in K} x_i \quad \text{and} \quad \bigwedge_{i=1}^n x_i = \sum_{\substack{K\subseteq N\\K\neq\emptyset}} (-1)^{k+1} \bigvee_{i\in K} x_i.$$

Proof. Let us prove the first identity. The other one can be treated similarly. Use induction on $n \in \mathbb{N}_0$. The result is true if $n = 1$ or $n = 2$. Suppose that it holds for $n \leq p$ and let us show that it holds for $n = p+1$. We have,

$$\begin{aligned}
\bigvee_{i=1}^{p+1} x_i &= \bigvee_{i=1}^{p} x_i + x_{p+1} - \Big(\bigvee_{i=1}^{p} x_i\Big) \wedge x_{p+1} \\
&= \bigvee_{i=1}^{p} x_i + x_{p+1} - \bigvee_{i=1}^{p} (x_i \wedge x_{p+1}) \\
&= \sum_{\substack{K\subseteq N_p\\K\neq\emptyset}} (-1)^{k+1} \bigwedge_{i\in K} x_i + x_{p+1} - \sum_{\substack{K\subseteq N_p\\K\neq\emptyset}} (-1)^{k+1} \bigwedge_{i\in K\cup\{p+1\}} x_i
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{p+1} x_i + \sum_{\substack{K\subseteq N_p\\ k\geq 2}} (-1)^{k+1} \bigwedge_{i\in K} x_i + \sum_{\substack{K\subseteq N_p\\ k\geq 1}} (-1)^{|K\cup\{p+1\}|+1} \bigwedge_{i\in K\cup\{p+1\}} x_i \\
&= \sum_{i=1}^{p+1} x_i + \sum_{\substack{K\subseteq N_{p+1}\\ k\geq 2}} (-1)^{k+1} \bigwedge_{i\in K} x_i \\
&= \sum_{\substack{K\subseteq N_{p+1}\\ K\neq\emptyset}} (-1)^{k+1} \bigwedge_{i\in K} x_i.
\end{aligned}
$$

■

Lemma 4.2.1 must be compared to the well-known Poincaré formula which can be found in probability theory: If $E_1, \ldots, E_n$ are random events then we have

$$\Pr(\bigcup_{i=1}^{n} E_i) = \sum_{\substack{K\subseteq N\\ K\neq\emptyset}} (-1)^{k+1} \Pr(\bigcap_{i\in K} E_i).$$

Proposition 4.2.3 *Let a and α be set functions $2^N \to \mathbb{R}$. Then the following three assertions are equivalent.*

$$
\begin{aligned}
&i) && \sum_{T\subseteq N} a_T \bigwedge_{i\in T} x_i = \sum_{T\subseteq N} \alpha_T \bigvee_{i\in T} x_i, \quad x\in E^n;\\
&ii) && a_\emptyset = \alpha_\emptyset \quad \textit{and} \quad a_T = (-1)^{t+1} \sum_{K\supseteq T} \alpha_K, \quad T\neq\emptyset;\\
&iii) && \alpha_\emptyset = a_\emptyset \quad \textit{and} \quad \alpha_T = (-1)^{t+1} \sum_{K\supseteq T} a_K, \quad T\neq\emptyset.
\end{aligned}
$$

Proof. Let us prove the equivalence between $i)$ and $ii)$. The equivalence between $i)$ and $iii)$ can be treated similarly.

We simply have, by Lemma 4.2.1,

$$
\begin{aligned}
\sum_{T\subseteq N} \alpha_T \bigvee_{i\in T} x_i &= \alpha_\emptyset + \sum_{\substack{T\subseteq N\\ T\neq\emptyset}} \alpha_T \sum_{\substack{K\subseteq T\\ K\neq\emptyset}} (-1)^{k+1} \bigwedge_{i\in K} x_i \\
&= \alpha_\emptyset + \sum_{\substack{K\subseteq N\\ K\neq\emptyset}} \Big[(-1)^{k+1} \sum_{T\supseteq K} \alpha_T\Big] \bigwedge_{i\in K} x_i.
\end{aligned}
$$

We conclude by the uniqueness of the coefficients in (4.11). ■

It is also worth noting that if μ is a possibility measure π, defined by

$$\pi(S) = \bigvee_{i\in S} \pi(i), \quad S\subseteq N,$$

then, by (4.4) and Lemma 4.2.1, we have, for all $T\subseteq N$,

$$a_T = (-1)^{t+1} \sum_{\substack{K\subseteq T\\ K\neq\emptyset}} (-1)^{k+1} \bigvee_{i\in K} \pi(i) = (-1)^{t+1} \bigwedge_{i\in T} \pi(i),$$

but this result has been known for a long time in possibility theory and evidence theory, see Shafer [167].

As it can be easily verified, the Choquet integral fulfils the following aggregation properties [80]: (Co), (In), (UIn), (Id), (Comp), (SPL). We shall show below that it also fulfils (CoAdd) and (BOM).

4.2.3 Some axiomatic characterizations

The Choquet integrals have become very popular in the field of fuzzy sets and multicriteria decision making. This is the reason why searching interpretable characterizations of this class of operators seems relevant. As already mentioned, an axiomatic characterization of any class of operators should not be reduced to a mathematical game. It must either reveal an underlying behavior that was not explicit in the operators, or provide a useful characterization.

Before going on, recall the following notations, already used in Section 3.5: for any $M \in A_n(E, \mathbb{R})$, with $E \supseteq [0,1]$, we set

$$\theta_S := M(e_S) \quad \text{and} \quad \overline{\theta}_S := M(\overline{e}_S), \quad S \subseteq N.$$

First, we start with Proposition 3.5.2, which can be rewritten as follows (see also Figure 3.9).

Theorem 4.2.2 *Assume $E \supseteq [0,1]$. A two-place function $M : E^2 \to \mathbb{R}$ fulfils (In, SPL) if and only if there exists a fuzzy measure μ on $\{1,2\}$ such that $M = \mathcal{C}_\mu$.*

Next, by adapting Theorems 3.5.1 and 3.5.6, we also have the following result.

Theorem 4.2.3 *Assume $E \supseteq [0,1]$ and consider a two-place Choquet integral $\mathcal{C}_\mu : E^2 \to \mathbb{R}$. Then*

i) $\mathcal{C}_\mu$ *fulfils (B) or (AD) if and only if* $\mathcal{C}_\mu \in \{\min, \max\} \cup \{\mathrm{WAM}_\omega \mid \omega \in [0,1]^2\}$

ii) $\mathcal{C}_\mu$ *fulfils (A) if and only if* $\mathcal{C}_\mu \in \{\min, \max, \mathrm{P}_1, \mathrm{P}_2\}$.

Of course, all the n-place operators fulfilling (In, SPL) are not Choquet integrals. For instance, the operator

$$M(x) = \Big(\frac{x_1 + x_2}{2}\Big) \wedge x_3, \quad x \in E^3,$$

fulfils (In, SPL) but corresponds to no Choquet integral; indeed, the Lovász extension of the pseudo-Boolean function $f : \{0,1\}^3 \to \mathbb{R}$ defined by

$$f(x) = (\frac{x_1 + x_2}{2}) \wedge x_3 = \frac{1}{2}(x_1 \wedge x_3) + \frac{1}{2}(x_2 \wedge x_3), \quad x \in \{0,1\}^3,$$

is given by

$$\hat{f}(x) = \frac{1}{2}(x_1 \wedge x_3) + \frac{1}{2}(x_2 \wedge x_3)$$

and we have $\hat{f}(0,1,1/2) = 1/4 \neq 1/2 = M(0,1,1/2)$.

The class of n-place Choquet integrals has been characterized by Schmeidler [160], see also [33], [77], and [89, Theorem 8.6]. We present below a slightly different statement.

Theorem 4.2.4 *Assume $E \supseteq [0,1]$. $M \in A_n(E, \mathbb{R})$ fulfils (In, SPL, CoAdd) if and only if there exists a fuzzy measure μ on N such that $M = \mathcal{C}_\mu$.*

Proof. (Sufficiency) Let us show that $\mathcal{C}_\mu$ fulfils (CoAdd), see also [148]. Consider two comonotonic vectors $x, x' \in E^n$. Then we have

$$\begin{aligned}
\mathcal{C}_\mu(x + x') &= \sum_{i=1}^{n} (x_{(i)} + x'_{(i)})\,[\mu_{\{(i),\ldots,(n)\}} - \mu_{\{(i+1),\ldots,(n)\}}] \\
&= \sum_{i=1}^{n} x_{(i)}\,[\mu_{\{(i),\ldots,(n)\}} - \mu_{\{(i+1),\ldots,(n)\}}] + \sum_{i=1}^{n} x'_{(i)}\,[\mu_{\{(i),\ldots,(n)\}} - \mu_{\{(i+1),\ldots,(n)\}}] \\
&= \mathcal{C}_\mu(x) + \mathcal{C}_\mu(x').
\end{aligned}$$

(Necessity) By Proposition 3.5.1, we can assume that $M \in A_n([0,1], \mathbb{R})$. Fix $\pi \in \Pi_n$ and set $x \in \mathcal{B}_\pi$. By (SPL), we have

$$\begin{aligned} M(x) &= M([x_{\pi(1)}, \ldots, x_{\pi(n)}]_{\pi^{-1}}) \\ &= M([0, x_{\pi(2)} - x_{\pi(1)}, \ldots, x_{\pi(n)} - x_{\pi(1)}]_{\pi^{-1}}) + x_{\pi(1)}. \end{aligned}$$

By (CoAdd) and (SPL), we have, for all $i \in \{2, \ldots, n\}$,

$$\begin{aligned} & M([\underbrace{0, \ldots, 0}_{i-1}, x_{\pi(i)} - x_{\pi(i-1)}, \ldots, x_{\pi(n)} - x_{\pi(i-1)}]_{\pi^{-1}}) \\ = \; & M([\underbrace{0, \ldots, 0}_{i-1}, x_{\pi(i)} - x_{\pi(i-1)}, \ldots, x_{\pi(i)} - x_{\pi(i-1)}]_{\pi^{-1}}) \\ & + M([\underbrace{0, \ldots, 0, 0}_{i}, x_{\pi(i+1)} - x_{\pi(i)}, \ldots, x_{\pi(n)} - x_{\pi(i)}]_{\pi^{-1}}) \\ = \; & (x_{\pi(i)} - x_{\pi(i-1)})\, \theta_{\{\pi(i), \ldots, \pi(n)\}} + M([\underbrace{0, \ldots, 0, 0}_{i}, x_{\pi(i+1)} - x_{\pi(i)}, \ldots, x_{\pi(n)} - x_{\pi(i)}]_{\pi^{-1}}) \end{aligned}$$

and thus, recursively,

$$M(x) = \sum_{i=2}^{n} (x_{\pi(i)} - x_{\pi(i-1)})\, \theta_{\{\pi(i), \ldots, \pi(n)\}} + x_{\pi(1)} = \mathcal{C}_\mu(x),$$

with $\mu_T = M(e_T) = \theta_T$ for all $T \subseteq N$. Moreover, by (In), μ is a fuzzy measure. ∎

We now intend to show that the class of Choquet integrals can also be characterized by (In, SPL, BOM). For this purpose, we need two lemmas.

Lemma 4.2.2 *Let $n \in \mathbb{N}_0$, $n \geq 2$. If $M \in A_n([0,1],[0,1])$ fulfils (In, SPL, BOM) and if there exists $S \subseteq N$ such that $\theta_S \in]0,1[$ then M fulfils (CoAdd).*

Proof. The proof is simply an adaptation of that of Lemma 3.5.1. Let x, y be two comonotonic vectors in $[0,1]^n$ with $x_i + y_i \in [0,1]$ for all $i \in N$. Thus there exists $\pi \in \Pi_n$ such that

$$x_{\pi(1)} \leq \cdots \leq x_{\pi(n)} \quad \text{and} \quad y_{\pi(1)} \leq \cdots \leq y_{\pi(n)}.$$

Let us show that

$$M(x+y) = M(x) + M(y). \tag{4.12}$$

As in the proof of Lemma 3.5.1, we can assume that $x_i, y_i \leq \inf\{\theta_S, 1 - \theta_S\}$ for all $i \in N$. Let $\pi' \in \Pi_n$ such that $\pi'(i) > \pi'(j)$ for all $i \in S$ and all $j \notin S$, and consider the ordered square matrix X of n rows r_i and n columns c_j $(i, j \in N)$, where r_i is defined as follows (for $i \in N$):

$$r_i = \Big[\frac{1}{2}\Big(\frac{x_{\pi(i)}}{\theta_S} + 1\Big) e_S + \frac{1}{2}\,\frac{y_{\pi(i)}}{1 - \theta_S}\, \overline{e}_S\Big]_{\pi'}.$$

On the one hand, by (SPL), we have, for all $i \in N$:

$$M([r_i]_{\pi'^{-1}}) = \frac{1}{2}\,\frac{y_{\pi(i)}}{1-\theta_S} + \frac{1}{2}\Big(\frac{x_{\pi(i)}}{\theta_S} + 1 - \frac{y_{\pi(i)}}{1-\theta_S}\Big)\theta_S = \frac{x_{\pi(i)} + y_{\pi(i)}}{2} + \frac{\theta_S}{2},$$

and thus

$$M([M([r_1]_{\pi'^{-1}}), \ldots, M([r_n]_{\pi'^{-1}})]_{\pi^{-1}}) = \frac{1}{2}\, M(x+y) + \frac{\theta_S}{2}.$$

On the other hand, for all $j \in N$, we have

$$M([c_j]_{\pi^{-1}}) = \begin{cases} \frac{1}{2}(\frac{M(x)}{\theta_S} + 1), & \text{if } j \in \pi'(S), \\ \frac{1}{2}\frac{M(y)}{1-\theta_S}, & \text{otherwise.} \end{cases}$$

However, by (2.6) and (2.2), M fulfils (Comp). Hence, by (SPL), we have,

$$\begin{aligned} M([M([c_1]_{\pi^{-1}}), \ldots, M([c_n]_{\pi^{-1}})]_{\pi'^{-1}}) &= M\Big[\frac{1}{2}\Big(\frac{M(x)}{\theta_S} + 1\Big)e_S + \frac{1}{2}\frac{M(y)}{1-\theta_S}\,\overline{e}_S\Big] \\ &= \frac{1}{2}\,[M(x) + M(y)] + \frac{\theta_S}{2}. \end{aligned}$$

Since M fulfils (BOM), we have (4.12). ∎

Lemma 4.2.3 *Let $n \in \mathbb{N}_0$, $n \geq 2$. If $M \in A_n([0,1], \mathbb{R})$ fulfils (In, SPL) and if $\theta_S \in \{0,1\}^n$ for all $S \subseteq N$, then there exists a 0-1 fuzzy measure μ on N such that $M = \mathcal{C}_\mu$.*

Proof. Let $\pi \in \Pi_n$. By (In), there exists $k \in \{0, \ldots, n-1\}$ such that

$$\theta_{\{\pi(i), \ldots, \pi(n)\}} = \begin{cases} 1, & \forall i \in \{1, \ldots, k\}, \\ 0, & \forall i \in \{k+1, \ldots, n\}. \end{cases}$$

Let $x \in \mathcal{B}_\pi$. By (In) and (SPL), we have

$$\begin{aligned} M([x_{\pi(1)}, \ldots, x_{\pi(n)}]_{\pi^{-1}}) &\leq M([k \odot x_{\pi(k)}, (n-k) \odot x_{\pi(n)}]_{\pi^{-1}}) \\ &= x_{\pi(k)} + M([k \odot 0, (n-k) \odot (x_{\pi(n)} - x_{\pi(k)})]_{\pi^{-1}}) \\ &= x_{\pi(k)} \end{aligned}$$

and

$$\begin{aligned} M([x_{\pi(1)}, \ldots, x_{\pi(n)}]_{\pi^{-1}}) &\geq M([(k-1) \odot x_{\pi(1)}, (n-k+1) \odot x_{\pi(k)}]_{\pi^{-1}}) \\ &= M([(k-1) \odot 0, (n-k+1) \odot (x_{\pi(k)} - x_{\pi(1)})]_{\pi^{-1}}) + x_{\pi(1)} \\ &= x_{\pi(k)} - x_{\pi(1)} + x_{\pi(1)} \\ &= x_{\pi(k)}. \end{aligned}$$

Thus, we have

$$M(x) = M([x_{\pi(1)}, \ldots, x_{\pi(n)}]_{\pi^{-1}}) = x_{\pi(k)} = \mathcal{C}_\mu(x)$$

with $\mu_{\{\pi(k), \ldots, \pi(n)\}} - \mu_{\{\pi(k+1), \ldots, \pi(n)\}} = 1$. ∎

Theorem 4.2.5 *Assume $E \supseteq [0,1]$. $M \in A_n(E, E)$ fulfils (In, SPL, BOM) if and only if there exists a fuzzy measure μ on N such that $M = \mathcal{C}_\mu$.*

Proof. (Sufficiency) Let us show that $\mathcal{C}_\mu$ fulfils (BOM). Observe first that if $x_1 \leq \cdots \leq x_n$ then, for any $\pi \in \Pi_n$, we have

$$\mathcal{C}_\mu([x_1, \ldots, x_n]_\pi) = \sum_{i=1}^{n} x_i\,[\mu_{\{\pi^{-1}(i), \ldots, \pi^{-1}(n)\}} - \mu_{\{\pi^{-1}(i+1), \ldots, \pi^{-1}(n)\}}].$$

Now consider an ordered square matrix $X \in E^{n\times n}$ and $\pi, \pi' \in \Pi_n$. By (In), we have

$$\mathcal{C}_\mu([x_{11},\ldots,x_{1n}]_{\pi'}) \leq \cdots \leq \mathcal{C}_\mu([x_{n1},\ldots,x_{nn}]_{\pi'})$$

and by the previous identity, we have

$$\begin{aligned}
&\mathcal{C}_\mu([\mathcal{C}_\mu([x_{11},\ldots,x_{1n}]_{\pi'}),\ldots,\mathcal{C}_\mu([x_{n1},\ldots,x_{nn}]_{\pi'})]_\pi)\\
=\;& \sum_{i=1}^n \mathcal{C}_\mu([x_{i1},\ldots,x_{in}]_{\pi'})\,[\mu_{\{\pi^{-1}(i),\ldots,\pi^{-1}(n)\}} - \mu_{\{\pi^{-1}(i+1),\ldots,\pi^{-1}(n)\}}]\\
=\;& \sum_{i,j=1}^n x_{ij}\,[\mu_{\{\pi'^{-1}(j),\ldots,\pi'^{-1}(n)\}} - \mu_{\{\pi'^{-1}(j+1),\ldots,\pi'^{-1}(n)\}}][\mu_{\{\pi^{-1}(i),\ldots,\pi^{-1}(n)\}} - \mu_{\{\pi^{-1}(i+1),\ldots,\pi^{-1}(n)\}}]
\end{aligned}$$

which does not change when permuting π and π'. Hence $\mathcal{C}_\mu$ fulfils (BOM).

(Necessity) By Proposition 3.5.1, we can assume that $M \in A_n([0,1],[0,1])$. If there exists $S \subseteq N$ such that $\theta_S \in]0,1[$ then, by Lemma 4.2.2, M fulfils (CoAdd) and, by Theorem 4.2.4, M is a Choquet integral.

Otherwise, if $\theta_S \in \{0,1\}^n$ for all $S \subseteq N$, then we can immediately conclude by Lemma 4.2.3. ■

The following characterization has been suggested by D. Dubois.

Theorem 4.2.6 *Assume* $[0,1] \subseteq E \subseteq \mathbb{R}^+$. *The Choquet integrals on* E^n *are exactly those* $M \in A_n(E,\mathbb{R})$ *which fulfil (In, WId) and*

$$M(\lambda\, x + (1-\lambda)\, x') = \lambda\, M(x) + (1-\lambda)\, M(x'), \quad \lambda \in [0,1], \tag{4.13}$$

for all comonotonic vectors $x, x' \in E^n$.

Proof. (Sufficiency) Trivial.

(Necessity) It is clear that, by (4.13), M is an affine function on each set $\mathcal{K}_\pi \cap E^n$ $(\pi \in \Pi_n)$, and thus is the Lovász extension of a pseudo-Boolean function f. By (In, WId), f corresponds to a fuzzy measure, which is sufficient. ■

By Theorem 3.5.2, we also have the following result.

Theorem 4.2.7 *Assume* $E \supseteq [0,1]$. *The Choquet integral* $\mathcal{C}_\mu \in A_n(E,E)$ *fulfils (B) if and only if*

$$\mathcal{C}_\mu \in \{\min_S, \max_S \mid S \subseteq N\} \cup \{\mathrm{WAM}_\omega \mid \omega \in [0,1]^n\}.$$

An *extended Choquet integral* $\mathcal{C}$ is an extended aggregation operator $M \in A(E,\mathbb{R})$ such that, for all $n \in \mathbb{N}_0$, $M^{(n)}$ is a Choquet integral. Concerning such extended aggregation operators, we present the following result.

Theorem 4.2.8 *Assume* $E \supseteq [0,1]$. $M \in A(E,E)$ *fulfils (In, SPL, GBOM) if and only if* M *is an extended Choquet integral.*

Proof. (Sufficiency) The proof is similar to that of Theorem 4.2.5.

(Necessity) It is clear that, for all $n \in \mathbb{N}_0$, $M^{(n)}$ fulfils (In, SPL, BOM). We then conclude by Theorem 4.2.5. ■

Theorems 3.5.3, 3.5.4, 3.5.5 and 3.5.7 provide successively the following results.

Theorem 4.2.9 *Assume $E \supseteq [0,1]$. $\mathcal{C} \in A(E,E)$ is an extended Choquet integral and fulfils (GB) if and only if*

- *either: for all $n \in \mathbb{N}_0$, there exists $S \subseteq N_n$ such that $\mathcal{C}^{(n)} = \min_S$,*
- *or: for all $n \in \mathbb{N}_0$, there exists $S \subseteq N_n$ such that $\mathcal{C}^{(n)} = \max_S$,*
- *or: for all $n \in \mathbb{N}_0$, there exists $\omega \in [0,1]^n$ such that $\mathcal{C}^{(n)} = \mathrm{WAM}_\omega$.*

Theorem 4.2.10 *Assume $E \supseteq [0,1]$. $\mathcal{C} \in A(E,E)$ is an extended Choquet integral and fulfils (D) if and only if*

- *either: $\mathcal{C} = (\min^{(n)})_{n\in\mathbb{N}_0}$,*
- *or: $\mathcal{C} = (\max^{(n)})_{n\in\mathbb{N}_0}$,*
- *or: there exists $\theta \in [0,1]$ such that, for all $n \in \mathbb{N}_0$, we have $\mathcal{C}^{(n)} = \mathrm{WAM}_\omega$ with*

$$\omega_i = \frac{(1-\theta)^{n-i}\,\theta^{i-1}}{\sum_{j=1}^{n}(1-\theta)^{n-j}\,\theta^{j-1}}, \quad \forall i \in N_n.$$

Theorem 4.2.11 *Assume $E \supseteq [0,1]$. $\mathcal{C} \in A(E,E)$ is an extended Choquet integral and fulfils (SD) if and only if*

$$\mathcal{C} = (\min^{(n)})_{n\in\mathbb{N}_0} \text{ or } (\max^{(n)})_{n\in\mathbb{N}_0} \text{ or } (P_1^{(n)})_{n\in\mathbb{N}_0} \text{ or } (P_n^{(n)})_{n\in\mathbb{N}_0} \text{ or } (\mathrm{AM}^{(n)})_{n\in\mathbb{N}_0}.$$

Theorem 4.2.12 *Assume $E \supseteq [0,1]$. $\mathcal{C} \in A(E,E)$ is an extended Choquet integral and fulfils (A) if and only if*

$$\mathcal{C} = (\min^{(n)})_{n\in\mathbb{N}_0} \text{ or } (\max^{(n)})_{n\in\mathbb{N}_0} \text{ or } (P_1^{(n)})_{n\in\mathbb{N}_0} \text{ or } (P_n^{(n)})_{n\in\mathbb{N}_0}.$$

4.2.4 Weighted arithmetic means

The best known and most often used weighted mean operator in many applications is the weighted arithmetic mean operator (WAM), defined by (1.2).

It is easy to prove that WAM operators fulfil the following properties: (Co), (In), (UIn), (Id), (Comp), (SPL), (SSN), (Add), (B), (Sep). Moreover, one can readily see that any WAM_ω is a Choquet integral $\mathcal{C}_\mu$ with respect to an additive fuzzy measure (probability measure):

$$\mu_T = \sum_{i\in T}\omega_i, \quad T \subseteq N.$$

The corresponding Möbius representation is given by:

$$\begin{cases} a_i = \omega_i, & \forall i \in N, \\ a_T = 0, & \forall T \subseteq N \text{ such that } |T| \neq 1, \end{cases}$$

As a consequence, we can see that the weighted arithmetic means are the additive Choquet integrals.

Theorem 4.2.13 *The Choquet integral $\mathcal{C}_\mu \in A_n(E,\mathbb{R})$ fulfils (Add) if and only if there exists $\omega \in [0,1]^n$ such that $\mathcal{C}_\mu = \mathrm{WAM}_\omega$.*

Using Theorem 4.2.4, we also have the following corollary.

Corollary 4.2.1 *Assume $E \supseteq [0,1]$. $M \in A_n(E,\mathbb{R})$ fulfils (In, SPL, Add) if and only if there exists $\omega \in [0,1]^n$ such that $M = \mathrm{WAM}_\omega$.*

The class of WAM operators includes two important special cases, namely:

- the arithmetic mean AM, when $\omega_i = 1/n$ for all i,
- the k-th projection P_k, when $\omega_k = 1$.

It is clear that a WAM_ω function fulfils (Sy) if and only if $\omega_i = 1/n$ for all i (arithmetic mean). It fulfils (SIn) if and only if $\omega_i > 0$ for all i.

The family of weighted arithmetic means can be characterized from some results of Section 3.5. For example, adding (SSN) to the characterizations of Proposition 3.5.2 and Theorem 3.5.2 leads to the following two corollaries (recall that, by Proposition 2.2.2, (SSi, SSN) implies (SPL)).

Corollary 4.2.2 *$M \in A_2([0,1],[0,1])$ fulfils (In, SSi, SSN) if and only if there exists $\omega \in [0,1]^2$ such that $M = \mathrm{WAM}_\omega$.*

Corollary 4.2.3 *$M \in A_n([0,1],[0,1])$ fulfils (In, SSi, SSN, B) if and only if there exists $\omega \in [0,1]^n$ such that $M = \mathrm{WAM}_\omega$.*

From Corollary 3.5.3, we immediately deduce the following characterization of the arithmetic mean.

Corollary 4.2.4 *Assume $E \supseteq [0,1]$. $M \in A_n(E,E)$ fulfils (Sy, SIn, SPL, B) if and only if $M = \mathrm{AM}$.*

In addition to the previous results, some characterizations of the family of weighted arithmetic means are presented in Aczél [4, Sect. 5.3.1] (see also Proposition 9 in [7, Chapter 15]).

Proposition 4.2.4 *The function $M \in A_2(\mathbb{R},\mathbb{R})$ has the properties*

$$M(x+s, y+s) = M(x,y) + s \quad \textit{and} \quad M(r\,x, r\,y) = r\,M(x,y) \quad (x,y,s,r \in \mathbb{R}, r \neq 0)$$

if and only if, there exists $\theta \in \mathbb{R}$ such that

$$M(x,y) = (1-\theta)\,x + \theta\,y \quad (x,y \in \mathbb{R}).$$

If M is symmetric, then $\theta = 1/2$.

Proposition 4.2.5 *$M \in A_n(\mathbb{R},\mathbb{R})$ fulfils (Co, Add) and is such that $M(x_0,\ldots,x_0) = x_0$ for an $x_0 \in \mathbb{R}_0$ if and only if there exists $\omega \in \mathbb{R}^n$ with $\sum_i \omega_1 = 1$ such that*

$$M(x) = \sum_{i=1}^{n} \omega_i\,x_i, \quad x \in \mathbb{R}^n.$$

In the previous two propositions, it is clear that if M fulfils (In) then the weights θ and ω_i are non-negative.

Proposition 4.2.6 *$M \in A_n(\mathbb{R},\mathbb{R})$ fulfils (Sy, Id, Add) if and only if $M = \mathrm{AM}$.*

The following result can be extracted from [9, pp. 413–414].

Proposition 4.2.7 *Let $E = \mathbb{R}$ or $\mathbb{R}^+$. If $M \in A_n(E,\mathbb{R})$ fulfils (SSi, Add) then*

$$M(x) = \sum_{i=1}^{n} a_i\,x_i, \quad \forall x \in E^n,$$

where $a_1,\ldots,a_n$ are arbitrary constants.

From Theorem 3.4.7, we can deduce the following.

Proposition 4.2.8 *$M \in A_n(\mathbb{R}, \mathbb{R})$ fulfils (Id, ISUII) if and only if*

$$M(x) = \sum_{i=1}^{n} a_i\, x_i, \quad x \in \mathbb{R}^n,$$

with $\sum_i a_i = 1$, $a_i \in \mathbb{R}$. If, moreover, M fulfils (In) then $a_i \geq 0$. If M fulfils (Sy) then $a_i = 1/n$.

For extended aggregation operators, we also have the following two corollaries. They can be immediately deduced from Theorem 3.5.3 and Corollaries 3.5.4 and 3.5.5.

Corollary 4.2.5 *$M \in A([0,1],[0,1])$ fulfils (In, SSi, SSN, GB) if and only if, for all $n \in \mathbb{N}_0$, there exists $\omega \in [0,1]^n$ such that $M^{(n)} = \mathrm{WAM}_\omega$.*

Corollary 4.2.6 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (Sy, SIn, SPL) and (D or SD or GB) if and only if, for all $n \in \mathbb{N}_0$, $M^{(n)} = \mathrm{AM}$.*

Theorem 4.2.14 *Let $\theta \in \mathbb{R}$. $M \in A(\mathbb{R}, \mathbb{R})$ fulfils (Id, D) and is such that*

$$M(x_1, x_2) = (1-\theta)\, x_1 + \theta\, x_2, \quad (x_1, x_2) \in \mathbb{R}^2,$$

if and only if, for all $n \in \mathbb{N}_0$,

$$M(x) = \frac{1}{D_n} \sum_{i=1}^{n} (1-\theta)^{n-i}\, \theta^{i-1} x_i, \quad x \in \mathbb{R}^n,$$

where $D_n = \sum_{j=1}^{n} (1-\theta)^{n-j}\, \theta^{j-1}$.

Proof. (Sufficiency) We have already observed in Theorem 3.5.4 that M fulfils (Id, D).

(Necessity) The proof (by induction over n) is an adaptation of that of Theorem 3 in [4, Sect. 5.3.1]. Suppose the result true for $n \geq 2$. By (D), we have

$$\begin{aligned} M^{(n+1)}(x_1, \ldots, x_{n+1}) &= M^{(n+1)}(n \odot M^{(n)}(x_1, \ldots, x_n), x_{n+1}) \\ &= G(M^{(n)}(x_1, \ldots, x_n), x_{n+1}) \\ &= M^{(n+1)}(x_1, n \odot M^{(n)}(x_2, \ldots, x_{n+1})), \end{aligned}$$

or (if, for example, $\theta \neq 0$) with $x_1 = \cdots = x_{n-1} = 0$,

$$\begin{aligned} G\Big(\frac{\theta^{n-1} x_n}{D_n}, x_{n+1}\Big) &= M^{(n+1)}\Big(0, n \odot \frac{(1-\theta)\,\theta^{n-2} x_n + \theta^{n-1} x_{n+1}}{D_n}\Big) \\ &= f\Big(\frac{(1-\theta)\,\theta^{n-2} x_n + \theta^{n-1} x_{n+1}}{D_n}\Big), \end{aligned}$$

therefore

$$G(x, x_{n+1}) = f\Big(\frac{1-\theta}{\theta}\, x + \frac{\theta^{n-1} x_{n+1}}{D_n}\Big).$$

Thus

$$\begin{aligned} M^{(n+1)}(x_1, \ldots, x_{n+1}) &= G(M^{(n)}(x_1, \ldots, x_n), x_{n+1}) \\ &= f\Big(\frac{1-\theta}{\theta}\, M^{(n)}(x_1, \ldots, x_n) + \frac{\theta^{n-1} x_{n+1}}{D_n}\Big) \\ &= f\Big(\frac{1}{D_n} \sum_{i=1}^{n+1} (1-\theta)^{n+1-i} \theta^{i-2} x_i\Big) \end{aligned}$$

and finally because of (Id)

$$x = M^{(n+1)}(x, \dots, x) = f\Big(\frac{D_{n+1}}{\theta\, D_n}\, x\Big),$$

that is,

$$f(t) = \frac{\theta\, D_n}{D_{n+1}}\, t,$$
$$M^{(n+1)}(x_1, \dots, x_{n+1}) = \frac{1}{D_{n+1}} \sum_{i=1}^{n+1} (1-\theta)^{n+1-i} \theta^{i-1} x_i,$$

which was to be proved. ■

Yager and Rybalov [198] investigated the extended weighted arithmetic means that fulfil the self-identity property. Recall that self-identity (SId) relates the aggregation process at cardinalities n and $n+1$ in such a way that adding as the $(n+1)$-th element the aggregated value of the previous n elements must give the original aggregated value.

Theorem 4.2.15 *Let $M \in A(E, E)$ defined by $M^{(n)} = \mathrm{WAM}_{\omega^{(n)}}$ for all $n \in \mathbb{N}_0$. Then M fulfils (SId) if and only if, for all $n \in \mathbb{N}_0$, we have*

$$1 = \omega_1^{(1)} \geq \ldots \geq \omega_1^{(n)}$$

and

$$\omega_i^{(n)} = \omega_1^{(n)} \Big(\frac{1}{\omega_1^{(i)}} - \frac{1}{\omega_1^{(i-1)}}\Big), \quad i = 2, \dots, n.$$

We see that under (SId), the extended weighted arithmetic mean is uniquely determined by selecting $\omega_1^{(2)}, \dots, \omega_1^{(n)}$. The weights $\omega_i^{(n)}$ are then defined by the ratio of $\omega_1^{(n)}$ with first weights in the previous aggregations.

As a particular case, we can consider the extended arithmetic mean, for which we have $\omega_i^{(n)} = \frac{1}{n}$. We might also imagine that there exists $\alpha \in [0, 1]$ such that

$$\omega_1^{(i)} = \alpha\, \omega_1^{(i-1)}, \quad i = 2, \dots, n.$$

That is, we require that the first weight in each succeeding aggregation is reduced by α proportion. In this case we have $\omega_1^{(n)} = \alpha^{n-1}$ and

$$\omega_i^{(n)} = \alpha^{n-i}(1 - \alpha), \quad i = 2, \dots, n.$$

The following recursive formula also holds:

$$M(x_1, \dots, x_{n+1}) = \alpha\, M(x_1, \dots, x_n) + (1 - \alpha)\, x_{n+1}.$$

We note this form of aggregation known as exponential smoothing [22] is a classic non-symmetric type of linear aggregation.

4.2.5 Ordered weighted averaging operators

The ordered weighted averaging aggregation operators (OWA), defined by (1.3), were proposed in 1988 by Yager [192]. Since their introduction, they have been applied to many fields as neural networks [190, 194], data base systems [191, 195], fuzzy logic controllers [60, 193], and group decision making [31, 192]. The OWA operators can also be used in decision making under uncertainty to modelize the anticipated utility [147, 166].

Their structural properties [174] and their links with fuzzy integrals [77, 86] were also investigated. For a recent list of references, see [103].

Thus defined, an OWA operator is given explicitly by a weighted arithmetic mean of order statistics. A fundamental aspect of such an operator is the re-ordering step, in particular a score x_i is not associated with a particular weight ω_i, but rather a weight is associated with a particular ordered position of score. This ordering step introduces a non-linearity into the aggregation process.

Example 4.2.1 Assume $\omega = (0.4, 0.3, 0.2, 0.1)$ and $x = (0.7, 1, 0.3, 0.6)$. In this case carrying out the reordering process we get that $x_{(1)} = 0.3$, $x_{(2)} = 0.6$, $x_{(3)} = 0.7$ and $x_{(4)} = 1$. Then performing the aggregation $\sum_{i=1}^n \omega_i \, x_{(i)}$ we get

$$\mathrm{OWA}_\omega(0.7, 1, 0.3, 0.6) = (0.4)(0.3) + (0.3)(0.6) + (0.2)(0.7) + (0.1)(1) = 0.54$$

OWA operators fulfil a number of well-known and easy-to-prove properties [31, 192], namely: (Sy), (Co), (In), (UIn), (Id), (Comp), (SPL), (CoAdd), (BOM). More precisely, it is a well known fact (see e.g. [68, 135]) that OWA operators are a particular case of discrete Choquet integrals with respect to a fuzzy measure depending only on the cardinal of subsets. In fact, the class of OWA operators coincides with the class of Choquet integrals which fulfil (Sy), see [78, 79]. This result can be stated as follows.

Theorem 4.2.16 *Let μ be a fuzzy measure on N. Then the following assertions are equivalent.*

i) *μ depends only on the cardinality of subsets*

ii) *there exists $\omega \in [0,1]^n$ such that $\mathcal{C}_\mu = \mathrm{OWA}_\omega$*

iii) *$\mathcal{C}_\mu$ fulfils (Sy).*

As Choquet integrals have been thoroughly studied in the context of multicriteria decision problems, the OWA operators can benefit from these studies.

The fuzzy measure μ associated to an OWA_ω is given by

$$\mu_T = \sum_{i=n-t+1}^{n} \omega_i, \quad T \subseteq N, T \neq \emptyset, \tag{4.14}$$

and its Möbius representation by [86, Theorem 1]

$$a_T = \sum_{j=0}^{t-1} \binom{t-1}{j} (-1)^{t-1-j} \, \omega_{n-j}, \quad T \subseteq N, T \neq \emptyset. \tag{4.15}$$

Both representations depend only on the cardinal of the subsets.

Conversely, the weights associated to OWA_ω are given by

$$\omega_{n-t} = \mu_{T \cup i} - \mu_T = \sum_{K \subseteq T} a_{K \cup i}, \quad i \in N, \;\; T \subseteq N \setminus i. \tag{4.16}$$

Now, let us show that the Möbius representation can take a very simple form. Consider the difference operator

$$\Delta_k\, x_k := x_{k+1} - x_k$$

for sequences $(x_k)_{k\in\mathbb{N}}$. It is well known that we have (cf. Berge [19, Chap. 1, Sect. 8])

$$(\Delta_k^t\, x_k)_{k=0} = \sum_{j=0}^{t} \binom{t}{j} (-1)^{t-j}\, x_j, \quad t \in \mathbb{N}. \tag{4.17}$$

By (4.17), we have

$$a_T = (\Delta_k^{t-1}\, \omega_{n-k})_{k=0}, \quad T \subseteq N,\ T \neq \emptyset. \tag{4.18}$$

The class of OWA operators includes some important special cases, namely:

- the min operator, when $\omega_1 = 1$,
- the max operator, when $\omega_n = 1$,
- the arithmetic mean AM, when $\omega_i = 1/n$ for all i,
- the k-th order statistic OS_k, when $\omega_k = 1$,
- the median $(x_{(n/2)} + x_{(n/2+1)})/2$, when n is even and $\omega_{n/2} = \omega_{n/2+1} = 1/2$,
- the median $x_{(\frac{n+1}{2})}$, when n is odd and $\omega_{\frac{n+1}{2}} = 1$,
- the mean excluding the extremes as used by some jury of international olympic competitions, when $\omega_1 = \omega_n = 0$ and $\omega_i = \frac{1}{n-2}$ for $i \neq 1, n$.

The two-place OWA operators have been already characterized in Corollary 3.5.1 by (Sy, In, SPL). For n-place operators, some characterizations can be deduced from those of the Choquet integrals, see also [119].

Theorem 4.2.17 *Assume $E \supseteq [0,1]$ and let $M \in A_n(E, \mathbb{R})$. The following statements are equivalent.*

i) *M fulfils (Sy, In, SPL, CoAdd)*
ii) *M fulfils (Sy, In, SPL, BOM)*
iii) *there exists $\omega \in [0,1]^n$ such that $M = \mathrm{OWA}_\omega$.*

Theorem 4.2.18 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ fulfils (Sy, In, SPL, GBOM) if and only if, for all $n \in \mathbb{N}_0$, there exists $\omega \in [0,1]^n$ such that $M^{(n)} = \mathrm{OWA}_\omega$.*

The following result can also be useful.

Theorem 4.2.19 *i) The OWA_ω operator fulfils (SIn) if and only if $\omega_i > 0$ for all i.*
ii) The OWA_ω operator fulfils (SSN) on $[0,1]^n$ if and only if $\omega_{n+1-i} = \omega_i$ for all i.

Proof. *i*) Trivial.
ii) Let $x \in [0,1]^n$ such that $x_1 \leq \cdots \leq x_n$. Using (SSN), we have

$$\begin{aligned} M(1-x_1,\ldots,1-x_n) = 1 - M(x_1,\ldots,x_n) \quad &\Leftrightarrow \quad \sum_{i=1}^{n} \omega_i\,(1 - x_{n+1-i}) = 1 - \sum_{i=1}^{n} \omega_i\, x_i \\ &\Leftrightarrow \quad \sum_{i=1}^{n} (\omega_{n+1-i} - \omega_i)\, x_i = 0. \end{aligned}$$

Hence the result. ∎

Fodor, Marichal and Roubens [68] defined the quasi-OWA operators as a generalization of the OWA operators. These operators are of the form

$$M(x) = f^{-1}\Big[\sum_{i=1}^{n} \omega_i\, f(x_{(i)})\Big], \quad \text{with} \quad \sum_{i=1}^{n} \omega_i = 1, \;\; \omega_i \geq 0, \tag{4.19}$$

where f is a continuous strictly monotonic function. Properties which are fulfilled by the quasi-OWA operators are: (Sy), (Co), (In), (UIn), (Comp), (BOM).

We then define the extended quasi-OWA operators as follows: M is an extended quasi-OWA operator if there exists a continuous strictly monotonic function f such that, for all $n \in \mathbb{N}_0$, $M^{(n)}$ is of the form (4.19). Note that the extended quasi-OWA operators fulfil (GBOM).

Quasi-OWA operators and extended quasi-OWA operators have still to be characterized but we can prove the following theorem [68].

Theorem 4.2.20 *Assume $E \supseteq [0,1]$. $M \in A(E,E)$ is an extended quasi-*OWA *operator fulfilling (D) if and only if it corresponds to the extended* min *operator or extended* max *operator or extended quasi-arithmetic mean operators.*

Proof. (Sufficiency) Trivial.

(Necessity) Assume that there exists a continuous strictly monotonic function $f : E \to \mathbb{R}$ such that, for each $n \in \mathbb{N}_0$, $M^{(n)}$ is given by (4.19). Define $\Omega := f(E) = \{f(x)\,|\,x \in E\}$. The extended operator $F \in A(\Omega,\Omega)$ defined by

$$F(z_1,\ldots,z_n) := f[M(f^{-1}(z_1),\ldots,f^{-1}(z_n))] = \sum_{i=1}^{n} \omega_i\, z_{(i)}, \quad \forall z \in \Omega^n, \quad \forall n \in \mathbb{N}_0,$$

fulfils (Sy, In, SPL, D). We then can conclude by using Corollary 3.5.5. ■

Ovchinnikov [142] introduced the concept of *weighted order statistic averaging* (WOSA) operator as a compensative operator (Comp) of the form

$$M_\omega(x) = \sum_{i=1}^{n} \omega_i\, x_{(i)}, \tag{4.20}$$

where weights ω_i's are real numbers.

It is evident that not every operator in the form (4.20) is compensative. The following theorem describes the class of WOSA operators [142].

Theorem 4.2.21 *Assume $E \supseteq [0,1]$. An operator $M \in A_n(E,\mathbb{R})$ of the form (4.20) fulfils (Comp) if and only if $\sum_{i=1}^{n} \omega_i = 1$ and, for any k, $0 \leq \sum_{i=k}^{n} \omega_i \leq 1$.*

The class of WOSA operators includes all OWA operators and both classes have been characterized in [142] by means of an analytic property which is the directional differentiability at zero.

Actually the OWA operators are exactly those WOSA operators which fulfil (In). More precisely, we have the following characterization, which is to be compared with Theorem 4.2.17.

Theorem 4.2.22 *Assume $E \supseteq [0,1]$. $M \in A_n(E,\mathbb{R})$ fulfils (Sy, Comp, SPL, CoAdd) if and only if there exists $\omega \in \mathbb{R}^n$ such that $M = \mathrm{WOSA}_\omega$.*

Proof. (Sufficiency) Trivial.

(Necessity) In the proof of Theorem 4.2.4, it has been proved that if $M \in A_n(E, \mathbb{R})$ fulfils (SPL, CoAdd) then it is a Choquet integral with respect to a measure μ satisfying $\mu_\emptyset = 0$ and $\mu_N = 1$, but not necessarily monotonic. By (Sy), M is of the form (4.20). We then conclude by adding (Comp). ∎

4.3 The Sugeno integral

In this section, we investigate the Sugeno integral under the viewpoint of aggregation. In particular, it will be shown that this integral can be written in the form of a weighted max-min function, which will be introduced and studied below. Although the coefficients involved in these functions are not really weights, but rather thresholds or aspiration degrees (see Section 6.5.3), we will speak in terms of weights.

The formal analogy between the weighted max-min function and the multilinear polynomial is obvious: minimum corresponds to product, maximum does to sum. Moreover, it is emphasized that weighted max-min functions can be calculated as medians, i.e., the qualitative counterparts of multilinear polynomials.

This section aims at offering a better understanding of the nature of the Sugeno integral as an aggregation operator. All the results proved by the author can be found in Marichal [115].

The notations $\theta_S := M(e_S)$ and $\overline{\theta}_S := M(\overline{e}_S)$ will sometimes be used.

4.3.1 Weighted max-min functions

If f_μ is the pseudo-Boolean function which represents a given fuzzy measure μ, then we can write[6]

$$f_\mu(x) = \bigvee_{T \subseteq N} \Big[\mu_T \wedge (\bigwedge_{i \in T} x_i)\Big], \quad x \in \{0,1\}^n.$$

However, such an expression can sometimes be simplified as the following example shows: assuming that $N = \{1,2\}$ and $\mu_1 = 1$, $\mu_2 = 0$, we have

$$f_\mu(x) = x_1 \vee (x_1 \wedge x_2) = x_1. \tag{4.21}$$

Thus, in a more general way, we see that there exist several set functions $c : 2^N \to [0,1]$ fulfilling $c_\emptyset = 0$ and

$$\bigvee_{T \subseteq N} c_T = 1$$

such that

$$f_\mu(x) = \bigvee_{T \subseteq N} \Big[c_T \wedge (\bigwedge_{i \in T} x_i)\Big], \quad x \in \{0,1\}^n.$$

We now investigate a natural extension on $[0,1]^n$ of such pseudo-Boolean functions: the weighted max-min function.

Definition 4.3.1 For any set function $c : 2^N \to [0,1]$ such that $c_\emptyset = 0$ and

$$\bigvee_{T \subseteq N} c_T = 1,$$

[6] This relies heavily on the assumption that μ is monotonic (as opposed to (4.2), which is always valid).

the *weighted max-min function* $\mathrm{W}_c^{\vee\wedge} : [0,1]^n \to [0,1]$ associated to c is defined by

$$\mathrm{W}_c^{\vee\wedge}(x) = \bigvee_{T\subseteq N} \Big[c_T \wedge (\bigwedge_{i\in T} x_i)\Big], \quad x \in [0,1]^n.$$

Observe that we have

$$\mathrm{W}_c^{\vee\wedge}(e_S) = \bigvee_{T\subseteq S} c_T, \quad S \subseteq N. \tag{4.22}$$

As already observed in (4.21), the set function c which defines $\mathrm{W}_c^{\vee\wedge}$ is not uniquely determined. The next proposition precises conditions under which two weighted max-min functions are identical.

Proposition 4.3.1 *Let c and c' be set functions defining $\mathrm{W}_c^{\vee\wedge}$ and $\mathrm{W}_{c'}^{\vee\wedge}$ respectively. Then the following three assertions are equivalent:*

i) $\mathrm{W}_{c'}^{\vee\wedge} = \mathrm{W}_c^{\vee\wedge}$

ii) $\mathrm{W}_{c'}^{\vee\wedge}(e_S) = \mathrm{W}_c^{\vee\wedge}(e_S), \quad S \subseteq N$

iii) for all $T \subseteq N$, $T \neq \emptyset$, we have

$$\begin{cases} c'_T = c_T, & \text{if } c_T > \bigvee_{K\subsetneq T} c_K, \\ 0 \le c'_T \le \bigvee_{K\subseteq T} c_K, & \text{if } c_T \le \bigvee_{K\subsetneq T} c_K. \end{cases}$$

Proof. *i)* $\Rightarrow$ *ii)* Trivial.

ii) $\Rightarrow$ *iii)*. Let $T \subseteq N$, $T \neq \emptyset$. On the one hand, we have, by (4.22),

$$0 \le c'_T \le \bigvee_{K\subseteq T} c'_K = \bigvee_{K\subseteq T} c_K.$$

On the other hand, assuming that $c_T > \bigvee_{K\subsetneq T} c_K$, we obtain

$$c_T = \bigvee_{K\subseteq T} c_K = \bigvee_{K\subseteq T} c'_K$$

implying $c_T = c'_T$; indeed, otherwise there would exist $K^* \subsetneq T$ such that

$$c_T = c'_{K^*} \le \bigvee_{L\subseteq K^*} c'_L = \bigvee_{L\subseteq K^*} c_L \le \bigvee_{K\subsetneq T} c_K < c_T$$

which is a contradiction.

iii) $\Rightarrow$ *i)*. Assume $c_T \le \bigvee_{K\subsetneq T} c_K$ and let $K^* \subsetneq T$ such that $c_{K^*} = \bigvee_{K\subsetneq T} c_K$. Then we have $c_{K^*} \ge c_T$ and

$$c_{K^*} \wedge (\bigwedge_{i\in K^*} x_i) \ge c_T \wedge (\bigwedge_{i\in T} x_i)$$

and so c_T can be replaced by any number lying between 0 and $c_{K^*} = \bigvee_{K\subseteq T} c_K$ without altering $\mathrm{W}_c^{\vee\wedge}$. ∎

Let c be any set function defining $\mathrm{W}_c^{\vee\wedge}$ and let $T \subseteq N$, $T \neq \emptyset$. If $c_T > \bigvee_{K\subsetneq T} c_K$ then c_T cannot be modified without altering $\mathrm{W}_c^{\vee\wedge}$. In the other case, it can be replaced by any value lying between 0 and $\bigvee_{K\subseteq T} c_K$.

If c is such that

$$\forall T \subseteq N, T \neq \emptyset: \; c_T = 0 \;\; \Leftrightarrow \;\; c_T \leq \bigvee_{K \subsetneq T} c_K$$

then all the c_T's are taken as small as possible and we say that $\mathrm{W}_c^{\vee\wedge}$ is put in its *canonical* form. By contrast, if c is such that

$$\forall T \subseteq N: \; c_T = \bigvee_{K \subseteq T} c_K$$

then the c_T's are taken as large as possible and we say that $\mathrm{W}_c^{\vee\wedge}$ is put in its *complete* form. In this case, c is a fuzzy measure since it is increasing. In fact, $\mathrm{W}_c^{\vee\wedge}$ is put in its complete form if and only if c is increasing.

For instance, all the possible expressions of $x_1 \vee (x_1 \wedge x_2)$ as a two-place weighted max-min function are given by

$$x_1 \vee (\lambda \wedge x_1 \wedge x_2), \quad \lambda \in [0,1].$$

The case $\lambda = 0$ corresponds to the canonical form and the case $\lambda = 1$ corresponds to the complete form.

Proposition 4.3.2 *We can determine the complete form of any function* $\mathrm{W}_c^{\vee\wedge}$ *by taking* $c_T = \mathrm{W}_c^{\vee\wedge}(e_T)$ *for all* $T \subseteq N$. *We then get its canonical form by considering successively the* T*'s in decreasing cardinality order and setting* $c_T = 0$ *whenever*

$$c_T \leq \bigvee_{i \in T} c_{T \setminus i}, \quad T \neq \emptyset.$$

Proof. Obtaining the complete form follows from (4.22). Now, fix $T \subseteq N$, $T \neq \emptyset$. If $c_K = \bigvee_{L \subseteq K} c_L$ for all $K \subseteq T$, then

$$\bigvee_{K \subsetneq T} c_K = \bigvee_{K \subsetneq T} \bigvee_{L \subseteq K} c_L = \bigvee_{\substack{K \subsetneq T \\ k=t-1}} \bigvee_{L \subseteq K} c_L = \bigvee_{\substack{K \subsetneq T \\ k=t-1}} c_K,$$

which is sufficient. ■

Proposition 4.3.3 *Let* c *and* c' *be set functions defining* $\mathrm{W}_c^{\vee\wedge}$ *and* $\mathrm{W}_{c'}^{\vee\wedge}$ *respectively. Then we have*

$$\mathrm{W}_{c'}^{\vee\wedge} = \mathrm{W}_c^{\vee\wedge} \quad \Rightarrow \quad \bigvee_{K \subsetneq T} c'_K = \bigvee_{K \subsetneq T} c_K, \quad T \neq \emptyset.$$

Proof. It suffices to show that

$$\bigvee_{K \subsetneq T} c_K = \bigvee_{K \subsetneq T} \mathrm{W}_c^{\vee\wedge}(e_K) \;\; \Big(= \bigvee_{K \subsetneq T} \bigvee_{L \subseteq K} c_L \Big),$$

which is trivial since $\{c_K \,|\, K \subsetneq T\} = \{c_L \,|\, L \subseteq K, K \subsetneq T\}$. ■

4.3.2 Weighted min-max functions

By exchanging the position of the max and min operations in Definition 4.3.1, we can define the weighted min-max functions as follows.

Definition 4.3.2 For any set function $d : 2^N \to [0,1]$ such that $d_\emptyset = 1$ and

$$\bigwedge_{T\subseteq N} d_T = 0,$$

the *weighted min-max function* $\mathrm{W}_d^{\wedge\vee} : [0,1]^n \to [0,1]$ associated to d is defined by

$$\mathrm{W}_d^{\wedge\vee}(x) = \bigwedge_{T\subseteq N} \Big[d_T \vee (\bigvee_{i\in T} x_i)\Big], \quad x \in [0,1]^n.$$

Observe that we have

$$\mathrm{W}_d^{\wedge\vee}(e_S) = \bigwedge_{T\subseteq N\setminus S} d_T, \quad S \subseteq N. \tag{4.23}$$

Moreover, the set function d which defines $\mathrm{W}_d^{\wedge\vee}$ is not uniquely determined; indeed, we have, for instance, $x_1 \wedge (x_1 \vee x_2) = x_1$. We then have a result similar to Proposition 4.3.1.

Proposition 4.3.4 *Let d and d' be set functions defining $\mathrm{W}_d^{\wedge\vee}$ and $\mathrm{W}_{d'}^{\wedge\vee}$ respectively. Then the following three assertions are equivalent:*

i) $\mathrm{W}_{d'}^{\wedge\vee} = \mathrm{W}_d^{\wedge\vee}$

ii) $\mathrm{W}_{d'}^{\wedge\vee}(e_S) = \mathrm{W}_d^{\wedge\vee}(e_S), \quad S \subseteq N$

iii) for all $T \subseteq N$, $T \neq \emptyset$, we have

$$\begin{cases} d'_T = d_T, & \text{if } d_T < \bigwedge_{K\subsetneq T} d_K, \\ \bigwedge_{K\subseteq T} d_K \leq d'_T \leq 1, & \text{if } d_T \geq \bigwedge_{K\subsetneq T} d_K. \end{cases}$$

Let d be any set function defining $\mathrm{W}_d^{\wedge\vee}$ and let $T \subseteq N$, $T \neq \emptyset$. If $d_T < \bigwedge_{K\subsetneq T} d_K$ then d_T cannot be modified without altering $\mathrm{W}_d^{\wedge\vee}$. In the other case, it can be replaced by any value lying between $\bigwedge_{K\subseteq T} d_K$ and 1.

If d is such that

$$\forall\, T \subseteq N, T \neq \emptyset : \; d_T = 1 \;\Leftrightarrow\; d_T \geq \bigwedge_{K\subsetneq T} d_K$$

then all the d_T's are taken as large as possible and we say that $\mathrm{W}_d^{\wedge\vee}$ is put in its *canonical* form. By contrast, if d is such that

$$\forall\, T \subseteq N : \; d_T = \bigwedge_{K\subseteq T} d_K$$

then the d_T's are taken as small as possible and we say that $\mathrm{W}_d^{\wedge\vee}$ is put in its *complete* form. In this case, d is decreasing. In fact, $\mathrm{W}_d^{\wedge\vee}$ is put in its complete form if and only if d is decreasing.

Proposition 4.3.5 *We can determine the complete form of any function $\mathrm{W}_d^{\wedge\vee}$ by taking $d_T = \mathrm{W}_d^{\wedge\vee}(\bar{e}_T)$ for all $T \subseteq N$. We then get its canonical form by considering successively the T's in decreasing cardinality order and setting $d_T = 1$ whenever*

$$d_T \geq \bigwedge_{i\in T} d_{T\setminus i}, \quad T \neq \emptyset.$$

4.3.3 Correspondence formulas and equivalent forms

We now prove that any weighted max-min function can be put in the form of a weighted min-max function and conversely. The next proposition gives the correspondence formulas.

Proposition 4.3.6 *Let c and d be set functions defining $\mathrm{W}_c^{\vee\wedge}$ and $\mathrm{W}_d^{\wedge\vee}$ respectively. Then we have*

$$\mathrm{W}_c^{\vee\wedge} = \mathrm{W}_d^{\wedge\vee} \quad \Leftrightarrow \quad \bigvee_{K\subseteq T} c_K = \bigwedge_{K\subseteq N\setminus T} d_K \quad \forall T \subseteq N.$$

Proof. ($\Rightarrow$) By (4.22) and (4.23), we have, for all $T \subseteq N$,

$$\bigvee_{K\subseteq T} c_K = \mathrm{W}_c^{\vee\wedge}(e_T) = \mathrm{W}_d^{\wedge\vee}(e_T) = \bigwedge_{K\subseteq N\setminus T} d_K.$$

($\Leftarrow$) Let d be any set function defining $\mathrm{W}_d^{\wedge\vee}$. Using classical distributivity, we can find a set function c' defining a $\mathrm{W}_{c'}^{\vee\wedge}$ such that $\mathrm{W}_{c'}^{\vee\wedge} = \mathrm{W}_d^{\wedge\vee}$. We then observe that, for all $T \subseteq N$,

$$\bigvee_{K\subseteq T} c'_K = \mathrm{W}_{c'}^{\vee\wedge}(e_T) = \mathrm{W}_d^{\wedge\vee}(e_T) = \bigwedge_{K\subseteq N\setminus T} d_K = \bigvee_{K\subseteq T} c_K.$$

By Proposition 4.3.1, we simply have $\mathrm{W}_c^{\vee\wedge} = \mathrm{W}_d^{\wedge\vee}$. ∎

When $\mathrm{W}_c^{\vee\wedge}$ and $\mathrm{W}_d^{\wedge\vee}$ are put in their complete forms, the correspondence formulas become simpler.

Corollary 4.3.1 *For any increasing set function c defining $\mathrm{W}_c^{\vee\wedge}$ and any decreasing set function d defining $\mathrm{W}_d^{\wedge\vee}$, we have*

$$\mathrm{W}_c^{\vee\wedge} = \mathrm{W}_d^{\wedge\vee} \quad \Leftrightarrow \quad c_T = d_{N\setminus T} \quad \forall T \subseteq N.$$

The following example illustrates the use of the correspondance formulas.

Example 4.3.1 Let $N = \{1, 2, 3\}$. We have

$$(0.1 \wedge x_1) \vee (0.3 \wedge x_2) \vee (x_2 \wedge x_3) = (0.1 \vee x_2) \wedge (0.3 \vee x_3) \wedge (x_1 \vee x_2), \quad x \in [0,1]^3.$$

Indeed, starting from the left-hand side which is the canonical form of a weighted max-min function, we can compute its complete form, then the complete form of the corresponding weighted min-max function and finally the canonical form:

$$\begin{aligned}
& (0.1 \wedge x_1) \vee (0.3 \wedge x_2) \vee (x_2 \wedge x_3) \\
= \;& 0 \vee (0.1 \wedge x_1) \vee (0.3 \wedge x_2) \vee (0 \wedge x_3) \vee (0.3 \wedge x_1 \wedge x_2) \vee (0.1 \wedge x_1 \wedge x_3) \vee (1 \wedge x_2 \wedge x_3) \\
& \vee (1 \wedge x_1 \wedge x_2 \wedge x_3) \\
= \;& 1 \wedge (1 \vee x_1) \wedge (0.1 \vee x_2) \wedge (0.3 \vee x_3) \wedge (0 \vee x_1 \vee x_2) \wedge (0.3 \vee x_1 \vee x_3) \wedge (0.1 \vee x_2 \vee x_3) \\
& \wedge (0 \vee x_1 \vee x_2 \vee x_3) \\
= \;& (0.1 \vee x_2) \wedge (0.3 \vee x_3) \wedge (x_1 \vee x_2).
\end{aligned}$$

We now show that, on each simplex $\mathcal{B}_\pi$, $\mathrm{W}_c^{\vee\wedge}$ and $\mathrm{W}_d^{\wedge\vee}$ are medians weighted by $n-1$ coefficients. To present this, we need a technical lemma which was established by Dubois and Prade [47].

Lemma 4.3.1 *Let $x, x' \in [0,1]^n$ with $x_1 \leq \ldots \leq x_n$ and $x'_1 \geq \ldots \geq x'_n$.*
(i) If $x'_1 = 1$ then

$$\bigvee_{i=1}^{n} (x_i \wedge x'_i) = \mathrm{median}(x_1, \ldots, x_n, x'_2, \ldots, x'_n).$$

(ii) If $x'_n = 0$ then

$$\bigwedge_{i=1}^{n} (x_i \vee x'_i) = \mathrm{median}(x_1, \ldots, x_n, x'_1, \ldots, x'_{n-1}).$$

Theorem 4.3.1 *i) For any increasing set function c defining $\mathrm{W}_c^{\vee\wedge}$, we have, for all $x \in [0,1]^n$,*

$$\begin{aligned}\mathrm{W}_c^{\vee\wedge}(x) &= \bigvee_{i=1}^{n} [x_{(i)} \wedge c_{\{(i),\ldots,(n)\}}] \\ &= \mathrm{median}(x_1, \ldots, x_n, c_{\{(2),\ldots,(n)\}}, c_{\{(3),\ldots,(n)\}}, \ldots, c_{\{(n)\}}).\end{aligned}$$

ii) For any decreasing set function d defining $\mathrm{W}_d^{\wedge\vee}$, we have, for all $x \in [0,1]^n$,

$$\begin{aligned}\mathrm{W}_d^{\wedge\vee}(x) &= \bigwedge_{i=1}^{n} [x_{(i)} \vee d_{\{(1),\ldots,(i)\}}] \\ &= \mathrm{median}(x_1, \ldots, x_n, d_{\{(1)\}}, d_{\{(1),(2)\}}, \ldots, d_{\{(1),\ldots,(n-1)\}}).\end{aligned}$$

Proof. *i)* Let $x \in [0,1]^n$. Since c is increasing, we have

$$\begin{aligned}\bigvee_{i=1}^{n} [x_{(i)} \wedge c_{\{(i),\ldots,(n)\}}] &= \bigvee_{i=1}^{n} \bigvee_{\substack{T \subseteq \{(i),\ldots,(n)\} \\ T \ni (i)}} [c_T \wedge x_{(i)}] \\ &= \bigvee_{i=1}^{n} \bigvee_{\substack{T \subseteq \{(i),\ldots,(n)\} \\ T \ni (i)}} \Big[c_T \wedge (\bigwedge_{j \in T} x_j)\Big] \\ &= \bigvee_{T \subseteq N} \Big[c_T \wedge (\bigwedge_{j \in T} x_j)\Big]\end{aligned}$$

which proves the first equality. The second one follows from Lemma 4.3.1.

ii) Let c be an increasing set function defined by $c_T = d_{N \setminus T}$ for all $T \subseteq N$. For all $x \in [0,1]^n$, we have

$$\begin{aligned}\mathrm{W}_d^{\wedge\vee}(x) &= \mathrm{W}_c^{\vee\wedge}(x) \quad \text{(by Corollary 4.3.1)} \\ &= \mathrm{median}(x_1, \ldots, x_n, c_{\{(2),\ldots,(n)\}}, c_{\{(3),\ldots,(n)\}}, \ldots, c_{\{(n)\}}) \quad \text{(by } i)) \\ &= \mathrm{median}(x_1, \ldots, x_n, d_{\{(1)\}}, d_{\{(1),(2)\}}, \ldots, d_{\{(1),\ldots,(n-1)\}}) \\ &= \bigwedge_{i=1}^{n} [x_{(i)} \vee d_{\{(1),\ldots,(i)\}}] \quad \text{(by Lemma 4.3.1)},\end{aligned}$$

and the proof is complete. ∎

4.3.4 Alternative expressions of the Sugeno integral

Theorem 4.3.1 shows that the class of the Sugeno integrals coincides with the family of weighted max-min functions which, in turn, coincides with the family of weighted min-max functions. This allows to derive alternative expressions of the Sugeno integral. Note that the expression in terms of median was already established in 1978 by Kandel and Byatt [104].

Theorem 4.3.2 *Let $x \in [0,1]^n$ and μ be a fuzzy measure on N. Then we have*

$$
\begin{aligned}
\mathcal{S}_\mu(x) &= \bigvee_{i=1}^{n} [x_{(i)} \wedge \mu_{\{(i),\ldots,(n)\}}] \\
&= \bigwedge_{i=1}^{n} [x_{(i)} \vee \mu_{\{(i+1),\ldots,(n)\}}] \\
&= \bigvee_{T \subseteq N} \Big[\mu_T \wedge \big(\bigwedge_{i\in T} x_i\big)\Big] \\
&= \bigwedge_{T \subseteq N} \Big[\mu_{N\setminus T} \vee \big(\bigvee_{i\in T} x_i\big)\Big] \\
&= \operatorname{median}(x_1,\ldots,x_n,\mu_{\{(2),\ldots,(n)\}},\mu_{\{(3),\ldots,(n)\}},\ldots,\mu_{\{(n)\}}).
\end{aligned}
$$

Let us consider an example. Assume that $N = \{1,2,3\}$ and $x \in [0,1]^3$ with $x_3 \leq x_1 \leq x_2$. Then

$$
\begin{aligned}
\mathcal{S}_\mu(x_1,x_2,x_3) &= x_3 \vee (x_1 \wedge \mu_{\{1,2\}}) \vee (x_2 \wedge \mu_{\{2\}}) \\
&= (x_3 \vee \mu_{\{1,2\}}) \wedge (x_1 \vee \mu_{\{2\}}) \wedge x_2 \\
&= \operatorname{median}(x_1,x_2,x_3,\mu_{\{1,2\}},\mu_{\{2\}}).
\end{aligned}
$$

We can observe that, as an aggregation operator, the Sugeno integral with respect to a measure μ is an extension on the entire hypercube $[0,1]^n$ of the pseudo-Boolean function f_μ which defines μ. The same conclusion had been obtained for the Choquet integral (see Section 4.2.2). Notice also that, on each simplex $\mathcal{B}_\pi$, the Sugeno integral identifies with a weighted median and the Choquet integral with a weighted arithmetic mean.

4.3.5 Some axiomatic characterizations

The class of Sugeno integrals can be characterized by means of some selected properties, see also [33, 149].

Theorem 4.3.3 *Let $M : [0,1]^n \to \mathbb{R}$. Then the following assertions are equivalent:*
- *i) M fulfils (In, Id, CoMin, CoMax)*
- *ii) M fulfils (In, SMin, SMax)*
- *iii) M fulfils (In, Id, SMinB, SMaxB)*
- *iv) There exists a set function $c : 2^N \to [0,1]$ such that $M = \mathrm{W}_c^{\vee\wedge}$*
- *v) There exists a set function $d : 2^N \to [0,1]$ such that $M = \mathrm{W}_d^{\wedge\vee}$*
- *vi) There exists a fuzzy measure μ on N such that $M = \mathcal{S}_\mu$*

Proof. *i)* $\Rightarrow$ *ii)* Let $x \in [0,1]^n$ and $r \in [0,1]$. Since x and $(r,\ldots,r) \in [0,1]^n$ are comonotonic, we have,

$$M(x_1 \wedge r,\ldots,x_n \wedge r) \overset{\text{(CoMin)}}{=} M(x_1,\ldots,x_n) \wedge M(r,\ldots,r) \overset{\text{(Id)}}{=} M(x_1,\ldots,x_n) \wedge r,$$

and M fulfils (SMin). One can prove similarly that M also fulfils (SMax).

ii) $\Rightarrow$ *iii)* Trivial (see also Proposition 2.2.5).

iii) $\Rightarrow$ *iv)* The proof is very similar to that of Theorem 3.4.10.

Let $x \in [0,1]^n$. On the one hand, for all $T \subseteq N$, we have

$$M(x) \overset{\text{(In)}}{\geq} M[(\bigwedge_{i\in T} x_i)\, e_T] \overset{\text{(by (2.8))}}{=} \theta_T \wedge (\bigwedge_{i\in T} x_i)$$

and thus

$$M(x) \geq \bigvee_{T\subset N} \Big[\theta_T \wedge (\bigwedge_{i\in T} x_i)\Big].$$

On the other hand, let $T^* \subseteq N$ such that $\theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$ is maximum and set

$$J := \{j \in N \,|\, x_j \leq \theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)\}.$$

We should have $J \neq \emptyset$; indeed, if $x_j > \theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$ for all $j \in N$, then we have, since $\theta_N = 1$,

$$\theta_N \wedge (\bigwedge_{i\in N} x_i) > \theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$$

which contradicts the definition of T^*. Moreover, we should have $\overline{\theta}_J \leq \theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)$, for otherwise we would have, assuming $N \setminus J \neq \emptyset$,

$$\overline{\theta}_J \wedge (\bigwedge_{i\in N\setminus J} x_i) > \theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i),$$

a contradiction. Finally, we have,

$$\begin{aligned} M(x) &\overset{\text{(In)}}{\leq} M\Big[\Big(\theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)\Big)\, e_J + \overline{e}_J\Big] \\ &\overset{\text{(by (2.9))}}{=} \Big[\theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i)\Big] \vee \overline{\theta}_J \\ &= \theta_{T^*} \wedge (\bigwedge_{i\in T^*} x_i) \\ &= \bigvee_{T\subseteq N} \Big[\theta_T \wedge (\bigwedge_{i\in T} x_i)\Big]. \end{aligned}$$

iv) $\Rightarrow$ *v)* $\Rightarrow$ *vi)* See Propositions 4.3.6 and 4.3.1.

vi) $\Rightarrow$ *i)* Of course, $\mathcal{S}_\mu$ fulfils (In, Id). Next, consider two comonotonic vectors $x, x' \in [0,1]^n$. Then we have

$$\begin{aligned} \mathcal{S}_\mu(x_1 \wedge x'_1, \ldots, x_n \wedge x'_n) &= \bigvee_{i=1}^{n} [(x_{(i)} \wedge x'_{(i)}) \wedge \mu_{\{(i),\ldots,(n)\}}] \\ &= \Big[\bigvee_{i=1}^{n} [x_{(i)} \wedge \mu_{\{(i),\ldots,(n)\}}]\Big] \wedge \Big[\bigvee_{i=1}^{n} [x'_{(i)} \wedge \mu_{\{(i),\ldots,(n)\}}]\Big] \\ &= \mathcal{S}_\mu(x_1, \ldots, x_n) \wedge \mathcal{S}_\mu(x'_1, \ldots, x'_n) \end{aligned}$$

and $\mathcal{S}_\mu$ fulfils (CoMin). The same argument allows to show that $\mathcal{S}_\mu$ also fulfils (CoMax). See also [148]. ∎

Another characterization of the class of Sugeno integrals is proposed in Section 6.5. In this characterization, it is assumed that the aggregation operator *a priori* depends on a fuzzy measure, see Theorem 6.5.1.

The following characterization, restricted to the case $n = 2$, is a particular case of Theorem 3.3.7 (see also Figure 3.7). It shows that, under (In, Id), the (A) property combined with (Co) produce exactly the same effect as that of (CoMin, CoMax) or (SMinB, SMaxB).

Theorem 4.3.4 *A two-place function $M : [0,1]^2 \to [0,1]$ fulfils (Co, In, Id, A) if and only if there exists a fuzzy measure μ on $\{1,2\}$ such that $M = \mathcal{S}_\mu$.*

This result shows that any two-place Sugeno integral fulfils (A). One can easily verify that it also fulfils (AD) and (B). This is not the case for the Choquet integral (see Theorem 4.2.3).

Theorem 4.3.5 *Any two-place Sugeno integral fulfils (A, AD, B).*

The form of n-place Sugeno integrals fulfilling (B) is not known yet. However, when $n \geq 3$, one can say that not all the Sugeno integrals fulfil (B), as we can easily verify for the following Sugeno integral

$$M(x_1, x_2, x_3) = (x_1 \wedge x_2) \vee (x_2 \wedge x_3)$$

and the matrix

$$X = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

An *extended Sugeno integral* $\mathcal{S}$ is an extended aggregation operator $M \in A(E, \mathbb{R})$ such that, for all $n \in \mathbb{N}_0$, $M^{(n)}$ is a Sugeno integral. Concerning such extended aggregation operators, we present the following result.

Theorem 4.3.6 *Let $M \in A([0,1],[0,1])$. Then the following assertions are equivalent.*

i) *M is an extended Sugeno integral and fulfils (SD)*
ii) *M is an extended Sugeno integral and fulfils (D)*
iii) *M is an extended Sugeno integral and fulfils (A)*
iv) *M fulfils (Co, In, Id, A)*
v) *there exist $\alpha, \beta \in [0,1]$ such that, for all $n \in \mathbb{N}_0$, we have*

$$M(x) = (\alpha \wedge x_1) \vee \Big(\bigvee_{i=2}^{n-1} (\alpha \wedge \beta \wedge x_i)\Big) \vee (\beta \wedge x_n) \vee \Big(\bigwedge_{i=1}^{n} x_i\Big), \quad x \in [0,1]^n.$$

Proof. *iv)* $\Leftrightarrow$ *v)* See Theorem 3.3.8.

iii) $\Rightarrow$ *iv)* Trivial.

iv) & *v)* $\Rightarrow$ *iii)* See Theorems 3.3.8 and 4.3.3.

i) $\Rightarrow$ *ii)* See (2.12).

ii) $\Rightarrow$ *v)* The proof can be compared with that of Theorem 3.5.4. By Theorem 4.3.3, the result holds true for $n = 2$. Suppose it holds for a fixed $n \geq 2$ and show it still holds for $n+1$. Let $(x_1, \ldots, x_{n+1}) \in [0,1]^{n+1}$ and set $x := (x_1, \ldots, x_n)$. Since M fulfils (D) and (SMin, SMax), we have

$$\begin{aligned} M^{(n+1)}(x_1, \ldots, x_{n+1}) &= M^{(n+1)}(n \odot M^{(n)}(x), x_{n+1}) \\ &= \begin{cases} x_{n+1} \vee [M^{(n)}(x) \wedge \overline{\theta}_{n+1}^{(n+1)}] & \text{if } x_{n+1} \leq M^{(n)}(x), \\ M^{(n)}(x) \vee [x_{n+1} \wedge \theta_{n+1}^{(n+1)}] & \text{if } x_{n+1} \geq M^{(n)}(x). \end{cases} \end{aligned}$$

Let us show that $\overline{\theta}_{n+1}^{(n+1)}$ is uniquely determined. The same can be done for $\theta_{n+1}^{(n+1)}$. Using induction, we have

$$\overline{\theta}_{n+1}^{(n+1)} \overset{\text{(D)}}{=} M^{(n+1)}(1, n \odot \overline{\theta}_n^{(n)}) \overset{\text{(ind.)}}{=} M^{(n+1)}(1, n \odot \alpha) = \alpha \vee \theta_1^{(n+1)}$$

and

$$\theta_1^{(n+1)} \overset{\text{(D)}}{=} M^{(n+1)}(n \odot \theta_1^{(n)}, 0) \overset{\text{(ind.)}}{=} M^{(n+1)}(n \odot \alpha, 0) = \alpha \wedge \overline{\theta}_{n+1}^{(n+1)}.$$

Hence $\overline{\theta}_{n+1}^{(n+1)} = \alpha$.

Consequently, $M^{(n+1)}(x_1, \ldots, x_{n+1})$ is uniquely determined.

v) $\Rightarrow$ *i)* By Theorem 4.3.3, we only have to prove that M fulfils (SD). Fix $n \in \mathbb{N}_0$ and $K = \{i_1, \ldots, i_k\} \subseteq N_n$ with $i_1 < \cdots < i_k$. Let $x \in [0,1]^n$. Since M fulfils (A), we have, setting $r := n + k^2 - k$,

$$M\Big(M^{(k)}(x_{i_1}, \ldots, x_{i_k})\, e_K + \sum_{i \notin K} x_i\, e_i\Big) = M(y_1, \ldots, y_r)$$

with $y_1 = x_1$, $y_r = x_n$ and $\{y_1, \ldots, y_r\} = \{x_1, \ldots, x_n\}$. We then have

$$\begin{aligned} M(y_1, \ldots, y_r) &= (\alpha \wedge y_1) \vee \Big(\bigvee_{i=1}^{r} (\alpha \wedge \beta \wedge y_i)\Big) \vee (\beta \wedge y_r) \vee \Big(\bigwedge_{i=1}^{r} y_i\Big) \\ &= (\alpha \wedge x_1) \vee \Big(\bigvee_{i=1}^{n} (\alpha \wedge \beta \wedge x_i)\Big) \vee (\beta \wedge x_n) \vee \Big(\bigwedge_{i=1}^{n} x_i\Big) \\ &= M(x_1, \ldots, x_n). \end{aligned}$$

Hence M fulfils (SD). ■

4.3.6 Weighted maximum and minimum operators

Min and max operators have been extended by Dubois and Prade [47], in a way which is consistent with possibility theory: the weighted minimum (wmin) and maximum (wmax).

Using the concept of possibility and necessity of fuzzy events [45, 200], one can evaluate the possibility that a relevant goal is attained, and the necessity that all the relevant goals are attained by the help of wmin and wmax operators. The formal analogy with the weighted arithmetic mean (WAM) is obvious.

Definition 4.3.3 For any weight vector $\omega = (\omega_1, \ldots, \omega_n) \in [0,1]^n$ such that

$$\bigvee_{i=1}^{n} \omega_i = 1,$$

the *weighted maximum* operator wmax_ω associated to ω is defined by

$$\text{wmax}_\omega(x) = \bigvee_{i=1}^{n} (\omega_i \wedge x_i), \quad x \in [0,1]^n.$$

For any weight vector $\omega = (\omega_1, \ldots, \omega_n) \in [0,1]^n$ such that

$$\bigwedge_{i=1}^{n} \omega_i = 0,$$

the *weighted minimum* operator wmin_ω associated to ω is defined by

$$\text{wmin}_\omega(x) = \bigwedge_{i=1}^{n} (\omega_i \vee x_i), \quad x \in [0,1]^n.$$

Any wmax_ω operator is a $\mathrm{W}_c^{\vee\wedge}$ function whose canonical form is defined by:

$$\begin{cases} c_i = \omega_i, & \forall i \in N, \\ c_T = 0, & \forall T \subseteq N \text{ such that } |T| \neq 1, \end{cases}$$

and complete form by ($c_T = \mu_T$):

$$c_T = \bigvee_{i \in T} \omega_i, \quad \forall T \subseteq N.$$

In this case, if c is increasing then it represents a *possibility measure* π which is characterized by the following property:

$$\pi(S \cup T) = \pi(S) \vee \pi(T), \quad \forall S, T \subseteq N.$$

Likewise, any wmin_ω operator is a $\mathrm{W}_d^{\wedge\vee}$ function whose canonical form is defined by:

$$\begin{cases} d_i = \omega_i, & \forall i \in N, \\ d_T = 1, & \forall T \subseteq N \text{ such that } |T| \neq 1, \end{cases}$$

and complete form by ($d_T = \mu_{N\setminus T}$):

$$d_T = \bigwedge_{i \in T} \omega_i, \quad \forall T \subseteq N.$$

In this case, if d is decreasing then the set function c', defined by $c'_T = d_{N\setminus T}$ for all $T \subseteq N$, represents a *necessity measure* $\mathcal{N}$ which is characterized by the following property:

$$\mathcal{N}(S \cap T) = \mathcal{N}(S) \wedge \mathcal{N}(T), \quad \forall S, T \subseteq N.$$

The operators wmax_ω and wmin_ω have been characterized by Fodor and Roubens [71]. We present hereafter a slightly more general statement.

Theorem 4.3.7 *i)* $M \in A_n([0,1], \mathbb{R})$ *fulfils (WId, SMinB, Max) if and only if there exists* $\omega \in [0,1]^n$ *such that* $M = \mathrm{wmax}_\omega$.

ii) $M \in A_n([0,1], \mathbb{R})$ *fulfils (WId, SMaxB, Min) if and only if there exists* $\omega \in [0,1]^n$ *such that* $M = \mathrm{wmin}_\omega$.

Proof. *i)* (Sufficiency) Trivial.

(Necessity). For all $i \in N$, we have $\theta_i \in [0,1]$; indeed, by (SMinB), we have $\theta_i = \theta_i \wedge 1 \leq 1$, and by (Max), $\theta_i = \theta_i \vee \theta_\emptyset = \theta_i \vee 0 \geq 0$. On the other hand, for all $x \in [0,1]^n$, we have, setting $\omega_i = \theta_i \in [0,1]$,

$$M(x) \overset{(Max)}{=} \bigvee_{i=1}^n M(x_i\, e_i) \overset{(SMinB)}{=} \bigvee_{i=1}^n (\omega_i \wedge x_i).$$

Moreover, $\bigvee_{i=1}^n \omega_i = M(1, \ldots, 1) = 1$ as required.

ii) Similar to *i)*. ∎

We know that the weighted minimum and maximum operators are particular Sugeno integrals. More precisely, we have proved the following, see also [47, 149] .

Theorem 4.3.8 *Let* μ *be a fuzzy measure on* N. *Then the following assertions are equivalent.*

i) μ *is a possibility measure*

ii) *there exists* $\omega \in [0,1]^n$ *such that* $\mathcal{S}_\mu = \mathrm{wmax}_\omega$

iii) $\mathcal{S}_\mu$ *fulfils (Max).*

The following assertions are equivalent.

iv) μ *is a necessity measure*

v) *there exists* $\omega \in [0,1]^n$ *such that* $\mathcal{S}_\mu = \mathrm{wmin}_\omega$

vi) $\mathcal{S}_\mu$ *fulfils (Min).*

4.3.7 Ordered weighted maximum and minimum operators

Dubois *et al.* [53] used the ordered weighted maximum (owmax) and minimum (owmin) for modelling soft partial matching. The basic idea of owmax (and owmin) is the same as in the OWA operator introduced by Yager [192]. That is, in both papers weights are associated with a particular rank rather than a particular element. The main difference between OWA and owmax (and owmin) is in the underlying non-ordered aggregation operation. OWA uses weighted arithmetic mean while owmax and owmin apply weighted maximum and minimum. At first glance, this does not seem to be an essential difference. However, Dubois and Prade [47] proved that owmax and owmin are equivalent to the median of the ordered values and some appropriately chosen additional numbers used instead of the original weights.

Definition 4.3.4 For any weight vector $\omega = (\omega_1, \dots, \omega_n) \in [0,1]^n$ such that

$$1 = \omega_1 \geq \ldots \geq \omega_n,$$

the *ordered weighted maximum* operator owmax_ω associated to ω is defined by

$$\mathrm{owmax}_\omega(x) = \bigvee_{i=1}^{n} (\omega_i \wedge x_{(i)}), \quad x \in [0,1]^n.$$

For any weight vector $\omega' = (\omega'_1, \dots, \omega'_n) \in [0,1]^n$ such that

$$\omega'_1 \geq \ldots \geq \omega'_n = 0,$$

the *ordered weighted minimum* operator $\mathrm{owmin}_{\omega'}$ associated to ω' is defined by

$$\mathrm{owmin}_{\omega'}(x) = \bigwedge_{i=1}^{n} (\omega'_i \vee x_{(i)}), \quad x \in [0,1]^n.$$

In Definition 4.3.4, the inequalities $\omega_1 \geq \ldots \geq \omega_n$ and $\omega'_1 \geq \ldots \geq \omega'_n$ are not restrictive. Indeed, if there exists $i \in \{1, \dots, n-1\}$ such that $\omega_i \leq \omega_{i+1}$ and $\omega'_i \leq \omega'_{i+1}$ then we have

$$\begin{aligned} (\omega_i \wedge x_{(i)}) \vee (\omega_{i+1} \wedge x_{(i+1)}) &= \omega_{i+1} \wedge x_{(i+1)}, \\ (\omega'_i \vee x_{(i)}) \wedge (\omega'_{i+1} \vee x_{(i+1)}) &= \omega'_i \vee x_{(i)}. \end{aligned}$$

This means that ω_i can be replaced by ω_{i+1} in owmax_ω and ω'_{i+1} by ω'_i in $\mathrm{owmin}_{\omega'}$.

Any owmax_ω operator is a $\mathrm{W}_c^{\vee\wedge}$ function whose canonical form is defined by:

$$\forall T \subseteq N, T \neq \emptyset : c_T = \begin{cases} 0, & \text{if } \omega_{n-t+1} = \omega_{n-t+2}, \\ \omega_{n-t+1}, & \text{else,} \end{cases}$$

and complete form by ($c_T = \mu_T$):

$$\forall T \subseteq N, T \neq \emptyset : c_T = \omega_{n-t+1}.$$

Likewise, any $\mathrm{owmin}_{\omega'}$ operator is a $\mathrm{W}_d^{\wedge\vee}$ function whose canonical form is defined by:

$$\forall T \subseteq N, T \neq \emptyset : d_T = \begin{cases} 1, & \text{if } \omega'_t = \omega'_{t-1}, \\ \omega'_t, & \text{else,} \end{cases}$$

and complete form by ($d_T = \mu_{N \setminus T}$):

$$\forall T \subseteq N, T \neq \emptyset : d_T = \omega'_t.$$

The next proposition shows that any ordered weighted maximum operator can be put in the form of an ordered weighted minimum operator and conversely.

Proposition 4.3.7 *Let ω and ω' be weight vectors defining* owmax_ω *and* $\mathrm{owmin}_{\omega'}$ *respectively. Then we have*

$$\mathrm{owmin}_{\omega'} = \mathrm{owmax}_\omega \quad \Leftrightarrow \quad \omega'_i = \omega_{i+1} \quad \forall i \in \{1, \ldots, n-1\}.$$

Proof. If the fuzzy measure μ defines the complete form of owmax_ω then we have

$$\mu_{\{(i+1),\ldots,(n)\}} = \omega_{i+1}, \quad i \in \{1, \ldots, n-1\}.$$

Theorem 4.3.2 then allows to conclude. ∎

It is interesting to note that, according to Lemma 4.3.1, we have, for all $x \in [0,1]^n$,

$$\begin{aligned} \mathrm{owmax}_\omega(x) &= \mathrm{median}(x_1, \ldots, x_n, \omega_2, \ldots, \omega_n), \\ \mathrm{owmin}_{\omega'}(x) &= \mathrm{median}(x_1, \ldots, x_n, \omega'_1, \ldots, \omega'_{n-1}). \end{aligned}$$

We now show that the owmax_ω and $\mathrm{owmin}_{\omega'}$ operators are exactly those weighted max-min functions (or Sugeno integrals) which fulfil (Sy). To do this, we need a lemma which is due to Grabisch [78].

Lemma 4.3.2 *The Sugeno integral $\mathcal{S}_\mu$ fulfils (Sy) if and only if*

$$\mu_S = \mu_T \text{ whenever } |S| = |T|.$$

We then have the following characterization.

Theorem 4.3.9 *Let μ be a fuzzy measure on N. Then the following assertions are equivalent.*
i) *μ depends only on the cardinality of subsets*
ii) *there exists $\omega \in [0,1]^n$ such that $\mathcal{S}_\mu = \mathrm{owmax}_\omega$*
iii) *there exists $\omega' \in [0,1]^n$ such that $\mathcal{S}_\mu = \mathrm{owmin}_{\omega'}$*
iv) *$\mathcal{S}_\mu$ fulfils (Sy).*

Proof. *i)* $\Rightarrow$ *ii)* Let $x \in [0,1]^n$. Setting $\omega_i := \mu_{\{(i),\ldots,(n)\}}$ for all $i \in N$, we have

$$\mathcal{S}_\mu(x) = \bigvee_{i=1}^n (x_{(i)} \wedge \omega_i).$$

ii) $\Rightarrow$ *iii)* See Proposition 4.3.7.
iii) $\Rightarrow$ *iv)* Trivial.
iv) $\Rightarrow$ *i)* See Lemma 4.3.2. ∎

Note that other characterizations of these families have been obtained in [71] by means of ordered versions of (SMin), (SMax), (Min) and (Max), which seem to be unappealing properties.

4.3.8 Associative medians

Definition 4.3.5 For any $\alpha \in [0,1]$, the n-place *associative median* operator $\mathrm{amed}_\alpha^{(n)}$ associated to α is defined by

$$\mathrm{amed}_\alpha^{(n)}(x) = \mathrm{median}\Big(\bigwedge_{i=1}^n x_i, \bigvee_{i=1}^n x_i, \alpha\Big), \quad x \in [0,1]^n.$$

Observe that, for all $\alpha \in [0,1]$ and all $x \in [0,1]^n$, we have

$$\text{median}\Big(\bigwedge_{i=1}^{n} x_i, \bigvee_{i=1}^{n} x_i, \alpha\Big) = \text{median}(x_1, \ldots, x_n, \underbrace{\alpha, \ldots, \alpha}_{n-1}).$$

Thus, any amed_α operator with n variables is a $\mathrm{W}_c^{\vee\wedge}$ function (or a Sugeno integral) whose canonical form is defined by:

$$\begin{cases} c_i = \alpha, & \text{for all } i \in N, \\ c_T = 0, & \text{for all } T \subseteq N \text{ such that } 1 < t < n, \\ c_N = \begin{cases} 1, & \text{if } \alpha < 1, \\ 0, & \text{if } \alpha = 1, \end{cases} \end{cases}$$

and complete form by ($c_T = \mu_T$):

$$\begin{cases} c_T = \alpha, & \text{for all } T \subseteq N \text{ such that } 1 \leq t < n, \\ c_N = 1. \end{cases}$$

Moreover, from Theorem 4.3.2, we immediately have, for all $x \in [0,1]^n$,

$$\text{amed}_\alpha^{(n)}(x) = x_{(1)} \vee \Big[\bigvee_{i=2}^{n} (x_{(i)} \wedge \alpha)\Big] = x_{(1)} \vee (x_{(n)} \wedge \alpha) = x_{(n)} \wedge (x_{(1)} \vee \alpha).$$

It follows that any Sugeno integral $\mathcal{S}_\mu$ is an associative median if and only if the fuzzy measure μ is constant on $2^N \setminus \{\emptyset, N\}$.

Coming back to associative operators defined on $[0,1]^2$, we have the following characterization. It follows from Theorems 3.3.9 and 4.3.9.

Theorem 4.3.10 *Let $M \in A_2([0,1],[0,1])$. Then the following assertions are equivalent:*
i) M fulfils (Sy, Co, In, Id, A)
ii) there exists $\alpha \in [0,1]$ such that $M = \text{amed}_\alpha^{(2)}$
iii) there exists a fuzzy measure μ on $\{1,2\}$ with $\mu_{\{1\}} = \mu_{\{2\}}$ such that $M = \mathcal{S}_\mu$.

It should be mentioned that the equivalence *i)* $\Leftrightarrow$ *ii)* was already established by Fung and Fu [75] in 1975 and in a revisited way by Dubois and Prade [44] in 1984.

In this context, the second part of Theorem 3.3.9 takes the following form.

Corollary 4.3.2 *$M \in A([0,1],[0,1])$ fulfils (Sy, Co, In, Id, A) if and only if there exists $\alpha \in [0,1]$ such that $M = (\text{amed}_\alpha^{(n)})_{n \in \mathbb{N}_0}$.*

4.4 Common area between the two classes of integrals

We now investigate the intersection between the class of Choquet integrals and that of Sugeno integrals. We prove that this intersection corresponds to the family of Boolean max-min functions. Some subfamilies are also studied.

4.4.1 Boolean max-min and min-max functions

Recall the definition of a Boolean max-min function (see Definition 3.4.1) and introduce its dual form: the Boolean min-max function.

Definition 4.4.1 i) For any set function $c : 2^N \to \{0,1\}$ such that $c_\emptyset = 0$ and

$$\bigvee_{T\subseteq N} c_T = 1,$$

the *Boolean max-min function* $\mathrm{B}_c^{\vee\wedge}$ associated to c is defined by

$$\mathrm{B}_c^{\vee\wedge}(x) = \mathrm{W}_c^{\vee\wedge}(x) = \bigvee_{T\subseteq N} \Big[c_T \wedge (\bigwedge_{i\in T} x_i)\Big].$$

ii) For any set function $d : 2^N \to \{0,1\}$ such that $d_\emptyset = 1$ and

$$\bigwedge_{T\subseteq N} d_T = 0,$$

the *Boolean min-max function* $\mathrm{B}_d^{\wedge\vee}$ associated to d is defined by

$$\mathrm{B}_d^{\wedge\vee}(x) = \mathrm{W}_d^{\wedge\vee}(x) = \bigwedge_{T\subseteq N} \Big[d_T \vee (\bigvee_{i\in T} x_i)\Big].$$

Thus defined, a Boolean max-min function (resp. Boolean min-max function) is nothing else than a weighted max-min function (resp. weighted min-max function) whose canonical and complete forms are defined by set functions taking their values in $\{0,1\}$. Moreover, we can write, for any $x \in [0,1]^n$,

$$\begin{aligned} \mathrm{B}_c^{\vee\wedge}(x) &= \bigvee_{\substack{T\subseteq N\\ c_T=1}} \bigwedge_{i\in T} x_i \in \{x_1,\ldots,x_n\}, && \text{(disjunctive normal form)}\\ \mathrm{B}_d^{\wedge\vee}(x) &= \bigwedge_{\substack{T\subseteq N\\ d_T=0}} \bigvee_{i\in T} x_i \in \{x_1,\ldots,x_n\}, && \text{(conjunctive normal form).} \end{aligned}$$

In terms of fuzzy measures, if the set function c is increasing, it represents a 0-1 fuzzy measure. More precisely, Murofushi and Sugeno [135, Sect. 2] showed the following result.

Proposition 4.4.1 *If μ is a 0-1 fuzzy measure on N then the Choquet and the Sugeno integral take the following form:*

$$\mathcal{C}_\mu = \mathcal{S}_\mu = \mathrm{B}_\mu^{\vee\wedge}.$$

We now show a stronger result: the common part between the class of Choquet integrals and that of Sugeno integrals coincides with the class of Boolean max-min functions.

Theorem 4.4.1 *Let $M \in A_n([0,1],\mathbb{R})$. Then the following assertions are equivalent.*

i) There exists a 0-1 fuzzy measure μ on N such that $M = \mathcal{S}_\mu$.
ii) There exist fuzzy measures μ and ν on N such that $M = \mathcal{C}_\mu = \mathcal{S}_\nu$.
iii) M fulfils (UIn) and there exists a fuzzy measure μ on N such that $M = \mathcal{S}_\mu$.
iv) There exists a set function $c : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_c^{\vee\wedge}$.
v) There exists a set function $d : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_d^{\wedge\vee}$.

Proof. *i)* $\Rightarrow$ *ii)* See Proposition 4.4.1.

ii) $\Rightarrow$ *iii)* Evident, since $\mathcal{C}_\mu$ fulfils (UIn).

iii) $\Rightarrow$ *iv)* By Proposition 4.4.1, it suffices to show that μ is a 0-1 fuzzy measure. Suppose that there exists $T \subseteq N$ such that $\mu_T \in]0,1[$. We can write $N = \{t_1, \dots, t_n\}$ and $T = \{t_k, \dots, t_n\}$, with $k \in \{2, \dots, n\}$. Let $x \in [0,1]^n$ such that

$$x_{t_1} \leq \cdots \leq x_{t_{k-1}} < \mu_T < x_{t_k} \leq \cdots \leq x_{t_n}.$$

By Theorem 4.3.2, we always have

$$\mathcal{S}_\mu(x) = \text{median}(x_1, \dots, x_n, \mu_{\{t_2,\dots,t_n\}}, \dots, \mu_T, \dots, \mu_{\{t_n\}}) = \mu_T$$

and $\mathcal{S}_\mu$ does not fulfils (UIn).

iv) $\Leftrightarrow$ *v)* See Corollary 4.3.1.

iv) $\Rightarrow$ *i)* Evident. ∎

Since any $\mathrm{B}_c^{\vee\wedge}$ function is a Choquet integral (see Proposition 4.4.1), it fulfils (SPL) and, by Proposition 3.5.1, it can be defined on any E^n, where $E \supseteq [0,1]$. By using Lemma 4.2.3 and Theorem 3.5.2, we deduce the following result.

Theorem 4.4.2 *Let $M \in A_n(E, \mathbb{R})$, with $E \supseteq [0,1]$. Then the following assertions are equivalent.*

i) There exists a 0-1 fuzzy measure on N such that $M = \mathcal{C}_\mu$.

ii) M fulfils (In, SPL) and $M(e_T) \in \{0,1\}$ for all $T \subseteq N$.

iii) There exists a set function $c : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_c^{\vee\wedge}$.

iv) There exists a set function $d : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_d^{\wedge\vee}$.

Moreover, if $M = \mathrm{B}_c^{\vee\wedge}$ then M fulfils (B) if and only if there exists a non-empty subset $S \subseteq N$ such that $M = \min_S$ or $\max_S$.

We already noticed that the class of Sugeno integrals fulfilling (B) is not described yet. But the second part of the previous theorem describes the subclass of Boolean max-min functions fulfilling this property.

According to Theorems 3.4.10, 3.4.16 and 3.4.20, we also have the following result.

Theorem 4.4.3 *Let $M \in A_n([a,b],[a,b])$. Then the following assertions are equivalent.*

i) M fulfils (In, OS').

ii) M fulfils (In, Id, CM').

iii) M fulfils (Co, Id, OS).

iv) There exists a set function $c : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_c^{\vee\wedge}$.

v) There exists a set function $d : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_d^{\wedge\vee}$.

Recall also the statement of Theorem 3.4.12

Theorem 4.4.4 *Let $M \in A_n(E, \mathbb{R})$, where E is any real interval, finite or infinite. Then the following assertions are equivalent.*

i) M fulfils (Co, Id, CM).

ii) M fulfils (Co, Id, CM').

iii) There exists a set function $c : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_c^{\vee\wedge}$.

iv) There exists a set function $d : 2^N \to \{0,1\}$ such that $M = \mathrm{B}_d^{\wedge\vee}$.

4.4.2 Partial maximum and minimum operators

The partial minimum and maximum operators ($\min_S$ and $\max_S$) are defined by (1.6) and (1.7).

Any $\min_S$ operator is a $\mathrm{B}_c^{\vee\wedge}$ whose canonical form is defined by:

$$\forall T \subseteq N, T \neq \emptyset : c_T = \begin{cases} 1, & \text{if } T = S, \\ 0, & \text{otherwise,} \end{cases}$$

and complete form by ($c_T = \mu_T$):

$$\forall T \subseteq N, T \neq \emptyset : c_T = \begin{cases} 1, & \text{if } T \supseteq S, \\ 0, & \text{otherwise.} \end{cases}$$

As a particular Choquet integral (4.11), the associated Möbius representation identifies with the above canonical form:

$$\forall T \subseteq N, T \neq \emptyset : a_T = \begin{cases} 1, & \text{if } T = S, \\ 0, & \text{otherwise.} \end{cases}$$

Likewise, any $\max_S$ operator is a $\mathrm{B}_d^{\wedge\vee}$ whose canonical form is defined by:

$$\forall T \subseteq N, T \neq \emptyset : d_T = \begin{cases} 0, & \text{if } T = S, \\ 1, & \text{otherwise,} \end{cases}$$

and complete form by ($d_T = \mu_{N \setminus T}$):

$$\forall T \subseteq N, T \neq \emptyset : d_T = \begin{cases} 0, & \text{if } T \supseteq S, \\ 1, & \text{otherwise.} \end{cases}$$

One can easily show that the associated Möbius representation is given by (see Lemma 4.2.1):

$$\forall T \subseteq N, T \neq \emptyset : a_T = \begin{cases} (-1)^{t+1}, & \text{if } T \subseteq S, \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, for all non-empty subset $S \subseteq N$, we have

$$\min_S = \mathrm{wmin}_{\overline{e}_S} \text{ and } \max_S = \mathrm{wmax}_{e_S}.$$

It follows that any Sugeno integral $\mathcal{S}_\mu$ is a partial maximum operator (resp. partial minimum operator) if and only if μ is a 0-1 possibility measure (resp. 0-1 necessity measure).

It should be noted that the unanimity game v_S for $S \subseteq N$, as a particular fuzzy measure (see Section 4.1.1), defines the partial minimum $\min_S$.

The following characterization can be easily deduced from Theorems 4.3.7 and 4.4.1.

Theorem 4.4.5 *i)* $M \in A_n([0,1], \mathbb{R})$ *fulfils (UIn, WId, SMinB, Max) if and only if there exists a non-empty subset* $S \subseteq N$ *such that* $M = \max_S$.

ii) $M \in A_n([0,1], \mathbb{R})$ *fulfils (UIn, WId, SMaxB, Min) if and only if there exists a non-empty subset* $S \subseteq N$ *such that* $M = \min_S$.

4.4.3 Projections, order statistics and medians

The projections P_k and the order statistics OS_k are respectively defined by (1.4) and (1.5).

Any projection P_k is a $\mathrm{B}_c^{\vee\wedge}$ whose canonical form is given by $c_T = 1$ if $T = \{k\}$, and 0 otherwise. This set function c also represents the associated Möbius representation. The complete form is then given by $c_T = 1$ if $T \ni k$, and 0 otherwise.

By using Theorem 3.4.17, we can deduce the following result.

Theorem 4.4.6 *$M \in A_n(\mathbb{R}, \mathbb{R})$ fulfils (Co, Id, CMIS) if and only if there exists $k \in N$ such that $M = \mathrm{P}_k$.*

Any order statistic OS_k is a $\mathrm{B}_c^{\vee\wedge}$ whose canonical form is defined by

$$\forall\, T \subseteq N, T \neq \emptyset : c_T = \begin{cases} 1, & \text{if } t = n-k+1, \\ 0, & \text{otherwise,} \end{cases}$$

and complete form by ($c_T = \mu_T$):

$$\forall\, T \subseteq N, T \neq \emptyset : c_T = \begin{cases} 1, & \text{if } t \geq n-k+1, \\ 0, & \text{otherwise.} \end{cases}$$

Of course, it is also a $\mathrm{B}_d^{\wedge\vee}$ function whose canonical form is defined by:

$$\forall\, T \subseteq N, T \neq \emptyset : d_T = \begin{cases} 0, & \text{if } t = k, \\ 1, & \text{otherwise,} \end{cases}$$

and complete form by ($d_T = \mu_{N \setminus T}$):

$$\forall\, T \subseteq N, T \neq \emptyset : d_T = \begin{cases} 0, & \text{if } t \geq k, \\ 1, & \text{otherwise.} \end{cases}$$

Thus, we retrieve formulas (3.51):

$$x_{(k)} = \bigvee_{\substack{T \subseteq N \\ t=n-k+1}} \bigwedge_{i \in T} x_i = \bigwedge_{\substack{T \subseteq N \\ t=k}} \bigvee_{i \in T} x_i, \quad k \in N.$$

By Theorem 4.3.2, we also have, if $x \in [0,1]^n$,

$$x_{(k)} = \mathrm{median}(x_1, \ldots, x_n, \underbrace{1, \ldots, 1}_{k-1}, \underbrace{0, \ldots, 0}_{n-k}), \quad k \in N.$$

Note also that any order statistic OS_k is an OWA_ω associated to $\omega = e_k$. In particular, by (4.15), the associated Möbius representation is given by

$$\forall T \subseteq N : a_T = \begin{cases} (-1)^{t-n+k-1}\binom{t-1}{n-k}, & \text{if } t \geq n-k+1, \\ 0, & \text{otherwise.} \end{cases}$$

According to Theorem 3.4.14, we know that the order statistics on any E^n form the class of functions satisfying (Sy, Co, Id, CM) and are also the Boolean max-min functions fulfilling (Sy). As a consequence, we have the following result.

Theorem 4.4.7 *Let $M \in A_n([0,1],\mathbb{R})$. Then the following assertions are equivalent.*

i) *M fulfils (Sy) and there exists a 0-1 fuzzy measure μ such that $M = \mathcal{C}_\mu$.*

ii) *M fulfils (Sy) and there exists a 0-1 fuzzy measure μ such that $M = \mathcal{S}_\mu$.*

iii) *M fulfils (UIn) and there exists $\omega \in [0,1]^n$ such that $M = \mathrm{owmax}_\omega$.*

iv) *M fulfils (Sy) and there exists a set function c such that $M = \mathrm{B}_c^{\vee\wedge}$.*

v) *There exists $k \in N$ such that $M = \mathrm{OS}_k$.*

A particular case of order statistic is the so-called median of an odd number of scores. When the scores are ordered, the median corresponds to the middle value: if $x_1, \ldots, x_{2k-1} \in E$, we have

$$\begin{aligned}\mathrm{median}(x_1,\ldots,x_{2k-1}) = x_{(k)} &= \bigvee_{1\le i_1<\cdots<i_k\le 2k-1} (x_{i_1}\wedge\cdots\wedge x_{i_k}) \\ &= \bigwedge_{1\le i_1<\cdots<i_k\le 2k-1} (x_{i_1}\vee\cdots\vee x_{i_k}).\end{aligned}$$

For instance, we have

$$\begin{aligned}\mathrm{median}(x_1,x_2,x_3) &= (x_1\wedge x_2)\vee(x_1\wedge x_3)\vee(x_2\wedge x_3) \\ &= (x_1\vee x_2)\wedge(x_1\vee x_3)\wedge(x_2\vee x_3).\end{aligned}$$

Regarding medians, we have the following immediate characterization.

Theorem 4.4.8 *Let $k \in \mathbb{N}_0$ and $M \in A_{2k-1}(E,\mathbb{R})$.*

i) If $E \subseteq \mathbb{R}^+$ contains x and $1/x$ simultaneously, then M is an order statistic fulfilling (Rec) if and only if $M = \mathrm{median}$.

ii) If $E = [0,1]$ then M is an order statistic fulfilling (SSN) if and only if $M = \mathrm{median}$.

4.5 Set relations between some subclasses of integrals

Before closing the chapter, we summarize some set relations between subfamilies of Choquet and Sugeno integrals on $[0,1]^n$, see Figure 4.1.

The following relations can be deduced from the results presented in this chapter.

$\{\mathcal{C}_\mu \mid \mu : 2^N \to [0,1]\} \cap \{\mathcal{S}_\mu \mid \mu : 2^N \to [0,1]\} = \{\mathrm{B}_c^{\vee\wedge} \mid c : 2^N \to \{0,1\}\}$

$\{\mathrm{WAM}_\omega \mid \omega \in [0,1]\} \subset \{\mathcal{C}_\mu \mid \mu : 2^N \to [0,1]\}$

$\{\mathrm{wmax}_\omega \mid \omega \in [0,1]\} \cup \{\mathrm{wmin}_\omega \mid \omega \in [0,1]\} \subset \{\mathcal{S}_\mu \mid \mu : 2^N \to [0,1]\}$

$\{\mathrm{wmax}_\omega \mid \omega \in [0,1]\} \cap \{\mathrm{B}_c^{\vee\wedge} \mid c : 2^N \to \{0,1\}\} = \{\max_S \mid S \subseteq N\}$

$\{\mathrm{wmin}_\omega \mid \omega \in [0,1]\} \cap \{\mathrm{B}_c^{\vee\wedge} \mid c : 2^N \to \{0,1\}\} = \{\min_S \mid S \subseteq N\}$

$\{\mathrm{WAM}_\omega \mid \omega \in [0,1]\} \cap \{\mathrm{B}_c^{\vee\wedge} \mid c : 2^N \to \{0,1\}\} = \{\mathrm{P}_i \mid i \in N\}$

$\{\max_S \mid S \subseteq N\} \cap \{\min_S \mid S \subseteq N\} = \{\mathrm{P}_i \mid i \in N\}$

$\{\mathcal{C}_\mu \mid \mu : 2^N \to [0,1]\} \cap \{M \mid M \text{ fulfils (Sy)}\} = \{\mathrm{OWA}_\omega \mid \omega \in [0,1]\}$

$\{\mathcal{S}_\mu \mid \mu : 2^N \to [0,1]\} \cap \{M \mid M \text{ fulfils (Sy)}\} = \{\mathrm{owmax}_\omega \mid \omega \in [0,1]\} = \{\mathrm{owmin}_{\omega'} \mid \omega' \in [0,1]\}$

$\{\mathrm{B}_c^{\vee\wedge} \mid c : 2^N \to \{0,1\}\} \cap \{M \mid M \text{ fulfils (Sy)}\} = \{\mathrm{OS}_i \mid i \in N\}$

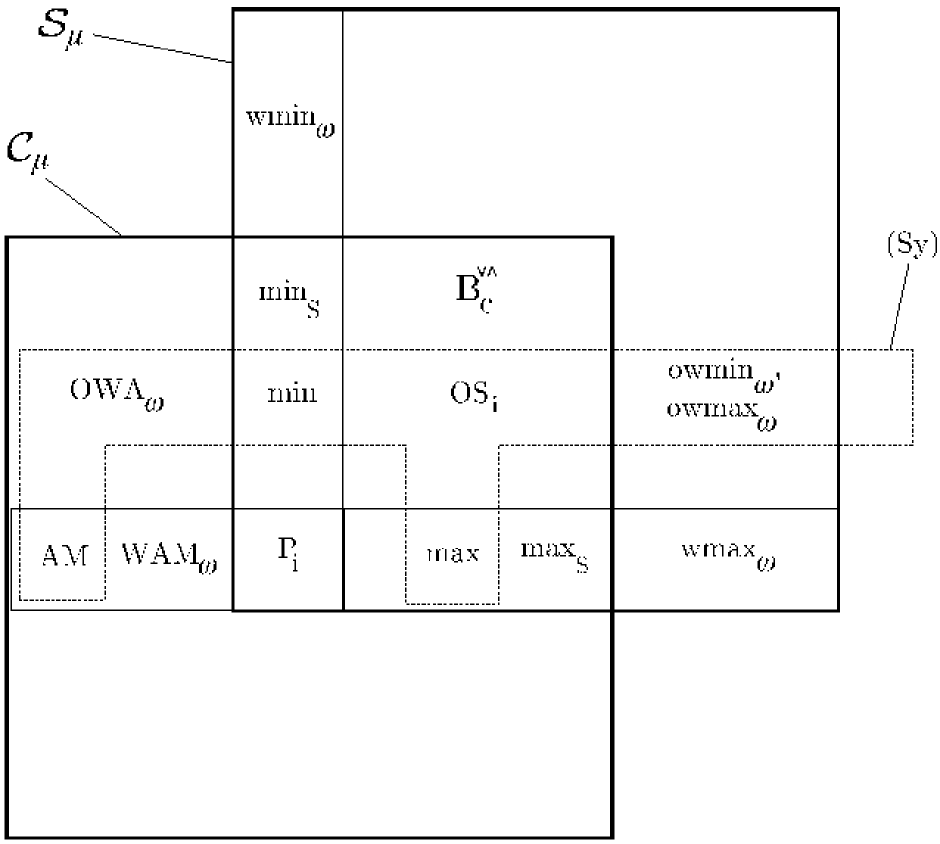


Figure 4.1: Set relations between some subclasses of integrals

Notes

1. The equivalence between the discrete Choquet integral and the Lovász extension is presented in Section 4.2.1. It seems that this connection was previously unknown. It allows us to have a geometrical interpretation of the graph of this function. Moreover, it leads to a practical representation by means of the Möbius transform (see Proposition 4.2.1).

2. Some axiomatic characterizations of the class of Choquet integrals have been presented. Unfortunately, those involving (CoAdd) and (BOM) are of little interest since these axioms are not very appealing in multicriteria decision making. In Section 6.1.3, we present another characterization, which is much more natural.

3. The class of the Sugeno integrals has been characterized in Theorem 4.3.3 by means of rather technical conditions. We present in Section 6.5.3 a more interesting characterization, in which it is assumed that the fuzzy measure is ordinal in nature.

Chapter 5

Power indices and interactions between criteria

This chapter is devoted to the investigation of interaction phenomena between criteria in MCDM problems. As we make use of several concepts borrowed from cooperative game theory, we use the related terminology. This is not a restriction since any fuzzy measure is a particular game.

Let v^N be a cooperative game on the finite set of players N, that is, a set function $v^N : 2^N \to \mathbb{R}$ such that $v^N(\emptyset) = 0$. For any coalition $S \subseteq N$, the real number $v^N(S)$ represents the *worth* of S. The superscript N will be omitted if there is no ambiguity. The set of all games defined on N is denoted $\mathcal{G}^N$. The pseudo-Boolean function f_v which defines v will simply be denoted by f. We also define

$$\mathcal{G} := \bigcup_{n \in \mathbb{N}_0} \mathcal{G}^{N_n}.$$

In multicriteria decision making, when N represents a set of criteria, v is a fuzzy measure acting as a *weight function* and the value $v(S)$ represents the *weight* assigned to the combination S of criteria.

The outline of this chapter is as follows. The concept of power indices is presented in Section 5.1. Axiomatic characterizations of Shapley and Banzhaf power indices are also presented. In Section 5.2 the notion of interaction indices is presented in a formal way. The Shapley and Banzhaf interaction indices appear as extensions of the corresponding power indices. Some characterizations are also mentioned. In Section 5.3 we show that these interaction indices are equivalent representations of the set function v. This is done through the use of the so-called multilinear extensions, but also by means of fractal and cardinality matrices. Finally, in Section 5.4 a new definition of interaction is proposed and studied. It is built from the concept of maximal chains defined in the lattice related to the power set of N.

5.1 Power indices

5.1.1 Shapley and Banzhaf values

A *simple game* on N is a monotonic game v such that $v(S) \in \{0,1\}$ for all $S \subseteq N$. Coalitions with $v(S) = 1$ are called *winning*, the rest are *losing*. Simple games model the allocation of power in committees: a coalition is winning if it controls the decisions.

For solving simple games, Shapley [169] assigned to each player $i \in N$ a payoff $\phi^v_{\mathrm{Sh}}(i)$ which indicates the individual power of i in the game[1]. Such a number is called "Shapley value" or "Shapley power index" of player i with respect to v and is defined by ($t = |T|$):

$$\phi_{\mathrm{Sh}}^v(i) := \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\,t!}{n!}\,[v(T \cup i) - v(T)]. \tag{5.1}$$

If i joins a coalition $T \subseteq N \setminus i$, he might receive a payoff that corresponds to $v(T \cup i) - v(T)$. A swing occurs in a simple game when a player i joining T obtains $v(T \cup i) - v(T) = 1$ and transforms the coalition from losing to winning. We quote from Shapley and Shubik [171]:

> ...our definition of the power of an individual member depends on the chance he has of being critical to the success of a winning coalition.

The Shapley power index (5.1) can be interpreted as the weighted proportion of the number of coalitions which are winning in presence of i and losing in its absence. To make this clearer, it is interesting to rewrite the index as follows:

$$\phi_{\mathrm{Sh}}^v(i) = \frac{1}{n} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus i \\ |T| = t}} [v(T \cup i) - v(T)]. \tag{5.2}$$

Thus, the average value of $v(T \cup i) - v(T)$ is computed first over the coalitions of same size $t \in \{0, \ldots, n-1\}$ and then over all the possible sizes. Consequently, the coalitions containing about $n/2$ players are the less important in the average, since they are numerous and a same player j is very often involved into them.

Note that in terms of the Möbius representation the Shapley power index is written as [169]:

$$\phi_{\mathrm{Sh}}^v(i) = \sum_{T \ni i} \frac{1}{t}\, a(T), \quad i \in N. \tag{5.3}$$

There is in fact another common way of defining a power index, due to Banzhaf [15] (see also Dubey and Shapley [40]). The so-called "Banzhaf value" or "Banzhaf power index", defined as

$$\phi_{\mathrm{B}}^v(i) := \frac{1}{2^{n-1}} \sum_{T \subseteq N \setminus i} [v(T \cup i) - v(T)], \tag{5.4}$$

can be viewed as an alternative to the Shapley value. It assigns to player i the probability that a swing occurs in a simple game when i joins a coalition picked at random from among the 2^{n-1} coalitions not including i.

This power index has been applied to settle constitutional issues in the courts. We quote from Banzhaf [15]:

> ...The voting power of a legislator is not necessary proportional to the number of voters he can cast... The power of a legislator X...(corresponds)...to the number of possible combinations of the entire legislature in which X can alter the outcome by changing his vote...

In terms of the Möbius representation the Banzhaf power index is written as:

$$\phi_{\mathrm{B}}^v(i) = \sum_{T \ni i} \frac{1}{2^{t-1}}\, a(T), \quad i \in N.$$

[1] This non-standard notation will be justified in Section 5.2.

The definition of Shapley and Banzhaf values can be extended to non-simple games[2]. In that case, the notion of swing is replaced by what could be called the *marginal contribution*, i.e. the difference of worth $v(T \cup i) - v(T)$ when player i joins coalition T. The Shapley (or Banzhaf) value related to player i is then a weighted average value of the marginal contribution of i alone in all coalitions. Such a value expresses a power index.

In this context, the parallelism between game theory and multicriteria decision making is clear and has been pointed out in 1992 by Murofushi [129]. The overall importance of a criterion $i \in N$ into a decision problem is not solely determined by the number $v(i)$, but also by all $v(T)$ such that $i \in T$. Indeed, we may have $v(i) = 0$, suggesting that element i is unimportant, but it may happen that for many subsets $S \subseteq N$, $v(S \cup i)$ is much greater than $v(S)$, suggesting that i is actually important. Thus, the importance of criteria can be represented by a power index.

Note that, when v is additive, we clearly have $v(T \cup i) - v(T) = v(i)$ for all $i \in N$ and all $T \subseteq N \setminus i$, and hence

$$\phi^v_{\mathrm{Sh}}(i) = \phi^v_{\mathrm{B}}(i) = v(i), \quad i \in N. \tag{5.5}$$

If v is non-additive then some criteria are dependent and (5.5) generally does not hold anymore. This shows that it is reasonable to search for a coefficient of overall importance for each criterion.

5.1.2 Axiomatic and approximation approaches

We define a *value* or *power index* of the game $v \in \mathcal{G}^N$ to be a function $\phi^v : N \to \mathbb{R}$. Of course, this can be viewed as a vector $(\phi^v(1), \ldots, \phi^v(n)) \in \mathbb{R}^n$ called *payoff vector*. Such a vector may be identified with the additive game $w \in \mathcal{G}^N$ in which the worth of every coalition is the sum of payoffs of its members:

$$w(S) = \sum_{i \in S} \phi^v(i), \quad S \subseteq N. \tag{5.6}$$

A major problem in game theory is how to distribute the worth $v(N)$ of the total coalition among its members in a way that takes into account the worths of the various coalitions. Formally, such a distribution is a payoff vector satisfying $\sum_{i \in N} \phi^v(i) = v(N)$. As we will see later, the Shapley value, which was obtained with an axiomatic approach [169], provides a solution to this problem.

Before presenting an axiomatic characterization which is due to Weber [187], we introduce some definitions.

An element $i \in N$ is said to be a *dummy* player if $v(T \cup i) = v(T) + v(i)$ for all $T \subseteq N \setminus i$. In other words, his marginal contribution to any coalition is simply his individual worth.

For any permutation $\pi \in \Pi_n$, the game πv is defined by $\pi v(\pi(S)) = v(S)$, where $\pi(S) = \{\pi(i) \,|\, i \in S\}$.

Let v^N be a game on N. The *reduced game with respect to* T, $\emptyset \neq T \subseteq N$, is a game denoted $v_{[T]}^{(N \setminus T) \cup [T]}$ defined on the set $(N \setminus T) \cup [T]$ of $(n - t + 1)$ players where $[T]$ indicates a single hypothetical player, which is the representative of the players in T. The reduced game $v_{[T]}$ is defined as follows for any $S \subseteq N \setminus T$:

$$\begin{aligned} v_{[T]}(S) &:= v(S) \\ v_{[T]}(S \cup [T]) &:= v(S \cup T). \end{aligned}$$

Let us consider the following axioms.

[2]Note that, historically, the Shapley value was first introduced [169] for non-simple games. It was then applied in the particular case of simple games [171].

- *Linearity axiom* (L): ϕ^v is a linear function on $\mathcal{G}^N$, that is $\phi^{v+v'} = \phi^v + \phi^{v'}$, and $\phi^{r\,v} = r\,\phi^v$, for any $v, v' \in \mathcal{G}^N$ and any $r > 0$.
- *Dummy axiom* (D'): if $i \in N$ is a dummy player, then $\phi^v(i) = v(i)$.
- *Monotonicity axiom* (M): if v is monotonic, then $\phi^v(i) \geq 0$, for all $i \in N$.
- *Symmetry axiom* (S): for all $v \in \mathcal{G}^N$ and all $\pi \in \Pi_n$, we have $\phi^v(i) = \phi^{\pi v}(\pi(i))$ for all $i \in N$.
- *Efficiency axiom* (E) [169]: for any $v \in \mathcal{G}^N$, we have $\sum_{i\in N} \phi^v(i) = v(N)$.
- *2-efficiency axiom* (2-E) [137]: for any $v \in \mathcal{G}^N$, we have $\phi^v(i) + \phi^v(j) = \phi^{v_{[ij]}}([ij])$.

Let us comment on these axioms. (L) implies that values are linear combinations of the basic information related to the game. (D') means clearly that a dummy player has a value equal to its own worth. (M) says that if the presence of a player in a coalition never "hurts" it, then his power index must be non-negative. (S) requires that the names of the players play no role in determining the value, which should be sensitive only to how the game responds to the presence of a player in a coalition. (E) assures the players share the total amount $v(N)$ among them in terms of their respective values. (2-E) expresses the fact that the sum of the values of two players should be equal to the value of these players considered as twins in the corresponding reduced game.

On the basis of these axioms, Weber [187] showed the following result.

Theorem 5.1.1 *Let ϕ^v be a value defined for any $v \in \mathcal{G}^N$.*

(i) If ϕ^v fulfils (L) then there exists a family of real constants[3] $\{a^i_T \,|\, T \subseteq N\}$ *such that*

$$\phi^v(i) = \sum_{T\subseteq N} a^i_T\, v(T).$$

(ii) If ϕ^v fulfils (L, D') then there exists a family of real constants $\{p^i_T \,|\, T \subseteq N \setminus i\}$ satisfying $\sum_{T\subseteq N\setminus i} p^i_T = 1$ such that

$$\phi^v(i) = \sum_{T\subseteq N\setminus i} p^i_T\,[v(T\cup i) - v(T)]. \tag{5.7}$$

(iii) If ϕ^v fulfils (L, D', M) then, in addition, $p^i_T \geq 0$, for all $i \in N$ and all $T \subseteq N \setminus i$.

(iv) Let ϕ^v be a value of the form (5.7) for all $i \in N$, $v \in \mathcal{G}^N$. If ϕ^v fulfils (S) then there exists a family of constants $p_0, \ldots, p_{n-1}$ such that $p^i_T = p_{|T|}$ for all $i \in N$ and all $T \subseteq N \setminus i$.

(v) If ϕ^v fulfils (L, D', S, E) then it is the Shapley value.

According to Weber [187], the values fulfilling (L, D', M) form the class of *probabilistic values*. As a justification, he proposed the following probabilistic interpretation of these values. Assume that $\{p^i_T \,|\, T \subseteq N \setminus i\}$ is a probability distribution over the collection of coalitions not containing i and suppose that the participation of player i consists merely of joining some coalition S, and then receiving as a reward his marginal contribution $v(S \cup i) - v(S)$ to the coalition. If, for each

[3] Here, constant means independent of v.

$T \subseteq N \setminus i$, p_T^i is the (subjective) probability that he joins coalition T, then (5.7) is simply his expected payoff from the game.

Both the Shapley and Banzhaf values are instances of probabilistic values. $\phi_{\mathrm{B}}^v(i)$ arises from the subjective belief that i is equally likely to join any coalition whereas $\phi_{\mathrm{Sh}}^v(i)$ arises from the belief that the coalition he joins is equally likely to be of any size t $(0 \leq t \leq n-1)$ and that all coalitions of size t are equally likely.

According to Grabisch and Roubens [93], the values fulfilling (L, D', M, S) form the class of *cardinal-probabilistic values*. This is justified by the fact that each coefficient p_T^i only depends on the cardinal of T.

Concerning the axiomatization of Banzhaf value, the following result has been shown by Grabisch and Roubens [92].

Theorem 5.1.2 *Let ϕ^v be a value defined for any $v \in \mathcal{G}$. If ϕ^v fulfils (L, D', S, 2-E) then it is the Banzhaf value.*

Notice that Nowak [137] showed a similar result, but without using (L). Lehrer [109] also axiomatized the Banzhaf value using a weaker form of (2-E).

In addition to the above axiomatic approach, it is natural to ask whether the Shapley and Banzhaf values can be obtained by an approximation approach. The best linear approximation to a game v is a set function $\overline{v}$ defined by $\overline{v}(S) = \overline{a}_0 + \sum_{i \in S} \overline{a}_i$, and which minimizes

$$\sum_{T \subseteq N} [v(T) - \overline{v}(T)]^2.$$

The best linear approximation has been characterized by Hammer and Holzman [95]. It is such that $\overline{a}_i = \phi_{\mathrm{B}}^v(i)$ for all $i \in N$.

Generally, $\sum_{i \in N} \phi_{\mathrm{B}}^v(i) \neq v(N)$ and condition (E) is not satisfied. We can easily overcome this drawback by normalization. Should we normalize by an additive amount or a multiplicative factor? The answer can be found in Hammer and Holzman [95]:

- $\overline{a}_i = \phi_{\mathrm{B}}^v(i) + \frac{1}{n}\,[v(N) - \sum_j \phi_{\mathrm{B}}^v(j)]$ defines the best linear approximation under the constraints $\overline{a}_0 = 0$ and $\sum_{i \in N} \overline{a}_i = v(N)$ but fails to satisfy (D') and the monotonicity of $\overline{v}$ is not ensured.

- $\overline{a}_i = \dfrac{\phi_{\mathrm{B}}^v(i)}{\sum_j \phi_{\mathrm{B}}^v(j)}\, v(N)$ satisfies (E) but also has its problems (see [40] for a discussion).

Note that the Shapley value can also be obtained by the approximation approach. But the least squares criterion in choosing a best approximation has to be replaced by a suitable weighted least squares criterion. The next result is due to Charnes *et al.* [24].

Theorem 5.1.3 *For any game $v \in \mathcal{G}^N$, the additive game $w \in \mathcal{G}^N$ that corresponds to the Shapley value of v (i.e., $w(S) = \sum_{i \in S} \phi_{\mathrm{Sh}}^v(i)$) minimizes*

$$\sum_{T \subseteq N} \rho_T \, [v(T) - w(T)]^2$$

among all additive games w satisfying $w(N) = v(N)$, provided $\rho_T = \binom{n-2}{t-1}^{-1}$ for $T \neq \emptyset, N$.

We elaborate on these approximation considerations in Chapter 7.

5.2 The concept of interaction

In addition to the notion of power index, an interesting concept is that of *interaction* among players or criteria. Actually, the problem of modelling interaction remains a difficult question in multicriteria decision making, often overlooked in practical applications. Although everybody agrees that interaction phenomena do exist in real situations, the lack of suitable tool to model them frequently causes the practitioner to assume that his criteria are independent and exhaustive. This comes primarily from the absence of a precise definition of interaction.

However, the problem has recently been addressed under the viewpoint of cooperative game theory and multicriteria decision making, and an approach which seems suitable has been pointed out. The origin of the idea is due to Murofushi and Soneda [130], who propose an interaction index among a pair of criteria, based on multiattribute utility theory. Later, Grabisch [81] and Roubens [156] generalized this index to any subset S, thus giving rise to the so-called Shapley and Banzhaf interaction indices.

5.2.1 The Shapley and Banzhaf interaction indices

We have observed in Section 5.1 that when a game is not additive, then some players interact. Of course, it would be interesting to appraise the degree of interaction among any subset of players.

Consider a pair $\{i, j\} \subseteq N$ of players. It may happen that $v(i)$ and $v(j)$ are small and at the same time $v(ij)$ is large. The converse could have happened as well, and in the latter case, players i and j have little incentive to cooperate. Clearly, the number $\phi^v(i)$ merely measures the average added worth that player i brings to all possible coalitions, but it does not explain why player i may have a large importance. In other words, it gives no information on the phenomena of interaction or cooperation existing among players. Taking again the above example of players i and j, the difference

$$a(ij) = v(ij) - v(i) - v(j) \tag{5.8}$$

seems to reflect the degree of interaction between these players. Actually we could say:

- players i and j have interest to cooperate, or exhibit a *positive interaction* when the worth of coalition ij is *more* than the sum of individual worths: $a(ij) > 0$,
- players i and j have no interest to cooperate, or exhibit a *negative interaction* when the worth of coalition ij is *less* than the sum of individual worths: $a(ij) < 0$,
- players i and j can act independently in case of equality: $a(ij) = 0$.

Of course, a similar interpretation exists for criteria in multicriteria decision making: the difference (5.8) is positive if there is a synergy effect between i and j. These two criteria then interfere in a positive way. The difference is negative in case of overlap effect between i and j. The criteria then interfere in a negative way. Finally, the difference is zero when the individual importances $v(i)$ and $v(j)$ add up without interfering. In this case, there is no interaction between i and j.

As for power indices, a proper definition of interaction should consider not only $v(i), v(j), v(ij)$ but also the worths of all subsets containing i and j. We may say that i and j have incentives to cooperate when the marginal contribution of j to every subset that contains i is greater that the marginal contribution of j to the same subset when i is excluded, i.e. when

$$v(T \cup ij) - v(T \cup i) > v(T \cup j) - v(T), \quad T \subseteq N \setminus ij.$$

Players i and j can act independently in case of equality for all T and have no interest to cooperate when the inequality is reversed for all T.

Thus the interaction between players i an j can be considered as the average of the marginal contributions of j in the presence of i minus the average of the marginal contributions of j in the absence of i which corresponds to the weighted sum over all coalitions $T \subseteq N \setminus ij$ of

$$v(T \cup ij) - v(T \cup i) - v(T \cup j) + v(T).$$

Murofushi and Soneda [130] have proposed the following definition, borrowing concepts from multiattribute utility theory. The interaction index of elements i, j is defined by

$$I^v(ij) := \sum_{T \subseteq N \setminus ij} \frac{(n-t-2)!\,t!}{(n-1)!}\,[v(T \cup ij) - v(T \cup i) - v(T \cup j) + v(T)],$$

and can be interpreted as a weighted average of the *added value* produced by putting i and j together, all coalitions being considered. When $I^v(ij)$ is positive (resp. negative), then the interaction between i and j is said to be positive (resp. negative).

The interaction index among a combination S of players or criteria has been introduced by Grabisch [81] as a natural extension of the case $|S| = 2$. The *Shapley interaction index* related to v, is defined by

$$I^v_{\mathrm{Sh}}(S) := \sum_{T \subseteq N \setminus S} \frac{(n-t-s)!\,t!}{(n-s+1)!} \sum_{L \subseteq S} (-1)^{s-l} v(L \cup T), \quad S \subseteq N, \tag{5.9}$$

that is, in terms of the Möbius representation [81],

$$I^v_{\mathrm{Sh}}(S) = \sum_{T \supseteq S} \frac{1}{t-s+1}\,a(T), \quad S \subseteq N. \tag{5.10}$$

Viewed as a set function, the Shapley interaction index coincides on singletons with the Shapley value (5.1).

Roubens [156] developed a parallel notion of interaction index, based on the Banzhaf value (5.4): the *Banzhaf interaction index*, defined by

$$I^v_{\mathrm{B}}(S) := \frac{1}{2^{n-s}} \sum_{T \subseteq N \setminus S} \sum_{L \subseteq S} (-1)^{s-l} v(L \cup T), \quad S \subseteq N, \tag{5.11}$$

that is, in terms of the Möbius representation [156],

$$I^v_{\mathrm{B}}(S) = \sum_{T \supseteq S} (\frac{1}{2})^{t-s}\,a(T), \quad S \subseteq N. \tag{5.12}$$

We can see that the interaction indices I^v_{Sh} and I^v_{B} provide extensions of the notion of value, where a value is a function over the set of players (hence the notation $\phi^v(i)$) while the extension is defined over all subsets of players. Clearly, the notion of interaction among players should have a meaning only for at least two players. The interaction of a single player, or of the empty set, has no meaning with regard to the intuitive idea of interaction. However, from a mathematical point of view, I^v_{Sh} and I^v_{B} can be considered as set functions, i.e. defined for *all* subsets of N.

In Section 5.4, we will introduce a third interaction index, namely the *chaining interaction index*

$$I_{\mathrm{R}}^v(S) := \sum_{T\subseteq N\setminus S} s\,\frac{(n-s-t)!\,(s+t-1)!}{n!}\sum_{L\subseteq S}(-1)^{s-l}v(L\cup T), \quad \emptyset\neq S\subseteq N,$$

which also extends the Shapley value. In terms of the Möbius representation, it takes a very simple form:

$$I_{\mathrm{R}}^v(S) = \sum_{T\supseteq S}\frac{s}{t}\,a(T), \quad \emptyset\neq S\subseteq N.$$

Let $\delta_S v : 2^N \to \mathbb{R}$ be the s-th order derivative of v at S defined recursively by $\delta_i v(T) = v(T) - v(T\setminus i)$, $\delta_{ij}v(T) = \delta_i(\delta_j v(T)) = \delta_j(\delta_i v(T))$, etc., for all $T\subseteq N$.

It is easy to show by induction over s that

$$\delta_S v(T\cup S) = \sum_{L\subseteq S}(-1)^{s-l}v(L\cup T), \quad \forall S\subseteq N,\ \forall T\subseteq N\setminus S. \tag{5.13}$$

In particular, for $S=\{i\}$, we obtain the marginal contribution of player i to the coalition $T\subseteq N\setminus i$:

$$\delta_i v(T\cup i) = v(T\cup i) - v(T).$$

For $S=\{i,j\}$, we obtain the marginal interaction between i and j, conditioned to the presence of elements of the coalition $T\subseteq N\setminus ij$:

$$\delta_{ij}v(T\cup ij) = v(T\cup ij) - v(T\cup i) - v(T\cup j) + v(T).$$

More generally, $\delta_S v(T\cup S)$ represents the marginal interaction between the elements of the coalition $S\subseteq N$ in the presence of elements of the coalition $T\subseteq N\setminus S$.

Thus we can see that I_{Sh}^v, I_{B}^v and I_{R}^v are interaction indices of the form

$$I^v(S) = \sum_{T\subseteq N\setminus S} p_T^S\,\delta_S v(T\cup S), \quad \text{with } p_T^S\geq 0,\ \sum_{T\subseteq N\setminus S} p_T^S = 1.$$

Such interactions are called *probabilistic interactions.*

It should be noted that the concept of interaction between elements is not really new. It had already been introduced in statistical analysis of factorial experiments, where the main effects (average contributions) of a number of different factors are investigated simultaneously. The interactions among factors have then been defined to model a degree of dependence between them (see e.g. [30, Chap. 5]).

5.2.2 Axiomatic characterizations

We define an *interaction index* of the game $v\in\mathcal{G}^N$ to be a function $I^v : 2^N \to \mathbb{R}$. Thus, $I^v(S)$ expresses the amount of interaction among s players in coalition S for the game v. $I^v(i)$ represents the value related to player i.

In this section, we present a characterization similar to that of Section 5.1.2 for power indices. In addition to the definitions mentioned there, we introduce the following concepts.

Recall that the *unanimity game* for $T\subseteq N$ is the game v_T such that $v_T(S)=1$ if and only if $S\supseteq T$ and $v_T(S)=0$ otherwise. A slightly different type of simple game is denoted $\hat{v}_T$ and is defined by $\hat{v}_T(S)=1$ if and only if $S\supsetneq T$ and $v_T(S)=0$ otherwise.

Let v^N be a game on N, and $i \in N$. The restriction of v^N to $N \setminus i$, denoted $v^{N\setminus i}$, is defined by $v^{N\setminus i}(S) = v^N(S)$ for all $S \subseteq N \setminus i$. In fact, this is equivalent to consider for v^N only coalitions not containing i. The game on $N \setminus i$ in the presence of i, denoted $v_{\cup i}^{N\setminus i}$, is defined by

$$v_{\cup i}^{N\setminus i}(S) = v^N(S \cup i) - v^N(i), \quad S \subseteq N \setminus i.$$

This is roughly equivalent to consider for v^N only coalitions containing i. Substraction of $v^N(i)$ is introduced only to satisfy the constraint $v_{\cup i}^{N\setminus i}(\emptyset) = 0$.

Let us consider some properties that might be satisfied by an interaction index.

- *Linearity axiom* (L): I^v is a linear function on $\mathcal{G}^N$, that is $I^{v+v'} = I^v + I^{v'}$, and $I^{r\,v} = r\,I^v$, for any $v, v' \in \mathcal{G}^N$ and any $r > 0$.
- *Dummy axiom* (D): if $i \in N$ is a dummy player for $v \in \mathcal{G}^N$, then
  (D') $I^v(i) = v(i)$.
  (D") $I^v(S \cup i) = 0, \quad \forall S \subseteq N \setminus i,\ S \neq \emptyset$.
- *Symmetry axiom* (S): for all $v \in \mathcal{G}^N$ and all $\pi \in \Pi_n$, we have $I^v(i) = I^{\pi v}(\pi(S))$ for all $S \subseteq N$.
- *Recursivity axiom* (R): I^v obeys the following recurrence formula, for any $v \in \mathcal{G}^N$,
$$I^{v^N}(S) = I^{v_{\cup i}^{N\setminus i}}(S \setminus i) - I^{v^{N\setminus i}}(S \setminus i), \quad \forall S \subseteq N,\ S \neq \emptyset,\ \forall i \in S. \tag{5.14}$$
- *Positive interaction axiom* (P): For any game $\hat{v}_T$, $T \subseteq N$, we have
$$I^{\hat{v}_T}(S \cup i) \geq 0, \quad \forall S \subseteq T,\ \forall i \notin T.$$
- *Unanimity games axiom* (U): For any unanimity game v_T, $I^{v_T}(S)$ is maximal for $S = T$. The maximal value is taken equal to one.
- *Efficiency axiom* (E) [169]: for any $v \in \mathcal{G}^N$, we have $\sum_{i\in N} I^v(i) = v(N)$.
- *2-efficiency axiom* (2-E) [137]: for any $v \in \mathcal{G}^N$, we have $I^v(i) + I^v(j) = I^{v_{[ij]}}([ij])$.

(L) implies that values and interactions are linear combinations of the basic information related to the game: the worth of each subset of players.

(D) means that a dummy player has a value equal to its worth and that he/she does not interact with any outside coalition.

(S) indicates that the names of the players play no role in determining the values and interactions.

(R) postulates that interaction at level s is linked to the difference of interactions defined at level $s-1$. More precisely, the interaction between the players in S is equal to the interaction between the players in $S \setminus i$ in the omnipresence of i, minus the interaction between the players of $S \setminus i$ in the absence of i.

In the case of coalition of two players, relation (5.14) gives:

$$I^{v^N}(ij) = I^{v_{\cup i}^{N\setminus i}}(j) - I^{v^{N\setminus i}}(j) = I^{v_{\cup j}^{N\setminus j}}(i) - I^{v^{N\setminus j}}(i)$$

which means that the interaction between i and j is equal to the value of j in the omnipresence of i minus the value of j in the absence of i. Of course, the interaction is a symmetric function of i and j.

(P) indicates that there exists a positive interaction or synergy between $i \notin T$ and $S \subseteq T$ within the simple game $\hat{v}_T$. Indeed, neither i nor S can make $T \setminus S$ to be a winning coalition, but only the contribution of both.

(U) corresponds to the following observation. Consider any unanimity game v_T. If a player i does not belong to T, he/she is a dummy in v_T and plays no active role in that game. On the contrary, any player i that belongs to T plays a major role transforming a losing coalition $T \setminus i$ into a winning one $T : v_T(T \setminus i) = 0$, $v_T((T \setminus i) \cup i) = v_T(T) = 1$. As this is true for any $i \in T$ and only for these elements, the interaction $I^{v_T}(S)$ should be maximal for $S = T$.
For normalization reasons, we consider that the maximal value should be equal to one.

(E) and (2-E) are dedicated to values and have been already explained in Section 5.1.2.

Adopting an approach similar to that of Weber [187], Grabisch and Roubens [92] proved the following.

Theorem 5.2.1 *Let I^v be an interaction index defined for any $v \in \mathcal{G}$.*
(i) If I^v fulfils (L) then, for every $S \subseteq N$, there exists a family of real constants $\{a_T^S \,|\, T \subseteq N\}$ such that

$$I^v(S) = \sum_{T \subseteq N} a_T^S \, v(T).$$

(ii) If I^v fulfils (L, D) then, for every $S \subseteq N$, there exists a family of real constants $\{p_T^S \,|\, T \subseteq N \setminus S\}$ such that

$$I^v(S) = \sum_{T \subseteq N \setminus S} p_T^S \, \delta_S v(T \cup S). \tag{5.15}$$

(iii) If I^v fulfils (L, D, S) then there exists a family of real constants $\{p_t^s(n) \,|\, s = 0, \ldots, n;\ t = 0, \ldots, n - s\}$ such that

$$I^v(S) = \sum_{T \subseteq N \setminus S} p_t^s(n) \, \delta_S v(T \cup S).$$

(iv) If I^v fulfils (L, D, S, R) then, in addition,

$$p_t^s(n) = p_t^1(n - s + 1).$$

In other words, the coefficients $p_t^s(n)$ depend only on t and $n - s$.
(v) If I^v fulfils (L, D, S, R, E) then it is the Shapley interaction index.
(vi) If I^v fulfils (L, D, S, R, 2-E) then it is the Banzhaf interaction index.

Note that the recursivity axiom (R) permits to link interaction indices to values in a unique way. That is, if for example the Shapley value is chosen, the interaction index based on the Shapley value is uniquely determined, and the coefficients $p_t^s(n)$ are known. The same will be true for any value, provided it satisfies (L, D, S).

In addition to the previous result, Grabisch and Roubens [91, 93] proved the following.

Theorem 5.2.2 *Let I^v be an interaction index defined for any $v \in \mathcal{G}^N$.*
(i) If I^v fulfils (L, D, P) then $p_T^S \geq 0$ for all $S \subseteq N$, $S \neq \emptyset$, and all $T \subseteq N \setminus S$.
(ii) If I^v fulfils (L, D, P, U) then, in addition, $\sum_{T \subseteq N \setminus S} p_T^S = 1$ for all $S \subseteq N$.

According to Grabisch and Roubens [93], the interaction indices fulfilling (L, D, P, U) form the class of *probabilistic interactions.* Those fulfilling (L, D, P, U, S) form the class of *cardinal-probabilistic interactions.* I^v_{B} and I^v_{Sh} belong to these classes and have the following probabilistic interpretation: let us suppose that any coalition $S \subseteq N$ joins a coalition $T \subseteq N \setminus S$ at random with a probability p^S_T. Then the interaction index (5.15) can be thought of as the mathematical expectation of the marginal interaction $\delta_S v(T \cup S)$. Depending on the given randomization scheme, this interaction index takes a well defined form:

- if the coalition S is equally likely to join any coalition $T \subseteq N \setminus S$, its probability to join is $p^S_T = \frac{1}{2^{n-s}}$ and we get I_{B}.
- if the coalition S is equally likely to join any coalition $T \subseteq N \setminus S$ of size t $(0 \leq t \leq n-s)$ and that all coalitions of size t are equally likely, its probability to join is $p^S_T = \frac{1}{n-s+1}\binom{n-s}{t}^{-1}$ and we get I_{Sh}.

The results of Theorems 5.2.1 and 5.2.2 are summarized in Figure 5.1.

5.3 Equivalent representations of a set function

As already mentioned, real valued set functions, which are not necessarily additive, are extensively used in decision theory. This section mostly concentrates on some alternative representations of set functions and on their usefulness in game theory and in multicriteria decision making. All the results presented in this section are written in Grabisch, Marichal and Roubens [87].

Consider a real valued set function $v : 2^N \to \mathbb{R}$. There exist several equivalent ways to define v. The first one is to give for any subset S the number $v(S)$. The second one is to observe that v can be expressed in a unique way as:

$$v(S) = \sum_{T \subseteq S} a(T), \quad S \subseteq N,$$

where a is the Möbius transform of v, see Section 4.1.2.

We know that the set function a is a *representation* of v. More formally, a set function $w : 2^N \to \mathbb{R}$ is a representation of v if there exists an invertible transform $\mathcal{T}$ such that

$$w = \mathcal{T}(v) \quad \text{and} \quad v = \mathcal{T}^{-1}(w).$$

In addition to the Möbius representation of v, we introduce the following definitions:

- The *dual representation* of v, denoted v^*, is defined by

$$v^*(S) := v(N) - v(N \setminus S), \quad S \subseteq N.$$

- The *co-Möbius representation* of v, denoted b, is defined by

$$b(S) := \sum_{T \supseteq N \setminus S} (-1)^{n-t} v(T) = \sum_{T \subseteq S} (-1)^t \, v(N \setminus T), \quad S \subseteq N. \tag{5.16}$$

In evidence theory (Shafer [167]), v corresponds to the belief function, v^* is called the plausibility function, a corresponds to the mass or basic probability assignment and b is called the commonality function.

(L): $I^v(S) = \sum_{T\subseteq N} a_T^S\, v(T), \;\; a_T^S \in \mathbb{R}$

(L, D): $I^v(S) = \sum_{T\subseteq N\setminus S} p_T^S\, \delta_S v(T\cup S), \;\; p_T^S \in \mathbb{R}$

(L, D, S)
$p_T^S = p_t^s$
(cardinal interaction)

(L, D, P)
$p_T^S \geq 0$

(L, D, S, R)
$p_t^s(n) = p_t^1(n-s+1)$

(L, D, P, U)
$p_T^S \geq 0, \;\; \sum_{T\subseteq N\setminus S} p_T^S = 1$
(probabilistic interaction)

(L, D, S, R, E)
Shapley interaction $I_{\mathrm{Sh}}^v(S)$
$p_t^s(n) = \dfrac{1}{n-s+1}\dbinom{n-s}{t}^{-1}$

(L, D, P, U, S)
$p_t^s \geq 0, \;\; \sum_{T\subseteq N\setminus S} p_t^s = 1$
(cardinal-probabilistic interaction)

(L, D, S, R, 2-E)
Banzhaf interaction $I_{\mathrm{B}}^v(S)$
$p_t^s(n) = \dfrac{1}{2^{n-s}}$

Figure 5.1: Classification of interactions

If $v(\emptyset) = 0$ then v^* is a representation of v since $(v^*)^* = v$. Moreover, for any v, it is already known that the Möbius transform is invertible and thus is a representation of v. The main aim of this section is to show that b is also a representation of v, as well as the interaction indices I_{B} and I_{Sh}. All these representations are linear, that is, such that $\mathcal{T}$ is a linear operator. We also give all the conversion formulas between v, a, b, I_{B} and I_{Sh}.

In the following three subsections, we will give some technical results involving pseudo-Boolean functions and some of their extensions. These results will be very useful as we continue.

5.3.1 The use of pseudo-Boolean functions

Let us introduce the concept of derivatives of pseudo-Boolean functions, which will be useful in the sequel, see e.g. [95].

Definition 5.3.1 Given $S = \{i_1, \dots, i_s\} \subseteq N$, the s-th order derivative of a pseudo-Boolean function $f : \{0,1\}^n \to \mathbb{R}$ with respect to $x_{i_1}, \dots, x_{i_s}$ is the function $\Delta_S f : \{0,1\}^n \to \mathbb{R}$ defined inductively as

$$\Delta_S f(x) = \Delta_{i_1}(\Delta_{S\setminus i_1} f)(x),$$

where $\Delta_i f(x)$ $(i \in N)$ is the (first) derivative defined by

$$\Delta_i f(x) := f(x\,|\,x_i = 1) - f(x\,|\,x_i = 0), \quad x \in \{0,1\}^n,$$

and, as usual, $\Delta_\emptyset f(x) = f(x)$ for all $x \in \{0,1\}^n$. For all $S \subseteq N$, $\Delta_S f(x)$ will be called the S-derivative of $f(x)$.

Thus defined, $\Delta_S f(x)$ depends only on the variables x_i for $i \notin S$, but we still regard it as a function on $\{0,1\}^n$. For instance, we have, for all $T \subseteq N$,

$$(\Delta_i f)(e_T) = \begin{cases} f(e_T) - f(e_{T\setminus i}), & \text{if } i \in T, \\ f(e_{T\cup i}) - f(e_T), & \text{if } i \notin T. \end{cases}$$

If f is given under the form (4.2) then we can easily see that:

$$\Delta_S f(x) = \sum_{T \supseteq S} a(T) \prod_{i \in T\setminus S} x_i, \quad \forall x \in \{0,1\}^n, \ \forall S \subseteq N. \tag{5.17}$$

Hence we have

$$(\Delta_S f)(e_T) = \sum_{L \subseteq T} a(L \cup S), \quad \forall S \subseteq N, \ \forall T \subseteq N \setminus S. \tag{5.18}$$

In fact, we can easily see that

$$(\Delta_S f)(e_T) = \delta_S v(T \cup S), \quad \forall S, T \subseteq N, \tag{5.19}$$

and, by (5.13), we have

$$(\Delta_S f)(e_T) = \sum_{L \subseteq S} (-1)^{s-l} v(L \cup T), \quad \forall S \subseteq N, \ \forall T \subseteq N \setminus S. \tag{5.20}$$

Moreover, by combining (5.20) and (5.16), we obtain

$$b(S) = (\Delta_S f)(\overline{e}_S) = (\Delta_S f)(e_{N\setminus S}), \quad S \subseteq N, \tag{5.21}$$

and by using (5.17),

$$b(S)=\sum_{T\supseteq S} a(T), \quad S\subseteq N. \tag{5.22}$$

Now, by (5.19), we can see that the interaction indices I_{B} and I_{Sh} are of the form:

$$I(S)=\sum_{T\subseteq N\setminus S} p_T^S\,(\Delta_S\, f)(e_T), \quad S\subseteq N,$$

and equations (5.11) and (5.9) become:

$$I_{\mathrm{B}}(S) = \frac{1}{2^{n-s}}\sum_{T\subseteq N\setminus S}(\Delta_S\, f)(e_T), \quad S\subseteq N, \tag{5.23}$$

$$I_{\mathrm{Sh}}(S) = \frac{1}{n-s+1}\sum_{T\subseteq N\setminus S}\binom{n-s}{t}^{-1}(\Delta_S\, f)(e_T), \quad S\subseteq N. \tag{5.24}$$

The following result shows that equations (5.23) and (5.24) can be rewritten in another form.

Proposition 5.3.1 *We have*

$$I_{\mathrm{B}}(S) = \frac{1}{2^{n}}\sum_{x\in\{0,1\}^n}(\Delta_S\, f)(x), \quad S\subseteq N, \tag{5.25}$$

$$I_{\mathrm{Sh}}(S) = \frac{1}{n+1}\sum_{x\in\{0,1\}^n}\binom{n}{\sum_i x_i}^{-1}(\Delta_S\, f)(x), \quad S\subseteq N. \tag{5.26}$$

Proof. Given $S\subseteq N$, we simply have

$$\begin{aligned}
\frac{1}{2^n}\sum_{x\in\{0,1\}^n}(\Delta_S\, f)(x) &= \frac{1}{2^n}\sum_{T\supseteq S}a(T)\sum_{x\in\{0,1\}^n}\prod_{i\in T\setminus S}x_i \quad \text{(by (5.17))}\\
&= \frac{1}{2^n}\sum_{T\supseteq S}a(T)\sum_{K\subseteq N\setminus(T\setminus S)}1\\
&= \sum_{T\supseteq S}(\frac{1}{2})^{t-s}a(T)\\
&= I_{\mathrm{B}}(S) \quad \text{(by (5.12))},
\end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{n+1}\sum_{x\in\{0,1\}^n}\binom{n}{\sum_i x_i}^{-1}(\Delta_S\, f)(x)\\
&= \frac{1}{n+1}\sum_{T\supseteq S}a(T)\sum_{x\in\{0,1\}^n}\binom{n}{\sum_i x_i}^{-1}\prod_{i\in T\setminus S}x_i \quad \text{(by (5.17))}\\
&= \frac{1}{n+1}\sum_{T\supseteq S}a(T)\sum_{K\subseteq N\setminus(T\setminus S)}\binom{n}{k}^{-1}\\
&= \frac{1}{n+1}\sum_{T\supseteq S}a(T)\sum_{k=0}^{n-t+s}\binom{n-t+s}{k}\binom{n}{k}^{-1}\\
&= \sum_{T\supseteq S}\frac{1}{t-s+1}a(T)\\
&= I_{\mathrm{Sh}}(S) \quad \text{(by (5.10))},
\end{aligned}$$

which proves the result. ∎

It should be noted that equations (5.12) and (5.10) can be easily obtained from (5.23) and (5.24) respectively by using the following formula:

$$\sum_{T\subseteq N\setminus S} p_t^s\,(\Delta_S f)(e_T) = \sum_{T\supseteq S}\Big[\sum_{k=0}^{n-t}\binom{n-t}{k}p_{k+t-s}^s\Big]a(T),\quad S\subseteq N. \tag{5.27}$$

The proof of this formula is simple: setting $L' := L\cup S$, we have, from (5.18),

$$\begin{aligned}
\sum_{T\subseteq N\setminus S} p_t^s\,(\Delta_S f)(e_T) &= \sum_{T\subseteq N\setminus S} p_t^s \sum_{L\subseteq T} a(L\cup S)\\
&= \sum_{L'\supseteq S}\Big[\sum_{T:L'\setminus S\subseteq T\subseteq N\setminus S} p_t^s\Big]a(L')\\
&= \sum_{L'\supseteq S}\Big[\sum_{t=l'-s}^{n-s}\binom{n-l'}{t-l'+s}p_t^s\Big]a(L')\\
&= \sum_{L'\supseteq S}\Big[\sum_{k=0}^{n-l'}\binom{n-l'}{k}p_{k+l'-s}^s\Big]a(L'),
\end{aligned}$$

which proves (5.27).

5.3.2 Multilinear extension of pseudo-Boolean functions

From any pseudo-Boolean function $f:\{0,1\}^n\to\mathbb{R}$, we can define a variety of extensions $\bar f:[0,1]^n\to\mathbb{R}$ which interpolate f at the 2^n vertices of $[0,1]^n$, that is $\bar f(e_S)=f(e_S)=v(S)$ for all $S\subseteq N$. For instance, we have seen that the Choquet integral with respect to a set function v is the Lovász extension of the pseudo-Boolean function which defines v, see Section 4.2.

The S-derivative of any extension $\bar f$ is defined inductively in the same way as for f. In particular, we have

$$\Delta_S\,\bar f(x) = \Delta_S\, f(x),\quad \forall x\in\{0,1\}^n,\ \forall S\subseteq N.$$

Let us introduce the notation $\underline{x} := (x,\dots,x)\in[0,1]^n$ for all $x\in[0,1]$. By (5.18) and (5.21), we immediately have

$$\begin{aligned}
a(S) &= (\Delta_S\,\bar f)(\underline{0}),\quad S\subseteq N\\
b(S) &= (\Delta_S\,\bar f)(\underline{1}),\quad S\subseteq N
\end{aligned}$$

for any extension $\bar f$ of f.

The polynomial expression (4.2) was used in game theory in 1972 by Owen [144] as the *multilinear extension* of a game.

Definition 5.3.2 If the pseudo-Boolean function f has the unique multilinear expression (4.2) then the *multilinear extension* of f (MLE) is the function $g:[0,1]^n\to\mathbb{R}$ defined by

$$g(x) := \sum_{T\subseteq N} f(e_T)\prod_{i\in T}x_i\prod_{i\notin T}(1-x_i) = \sum_{T\subseteq N} a(T)\prod_{i\in T}x_i,\quad x\in[0,1]^n. \tag{5.28}$$

It has been proved by Owen [145] that g is the only multilinear function (i.e. linear in each of the variables x_i) on $[0,1]^n$ that coincides with f on $\{0,1\}^n$. More precisely, g corresponds to the classical linear interpolation (with respect to each of the n variables) of f.

It is easy to see that:

$$\Delta_S\, g(x) = \sum_{T \supseteq S} a(T) \prod_{i \in T \setminus S} x_i, \quad \forall x \in [0,1]^n, \ \forall S \subseteq N, \tag{5.29}$$

and, by (5.17), we can observe that $\Delta_S\, g$ is the MLE of $\Delta_S\, f$.

From (5.12) and (5.29), we can readily see that the Banzhaf interaction index related to S is obtained by integrating the S-derivative of the MLE of game v over the hypercube. Formally, this result can be stated as follows.

Proposition 5.3.2 *We have*

$$I_{\mathrm{B}}(S) = \int_{[0,1]^n} (\Delta_S\, g)(x)\, dx, \quad S \subseteq N. \tag{5.30}$$

This result can be interpreted by analogy with (5.25): $I_{\mathrm{B}}(S)$ is the average value of $\Delta_S\, f$ over $\{0,1\}^n$, but also the average value of its MLE over $[0,1]^n$.

From (5.29), we immediately have:

$$(\Delta_S\, g)(\underline{x}) = \sum_{T \supseteq S} a(T)\, x^{t-s}, \quad \forall x \in [0,1], \ \forall S \subseteq N. \tag{5.31}$$

Consequently, we have, using (5.22), (5.12), and (5.10):

$$\begin{aligned}
a(S) &= (\Delta_S\, g)(\underline{0}), \quad S \subseteq N & (5.32)\\
b(S) &= (\Delta_S\, g)(\underline{1}), \quad S \subseteq N & (5.33)\\
I_{\mathrm{B}}(S) &= (\Delta_S\, g)(\underline{1/2}), \quad S \subseteq N & (5.34)\\
I_{\mathrm{Sh}}(S) &= \int_0^1 (\Delta_S\, g)(\underline{x})\, dx, \quad S \subseteq N. & (5.35)
\end{aligned}$$

We see that the Banzhaf interaction index related to S is the value of the S-derivative of the MLE of game v on the center of the hypercube $[0,1]^n$, while the Shapley interaction index related to S is obtained by integrating the S-derivative of the MLE of game v along the main diagonal of the hypercube. This latter result has been proved by Owen [145] when $|S| = 1$.

5.3.3 Links with the Lovász extension

Let $\hat{f}$ be the Lovász extension of f on $[0,1]^n$. It is easy to see that:

$$\Delta_S\, \hat{f}(x) = \sum_{T \supseteq S} a(T) \bigwedge_{i \in T \setminus S} x_i, \quad \forall x \in [0,1]^n, \ \forall S \subseteq N, \tag{5.36}$$

and, by (5.17), we can observe that $\Delta_S\, \hat{f}$ is the Lovász extension of $\Delta_S\, f$.

The following lemma will be very useful in the sequel.

Lemma 5.3.1 *We have*

$$\int_{[0,1]^n} \bigwedge_{i \in S} x_i\, dx = \frac{1}{s+1}, \quad S \subseteq N. \tag{5.37}$$

Proof. Observe first that we can assume $S = N$. Next, we have

$$\begin{aligned}\int_{[0,1]^n} \bigwedge_{i\in N} x_i\,dx &= \sum_{\pi\in\Pi_n} \int_{\mathcal{B}_\pi} x_{\pi(1)}\,dx\\ &= \sum_{\pi\in\Pi_n} \int_0^1 \int_0^{x_{\pi(n)}} \cdots \int_0^{x_{\pi(2)}} x_{\pi(1)}\,dx_{\pi(1)}\cdots dx_{\pi(n)}\\ &= \sum_{\pi\in\Pi_n} \frac{1}{(n+1)!} = \frac{1}{n+1}.\end{aligned}$$

■

From (5.10), (5.36) and (5.37), we can readily see that the Shapley interaction index related to S is obtained by integrating the S-derivative of the Lovász extension of game v over the hypercube. This result, which is to be compared with (5.30), can be stated as follows.

Proposition 5.3.3 *We have*

$$I_{\mathrm{Sh}}(S) = \int_{[0,1]^n} \Delta_S\,\hat{f}(x)\,dx, \quad S \subseteq N. \tag{5.38}$$

From (5.36), we immediately have

$$(\Delta_S\,\hat{f})(\underline{x}) = a(S) + x \sum_{\substack{T\supset S\\ T\neq S}} a(T), \quad \forall x \in [0,1],\ \forall S \subseteq N.$$

Consequently, we have, using (5.22):

$$\begin{aligned} a(S) &= (\Delta_S\,\hat{f})(\underline{0}), \quad S\subseteq N\\ b(S) &= (\Delta_S\,\hat{f})(\underline{1}), \quad S\subseteq N\\ \frac{a(S)+b(S)}{2} &= (\Delta_S\,\hat{f})(\underline{1/2}), \quad S\subseteq N\\ \frac{a(S)+b(S)}{2} &= \int_0^1 (\Delta_S\,\hat{f})(\underline{x})\,dx, \quad S\subseteq N\end{aligned}$$

5.3.4 Some conversion formulas derived from the MLE

It is easy to see that, for any function g of the form (5.28), the operator Δ_S identifies with the classical S-derivative, that is,

$$\Delta_S\,g(x) = \frac{\partial^s\,g(x)}{\partial x_{i_1}\cdots\partial x_{i_s}} \quad \text{where } S = \{i_1,\ldots,i_s\}.$$

The Taylor formula for functions of several variables then can be applied to g. This leads to the equality:

$$g(x) = \sum_{T\subseteq N}\prod_{i\in T}(x_i - y_i)\,\Delta_T\,g(y), \quad x,y\in[0,1]^n. \tag{5.39}$$

Replacing x by e_S and y by $\underline{y}$ provides:

$$v(S) = \sum_{T\subseteq N}\prod_{i\in T}((e_S)_i - y)\,(\Delta_T\,g)(\underline{y}), \quad \forall y\in[0,1],\ \forall S\subseteq N. \tag{5.40}$$

On the basis of (5.32)–(5.34), we can obtain the conversions from a, b, I_{B} to v by replacing y respectively by 0, 1 and 1/2 in (5.40). The corresponding formulas can be found in Tables 5.3 and 5.4 (Section 5.3.7).

By successive derivations of (5.39), we obtain:

$$\Delta_S\, g(x) = \sum_{T\supseteq S} \prod_{i\in T\setminus S} (x_i - y_i)\, \Delta_T g(y), \quad \forall x, y \in [0,1]^n,\ \forall S \subseteq N.$$

In particular, we have:

$$(\Delta_S\, g)(\underline{x}) = \sum_{T\supseteq S} (x-y)^{t-s}\, (\Delta_T g)(\underline{y}), \quad \forall x, y \in [0,1],\ \forall S \subseteq N. \tag{5.41}$$

We can get all the conversions between a, b, and I_{B} by replacing x and y by 0, 1, and 1/2 in (5.41). The corresponding formulas are written in Tables 5.3 and 5.4.

Combining (5.35) and (5.41), we immediately have:

$$\begin{aligned} I_{\mathrm{Sh}}(S) &= \sum_{T\supseteq S} \left[\int_0^1 (x-y)^{t-s}\, dx\right] (\Delta_T g)(\underline{y}) \\ &= \sum_{T\supseteq S} \frac{(1-y)^{t-s+1} - (-y)^{t-s+1}}{t-s+1}\, (\Delta_T g)(\underline{y}), \quad \forall y \in [0,1],\ \forall S \subseteq N. \end{aligned} \tag{5.42}$$

We then obtain the conversions from a, b, I_{B} to I_{Sh} by replacing y successively by 0, 1, and 1/2 in (5.42), see Tables 5.3 and 5.4.

The conversions from I_{Sh} to a, b, I_{B}, are a little bit more delicate. Let $\{B_n\}_{n\in\mathbb{N}}$ be the sequence of Bernoulli numbers defined recursively by

$$\begin{cases} B_0 = 1, \\ \displaystyle\sum_{k=0}^{n} \binom{n+1}{k} B_k = 0, \quad n \in \mathbb{N}_0. \end{cases}$$

The first elements of the sequence are:

$$B_0 = 1, B_1 = -\frac{1}{2}, B_2 = \frac{1}{6}, B_3 = 0, B_4 = -\frac{1}{30}, B_5 = 0, B_6 = \frac{1}{42}, \ldots. \tag{5.43}$$

The Bernoulli polynomials are then defined by

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k\, x^{n-k}, \quad \forall n \in \mathbb{N},\ \forall x \in \mathbb{R}.$$

It is well known that these polynomials fulfil the following properties (see e.g. [1]):

$$B_n(0) = B_n, \quad \forall n \in \mathbb{N} \tag{5.44}$$

$$B_n(1) = (-1)^n\, B_n, \quad \forall n \in \mathbb{N} \tag{5.45}$$

$$B_n(1/2) = (\frac{1}{2^{n-1}} - 1)\, B_n, \quad \forall n \in \mathbb{N} \tag{5.46}$$

$$B_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} B_k(x)\, y^{n-k}, \quad \forall n \in \mathbb{N},\ \forall x, y \in \mathbb{R} \tag{5.47}$$

$$\int_0^1 B_n(x)\, dx = 0, \quad \forall n \in \mathbb{N}_0. \tag{5.48}$$

We have the following lemma.

Lemma 5.3.2 *For all $S, K \subseteq N$ such that $S \subseteq K$, we have:*

$$\sum_{T:S\subseteq T\subseteq K} B_{t-s}(x)\,\frac{1}{k-t+1} = x^{k-s}, \quad x\in[0,1]; \tag{5.49}$$

Proof. We have

$$\begin{aligned}
\sum_{T:S\subseteq T\subseteq K} B_{t-s}(x)\,\frac{1}{k-t+1} &= \sum_{t=s}^{k}\binom{k-s}{t-s}B_{t-s}(x)\,\frac{1}{k-t+1}\\
&= \sum_{u=0}^{k-s}\binom{k-s}{u}B_u(x)\,\frac{1}{k-s-u+1}\\
&= \int_0^1\sum_{u=0}^{k-s}\binom{k-s}{u}B_u(x)\,y^{k-s-u}\,dy\\
&= \int_0^1\sum_{u=0}^{k-s}\binom{k-s}{u}B_u(y)\,x^{k-s-u}\,dy \quad \text{(by (5.47))}\\
&= x^{k-s} \quad \text{(by (5.48))}.
\end{aligned}$$

■

We then have the following result.

Proposition 5.3.4 *We have*

$$(\Delta_S\, g)(\underline{x}) = \sum_{T\supseteq S} B_{t-s}(x)\, I_{\mathrm{Sh}}(T), \quad \forall x\in[0,1],\ \forall S\subseteq N; \tag{5.50}$$

Proof. We have

$$\begin{aligned}
\sum_{T\supseteq S} B_{t-s}(x)\, I_{\mathrm{Sh}}(T) &= \sum_{T\supseteq S} B_{t-s}(x)\sum_{K\supseteq T}\frac{1}{k-t+1}\,a(K) \quad \text{(by (5.10))}\\
&= \sum_{K\supseteq S} a(K)\sum_{T:S\subseteq T\subseteq K} B_{t-s}(x)\,\frac{1}{k-t+1}\\
&= \sum_{K\supseteq S} a(K)\,x^{k-s} \quad \text{(by (5.49))}\\
&= (\Delta_S\, g)(\underline{x}) \quad \text{(by (5.31))}.
\end{aligned}$$

■

We then obtain the conversions from I_{Sh} to a, b, I_{B} by replacing x successively by 0, 1, and 1/2 in (5.50), and by using (5.44)–(5.46).

5.3.5 Fractal and cardinality transformations

We now present the matrix form of the linear transforms that allow to go from a representation to another one. We also show that the corresponding matrices have remarkable properties.

Any pair (x, y) extracted from the set $\{v, a, b, I_{\mathrm{B}}, I_{\mathrm{Sh}}\}$ can produce a matricial relation

$$y = T\circ x$$

where $x, y : 2^N \to \mathrm{I\!R}$ and where T is a *transformation matrix* of dimension $2^n \times 2^n$ if the 2^n elements of 2^N are ordered according to some sequence.

Let us consider the following total ordering of the elements of 2^N,

$$\mathcal{O} : \emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{1,3\}, \{2,3\}, \{1,2,3\}, \{4\}, \ldots, N.$$

This order is obtained as follows. We consider the natural sequence of integers from 0 to $2^n - 1$, that is $0, 1, 2, \ldots, i, \ldots, 2^n - 1$, and its binary notation $[0]_2, [1]_2, \ldots, [i]_2, \ldots, [2^n - 1]_2$, which is (with n digits) $000\cdots00, 000\cdots01, 000\cdots10, \ldots, 111\cdots11$. To any number $[i]_2$ in binary notation corresponds a unique subset $I \subseteq N$ such that $j \in I$ if and only if the $(n+1-j)$-th digit in $[i]_2$ is 1.

We obtain the vectors of $\mathrm{I\!R}^{2^n}$:

$$\begin{aligned} x^t_{(n)} &= (\, x(\emptyset) \quad x(\{1\}) \quad x(\{2\}) \quad x(\{1,2\}) \quad \ldots \quad x(N)\,) \\ y^t_{(n)} &= (\, y(\emptyset) \quad y(\{1\}) \quad y(\{2\}) \quad y(\{1,2\}) \quad \ldots \quad y(N)\,) \end{aligned}$$

(here the superscript t represents the transposition operation) and we determine the matricial relation

$$y_{(n)} = T_{(n)} \circ x_{(n)}$$

with

$$T_{(n)} = \begin{array}{c} \\ \emptyset \\ \{1\} \\ \vdots \\ N \end{array} \begin{array}{c} \begin{array}{cccc} \emptyset & \{1\} & \cdots & N \end{array} \\ \begin{pmatrix} T(\emptyset,\emptyset) & T(\emptyset,\{1\}) & \cdots & T(\emptyset,N) \\ T(\{1\},\emptyset) & T(\{1\},\{1\}) & \cdots & T(\{1\},N) \\ \vdots & \vdots & & \vdots \\ T(N,\emptyset) & T(N,\{1\}) & \cdots & T(N,N) \end{pmatrix} \end{array}.$$

Three particular transformations will be considered:

(i) the *fractal* transformation linked to a "fractal matrix" $T = F$ defined with the help of one "basic fractal matrix" $F_{(1)}$ which is supposed to be invertible.

$$\begin{aligned} F_{(1)} &:= \begin{pmatrix} f_1 & f_2 \\ f_3 & f_4 \end{pmatrix}, \quad f_i \in \mathrm{I\!R}, \ \ i = 1,2,3,4 \\ F^{-1}_{(1)} &:= \begin{pmatrix} g_1 & g_2 \\ g_3 & g_4 \end{pmatrix} \\ F_{(k)} &:= \begin{pmatrix} f_1\, F_{(k-1)} & f_2\, F_{(k-1)} \\ f_3\, F_{(k-1)} & f_4\, F_{(k-1)} \end{pmatrix}, \quad k = 2, \ldots, n. \end{aligned}$$

It can be shown that the inverse matrix is also fractal. In general we have:

$$F^{-1}_{(k)} = \begin{pmatrix} g_1\, F^{-1}_{(k-1)} & g_2\, F^{-1}_{(k-1)} \\ g_3\, F^{-1}_{(k-1)} & g_4\, F^{-1}_{(k-1)} \end{pmatrix}, \quad k = 2, \ldots, n.$$

(ii) the *upper-cardinality* transformation linked to an "upper-cardinality matrix" $T = C$ based on a sequence of real numbers $(c_0, c_1, \ldots, c_k, \ldots, c_n)$, $c_0 = 1$, and

$$\begin{aligned} C_{(1)} &:= \begin{pmatrix} c_0 & c_1 \\ 0 & c_0 \end{pmatrix}, \quad C^l_{(1)} := \begin{pmatrix} c_{l-1} & c_l \\ 0 & c_{l-1} \end{pmatrix}, \ l = 1, \ldots, n \\ C_{(2)} &:= \begin{pmatrix} C^1_{(1)} & C^2_{(1)} \\ 0 & C^1_{(1)} \end{pmatrix}, \quad C^l_{(2)} := \begin{pmatrix} C^l_{(1)} & C^{l+1}_{(1)} \\ 0 & C^l_{(1)} \end{pmatrix}, \ l = 1, \ldots, n-1 \\ C_{(k)} &:= \begin{pmatrix} C^1_{(k-1)} & C^2_{(k-1)} \\ 0 & C^1_{(k-1)} \end{pmatrix}, \ k = 2, \ldots, n. \end{aligned}$$

Using the sequence $\mathcal{O}$ to order the rows and the columns of $C_{(n)}$, one obtains (blanks replace zeroes):

$$C_{(n)} = \begin{array}{c} \\ \emptyset \\ \{1\} \\ \{2\} \\ \{1,2\} \\ \{3\} \\ \{1,3\} \\ \{2,3\} \\ \{1,2,3\} \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccccccc} \emptyset & \{1\} & \{2\} & \{1,2\} & \{3\} & \{1,3\} & \{2,3\} & \{1,2,3\} & \cdots \end{array} \\ \left(\begin{array}{cccccccc} c_0 & c_1 & c_1 & c_2 & c_1 & c_2 & c_2 & c_3 \\ & c_0 & & c_1 & & c_1 & & c_2 \\ & & c_0 & c_1 & & & c_1 & c_2 \\ & & & c_0 & & & & c_1 \\ & & & & c_0 & c_1 & c_1 & c_2 \\ & & & & & c_0 & & c_1 \\ & & & & & & c_0 & c_1 \\ & & & & & & & c_0 \\ & & & & & & & \end{array}\right) \end{array}.$$

(iii) the *lower-cardinality* transformation linked to a "lower-cardinality matrix" $T = C^t$ based on a sequence of real numbers $(c_0, c_1, \ldots, c_k, \ldots, c_n)$, $c_0 = 1$, and

$$\begin{aligned} C^t_{(1)} &:= \begin{pmatrix} c_0 & 0 \\ c_1 & c_0 \end{pmatrix}, \ldots \\ C^t_{(k)} &:= \begin{pmatrix} C^{1\,t}_{(k-1)} & 0 \\ C^{2\,t}_{(k-1)} & C^{1\,t}_{(k-1)} \end{pmatrix}, \ k = 2, \ldots, n. \end{aligned}$$

Both fractal and cardinality transformations correspond to a two-place real valued set function Φ. We introduce the product of two such transformations Φ and Ψ to define:

$$(\Phi \circ \Psi)(A, B) := \sum_{C \subseteq N} \Phi(A, C)\, \Psi(C, B), \quad A, B \subseteq N.$$

In the case of the upper-cardinality transformation (see Denneberg and Grabisch [37])

$$\Phi(A, B) = \Phi(\emptyset, B \setminus A) = \begin{cases} c_{|B \setminus A|}, & \text{if } A \subseteq B, \\ 0, & \text{otherwise,} \end{cases}$$

and this definition justifies the terminology used.

If we are concerned with a lower-cardinality transformation,

$$\Phi(A, B) = \Phi(A \setminus B, \emptyset) = \begin{cases} c_{|A \setminus B|}, & \text{if } B \subseteq A, \\ 0, & \text{otherwise.} \end{cases}$$

Let us now consider the families:

$$\begin{aligned} \mathcal{G}_F &:= \{F : 2^N \times 2^N \to \mathbb{R} \mid F \text{ is built on a basic invertible fractal matrix } F_{(1)}\} \\ \mathcal{G}_{\overline{C}} &:= \{C : 2^N \times 2^N \to \mathbb{R} \mid C \text{ is determined by a sequence } (c_k)\} \\ \mathcal{G}_{\underline{C}} &:= \{C^t : 2^N \times 2^N \to \mathbb{R} \mid C^t \text{ is determined by a sequence } (c_k)\} \end{aligned}$$

The three families form a multiplicative group for the composition law $(\circ)$ with neutral element

$$I(A, B) := \begin{cases} 1, & \text{if } A = B, \\ 0, & \text{else.} \end{cases}$$

The families $\mathcal{G}_{\overline{C}}$ and $\mathcal{G}_{\underline{C}}$ form an Abelian group (i.e. commutative) but the property of commutativity is generally not satisfied for $\mathcal{G}_F$.

In the case of the upper-cardinality transformation, $y_{(n)} = C_{(n)}\, x_{(n)}$ can be rewritten as

$$y(S) = \sum_{T \supseteq S} c_{t-s}\, x(T), \quad S \subseteq N. \tag{5.51}$$

Moreover, if C^1 and C^2 represent two upper-cardinality transformations, the sequence (c_k) related to $C^1 \circ C^2$ corresponds to (see [37])

$$c_k = \sum_{l=0}^{k} \binom{k}{l} c^1_{k-l}\, c^2_l = \sum_{l=0}^{k} \binom{k}{l} c^2_{k-l}\, c^1_l, \quad k = 0, \dots, n. \tag{5.52}$$

The inverse C^{-1} of the upper-cardinality transformation C is obtained with $c_0^{-1} = 1$ and

$$c_k^{-1} = -\sum_{l=0}^{k-1} \binom{k}{l} c_{k-l}\, c_l^{-1}, \quad k = 1, \dots, n. \tag{5.53}$$

It is obvious that the lower-cardinality transformation $y_{(n)} = C^t_{(n)}\, x_{(n)}$ can be rewritten as

$$y(S) = \sum_{T \subseteq S} c_{s-t}\, x(T), \quad S \subseteq N.$$

If $C^{1\,t}$ and $C^{2\,t}$ represent two lower-cardinality transformations, the sequence (c_k) related to $C^{1\,t} \circ C^{2\,t}$ corresponds to the formula (5.52) and the inverse $(C^t)^{-1}$ of the lower-cardinality transformation C^t is obtained with (5.53).

If a fractal transformation F is considered, $y_{(n)} = F_{(n)}\, x_{(n)}$ can be rewritten as

$$y(S) = \sum_{T \subseteq N} F(S,T)\, x(T), \quad S \subseteq N.$$

We know that F^{-1} is also a fractal transformation and we can easily check that

$$F^{-1}_{(n)}(S,T) = \frac{(-1)^{t-s}}{(\det F_{(1)})^n}\, F_{(n)}(N \setminus T, N \setminus S), \quad \forall S, T \subseteq N. \tag{5.54}$$

Moreover, the composition of two fractal transformations F^1 and F^2 corresponds to a fractal transformation with basic fractal matrix $F_{(1)} = F^1_{(1)} \circ F^2_{(1)}$.

It should be noted that any fractal transformation with a basic fractal matrix:

$$F_{(1)} = \begin{pmatrix} 1 & \rho \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 0 \\ \rho & 1 \end{pmatrix}$$

is an upper (lower)-cardinality transformation with the sequence $c_k = \rho^k$. The converse is also true.

From classical results in combinatorics [154], all conversion formulas between v, a and b are well known. We can observe that all the transformations between v, a, b and I_{B} are fractal. For instance, the Möbius representation (4.4) can be rewritten under the fractal form

$$a_{(n)} = M_{(n)} \circ v_{(n)}$$

with the use of the basic fractal matrix:

$$M_{(1)} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}. \tag{5.55}$$

We see that transformation M also corresponds to a lower-cardinality transformation with $c_k = (-1)^k$ and we immediately obtain that

$$v_{(n)} = M^{-1}_{(n)} \circ a_{(n)}$$

where M^{-1} corresponds to the basic fractal matrix:

$$M^{-1}_{(1)} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

or the lower-cardinality transformation with sequence $c_k = 1$, which gives (4.5).

More generally, one can easily see that the generating conversion formula (5.40) corresponds, for any fixed $y \in [0,1]$, to a fractal transformation whose basic fractal matrix is

$$F_{(1)} = \begin{pmatrix} 1 & -y \\ 1 & 1-y \end{pmatrix}.$$

By (5.54), the formula (5.40) can immediately be inverted into

$$(\Delta_S\, g)(\underline{y}) = \sum_{T \subseteq N} (-1)^{t-s} \prod_{i \in N \setminus S} ((\overline{e}_T)_i - y)\, v(T), \quad y \in [0,1]. \tag{5.56}$$

Replacing y respectively by 0, 1 and 1/2 in (5.56), we obtain the conversions from v to a, b and I_{B}, see Table 5.3.

The generating conversion formula (5.41) corresponds, for any fixed $x, y \in [0,1]$, to a fractal transformation with basic fractal matrix:

$$F_{(1)} = \begin{pmatrix} 1 & x-y \\ 0 & 1 \end{pmatrix}.$$

Observe that this transformation also corresponds to an upper-cardinality transformation with sequence $c_k = (x-y)^k$.

We have just shown that all the transformations between v, a, b and I_{B} are fractal. The corresponding basic fractal matrices are summarized in Table 5.1.

Due to (5.51), it is clear that the generating conversion formula (5.42) corresponds, for any fixed $y \in [0,1]$, to an upper-cardinality transformation with sequence

$$c_k = \int_0^1 (x-y)^k\, dx = \frac{(1-y)^{k+1} - (-y)^{k+1}}{k+1},$$

whereas the inverse transformation (5.50) corresponds to an upper-cardinality transformation with sequence $c_k^{-1} = B_k(y)$. Thus, all the transformations between a, b, I_{B} and I_{Sh} are upper-cardinality transformations. The corresponding sequences are summarized in Table 5.2.

5.3.6 Pascal matrices

Now, let us turn to the two remaining cases: the transformations from v to I_{Sh} and the converse, which are neither fractal, nor upper-cardinal. From (5.9), we obtain, by setting $T' := T \cup L$ (which implies $L = T' \cap S$ and $T = T' \setminus S$):

$$I_{\mathrm{Sh}}(S) = \sum_{T' \subseteq N} \frac{|N \setminus (S \cup T')|!\, |T' \setminus S|!}{(n-s+1)!} (-1)^{|S \setminus T'|}\, v(T'), \quad S \subseteq N,$$

	v	a	b	I_{B}
v	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$	$\begin{pmatrix}1 & -1\\ 1 & 0\end{pmatrix}$	$\begin{pmatrix}1 & -1/2\\ 1 & 1/2\end{pmatrix}$
a	$\begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & -1/2\\ 0 & 1\end{pmatrix}$
b	$\begin{pmatrix}0 & 1\\ -1 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 1/2\\ 0 & 1\end{pmatrix}$
I_{B}	$\begin{pmatrix}1/2 & 1/2\\ -1 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 1/2\\ 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & -1/2\\ 0 & 1\end{pmatrix}$	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$

Table 5.1: Basic fractal matrices for the equivalent representations $(v, a, b, I_{\mathrm{B}})$.

	a	b	I_{B}	I_{Sh}
a	$\begin{cases} c_0 = 1 \\ c_{k>0} = 0 \end{cases}$	$c_k = (-1)^k$	$c_k = (-\frac{1}{2})^k$	$c_k = B_k$
b	$c_k = 1$	$\begin{cases} c_0 = 1 \\ c_{k>0} = 0 \end{cases}$	$c_k = (\frac{1}{2})^k$	$c_k = (-1)^k B_k$
I_{B}	$c_k = (\frac{1}{2})^k$	$c_k = (-\frac{1}{2})^k$	$\begin{cases} c_0 = 1 \\ c_{k>0} = 0 \end{cases}$	$c_k = (\frac{1}{2^{k-1}} - 1)B_k$
I_{Sh}	$c_k = \frac{1}{k+1}$	$c_k = \frac{(-1)^k}{k+1}$	$c_k = \frac{1+(-1)^k}{(k+1)2^{k+1}}$	$\begin{cases} c_0 = 1 \\ c_{k>0} = 0 \end{cases}$

Table 5.2: Cardinality sequences for the equivalent representations $(a, b, I_{\mathrm{B}}, I_{\mathrm{Sh}})$.

which can also be written under the form

$$I_{\mathrm{Sh}}(S) = \sum_{T\subseteq N} \frac{(-1)^{|S\setminus T|}}{(n-s+1)\binom{n-s}{|T\setminus S|}}\, v(T), \quad S \subseteq N. \tag{5.57}$$

With matricial notations, this identity is written:

$$I_{S\,(n)} = H_{(n)} \circ a_{(n)} = H_{(n)} \circ M_{(n)} \circ v_{(n)} \tag{5.58}$$

where $H_{(n)}$ is an upper-cardinality matrix based on the sequence $h_k = \frac{1}{k+1}$, and $M_{(n)}$ is the fractal matrix generated by (5.55).

The inverse formula can be found in [37, 83]: for all $S \subseteq N$, we have, using adequate correspondance formulas,

$$\begin{aligned} v(S) &= \sum_{K\subseteq S} a(K) = \sum_{K\subseteq S}\sum_{T\supseteq K} B_{t-k}\, I_{\mathrm{Sh}}(T) = \sum_{T\subseteq N} I_{\mathrm{Sh}}(T) \sum_{K\subseteq T\cap S} B_{t-k} \\ &= \sum_{T\subseteq N} I_{\mathrm{Sh}}(T) \sum_{k=0}^{|T\cap S|} \binom{|T\cap S|}{k} B_{t-k}, \end{aligned}$$

that is

$$v(S) = \sum_{T\subseteq N} \beta^{|T|}_{|T\cap S|}\, I_{\mathrm{Sh}}(T), \quad S \subseteq N, \tag{5.59}$$

with

$$\beta^l_k := \sum_{j=0}^{k} \binom{k}{j} B_{l-j}. \tag{5.60}$$

The first values of β^l_k are:

$k\backslash l$	0	1	2	3	4
0	1	$-1/2$	$1/6$	0	$-1/30$
1		$1/2$	$-1/3$	$1/6$	$-1/30$
2			$1/6$	$-1/6$	$2/15$
3				0	$-1/30$
4					$-1/30$

Some properties of the β^l_k are shown in [37, 83]. Note that the inverse formula (5.59) corresponds to

$$v_{(n)} = M^{-1}_{(n)} \circ H^{-1}_{(n)} \circ I_{S\,(n)}.$$

Although these transformations between I_{Sh} and v are neither fractal nor cardinal, their associated matrices have nevertheless a remarkable structure, and we call them *Pascal matrices*:

(i) a *direct Pascal matrix* P based on a sequence of real numbers $(p_0, p_1, \ldots, p_k, \ldots, p_n)$, such that:

$$\begin{aligned} P_{(1)} &:= \begin{pmatrix} p_0 & p_1 \\ p_0 & p_0+p_1 \end{pmatrix}, \quad P^l_{(1)} := \begin{pmatrix} p_{l-1} & p_l \\ p_{l-1} & p_{l-1}+p_l \end{pmatrix},\ l = 1,\ldots,n \\ P_{(2)} &:= \begin{pmatrix} P^1_{(1)} & P^2_{(1)} \\ P^1_{(1)} & P^1_{(1)}+P^2_{(1)} \end{pmatrix}, \quad P^l_{(2)} := \begin{pmatrix} P^l_{(1)} & P^{l+1}_{(1)} \\ P^l_{(1)} & P^l_{(1)}+P^{l+1}_{(1)} \end{pmatrix},\ l = 1,\ldots,n-1 \\ P_{(k)} &:= \begin{pmatrix} P^1_{(k-1)} & P^2_{(k-1)} \\ P^1_{(k-1)} & P^1_{(k-1)}+P^2_{(k-1)} \end{pmatrix},\ k = 2,\ldots,n. \end{aligned}$$

(ii) an *inverse Pascal matrix* Q based on a sequence of real numbers $(q_0, q_1, \ldots, q_k, \ldots, q_n)$, such that:

$$
\begin{aligned}
Q_{(1)} &:= \begin{pmatrix} q_0 - q_1 & q_1 \\ -q_0 & q_0 \end{pmatrix}, \quad Q^l_{(1)} := \begin{pmatrix} q_{l-1} - q_l & q_l \\ -q_{l-1} & q_{l-1} \end{pmatrix}, \ l = 1, \ldots, n \\
Q_{(2)} &:= \begin{pmatrix} Q^1_{(1)} - Q^2_{(1)} & Q^2_{(1)} \\ -Q^1_{(1)} & Q^1_{(1)} \end{pmatrix}, \quad Q^l_{(2)} := \begin{pmatrix} Q^l_{(1)} - Q^{l+1}_{(1)} & Q^{l+1}_{(1)} \\ -Q^l_{(1)} & Q^l_{(1)} \end{pmatrix}, \ l = 1, \ldots, n-1 \\
Q_{(k)} &:= \begin{pmatrix} Q^1_{(k-1)} - Q^2_{(k-1)} & Q^2_{(k-1)} \\ -Q^1_{(k-1)} & Q^1_{(k-1)} \end{pmatrix}, \ k = 2, \ldots, n.
\end{aligned}
$$

The name 'Pascal matrix' comes from the fact that, as in the Pascal triangle, elements are obtained by the sum of two preceding elements. Direct Pascal matrices are constructed from the upper left-hand corner, while inverse Pascal matrices start from the lower right-hand corner. An example of each kind is shown below ($n = 2$), where for $P_{(2)}$ the sequence of Bernoulli numbers have been chosen (thus retrieving the β^l_k coefficients and all their properties shown in [37, 83]), and for $Q_{(2)}$ the sequence $h_k = \frac{1}{k+1}$, $k = 0, 1, 2$, defined above (see (5.58)) (thus retrieving the coefficients of (5.57)):

$$
\begin{aligned}
P_{(2)} &= M^{-1}_{(2)} \circ H^{-1}_{(2)} = \begin{pmatrix} 1 & -1/2 & -1/2 & 1/6 \\ 1 & 1/2 & -1/2 & -1/3 \\ 1 & -1/2 & 1/2 & -1/3 \\ 1 & 1/2 & 1/2 & 1/6 \end{pmatrix} \\
Q_{(2)} &= H_{(2)} \circ M_{(2)} = \begin{pmatrix} 1/3 & 1/6 & 1/6 & 1/3 \\ -1/2 & 1/2 & -1/2 & 1/2 \\ -1/2 & -1/2 & 1/2 & 1/2 \\ 1 & -1 & -1 & 1 \end{pmatrix}
\end{aligned}
$$

Any Pascal matrix can be written as the product of an upper-cardinal matrix and either the Möbius matrix M or its inverse:

$$
\begin{aligned}
P_{(n),\, p_0,\ldots,p_n} &= M^{-1}_{(n)} \circ C_{(n),\, p_0,\ldots,p_n} \\
Q_{(n),\, q_0,\ldots,q_n} &= C_{(n),\, q_0,\ldots,q_n} \circ M_{(n)}
\end{aligned}
$$

(the generating sequence is written in subscript), as it can be easily verified. The set of (direct or inverse) Pascal matrices does not form a group since the product of two such matrices is no more a Pascal matrix. However, since the inverse of an upper-cardinality transformation is again upper-cardinal, it follows that the inverse of a direct (resp. inverse) Pascal matrix is an inverse (resp. direct) Pascal matrix.

5.3.7 Explicit transformation formulas

We now give all the conversion formulas between the five representations v, a, b, I_{B}, I_{Sh} of a set function v. All these representations are linear, that is, such that the transform $\mathcal{T}$ is a linear operator which can be written under a matrix form. The explicit transformation formulas are gathered in Tables 5.3 and 5.4.

	v	a	b
$v(S) =$	$v(S)$	$\sum_{T\subseteq S} a(T)$	$\sum_{T\subseteq N\setminus S} (-1)^t b(T)$
$a(S) =$	$\sum_{T\subseteq S} (-1)^{s-t} v(T)$	$a(S)$	$\sum_{T\supseteq S} (-1)^{t-s} b(T)$
$b(S) =$	$\sum_{T\supseteq N\setminus S} (-1)^{n-t} v(T)$	$\sum_{T\supseteq S} a(T)$	$b(S)$
$I_{\mathrm{B}}(S) =$	$(\frac{1}{2})^{n-s} \sum_{T\subseteq N} (-1)^{\lvert S\setminus T\rvert} v(T)$	$\sum_{T\supseteq S} (\frac{1}{2})^{t-s} a(T)$	$\sum_{T\supseteq S} (-\frac{1}{2})^{t-s} b(T)$
$I_{\mathrm{Sh}}(S) =$	$\sum_{T\subseteq N} \frac{(-1)^{\lvert S\setminus T\rvert}}{(n-s+1)\binom{n-s}{\lvert T\setminus S\rvert}} v(T)$	$\sum_{T\supseteq S} \frac{1}{t-s+1} a(T)$	$\sum_{T\supseteq S} \frac{(-1)^{t-s}}{t-s+1} b(T)$

Table 5.3: Alternative representations in terms of v, a, b

	I_{B}	I_{Sh}
$v(S) =$	$\sum_{T\subseteq N} (\frac{1}{2})^t (-1)^{\lvert T\setminus S\rvert} I_{\mathrm{B}}(T)$	$\sum_{T\subseteq N} \Big[\sum_{k=0}^{\lvert T\cap S\rvert} \binom{\lvert T\cap S\rvert}{k} B_{t-k} \Big] I_{\mathrm{Sh}}(T)$
$a(S) =$	$\sum_{T\supseteq S} (-\frac{1}{2})^{t-s} I_{\mathrm{B}}(T)$	$\sum_{T\supseteq S} B_{t-s}\, I_{\mathrm{Sh}}(T)$
$b(S) =$	$\sum_{T\supseteq S} (\frac{1}{2})^{t-s} I_{\mathrm{B}}(T)$	$\sum_{T\supseteq S} (-1)^{t-s} B_{t-s}\, I_{\mathrm{Sh}}(T)$
$I_{\mathrm{B}}(S) =$	$I_{\mathrm{B}}(S)$	$\sum_{T\supseteq S} (\frac{1}{2^{t-s-1}} - 1) B_{t-s}\, I_{\mathrm{Sh}}(T)$
$I_{\mathrm{Sh}}(S) =$	$\sum_{T\supseteq S} \frac{1+(-1)^{t-s}}{(t-s+1)\, 2^{t-s+1}} I_{\mathrm{B}}(T)$	$I_{\mathrm{Sh}}(S)$

Table 5.4: Alternative representations in terms of I_{B} and I_{Sh}

5.4 The chaining interaction index

The purpose of this section is to introduce a new interaction index belonging to the family of cardinal-probabilistic interaction indices: the *chaining interaction index* I_{R}^v, for which one has

$$p_t^s(n) = s\,\frac{(n-s-t)!\,(s+t-1)!}{n!}. \tag{5.61}$$

We can already notice that $I_{\mathrm{R}}^v(i) = I_{\mathrm{Sh}}^v(i)$ for all $i \in N$. Thus $I_{\mathrm{R}}^v(S)$ can be considered as an extension of the Shapley value to determine the interaction between the players of the coalition $S \subseteq N$.

The results presented in this section have been proved by Marichal and Roubens [122].

5.4.1 Definition

Let us consider the lattice $\mathcal{L}(N)$ related to the power set of N. We can represent $\mathcal{L}(N)$ as a graph called Hasse Diagram $H(N)$ whose nodes correspond to the coalitions $S \subseteq N$ and the edges represent adding a player to the bottom coalition to get the top coalition. A *maximal chain* of $H(N)$ is an ordered collection of $n+1$ nested distinct coalitions

$$\mathcal{M} = (\emptyset = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_{n-1} \subsetneq M_n = N).$$

The set of maximal chains of $H(N)$ is denoted $C(N)$. Let $\mathcal{M}$ be an element of $C(N)$ and M^S the minimal coalition belonging to $\mathcal{M}$ that contains S. The cardinality of $C(N)$ is equal to $n!$ and we define

$$I_{\mathrm{R}}^v(S) = \frac{1}{n!}\sum_{\mathcal{M}\subseteq C(N)} \delta_S v(M^S), \quad \emptyset \neq S \subseteq N.$$

The value $I_{\mathrm{R}}^v(i)$ corresponds to the Shapley value related to i as it was mentioned by Edelman [58] dealing with cooperative games in which only certain coalitions are allowable.

We now prove that I_{R}^v is a cardinal-probabilistic interaction

$$I^v(S) = \sum_{T\subseteq N\setminus S} p_t^s(n)\,\delta_S v(T\cup S) \qquad (p_t^s(n) \geq 0, \ \sum_{T\subseteq N\setminus S} p_t^s(n) = 1),$$

for which $p_t^s(n)$ is defined by (5.61).

If $C^{S,S\cup T}$ represents the subclass included in $C(N)$ of maximal chains that have $\{S\cup T\}$ as minimal coalition including S, we have

$$\begin{aligned} I_{\mathrm{R}}^v(S) &= \frac{1}{n!}\sum_{T\subseteq N\setminus S} \left|C^{S,S\cup T}\right| \delta_S v(S\cup T) \\ &= \sum_{T\subseteq N\setminus S} p_t^s(n)\,\delta_S v(T\cup S) \end{aligned}$$

with $p_t^s(n) = \frac{1}{n!}\left|C^{S,S\cup T}\right|$, $s = 1,\ldots,n$, $t = 0,\ldots,n-s$ (notice that $\left|C^{S,S\cup T}\right|$ only depends on s and t).

For example, if $N = \{1,2,3\}$, we have

$$\begin{aligned} C^{\{1\},\{1\}} &= \{(\emptyset \subsetneq \{1\} \subsetneq \{1,2\} \subsetneq \{1,2,3\}), (\emptyset \subsetneq \{1\} \subsetneq \{1,3\} \subsetneq \{1,2,3\})\} \\ C^{\{1\},\{1,2\}} &= \{(\emptyset \subsetneq \{2\} \subsetneq \{1,2\} \subsetneq \{1,2,3\})\} \\ C^{\{1\},\{1,3\}} &= \{(\emptyset \subsetneq \{3\} \subsetneq \{1,3\} \subsetneq \{1,2,3\})\} \\ C^{\{1\},\{1,2,3\}} &= \{(\emptyset \subsetneq \{2\} \subsetneq \{2,3\} \subsetneq \{1,2,3\}), (\emptyset \subsetneq \{3\} \subsetneq \{2,3\} \subsetneq \{1,2,3\})\} \end{aligned}$$

$$p_0^1(3) = \frac{2}{6}, \quad p_1^1(3) = \frac{1}{6}, \quad p_2^1(3) = \frac{2}{6}$$

$$\begin{aligned} C^{\{1,2\},\{1,2\}} &= \{(\emptyset \subsetneq \{1\} \subsetneq \{1,2\} \subsetneq \{1,2,3\}), (\emptyset \subsetneq \{2\} \subsetneq \{1,2\} \subsetneq \{1,2,3\})\} \\ C^{\{1,2\},\{1,2,3\}} &= \{(\emptyset \subsetneq \{1\} \subsetneq \{1,3\} \subsetneq \{1,2,3\}), (\emptyset \subsetneq \{2\} \subsetneq \{2,3\} \subsetneq \{1,2,3\}), \\ &\quad (\emptyset \subsetneq \{3\} \subsetneq \{1,3\} \subsetneq \{1,2,3\}), (\emptyset \subsetneq \{3\} \subsetneq \{2,3\} \subsetneq \{1,2,3\})\}. \end{aligned}$$

It is easy to observe that $C^{S,S\cup T}$ corresponds to a disjoint union of sets of maximal chains defined in sublattices of $\mathcal{L}(N)$. In particular, we can see that

$$\left|C^{S,S\cup T}\right| = \left|C(N \setminus (S \cup T))\right| \times \left|\bigcup_{i\in S} C((S \setminus i) \cup T)\right|.$$

We then have, for all $s = 1, \ldots, n$ and all $t = 0, \ldots, n-s$,

$$\begin{aligned} p_t^s(n) &= \frac{1}{n!}\left|C^{S,S\cup T}\right| = \frac{s}{n!}\left|C(N \setminus (S \cup T))\right| \times \left|C((S \setminus 1) \cup T)\right| \\ &= s\,\frac{(n-s-t)!\,(s+t-1)!}{n!}. \end{aligned}$$

The chaining interaction index I_{R}^v is of cardinal-probabilistic type since, if $S \neq \emptyset$,

$$\begin{aligned} \sum_{T\subseteq N\setminus S} p_t^s(n) &= \sum_{t=0}^{n-s} \binom{n-s}{t} s\,\frac{(n-s-t)!\,(s+t-1)!}{n!} \\ &= \frac{s!\,(n-s)!}{n!} \sum_{t=0}^{n-s} \binom{s+t-1}{s-1} = 1. \end{aligned}$$

5.4.2 Some equivalent representations

It has been proved in Section 5.3 that I_{Sh}^v and I_{B}^v are representations of v. We prove here that I_{R}^v also is a representation of v. More precisely, we prove the following two identities.

$$I_{\mathrm{R}}(S) = \sum_{T\supseteq S} \frac{s}{t}\, a(T), \quad \emptyset \neq S \subseteq N \tag{5.62}$$

$$a(S) = \sum_{T\supseteq S} (-1)^{t-s}\, \frac{s}{t}\, I_{\mathrm{R}}(T), \quad \emptyset \neq S \subseteq N. \tag{5.63}$$

On the one hand, we have (see (5.27))

$$\sum_{T\subseteq N\setminus S} p_t^s(n)\,\delta_S v(T\cup S) = \sum_{T\supseteq S} \xi_t^s(n)\,a(T)$$

with

$$\xi_t^s(n) = \sum_{k=0}^{n-t} \binom{n-t}{k} p_{k+t-s}^s(n), \quad t = s, \ldots, n.$$

When $p_t^s(n)$ is given by (5.61), the previous identity becomes

$$\xi_t^s(n) = \frac{s\,(n-t)!\,(t-1)!}{n!} \sum_{k=0}^{n-t} \binom{k+t-1}{t-1} = \frac{s}{t}, \quad t = s, \ldots, n,$$

which proves (5.62). On the other hand, we have, by (5.62),

$$\begin{aligned}
\sum_{T\supseteq S}(-1)^{t-s}\,\frac{s}{t}\,I_{\mathrm{R}}(T) &= \sum_{T\supseteq S}(-1)^{t-s}\,\frac{s}{t}\,\sum_{K\supseteq T}\frac{t}{k}\,a(K)\\
&= \sum_{K\supseteq S}\frac{s}{k}\Big[\sum_{T:S\subseteq T\subseteq K}(-1)^{t-s}\Big]\,a(K)\\
&= \sum_{K\supseteq S}\frac{s}{k}\Big[\sum_{t=s}^{k}\binom{k-s}{t-s}(-1)^{t-s}\Big]\,a(K)\\
&= \sum_{K\supseteq S}\frac{s}{k}\,(1-1)^{k-s}\,a(K)=a(S),
\end{aligned}$$

which proves (5.63).

It is interesting to note how simple formula (5.62) is. Moreover, comparing I_{R}^v and I_{Sh}^v one can see that the terms of the summation are weighted by elements which are linearly decreasing with the argument $t=|T|$, whereas these elements are exponentially decreasing for I_{B}^v.

The conversion formulas between I_{R} and v can be given as follows:

$$\begin{aligned}
I_{\mathrm{R}}(S) &= \sum_{T\subseteq N}\frac{s\,(-1)^{|S\setminus T|}}{n\binom{n-1}{|S\cup T|-1}}\,v(T),\quad \emptyset\neq S\subseteq N, &(5.64)\\
v(S) &= \sum_{T\subseteq N\setminus S}\frac{(-1)^t}{t+1}\sum_{i\in S}I_{\mathrm{R}}(T\cup i),\quad \emptyset\neq S\subseteq N. &(5.65)
\end{aligned}$$

Indeed, using (5.13), we obtain, by setting $T' := T\cup L$ (which implies $L=T'\cap S$ and $T=T'\setminus S$):

$$\begin{aligned}
I_{\mathrm{R}}(S) &= \sum_{T\subseteq N\setminus S} s\frac{(n-s-t)!\,(s+t-1)!}{n!}\sum_{L\subseteq S}(-1)^{s-l}v(L\cup T)\\
&= \sum_{T'\subseteq N}\frac{s}{n}\,\frac{(n-|S\cup T'|)!\,(|S\cup T'|-1)!}{(n-1)!}\,(-1)^{|S\setminus T'|}v(T')
\end{aligned}$$

which proves (5.64). On the other hand, we have, by (4.5) and (5.63),

$$\begin{aligned}
v(S) &= \sum_{T\subseteq S}\sum_{K\supseteq T}(-1)^{k-t}\,\frac{t}{k}\,I_{\mathrm{R}}(K)\\
&= \sum_{K\subseteq N}\frac{1}{k}\,I_{\mathrm{R}}(K)\sum_{T\subseteq K\cap S}(-1)^{k-t}t\\
&= \sum_{K\subseteq N}\frac{1}{k}\,I_{\mathrm{R}}(K)\underbrace{\sum_{t=0}^{|K\cap S|}\binom{|K\cap S|}{t}(-1)^{k-t}t}_{(*)}
\end{aligned}$$

where the second sum $(*)$ equals $(-1)^{k+1}$ if $|K\cap S|=1$ and 0 otherwise. Therefore

$$v(S)=\sum_{\substack{K\subseteq N\\ |K\cap S|=1}}\frac{(-1)^{k+1}}{k}\,I_{\mathrm{R}}(K)$$

which proves (5.65).

The conversion formulas between I_{R} and I_{Sh} can also be given. We have

$$I_{\mathrm{R}}(S) = \sum_{T\supseteq S} \gamma_t^s\, I_{\mathrm{Sh}}(T), \quad \emptyset \neq S \subseteq N, \tag{5.66}$$

$$I_{\mathrm{Sh}}(S) = I_{\mathrm{R}}(S) + \sum_{\substack{T\supset S\\ T\neq S}} (-1)^{t-s} \frac{s-1}{t\,(t-s+1)}\, I_{\mathrm{R}}(T), \quad \emptyset \neq S \subseteq N, \tag{5.67}$$

with

$$\gamma_t^s = \int_0^1 s\, x^{s-1}\, B_{t-s}(x)\, dx = \sum_{k=0}^{t-s} \binom{t-s}{k} \frac{s}{t-k}\, B_k, \quad t = s, \dots, n,$$

where $\{B_n\}_{n\in\mathbb{N}}$ is the sequence of Bernoulli numbers (5.43) and $B_n(x)$ is the n-th Bernoulli polynomial, see Section 5.3.4.

The first values of γ_t^s are:

$s\backslash t$	1	2	3	4	5	6	7
1	1	0	0	0	0	0	0
2		1	1/6	0	−1/60	0	1/126
3			1	1/4	1/60	−1/40	−1/210
4				1	3/10	1/30	−1/35
5					1	1/3	1/21
6						1	5/14
7							1

Let us prove (5.66). It has been established in Section 5.3 that (see Table 5.4):

$$a(S) = \sum_{T\supseteq S} B_{t-s}\, I_{\mathrm{Sh}}(T), \quad S \subseteq N.$$

Therefore, by (5.62),

$$I_{\mathrm{R}}(S) = \sum_{K\supseteq S} \frac{s}{k} \sum_{T\supseteq K} B_{t-k}\, I_{\mathrm{Sh}}(T) = \sum_{T\supseteq S} \gamma_t^s\, I_{\mathrm{Sh}}(T)$$

with

$$\begin{aligned}
\gamma_t^s &= \sum_{K:S\subseteq K\subseteq T} \frac{s}{k}\, B_{t-k} = \sum_{k=s}^{t} \binom{t-s}{k-s} \frac{s}{k}\, B_{t-k}\\
&= \sum_{k=0}^{t-s} \binom{t-s}{k} \frac{s}{k+s}\, B_{t-s-k} = \sum_{k=0}^{t-s} \binom{t-s}{k} \frac{s}{t-k} B_k\\
&= \int_0^1 s \sum_{k=0}^{t-s} \binom{t-s}{k} x^{t-k-1} B_k\, dx\\
&= \int_0^1 s\, x^{s-1}\, B_{t-s}(x)\, dx.
\end{aligned}$$

Let us prove (5.67). It has been established in Section 5.3 that (see Table 5.3):

$$I_{\mathrm{Sh}}(S) = \sum_{T\supseteq S} \frac{1}{t-s+1}\, a(T), \quad S \subseteq N.$$

Therefore, by (5.63),

$$\begin{aligned} I_{\mathrm{Sh}}(S) &= \sum_{K\supseteq S} \frac{1}{k-s+1} \sum_{T\supseteq K} (-1)^{t-k}\,\frac{k}{t}\, I_{\mathrm{R}}(T) \\ &= \sum_{T\supseteq S} \frac{(-1)^t}{t} \Big[\sum_{K:S\subseteq K\subseteq T} (-1)^k \,\frac{k}{k-s+1}\Big]\, I_{\mathrm{R}}(T) \end{aligned}$$

and

$$\begin{aligned} \sum_{K:S\subseteq K\subseteq T} (-1)^k\,\frac{k}{k-s+1} &= \sum_{k=s}^{t} \binom{t-s}{k-s} (-1)^k\,\frac{k}{k-s+1} \\ &= (-1)^s \sum_{k=0}^{t-s} \binom{t-s}{k} (-1)^k\,[1+\frac{s-1}{k+1}] \\ &= (-1)^s \Big[(1-1)^{t-s} + (s-1)\int_0^1 \sum_{k=0}^{t-s}\binom{t-s}{k}(-x)^k\,dx\Big] \\ &= (-1)^s \Big[(1-1)^{t-s} + (s-1)\int_0^1 (1-x)^{t-s}\,dx\Big] \\ &= (-1)^s \Big[(1-1)^{t-s} + \frac{s-1}{t-s+1}\Big]. \end{aligned}$$

Hence the result.

Using similar arguments, we can also prove that

$$I_{\mathrm{B}}(S) = \sum_{T\supseteq S} \Big(-\frac{1}{2}\Big)^{t-s}\,\frac{2s-t}{t}\, I_{\mathrm{R}}(T), \quad \emptyset \neq S \subseteq N;$$

indeed, we simply have

$$\begin{aligned} I_{\mathrm{B}}(S) &= \sum_{K\supseteq S} \Big(\frac{1}{2}\Big)^{k-s} a(K) \\ &= \sum_{K\supseteq S} \Big(\frac{1}{2}\Big)^{k-s} \sum_{T\supseteq K} (-1)^{t-k}\,\frac{k}{t}\, I_{\mathrm{R}}(T) \\ &= \sum_{T\supseteq S} \frac{(-1)^t}{t} \Big[\sum_{K:S\subseteq K\subseteq T} (-1)^k \frac{k}{2^{k-s}}\Big]\, I_{\mathrm{R}}(T) \end{aligned}$$

and

$$\begin{aligned} \sum_{K:S\subseteq K\subseteq T} (-1)^k \frac{k}{2^{k-s}} &= \sum_{k=s}^{t} \binom{t-s}{k-s} (-1)^k \frac{k}{2^{k-s}} \\ &= (-1)^s \sum_{k=0}^{t-s} \binom{t-s}{k} (-1)^k\,\frac{k+s}{2^k} \\ &= (-1)^s\, 2^{s-t} (2s-t). \end{aligned}$$

5.4.3 Links with multilinear extensions and potentials

Let g be the multilinear extension of v (MLE). It has been proved in Section 5.3.2 that

$$\begin{aligned} I^v_{\mathrm{Sh}}(S) &= \int_0^1 (\Delta_S g)(\underline{x})\,dx, \quad S\subseteq N, \\ I^v_{\mathrm{B}}(S) &= (\Delta_S g)(\underline{1/2}), \quad S\subseteq N. \end{aligned}$$

There is also a close link between $\Delta_S g$ and I_{R}. We can easily prove that

$$I_{\mathrm{R}}^v(S) = \int_0^1 s\, x^{s-1}\, (\Delta_S g)(\underline{x})\, dx, \quad \emptyset \neq S \subseteq N; \tag{5.68}$$

indeed, it has been proved in Section 5.3.4 that

$$(\Delta_S g)(\underline{x}) = \sum_{T \supseteq S} x^{t-s}\, a(T),$$

and hence we have

$$\begin{aligned} I_{\mathrm{R}}^v(S) = \sum_{T \supseteq S} \frac{s}{t}\, a(T) &= \int_0^1 \sum_{T \supseteq S} s\, x^{t-1}\, a(T)\, dx \\ &= \int_0^1 s\, x^{s-1} \sum_{T \supseteq S} x^{t-s}\, a(T)\, dx \\ &= \int_0^1 s\, x^{s-1}\, (\Delta_S g)(\underline{x})\, dx. \end{aligned}$$

We also have

$$I_{\mathrm{R}}(S) = \sum_{T \supseteq S} \Big[\int_0^1 s\, x^{s-1}\, (x - \frac{1}{2})^{t-s}\, dx\Big]\, I_{\mathrm{B}}(T), \quad \emptyset \neq S \subseteq N;$$

indeed, it has been shown in Section 5.3.4 that

$$(\Delta_S g)(\underline{x}) = \sum_{T \supseteq S} (x - \frac{1}{2})^{t-s}\, I_{\mathrm{B}}(T),$$

and hence we can conclude by (5.68).

It is worth noting that, by using the same argument, formula (5.66) can be retrieved by means of equation (5.50).

The chaining interaction index I_{R}^v can be expressed easily in term of *potential*. If P^v is a functional that represents the potential for the Shapley value (see Hart and Mas-Colell [99, 100]) defined as $P^v(\emptyset) = 0$ and $\sum_{i \in N} \delta_i P^v(N) = v(N)$, then we have

$$P^v(N) = \sum_{T \subseteq N} \frac{1}{t}\, a(T),$$

and

$$\delta_i P^v(N) = P^v(N) - P^v(N \setminus i) = \phi_{\mathrm{Sh}}^v(i), \quad i \in N.$$

More generally, one can easily see that

$$s\, \delta_S P^v(N) = I_{\mathrm{R}}(S), \quad \emptyset \neq S \subseteq N.$$

Likewise, if Q^v represents the potential for the Banzhaf value (see Dragan [39]) defined as $Q^v(\emptyset) = 0$ and $\sum_{i \in N} \delta_i Q^v(N) = \sum_{i \in N} \phi_{\mathrm{B}}^v(i)$, then we have

$$Q^v(N) = \sum_{T \subseteq N} \frac{1}{2^{t-1}}\, a(T),$$

and

$$\delta_i Q^v(N) = \phi_{\mathrm{B}}^v(i), \quad i \in N.$$

More generally, we can readily verify that

$$2^{s-1}\, \delta_S Q^v(N) = I_{\mathrm{B}}(S), \quad \emptyset \neq S \subseteq N.$$

Notes

1. The concept of interaction indices was essentially introduced and characterized by Grabisch and Roubens, see [81, 91, 92, 93, 156].

2. In Section 5.3, we have shown that these interaction indices are equivalent representations of the set function v. For this purpose, we have introduced several tools and we have proved technical results based on pseudo-Boolean functions as well as their multilinear extensions (MLE). We have also introduced the concept of fractal and cardinality transformations to derive correspondence formulas between the different representations. In particular, we have shown that all the transformations are linear.

3. We have introduced a new type of interaction index: the chaining interaction index. Correspondence formulas have also been presented by means of the multilinear extensions.

Chapter 6

Applications to multicriteria decision making

In this chapter we study the aggregation problem in the presence of interacting criteria. We represent the process of aggregation by means of a fuzzy measure on the set of criteria. Thus, all the aggregation operators we consider depend on a fuzzy measure.

When scores are given on the same cardinal scale, we suggest using the Choquet integral as an aggregation tool, unless specific properties are required. This suggestion is actually based on a characterization of the Choquet integral involving only natural properties.

The interaction indices presented in Chapter 5 are then used to express the dependence between criteria. On this matter, it appears that the Shapley indices are much more suitable than the Banzhaf indices. Other indices such as veto and favor indices or degree of orness are also introduced to have a better understanding of the behavioral properties of the Choquet integral. When the fuzzy measure is not completely known, such indices can help the decision maker to assess it. This corresponds to the inverse problem of identifying the weights from parametric specifications on criteria. Since the meaning of these weights is not very clear for subsets of at least three criteria, it is suggested to use the concept of k-order fuzzy measure.

We also examine the case where scores are given on the same ordinal scale. When there is commensurability between the scores and the fuzzy measure, the Sugeno integral is pointed out through a natural characterization.

This chapter is organized as follows. In Section 6.1 we examine the case of cardinal scales. Two aggregation operators are compared: the Shapley integral and the Choquet integral. We give a justification for the use of the latter and we abandon the former. In Section 6.2 we present several indices related to the Choquet integral or its underlying fuzzy measure, namely degree of tolerance, veto and favor degrees, and dispersion measure. These indices as well as the power and interaction indices form a kind of identity card of the fuzzy measure. In Section 6.3, we introduce the concept of k-order fuzzy measure and show the usefulness of confining oneself to 2-order measures. Section 6.4 is devoted to the inverse problem of determining the fuzzy measure. We propose assessing a 2-order fuzzy measure from two types of information: parametric specifications based on importance and interaction indices, and learning data based on a predefined set of prototypes whose profiles are known. Some illustrative examples are also presented. Finally, Section 6.5 deals with the Sugeno integral and the ordinal scales.

6.1 Aggregation of scores defined on cardinal scales

We have presented in Section 1.3.1 the cardinal setting of the aggregation of multiple criteria. In the present section, we analyze in more details some operators that can be suitable for the aggregation phase, and we justify the use of the Choquet integral.

6.1.1 The commensurability assumption

In many practical applications, the decision criteria are defined on independent measurement scales $E_i \subseteq \mathbb{R}$, as in Example 1.3.1. However, aggregating values defined on independent scales has good chances to lead to a dictatorial aggregation. For instance, when the E_i are independent interval scales, the only allowed aggregation operators are of the form

$$M(x) = a\,x_j + b, \quad \text{for some } j \in N,$$

(see Theorem 3.4.8), and these operators are strongly equivalent to the dictator x_j (see Theorem 1.3.1). Therefore, it seems necessary to express all the criteria on a same measurement scale. Actually, this is the role played by the utility functions u_i introduced in multiattribute utility theory.

Back to Example 1.3.1, it is clear that the consumer cannot compare directly consumption and comfort since they are defined on different scales. Nevertheless, he/she can affect, to both values, degrees of satisfaction which are comparable. According to the notation introduced in Section 1.3.1, this means that, if the consumer gives a degree of satisfaction $u_2(x_2^2)$ for the car 2 and $u_3(x_3^1)$ for the car 1 and $u_2(x_2^2) \geq u_3(x_3^1)$, he/she can say that he/she would rather have a car that consumes $9\,\ell/100$ km than a car with very good comfort (without prior knowledge of the other values of criteria).

Therefore, we shall assume that any score on a criterion can be compared with any other score on another criterion. This is the *commensurability assumption*. In particular, all the criteria are expressed on the same scale E. A more formal approach of commensurability can be found in [125].

6.1.2 The Shapley integral

Suppose that all the criteria are defined on a same ratio or interval scale E. Of course, we can assume without loss of generality that this common scale is $[0, 1]$.

The most common aggregation tool used in multicriteria decision making is the weighted arithmetic mean $\sum_i \omega_i\, x_i$ where ω_i represents the weight of importance of criterion i. However, we know that such an operator cannot express any interaction between criteria. So, it is better to consider fuzzy integrals fulfilling either (SSi) or (SPL), like the Choquet integral. But in order to keep the aggregation step as simple and intuitive as possible, we propose a new type of fuzzy integral: the Shapley integral.

Let us consider the additive fuzzy measure derived from the Shapley value, see (5.6):

$$p(S) = \sum_{i \in S} \phi^{\mu}_{\mathrm{Sh}}(i), \quad S \subseteq N.$$

We can define a new fuzzy integral as a natural extension of p.

Definition 6.1.1 Let μ be a fuzzy measure on N. The *Shapley integral* of a function $x : N \to [0,1]$ with respect to μ is defined by

$$\mathrm{Sh}_\mu(x) = \sum_{i\in N} \phi^\mu_{\mathrm{Sh}}(i)\, x_i.$$

Thus defined, the Shapley integral is a weighted arithmetic mean operator WAM_ω whose weights are the Shapley power indices: $\omega_i = \phi^\mu_{\mathrm{Sh}}(i)$ for all $i \in N$. Starting from any fuzzy measure, we can define the Shapley additive measure and aggregate by the corresponding weighted arithmetic mean.

Note that, contrary to the Choquet and Sugeno integrals, the Shapley integral w.r.t. the fuzzy measure μ is not an extension of μ (i.e. of the pseudo-Boolean function which defines μ); indeed, for any $S \subseteq N$, we generally have

$$\mathrm{Sh}_\mu(e_S) = \sum_{i\in S} \phi^\mu_{\mathrm{Sh}}(i) \neq \mu(S).$$

In terms of the Möbius representation, the Shapley integral has an interesting form.

Proposition 6.1.1 *Any Shapley integral* $\mathrm{Sh}_\mu : [0,1]^n \to [0,1]$ *can be written as*

$$\mathrm{Sh}_\mu(x) = \sum_{\substack{T\subseteq N\\ T\neq\emptyset}} a_T \Big(\frac{1}{t}\sum_{i\in T} x_i\Big), \quad x \in [0,1]^n, \tag{6.1}$$

where a is the Möbius representation of μ.

Proof. By (5.3), we simply have

$$\mathrm{Sh}_\mu(x) = \sum_{i\in N} \Big(\sum_{T\ni i} \frac{1}{t}\, a_T\Big)\, x_i.$$

Permuting the sums leads immediately to the result. ■

Although the Shapley integral takes into account the dependence between criteria expressed in the underlying fuzzy measure, it remains a uninteresting aggregation operator. Indeed, since the Shapley integral is nothing less than a weighted arithmetic mean, it is not suitable to aggregate criteria when mutual preferential independence is violated.

We show below that the Choquet integral is a rather natural aggregation operator, which can be used in many applications.

6.1.3 The Choquet integral revisited

In Section 4.2.3 we have characterized the Choquet integral by means of either (In, SPL, CoAdd) or (In, SPL, BOM) properties. However, both (CoAdd) and (BOM) seem to be unattractive properties in multicriteria decision making.

We now propose a new characterization of the Choquet integral, which involves rather natural properties. The first one is linearity with respect to the fuzzy measure.

Definition 6.1.2 (LM) An aggregation operator $M_\mu \in A_n(E, \mathbb{R})$ depending on a fuzzy measure μ on N is *linear w.r.t.* μ if the following holds:

For any $k \in \mathbb{N}_0$, if $\mu^1, \ldots, \mu^k$ and $\mu = \sum_{i=1}^k \alpha_i\, \mu^i$ ($\alpha_i \in \mathbb{R}$) are fuzzy measures on N then

$$M_\mu = \sum_{i=1}^k \alpha_i\, M_{\mu^i}.$$

Proposition 6.1.2 $M_\mu \in A_n(E, \mathbb{R})$ *fulfils (LM) if and only if there exist functions* $g_T : E^n \to \mathbb{R}, T \subseteq N$, *such that, for any fuzzy measure* μ *with Möbius representation* a, *we have*

$$M_\mu(x) = \sum_{T \subseteq N} a_T\, g_T(x), \quad x \in E^n.$$

Proof. (Sufficiency) Follows from the linearity of the expression of μ in terms of a.

(Necessity) For every $T \subseteq N$, consider the fuzzy measure $\mu^{(T)}$ defined by $\mu_S^{(T)} = 1$ if and only if $S \supseteq T$, and 0 otherwise (unanimity game for T). Let μ be a fuzzy measure on N. By (4.3), we have

$$\mu = \sum_{T \subseteq N} a_T\, \mu^{(T)},$$

and by Definition 6.1.2, we have

$$M_\mu(x) = \sum_{T \subseteq N} a_T\, M_{\mu^{(T)}}(x),$$

for all $x \in E^n$. ■

As the previous proposition shows, the (LM) property constitutes a natural step towards a possible generalization of the weighted arithmetic mean. Indeed, it allows to take into account in a very elementary way not only the weights of criteria, but also the interactions among them[1].

We can immediately see that the Choquet integral (4.11), the Shapley integral (6.1), and the MLE of a fuzzy measure (5.28) fulfil (LM). Moreover, the Choquet and Shapley integrals fulfil (SPL) whereas the MLE does not. Also, the Choquet integral is an extension of the associated fuzzy measure whereas the Shapley integral is not.

Definition 6.1.3 (Ext) $M_\mu \in A_n(E, \mathbb{R})$ is *an extension of* μ if $M_\mu(e_T) = \mu_T$ for all $T \subseteq N$.

Notice that (Ext) can be viewed as a proper definition of the importance of a subset of criteria: for any subset $T \subseteq N$, the weight μ_T is the global score obtained by an alternative having e_T as profile.

We thus see that the Choquet integral, as an operator depending on a fuzzy measure, fulfils (In, SPL, LM, Ext). The following result shows that these properties, which are natural enough, characterize the Choquet integral.

Theorem 6.1.1 *Assume that* $E \supseteq [0, 1]$. *An aggregation operator* $M_\mu \in A_n(E, \mathbb{R})$ *depending on a fuzzy measure* μ *fulfils (In, SPL, LM, Ext) if and only if* $M_\mu = \mathcal{C}_\mu$.

Proof. (Sufficiency) Trivial.

(Necessity) By Proposition 6.1.2, there exist functions $g_T : E^n \to \mathbb{R}$ with $T \subseteq N$, such that, for any fuzzy measure μ with Möbius representation a, we have

$$M_\mu(x) = \sum_{T \subseteq N} a_T\, g_T(x), \quad x \in E^n.$$

[1]Due to the linear representation of a in terms of I_{Sh}, we can see that the (LM) property is equivalent to the following assertion: there exist functions $h_T : E^n \to \mathbb{R}$ ($T \subseteq N$) such that, for any fuzzy measure μ with Shapley interaction representation I_{Sh}, we have

$$M_\mu(x) = \sum_{T \subseteq N} I_{\mathrm{Sh}}(T)\, h_T(x), \quad x \in E^n.$$

Let $\mu^{(T)}$ be the unanimity game for T. As in Proposition 6.1.2, we have

$$M_\mu(x) = \sum_{T\subseteq N} a_T\, M_{\mu^{(T)}}(x),$$

for all $x \in E^n$, i.e. $g_T = M_{\mu^{(T)}}$, and g_T fulfils (In, SPL) for all $T \subseteq N$.

Fix $T \subseteq N$. By (Ext), we have $g_T(e_S) = M_{\mu^{(T)}}(e_S) = \mu_S^{(T)} \in \{0,1\}$, and by Theorem 4.4.2, there exists a set function $c : 2^N \to \{0,1\}$ such that $g_T = \mathrm{B}_c^{\vee\wedge}$.

We then have, for all $S \subseteq N$,

$$c_S = g_T(e_S) = \mu_S^{(T)} = \begin{cases} 1, & \text{if } T \subseteq S, \\ 0, & \text{otherwise,} \end{cases}$$

and hence

$$g_T(x) = \bigvee_{S\supseteq T} \bigwedge_{i\in S} x_i = \bigwedge_{i\in T} x_i, \quad x \in E^n,$$

which proves the theorem. ■

Note that the MLE fulfils (LM, In, Ext) but not (SPL), and the Shapley integral fulfils (LM, In, SPL) but not (Ext).

Corollary 6.1.1 *Assume that $E \supseteq [0,1]$. An aggregation operator $M_\omega \in A_n(E,\mathbb{R})$ depending on an additive measure ω fulfils (In, SPL, LM, Ext) if and only if $M_\omega = \mathrm{WAM}_\omega$.*

Corollary 6.1.2 *Assume that $E \supseteq [0,1]$. An aggregation operator $M_\mu \in A_n(E,\mathbb{R})$ depending on a fuzzy measure μ fulfils (Sy, In, SPL, LM, Ext) if and only if there exists $\omega \in [0,1]^n$ such that $M_\mu = \mathrm{OWA}_\omega$.*

Theorem 6.1.1 shows that the Choquet integral seems to be a suitable operator for aggregation of interacting criteria. Of course, it would be interesting to have a characterization of the Choquet integral as a utility function, that is defined up to an increasing bijection (see Section 1.3.3). Note that an attempt along this line was done in [124, 125], using the formal parallelism between multicriteria decision making and decision under uncertainty.

It is worth noting that, contrary to the Shapley integral, the Choquet integral is able to perform aggregation of criteria, even when mutual preferential independence is violated. We will see this in Example 6.4.1. Moreover, Murofushi and Sugeno [134] have proved a fundamental result relating preferential independence and additivity of the fuzzy measure. To present this result, we need a definition.

Definition 6.1.4 An attribute $i \in N$ is called essential if there exist $x_i, y_i \in E_i$ and $x_{N\setminus i} \in E_{N\setminus i}$ such that

$$(x_i, x_{N\setminus i}) \succ (y_i, x_{N\setminus i}).$$

Theorem 6.1.2 *Consider a multicriteria decision making problem and assume that there exists a fuzzy measure μ on N such that the utility function u is given by the Choquet integral:*

$$u(x_1, \ldots, x_n) = \mathcal{C}_\mu(u_1(x_1), \ldots, u_n(x_n))$$

where the u_i's are uni-dimensional utilities. If there are at least three essential attributes then the following assertions are equivalent:

i) The attributes are mutually preferentially independent.

ii) μ is additive.

6.2 Behavioral analysis of aggregation

Thus far we have focussed on mathematical properties of aggregation operators, and neglected somewhat the behavioral properties of these operators. Such properties can reflect the behavior of the decision maker in the aggregation phase. This is of course what (1.11) and (1.12) do, but more specifically, this behavior can be expressed through different concepts, such as the degree of tolerance, veto effects, importance of criteria, interaction between criteria, degree of use of the data, etc.

It seems difficult to relate these behavioral properties which are not precisely defined, to well cut mathematical properties. We know that order statistics OS_k are more or less tolerant depending on the value of k, we know that a weighted arithmetic mean WAM is able to represent importance on criteria, but not interaction, but what about OWA, quasi-arithmetic means, weighted minimum, etc.?

6.2.1 Importance and interaction among criteria

We recall that a fuzzy measure μ defines weights on individual criteria by means of the coefficients μ_i, but also on any group S of criteria by means of μ_S. This makes possible the representation of interaction between criteria.

The Shapley and Banzhaf power indices presented in Section 5.1 seem to be suitable for modelling global importance of criteria. We know that the importance of a singleton i may be low although for most $S \subseteq N \setminus i$, $\mu_{S\cup i}$ could be high, showing that i is however an important element in the decision. For a given criterion i, we will call μ_i the *apparent weight* of criterion i, and $\phi_{\mathrm{Sh}}(i)$ or $\phi_{\mathrm{B}}(i)$ the *real weight* of criterion i. Thus the apparent weights can be very different from the real weights.

Likewise, the Shapley and Banzhaf interaction indices presented in Section 5.2 allow to model the degree of interaction that exists between two criteria or even among a combination of criteria. Recall that, for two criteria $i, j \in N$, the following cases can happen:

- $I(ij) > 0$: i and j are complementary or have a positive synergy.
- $I(ij) < 0$: i and j are substitutive or have a negative synergy (overlap effect).
- $I(ij) = 0$: i and j are independent or have no interaction.

We have to mention that in case of negative interaction, there are actually two possible interpretations quite different from each other:

1. criteria i and j are *correlated*. They provide the same information. This kind of phenomenon can be detected by observing the profiles of several actions.
2. criteria i and j are *interchangeable*. The satisfaction of one of the two is sufficient. This is simply an opinion on the importance of the criteria, which is independent of scores obtained by actions on these two criteria.

Power and interaction indices can be computed for fuzzy measures corresponding to various fuzzy integrals: WAM, OWA, etc. Table 6.1 summarizes the Shapley indices for some fuzzy integrals.

fuzzy integral	$\phi_{\mathrm{Sh}}(i)$	$I_{\mathrm{Sh}}(ij)$
WAM_ω	ω_i	0
OWA_ω	$1/n$	$\dfrac{\omega_1-\omega_n}{n-1}$
OS_k	$1/n$	0 if $k \neq 1, n$
$\min_S$	$\begin{cases} 1/s & \text{if } i \in S \\ 0 & \text{otherwise} \end{cases}$	$\begin{cases} 1/(s-1) & \text{if } i,j \in S \\ 0 & \text{otherwise} \end{cases}$
$\max_S$	$\begin{cases} 1/s & \text{if } i \in S \\ 0 & \text{otherwise} \end{cases}$	$\begin{cases} -1/(s-1) & \text{if } i,j \in S \\ 0 & \text{otherwise} \end{cases}$

Table 6.1: Shapley power and interaction indices for various fuzzy integrals

It seems that for OWA operators, the Shapley indices are very much simpler than the Banzhaf indices; indeed, for such operators, we have (see also [86])

$$\begin{aligned}
\phi_{\mathrm{B}}(i) &= \frac{1}{2^{n-1}} \sum_{t=0}^{n-1} \binom{n-1}{t} \omega_{n-t}, \\
I_{\mathrm{B}}(ij) &= \frac{1}{2^{n-2}} \sum_{t=0}^{n-2} \binom{n-2}{t} (\omega_{n-t-1} - \omega_{n-t}).
\end{aligned}$$

We also note that $I_{\mathrm{Sh}}(ij)$ is the same for all pairs of elements, so that the interaction effect between elements is constant on all pairs of elements. $I_{\mathrm{Sh}}(ij)$ is strictly positive if $\omega_1 > \omega_n$. In this case OWA has a conjunctive behavior; indeed, if ω_1 is high (or at least higher than ω_n which is the weight on the highest value) then clearly OWA behaves like a min. Conversely, $I_{\mathrm{Sh}}(ij)$ is strictly negative if $\omega_1 < \omega_n$. In this case OWA has a disjunctive behavior.

An interesting fact is the following. It can happen that no pair of criteria interact while having $\mu_i \neq \phi_{\mathrm{Sh}}(i)$ for at least one $i \in N$. As shown in Table 6.1, it is the case for OS_k for $k \neq 1, n$.

6.2.2 Degree of disjunction

The Choquet integrals allows us, by appropriate choice of the fuzzy measure, to move continuously from min to max. To classify these Choquet integrals in regard to their location on this continuum a measure of disjunction can be introduced.

Let us define the *average value* of the Choquet integral as

$$m(\mathcal{C}_\mu) := \int_{[0,1]^n} \mathcal{C}_\mu(x)\, dx.$$

A *degree of orness* of $\mathcal{C}_\mu$ corresponds to

$$\mathrm{orness}(\mathcal{C}_\mu) := \frac{m(\mathcal{C}_\mu) - m(\min)}{m(\max) - m(\min)}.$$

We note that orness($\mathcal{C}_\mu$) is always in the unit interval. Moreover, we have orness(min) = 0 and orness(max) = 1. Furthermore, it is noted that the closer orness($\mathcal{C}_\mu$) is to 0, the nearer $\mathcal{C}_\mu$ is to min and has a conjunctive behavior, while the closer orness($\mathcal{C}_\mu$) is to 1, the nearer $\mathcal{C}_\mu$ is to max and has a disjunctive behavior.

Thus, the degree of orness is a measure of the tolerance of the decision maker. Tolerant decision makers can accept that only *some* criteria are satisfied; this corresponds to a disjunctive behavior (orness($\mathcal{C}_\mu$) > 0.5), whose extreme exemple is max. On the other hand, intolerant decision makers demand that *most* criteria be equally satisfied; this corresponds to a conjunctive behavior (orness($\mathcal{C}_\mu$) < 0.5), whose extreme example is min. Of course orness($\mathcal{C}_\mu$) $= 0.5$ corresponds to equitable decision makers.

The concept of orness is very useful to get information about the behavior of the decision maker. In fact, two decision makers with same partial scores $x_1, \ldots, x_n$, same weights on criteria, could still have different behaviors in the sense that one of them can be tolerant and the other intolerant.

Notice that the degree of orness defined here corresponds to decision making problems that are modelled by the Choquet integral. Of course it can be defined for any compensative aggregation operator. It should be noted that this concept has been introduced as early as 1974 by Dujmovic [54, 55] in the particular case of root-power means, i.e. operators of the form (3.2). Here, we have merely applied the definition to Choquet integrals.

Theorem 6.2.1 *For any Choquet integral $\mathcal{C}_\mu$, we have*

$$\mathrm{orness}(\mathcal{C}_\mu) = \frac{1}{n-1} \sum_{T \subseteq N} \frac{n-t}{t+1}\, a_T \tag{6.2}$$

where a is the Möbius representation of μ.

Proof. By (5.37), we immediately have $m(\min) = 1/(n+1)$ and $m(\max) = n/(n+1)$. Moreover, by (4.11), we have

$$m(\mathcal{C}_\mu) = \sum_{T \subseteq N} \frac{1}{t+1}\, a_T.$$

We then can conclude since $\sum_{T \subseteq N} a_T = 1$. ∎

Using the fact that $m(\mathcal{C}_\mu) = I_{\mathrm{Sh}}(\emptyset)$ (cf. (5.38)), we can easily see that (see Table 5.3):

$$\mathrm{orness}(\mathcal{C}_\mu) = \frac{1}{n-1} \sum_{T \subsetneq N} \frac{(n-t)!\, t!}{n!}\, \mu_T = \frac{1}{n-1} \sum_{t=1}^{n-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu_T. \tag{6.3}$$

For the particular case of OWA functions, we recognize the degree of orness introduced intuitively in 1988 by Yager [192]:

$$\mathrm{orness}(\mathrm{OWA}_\omega) = \frac{1}{n-1} \sum_{i=1}^{n} (i-1)\, \omega_i.$$

It was proved in [13] that this degree of orness is essentially the expected value of the number of OR operators among the $n-1$ operators connecting n fuzzy sets/predicates in a statement.

For weighted arithmetic means, we have

$$\mathrm{orness}(\mathrm{WAM}_\omega) = \mathrm{orness}(\mathrm{Sh}_\mu) = \frac{1}{2}.$$

Proposition 6.2.1 min *(resp.* max*) is the only Choquet integral* $\mathcal{C}_\mu$ *such that* $\mathrm{orness}(\mathcal{C}_\mu) = 0$ *(resp.* 1*).*

Proof. Suppose $\mathrm{orness}(\mathcal{C}_\mu) = 0$. Then, by (6.3), we have $\mu_T = 0$ for all $T \neq N$ and $\mathcal{C}_\mu = \min$.

Now, suppose that $\mathrm{orness}(\mathcal{C}_\mu) = 1$. Then the Choquet integral $\mathcal{C}_{\mu'}(x) = 1 - \mathcal{C}_\mu(1-x)$, defined by

$$\sum_{T \subseteq N} a'_T \bigwedge_{i \in T} x_i = \sum_{T \subseteq N} a_T \bigvee_{i \in T} x_i,$$

is such that $\mathrm{orness}(\mathcal{C}_{\mu'}) = 0$ and $\mathcal{C}_{\mu'} = \min$, which is sufficient. ■

Table 6.2 gives the degree of orness of some particular Choquet integrals.

Choquet integral $\mathcal{C}_\mu$	degree of orness
WAM_ω	$\frac{1}{2}$
OWA_ω	$\frac{1}{n-1}\sum_{i=1}^{n}(i-1)\,\omega_i$
$\min_S \quad (S \neq \emptyset)$	$\frac{n-s}{(n-1)(s+1)}$
$\max_S \quad (S \neq \emptyset)$	$\frac{ns-1}{(n-1)(s+1)}$
$\mathrm{OS}_k \quad (k \in N)$	$\frac{k-1}{n-1}$
median	$\frac{1}{2}$

Table 6.2: Degree of orness of some Choquet integrals

Note that a degree of intolerance (conjunction, andness) can also be defined. It simply corresponds to

$$\mathrm{andness}(\mathcal{C}_\mu) := \frac{m(\max) - m(\mathcal{C}_\mu)}{m(\max) - m(\min)},$$

and we have $\mathrm{andness}(\mathcal{C}_\mu) = 1 - \mathrm{orness}(\mathcal{C}_\mu)$ for all fuzzy measures μ on N.

6.2.3 Veto and favor effects

An interesting behavioral phenomenon in aggregation is the veto effect, and its counterpart, the favor effect. Suppose M is an aggregation operator being used for a multicriteria decision making problem. A criterion k is said to be a *veto* or a *blocker* for M if its non satisfaction entails necessarily a low global score. Formally, k is a veto for M if

$$M(x_1, \ldots, x_n) \leq x_k, \quad x \in E^n. \tag{6.4}$$

Similarly, the criterion k is a *favor* or a *pusher* for M if its satisfaction entails necessarily a high global score:

$$M(x_1,\dots,x_n)\geq x_k,\quad x\in E^n. \tag{6.5}$$

The concepts of veto and favor have been already proposed by Dubois and Koning [42] in the context of social choice functions, where "favor" was called "dictator".

A consequence of the definition is that no aggregation operator can model simultaneously a veto on a criterion and a favor on another one; indeed it is not possible to have

$$x_i\leq M(x_1,\dots,x_n)\leq x_j\quad\text{for all } x\in E^n.$$

Note that if the decision maker considers that a given criterion must absolutely be satisfied (veto criterion), then he/she is conjunctive oriented. Indeed, by (6.4) we have $m(\mathcal{C}_\mu)\leq 0.5$, which is sufficient. Similarly, if the decision maker considers that a given criterion is sufficient to be satisfied (favor criteria) then he/she is disjunctive oriented: by (6.5) we have $m(\mathcal{C}_\mu)\geq 0.5$.

Proposition 6.2.2 *For the Choquet integral $\mathcal{C}_\mu$ and the Sugeno integral $\mathcal{S}_\mu$, k is a veto if and only if $\mu_T=0$ whenever $k\notin T$. Similarly, k is a favor if and only if $\mu_T=1$ whenever $k\in T$.*

Proof. The case of the Choquet integral has been proved by Grabisch [83]. Let us consider the case of the Sugeno integral.

(Necessity) Trivial, since $\mu_T=\mathcal{S}_\mu(e_T)$ for all $T\subseteq N$.

(Sufficiency) Let $x\in[0,1]^n$. If $\mu_T=0$ whenever $k\notin T$ then we have, by Theorem 4.3.2,

$$\mathcal{S}_\mu(x)=\bigvee_{T\subseteq N\setminus k}[\mu_{T\cup k}\wedge(\bigwedge_{i\in T}x_i)\wedge x_k]\leq x_k.$$

If $\mu_T=1$ whenever $k\in T$, we have, by Theorem 4.3.2,

$$\mathcal{S}_\mu(x)=\bigwedge_{T\subseteq N\setminus k}[\mu_{N\setminus(T\cup k)}\vee(\bigvee_{i\in T}x_i)\vee x_k]\geq x_k.$$

■

Using monotonicity of the fuzzy measure, we immediately see that k is a veto for the Choquet and Sugeno integrals if and only if $\mu_{N\setminus k}=0$. Similarly, it is a favor if and only if $\mu_k=1$.

It seems reasonable to define indices that measure the degree of veto or favor of a given criterion. If the Choquet integral is considered, a natural definition of a degree of veto (resp. favor) consists in considering the probability

$$\Pr[\mathcal{C}_\mu(x)\leq x_k]\qquad(\text{resp. }\Pr[\mathcal{C}_\mu(x)\geq x_k]) \tag{6.6}$$

where $x\in[0,1]^n$ is a multi-dimensional random variable uniformly distributed[2]. Unfortunately, such an index does not depend always continuously on μ, see (6.7) below.

[2]For convenience sake, we consider the uniform distribution. It is an approximation since, in practical applications, all the profiles do not occur with the same frequency.

Proposition 6.2.3 *Assume that $x \in [0,1]^n$ is a multi-dimensional random variable uniformly distributed. Then we have*

$$\Pr[\mathrm{WAM}_\omega(x) \le x_k] = \begin{cases} 1, & \text{if } \omega_k = 1, \\ 1/2, & \text{otherwise.} \end{cases} \tag{6.7}$$

$$\Pr[\min_S(x) \le x_k] = \begin{cases} 1, & \text{if } k \in S, \\ s/(s+1), & \text{otherwise.} \end{cases} \tag{6.8}$$

$$\Pr[\min_S(x) \ge x_k] = \begin{cases} 1/s, & \text{if } k \in S, \\ 1/(s+1), & \text{otherwise.} \end{cases} \tag{6.9}$$

Proof. Let us prove (6.7). The case $\omega_k = 1$ is trivial. If $\omega_k < 1$ then we have

$$\begin{aligned}
\Pr\Big[\sum_{i=1}^n \omega_i\, x_i \le x_k\Big] &= \Pr\Big[\sum_{\substack{i=1\\ i\neq k}}^n \omega_i\, x_i \le x_k(1-\omega_k)\Big] \\
&= \Pr\Big[\frac{1}{1-\omega_k}\sum_{\substack{i=1\\ i\neq k}}^n \omega_i\, x_i \le x_k\Big] \\
&= 1 - \Pr\Big[\frac{1}{1-\omega_k}\sum_{\substack{i=1\\ i\neq k}}^n \omega_i\, x_i > x_k\Big] \\
&= 1 - \int_0^1 \cdots \int_0^1 \frac{1}{1-\omega_k}\sum_{\substack{i=1\\ i\neq k}}^n \omega_i\, x_i \; dx_1 \cdots dx_{k-1} dx_{k+1} \cdots dx_n \\
&= 1 - \frac{1}{2(1-\omega_k)}\sum_{\substack{i=1\\ i\neq k}}^n \omega_i \\
&= 1/2.
\end{aligned}$$

Let us prove (6.8). The case $k \in S$ is trivial. If $k \notin S$ then we have by (5.37),

$$\begin{aligned}
\Pr[\min_S(x) \le x_k] &= 1 - \Pr[\min_S(x) > x_k] \\
&= 1 - \int_{[0,1]^n} \min_S(x)\, dx \\
&= \frac{s}{s+1}.
\end{aligned}$$

Now let us prove (6.9). Suppose that $k \in S$. We immediately have

$$\Pr[\min_S(x) \ge x_k] = \Pr[\min_S(x) = x_k] = \frac{1}{s}.$$

When $k \notin S$, we have

$$\Pr[\min_S(x) \ge x_k] = \int_{[0,1]^n} \min_S(x)\, dx = \frac{1}{s+1}$$

Thus the result is proved. ∎

As for power and interaction indices, we search for veto and favor indices which are linear with respect to the fuzzy measure. If $\mathrm{veto}'(\mathcal{C}_\mu; i)$ and $\mathrm{favor}'(\mathcal{C}_\mu; i)$ denote these indices, we have, as in Proposition 6.1.2:

$$\begin{aligned}
\mathrm{veto}'(\mathcal{C}_\mu;i) &= \sum_{T\subseteq N} a_T\,\mathrm{veto}'(\mathcal{C}_{\mu^{(T)}};i)\\
\mathrm{favor}'(\mathcal{C}_\mu;i) &= \sum_{T\subseteq N} a_T\,\mathrm{favor}'(\mathcal{C}_{\mu^{(T)}};i)
\end{aligned}$$

where a is the Möbius representation of μ and $\mu^{(T)}$ is the unanimity game for T. For $\mathcal{C}_{\mu^{(T)}}$ ($=\min_T$), we might define

$$\begin{aligned}
\mathrm{veto}'(\mathcal{C}_{\mu^{(T)}};i) &:= \Pr[\min_T(x)\leq x_i]\\
\mathrm{favor}'(\mathcal{C}_{\mu^{(T)}};i) &:= \Pr[\min_T(x)\geq x_i],
\end{aligned}$$

so that we have, by (6.8) and (6.9),

$$\begin{aligned}
\mathrm{veto}'(\mathcal{C}_\mu;i) &= 1-\sum_{T\subseteq N\setminus i}\frac{1}{t+1}\,a_T = 1-\sum_{T\subseteq N\setminus i}\frac{(n-t-1)!\,t!}{n!}\,\mu_T, \quad i\in N,\\
\mathrm{favor}'(\mathcal{C}_\mu;i) &= \sum_{T\subseteq N\setminus i}\frac{1}{t+1}\,(a_T+a_{T\cup i}) = \sum_{T\subseteq N\setminus i}\frac{(n-t-1)!\,t!}{n!}\,\mu_{T\cup i}, \quad i\in N.
\end{aligned}$$

Of course, for all $i\in N$, we have $\mathrm{veto}'(\mathcal{C}_\mu;i),\mathrm{favor}'(\mathcal{C}_\mu;i)\in[0,1]$. Moreover, these indices are closely related to the Shapley value, as the following immediate identity shows:

$$\mathrm{veto}'(\mathcal{C}_\mu;i)+\mathrm{favor}'(\mathcal{C}_\mu;i)=\phi_{\mathrm{Sh}}(i)+1.$$

Furthermore we can see that the closer $\mathrm{veto}'(\mathcal{C}_\mu;i)$ is to 1, the more i is a veto criterion (for $\mathcal{C}_\mu$). Likewise, the closer $\mathrm{favor}'(\mathcal{C}_\mu;i)$ is to 1, the more i is a favor criterion. Note that a given criterion i can be both a veto and a favor. In this case, i is a dictator: the Choquet integral reduces to the projection P_i, and we have $\mathrm{veto}'(\mathrm{P}_i;i)=\mathrm{favor}'(\mathrm{P}_i;i)=1$.

Now, consider the operator $M=\min$. It is such that any criterion is a veto and we have $\mathrm{veto}'(\mathcal{C}_\mu;i)=1$ and $\mathrm{favor}'(\mathcal{C}_\mu;i)=1/n$ for all $i\in N$. On the other hand, $M=\max$ is such that any criterion is a favor and, in this case, $\mathrm{veto}'(\mathcal{C}_\mu;i)=1/n$ and $\mathrm{favor}'(\mathcal{C}_\mu;i)=1$ for all $i\in N$. Thus, it seems better to replace the veto and favor indices by

$$\frac{n\,\mathrm{veto}'(\mathcal{C}_\mu;i)-1}{n-1}\quad\text{and}\quad\frac{n\,\mathrm{favor}'(\mathcal{C}_\mu;i)-1}{n-1}$$

respectively, so that they range in $[0,1]$. We then present the following new definitions:

$$\begin{aligned}
\mathrm{veto}(\mathcal{C}_\mu;i) &:= 1-\frac{n}{n-1}\sum_{T\subseteq N\setminus i}\frac{1}{t+1}\,a_T, \quad i\in N,\\
\mathrm{favor}(\mathcal{C}_\mu;i) &:= \frac{n}{n-1}\sum_{T\subseteq N\setminus i}\frac{1}{t+1}\,(a_{T\cup i}+a_T)-\frac{1}{n-1}, \quad i\in N,
\end{aligned}$$

In terms of the fuzzy measure, these indices become:

$$\begin{aligned}
\mathrm{veto}(\mathcal{C}_\mu;i) &= 1-\frac{1}{n-1}\sum_{T\subseteq N\setminus i}\frac{(n-t-1)!\,t!}{(n-1)!}\,\mu_T, \quad i\in N,\\
\mathrm{favor}(\mathcal{C}_\mu;i) &= \frac{1}{n-1}\sum_{T\subseteq N\setminus i}\frac{(n-t-1)!\,t!}{(n-1)!}\,\mu_{T\cup i}-\frac{1}{n-1}, \quad i\in N.
\end{aligned}$$

and we have

$$\mathrm{veto}(\mathcal{C}_\mu; i) + \mathrm{favor}(\mathcal{C}_\mu; i) = \frac{1}{n-1}(n\,\phi_{\mathrm{Sh}}(i) + n - 2), \quad i \in N.$$

Thus defined, we see that $\mathrm{veto}(\mathcal{C}_\mu; i)$ is more or less the degree to which the decision maker demands that criterion i is satisfied. Notice that such a concept is different from the weights of criteria: we might have a high degree of veto on a not very important criterion.

Likewise, $\mathrm{favor}(\mathcal{C}_\mu; i)$ is the degree to which the decision maker considers that a good score along criterion i is sufficient to be satisfied. Table 6.3 gives the veto and favor indices for some Choquet integrals.

Choquet integral $\mathcal{C}_\mu$	$\mathrm{veto}(\mathcal{C}_\mu; i)$	$\mathrm{favor}(\mathcal{C}_\mu; i)$
AM	$1/2$	$1/2$
WAM_ω	$\frac{1}{2} + \frac{n(\omega_i - 1/n)}{2(n-1)}$	$\frac{1}{2} + \frac{n(\omega_i - 1/n)}{2(n-1)}$
OWA_ω	$1 - \frac{1}{n-1}\sum_{j=1}^{n}(j-1)\,\omega_j$	$\frac{1}{n-1}\sum_{j=1}^{n}(j-1)\,\omega_j$
OS_k	$\frac{n-k}{n-1}$	$\frac{k-1}{n-1}$
median	$1/2$	$1/2$
$\min_S$	$\begin{cases} 1 & \text{if } i \in S \\ \frac{ns-s-1}{(n-1)(s+1)} & \text{otherwise} \end{cases}$	$\begin{cases} \frac{n-s}{(n-1)s} & \text{if } i \in S \\ \frac{n-s-1}{(n-1)(s+1)} & \text{otherwise} \end{cases}$
$\max_S$	$\begin{cases} \frac{n-s}{(n-1)s} & \text{if } i \in S \\ \frac{n-s-1}{(n-1)(s+1)} & \text{otherwise} \end{cases}$	$\begin{cases} 1 & \text{if } i \in S \\ \frac{ns-s-1}{(n-1)(s+1)} & \text{otherwise} \end{cases}$

Table 6.3: Veto and favor indices for various Choquet integrals

To justify the use of the veto and favor indices introduced above, we now propose an axiomatic characterization. Before presenting it, we need a lemma.

Lemma 6.2.1 *Let ψ be a real-valued function defined on the set of fuzzy measures on N. Assume that ψ is linear w.r.t. μ, that is, there exist real constants p_T, $T \subseteq N$, such that*

$$\psi(\mu) = \sum_{T \subseteq N} p_T\,\mu_T$$

for any fuzzy measure on N. Then, if $\mu_{\min_S} = \mu^{(S)}$ (resp. $\mu_{\max_S}$) is the fuzzy measure on N corresponding to $\min_S$ (resp. $\max_S$) then

$$\psi(\mu_{\min_S}) \;=\; \sum_{\substack{T \subseteq S \\ T \neq \emptyset}} (-1)^{t+1}\,\psi(\mu_{\max_T}) \tag{6.10}$$

$$\psi(\mu_{\max_S}) \;=\; \sum_{\substack{T\subseteq S\\ T\neq\emptyset}} (-1)^{t+1}\,\psi(\mu_{\min_T}) \tag{6.11}$$

Proof. Let us prove (6.11). By the linearity of ψ, we have

$$\psi(\mu) = \sum_{T\subseteq N} a_T\,\psi(\mu_{\min_T}), \quad \text{for all } \mu.$$

Moreover, the Möbius representation of $\mu_{\max_S}$ is given by (see Section 4.4.2):

$$\forall\, T\subseteq N, T\neq\emptyset : a_T = \begin{cases} (-1)^{t+1}, & \text{if } T\subseteq S,\\ 0, & \text{otherwise,}\end{cases}$$

which is sufficient.

Now, let us prove (6.10). By (6.11), we have

$$\begin{aligned}\sum_{\substack{T\subseteq S\\ T\neq\emptyset}} (-1)^{t+1}\,\psi(\mu_{\max_T}) &= \sum_{\substack{T\subseteq S\\ T\neq\emptyset}} (-1)^{t+1} \sum_{\substack{K\subseteq T\\ K\neq\emptyset}} (-1)^{k+1}\,\psi(\mu_{\min_K})\\ &= \sum_{\substack{K\subseteq S\\ K\neq\emptyset}} \psi(\mu_{\min_K}) \sum_{T: K\subseteq T\subseteq S} (-1)^{t-k}\\ &= \psi(\mu_{\min_S}).\end{aligned}$$

■

Theorem 6.2.2 *A function $\psi(\mathcal{C}_\mu; i)$ defined on N and the set of Choquet integrals w.r.t. a fuzzy measure on N satisfies the following axioms:*

- linearity axiom: *there exist real constants p^i_T, $T\subseteq N$, such that*

$$\psi(\mathcal{C}_\mu; i) = \sum_{T\subseteq N} p^i_T\,\mu_T, \quad \textit{for all } \mu,$$

- symmetry axiom: *for all $\pi\in\Pi_n$, we have*

$$\psi(\mathcal{C}_\mu; i) = \psi(\mathcal{C}_{\pi\mu}; \pi(i)), \quad \textit{for all } \mu,$$

- boundary axiom: *for all $S\subseteq N$ and all $i\in S$,*

$$\psi(\min_S; i) = 1, \quad (\textit{resp. } \psi(\max_S; i) = 1)$$

- normalization axiom:

$$\frac{1}{n}\sum_{i=1}^{n} \psi(\mathcal{C}_\mu; i) = \text{andness}(\mathcal{C}_\mu), \quad \textit{for all } \mu,$$

$$(\textit{resp. } \frac{1}{n}\sum_{i=1}^{n} \psi(\mathcal{C}_\mu; i) = \text{orness}(\mathcal{C}_\mu), \quad \textit{for all } \mu),$$

if and only if $\psi(\mathcal{C}_\mu; i) = \text{veto}(\mathcal{C}_\mu; i)$ (resp. $\text{favor}(\mathcal{C}_\mu; i)$).

Proof. (Sufficiency) One can verify that veto and favor fulfil the corresponding axioms.

(Necessity) Consider first the case of veto. Let $S \subseteq N$ and $\pi \in \Pi_n$ such that $\pi(S) = S$. Then we can easily see that $\mu_{\min_S} = \pi\mu_{\min_S}$. Moreover, by the symmetry axiom, we have

$$\psi(\min_S; i) = \psi(\min_S; j), \quad \forall i, j \notin S.$$

By the normalization axiom, we have

$$\frac{1}{n}\sum_{i\in S}\psi(\min_S; i) + \frac{1}{n}\sum_{i\notin S}\psi(\min_S; i) = \frac{ns-1}{(n-1)(s+1)}.$$

By the boundary axiom, this identity becomes

$$\frac{s}{n} + \frac{n-s}{n}\,\psi(\min_S; i) = \frac{ns-1}{(n-1)(s+1)}, \quad \forall i \notin S,$$

and we have

$$\psi(\min_S; i) = \frac{ns-s-1}{(n-1)(s+1)}, \quad \forall i \notin S.$$

By the linearity, we then have, for all fuzzy measures μ on N,

$$\begin{aligned}
\psi(\mathcal{C}_\mu; i) &= \sum_{T\subseteq N} a_T\,\psi(\min_T; i) \\
&= \sum_{T\ni i} a_T + \sum_{T\not\ni i}\frac{nt-t-1}{(n-1)(t+1)}\,a_T \\
&= 1 - \frac{n}{n-1}\sum_{T\not\ni i}\frac{1}{t+1}\,a_T, \quad (\text{since } \sum_{T\subseteq N} a_T = 1).
\end{aligned}$$

Now consider the case of favor. As in the case of veto, we can show that

$$\psi(\max_S; i) = \frac{ns-s-1}{(n-1)(s+1)}, \quad \forall i \notin S.$$

By Lemma 6.2.1, we have, for all $i \notin S$,

$$\begin{aligned}
\psi(\min_S; i) &= \sum_{\substack{T\subseteq S\\ T\neq\emptyset}} (-1)^{t+1}\,\psi(\max_T; i) \\
&= \sum_{\substack{T\subseteq S\\ T\neq\emptyset}} (-1)^{t+1}\,\frac{nt-t-1}{(n-1)(t+1)} \\
&= \sum_{t=1}^{s}\binom{s}{t}(-1)^{t+1}\,\frac{nt-t-1}{(n-1)(t+1)} \\
&= \begin{cases} 0, & \text{if } s = 0, \\[2ex] \dfrac{n-s-1}{(n-1)(s+1)}, & \text{otherwise.}\end{cases}
\end{aligned}$$

When $i \in S$, we have, by Lemma 6.2.1,

$$\psi(\min_S; i) = \sum_{\substack{T\subseteq S\\ T\neq\emptyset}} (-1)^{t+1}\,\psi(\max_T; i)$$

$$
\begin{aligned}
&= \sum_{\substack{T\subseteq S\\ T\ni i}} (-1)^{t+1} + \sum_{\substack{\emptyset\neq T\subseteq S\\ T\not\ni i}} (-1)^{t+1}\,\frac{nt-t-1}{(n-1)(t+1)}\\
&= \underbrace{\sum_{t=1}^{s}\binom{s-1}{t-1}(-1)^{t+1}}_{\begin{cases}1, & \text{if } s=1\\ 0, & \text{otherwise}\end{cases}} + \underbrace{\sum_{t=1}^{s-1}\binom{s-1}{t}(-1)^{t+1}\,\frac{nt-t-1}{(n-1)(t+1)}}_{\begin{cases}0, & \text{if } s=1\\ \frac{n-s}{s(n-1)}, & \text{otherwise}\end{cases}}\\
&= \begin{cases}1, & \text{if } s=1,\\[1ex] \dfrac{n-s}{s(n-1)}, & \text{otherwise}\end{cases}\\
&= \frac{n-s}{s(n-1)}
\end{aligned}
$$

Now, by the linearity, we have, for all fuzzy measures μ on N,

$$
\begin{aligned}
\psi(\mathcal{C}_\mu;i) &= \sum_{T\subseteq N} a_T\,\psi(\min_T;i)\\
&= \sum_{T\ni i} a_T\,\frac{n-t}{t(n-1)} + \sum_{T\not\ni i} a_T\,\frac{n-t-1}{(n-1)(t+1)}\\
&= \frac{1}{n-1}\sum_{T\not\ni i}(a_{T\cup i}+a_T)\,\frac{n-t-1}{t+1}\\
&= \frac{n}{n-1}\sum_{T\not\ni i}\frac{1}{t+1}\,(a_{T\cup i}+a_T) - \frac{1}{n-1}
\end{aligned}
$$

and the proof is complete. ■

It is possible to generalize the concept of veto to several criteria [83]: a veto for criteria $K\subseteq N$, which means $M(x)\leq \bigwedge_{k\in K} x_k$, is obtained for the Choquet and Sugeno integrals by any fuzzy measure μ such that $\mu_T=0$ whenever $K\not\subseteq T$. Similarly, a favor for criteria $K\subseteq N$, which means $M(x)\geq \bigvee_{k\in K} x_k$, is obtained by any fuzzy measure μ such that $\mu_T=1$ whenever $K\cap T\neq\emptyset$.

We can also extend all these concepts to a variety of situations. As example, let us imagine that at least one criterion must be satisfied among a subset K of criteria. In this case, we clearly have

$$M(x_1,\ldots,x_n)\leq \bigvee_{k\in K} x_k,\quad x\in E^n.$$

6.2.4 Measure of dispersion

Consider the median and the arithmetic mean, both of which are OWA operators with weight vectors of the form

$$(0,\ldots,1,\ldots,0)\quad\text{and}\quad(\frac{1}{n},\ldots,\frac{1}{n}),$$

respectively. We note that these operators have the same degree of orness, 1/2, but we can see that they are different in the sense that the first one focuses all the weight on only one argument.

In order to capture this idea, Yager [192] proposed a *measure of dispersion* associated to the weight vector ω of an OWA operator:

$$\mathrm{disp}(\omega) = -\sum_{i=1}^{n} \omega_i \ln \omega_i, \tag{6.12}$$

where ln is the Neperian natural logarithm, and $0 \ln 0 := 0$ by convention.

This dispersion is actually a measure of *entropy*, a well-known concept introduced as early as 1949 in the Shannon information theory [168]. It allows to measure how much of the information in the arguments is used. In a certain sense the more disperse the ω the more the information about the individual criteria is being used in the aggregation process.

The dispersion is maximum only when the weight vector corresponds to that of the arithmetic mean [189]:

$$\mathrm{disp}(\omega_{\mathrm{AM}}) = \ln n,$$

and minimum only when it corresponds to that of an order statistic:

$$\mathrm{disp}(\omega_{\mathrm{OS}_k}) = 0, \quad k \in N.$$

Thus, we always have

$$0 \leq \mathrm{disp}(\omega) \leq \ln n$$

and the dispersion can be normalized into

$$\mathrm{disp}'(\omega) = -\frac{1}{\ln n} \sum_{i=1}^{n} \omega_i \ln \omega_i.$$

We now intend to generalize this concept to any fuzzy measure. On the one hand, comparing the operators

$$\mathrm{OWA}_\omega(x) = \sum_{i=1}^{n} x_{(i)}\, \omega_i \quad \text{and} \quad \mathcal{C}_\mu(x) = \sum_{i=1}^{n} x_{(i)}\, [\delta_{(i)}\, \mu_{\{(i),\dots,(n)\}}]$$

suggests to propose as measure of dispersion a sum over $i \in N$ of an average value of

$$[\delta_i\, \mu_{T\cup i}] \ln[\delta_i\, \mu_{T\cup i}], \qquad T \subseteq N \setminus i.$$

On the other hand, thinking of the concept of cardinal-probabilistic value (see Section 5.1.2), we can propose a measure of dispersion of the form

$$\mathrm{disp}(\mu) = -\sum_{i=1}^{n} \sum_{T\subseteq N\setminus i} p_t\, [\delta_i\, \mu_{T\cup i}] \ln[\delta_i\, \mu_{T\cup i}], \tag{6.13}$$

with $\sum_{T\subseteq N\setminus i} p_t = 1$ and $p_t \geq 0$.

In the particular case of OWA operators, we have $\delta_i\, \mu_{T\cup i} = \omega_{n-t}$ (cf. (4.14)) and the measure of dispersion (6.13) becomes

$$\begin{aligned}
\mathrm{disp}(\mu_{\mathrm{OWA}_\omega}) &= -\sum_{i=1}^{n} \sum_{T\subseteq N\setminus i} p_t\, \omega_{n-t} \ln \omega_{n-t} \\
&= -\sum_{i=1}^{n} \sum_{t=0}^{n-1} \binom{n-1}{t} p_t\, \omega_{n-t} \ln \omega_{n-t} \\
&= -\sum_{t=0}^{n-1} n \binom{n-1}{t} p_t\, \omega_{n-t} \ln \omega_{n-t}
\end{aligned}$$

Since this expression should coincide with (6.12), we must set

$$p_t := \frac{1}{n\binom{n-1}{t}} = \frac{(n-t-1)!\,t!}{n!},$$

and the coefficients p_t are those of the Shapley value.

We then propose the following definition.

Definition 6.2.1 The dispersion of a fuzzy measure μ on N is defined by

$$\mathrm{disp}(\mu) = -\sum_{i=1}^{n} \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\,t!}{n!} [\delta_i\, \mu_{T \cup i}] \ln[\delta_i\, \mu_{T \cup i}]. \tag{6.14}$$

In a decision problem modelled by the Choquet integral, the dispersion (6.14) can be interpreted as the degree to which we use all the information contained in the arguments $(x_1, \ldots, x_n)$ when calculating the aggregated value $\mathcal{C}_\mu(x_1, \ldots, x_n)$.

In the particular case of WAM operators, we have $\delta_i\, \mu_{T \cup i} = \omega_i$, and hence

$$\mathrm{disp}(\mu_{\mathrm{WAM}_\omega}) = \mathrm{disp}(\mu_{\mathrm{OWA}_\omega}) = -\sum_{i=1}^{n} \omega_i \ln \omega_i,$$

which do not contradict the idea of a dispersion measure: the dispersion should not depend on a reordering of the arguments.

The dispersion (6.14) is still to be characterized. However, we believe that many properties of the classical entropy can be adapted to this new dispersion. For instance, we have the following.

Proposition 6.2.4 *Let μ be a fuzzy measure on N. Then* $\mathrm{disp}(\mu) = 0$ *if and only if μ is a 0-1 fuzzy measure.*

Proof. We simply have

$$\begin{aligned}
\mathrm{disp}(\mu) = 0 \quad &\Leftrightarrow \quad [\delta_i\, \mu_{T \cup i}] \ln[\delta_i\, \mu_{T \cup i}] = 0, \quad \forall i \in N,\ \forall T \subseteq N \setminus i, \\
&\Leftrightarrow \quad \delta_i\, \mu_{T \cup i} \in \{0, 1\}, \quad \forall i \in N,\ \forall T \subseteq N \setminus i, \\
&\Leftrightarrow \quad \mu_T \in \{0, 1\}, \quad \forall T \subseteq N.
\end{aligned}$$

∎

Proposition 6.2.4 is in agreement with the idea of a dispersion measure. Indeed, by Theorem 4.4.2, μ is a 0-1 fuzzy measure if and only if there exists $c : 2^N \to \{0, 1\}$ such that $\mathcal{C}_\mu = \mathrm{B}_c^{\vee\wedge}$. Since $\mathrm{B}_c^{\vee\wedge}(x) \in \{x_1, \ldots, x_n\}$, only one piece of information is essentially used in the aggregation.

The dispersion can be computed for various aggregation operators. Table 6.4 summarizes the dispersion for some operators.

We now show that $0 \leq \mathrm{disp}(\mu) \leq \ln n$, and that the dispersion is maximum only when $\mathcal{C}_\mu = \mathrm{AM}$. The proof is a straightforward adaptation of that used for the classical entropy.

Lemma 6.2.2 *If the numbers* $\{c_T^i > 0, d_T^i \geq 0 \,|\, i \in N, T \subseteq N \setminus i\}$ *are such that*

$$\sum_{i=1}^{n} \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\,t!}{n!}\, c_T^i \leq 1 \quad \textit{and} \quad \sum_{i=1}^{n} \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\,t!}{n!}\, d_T^i = 1$$

operator	dispersion
WAM_ω	$-\sum_{i=1}^{n}\omega_i\ln\omega_i$
OWA_ω	$-\sum_{i=1}^{n}\omega_i\ln\omega_i$
owmax_ω	$-\sum_{i=1}^{n}(\omega_i-\omega_{i+1})\ln(\omega_i-\omega_{i+1}),\ \text{ with }\ \omega_{n+1}=0$
amed_α	$-\alpha\ln\alpha-(1-\alpha)\ln(1-\alpha)$
AM	$\ln n$
$\mathrm{B}_c^{\vee\wedge}$	0

Table 6.4: Dispersion for some operators

then

$$\sum_{i=1}^{n}\sum_{T\subseteq N\setminus i}\frac{(n-t-1)!\,t!}{n!}\,d_T^i\,\ln(\frac{c_T^i}{d_T^i})\leq 0,$$

Moreover, the equality holds if and only if

$$d_T^i\neq 0\ \Rightarrow\ d_T^i=c_T^i,\qquad \forall i\in N,\forall T\subseteq N\setminus i.$$

Proof. Since $\ln x\leq x-1$ for all $x>0$, we simply have

$$\begin{aligned}\sum_{i=1}^{n}\sum_{\substack{T\subseteq N\setminus i\\ d_T^i\neq 0}}\frac{(n-t-1)!\,t!}{n!}\,d_T^i\,\ln(\frac{c_T^i}{d_T^i}) &\leq \sum_{i=1}^{n}\sum_{\substack{T\subseteq N\setminus i\\ d_T^i\neq 0}}\frac{(n-t-1)!\,t!}{n!}\,d_T^i\,(\frac{c_T^i}{d_T^i}-1)\\ &= \sum_{i=1}^{n}\sum_{\substack{T\subseteq N\setminus i\\ d_T^i\neq 0}}\frac{(n-t-1)!\,t!}{n!}\,c_T^i-1\\ &\leq 1-1=0.\end{aligned}$$

The second part follows from the fact that $\ln x=x-1$ if and only if $x=1$. ■

Theorem 6.2.3 *For any fuzzy measure μ on N, we have $0\leq\mathrm{disp}(\mu)\leq\ln n$. Moreover,* $\mathrm{disp}(\mu)=\ln n$ *if and only if $\delta_i\,\mu_{T\cup i}=1/n$ for all $i\in N$ and all $T\subseteq N\setminus i$.*

Proof. For all $i\in N$ and all $T\subseteq N\setminus i$, we have $\delta_i\,\mu_{T\cup i}\in[0,1]$ and hence $-\ln[\delta_i\,\mu_{T\cup i}]\geq 0$ and $\mathrm{disp}(\mu)\geq 0$.

For the second inequality, we have, since $\sum_{i=1}^n \phi_{\rm Sh}(i) = 1$,

$$\begin{aligned}
\mathrm{disp}(\mu) - \ln n &= \sum_{i=1}^n \sum_{T\subseteq N\setminus i} \frac{(n-t-1)!\,t!}{n!}\,[\delta_i\,\mu_{T\cup i}]\ln\frac{1}{\delta_i\,\mu_{T\cup i}} + \Big[\sum_{i=1}^n \phi_{\rm Sh}(i)\Big]\ln\frac{1}{n} \\
&= \sum_{i=1}^n \sum_{T\subseteq N\setminus i} \frac{(n-t-1)!\,t!}{n!}\,[\delta_i\,\mu_{T\cup i}]\Big[\ln\frac{1}{\delta_i\,\mu_{T\cup i}} + \ln\frac{1}{n}\Big] \\
&= \sum_{i=1}^n \sum_{T\subseteq N\setminus i} \frac{(n-t-1)!\,t!}{n!}\,[\delta_i\,\mu_{T\cup i}]\ln\frac{1}{n\,\delta_i\,\mu_{T\cup i}}.
\end{aligned}$$

Applying Lemma 6.2.2 with $c_T^i = 1/n$ and $d_T^i = \delta_i\,\mu_{T\cup i}$ leads to $\mathrm{disp}(\mu) - \ln n \leq 0$. The equality holds if and only if $\delta_i\,\mu_{T\cup i} = 1/n$. ■

Thus defined, $\mathrm{disp}(\mu)$ must be viewed as an *absolute* dispersion. Clearly, a *relative* dispersion can be defined by

$$\mathrm{disp}'(\mu) = \frac{1}{\ln n}\,\mathrm{disp}(\mu)$$

and by Theorem 6.2.3, we have $0 \leq \mathrm{disp}'(\mu) \leq 1$.

6.2.5 An illustrative example

We give here an example, borrowed from Grabisch [79, 80]. Let us consider the problem of evaluating students in high school with respect to three subjects: mathematics (M), physics (P) and literature (L). Usually, this is done by a simple weighted arithmetic mean, whose weights are the coefficients of importance of the differents subjects. Suppose that the school is more scientifically than literary oriented, so that weights could be for example 3, 3 and 2 respectively. Then the weighted arithmetic mean will give the following results for three students a, b and c (marks are given on a scale from 0 to 20):

student	M	P	L	global evaluation (weighted arithmetic mean)
a	18	16	10	15.25
b	10	12	18	12.75
c	14	15	15	14.625

If the school wants to favor well equilibrated students without weak points then student c should be considered better than student a, who has a severe weakness in literature. Unfortunately, no weight vector $(\omega_{\rm M}, \omega_{\rm P}, \omega_{\rm L})$ satisfying $\omega_{\rm M} = \omega_{\rm P} > \omega_{\rm L}$ is able to favor student c; indeed, we have:

$$c \succ a \iff \omega_{\rm L} > \omega_{\rm M}.$$

The reason of this problem is that too much importance is given to mathematics and physics, which present some overlap effect since, usually, students good at mathematics are also good at physics (and vice versa), so that the evaluation is overestimated (resp. underestimated) for students good (resp. bad) at mathematics and/or physics. This problem can be solved by using a suitable fuzzy measure μ and the Choquet integral.

- Since scientific subjects are more important than literature, the following weights can be put on subjects taken individually (apparent weights): $\mu_{\rm M} = \mu_{\rm P} = 0.45$ and $\mu_{\rm L} = 0.3$. Note that the initial ratio of weights $(3, 3, 2)$ is kept unchanged.

$$
\begin{aligned}
a_S &= \sum_{T\subseteq S} (-1)^{s-t}\mu_T, && S\subseteq N\\
\phi_{\mathrm{Sh}}(i) &= \sum_{T\ni i} \frac{1}{t}\, a_T, && i\in N\\
I_{\mathrm{Sh}}(S) &= \sum_{T\supseteq S} \frac{1}{t-s+1}\, a_T, && S\subseteq N\\
\mathcal{C}_\mu(x) &= \sum_{T\subseteq N} a_T \bigwedge_{i\in T} x_i, && x\in[0,1]^n\\
\mathrm{veto}(\mathcal{C}_\mu;i) &= 1-\frac{n}{n-1}\sum_{T\not\ni i}\frac{1}{t+1}\, a_T, && i\in N\\
\mathrm{favor}(\mathcal{C}_\mu;i) &= \frac{n}{n-1}\sum_{T\not\ni i}\frac{1}{t+1}\,(a_{T\cup i}+a_T)-\frac{1}{n-1}\\
&= \frac{1}{n-1}(n\,\phi_{\mathrm{Sh}}(i)+n-2)-\mathrm{veto}(\mathcal{C}_\mu;i), && i\in N\\
m(\mathcal{C}_\mu) &= \sum_{T\subseteq N}\frac{1}{t+1}\, a_T = I_{\mathrm{Sh}}(\emptyset)\\
\mathrm{orness}(\mathcal{C}_\mu) &= \frac{1}{n-1}\Big[(n+1)\,m(\mathcal{C}_\mu)-1\Big] = \frac{1}{n}\sum_{i=1}^n \mathrm{favor}(\mathcal{C}_\mu;i)\\
\mathrm{disp}(\mu) &= -\sum_{i=1}^n\sum_{T\not\ni i}\frac{(n-t-1)!\,t!}{n!}\,[\delta_i\mu_{T\cup i}]\ln[\delta_i\mu_{T\cup i}]\\
&\quad \text{with}\quad \delta_i\mu_{T\cup i}=\sum_{K\subseteq T} a_{K\cup i}, && i\in N,\ T\not\ni i\\
\mathrm{disp}'(\mu) &= \frac{1}{\ln n}\,\mathrm{disp}(\mu)
\end{aligned}
$$

Table 6.5: Compilation of some useful formulas

- Since mathematics and physics overlap, the weights attributed to the pair $\{\mathrm{M}, \mathrm{P}\}$ should be less than the sum of the weights of mathematics and physics: $\mu_{\mathrm{MP}} = 0.5$.
- Since students equally good at scientific subjects and literature must be favored, the weight attributed to the pair $\{\mathrm{L}, \mathrm{M}\}$ should be greater than the sum of individual weights (the same for physics and literature): $\mu_{\mathrm{ML}} = \mu_{\mathrm{PL}} = 0.9$.
- $\mu_{\emptyset} = 0$ and $\mu_{\mathrm{MPL}} = 1$ by definition.

The Möbius representation is then given by

$$\begin{array}{llll} a_{\emptyset} = 0 & a_{\mathrm{M}} = 0.45 & a_{\mathrm{MP}} = -0.40 & a_{\mathrm{MPL}} = -0.10 \\ & a_{\mathrm{P}} = 0.45 & a_{\mathrm{ML}} = 0.15 & \\ & a_{\mathrm{L}} = 0.30 & a_{\mathrm{PL}} = 0.15 & \end{array}$$

Applying Choquet integral with the above fuzzy measure leads to the following new global evaluation:

student	M	P	L	global evaluation (Choquet integral)
a	18	16	10	13.9
b	10	12	18	13.6
c	14	15	15	14.6

The expected result is then obtained. Also remark that student b has still the lowest rank, as requested by the scientific tendency of this high school.

Note that the Shapley integral provides the following global evaluations: 14.08333, 13.91666 and 14.70833.

Now, let us turn to a deeper analysis of the orientation of the school or its director. From the fuzzy measure proposed, we obtain the following Shapley value and degrees of veto and favor:

	M	P	L
veto$(\mathcal{C}_\mu; i)$	0.3625	0.3625	0.525
favor$(\mathcal{C}_\mu; i)$	0.575	0.575	0.6
$\phi_{\mathrm{Sh}}(i)$	0.29166	0.29166	0.41666
$n * \phi_{\mathrm{Sh}}(i)$	0.875	0.875	1.25

As we can see, it is convenient to scale the Shapley value by a factor n, so that an importance index greater than 1 indicates an attribute more important than the average. Moreover, looking at the veto and favor degrees, we observe that the school seems to favor slightly the students (disjunctive oriented). This is in accordance with the degree of orness

$$\mathrm{orness}(\mathcal{C}_\mu) = 0.58333$$

Now, regarding the Shapley interaction indices I_{Sh}, we have:

	P	L
M	-0.45	0.10
P		0.10

These numerical values are also in accordance with the interpretation of the fuzzy measure. Moreover, the absolute dispersion of the fuzzy measure is given by $\mathrm{disp}(\mu) = 0.90084$. This leads to a rather satisfactory relative dispersion:

$$\mathrm{disp}'(\mu) = 0.81998$$

Remark that all these behavioral parameters have been obtained from a given fuzzy measure. In practical situations, the fuzzy measure is not completely available. We might then fix its values from information on behavioral parameters. Section 6.4 deals with this topic.

6.3 k-order fuzzy measures

6.3.1 Motivation

We know that a problem involving n criteria requires 2^n coefficients in $[0, 1]$ in order to define the fuzzy measure μ on every coalition. Of course, a decision maker is not able to give such an amount of information. Moreover, the meaning of the numbers μ_S and a_S is not always clear for the decision maker. In most cases, he/she will be able to guess the importance of singletons, of pairs of elements, but not that of subsets of more elements. Reciprocally, if a fuzzy measure is given, no expert can tell exactly what it means in terms of behavior in decision making. Thus, although fuzzy measures constitute a flexible tool for modelling the importance of coalitions, they are not easy to handle and interpret in a practical problem.

To overcome this problem, Grabisch [81] proposed to use the concept of *k-order fuzzy measure*. Looking at the polynomial expression of a fuzzy measure (4.2), one can notice that additive measures have a linear representation $f(x) = \sum_{i=1}^{n} a_i\, x_i$. By extension, we may think of a fuzzy measure having a polynomial representation of degree 2, or 3, or any fixed integer k. Such a fuzzy measure is naturally called k-order fuzzy measure since it represents a k-order approximation of its polynomial expression in the neighborhood of the origin.

Definition 6.3.1 A fuzzy measure μ defined on N is said to be of order k if its corresponding pseudo-Boolean function is a multilinear polynomial of degree k, i.e. $a_T = 0$ for all T such that $t > k$, and there exists at least one subset K of k elements such that $a_K \neq 0$.

This concept allows to range freely between purely additive measures ($k = 1$), defined by n coefficients, and general fuzzy measures ($k = n$), defined by 2^n coefficients. Indeed, when varying k from 1 to n, we recover all possible fuzzy measures. Looking at the passage formulas between a and I_{Sh}, I_{B}, I_{R}, we see that considering k-order fuzzy measures merely amounts to assuming that the interactions $I(T)$ for $t > k$ are zero.

The following result is easy to prove (see e.g. [83]).

Proposition 6.3.1 *Let μ be a k-order measure on N. Then*

i) $I_{\mathrm{Sh}}(T) = I_{\mathrm{B}}(T) = I_{\mathrm{R}}(T) = a_T$ *for every T such that $|T| = k$,*

ii) $I_{\mathrm{Sh}}(T) = I_{\mathrm{B}}(T) = I_{\mathrm{R}}(T) = a_T = 0$ *for every T such that $|T| > k$.*

Although the additive model is very simple to handle (only n coefficients are needed), it is restrictive and leads to a very poor modelling tool for application in multicriteria decision making. Grabisch [81] then suggested to consider the 2-order case, which seems to be the most interesting in practical applications, since it permits to model interaction between criteria while remaining very simple. Indeed, only $n+\binom{n}{2}=\frac{n(n+1)}{2}$ coefficients are required to define the fuzzy measure, namely the coefficients

$$\begin{aligned}\mu_i &= a_i, \quad i\in N,\\ \mu_{ij} &= a_i+a_j+a_{ij}, \quad \{i,j\}\subseteq N.\end{aligned}$$

The other coefficients are then given by:

$$\mu_S=\sum_{i\in S}a_i+\sum_{\{i,j\}\subseteq S}a_{ij}=\sum_{\{i,j\}\subseteq S}\mu_{ij}-(s-2)\sum_{i\in S}\mu_i, \quad S\subseteq N,\ s\geq 2.$$

In this case, the Choquet integral becomes

$$C_\mu(x)=\sum_{i\in N}a_i\,x_i+\sum_{\{i,j\}\subseteq N}a_{ij}\,(x_i\wedge x_j), \quad x\in\mathbb{R}^n. \tag{6.15}$$

Moreover, the power and interaction indices coincide ($I_{\mathrm{Sh}}=I_{\mathrm{B}}=I_{\mathrm{R}}=I$) and we have immediately:

$$I(i) = a_i+\frac{1}{2}\sum_{j\in N\setminus i}a_{ij}, \quad i\in N, \tag{6.16}$$

$$I(ij) = a_{ij}, \quad i,j\in N, \tag{6.17}$$

and $I(S)=0$ for all $S\subseteq N$, $s>2$. We thus remark that the interaction index $I(ij)$ coincides with the second degree term of the Choquet integral.

In this context conditions (4.6) for the coefficients $a_\emptyset$, a_i ($i\in N$), a_{ij} ($i,j\in N$) to define a fuzzy measure become:

$$\begin{cases} a_\emptyset=0,\\ \displaystyle\sum_{i\in N}a_i+\sum_{\{i,j\}\subseteq N}a_{ij}=1,\\ a_i\geq 0, \quad \forall i\in N,\\ \displaystyle a_i+\sum_{j\in T}a_{ij}\geq 0, \quad \forall i\in N, \forall T\subseteq N\setminus i.\end{cases} \tag{6.18}$$

Notice that, for any fuzzy measure, we have $I(i), a_i\in[0,1]$ for all $i\in N$ and

$$\sum_{i\in N}I(i)=1.$$

Moreover, Roubens [156] and Grabisch [81] proved that $I(ij)\in[-1,1]$ for all $i,j\in N$, which can be viewed as a consequence of the recurrence formula (5.14). Hence, for any 2-order fuzzy measure, we have $a_{ij}\in[-1,1]$ for all $\{i,j\}\subseteq N$. See general results on the bounds of a_S and $I(S)$ by Miranda and Grabisch [123].

It is also interesting to note that, for 2-order fuzzy measures, the Shapley and Banzhaf indices coincide, so that we do not have to justify the use of only one of the two. Moreover, for general fuzzy measures, we know that the Banzhaf indices do not sum up to 1, nor to any fixed number.

6.3.2 Alternative representations of the Choquet integral

As the interaction representations have an interesting meaning in the framework of multicriteria decision making, it would be useful to express the Choquet integral with respect to such representations. This would help to have a better understanding of the meaning of the Choquet integral, or to have some more efficient computation of it. We restrict ourselves to the 2-order case, which seems to be the most interesting in practical applications. A Choquet integral w.r.t. a 2-order fuzzy measure will be called a 2-order Choquet integral.

Theorem 6.3.1 *Let μ be a 2-order fuzzy measure on N. Then the best weighted arithmetic mean* WAM_ω *that minimizes*

$$\int_{[0,1]^n} [\mathcal{C}_\mu(x) - \mathrm{WAM}_\omega(x)]^2 \, dx$$

is given by the Shapley integral Sh_μ. *Moreover, we have, if $E \supseteq [0,1]$,*

$$\mathcal{C}_\mu(x) = \mathrm{Sh}_\mu(x) - \frac{1}{2} \sum_{\{i,j\}\subseteq N} I(ij)\,[(x_i \vee x_j) - (x_i \wedge x_j)], \quad x \in E^n. \tag{6.19}$$

Proof. The first part will be proved in Section 7.3.2 (see Theorem 7.3.1).

Now, by using the passage formula from I_{Sh} to a (cf. Table 5.4), we have, for all $x \in E^n$,

$$\begin{aligned}
\mathcal{C}_\mu(x) &= \sum_{\substack{T\subseteq N\\ t\geq 1}} a_T \bigwedge_{i\in T} x_i \\
&= \sum_{\substack{T\subseteq N\\ t\geq 1}} \Big[\sum_{\substack{K\supseteq T\\ k\geq 1}} B_{k-t}\, I_{\mathrm{Sh}}(K)\Big] \bigwedge_{i\in T} x_i \\
&= \sum_{\substack{K\subseteq N\\ k\geq 1}} \Big[\sum_{\substack{T\subseteq K\\ t\geq 1}} B_{k-t} \bigwedge_{i\in T} x_i\Big]\, I_{\mathrm{Sh}}(K).
\end{aligned}$$

Since $I_{\mathrm{Sh}}(K) = 0$ whenever $k \geq 3$, we simply have

$$\mathcal{C}_\mu(x) = \mathrm{Sh}_\mu(x) + \sum_{\{i,j\}\subseteq N} [-\frac{1}{2}x_i - \frac{1}{2}x_j + x_i \wedge x_j]\, I(ij),$$

as expected. ■

Equation (6.19) shows a decomposition of the Choquet integral $\mathcal{C}_\mu$ into a linear part and a non-linear part. The linear part, namely the Shapley integral, appears to be a 1-order approximation of $\mathcal{C}_\mu$. The non-linear part brings some correction to the Shapley integral that we can interpret as follows. Consider a pair of criteria $\{i,j\} \subseteq N$.

- A negative $I(ij)$, that is $I(ij) \in [-1,0]$, implies a disjunctive behavior between i and j. In multicriteria decision making, this means that the satisfaction of either i or j is sufficient to have a significant effect on the global score. If such a unilateral satisfaction is observed, we can increase the linear average by a positive amount that should be proportional to the variation between the utilities x_i and x_j but also to the importance of the interaction $|I(ij)|$.

- A positive $I(ij)$, that is $I(ij) \in [0,1]$, implies a conjunctive behavior, which means that both criteria have to be satisfied to lead to a high global score. In case of unilateral satisfaction, the linear average should be decreased by the same quantity as explained in the first case.

An other interesting form of the 2-order Choquet integral is given by Grabisch [83]:

Theorem 6.3.2 *Let μ be a 2-order fuzzy measure on N and assume $E \supseteq [0,1]$. Then we have*

$$\mathcal{C}_\mu(x) = \sum_{i\in N}\Big(I(i) - \frac{1}{2}\sum_{j\in N\setminus i}|I(ij)|\Big)\,x_i + \sum_{\substack{\{i,j\}\subseteq N\\ I(ij)<0}}|I(ij)|(x_i \wedge x_j) + \sum_{\substack{\{i,j\}\subseteq N\\ I(ij)>0}}|I(ij)|(x_i \vee x_j). \quad (6.20)$$

for all $x \in E^n$. Moreover, we have $I(i) - \frac{1}{2}\sum_{j\in N\setminus i}|I(ij)| \geq 0$ for all $i \in N$.

The interesting point in the decomposition (6.20) is that all terms are positive so that the contribution of each of them in the Choquet integral can be interpreted: a 2-order measure having strongly negative (resp. positive) $I(ij)$ will enforce the Choquet integral to be strongly disjunctive (resp. conjunctive), and a 2-order measure having low values for $I(ij)$ will lead to an almost linear Choquet integral.

As shown by Grabisch [82], decomposition of the Choquet integral in the interaction representation becomes complicated as soon as the fuzzy measure is of order > 2. For such measures, an interpretation similar to the previous one seems to be lost.

Regarding OWA operators, we also have the following.

Proposition 6.3.2 *An OWA_ω operator is defined from a 2-order fuzzy measure if and only if ω_t is linear in t, that is $\omega_t = a\,t + b$ for some $a, b \in \mathbb{R}$.*

Proof. (Necessity) By (4.16), we simply have $\omega_t = a_i + (n-t)\,a_{ij}$ for all $i,j,t \in N$.

(Sufficiency) By (4.18), we have $a_T = [\Delta_k^{t-1} a\,(n-k) + b]_{k=0}$ for all $T \subseteq N$, $T \neq \emptyset$, and hence $a_T = 0$ whenever $t > 2$. ■

We have seen in Table 6.1 that the fuzzy measure associated to an OWA_ω operator is such that (see also Grabisch [86]):

$$I_{\mathrm{Sh}}(i) = 1/n, \quad i \in N, \quad (6.21)$$

$$I_{\mathrm{Sh}}(ij) = \frac{\omega_1 - \omega_n}{n-1}, \quad i,j \in N. \quad (6.22)$$

We now give the form of a 2-order OWA in terms of the indices $I_{\mathrm{Sh}}(i)$ and $I_{\mathrm{Sh}}(ij)$.

Theorem 6.3.3 *Assume $E \supseteq [0,1]$. Any $\mathrm{OWA}_\omega \in A_n(E,E)$ of order 2 is of the form*

$$\mathrm{OWA}_\omega(x) = \frac{1}{n}\sum_{i=1}^n x_i + \frac{1}{2}\,\frac{\omega_1 - \omega_n}{n-1}\sum_{i=1}^n (n+1-2\,i)\,x_{(i)}, \quad x \in E^n.$$

Proof. By (6.16), (6.17), (6.21) and (6.22), we immediately have

$$a_i = \frac{1}{n} - \frac{1}{2}(\omega_1 - \omega_n), \quad i \in N,$$

$$a_{ij} = \frac{\omega_1 - \omega_n}{n-1}, \quad i,j \in N,$$

and by (4.16), we have

$$\omega_t = a_i + (n-t)\,a_{ij} = \frac{1}{n} + \frac{\omega_1 - \omega_n}{2(n-1)}(n+1-2\,t), \quad i,j,t \in N,$$

which is sufficient. ■

6.3.3 Representation of boundary and monotonicity constraints

We address in this section what could be called an inverse problem. Suppose that a set function μ on N is given under its Möbius or interaction representation, denoted a and I, respectively. What are the conditions on the values a_T (or $I(T)$) so that μ is a fuzzy measure, satisfying boundary and monotonicity conditions? Concerning the Möbius representation, the answer has been presented in Proposition 4.1.1. For the Shapley interaction representation, the result is due to Grabisch [81, 83].

Proposition 6.3.3 *A set of 2^n coefficients $I_{\mathrm{Sh}}(T)$, $T \subseteq N$, corresponds to the Shapley interaction representation of a fuzzy measure if and only if*

$$\begin{array}{ll} i) & \displaystyle\sum_{T\subseteq N} B_t\, I_{\mathrm{Sh}}(T) = 0, \\ ii) & \displaystyle\sum_{i\in N} I_{\mathrm{Sh}}(i) = 1, \\ iii) & \displaystyle\sum_{T\subseteq N\setminus i} \beta^{|T|}_{|T\cap S|}\, I_{\mathrm{Sh}}(T\cup i) \geq 0, \quad \forall i\in N,\ \forall S\subseteq N\setminus i, \end{array}$$

where B_t is the t-th Bernoulli number (5.43) and β^l_k is defined by (5.60).

For the Banzhaf interaction representation, we present a comparable result, see also [83, 90].

Proposition 6.3.4 *A set of 2^n coefficients $I_{\mathrm{B}}(T)$, $T \subseteq N$, corresponds to the Banzhaf interaction representation of a fuzzy measure if and only if*

$$\begin{array}{ll} i) & \displaystyle\sum_{T\subseteq N} (-\frac{1}{2})^t\, I_{\mathrm{B}}(T) = 0, \\ ii) & \displaystyle\sum_{T\subseteq N} (\frac{1}{2})^t\, I_{\mathrm{B}}(T) = 1, \\ iii) & \displaystyle\sum_{T\subseteq N\setminus i} (\frac{1}{2})^t(-1)^{|T\setminus S|}\, I_{\mathrm{B}}(T\cup i) \geq 0, \quad \forall i\in N,\ \forall S\subseteq N\setminus i. \end{array}$$

Proof. Conditions $i)$ and $ii)$ immediately follow from the passage formula from I_{B} to v, see Table 5.4. For the condition $iii)$, we have, for all $i \in N$ and all $S \subseteq N \setminus i$,

$$\begin{array}{rcl} 0 & \leq & \mu_{S\cup i} - \mu_S \quad \text{(see (4.1))} \\ & = & \displaystyle\sum_{\substack{T\subseteq N\\ i\in T}} (\frac{1}{2})^t\, [(-1)^{|T\setminus(S\cup i)|} - (-1)^{|T\setminus S|}]\, I_{\mathrm{B}}(T), \quad \text{(see passage from } I_{\mathrm{B}} \text{ to } v) \\ & = & 2\displaystyle\sum_{\substack{T\subseteq N\\ i\in T}} (\frac{1}{2})^t\, (-1)^{|T\setminus S|-1}\, I_{\mathrm{B}}(T), \end{array}$$

which is sufficient. ∎

Now the following question arises: if a set of power and pairwise interaction indices is given, satisfying $\sum_i I(i) = 1$, what are the conditions on the $I(ij)$ such that the underlying fuzzy measure is a 2-order fuzzy measure?

In case of 2-order fuzzy measures, the conditions of Propositions 6.3.3 and 6.3.4 become respectively:

$$\begin{cases} I_{\mathrm{Sh}}(\emptyset) + \frac{1}{6} \sum_{\{i,j\} \subseteq N} I_{\mathrm{Sh}}(ij) = \frac{1}{2}, \\ \sum_{i \in N} I_{\mathrm{Sh}}(i) = 1, \\ I_{\mathrm{Sh}}(i) - \frac{1}{2} \sum_{j \in N \setminus (S \cup i)} I_{\mathrm{Sh}}(ij) + \frac{1}{2} \sum_{j \in S} I_{\mathrm{Sh}}(ij) \geq 0, \quad \forall i \in N, \ \forall S \subseteq N \setminus i, \end{cases}$$

and

$$\begin{cases} I_{\mathrm{B}}(\emptyset) + \frac{1}{4} \sum_{\{i,j\} \subseteq N} I_{\mathrm{B}}(ij) = \frac{1}{2}, \\ \sum_{i \in N} I_{\mathrm{B}}(i) = 1, \\ I_{\mathrm{B}}(i) - \frac{1}{2} \sum_{j \in N \setminus (S \cup i)} I_{\mathrm{B}}(ij) + \frac{1}{2} \sum_{j \in S} I_{\mathrm{B}}(ij) \geq 0, \quad \forall i \in N, \ \forall S \subseteq N \setminus i. \end{cases}$$

Thus the coefficients $I(ij)$ must fulfil only the monotonicity conditions.

6.3.4 Equivalence classes of fuzzy measures

In [81], Grabisch addressed the interesting problem of finding equivalence classes of fuzzy measures. Set $k \in N$ and suppose that the decision maker is able to give the interaction indices $I_{\mathrm{Sh}}(T)$ for $1 \leq t \leq k$. Then it is interesting to find all (general) fuzzy measures having the same specified $I_{\mathrm{Sh}}(T)$ values. Clearly, this defines an equivalence relation on the set of fuzzy measures on N.

Starting from the passage formula from I_{Sh} to μ (see (5.59)), we have, for all $S \subseteq N$,

$$\begin{aligned} \mu_S &= \sum_{T \subseteq N} \beta_{|T \cap S|}^{|T|} I_{\mathrm{Sh}}(T) \\ &= I_{\mathrm{Sh}}(\emptyset) + \sum_{\substack{T \subseteq N \\ 1 \leq t \leq k}} \beta_{|T \cap S|}^{|T|} I_{\mathrm{Sh}}(T) + \sum_{\substack{T \subseteq N \\ t > k}} \beta_{|T \cap S|}^{|T|} I_{\mathrm{Sh}}(T). \end{aligned}$$

Of course, $I_{\mathrm{Sh}}(\emptyset)$ has no clear meaning for the decision maker and we can eliminate it by using the first condition of Proposition 6.3.3, that is:

$$I_{\mathrm{Sh}}(\emptyset) = - \sum_{\substack{T \subseteq N \\ t \geq 1}} B_t \, I_{\mathrm{Sh}}(T).$$

Taking the other conditions into account leads to the following result.

Theorem 6.3.4 *All the fuzzy measures having the same $I_{\mathrm{Sh}}(T)$ for $1 \leq t \leq k$, with*

$$\sum_{i \in N} I_{\mathrm{Sh}}(i) = 1$$

are of the form:

$$\mu_S = \sum_{\substack{T \subseteq N \\ 1 \leq t \leq k}} \Big(\beta_{|T \cap S|}^{|T|} - B_t \Big) I_{\mathrm{Sh}}(T) + \sum_{\substack{T \subseteq N \\ t > k}} \Big(\beta_{|T \cap S|}^{|T|} - B_t \Big) I_{\mathrm{Sh}}(T), \quad S \subset N, \ S \neq \emptyset, N,$$

where the coefficients $I_{\rm Sh}(T)$ *with* $t > k$ *are free up to the monotonicity constraints:*

$$\sum_{\substack{T\ni i\\ t>k}} \beta^{|T|-1}_{|T\cap S|}\, I_{\rm Sh}(T) \geq - \sum_{\substack{T\ni i\\ 1\leq t\leq k}} \beta^{|T|-1}_{|T\cap S|}\, I_{\rm Sh}(T), \quad i\in N,\ S\subseteq N\setminus i.$$

Of course a similar result can be obtained for Banzhaf interaction indices.

Theorem 6.3.5 *All the fuzzy measures having the same* $I_{\rm B}(T)$ *for* $1\leq t\leq k$ *are of the form:*

$$\mu_S = \sum_{\substack{T\subseteq N\\ 1\leq t\leq k\\ |T\cap S|\ odd}} (-\frac{1}{2})^{t-1}\, I_{\rm B}(T) + \sum_{\substack{T\subseteq N\\ t>k\\ |T\cap S|\ odd}} (-\frac{1}{2})^{t-1}\, I_{\rm B}(T), \quad S\subset N,\ S\neq\emptyset,$$

where the coefficients $I_{\rm B}(T)$ *with* $t > k$ *are free up to the following constraints:*

$$\sum_{\substack{T\subseteq N\\ t>k}} (\frac{1}{2})^t I_{\rm B}(T) = 1 - \sum_{\substack{T\subseteq N\\ 1\leq t\leq k}} (\frac{1}{2})^t I_{\rm B}(T),$$

$$\sum_{\substack{T\ni i\\ t>k}} (-\frac{1}{2})^{t-1}(-1)^{|T\cap S|}\, I_{\rm B}(T) \geq - \sum_{\substack{T\ni i\\ 1\leq t\leq k}} (-\frac{1}{2})^{t-1}(-1)^{|T\cap S|}\, I_{\rm B}(T), \quad i\in N,\ S\subseteq N\setminus i.$$

As an illustration, let us search for the 2-order fuzzy measures having the same Shapley (or Banzhaf) indices $I(i)$ such that $\sum_{i\in N} I(i) = 1$. According to the previous results, these 2-order fuzzy measures are of the form:

$$\mu_S = \sum_{i\in S} I(i) - \frac{1}{2}\sum_{i\in S}\sum_{j\in N\setminus S} I(ij), \quad S\subset N,\ S\neq\emptyset, N,$$

where the $I(ij)$ are such that

$$\sum_{j\in N\setminus(S\cup i)} I(ij) - \sum_{j\in S} I(ij) \leq 2\, I(i), \quad i\in N,\ S\subseteq N\setminus i. \qquad (6.23)$$

Grabisch [81, Sect. 7] observed that the $n2^{n-1}$ inequality constraints (6.23) can be simply expressed by

$$-2\, I(i) \leq \pm\, I(i1) \pm I(i2) \pm \cdots \pm I(in) \leq 2\, I(i), \quad i\in N,$$

with all possible combinations of $+$ and $-$. Since this inequality is double, reversing systematically $+$ and $-$ on all positions leads to the same constraint so that the total number of constraints is divided by 2, thus is equal to 2^{n-2} for a given $I(i)$, and $n2^{n-2}$ for all the $I(i)$.

6.4 Identification of weights of interactive criteria

Once a particular aggregation operator has been chosen for aggregation, relying on specific properties for a particular application, it remains to identify the parameters of the chosen operator, if any. It will be the case in particular for any weighted operator, such as weighted means, OWA and fuzzy integrals.

In this section we will present three approches related to the Choquet integral. The problem thus consists in identifying the associated fuzzy measure[3]. The first approach is based only on a predefinite degree of orness. The second and the third methods are based on experimental data, i.e. examples given by a decision maker.

[3]Once obtained, this fuzzy measure may not be used in any aggregation procedure. Its nature depends on the aggregation operator used to construct it. Thus, following the approaches presented here, it will be always used in a Choquet integral.

6.4.1 Identification by parametric specification

In [138, 139], O'Hagan suggested a procedure to generate the OWA weights that have a predefinite degree of orness $\alpha \in [0, 1]$ and maximize the entropy. O'Hagan called them MEOWA operators. The approach suggested by O'Hagan is based on the solution of the following constrained optimization problem:

$$\text{maximize} \qquad -\sum_{i=1}^{n} \omega_i \ln \omega_i$$

$$\text{subject to} \quad \begin{cases} \dfrac{1}{n-1} \displaystyle\sum_{i=1}^{n} (i-1)\, \omega_i = \alpha \\ \displaystyle\sum_{i=1}^{n} \omega_i = 1 \\ \omega_i \geq 0, \quad i \in N \end{cases}$$

We note that by just specifying one parameter, the desired level of orness α, we uniquely get the weights. We can see that this approach is in the spirit of the maximum entropy techniques.

For a general Choquet integral associated to a fuzzy measure, an identification can be done in the same way as for OWA operators: select from among the Choquet integrals that attain a given level of orness the one whose fuzzy measure has the maximum dispersion (see (6.14), (6.3) and (4.1) for the formulas):

$$\text{maximize} \quad -\sum_{i=1}^{n} \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\, t!}{n!} [\delta_i\, \mu_{T \cup i}] \ln[\delta_i\, \mu_{T \cup i}]$$

$$\text{subject to} \quad \begin{cases} \dfrac{1}{n-1} \displaystyle\sum_{T \subsetneq N} \frac{(n-t)!\, t!}{n!} \mu_T = \alpha \\ \mu_\emptyset = 0 \\ \mu_N = 1 \\ \mu_{T \cup i} \geq \mu_T, \quad \forall i \in N,\ \forall T \subseteq N \setminus i \end{cases}$$

6.4.2 Identification based on learning data

Another approach to obtain the fuzzy measure is to learn the weights from data consisting of n-tuples of individual scores along with their aggregated value (see e.g. [180]). Suppose that (z_k, y_k), $k = 1, \ldots, l$ are such learning data where $z_k = (z_{k1}, \ldots, z_{kn})$ is an input vector, representing the profile of object k, and y_k is the global score of object k. Then, one can try to identify the best fuzzy measure μ which minimizes the total squared error of the model, i.e.

$$E^2 = \sum_{k=1}^{l} [\mathcal{C}_\mu(z_k) - y_k]^2.$$

under the boundary and monotonicity constraints of the fuzzy measure. It has been shown in [89, Chap. 10] that this problem can be put under a quadratic program form, that is

$$\begin{array}{rl} \text{minimize} & \frac{1}{2} u^t D u + c^t u \\ \text{subject to} & Au + b \geq 0 \end{array}$$

where u is a (2^n-2)-dimensional vector containing all the coefficients of the fuzzy measure μ (except $\mu_\emptyset$ and μ_N which are fixed), D is a (2^n-2)-dimensional square matrix, c a (2^n-2)-dimensional vector, A a $n(2^{n-1}-1)\times(2^n-2)$ matrix, and b a $n(2^{n-1}-1)$-dimensional vector. See [89] for an extensive study of such a quadratic program.

When the Choquet integral is reduced to an OWA, the problem can be simplified by taking advantage of the linearity with respect to the ordered arguments, see [61].

Notice that a major disadvantage of this model is that it requires a precise global score z_k for each given profile. In practical situation, such data are not always available. Moreover, no semantical specification about criteria is taken into account in this model.

6.4.3 Combination of learning data with semantical considerations

In this section, we propose a model allowing to identify the fuzzy measure on the basis of learning data consisting of a partial preorder over a reference set of alternatives (prototypes) whose profiles are known[4], and also from semantical considerations about criteria such as a partial preorder over the set of apparent weights related to each criterion, a partial preorder over interactions between pairs of criteria, and the knowledge of the sign of interactions between some pairs of criteria. As usually, we assume that all the partial scores x_i^a are given according to a same interval scale $E \subseteq \mathbb{R}$ (commensurability hypothesis) and that the Choquet integral is used for the aggregation. Moreover, to keep the model as simple as possible, we assume that the fuzzy measure is of order 2. We will see that the main advantage of our approach is that the fuzzy measure can be obtained simply by solving a linear program.

Apart from the partial ranking over the reference set of prototypes, we consider some semantical specifications about criteria:

- *Importance of criteria.* This can be properly done by giving a partial preorder on N, representing a ranking of the apparent weights μ_i, $i \in N$. One can also imagine that some exact values can be given.

- *Interaction between criteria.* The interaction index $I(ij) = a_{ij}$ is suitable for this. One can give a partial preorder on the set of pairs of criteria. The sign of each interaction a_{ij} can also be given, or even exact values.

- *Symmetric criteria.* Two criteria i and j are symmetric if they can be exchanged without changing the aggregation mode. Then $\mu_{T\cup i} = \mu_{T\cup j}$ for all $T \subseteq N \setminus ij$. This reduces the number of coefficients.

- *Degree of orness.* We have seen in (6.2) that orness$(\mathcal{C}_\mu)$ is a linear expression and thus might be integrated in our model. For a 2-order fuzzy measure, it becomes

$$\text{orness}(\mathcal{C}_\mu) = \frac{1}{2}\sum_{i\in N} a_i + \frac{1}{3}\,\frac{n-2}{n-1}\sum_{\{i,j\}\subseteq N} a_{ij}.$$

  A predefinite degree of orness could be demanded. More freely, one can only ask that the degree of orness lies in a given interval. For instance: orness$(\mathcal{C}_\mu) > 1/2$.

- *Veto and favor effects.* As for the degree of orness, the degrees of favor and veto are linear expressions and can be used.

[4]Contrary to the previous method, the global scores are not needed in this approach.

- *Dispersion.* The measure of dispersion could also be taken into account. Unfortunately, its expression contains non-linear factors making the model itself non-linear.

We thus suppose that we have at our disposal an expert or decision maker who is able to tell the relative importance of criteria, and the kind of interaction between them, if any. In fact, in practical applications the decision maker is able to give information on the apparent weights and interaction indices much more easily than to assess directly the values of the fuzzy measure. It is thus important to ask the decision maker the good questions that will allow to identify the fuzzy measure (elicitation from the decision maker).

Formally, the input data of the problem can be summarized as follows:

- The set A of alternatives and the set N of criteria,
- A table of individual scores (utilities) x_i^a given on a same interval scale $E \subseteq \mathbb{R}$,
- A partial preorder $\succeq_A$ on A (ranking of alternatives),
- A partial preorder $\succeq_N$ on N (ranking of criteria),
- A partial preorder $\succeq_P$ on the set of pairs of criteria (ranking of interaction indices),
- The sign of interactions between some pairs of criteria $a_{ij} : > 0, = 0, < 0$.

In addition to these data, any other information such that the degree of orness, veto and favor degrees, etc. can be integrated, provided that it can be given by a linear expression.

All these data can be formulated in terms of linear equalities or inequalities linking the unknown "weights" μ. The model then consists in finding a feasible 2-order fuzzy measure. Thus we are faced with a linear constraints satisfaction problem. Note that strict inequalities can be converted into vague inequalities by introducing a positive slack variable as the following immediate proposition shows.

Proposition 6.4.1 *$x \in \mathbb{R}^n$ is a solution of the linear system*

$$\begin{cases} \sum_{j=1}^{n} a_{ij}\, x_j \le b_i, & i = 1, \dots, p, \\ \sum_{j=1}^{n} c_{ij}\, x_j < d_i, & i = 1, \dots, q, \end{cases}$$

if and only if there exists $\varepsilon > 0$ such that

$$\begin{cases} \sum_{j=1}^{n} a_{ij}\, x_j \le b_i, & i = 1, \dots, p, \\ \sum_{j=1}^{n} c_{ij}\, x_j \le d_i - \varepsilon, & i = 1, \dots, q. \end{cases}$$

In particular, a solution exists if and only if the following linear program

$$\textit{maximize} \quad z = \varepsilon$$

$$\textit{subject to} \quad \begin{cases} \sum_{j=1}^{n} a_{ij}\, x_j \le b_i, & i = 1, \dots, p, \\ \sum_{j=1}^{n} c_{ij}\, x_j \le d_i - \varepsilon, & i = 1, \dots, q, \end{cases}$$

has an optimal solution $x^* \in \mathbb{R}^n$ with an optimal value $\varepsilon^* > 0$. In this case, x^* is a solution of the first system.

Thus, the problem of finding a 2-order fuzzy measure can be formalized with the help of a linear program. It is obvious that the poorer the input information, the bigger the solution set. Hence, it is desirable that the information is as complete as possible. However, if this information contains incoherences then the solution set could be empty[5].

Written in terms of the Möbius representation, a model for identifying weights could be given as follows:

$$\text{maximize } z = \varepsilon$$

subject to

$$\left.\begin{array}{ll} C(a) - C(b) \geq \delta + \varepsilon & \text{if } a \succ_A b \\ -\delta \leq C(a) - C(b) \leq \delta & \text{if } a \sim_A b \end{array}\right\} \text{partial semiorder with threshold } \delta$$

$$\left.\begin{array}{ll} a_i - a_j \geq \varepsilon & \text{if } i \succ_N j \\ a_i = a_j & \text{if } i \sim_N j \end{array}\right\} \text{ranking of criteria (apparent weights)}$$

$$\left.\begin{array}{ll} a_{ij} - a_{kl} \geq \varepsilon & \text{if } ij \succ_P kl \\ a_{ij} = a_{kl} & \text{if } ij \sim_P kl \end{array}\right\} \text{ranking of pairs of criteria (interactions)}$$

$$\left.\begin{array}{ll} a_{ij} \geq \varepsilon \text{ (resp. } \leq -\varepsilon) & \text{if } a_{ij} > 0 \text{ (resp. } < 0) \\ a_{ij} = 0 & \text{if } a_{ij} = 0 \end{array}\right\} \text{sign of some interactions}$$

$$\left.\begin{array}{ll} \sum_{i \in N} a_i + \sum_{\{i,j\} \subseteq N} a_{ij} = 1 & \\ a_i \geq 0 & \forall i \in N \\ a_i + \sum_{j \in T} a_{ij} \geq 0 & \forall i \in N,\ \forall T \subseteq N \setminus i \end{array}\right\} \begin{array}{l}\text{boundary and monotonicity} \\ \text{conditions (6.18)}\end{array}$$

$$C(a) = \sum_{i \in N} a_i\, x_i^a + \sum_{\{i,j\} \subseteq N} a_{ij}\, [x_i^a \wedge x_j^a] \quad \forall a \in A \ \Big\} \text{definition of } \mathcal{C}_\mu$$

It seems natural to assume that the ranking over A is translated into a partial semiorder over the set of the global evaluations given by the Choquet integral. This partial semiorder has a fixed preference threshold δ, which can be tuned as wished. Such a threshold level should be reached by the difference between global scores to consider that one object should be significantly preferred to another object.

Let us comment on the scale used to define the utilities. Since the Choquet integral is stable under the same admissible transformations of interval scales, using utilities on a $[0, 100]$ scale or on $[-2, 3]$ scale has no influence on the ranking of alternatives. Now suppose that the utilities x_i^a are defined in $E = [0, 1]$. Changing this scale into $[p, q]$, with $p < q$, and translating the utilities in the appropriate way amounts to replacing only the first set of constraints by

$$C(a) - C(b) \geq \delta + \frac{\varepsilon}{q - p} \quad \text{if } a \succ_A b. \tag{6.24}$$

It is clear that if ε is missing in the other constraints then the optimal solution of the linear program depends on no scale transformation. Otherwise, if ε is present both in constraints (6.24) and the other constraints then the optimal solution can be sensitive to any scale transformation,

[5]An empty solution set could be due to an incompatibility between the given information and the assumption that the fuzzy measure is of order 2. In this case, it can be useful to consider a 3-order fuzzy measure or, if necessary, a fuzzy measure of higher order.

although the feasibility of the system remains unaltered. In this case, since $a_i \in [0,1]$ and $a_{ij} \in [-1,1]$ for all $i,j \in N$, we necessarily have $\varepsilon \in [0,2]$. We then can make the following theoretical observations:

- If $q-p$ is large (for example $q-p=100$) then $\varepsilon/(q-p)$ is small and the global evaluations $C(a)$ have good chances to be not very contrasted in the optimal solution.
- If $q-p$ is small (for example $q-p=0.01$) then, multiplying inequation (6.24) by $q-p$, we see that the optimal value ε^* will be small too; this implies that the weights a_i or the interactions a_{ij} will be not very contrasted.

Taking these facts into account, it seems that a reasonable compromise would be to take $q-p=1$ and hence to define the utilities on the unit interval $[0,1]$.

The value of the threshold δ must also be chosen carefully. Indeed, the more δ is large the more ε^* will be small. A too large δ can even make the program infeasible. We will not suggest any rule to fix δ. It is better to compare the solutions obtained with different values of δ.

Let us make a last observation. It might happen that some equalities are very constraining and make the program infeasible. In this case, these equalities can be relaxed into some strict inequalities in a coherent way. For example, the equality in the conditions $a_1 = a_2 > a_3$ could be relaxed into the following strict inequalities:

$$\left.\begin{matrix} a_1 \\ a_2 \end{matrix}\right\} > a_3 \quad \text{and} \quad |a_1 - a_2| < \left\{\begin{matrix} |a_1 - a_3| \\ |a_2 - a_3| \end{matrix}\right.$$

that is

$$\begin{cases} a_1 > a_3 \\ a_2 > a_3 \\ 2\,a_1 - a_2 - a_3 > 0 \\ -a_1 + 2\,a_2 - a_3 > 0 \end{cases}$$

In order to illustrate the model, we now present three small examples. The first two are constructed in such a way that no weighted arithmetic mean can be used as utility function.

Example 6.4.1 Consider the problem of ranking cooks on the basis of their capacity for preparing three dishes: frogs'legs (FL), steak tartare (ST), and stuffed clams (SC). Four cooks a, b, c, d acting as prototypes are evaluated as follows[6] (marks are multiplied by 20):

cook	FL	ST	SC
a	18	15	19
b	15	18	19
c	15	18	11
d	18	15	11

The decision maker is asked to express its advice by giving a ranking over $A = \{a,b,c,d\}$. Of course, he/she immediately suggests $a \succ_A d$ and $b \succ_A c$. However, these preferences do not contribute to anything since they naturally follow from the monotonicity of the Choquet integral. The decision maker realizes that the other comparisons are not so obvious since the associated profiles interlace. He/she then proposes the following reasoning: when a cook is renowned for his stuffed clams, it is preferable that he/she is also better in cooking frogs'legs than steak tartare,

[6] This example is similar to Exemple 1.3.1. But here commensurable partial utilities are available.

so $a \succ_A b$. However, when a cook badly prepares stuffed clams, it is more important that he/she is better in preparing steak tartare than frogs'legs, and so $c \succ_A d$.

Now, the question arises: does there exist an additive model leading to this partial ranking? Let $\omega_1, \omega_2, \omega_3$ represent the weights of criteria FL, ST, SC respectively. By using the weighted arithmetic mean as utility function, we obtain:

$$\begin{aligned} a \succ_A b \quad &\Leftrightarrow \quad \omega_1 > \omega_2, \\ c \succ_A d \quad &\Leftrightarrow \quad \omega_1 < \omega_2. \end{aligned}$$

We immediately observe that no weighted arithmetic mean can yield the proposed ranking. This is not surprising since clearly the criteria are not mutually preferentially independent. Thus, it is necessary to take into account the interactions between criteria. Notice that, even in this case, the Shapley integral cannot be used, since it is also a weighted arithmetic mean.

By extending the weighted arithmetic mean to the 2-order Choquet integral (6.15), we are led to the following conditions:

$$\begin{aligned} a \succ_A b \quad &\Leftrightarrow \quad 0.15\, a_1 - 0.15\, a_2 + 0.15\, a_{13} - 0.15\, a_{23} > 0, \\ b \succ_A c \quad &\Leftrightarrow \quad 0.4\, a_3 + 0.2\, a_{13} + 0.35\, a_{23} > 0, \\ c \succ_A d \quad &\Leftrightarrow \quad -0.15\, a_1 + 0.15\, a_2 > 0. \end{aligned}$$

Of course, these three conditions imply $a \succ_A d$.

It is clear that the problem has now at least one solution, which can be given by an optimal solution of the following linear program ($N = \{1, 2, 3\}$):

maximize $z = \varepsilon$

subject to

$$\begin{cases} 0.15\, a_1 - 0.15\, a_2 + 0.15\, a_{13} - 0.15\, a_{23} \geq \delta + \varepsilon & \\ 0.4\, a_3 + 0.2\, a_{13} + 0.35\, a_{23} \geq \delta + \varepsilon & \\ -0.15\, a_1 + 0.15\, a_2 \geq \delta + \varepsilon & \\ a_1 + a_2 + a_3 + a_{12} + a_{13} + a_{23} = 1 & \\ a_i \geq 0 & i \in N \\ a_i + a_{ij} \geq 0 & i, j \in N \\ a_i + a_{ij} + a_{ik} \geq 0 & i, j, k \in N. \end{cases}$$

Using an appropriate software, the following solution was obtained (for $\delta = 0.05$ fixed):

- Objective function: $\varepsilon = 0.025$ (note that $\varepsilon = 0$ if $\delta = 0.075$)
- Apparent weights a_i ($= \mu_i$) and real weights $I(i)$:

	FL	ST	SC
a_i	0	0.5	0.5
$I(i)$	0.25	0.25	0.5

- Interaction indices a_{ij}:

	ST	SC
FL	0	0.5
ST		-0.5

- The global evaluations $C = \mathcal{C}_\mu$:

	a	b	c	d
$C(\cdot)$	0.925	0.85	0.725	0.65
$20\,C(\cdot)$	18.5	17	14.5	13

Of course, we must be very cautious when one wants to get general conclusions only from the obtained solution. Indeed, it is not clear at all that there is no solution such that $a_{23} > 0$. All that the model does is to find a 2-order fuzzy measure that is coherent with the available information. Thus, the interpretation of such a solution could be irrelevant.

Regarding the degree of orness of the solution, we immediately see that orness$(\mathcal{C}_\mu) = 0.5$. Thus we are in presence of an equitable decision maker. This is due to the visible symmetry of the problem.

Example 6.4.2 Consider the problem of the evaluation of students in an institute of Mathematics with respect to three subjects: linear algebra (Al), calculus (Ca), and statistics (St). Suppose that the institute is oriented towards statistics. More precisely, suppose that the decision maker suggests the following partial ranking of the criteria:

$$\text{St} \succ_N \begin{cases} \text{Al} \\ \text{Ca} \end{cases}$$

For instance, the weights of subjects could be proportional to 2, 2 and 3, respectively.

Three students a, b, c have been evaluated as follows (marks are multiplied by 20):

student	Al	Ca	St	global evaluation (weighted arithmetic mean)
a	12	12	19	15
b	16	16	15	15.57143
c	19	19	12	16

The decision maker then reasons as follows: if a student is excellent at statistics (mark of at least 18) then he/she is excellent, whatever the marks obtained elsewhere. However, if he/she is not excellent at statistics then it is necessary to take into account the mark obtained in the other courses.

Consequently, on the basis of the available profiles, the decision maker proposes the following ranking:

$$a \succ_A c \succ_A b.$$

Let us show that the additive model is not appropriate for this example. Let $\omega_1, \omega_2, \omega_3$ represent the weights of criteria Al, Ca, St respectively. By using the weighted arithmetic mean as utility function, we get:

$$\begin{aligned} a \succ_A c \quad &\Leftrightarrow \quad \omega_1 + \omega_2 - \omega_3 < 0, \\ c \succ_A b \quad &\Leftrightarrow \quad \omega_1 + \omega_2 - \omega_3 > 0. \end{aligned}$$

Now, let us turn to the 2-order model, which takes into account the interactions between pairs of criteria. On this matter, the decision maker has observed that there is some overlap between statistics and calculus courses. Hence, we may assume that there exists a negative interaction between them:

$$a_{\{\mathrm{St},\mathrm{Ca}\}} < 0.$$

Putting all these data into a linear program as explained above, and then solving it with the help of a software, the following solution was obtained (for $\delta = 0.025$ fixed):

- Objective function: $\varepsilon = 0.055769$
- Apparent weights a_i and real weights $I(i)$:

	Al	Ca	St
a_i	0.7135	0.05577	1
$I(i)$	0.3567	0.02788	0.6154

- Veto and favor degrees:

	Al	Ca	St
veto(i)	0.23606	0.07162	0.42305
favor(i)	0.79907	0.47020	1

- Interaction indices a_{ij}:

	Ca	St
Al	0	-0.7135
Ca		-0.05577

- The global evaluations $C = \mathcal{C}_\mu$:

	a	b	c
$C(\cdot)$	0.95	0.7885	0.8692
$20\,C(\cdot)$	19	15.77	17.384

For this solution, we observe an absolute dispersion of 0.38242 (cf. (6.14)), that is a relative dispersion of $0.38242/\ln 3 = 0.34810$. This small value shows that few marks are really taken into account in the aggregation phase. Such a phenomenon is compatible with the fact that criterion St is a favor ($a_3 = 1$ and favor(3) $= 1$).

Moreover, we have orness($\mathcal{C}_\mu$) $= 0.75642$ and the aggregation is of disjunctive type. This is not surprising since the rules given in the problem globally favor the students, although criterion St was not *a priori* considered as a favor in the sense of (6.5).

Suppose that to avoid such a favor effect, the decision maker proposes a degree of orness less than 0.6. Introducing the additional constraint[7]

$$\frac{1}{2}\sum_{i\in N} a_i + \frac{1}{6}\sum_{\{i,j\}\subseteq N} a_{ij} \leq 0.6$$

in the linear program yields the following solution (for $\delta = 0.025$ fixed):

[7] Of course, this value 0.6 has no precise signification. Our purpose is simply to observe the behavior of the solution when lowering the degree of tolerance.

- Objective function: $\varepsilon = 0.025909$
- Apparent weights a_i and real weights $I(i)$:

	Al	Ca	St
a_i	0.4589	0.02591	0.8152
$I(i)$	0.3219	0.1054	0.5727

- Veto and favor degrees:

	Al	Ca	St
veto(i)	0.38212	0.27387	0.54399
favor(i)	0.60065	0.38416	0.8152

- Interaction indices a_{ij}:

	Ca	St
Al	0.1848	-0.4589
Ca		-0.02591

- The global evaluations $C = \mathcal{C}_\mu$:

	a	b	c
$C(\cdot)$	0.8853	0.7835	0.8344
$20\,C(\cdot)$	17.706	15.67	16.688

For this new solution, the orness is 0.6 and the score of student a went down from 19 to 17.7. Also, criterion St is no more a favor (favor(3) = 0.8152) and the absolute dispersion increased up to 0.73056 (relative dispersion = 0.66499).

Example 6.4.3 (a real application) Consider the problem of the evaluation of 5 trainees learning to drive military tanks[8]. Trainees are evaluated by instructors according to 4 criteria: precision firing (1), swiftness of target detection (2), path choice (3), and communication (4). The available ratings (defined in $[0, 1]$) are given in the following table:

name	crit. 1	crit. 2	crit. 3	crit. 4
Arthur (A)	1	1	0.75	0.25
Lancelot (L)	0.75	0.75	0.75	0.75
Yvain (Y)	1	0.625	0.50	1
Perceval (P)	0.25	0.50	0.75	0.75
Erec (E)	0.375	1	0.50	0.75

The following rankings are then proposed by the expert (instructor):

$$L \succ_A Y \succ_A E \succ_A P \succ_A A$$
$$4 \succ_N \left\{ \begin{matrix} 1 \\ 2 \end{matrix} \right\} \succ_N 3$$
$$\left. \begin{matrix} 14 \\ 24 \end{matrix} \right\} \succ_P 0 \succ_P 12.$$

For $\delta = 0.05$, we obtain the following solution:

[8]This example was proposed by Grabisch (Thomson-CSF, Central Research Laboratory, Orsay, France). It is given in more details in a forthcoming paper [94].

- Objective function: $\varepsilon = 0.03890$ (note that $\varepsilon = 0$ if $\delta = 0.1$)
- Apparent weights a_i and real weights $I(i)$:

	crit. 1	crit. 2	crit. 3	crit. 4
a_i	0.0389	0.0389	0	0.0778
$I(i)$	0.1984	0.1755	0.22185	0.40425

- Veto and favor degrees:

	crit. 1	crit. 2	crit. 3	crit. 4
veto(i)	0.68869	0.66833	0.71818	0.86302
favor(i)	0.24251	0.23233	0.24429	0.34264

- Interaction indices a_{ij}:

	crit. 2	crit. 3	crit. 4
crit. 1	-0.0389	0.2304	0.1275
crit. 2		0	0.3121
crit. 3			0.2133

- The global evaluations $C = \mathcal{C}_\mu$:

	A	L	Y	P	E
$C(\cdot)$	0.3944	0.75	0.6611	0.4833	0.5722

For this solution we observe a degree of orness of 0.26544 and the instructor is conjunctive oriented. This is consistent with the rather high degrees of veto obtained above. Such a behavior can be a good thing: in high technology, it is often necessary to demand that most of criteria are satisfied.

Moreover, we observe an absolute dispersion of 0.99323, that is a relative dispersion of $0.99323/\ln 4 = 0.71646$.

We also observe that the global evaluation of Lancelot (0.75) could be predicted: it follows from the compensativeness (Comp) of the Choquet integral.

Another approach to identifying the 2-order fuzzy measure consists in considering ε as a variable parameter and maximizing the entropy over the same constraints, giving rise to a parametric convex programming problem. Since the entropy is a strictly convex function, the optimal solution obtained for each admissible value of ε is unique. Then, it suffices to choose, after optimization, an appropriate value of ε in such a way that the entropy is not too low and at the same time that the contrast wished in the solution be attained. Note that such an approach has yet to be studied.

6.5 Aggregation of scores defined on ordinal scales

Thus far, we have investigated aggregation of cardinal data. But what about aggregation of profiles defined on ordinal scales? The Choquet integral is no longer suitable since in general it does not fulfil (CM). In this section, we attempt to bring some solutions to this aggregation problem.

6.5.1 The commensurability assumption

As for cardinal scales, we can easily see that aggregating values defined on independent ordinal scales often leads to a dictatorial aggregation. For instance, Kim [106, Corollary 1.2] showed that the non-constant operators $M \in A_n(\mathbb{R}, \mathbb{R})$ satisfying (Co, CMIS) are strongly equivalent to a dictator x_k (see Theorem 3.4.17).

Instead of considering independent ordinal scales, described by admissible transformations $\phi_i \in \Phi(E)$ on each criterion i, we shall assume that all the partial scores x_i^a are defined according to a same ordinal scale, meaning that all the scales are commensurable. In particular, this means that the equivalence

$$x_i^a \leq x_j^b \iff \phi_i(x_i^a) \leq \phi_j(x_j^b)$$

holds for any pair of alternatives $a, b \in A$ and any pair of criteria $i, j \in N$. In this case, we can immediately see that $\phi_1 = \cdots = \phi_n = \phi$ and the aggregation operators must fulfil (CM).

6.5.2 Continuous or non-continuous operators?

We have seen in Section 3.4.2 that non-constant operators $M \in A_n([a,b], \mathbb{R})$ fulfilling (Co, CM) are strongly equivalent to a Boolean max-min function $\mathrm{B}_c^{\vee\wedge}$ (see Theorem 3.4.15). Moreover, those operators $M \in A_n(E, \mathbb{R})$ that fulfil (Co, Id, CM) are exactly the Boolean max-min functions (see Theorem 3.4.12). (Id) merely ensures that the scale of the global scores is the same that the one of partial scores.

Thus, when continuity is assumed, the functions $\mathrm{B}_c^{\vee\wedge}$ seem to be the only suitable operators for aggregation of ordinal values. We have seen in Section 4.4.1 that $\mathrm{B}_c^{\vee\wedge}$ is both a Choquet and Sugeno integral with respect to a 0-1 fuzzy measure. Unfortunately, it seems impossible to express the weights of criteria from such a 0-1 fuzzy measure.

Perhaps the continuity property is responsible for this deadlock. In fact, non-continuous operators M fulfilling (CM) are still to be characterized. However, we already know that some of them are pathological in the sense that they are Boolean max-min functions almost everywhere in the definition set (see Section 3.4.2). Note also that the class of Boolean max-min functions on $[a,b]^n$ have been characterized as the operators fulfilling (In, Id, CM'), see Theorem 3.4.20.

The problem of aggregating ordinal values seems to be a difficult one and has still to be investigated in details. We propose here a solution for the case of independent criteria. Assume that the weights $\omega_1, \ldots, \omega_n$ are rational numbers given according to a ratio scale. Then there exist $p_i, q \in \mathbb{N}$, $q \neq 0$, $\sum_i p_i = q$ such that

$$p_i = q\,\omega_i.$$

Then, starting with the median operator, which is a particular $\mathrm{B}_c^{\vee\wedge}$ (see Section 4.4.3), we can weight this operator in the following way[9]:

$$\mathrm{median}_\omega(x) = \mathrm{median}(p_1 \odot x_1, \ldots, p_n \odot x_n), \quad x \in E^n.$$

Note that this process has already been presented for quasi-arithmetic means, see Section 3.2.1.

When a cardinal fuzzy measure is given on the criteria, it seems preferable to use the relational approach presented in Section 1.3.2.

[9] Of course, when an even number of arguments is considered, the result can be a pair or an interval.

6.5.3 Commensurability between the partial scores and the fuzzy measure

Consider the very particular case where the fuzzy measure given on criteria is ordinal in nature and whose values can be compared with the partial scores. In this case, the x_i^a's and the μ_T's belong to a same ordinal scale. This means that μ_T can no longer be interpreted as the weight of coalition T, but rather as a threshold or an aspiration degree. For instance, in the Sugeno integral

$$\mathcal{S}_\mu(x) = \bigvee_{T\subseteq N} \Big[\mu_T \wedge (\bigwedge_{i\in T} x_i)\Big],$$

if the score $\bigwedge_{i\in T} x_i$ is less than the threshold μ_T then $\bigwedge_{i\in T} x_i$ is considered in the aggregation. Otherwise, if $\bigwedge_{i\in T} x_i$ is greater than μ_T then the threshold is attained and μ_T can be considered in the aggregation.

Such a situation also occurs in decision under uncertainty when only qualitative (ordinal) information is available, see e.g. [41, 51, 52].

Definition 6.5.1 (CMCFM) $M_\mu \in A_n([0,1],\mathbb{R})$ is *comparison meaningful for ordinal values with commensurable fuzzy measure scale* if, for all $\phi \in \Phi([0,1])$, all $x, x \in [0,1]^n$ and all fuzzy measures μ, μ' on N, we have

$$M_\mu(x) \leq M_{\mu'}(x') \iff M_{\phi(\mu)}(\phi(x)) \leq M_{\phi(\mu')}(\phi(x')).$$

We now show that, under continuity, the Sugeno integrals are the only suitable aggregation operators in this context.

Theorem 6.5.1 $M_\mu \in A_n([0,1],\mathbb{R})$ *depends continuously on a fuzzy measure* μ *on* N *and fulfils (Co, Id, CMCFM) if and only if* $M_\mu = \mathcal{S}_\mu$.

Proof. (Sufficiency) Trivial.

(Necessity) It is clear that M_μ can be considered as a function of the x_i's and the μ_T's ($T \neq \emptyset, N$), that is a function with $n + 2^n - 2$ arguments:

$$M_\mu(x) = M'(x_1,\ldots,x_n;\mu_1,\ldots,\mu_{N\setminus 1}), \quad x \in [0,1]^n.$$

This function fulfils (Co, Id, CM) on $[0,1]^{n+2^n-2}$. By Theorem 3.4.12, there exists a set function $c' : 2^{\{1,\ldots,n+2^n-2\}} \to \{0,1\}$ such that

$$M'(x;\mu) = \mathrm{B}_{c'}^{\vee\wedge}(x;\mu), \quad \forall x\,\forall\mu,$$

that is

$$M'(x;\mu) = \bigvee_{T_1\subseteq N} \bigvee_{\substack{T_2\subseteq 2^N\\ T_2\neq\emptyset,N}} c'_{T_1,T_2}\Big[\Big(\bigwedge_{k\in T_1} x_k\Big)\wedge\Big(\bigwedge_{K\subseteq T_2}\mu_K\Big)\Big], \quad \forall x\,\forall\mu.$$

Since M_μ fulfils (Id), we have $M'(0;\mu) = 0$ for all μ, that is

$$\bigvee_{\substack{T_2\subseteq 2^N\\ T_2\neq\emptyset,N}} c'_{\emptyset,T_2}\Big(\bigwedge_{K\subseteq T_2}\mu_K\Big) = 0, \quad \forall\mu,$$

and there is no 'term' independent of variables x_i in $M'(x;\mu)$. Now, we can easily see that M_μ fulfils (In, SMin, SMax), and we then can conclude by Theorem 4.3.3. ∎

Notes

1. The Shapley integral seems to be a previously unknown aggregation operator. Although its use is not suitable when interacting criteria are considered, it deserves to be characterized. Such a characterization has yet to be done.

2. The concept of degree of disjunction is due to Dujmovic [54, 55]. In fact, he only applied his definition to the particular case of root-power means. On the other hand, Yager [192] introduced independently the degree of orness of an OWA in an intuitive way. In this chapter, we have adapted the definition of Dujmovic to the Choquet integral and then observed that the degree of orness of an OWA is actually its degree of disjunction. This enabled us to establish a connection between totally independent papers. We then interpreted this concept as a degree of tolerance.

3. The definitions of veto and favor criteria have been proposed by Dubois and Koning [42] and in a revisited way by Grabisch [83]. Since such criteria occur rarely in applications, we have proposed to extend these concepts to degrees of veto and favor for any criterion. A first attempt was done using probabilities, and another by axiomatic characterization.

4. The concept of entropy defined on a probability measure is well known and its use in MCDM was proposed by Yager [192] for the OWA operators. In this chapter we have properly generalized this concept to a general fuzzy measure. This generalization has yet to be characterized.

Chapter 7

Approximations of set functions

When a function describing some complicated relationship is given, one often seeks to replace it by a simpler functional form, usually linear, which approximates the given one. In this final chapter, we intend to study this sort of operation for general set functions, fuzzy measures and Choquet integrals.

Following common practice in regression analysis, we shall use the least squares criterion to choose a best approximation. In a multicriteria decision problem modelled by a Choquet integral $\mathcal{C}_\mu$, it could be interesting to find the best k-order fuzzy measure $\mu^{(k)}$ that minimizes

$$\int_{[0,1]^n} [\mathcal{C}_\mu(x) - \mathcal{C}_{\mu^{(k)}}(x)]^2 \, dx. \tag{7.1}$$

Note however that the distance used here could be discussed; indeed, another way for approximating a given fuzzy measure consists in minimizing

$$\sum_{T \subseteq N} [\mu_T - \mu_T^{(k)}]^2. \tag{7.2}$$

Thus, the choice of the distance should be dictated by the nature of the decision problem taken in consideration. For instance, it is assumed in (7.1) that all the profiles in $[0,1]^n$ have the same importance and are uniformly distributed.

Note also that other methods of approximation can be used, as for example the so-called upper and lower approximations, see e.g. Dubois and Prade [48, 50] and Grabisch [85]. Although they are very interesting, we will not approach these problems here.

In Section 7.1 we present the problem of approximations of pseudo-Boolean functions which have been introduced and partially solved by Hammer and Holzman [95]. In Section 7.2, we investigate the approximations of Lovász extensions. Finally, in Section 7.3, we apply the results obtained to the problems (7.1) and (7.2). The links between the linear approximations ($k = 1$) and the Shapley power indices are also studied.

7.1 Approximations of pseudo-Boolean functions

Hammer and Holzman [95] investigated the approximation of a pseudo-Boolean function by a multilinear polynomial of (at most) a specified degree. According to them, fixing $k \in \mathbb{N}$ with $k \leq n$, the best k-th approximation of f is the multilinear polynomial $f^{(k)} : \{0,1\}^n \to \mathbb{R}$ of degree $\leq k$ defined by

$$f^{(k)}(x) = \sum_{\substack{T\subseteq N\\ t\leq k}} a_T^{(k)} \prod_{i\in T} x_i$$

which minimizes

$$\sum_{x\in\{0,1\}^n} [f(x) - f^{(k)}(x)]^2$$

among all multilinear polynomials of degree $\leq k$. They proved that the best k-th approximation $f^{(k)}$ is given by the unique solution $\{a_T^{(k)}\,|\,T\subseteq N, t\leq k\}$ of the triangular linear system:

$$\frac{1}{2^n}\sum_{x\in\{0,1\}^n} \Delta_S f^{(k)}(x) = \frac{1}{2^n}\sum_{x\in\{0,1\}^n} \Delta_S f(x), \quad \forall S\subseteq N,\ \ s\leq k. \tag{7.3}$$

They also solved this system for $k=1$ and $k=2$. In this section, we intend to solve the system for any $k\leq n$.

Let $I_{\mathrm{B}}^{(k)}$ be the Banzhaf interaction index related to $f^{(k)}$. By (5.25), the system (7.3) can be written as

$$I_{\mathrm{B}}^{(k)}(S) = I_{\mathrm{B}}(S), \quad \forall S\subseteq N,\ \ s\leq k, \tag{7.4}$$

and by (5.12), it becomes

$$\sum_{\substack{T\supseteq S\\ t\leq k}} (\frac{1}{2})^{t-s} a_T^{(k)} = I_{\mathrm{B}}(S), \quad \forall S\subseteq N,\ \ s\leq k.$$

In particular, we have $a_S^{(k)} = I_{\mathrm{B}}(S)$ for all $S\subseteq N$ such that $s=k$. Hammer and Holzman [95, Sect. 3] obtained this result for $k=1$: $a_i^{(1)} = \phi_{\mathrm{B}}(i)$ for all $i\in N$.

By (5.12), we observe immediately that $I_{\mathrm{B}}^{(k)}(S)=0$ for all $S\subseteq N$ such that $s>k$. Hence, by using the passage formula from I_{B} to a, we have

$$a_S^{(k)} = \sum_{\substack{J\supseteq S\\ j\leq k}} (-\frac{1}{2})^{j-s} I_{\mathrm{B}}^{(k)}(J), \quad \forall S\subseteq N,\ \ s\leq k.$$

The system (7.4) then becomes

$$a_S^{(k)} = \sum_{\substack{J\supseteq S\\ j\leq k}} (-\frac{1}{2})^{j-s} I_{\mathrm{B}}(J), \quad \forall S\subseteq N,\ \ s\leq k,$$

and by (5.12), we have, for all $S\subseteq N$ with $s\leq k$,

$$\begin{aligned}
a_S^{(k)} &= \sum_{\substack{J\supseteq S\\ j\leq k}} (-\frac{1}{2})^{j-s} \sum_{T\supseteq J} (\frac{1}{2})^{t-j} a_T \\
&= \sum_{T\supseteq S} (\frac{1}{2})^{t-s} a_T \sum_{\substack{J:S\subseteq J\subseteq T\\ j\leq k}} (-1)^{j-s} \\
&= \sum_{T\supseteq S} (\frac{1}{2})^{t-s} a_T \sum_{j=s}^{\min(k,t)} \binom{t-s}{j-s} (-1)^{j-s}.
\end{aligned}$$

However, we have

$$\sum_{j=s}^{\min(k,t)} \binom{t-s}{j-s}(-1)^{j-s} = \begin{cases} (1-1)^{t-s}, & \text{if } t \leq k, \\ (-1)^{k-s}\binom{t-s-1}{k-s}, & \text{if } t > k \text{ (use induction over } k \geq s). \end{cases}$$

Therefore, we obtain an explicit formula for $a_S^{(k)}$:

$$a_S^{(k)} = a_S + (-1)^{k-s} \sum_{\substack{T \supseteq S \\ t>k}} \binom{t-s-1}{k-s} (\frac{1}{2})^{t-s} a_T, \quad \forall S \subseteq N, \ \ s \leq k. \tag{7.5}$$

Some particular cases are shown in Table 7.1. We thus retrieve the solutions obtained by Hammer and Holzman for $k = 1$ and $k = 2$.

$$\begin{array}{rcll}
a_\emptyset^{(0)} & = & \displaystyle\sum_{T \subseteq N} \frac{1}{2^t}\, a_T & \\[2ex]
a_\emptyset^{(1)} & = & \displaystyle\sum_{T \subseteq N} \frac{-(t-1)}{2^t}\, a_T & \\
a_i^{(1)} & = & \displaystyle\sum_{T \ni i} \frac{1}{2^{t-1}}\, a_T, & i \in N \\[2ex]
a_\emptyset^{(2)} & = & \displaystyle\sum_{T \subseteq N} \frac{(t-1)(t-2)}{2^{t+1}}\, a_T & \\
a_i^{(2)} & = & \displaystyle\sum_{T \ni i} \frac{-(t-2)}{2^{t-1}}\, a_T, & i \in N \\
a_{ij}^{(2)} & = & \displaystyle\sum_{T \ni i,j} \frac{1}{2^{t-2}}\, a_T, & \{i,j\} \subseteq N \\[2ex]
a_S^{(n-1)} & = & a_S - (-\frac{1}{2})^{n-s}\, a_N, & S \subseteq N, \ \ s \leq n-1
\end{array}$$

Table 7.1: Coefficients of the best k-th approximation for some values of k

7.2 Approximations of Lovász extensions

7.2.1 Definition and computation

Adopting an approach similar to the one of Hammer and Holzman [95], we define the problem of approximations of Lovász extensions as follows.

Definition 7.2.1 Let $\hat{f}$ be the Lovász extension of a pseudo-Boolean function f, and let k be an integer, $0 \leq k \leq n$. The best k-th approximation of $\hat{f}$ is the min-polynomial $\hat{f}^{(k)} : [0,1]^n \to \mathbb{R}$ of degree $\leq k$ defined by

$$\hat{f}^{(k)}(x) = \sum_{\substack{T \subseteq N \\ t \leq k}} a_T^{(k)} \bigwedge_{i \in T} x_i \tag{7.6}$$

which minimizes

$$\int_{[0,1]^n} [\hat{f}(x) - \hat{f}^{(k)}(x)]^2 \, dx$$

among all min-polynomials of degree $\leq k$. We write $\hat{f}^{(k)} = A^{(k)}(\hat{f})$.

When $k = 0$, the definition gives the best constant approximation; when $k = 1$, it gives the best linear approximation; when $k = 2$, it gives the best min-quadratic approximation; finally, every $\hat{f}$ is its own best n-th approximation.

In this section we intend to compute $A^{(k)}(\hat{f})$ for any $k \leq n$. The result will be presented in Theorem 7.2.1 below.

Observe first that the set $V^{(n)}$ of all min-polynomials $\hat{f}$ of the form

$$\hat{f}(x) = \sum_{T \subseteq N} a_T \bigwedge_{i \in T} x_i, \quad x \in [0,1]^n,$$

is a linear (or vector) space isomorphic to $\mathbb{R}^{2^n}$: indeed, any function $\hat{f} \in V^{(n)}$ can be identified with the vector listing its 2^n coefficients a_T in $\mathbb{R}^{2^n}$ (assuming a fixed ordering of the elements of 2^N). In the sequel, we shall often make this identification.

Moreover, it is clear that the binary operation $\langle \cdot\,, \cdot \rangle : V^{(n)} \times V^{(n)} \to \mathbb{R}$ defined by

$$\langle \hat{f}_1, \hat{f}_2 \rangle := \int_{[0,1]^n} \hat{f}_1(x) \, \hat{f}_2(x) \, dx$$

is a scalar product in $V^{(n)}$. This scalar product allows to define a norm $\|\hat{f}\| := \langle \hat{f}, \hat{f} \rangle^{1/2}$ and then a distance $d(\hat{f}_1, \hat{f}_2) := \|\hat{f}_1 - \hat{f}_2\|$ in $V^{(n)}$.

Now, for any $k \in \mathbb{N}$ with $k \leq n$, define $V^{(k)}$ as the subset of all min-polynomials $\hat{f}^{(k)}$ of degree $\leq k$, i.e. of the form (7.6). Thus, $V^{(k)}$ is a linear subspace in $V^{(n)}$ and one of its bases can be given by

$$B^{(k)} = \Big\{ \bigwedge_{j \in S} x_j \;\Big|\; S \subseteq N, s \leq k \Big\}.$$

In particular, we have

$$\dim(V^{(k)}) = \sum_{s=0}^{k} \binom{n}{s}.$$

Existence and uniqueness of the best k-th approximation follow from the fact that $A^{(k)}$ is the orthogonal projection onto $V^{(k)}$. Moreover, since the subspaces $V^{(k)}$ are nested, the operators $A^{(k)}$ commute in the following sense:

$$k \leq k' \;\Rightarrow\; A^{(k)}(A^{(k')}(\hat{f})) = A^{(k)}(\hat{f}), \quad \hat{f} \in V^{(n)}. \tag{7.7}$$

We thus observe that $A^{(k)}(\hat{f})$ can be attained from $\hat{f}$ by carrying out successively the projections $A^{(n-1)}, A^{(n-2)}, \ldots, A^{(k)}$. We now search for the relation that links any two consecutive projections. This relation will allow us to compute gradually $A^{(k)}(\hat{f})$ from $\hat{f}$.

Fix $k \in \mathbb{N}$ with $k \leq n-1$, and assume that $\hat{f}^{(k+1)} = A^{(k+1)}(\hat{f})$ is given. By (7.7), $A^{(k)}(\hat{f})$ is the orthogonal projection onto $V^{(k)}$ of $A^{(k+1)}(\hat{f})$ and is thus characterized by the fact that

$$A^{(k)}(\hat{f}) - A^{(k+1)}(\hat{f})$$

is orthogonal to $V^{(k)}$ or equivalently to all functions of a basis of $V^{(k)}$. Consequently, $A^{(k)}(\hat{f})$ is given by the unique solution $\hat{f}^{(k)}$ of the following system:

$$\int_{[0,1]^n} [\hat{f}^{(k)}(x) - \hat{f}^{(k+1)}(x)]\,(\bigwedge_{j\in S} x_j)\,dx = 0, \quad S \subseteq N, s \leq k. \tag{7.8}$$

More precisely, given the coefficients $a_T^{(k+1)}$ of $A^{(k+1)}(\hat{f})$, the coefficients $a_T^{(k)}$ of $A^{(k)}(\hat{f})$ are given by the unique solution of the linear system:

$$\sum_{\substack{T\subseteq N\\ t\leq k}} I_{S,T}\, a_T^{(k)} = \sum_{\substack{T\subseteq N\\ t\leq k+1}} I_{S,T}\, a_T^{(k+1)}, \quad S \subseteq N, s \leq k, \tag{7.9}$$

where

$$I_{S,T} := \int_{[0,1]^n} (\bigwedge_{i\in T} x_i)\,(\bigwedge_{j\in S} x_j)\,dx. \tag{7.10}$$

To solve this system, we need to calculate the explicit value of the integral (7.10). It is the purpose of the following lemma.

Lemma 7.2.1 *For all $T, S \subseteq N$, there holds*

$$\int_{[0,1]^n} (\bigwedge_{i\in T} x_i)\,(\bigwedge_{j\in S} x_j)\,dx = \frac{t+s+2}{(|T\cup S|+2)(t+1)(s+1)}. \tag{7.11}$$

Proof. Observe first that we can assume that T and S are such that $|T \cup S| = n$, so that $|T \cap S| = t + s - n$. Moreover, suppose that T and S are non-empty. Define

$$\mathcal{A}_\pi := \{x \in [0,1]^n \,|\, x_{\pi(1)} < \cdots < x_{\pi(n)}\}, \quad \pi \in \Pi_n,$$
$$I_\pi := \int_{\mathcal{A}_\pi} (\bigwedge_{i\in T} x_i)\,(\bigwedge_{j\in S} x_j)\,dx, \quad \pi \in \Pi_n,$$
$$\Pi_n^{(p)}(R) := \{\pi \in \Pi_n \,|\, \pi(1), \ldots, \pi(p) \in R\}, \quad R \subseteq N, p \in N, p \leq |R|.$$

On the one hand, we have $|\Pi_n^{(1)}(T\cap S)| = |T\cap S|\,(n-1)!$, where $|T\cap S|$ corresponds to all the possible choices of $\pi(1)$ in $T \cap S$, and $(n-1)!$ to all the permutations of the remaining elements in N. Moreover, for all $\pi \in \Pi_n^{(1)}(T \cap S)$, we have

$$I_\pi = \int_0^1 \int_0^{x_{\pi(n)}} \cdots \int_0^{x_{\pi(2)}} x_{\pi(1)}^2\, dx_{\pi(1)} \cdots dx_{\pi(n)} = \frac{2}{(n+2)!}$$

and

$$\sum_{\pi\in\Pi_n^{(1)}(T\cap S)} I_\pi = \frac{2\,|T\cap S|\,(n-1)!}{(n+2)!} = \frac{2\,(t+s-n)}{n(n+1)(n+2)}. \tag{7.12}$$

On the other hand, we have

$$|\Pi_n^{(p)}(S \setminus T)| = \binom{n-t}{p}\, p!\, t\,(n-p-1)!$$

where $\binom{n-t}{p}$ corresponds to all the choices of $\pi(1), \ldots, \pi(p)$ in $S \setminus T$, $p!$ to all the permutations of these elements, t to all the choices of $\pi(p+1)$ in T, and $(n-p-1)!$ to all the permutations of the remaining elements in N. Moreover, for all $\pi \in \Pi_n^{(p)}(S \setminus T)$, we have

$$I_\pi = \int_0^1 \int_0^{x_{\pi(n)}} \cdots \int_0^{x_{\pi(p+2)}} x_{\pi(p+1)} \int_0^{x_{\pi(p+1)}} \cdots \int_0^{x_{\pi(2)}} x_{\pi(1)}\, dx_{\pi(1)} \cdots dx_{\pi(n)} = \frac{p+2}{(n+2)!},$$

and

$$\begin{aligned}
\sum_{p=1}^{n-t} \sum_{\pi\in\Pi_n^{(p)}(S\setminus T)} I_\pi &= \frac{t}{(n+2)!}\sum_{p=1}^{n-t}\binom{n-t}{p} p!\,(n-p-1)!\,(p+2) \\
&= \frac{t!\,(n-t)!}{(n+2)!}\sum_{p=1}^{n-t}\binom{n-p-1}{t-1}(p+2) \\
&= \frac{t!\,(n-t)!}{(n+2)!}\Big[(n+2)\sum_{p=1}^{n-t}\binom{n-p-1}{t-1} - t\sum_{p=1}^{n-t}\binom{n-p}{t}\Big] \\
&= \frac{t!\,(n-t)!}{(n+2)!}\Big[(n+2)\binom{n-1}{t} - t\binom{n}{t+1}\Big],
\end{aligned}$$

that is

$$\sum_{p=1}^{n-t} \sum_{\pi\in\Pi_n^{(p)}(S\setminus T)} I_\pi = \frac{(n-t)(n+2t+2)}{n(n+1)(n+2)(t+1)}. \tag{7.13}$$

Similarly, we can write

$$\sum_{p=1}^{n-s} \sum_{\pi\in\Pi_n^{(p)}(T\setminus S)} I_\pi = \frac{(n-s)(n+2s+2)}{n(n+1)(n+2)(s+1)}. \tag{7.14}$$

Finally, it is clear that we have

$$\Pi_n = \Big[\Pi_n^{(1)}(T\cap S)\Big] \cup \Big[\bigcup_{p=1}^{n-t}\Pi_n^{(p)}(S\setminus T)\Big] \cup \Big[\bigcup_{p=1}^{n-s}\Pi_n^{(p)}(T\setminus S)\Big]$$

and by (7.12)–(7.14), we have

$$\begin{aligned}
\int_{[0,1]^n} (\bigwedge_{i\in T} x_i)\,(\bigwedge_{j\in S} x_j)\,dx &= \sum_{\pi\in\Pi_n^{(1)}(T\cap S)} I_\pi + \sum_{p=1}^{n-t}\sum_{\pi\in\Pi_n^{(p)}(S\setminus T)} I_\pi + \sum_{p=1}^{n-s}\sum_{\pi\in\Pi_n^{(p)}(T\setminus S)} I_\pi \\
&= \frac{t+s+2}{(n+2)(t+1)(s+1)},
\end{aligned}$$

as desired. We can easily see that the result still holds if $T=\emptyset$ or $S=\emptyset$. ∎

Now, we present the solution of the system (7.9). This solution provides a recursive formula linking $A^{(k)}(\hat f)$ to $A^{(k+1)}(\hat f)$.

Lemma 7.2.2 *Let $k\in\mathbb{N}$ with $k\le n-1$. Given the coefficients $a_T^{(k+1)}$ of $A^{(k+1)}(\hat f)$, the coefficients $a_T^{(k)}$ of $A^{(k)}(\hat f)$ are given by*

$$a_S^{(k)} = a_S^{(k+1)} + (-1)^{k+s}\,\frac{\binom{k+s+1}{k+1}}{\binom{2k+2}{k+1}} \sum_{\substack{T\supseteq S\\ t=k+1}} a_T^{(k+1)}, \quad S\subseteq N, s\le k. \tag{7.15}$$

Proof. We only have to prove that the coefficients $a_S^{(k)}$ given by (7.15) satisfy the system (7.9). For this purpose, let us show that, for all $S,T\subseteq N$ with $s\le k$ and $t=k+1$, we have

$$\sum_{J\subseteq T} I_{S,J}\,(-1)^{k+j}\binom{k+j+1}{k+1} = 0. \tag{7.16}$$

Partitioning $J \subseteq T$ into $P \subseteq T \setminus S$ and $Q \subseteq T \cap S$, we have

$$\sum_{J\subseteq T} I_{S,J}\,(-1)^{k+j}\binom{k+j+1}{k+1} = \sum_{P\subseteq T\setminus S}\ \sum_{Q\subseteq T\cap S} I_{S,P\cup Q}\,(-1)^{k+p+q}\binom{k+p+q+1}{k+1}.$$

Next, setting $R := T \cap S$ ($r \leq s \leq k$) and using (7.10) and (7.11), this latter expression can be rewritten as

$$\sum_{p=0}^{k-r+1}\binom{k-r+1}{p}\sum_{q=0}^{r}\binom{r}{q}\frac{p+q+s+2}{(p+s+2)(p+q+1)(s+1)}\,(-1)^{k+p+q}\binom{k+p+q+1}{k+1}.$$

or equivalently

$$-\sum_{p=0}^{k-r+1}\binom{k-r+1}{p}(-1)^{k-r+1-p}\,\frac{1}{p+s+2}\sum_{q=0}^{r}\binom{r}{q}(-1)^{r-q}\,g(p+q), \tag{7.17}$$

where $g : \mathbb{N} \to \mathbb{R}$ is defined by

$$\begin{aligned} g(z) &= \frac{z+s+2}{(z+1)(s+1)}\binom{z+k+1}{k+1} \\ &= \Big(\frac{1}{z+1}+\frac{1}{s+1}\Big)\binom{z+k+1}{k+1} \\ &= \frac{1}{k+1}\binom{z+k+1}{k}+\frac{1}{s+1}\binom{z+k+1}{k+1}, \end{aligned}$$

for all $z \in \mathbb{N}$.

Regard the difference operator

$$\Delta_n\, f(n) := f(n+1) - f(n)$$

for functions on $\mathbb{N}$. It is well known that we have (cf. Berge [19, Chap. 1, Sect. 8])

$$\Delta_n^k\, f(n) = \sum_{j=0}^{k}\binom{k}{j}(-1)^{k-j}\, f(n+j), \quad k \in \mathbb{N}. \tag{7.18}$$

Applying this to g, we obtain

$$\sum_{q=0}^{r}\binom{r}{q}(-1)^{r-q}\, g(p+q) = \Delta_p^r\, g(p) = \frac{1}{k+1}\binom{p+k+1}{k-r}+\frac{1}{s+1}\binom{p+k+1}{k-r+1},$$

and (7.17) becomes

$$-\sum_{p=0}^{k-r+1}\binom{k-r+1}{p}(-1)^{k-r+1-p}\,\frac{1}{p+s+2}\,\Big(\frac{1}{k+1}\binom{p+k+1}{k-r}+\frac{1}{s+1}\binom{p+k+1}{k-r+1}\Big).$$

Applying (7.18) again, we see that this latter espression can be written as

$$-\Big[\Delta_z^{k-r+1}\,\frac{1}{z+s+2}\,\Big(\frac{1}{k+1}\binom{z+k+1}{k-r}+\frac{1}{s+1}\binom{z+k+1}{k-r+1}\Big)\Big]_{z=0}.$$

We now show that the expression in brackets is identically zero. This will finally prove (7.16). Let us consider two exclusive cases:

- If $s = k$ then

$$\begin{aligned}
&\frac{1}{z+s+2}\Big(\frac{1}{k+1}\binom{z+k+1}{k-r}+\frac{1}{s+1}\binom{z+k+1}{k-r+1}\Big)\\
=\;&\frac{1}{(k+1)(z+k+2)}\binom{z+k+2}{k-r+1}\\
=\;&\frac{1}{(k+1)(k-r+1)}\binom{z+k+1}{k-r}
\end{aligned}$$

is a polynomial $P_{k-r}(z)$ of degree $k-r$, and $\Delta_z^{k-r+1}\,P_{k-r}(z)=0$.

- If $r \le s \le k-1$ then

$$\begin{aligned}
&\frac{1}{z+s+2}\Big(\frac{1}{k+1}\binom{z+k+1}{k-r}+\frac{1}{s+1}\binom{z+k+1}{k-r+1}\Big)\\
=\;&\frac{1}{z+s+2}\Big(\frac{1}{(k+1)(k-r)!}\prod_{i=r+2}^{k+1}(z+i)+\frac{1}{(s+1)(k-r+1)!}\prod_{i=r+1}^{k+1}(z+i)\Big)
\end{aligned}$$

is a polynomial $P'_{k-r}(z)$ of degree $k-r$, and $\Delta_z^{k-r+1}\,P'_{k-r}(z)=0$.

Now, from (7.16), we have, for all $S,T\subseteq N$ with $s\le k$ and $t=k+1$,

$$\sum_{\substack{J\subseteq T\\ j\le k}} I_{S,J}(-1)^{k+j}\binom{k+j+1}{k+1}=I_{S,T}\binom{2k+2}{k+1}.$$

Multiplying this identity by $a_T^{(k+1)}$ and then summing over T, we obtain

$$\begin{aligned}
&\sum_{\substack{T\subseteq N\\ t=k+1}}\sum_{\substack{J\subseteq T\\ j\le k}} I_{S,J}\,(-1)^{k+j}\,\frac{\binom{k+j+1}{k+1}}{\binom{2k+2}{k+1}}a_T^{(k+1)} = \sum_{\substack{T\subseteq N\\ t=k+1}} I_{S,T}\,a_T^{(k+1)}\\
\Leftrightarrow\;&\sum_{\substack{J\subseteq N\\ j\le k}} I_{S,J}\,(-1)^{k+j}\,\frac{\binom{k+j+1}{k+1}}{\binom{2k+2}{k+1}}\sum_{\substack{T\supseteq J\\ t=k+1}} a_T^{(k+1)} = \sum_{\substack{T\subseteq N\\ t=k+1}} I_{S,T}\,a_T^{(k+1)}\\
\Leftrightarrow\;&\sum_{\substack{J\subseteq N\\ j\le k}} I_{S,J}\,(a_J^{(k)}-a_J^{(k+1)}) = \sum_{\substack{T\subseteq N\\ t=k+1}} I_{S,T}\,a_T^{(k+1)}\quad\text{(by (7.15))}\\
\Leftrightarrow\;&\sum_{\substack{J\subseteq N\\ j\le k}} I_{S,J}\,a_J^{(k)} = \sum_{\substack{T\subseteq N\\ t\le k+1}} I_{S,T}\,a_T^{(k+1)}.
\end{aligned}$$

This shows that the coefficients $a_T^{(k)}$ given in the statement satisfy the system (7.9). ■

Now, let us turn to the explicit formula giving the coefficients of $A^{(k)}(\hat f)$ in terms of the coefficients of $\hat f$. It is worth comparing this formula with the solution (7.5) of the Hammer-Holzman approximation problem.

Theorem 7.2.1 *The coefficients of $A^{(k)}(\hat f)$ are given from those of $\hat f$ by*

$$a_S^{(k)} = a_S + (-1)^{k+s}\sum_{\substack{T\supseteq S\\ t>k}}\frac{\binom{k+s+1}{s}\binom{t-s-1}{k-s}}{\binom{t+k+1}{k+1}}\,a_T,\quad S\subseteq N, s\le k. \tag{7.19}$$

Proof. Due to the uniqueness of $A^{(k)}(\hat{f})$, we only have to prove that the coefficients $a_S^{(k)}$ given by (7.19) fulfil the equation (7.15).

Let $k \in \mathbb{N}$, $k \leq n-1$, and set $S \subseteq N$, $s \leq k$. By substituting (7.19) into (7.15), we obtain, after removing the common term a_S,

$$\begin{aligned}
&(-1)^{k+s} \sum_{\substack{R \supseteq S \\ r>k}} \frac{\binom{k+s+1}{k+1}\binom{r-s-1}{k-s}}{\binom{r+k+1}{k+1}} a_R \\
= \ &(-1)^{k+s+1} \sum_{\substack{R \supseteq S \\ r>k+1}} \frac{\binom{k+s+2}{k+2}\binom{r-s-1}{k-s+1}}{\binom{r+k+2}{k+2}} a_R + (-1)^{k+s} \frac{\binom{k+s+1}{k+1}}{\binom{2k+2}{k+1}} \sum_{\substack{T \supseteq S \\ t=k+1}} \Big[a_T + \sum_{\substack{R \supseteq T \\ r>k+1}} \frac{\binom{2k+3}{k+2}}{\binom{r+k+2}{k+2}} a_R \Big].
\end{aligned}$$

Let us show that this equality holds. Dividing by $(-1)^{k+s}$ and then removing the terms corresponding to a_R with $r = k+1$, the equality becomes

$$\sum_{\substack{R \supseteq S \\ r>k+1}} \frac{\binom{k+s+1}{k+1}\binom{r-s-1}{k-s}}{\binom{r+k+1}{k+1}} a_R = - \sum_{\substack{R \supseteq S \\ r>k+1}} \frac{\binom{k+s+2}{k+2}\binom{r-s-1}{k-s+1}}{\binom{r+k+2}{k+2}} a_R + \sum_{\substack{R \supseteq S \\ r>k+1}} \frac{\binom{k+s+1}{k+1}\binom{2k+3}{k+2}}{\binom{2k+2}{k+1}\binom{r+k+2}{k+2}} \sum_{\substack{T : S \subseteq T \subseteq R \\ t=k+1}} a_R.$$

Fix $R \supseteq S$ with $r > k+1$ and consider the coefficients of a_R in the previous equality. By identification, we have

$$\frac{\binom{k+s+1}{k+1}\binom{r-s-1}{k-s}}{\binom{r+k+1}{k+1}} = - \frac{\binom{k+s+2}{k+2}\binom{r-s-1}{k-s+1}}{\binom{r+k+2}{k+2}} + \frac{\binom{k+s+1}{k+1}\binom{2k+3}{k+2}\binom{r-s}{k-s+1}}{\binom{2k+2}{k+1}\binom{r+k+2}{k+2}},$$

and this equality can be easily verified. ■

Theorem 7.2.1 gives the general solution of the approximation problem proposed in Definition 7.2.1. Some particular cases are shown in Table 7.2.

We also observe that setting $S = \emptyset$ in (7.8) provides

$$\text{orness}(A^{(k)}(\hat{f})) = \text{orness}(\hat{f}), \quad \hat{f} \in V^{(n)}.$$

Example 7.2.1 *Let $\hat{f} : [0,1]^4 \to \mathbb{R}$ be given by*

$$\begin{aligned}
\hat{f}(x) \ = \ &\frac{3}{10} \Big[x_1 + x_2 + x_3 + (x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3) \Big] \\
&- \frac{21}{25} (x_1 \wedge x_2 \wedge x_3) + \frac{1}{25} (x_1 \wedge x_2 \wedge x_3 \wedge x_4).
\end{aligned}$$

The best constant approximation is given by

$$A^{(0)}(\hat{f}) = \frac{137}{250},$$

the best linear approximation by

$$(A^{(1)}\hat{f})(x) = \frac{1}{100} + \frac{89}{250} (x_1 + x_2 + x_3) + \frac{1}{125} x_4$$

and the best min-quadratic approximation by

$$\begin{aligned}
(A^{(2)}\hat{f})(x) \ = \ &-\frac{27}{700} + \frac{803}{1750} (x_1 + x_2 + x_3) - \frac{8}{875} x_4 \\
&- \frac{19}{175} \Big[(x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3) \Big] \\
&+ \frac{2}{175} \Big[(x_1 \wedge x_4) + (x_2 \wedge x_4) + (x_3 \wedge x_4) \Big].
\end{aligned}$$

$$
\begin{array}{rcll}
a_\emptyset^{(0)} & = & \displaystyle\sum_{T\subseteq N} \frac{1}{t+1}\, a_T & \\[2ex]
a_\emptyset^{(1)} & = & \displaystyle\sum_{T\subseteq N} \frac{-2(t-1)}{(t+1)(t+2)}\, a_T & \\
a_i^{(1)} & = & \displaystyle\sum_{T\ni i} \frac{6}{(t+1)(t+2)}\, a_T, & i\in N \\[2ex]
a_\emptyset^{(2)} & = & \displaystyle\sum_{T\subseteq N} \frac{3(t-1)(t-2)}{(t+1)(t+2)(t+3)}\, a_T & \\
a_i^{(2)} & = & \displaystyle\sum_{T\ni i} \frac{-24(t-2)}{(t+1)(t+2)(t+3)}\, a_T, & i\in N \\
a_{ij}^{(2)} & = & \displaystyle\sum_{T\ni i,j} \frac{60}{(t+1)(t+2)(t+3)}\, a_T, & \{i,j\}\subseteq N \\[2ex]
a_S^{(n-1)} & = & \displaystyle a_S + (-1)^{n+s-1}\,\frac{(n+s)!\,n!}{s!\,(2n)!}\, a_N, & S\subseteq N, s\leq n-1
\end{array}
$$

Table 7.2: Coefficients of $A^{(k)}(\hat f)$ for some values of k

Remark: Let $S=\{i_1,\dots,i_s\}$ and suppose that $\hat f$ does not depend on $x_{i_1}\wedge\cdots\wedge x_{i_s}$, that is $a_T=0$ for all $T\supseteq S$. Then, by (7.19), $a_T^{(k)}=0$ whenever $T\supseteq S$ and $A^{(k)}(\hat f)$ does not depend on $x_{i_1}\wedge\cdots\wedge x_{i_s}$.

In other terms, for all $S\subseteq N$,

$$\Delta_S\,\hat f(x)=0\quad \forall x\in[0,1]^n \quad\Rightarrow\quad \Delta_S\,(A^{(k)}\hat f)(x)=0\quad \forall x\in[0,1]^n.$$

7.2.2 Approximations having one fixed value

In certain applications of the theory developed here, one may be interested only in approximations that have fixed values at $x=\underline{0}$ and $x=\underline{1}$ (where $\underline{0}$ and $\underline{1}$ are shorthand for $(0,\dots,0)$ and $(1,\dots,1)$ respectively). In the present section, we investigate the case where the approximations have a fixed value at $x=\underline{0}$.

For any $\alpha\in\mathbb{R}$, let us introduce $V^{(k,\alpha)}$ for the set of all min-polynomials $\hat f^{(k,\alpha)}$ of degree $\leq k$ which fulfil $\hat f^{(k,\alpha)}(\underline{0})=\alpha$. All these functions are of the form:

$$\hat f^{(k,\alpha)}(x)=\alpha+\sum_{\substack{T\subseteq N\\ 1\leq t\leq k}} a_T^{(k,\alpha)}\bigwedge_{i\in T} x_i,\quad x\in[0,1]^n. \tag{7.20}$$

Given any Lovász extension $\hat f$, we are searching for a min-polynomial $\hat f^{(k,\alpha)}$ which minimizes

$$\int_{[0,1]^n}[\hat f(x)-\hat f^{(k,\alpha)}(x)]^2\,dx$$

among all min-polynomials of the form (7.20), and we write $\hat f^{(k,\alpha)}=A^{(k,\alpha)}(\hat f)$. Of course, this problem is relevant only if $\alpha\neq(A^{(k)}\hat f)(\underline{0})$.

First, it is clear that $V^{(k,\alpha)}$ is an affine subspace in $V^{(k)}$ and one of its bases can be given by

$$B^{(k,\alpha)} = \Big\{ \bigwedge_{i\in S} x_i \,\Big|\, S \subseteq N, 1 \leq s \leq k \Big\}.$$

In particular, we have

$$\dim(V^{(k,\alpha)}) = \dim(V^{(k)}) - 1.$$

Next, as in Definition 7.2.1, existence and uniqueness follow from the observation that $A^{(k,\alpha)}(\hat{f})$ is the orthogonal projection of $\hat{f}$ onto $V^{(k,\alpha)}$. This projection may be realized by first projecting $\hat{f}$ onto $V^{(k)}$ and then projecting the obtained function onto $V^{(k,\alpha)}$. Thus, one can obtain $A^{(k,\alpha)}(\hat{f})$ from the best k-th approximation $\hat{f}^{(k)} = A^{(k)}(\hat{f})$ given in Theorem 7.2.1. The following result shows how the coefficients of $A^{(k,\alpha)}(\hat{f})$ can be calculated.

Theorem 7.2.2 *The coefficients of $A^{(k,\alpha)}(\hat{f})$ are given from those of $A^{(k)}(\hat{f})$ by*

$$a_T^{(k,\alpha)} = a_T^{(k)} + c_T^{(k,\alpha)}, \quad T \subseteq N,\ 1 \leq t \leq k,$$

where the values $c_T^{(k,\alpha)}$ correspond to the unique solution of the linear system:

$$\sum_{\substack{T\subseteq N\\ 1\leq t\leq k}} I_{S,T}\, c_T^{(k,\alpha)} = -\frac{1}{s+1}\,\sigma^{(k,\alpha)}, \quad S \subseteq N,\ 1 \leq s \leq k, \tag{7.21}$$

with $\sigma^{(k,\alpha)} = \alpha - (A^{(k)}\hat{f})(\underline{0})$.

Proof. Since the function $A^{(k,\alpha)}(\hat{f})$ is the orthogonal projection onto $V^{(k,\alpha)}$ of $\hat{f}^{(k)} = A^{(k)}(\hat{f})$, it is characterized by the fact that

$$A^{(k,\alpha)}(\hat{f}) - A^{(k)}(\hat{f})$$

is orthogonal to $V^{(k,\alpha)}$ or equivalently to all functions of $B^{(k,\alpha)}$. Therefore, $A^{(k,\alpha)}(\hat{f})$ is given by the unique solution $\hat{f}^{(k,\alpha)}$ of the following system:

$$\int_{[0,1]^n} [\hat{f}^{(k,\alpha)}(x) - \hat{f}^{(k)}(x)]\,(\bigwedge_{i\in S} x_i)\,dx = 0, \quad S \subseteq N,\ 1 \leq s \leq k.$$

Setting $c_T^{(k,\alpha)} := a_T^{(k,\alpha)} - a_T^{(k)}$, we can see from (7.20) that this system is equivalent to (7.21). ■

Consider first the case of $k = 1$. Setting $S = \{p\}$, $p \in N$, the system (7.21) is written

$$\frac{1}{4}\sum_{i\in N\setminus p} c_i^{(1,\alpha)} + \frac{1}{3}\,c_p^{(1,\alpha)} = -\frac{1}{2}\,\sigma^{(1,\alpha)}, \quad p \in N,$$

and the solution is

$$c_i^{(1,\alpha)} = -\frac{6}{3n+1}\,\sigma^{(1,\alpha)}, \quad i \in N.$$

Therefore, by Theorem 7.2.2, the coefficients of $A^{(1,\alpha)}(\hat{f})$ are given by

$$\begin{cases} a_\emptyset^{(1,\alpha)} = \alpha \\ a_i^{(1,\alpha)} = a_i^{(1)} - \frac{6}{3n+1}\,(\alpha - a_\emptyset^{(1)}), \quad i \in N. \end{cases}$$

For $k \geq 2$, the system (7.21) seems to be difficult to solve. However, we conjecture that its solution is such that

$$|T_1| = |T_2| \;\Rightarrow\; c_{T_1}^{(k,\alpha)} = c_{T_2}^{(k,\alpha)}. \tag{7.22}$$

We already know that this is true for $k = 1$. In case of $k = 2$, assuming that (7.22) holds, one can show that the solution of (7.21) is given by

$$\begin{aligned} c_i^{(2,\alpha)} &= \frac{-36n+12}{8n^2+5n+3}\,\sigma^{(2,\alpha)}, \quad i \in N,\\ c_{ij}^{(2,\alpha)} &= \frac{60}{8n^2+5n+3}\,\sigma^{(2,\alpha)}, \quad i,j \in N, \end{aligned}$$

and the conjecture remains true for $k = 2$. Theorem 7.2.2 then allows to deduce the coefficients of $A^{(2,\alpha)}(\hat{f})$.

Because of the complexity of the system (7.21), we restrict ourselves to the cases of $k = 1$ and $k = 2$. Table 7.3 gives the corresponding coefficients.

$$\boxed{\begin{array}{lcll} a_\emptyset^{(1,\alpha)} &=& \alpha & \\ a_i^{(1,\alpha)} &=& a_i^{(1)} - \dfrac{6}{3n+1}\,(\alpha - a_\emptyset^{(1)}), & i \in N \\ a_\emptyset^{(2,\alpha)} &=& \alpha & \\ a_i^{(2,\alpha)} &=& a_i^{(2)} + \dfrac{-36n+12}{8n^2+5n+3}\,(\alpha - a_\emptyset^{(2)}), & i \in N \\ a_{ij}^{(2,\alpha)} &=& a_{ij}^{(2)} + \dfrac{60}{8n^2+5n+3}\,(\alpha - a_\emptyset^{(2)}), & \{i,j\} \subseteq N \end{array}}$$

Table 7.3: Coefficients of $A^{(k,\alpha)}(\hat{f})$ for $k = 1$ and $k = 2$

Example 7.1 (continued) *Consider again the function $\hat{f} : [0,1]^4 \to \mathbb{R}$ defined by*

$$\begin{aligned} \hat{f}(x) &= \frac{3}{10}\Big[x_1 + x_2 + x_3 + (x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3)\Big]\\ &\quad -\frac{21}{25}\,(x_1 \wedge x_2 \wedge x_3) + \frac{1}{25}\,(x_1 \wedge x_2 \wedge x_3 \wedge x_4). \end{aligned}$$

Then we have

$$(A^{(1,0)}\hat{f})(x) = \frac{586}{1625}\,(x_1 + x_2 + x_3) + \frac{41}{3250}\,x_4$$

and

$$\begin{aligned} (A^{(2,0)}\hat{f})(x) &= \frac{16049}{37750}\,(x_1 + x_2 + x_3) - \frac{809}{18875}\,x_4\\ &\quad -\frac{352}{3775}\Big[(x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3)\Big]\\ &\quad +\frac{101}{3775}\Big[(x_1 \wedge x_4) + (x_2 \wedge x_4) + (x_3 \wedge x_4)\Big]. \end{aligned}$$

7.2.3 Approximations having two fixed values

We now investigate the case where the approximations have fixed values at $x = \underline{0}$ and $x = \underline{1}$.

For any $\alpha, \beta \in \mathbb{R}$, we define $V^{(k,\alpha,\beta)}$ as the set of all min-polynomials $\hat{f}^{(k,\alpha,\beta)}$ of degree $\leq k$ which fulfil $\hat{f}^{(k,\alpha,\beta)}(\underline{0}) = \alpha$ and $\hat{f}^{(k,\alpha,\beta)}(\underline{1}) = \beta$. All these functions are of the form

$$\hat{f}^{(k,\alpha,\beta)}(x) = \sum_{\substack{T \subseteq N \\ t \leq k}} a_T^{(k,\alpha,\beta)} \bigwedge_{i \in T} x_i, \quad x \in [0,1]^n,$$

with

$$\begin{cases} a_\emptyset^{(k,\alpha,\beta)} = \alpha \\ \displaystyle\sum_{\substack{T \subseteq N \\ t \leq k}} a_T^{(k,\alpha,\beta)} = \beta. \end{cases}$$

For a fixed $j \in N$, these functions can also be written as

$$\hat{f}^{(k,\alpha,\beta)}(x) = \alpha + (\beta - \alpha)\, x_j + \sum_{\substack{T \subseteq N, T \neq j \\ 1 \leq t \leq k}} a_T^{(k,\alpha,\beta)} \,(\bigwedge_{i \in T} x_i - x_j), \quad x \in [0,1]^n. \tag{7.23}$$

Given any Lovász extension $\hat{f}$, we are searching for a min-polynomial $\hat{f}^{(k,\alpha,\beta)}$ which minimizes

$$\int_{[0,1]^n} [\hat{f}(x) - \hat{f}^{(k,\alpha,\beta)}(x)]^2\, dx$$

among all min-polynomials of the form (7.23), and we write $\hat{f}^{(k,\alpha,\beta)} = A^{(k,\alpha,\beta)}(\hat{f})$. This problem is relevant only if $\alpha \neq (A^{(k)}\hat{f})(\underline{0})$ or $\beta \neq (A^{(k)}\hat{f})(\underline{1})$.

The situation is similar to the previous one. $V^{(k,\alpha,\beta)}$ is an affine subspace in $V^{(k)}$ and one of its bases can be given by

$$B_j^{(k,\alpha,\beta)} = \Big\{ \bigwedge_{i \in S} x_i - x_j \;\Big|\; S \subseteq N, S \neq j, 1 \leq s \leq k \Big\}$$

where $j \in N$. In particular, we have

$$\dim(V^{(k,\alpha,\beta)}) = \dim(V^{(k)}) - 2.$$

Moreover, $A^{(k,\alpha,\beta)}(\hat{f})$ is the orthogonal projection of $\hat{f}$ onto $V^{(k,\alpha,\beta)}$. This projection may be performed from $A^{(k)}(\hat{f})$ and we have the following.

Theorem 7.2.3 *The coefficients of $A^{(k,\alpha,\beta)}(\hat{f})$ are given from those of $A^{(k)}(\hat{f})$ by*

$$a_T^{(k,\alpha,\beta)} = a_T^{(k)} + c_T^{(k,\alpha,\beta)}, \quad T \subseteq N,\ 1 \leq t \leq k,$$

where the values $c_T^{(k,\alpha,\beta)}$ correspond to the unique solution of the linear system ($j \in N$ being fixed):

$$\begin{aligned} &\sum_{\substack{T \subseteq N, T \neq j \\ 1 \leq t \leq k}} (I_{S,T} - I_{j,T} - I_{S,j} + \frac{1}{3})\, c_T^{(k,\alpha,\beta)} \\ = \;& (\frac{1}{2} - \frac{1}{s+1})\, \sigma^{(k,\alpha)} + (\frac{1}{3} - I_{S,j})\, \sigma^{(k,\alpha,\beta)}, \quad S \subseteq N,\ S \neq j,\ 1 \leq s \leq k, \end{aligned} \tag{7.24}$$

with $\sigma^{(k,\alpha)} = \alpha - (A^{(k)}\hat{f})(\underline{0})$ and $\sigma^{(k,\alpha,\beta)} = \beta - \alpha - (A^{(k)}\hat{f})(\underline{1}) + (A^{(k)}\hat{f})(\underline{0})$.

Proof. The proof is similar to that of Theorem 7.2.2. $A^{(k,\alpha,\beta)}(\hat{f})$ is given by the unique solution $\hat{f}^{(k,\alpha,\beta)}$ of the following system:

$$\int_{[0,1]^n} [\hat{f}^{(k,\alpha,\beta)}(x) - \hat{f}^{(k)}(x)] \, (\bigwedge_{i\in S} x_i - x_j) \, dx = 0, \quad S \subseteq N, \ S \neq j, \ 1 \leq s \leq k. \tag{7.25}$$

From (7.23), we have, setting $c_T^{(k,\alpha,\beta)} := a_T^{(k,\alpha,\beta)} - a_T^{(k)}$,

$$\hat{f}^{(k,\alpha,\beta)}(x) - \hat{f}^{(k)}(x) = \sigma^{(k,\alpha)} + \sigma^{(k,\alpha,\beta)} \, x_j + \sum_{\substack{T\subseteq N, T\neq j\\ 1\leq t\leq k}} c_T^{(k,\alpha,\beta)} \, (\bigwedge_{i\in T} x_i - x_j).$$

It follows that the system (7.25) is equivalent to (7.24). ∎

Consider the case $k = 1$. Setting $S = \{p\}$, $p \in N \setminus j$, the system (7.24) is written

$$\frac{1}{12} \sum_{i\in N\setminus jp} c_i^{(1,\alpha,\beta)} + \frac{2}{12} \, c_p^{(1,\alpha,\beta)} = \frac{1}{12} \, \sigma^{(1,\alpha,\beta)}, \quad p \in N \setminus j,$$

and the solution is

$$c_i^{(1,\alpha,\beta)} = \frac{1}{n} \, \sigma^{(1,\alpha,\beta)}, \quad i \in N.$$

Using the same conjecture as in the previous section, one can show that, for $k = 2$, the solution of (7.24) is given by

$$\begin{aligned} c_i^{(2,\alpha,\beta)} &= \frac{-30n+30}{5n^2-4n+3} \, \sigma^{(2,\alpha)} + \frac{-10n^2+11n+3}{n(5n^2-4n+3)} \, \sigma^{(2,\alpha,\beta)}, \quad i \in N,\\ c_{ij}^{(2,\alpha,\beta)} &= \frac{60}{5n^2-4n+3} \, \sigma^{(2,\alpha)} + \frac{30}{5n^2-4n+3} \, \sigma^{(2,\alpha,\beta)}, \quad i,j \in N, \end{aligned}$$

Theorem 7.2.3 provides the coefficients of $A^{(1,\alpha,\beta)}(\hat{f})$ and $A^{(2,\alpha,\beta)}(\hat{f})$. These coefficients are presented in Table 7.4.

$$\begin{array}{lcll} a_\emptyset^{(1,\alpha,\beta)} &=& \alpha & \\ a_i^{(1,\alpha,\beta)} &=& a_i^{(1)} + \frac{1}{n}(\beta - \alpha - \sum_{p\in N} a_p^{(1)}), & i \in N \\ a_\emptyset^{(2,\alpha,\beta)} &=& \alpha & \\ a_i^{(2,\alpha,\beta)} &=& a_i^{(2)} + \frac{-30n+30}{5n^2-4n+3}\,(\alpha - a_\emptyset^{(2)}) + \frac{-10n^2+11n+3}{n(5n^2-4n+3)}\,(\beta - \alpha - \sum_{p\in N} a_p^{(2)} - \sum_{\{p,q\}\subseteq N} a_{pq}^{(2)}), & i \in N \\ a_{ij}^{(2,\alpha,\beta)} &=& a_{ij}^{(2)} + \frac{60}{5n^2-4n+3}\,(\alpha - a_\emptyset^{(2)}) + \frac{30}{5n^2-4n+3}\,(\beta - \alpha - \sum_{p\in N} a_p^{(2)} - \sum_{\{p,q\}\subseteq N} a_{pq}^{(2)}), & \{i,j\} \subseteq N \end{array}$$

Table 7.4: Coefficients of $A^{(k,\alpha,\beta)}(\hat{f})$ for $k = 1$ and $k = 2$

Example 7.1 (continued) *For the function $\hat{f} : [0,1]^4 \to \mathbb{R}$ defined by*

$$\begin{aligned} \hat{f}(x) &= \frac{3}{10} \Big[x_1 + x_2 + x_3 + (x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3) \Big] \\ &\quad -\frac{21}{25} \, (x_1 \wedge x_2 \wedge x_3) + \frac{1}{25} \, (x_1 \wedge x_2 \wedge x_3 \wedge x_4), \end{aligned}$$

we have

$$(A^{(1,0,1)}\hat{f})(x) = \frac{337}{1000}\,(x_1+x_2+x_3) - \frac{11}{1000}\,x_4$$

and

$$\begin{aligned}(A^{(2,0,1)}\hat{f})(x) \;&=\; \frac{29419}{67000}\,(x_1+x_2+x_3) - \frac{1937}{67000}\,x_4\\ &\quad -\frac{181}{1675}\Big[(x_1\wedge x_2)+(x_1\wedge x_3)+(x_2\wedge x_3)\Big]\\ &\quad +\frac{4}{335}\Big[(x_1\wedge x_4)+(x_2\wedge x_4)+(x_3\wedge x_4)\Big].\end{aligned}$$

7.2.4 Increasing approximations having two fixed values

It is clear that any Lovász extension $\hat{f}$ is increasing if and only if its restriction f to $\{0,1\}^n$ is also increasing. Moreover, the increasing monotonicity conditions are given by the following linear inequalities, see (4.6):

$$\sum_{T:i\in T\subseteq S} a_T \geq 0, \quad \forall S\subseteq N,\ \forall i\in S.$$

Now, for any $\alpha,\beta\in\mathbb{R}$, we define $V^{[k,\alpha,\beta]}$ as the set of all increasing min-polynomials $\hat{f}^{[k,\alpha,\beta]}$ of degree $\leq k$ which fulfil $\hat{f}^{[k,\alpha,\beta]}(\underline{0})=\alpha$ and $\hat{f}^{[k,\alpha,\beta]}(\underline{1})=\beta$. All these functions are of the form

$$\hat{f}^{[k,\alpha,\beta]}(x) = \sum_{\substack{T\subseteq N\\ t\leq k}} a_T^{[k,\alpha,\beta]} \bigwedge_{i\in T} x_i, \quad x\in[0,1]^n,$$

with

$$\begin{cases} a_\emptyset^{[k,\alpha,\beta]} = \alpha \\ \displaystyle\sum_{\substack{T\subseteq N\\ t\leq k}} a_T^{[k,\alpha,\beta]} = \beta \\ \displaystyle\sum_{\substack{T:i\in T\subseteq S\\ t\leq k}} a_T^{[k,\alpha,\beta]} \geq 0, \quad S\subseteq N,\ i\in S. \end{cases}$$

Given any Lovász extension $\hat{f}$, we are searching for a min-polynomial $\hat{f}^{[k,\alpha,\beta]}$ which minimizes

$$\int_{[0,1]^n} [\hat{f}(x) - \hat{f}^{[k,\alpha,\beta]}(x)]^2\,dx$$

among all min-polynomials belonging to $V^{[k,\alpha,\beta]}$, and we write $\hat{f}^{[k,\alpha,\beta]} = A^{[k,\alpha,\beta]}(\hat{f})$. Of course this problem is relevant only if $A^{(k,\alpha,\beta)}(\hat{f})$ is not increasing.

In fact, for fixed $\hat{f}\in V^{(n)}$, we consider the following problem:

$$\inf\Big\{\|\hat{f} - \hat{f}^{[k,\alpha,\beta]}\|^2 \;:\; \hat{f}^{[k,\alpha,\beta]} \in V^{[k,\alpha,\beta]}\Big\} \tag{7.26}$$

i.e. we are interested in those points (if any) of $V^{[k,\alpha,\beta]}$ that are closest to $\hat{f}$ for the distance d.

Due to the presence of the linear inequalities, the set $V^{[k,\alpha,\beta]}$ is clearly a non-empty closed convex polyhedron in $V^{(k,\alpha,\beta)}$. Hence, the *existence* of a closest point in $V^{[k,\alpha,\beta]}$ to $\hat{f}$ is ensured and the inf in (7.26) is a min. Moreover, the convexity of $V^{[k,\alpha,\beta]}$ implies *uniqueness* [101, Chap. 3, Sect. 3.1]. The point $A^{[k,\alpha,\beta]}(\hat{f})$ is called the *projection* of $\hat{f}$ onto the polyhedron $V^{[k,\alpha,\beta]}$.

Now, observe that, for all $\hat{f} \in V^{(n)}$ and all $\hat{g} \in V^{[k,\alpha,\beta]}$, we have

$$\|\hat{f} - \hat{g}\|^2 = \|\hat{f} - A^{(k,\alpha,\beta)}(\hat{f})\|^2 + \|A^{(k,\alpha,\beta)}(\hat{f}) - \hat{g}\|^2.$$

Hence, $A^{[k,\alpha,\beta]}(\hat{f})$ can be obtained by projecting $A^{(k,\alpha,\beta)}(\hat{f})$ onto $V^{[k,\alpha,\beta]}$. This projection is thus given by the unique solution $\hat{f}^{[k,\alpha,\beta]}$ of the following problem:

$$\begin{aligned}
\text{minimize} \quad & \|A^{(k,\alpha,\beta)}(\hat{f}) - \hat{f}^{[k,\alpha,\beta]}\|^2 \\
= \quad & \int_{[0,1]^n} \Big[\sum_{\substack{T \subseteq N \\ t \leq k}} (a_T^{(k,\alpha,\beta)} - a_T^{[k,\alpha,\beta]}) \bigwedge_{i \in T} x_i \Big]^2 dx \\
= \quad & \sum_{\substack{T \subseteq N \\ t \leq k}} I_{T,T}\, (a_T^{(k,\alpha,\beta)} - a_T^{[k,\alpha,\beta]}) + 2 \sum_{\substack{\{S,T\} \subseteq 2^N \\ s,t \leq k}} I_{S,T}\, (a_S^{(k,\alpha,\beta)} - a_S^{[k,\alpha,\beta]})(a_T^{(k,\alpha,\beta)} - a_T^{[k,\alpha,\beta]})
\end{aligned}$$

subject to

$$\begin{cases}
a_\emptyset^{[k,\alpha,\beta]} = \alpha \\
\displaystyle\sum_{\substack{T \subseteq N \\ t \leq k}} a_T^{[k,\alpha,\beta]} = \beta \\
\displaystyle\sum_{\substack{T : i \in T \subseteq S \\ t \leq k}} a_T^{[k,\alpha,\beta]} \geq 0, \quad S \subseteq N,\ i \in S.
\end{cases}$$

This problem can be easily rewritten under the form of a *quadratic program* (QP). Indeed, considering the vector[1]

$$y = (a_T^{(k,\alpha,\beta)} - a_T^{[k,\alpha,\beta]})_{\substack{T \subseteq N \\ t \leq k}}$$

and the square matrix

$$Q = \Big(I_{S,T}\Big)_{\substack{S,T \subseteq N \\ s,t \leq k}}$$

the problem becomes:

$$\text{minimize} \quad z = y^t Q\, y$$

subject to

$$\begin{cases}
y_\emptyset = 0 \\
\displaystyle\sum_{\substack{T \subseteq N \\ t \leq k}} y_T = 0 \\
\displaystyle\sum_{\substack{T : i \in T \subseteq S \\ t \leq k}} y_T \leq \sum_{\substack{T : i \in T \subseteq S \\ t \leq k}} a_T^{(k,\alpha,\beta)}, \quad S \subseteq N,\ i \in S.
\end{cases}$$

This is a quadratic program involving $\sum_{s=0}^{k} \binom{n}{s}$ variables, 2 equality constraints and $n2^{n-1}$ inequality constraints (only n inequalities when $k = 1$), see (4.7).

Several procedures for solving quadratic programming problems have been proposed in literature (see e.g. [16, 20]). We will not present them here.

Although it seems impossible to obtain a closed form for the general solution of this problem, we will see in the next section that the projection $A^{[k,\alpha,\beta]}(\hat{f})$ can be obtained very efficiently when $k = 1$, $\alpha = 0$ and $\beta = 1$.

[1] If we fix a linear order on 2^N we can identify 2^N with $\{1, 2, \ldots, 2^n\}$.

7.2.5 Closest weighted arithmetic mean to a Lovász extension

In this section, we intend to determine the particular approximation $A^{[1,0,1]}(\hat{f})$ of a given Lovász extension $\hat{f}$. Thus, we search for the closest weighted arithmetic mean $\sum_{i=1}^n a_i^{[1,0,1]} x_i$ to $\hat{f}$. Of course, we assume that $A^{(1,0,1)}(\hat{f})$ is not increasing.

Using the notations introduced in the previous section, the problem is written

$$\text{minimize} \quad \int_{[0,1]^n} \Big[\sum_{i=1}^n (a_i^{(1,0,1)} - a_i^{[1,0,1]})\, x_i\Big]^2 \, dx$$

subject to

$$\begin{cases} \displaystyle\sum_{i=1}^n a_i^{[1,0,1]} = 1 \\ a_i^{[1,0,1]} \geq 0, \qquad i \in N. \end{cases}$$

The corresponding quadratic program is of the form:

$$\text{minimize} \quad \frac{1}{3}\sum_{i=1}^n y_i^2 + \frac{1}{4}\sum_{\substack{i,j=1 \\ i\neq j}}^n y_i\, y_j$$

subject to

$$\begin{cases} \displaystyle\sum_{i=1}^n y_i = 0 \\ y_i \leq a_i^{(1,0,1)}, \quad i \in N, \end{cases}$$

where $y_i := a_i^{(1,0,1)} - a_i^{[1,0,1]}$.

We will propose an algorithm allowing to determine efficiently the approximation $A^{[1,0,1]}(\hat{f})$ from $A^{(1,0,1)}(\hat{f})$. This algorithm consists in performing successive projections onto at most $n-1$ affine subspaces.

Recall first that

$$\begin{aligned} V^{(1,0,1)} &= \Big\{\sum_{i=1}^n \omega_i\, x_i \,\Big|\, \sum_{i=1}^n \omega_i = 1\Big\} \\ V^{[1,0,1]} &= \Big\{\sum_{i=1}^n \omega_i\, x_i \,\Big|\, \sum_{i=1}^n \omega_i = 1 \text{ and } \omega_i \geq 0 \text{ for all } i \in N\Big\}. \end{aligned}$$

We can readily see that $V^{[1,0,1]}$ is a polytope in $V^{(1,0,1)}$ whose vertices are $x_1, \ldots, x_n$. The distance between any two vertices is constant

$$\|x_i - x_j\| = \Big[\int_{[0,1]^n} (x_i - x_j)^2 \, dx\Big]^{1/2} = \frac{1}{\sqrt{6}}, \quad i, j \in N,\ i \neq j,$$

and $V^{[1,0,1]}$ is a regular simplex. In the sequel, we will denote this simplex by P. We thus have $\dim(P) = \dim(V^{(1,0,1)}) = n-1$ (see Figure 7.1).

We have $\hat{f}^{(1,0,1)} = A^{(1,0,1)}(\hat{f}) \in V^{(1,0,1)} \setminus P$ if and only if there exists $i \in N$ such that $a_i^{(1,0,1)} < 0$. In this case, $A^{[1,0,1]}(\hat{f})$ can be obtained by projecting $A^{(1,0,1)}(\hat{f})$ onto P.

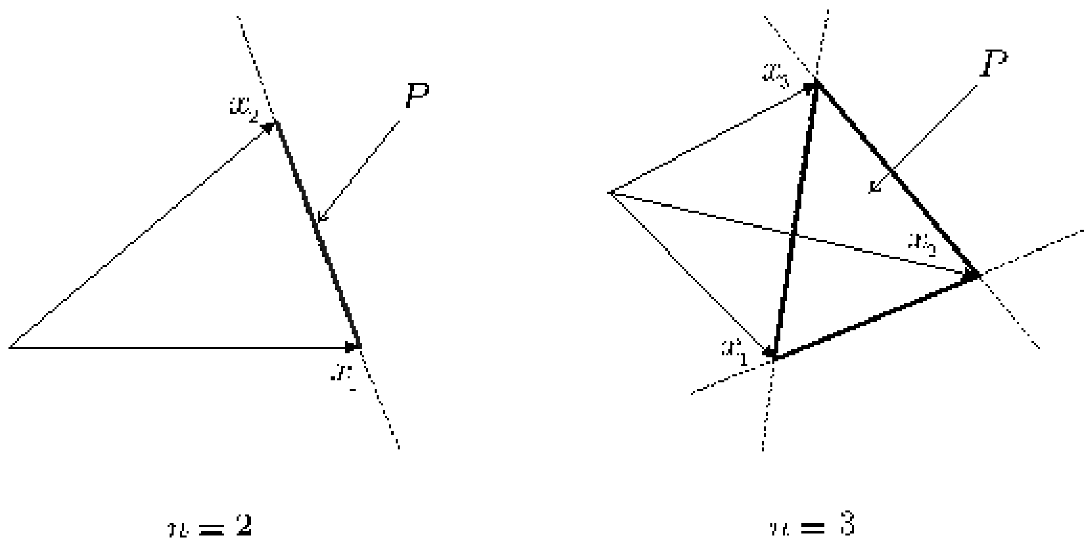


Figure 7.1: Regular simplex P

It is clear that the affine hull of $n-1$ vertices of P is an affine subspace of dimension $n-2$. Hence, a facet[2] F of P is the convex hull of $n-1$ of its vertices:

$$F = \operatorname{conv}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \quad \text{for a fixed } j \in N.$$

It is a regular simplex of dimension $n-2$. Obviously, the simplex P has n facets.

If $\hat{g} \in V^{(1,0,1)} \setminus P$ then there exists a facet $F_{\hat{g}}$ of P such that the affine hull $\operatorname{aff}(F_{\hat{g}})$ of its vertices contains $\hat{g}$ or separates $\hat{g}$ from P. We now show that the projection of $\hat{g}$ onto P is in $F_{\hat{g}}$.

Lemma 7.2.3 *Let $\hat{h} \in P$ and let F be a facet of P. Then the projection of $\hat{h}$ onto* $\operatorname{aff}(F)$ *is in F.*

Proof. We can assume without loss of generality that F is the convex hull of $x_1, \ldots, x_{n-1}$. To prove the result, it suffices to show that the projection of x_n onto $\operatorname{aff}(F)$ is in F. This is immediate since this projection is simply given by $\frac{1}{n-1}\sum_{i=1}^{n-1} x_i$. Indeed, it suffices to observe that

$$\langle x_n - \frac{1}{n-1}\sum_{i=1}^{n-1} x_i, x_j - x_k\rangle = 0, \quad j,k \in \{1, \ldots, n-1\}$$

as required. ∎

Theorem 7.2.4 *Let $\hat{g} \in V^{(1,0,1)} \setminus P$. Then the projection of $\hat{g}$ onto P is in $F_{\hat{g}}$.*

Proof. Let $\hat{h}$ be the projection of $\hat{g}$ onto P and assume that $\hat{h} \notin F_{\hat{g}}$. Let $\hat{h}'$ be the projection of $\hat{h}$ onto $\operatorname{aff}(F_{\hat{g}})$. By Lemma 7.2.3, we have $\hat{h}' \in F_{\hat{g}}$.

Next, since

$$\operatorname{aff}(F_{\hat{g}}) = \{\hat{w} \in V^{(1,0,1)} \mid \langle \hat{h} - \hat{h}', \hat{w} - \hat{h}'\rangle = 0\},$$

[2]According to the usual definition, a face F of P is a *facet* of P if $\dim(F) = \dim(P) - 1$.

we have $\langle \hat{h} - \hat{h}', \hat{g} - \hat{h}' \rangle \leq 0$ and

$$\begin{aligned} \|\hat{h} - \hat{g}\|^2 &= \|\hat{h} - \hat{h}'\|^2 + \|\hat{h}' - \hat{g}\|^2 - 2\langle \hat{h} - \hat{h}', \hat{g} - \hat{h}' \rangle \\ &> \|\hat{h}' - \hat{g}\|^2 \end{aligned}$$

with $\hat{h}, \hat{h}' \in P$. This is absurd. ■

By Theorem 7.2.4, the projection of $\hat{g}$ onto P can be obtained by first projecting $\hat{g}$ onto $\mathrm{aff}(F_{\hat{g}})$ and then projecting, if necessary, the obtained function onto $F_{\hat{g}}$. Note that if more than one affine hull contain $\hat{g}$ or separate it from P then the projection onto P is clearly in the intersection of the corresponding facets.

Therefore, if $A^{(1,0,1)}(\hat{f}) \in V^{(1,0,1)} \setminus P$, one can obtain $A^{[1,0,1]}(\hat{f})$ by means of the following algorithm:

Prestep. $P := V^{[1,0,1]}$, $\hat{g} := A^{(1,0,1)}(\hat{f})$.

Step 1. $F_{\hat{g}}^{\cap} :=$ intersection of all the facets of P whose affine hull contains $\hat{g}$ or separates it from P.

Step 2. $\hat{h} :=$ projection of $\hat{g}$ onto $\mathrm{aff}(F_{\hat{g}}^{\cap})$.

Step 3. If $\hat{h} \in F_{\hat{g}}^{\cap}$ then $\hat{h}$ is the projection of $A^{(1,0,1)}(\hat{f})$ onto P, $\longrightarrow$ stop, else $P \leftarrow F_{\hat{g}}^{\cap}$, $\hat{g} \leftarrow \hat{h}$, return to Step 1.

It is clear that this algorithm terminates with the desired projection. Furthermore, since the dimension of the problem decreases of at least one at each iteration, the projection will be attained with at most $n-1$ iterations, see Figure 7.2.

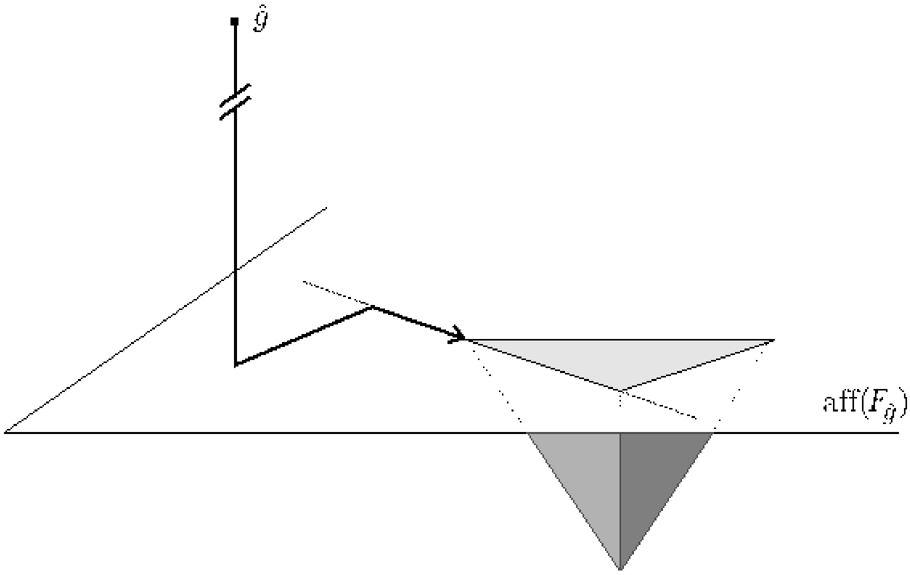


Figure 7.2: Successive projections onto affine hulls

Now, let us turn to the effective computation of the projection. The following lemma will be useful: it characterizes the affine hull of the facets of P.

Lemma 7.2.4 *Let F be a facet of P. Then, for any $j \in N$,*

$$F = \operatorname{conv}\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n\}$$

if and only if

$$\operatorname{aff}(F) = \Big\{\hat{w} \in V^{(1,0,1)} \,\Big|\, \hat{w}(x) = \sum_{i=1}^{n} c_i\, x_i \text{ and } c_j = 0\Big\}.$$

Proof. By definition of the affine hull, we immediately have

$$\operatorname{aff}(F) = \Big\{\hat{w} \in V^{(1,0,1)} \,\Big|\, \hat{w}(x) = \sum_{i \in N \setminus j} c_i\, x_i \text{ and } \sum_{i \in N \setminus j} c_i = 1\Big\}$$

which is sufficient. ■

Let $\hat{g} \in V^{(1,0,1)} \setminus P$ with

$$\hat{g}(x) = \sum_{i=1}^{n} a_i^*\, x_i, \quad x \in [0,1]^n.$$

Set $R := \{i \in N \,|\, a_i^* \leq 0\}$ and $r := |R|$. According to Lemma 7.2.4, there are r facets that contain $\hat{g}$ or separate it from P. If $F_{\hat{g}}^{\cap}$ denotes the intersection of these facets then we have

$$\operatorname{aff}(F_{\hat{g}}^{\cap}) = \Big\{\hat{w} \in V^{(1,0,1)} \,\Big|\, \hat{w}(x) = \sum_{i \in N \setminus R} c_i\, x_i\Big\}.$$

Now, let us fix $j \in N \setminus R$. A basis of $\operatorname{aff}(F_{\hat{g}}^{\cap})$ is given by

$$\{x_i - x_j \,|\, i \in N \setminus (R \cup j)\},$$

and hence, the projection of $\hat{g}$ onto $\operatorname{aff}(F_{\hat{g}}^{\cap})$ is given by the unique solution $\hat{h} \in \operatorname{aff}(F_{\hat{g}}^{\cap})$ of the linear system:

$$\langle \hat{g} - \hat{h}, x_i - x_j \rangle = 0, \quad i \in N \setminus (R \cup j). \tag{7.27}$$

However, for $i \in N \setminus (R \cup j)$, we have

$$\begin{aligned}
\langle \hat{g} - \hat{h}, x_i - x_j \rangle &= \int_{[0,1]^n} \Big[\sum_{\nu \in N} a_\nu^*\, x_\nu - \sum_{\nu \in N \setminus R} c_\nu\, x_\nu\Big](x_i - x_j)\, dx \\
&= \int_{[0,1]^n} \Big[\sum_{\nu \in R} a_\nu^*\, x_\nu + \sum_{\nu \in N \setminus R} (a_\nu^* - c_\nu)\, x_\nu\Big](x_i - x_j)\, dx \\
&= \frac{1}{3}(a_i^* - c_i) + \frac{1}{4} \sum_{\nu \in N \setminus (R \cup i)} (a_\nu^* - c_\nu) - \frac{1}{3}(a_j^* - c_j) - \frac{1}{4} \sum_{\nu \in N \setminus (R \cup j)} (a_\nu^* - c_\nu) \\
&= \frac{1}{12}(a_i^* - c_i) - \frac{1}{12}(a_j^* - c_j) \qquad (\text{since } \sum_{\nu \in N} a_\nu^* = \sum_{\nu \in N \setminus R} c_\nu = 1)
\end{aligned}$$

and the system (7.27) writes

$$c_i = a_i^* + c_j - a_j^*, \quad i \in N \setminus R.$$

Summing over $i \in N \setminus R$ provides

$$\begin{aligned} 1 = \sum_{i \in N \setminus R} c_i &= \sum_{i \in N \setminus R} a_i^* + (n-r)(c_j - a_j^*) \\ &= 1 - \sum_{i \in R} a_i^* + (n-r)(c_j - a_j^*) \end{aligned}$$

Hence, the solution of (7.27) is

$$\hat{h}(x) = \sum_{j \in N \setminus R} c_j \, x_j$$

with

$$c_j = a_j^* + \frac{1}{n-r} \sum_{i \in R} a_i^*, \quad j \in N \setminus R.$$

Consequently, the resolution of the problem can be done as follows:

Let $a_1^{(1,0,1)}, \ldots, a_n^{(1,0,1)}$ be the coefficients of $A^{(1,0,1)}(\hat{f})$, at least one of them being strictly negative. The projection $A^{[1,0,1]}(\hat{f})$ is obtained by means of the following algorithm, which has running time in $O(n^2)$:

Prestep. $a_i^* := a_i^{(1,0,1)}$ $(i \in N)$.

Step 1. $R := \{i \in N \mid a_i^* \leq 0\}$, $r := |R|$.

Step 2. $c_j := a_j^* + \dfrac{1}{n-r} \sum_{i \in R} a_i^* \quad (j \in N \setminus R)$.

Step 3. If $c_j \geq 0$ for all $j \in N \setminus R$

then $\hat{h}(x) = \sum_{j \in N \setminus R} c_j \, x_j$ is the desired projection, $\longrightarrow$ stop,

else $a_i^* \leftarrow \begin{cases} 0, & \text{if } i \in R, \\ c_i, & \text{if } i \in N \setminus R, \end{cases}$

return to Step 1.

This algorithm can be slightly simplified in its writing. Moreover, including it in the complete procedure of resolution leads to a flow chart which can be implemented very easily, see Figure 7.3.

Example 7.1 (continued) *For the function $\hat{f} : [0,1]^4 \to \mathbb{R}$ defined by*

$$\begin{aligned} \hat{f}(x) &= \frac{3}{10} \Big[x_1 + x_2 + x_3 + (x_1 \wedge x_2) + (x_1 \wedge x_3) + (x_2 \wedge x_3) \Big] \\ &\quad - \frac{21}{25} \, (x_1 \wedge x_2 \wedge x_3) + \frac{1}{25} \, (x_1 \wedge x_2 \wedge x_3 \wedge x_4), \end{aligned}$$

the closest weighted arithmetic mean to $\hat{f}$ is given by

$$(A^{[1,0,1]} \hat{f})(x) = \frac{1}{3} \, (x_1 + x_2 + x_3).$$

We can observe in Example 7.1 that $\hat{f}$ and all its approximations are symmetric functions in x_1, x_2, x_3. This is actually a general principle which is only due to the uniqueness of the best approximations, see [95, Proposition 2.5]:

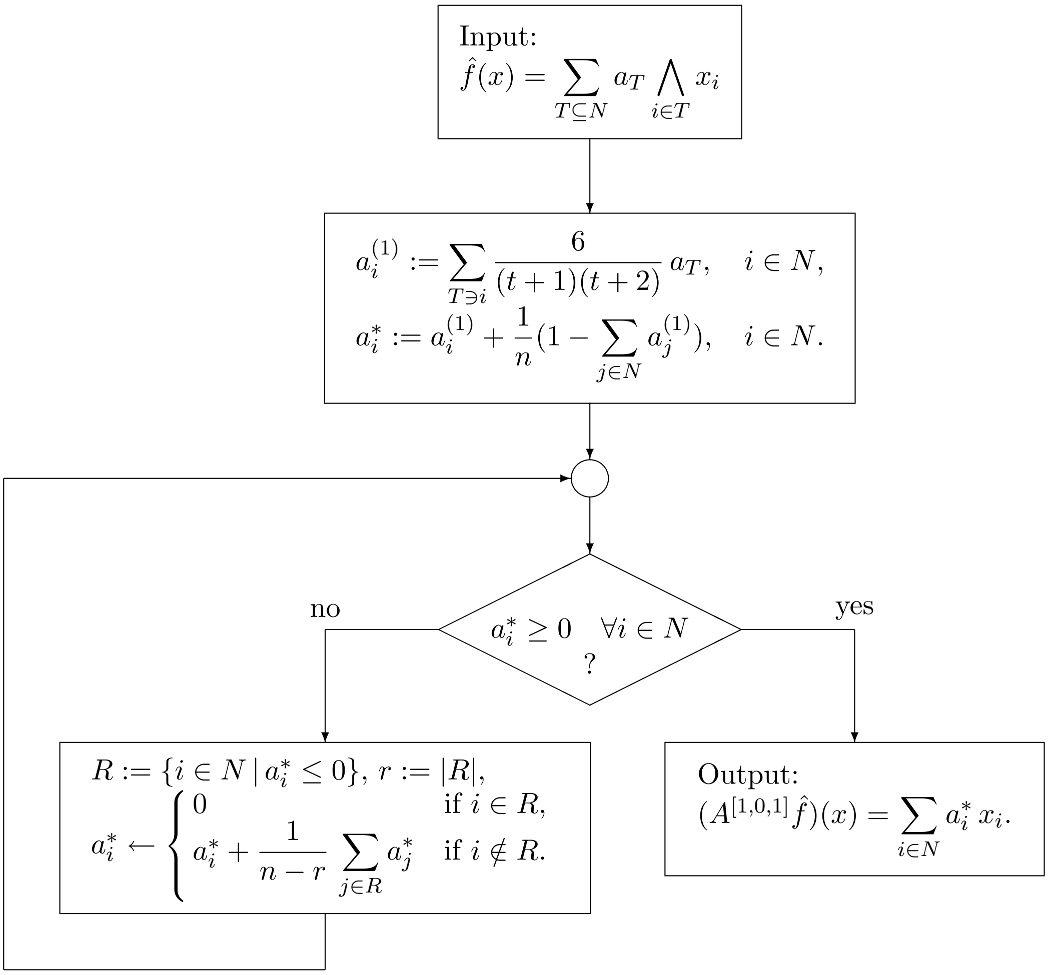


Figure 7.3: Closest weighted arithmetic mean to a Lovász extension (flow chart)

Proposition 7.2.1 *If $\pi \in \Pi_n$ is a symmetry of $\hat{f}$, that is $\hat{f}(x) = \hat{f}([x]_\pi)$ for all $x \in [0,1]^n$, then π is also a symmetry of $A^{(k)}(\hat{f}), A^{(k,\alpha)}(\hat{f}), A^{(k,\alpha,\beta)}(\hat{f})$ and $A^{[k,\alpha,\beta]}(\hat{f})$.*

Before closing this section, we present an easy test allowing to verify the rightness of the obtained projection $A^{[1,0,1]}(\hat{f})$. For this purpose, we need a result that characterizes the projection onto a convex set as solving a so-called variational inequality. The statement is the following, see Hiriart-Urruty and Lemaréchal [101, Chap. 3, Theorem 3.1.1].

Theorem 7.2.5 *Let C be any closed convex set in $\mathbb{R}^n$. A point $y_x \in C$ is the projection of $x \in \mathbb{R}^n \setminus C$ if and only if*

$$\langle x - y_x, y - y_x \rangle \leq 0 \quad \text{for all } y \in C. \tag{7.28}$$

The condition (7.28) simply expresses that the angle between $y - y_x$ and $x - y_x$ is obtuse, for any $y \in C$. When C is a polytope, it is clear that the inequality must be checked only over the vertices of the polytope.

Translating this to our projection problem leads to the following result: only n quadratic inequalities must be checked.

Theorem 7.2.6 *Let $\hat{f}$ be any Lovász extension and let $\hat{f}^{(1,0,1)} = A^{(1,0,1)}(\hat{f})$. Then the function*

$$\hat{h}(x) = \sum_{i=1}^{n} c_i\, x_i, \quad \text{with} \ \sum_{i=1}^{n} c_i = 1 \ \text{and} \ c_i \geq 0 \ \forall i,$$

corresponds to the approximation $A^{[1,0,1]}(\hat{f})$ if and only if

$$a_j^{(1,0,1)} - c_j \leq \sum_{i=1}^{n} (a_i^{(1,0,1)} - c_i)\, c_i, \quad j = 1, \ldots, n.$$

Proof. Since the vertices of P are the x_j's, the inequality in (7.28) becomes

$$\int_{[0,1]^n} [\hat{f}^{(1,0,1)}(x) - \hat{h}(x)][x_j - \hat{h}(x)]\, dx \leq 0, \quad j = 1, \ldots, n.$$

However, we have, for all $j = 1, \ldots, n$,

$$\begin{aligned}
&\int_{[0,1]^n} [\hat{f}^{(1,0,1)}(x) - \hat{h}(x)][x_j - \hat{h}(x)]\, dx \\
= \ & \int_{[0,1]^n} \Big[\sum_{i=1}^{n} (a_i^{(1,0,1)} - c_i)\, x_i\, x_j - \sum_{i,l=1}^{n} (a_i^{(1,0,1)} - c_i)\, c_l\, x_i\, x_l\Big]\, dx \\
= \ & \frac{1}{3}(a_j^{(1,0,1)} - c_j) + \frac{1}{4} \sum_{\substack{i=1 \\ i \neq j}}^{n} (a_i^{(1,0,1)} - c_i) - \frac{1}{4} \sum_{i=1}^{n} (a_i^{(1,0,1)} - c_i) \sum_{\substack{l=1 \\ l \neq i}}^{n} c_l - \frac{1}{3} \sum_{i=1}^{n} (a_i^{(1,0,1)} - c_i)\, c_i
\end{aligned}$$

Since $\sum_{i=1}^{n} a_i^{(1,0,1)} = \sum_{i=1}^{n} c_i = 1$, the latter expression becomes

$$\frac{1}{12}(a_j^{(1,0,1)} - c_j) - \frac{1}{12} \sum_{i=1}^{n} (a_i^{(1,0,1)} - c_i)\, c_i,$$

which leads to the result. ∎

7.3 Applications to Multicriteria Decision Making

Suppose that we want to approximate a given Choquet integral by a k-order Choquet integral for a fixed integer $k \leq n$. In a multicriteria decision problem modelled by a Choquet integral, the actually meaningful question is the following: *what is the best k-order fuzzy measure $\mu^{(k)}$ such that*

$$\int_{[0,1]^n} [\mathcal{C}_\mu(x) - \mathcal{C}_{\mu^{(k)}}(x)]^2\, dx$$

is minimized?

This corresponds to the least squares approximation problem of a particular Lovász extension under the boundary and monotonicity constraints.

When $k = 1$, this problem becomes very simple: the approximation under boundary contraints yields (see Tables 7.2 and 7.4):

$$a_i^{(1,0,1)} = a_i^{(1)} + \frac{1}{n}\,(1 - \sum_{j=1}^{n} a_j^{(1)}), \quad i \in N, \tag{7.29}$$

with

$$a_i^{(1)} = \sum_{T \ni i} \frac{6}{(t+1)(t+2)}\, a_T, \quad i \in N.$$

If necessary, the increasing approximation is then obtained by using the algorithm proposed in Section 7.2.5 (see Figure 7.3).

The following result gives sufficient conditions for obtaining $a_i^{(1,0,1)} \geq 0$ for all i.

Proposition 7.3.1 *If $\hat{f}$ is such that $a_T \geq 0$ for all $T \subseteq N$ then $a_i^{(1,0,1)} \geq 0$ for all $i \in N$.*

Proof. Consider the partial minimum $\min_S$ associated to a non-empty subset $S \subseteq N$. For such a function we have

$$\begin{aligned} i \in S &: \quad a_i^{(1,0,1)} = \frac{(s-1)(s-2)+6n}{n(s+1)(s+2)} \geq 0, \\ i \notin S &: \quad a_i^{(1,0,1)} = \frac{(s-1)(s-2)}{n(s+1)(s+2)} \geq 0. \end{aligned}$$

As $A^{(1,0,1)}$ is a projection, it is a linear operator. Hence, for $\hat{f}$, we have

$$a_i^{(1,0,1)} = \frac{1}{n} \sum_{T \ni i} \frac{(t-1)(t-2)+6n}{(t+1)(t+2)}\, a_T + \frac{1}{n} \sum_{T \not\ni i} \frac{(t-1)(t-2)}{(t+1)(t+2)}\, a_T, \tag{7.30}$$

which is positive. ■

The problem (7.2) can be solved in a very similar way as that of (7.1). Note that, in this discrete context, formula (7.29) is identically the same, see [95, Theorem 2.10].

7.3.1 An example

Let us go back to the example given in Section 6.2.5. This example involves 3 criteria, and the Möbius representation of the fuzzy measure is given by

$$\begin{array}{llll} a_\emptyset = 0 & a_{\mathrm{M}} = 0.45 & a_{\mathrm{MP}} = -0.40 & a_{\mathrm{MPL}} = -0.10 \\ & a_{\mathrm{P}} = 0.45 & a_{\mathrm{ML}} = 0.15 & \\ & a_{\mathrm{L}} = 0.30 & a_{\mathrm{PL}} = 0.15 & \end{array}$$

Computing the 2-order approximation $A^{[2,0,1]}(\mathcal{C}_\mu)$ (note that $A^{[2,0,1]}(\mathcal{C}_\mu) = A^{(2,0,1)}(\mathcal{C}_\mu)$) leads to the following coefficients $a' = a^{[2,0,1]}$:

$$\begin{array}{ll} a'_{\mathrm{M}} = 0.46666 & a'_{\mathrm{MP}} = -0.45 \\ a'_{\mathrm{P}} = 0.46666 & a'_{\mathrm{ML}} = 0.10 \\ a'_{\mathrm{L}} = 0.31666 & a'_{\mathrm{PL}} = 0.10 \end{array}$$

Starting from this 2-order approximation, we can compute the new global evaluation (with a 2-order Choquet integral):

student	M	P	L	global evaluation
a	18	16	10	13.83333
b	10	12	18	13.66666
c	14	15	15	14.88333

Surprisingly enough, we can see that the value of each of indices

$$\phi_{\mathrm{Sh}}(i), I_{\mathrm{Sh}}(ij), \mathrm{veto}(\mathcal{C}_\mu; i), \mathrm{favor}(\mathcal{C}_\mu; i), \mathrm{orness}(\mathcal{C}_\mu),$$

is unchanged. However, the dispersion has been slightly modified: we now have

$$\mathrm{disp}(\mu') = 0.89094 \quad \text{and} \quad \mathrm{disp}'(\mu') = 0.81097.$$

7.3.2 Links with the Shapley value

We have seen in Section 5.1.2 that the Shapley value corresponding to a given fuzzy measure can be viewed as a linear approximation of the fuzzy measure. Regarding the approximation of a Choquet integral, it can happen that the coefficients of the closest weighted arithmetic mean correspond to the Shapley value. The following theorem gives a useful example.

Theorem 7.3.1 *Let μ be a 2-order fuzzy measure on N. The closest weighted arithmetic mean to $\mathcal{C}_\mu$ is the Shapley integral* Sh_μ.

Proof. This is immediate: we simply have

$$a_i^{(1)} = a_i + \frac{1}{2}\sum_{j\in N\setminus i} a_{ij} = \phi_{\mathrm{Sh}}(i), \quad i \in N,$$

and $a_i^{(1,0,1)} = a_i^{(1)}$ for all $i \in N$. ■

More generally, we have the following.

Proposition 7.3.2 *If $n \leq 3$ or if $a_T = 0$ for all $T \subseteq N$ such that $3 \leq t \leq n-1$, then*

$$a_i^{(1,0,1)} = \phi_{\mathrm{Sh}}(i), \quad i \in N.$$

Proof. Since $\phi_{\mathrm{Sh}}(i) = \sum_{T\ni i}\frac{1}{t}a_T$, we have, using (7.30),

$$a_i^{(1,0,1)} - \phi_{\mathrm{Sh}}(i) = \frac{1}{n}\sum_{T\ni i}\frac{(t-1)(t-2)(t-n)}{t(t+1)(t+2)}\,a_T + \frac{1}{n}\sum_{T\not\ni i}\frac{(t-1)(t-2)}{(t+1)(t+2)}\,a_T.$$

This expression is 0 whenever $a_T = 0$ for all $T \subseteq N$ such that $3 \leq t \leq n-1$. ■

It should be noted that the Shapley value is not obtained in the general case. Example 7.1 concerns the approximation of a Choquet integral, and the Shapley value of the associated fuzzy measure is given by

$$\phi_{\mathrm{Sh}}(1) = \phi_{\mathrm{Sh}}(2) = \phi_{\mathrm{Sh}}(3) = \frac{33}{100} \text{ and } \phi_{\mathrm{Sh}}(4) = \frac{1}{100}$$

whereas the closest weighted arithmetic mean to the Choquet integral has the coefficients

$$a_1^{[1,0,1]} = a_2^{[1,0,1]} = a_3^{[1,0,1]} = \frac{1}{3} \text{ and } a_4^{[1,0,1]} = 0.$$

The following example shows that the Shapley value can be not obtained even when the approximation under boundary conditions is increasing.

Example 7.3.1 *Consider the Choquet integral $\mathcal{C}_\mu : [0,1]^4 \to [0,1]$ defined by*

$$\mathcal{C}_\mu(x) = \frac{1}{10}(x_1\wedge x_2\wedge x_3) + \frac{2}{10}(x_1\wedge x_2\wedge x_4) + \frac{3}{10}(x_1\wedge x_3\wedge x_4) + \frac{4}{10}(x_2\wedge x_3\wedge x_4).$$

We then have

$$\begin{aligned}(A^{(1)}\mathcal{C}_\mu)(x) &= -\frac{1}{5} + \frac{9}{50}x_1 + \frac{21}{100}x_2 + \frac{6}{25}x_3 + \frac{27}{100}x_4,\\ (A^{(1,0,1)}\mathcal{C}_\mu)(x) &= \frac{57}{650}x_1 + \frac{153}{1300}x_2 + \frac{48}{325}x_3 + \frac{231}{1300}x_4,\end{aligned}$$

and

$$\phi_{\mathrm{Sh}}(1) = \frac{1}{5}, \phi_{\mathrm{Sh}}(2) = \frac{7}{30}, \phi_{\mathrm{Sh}}(3) = \frac{4}{15}, \phi_{\mathrm{Sh}}(4) = \frac{3}{10}.$$

As already observed in Theorem 5.1.3, the Shapley value can be obtained in the discrete case by weighting the distance. Grabisch [86] observed that the distance used in (7.2) does not take into account the fact that there are elements inside each subset, and that a single element is involved several times in different subsets, especially with subsets of around $n/2$ elements, which are the most numerous. This means that a weighted distance taking into account this combinatorial aspect (as the one for the Shapley value (5.2)) should be used to avoid this effect. Thus, in the discrete case, we can have two attitudes:

- if in a given problem, we reason on elements, and set functions are involved, which should be approximated, then use a weighted distance as for the Shapley value (approximation on N).
- if in a given problem where set functions are involved, elements are unimportant or not relevant, then use a non weighted distance (approximation on 2^N).

In the continuous case, no weighted distance has been found yet for obtaining the Shapley integral as linear approximation. Although this is still an open problem, a suitable distance (if any) should have a reasonable interpretation.

7.3.3 Approximations of OWA operators

Let us consider an OWA_ω operator, that is a symmetric Choquet integral (see Theorem 4.2.16). By symmetry, we immediately see that (see Proposition 7.2.1)

$$(A^{(1,0,1)}\mathrm{OWA}_\omega)(x) = (A^{[1,0,1]}\mathrm{OWA}_\omega)(x) = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad x \in [0,1]^n.$$

Obviously, this is a rather crude approximation, so we turn to the second order approximation. Using Table 7.4, one can show that

$$(A^{(2,0,1)}\mathrm{OWA}_\omega)(x) = C\sum_{i\in N} x_i + D\sum_{\{i,j\}\subseteq N}(x_i \wedge x_j), \quad x \in [0,1]^n, \tag{7.31}$$

with

$$\begin{aligned} C &= \sum_{j=1}^{n} \frac{-10n^4+30jn^3-19n^3+16n^2+90j^2n-120jn+31n-60j^3+90j^2-30j+6}{n(n+1)(n+2)(5n^2-4n+3)}\,\omega_j \\ D &= \sum_{j=1}^{n} \frac{30(n-2j+1)(n^3+n^2+2jn-2n-2j^2+2j)}{(n-1)n(n+1)(n+2)(5n^2-4n+3)}\,\omega_j. \end{aligned}$$

We observe that such an approximation is heavy to compute. Moreover, by (4.6), the approximation (7.31) is increasing if and only if

$$C + t\,D \geq 0, \quad t = 0, \dots, n-1,$$

that is, if and only if

$$C \geq 0 \quad \text{and} \quad C + (n-1)\,D \geq 0.$$

Example 7.3.2 *The* OWA_ω *operator defined on* $[0,1]^4$ *by the weight vector*

$$\omega = (\frac{1}{15}, \frac{2}{15}, \frac{4}{15}, \frac{8}{15})$$

is such that

$$(A^{(2,0,1)}\mathrm{OWA}_\omega)(x) = (A^{[2,0,1]}\mathrm{OWA}_\omega)(x) = \frac{643}{1340}\sum_{i=1}^{4} x_i - \frac{154}{1005}\sum_{\{i,j\}\subseteq\{1,2,3,4\}}(x_i \wedge x_j).$$

Notice that Grabisch [86] proposed the following 2-nd order approximation:

$$\mathrm{OWA}_\omega(x) \approx \frac{1}{n}\sum_{i\in N} x_i + \frac{\omega_1-\omega_n}{n-1}\sum_{\{i,j\}\subseteq N}(x_i \wedge x_j).$$

However, such an approximation fails to agree with OWA_ω at $x = \underline{1}$ and fails to be increasing.

On the basis of Theorem 6.3.3, we propose the following form:

$$\begin{aligned}\mathrm{OWA}_\omega(x) &\approx \frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{2}\,\frac{\omega_1-\omega_n}{n-1}\sum_{i=1}^{n}(n+1-2\,i)\,x_{(i)}\\ &= \Big(\frac{1}{n}-\frac{\omega_1-\omega_n}{2}\Big)\sum_{i\in N} x_i + \frac{\omega_1-\omega_n}{n-1}\sum_{\{i,j\}\subseteq N}(x_i \wedge x_j).\end{aligned}$$

Of course, this approximation corresponds to a 2-order OWA, but fails to be the closest to OWA_ω. Moreover, it only involves the weights ω_1 and ω_n.

According to Proposition 6.3.2, we note that this approximation consists in replacing the weight function ω_i by a linear weight function $\omega'_i = a\,i + b$ with slope

$$a = \frac{\omega_n - \omega_1}{n-1}.$$

For the operator proposed in Example 7.3.2, this approximation is given by

$$\frac{29}{60}\sum_{i=1}^{4} x_i - \frac{7}{45}\sum_{\{i,j\}\subseteq\{1,2,3,4\}}(x_i \wedge x_j).$$

It is a 2-order OWA operator with weight vector $\omega' = (\frac{1}{60}, \frac{31}{180}, \frac{59}{180}, \frac{29}{60})$, see Figure 7.4.

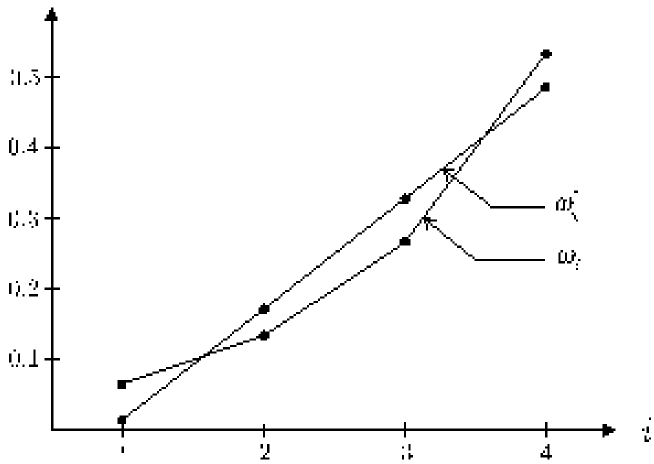


Figure 7.4: Linear approximation of the weights ω_i

Notation and symbols

Miscellaneous symbols

The symbol := means that the left-hand side of the equation is *defined* by the right-hand side.

The set of *non-negative integers* is denoted by $\mathbb{N}$, that of *strictly positive integers* by $\mathbb{N}_0$, that of real numbers by $\mathbb{R}$. We then set $\mathbb{R}^+ := \{x \in \mathbb{R} \,|\, x \geq 0\}$ and $\mathbb{R}_0^+ := \{x \in \mathbb{R} \,|\, x > 0\}$.

For any $n \in \mathbb{N}_0$, $N_n := \{1, \ldots, n\}$.

Π_n denotes the set of permutations of N_n. For any $\pi \in \Pi_n$ and any $S \subseteq N_n$, we set $\pi(S) := \{\pi(i) \,|\, i \in S\}$.

For all $S, T \subseteq N_n$, set difference of S and T is denoted by $S \setminus T$. 2^N indicates the power set of N, i.e. the set of all subsets in N. Cardinality of sets $S, T, \ldots$ is denoted whenever possible by corresponding lower cases $s, t, \ldots$, otherwise by the standard notation $|S|, |T|, \ldots$.

The notation $S \subsetneq T$ means $S \subset T$ and $S \neq T$.

$\wedge, \vee$ denote respectively the minimum and maximum operations.

Aggregation operators

E, F denote real intervals, finite or infinite. E represents the definition set of values to be aggregated.

E° denotes the *interior* of E, that is the corresponding open set.

For all $S \subseteq N_n$ and all real intervals E_i $(i \in N_n)$, we set $E_S := \times_{i \in S} E_i$.

$M^{(n)}$ represents an *aggregation operator*, that is a function $M^{(n)} : E^n \to F$. A sequence $M = (M^{(n)})_{n \in \mathbb{N}_0}$ of aggregation operators $M^{(n)} : E^n \to F$ is called an *extended aggregation operator.*

$A_n(E, F)$ denotes the set of all aggregation operators from E^n to F. $A(E, F)$ denotes the set of all extended aggregation operators whose the n-th element is in $A_n(E, F)$.

For all $k \in \mathbb{N}_0$, we set $k \odot x := x, \ldots, x$ (k times). For instance,

$$M(3 \odot x, 2 \odot y) = M(x, x, x, y, y).$$

The vector $(x, \ldots, x)$ in $\mathbb{R}^n$ is simply denoted by $\underline{x}$.

For all $S \subseteq N_n$, the *characteristic vector* (or *incidence vector*) of S in $\{0,1\}^n$ is defined by

$$e_S^{(n)} := (x_1, \ldots, x_n) \in \{0,1\}^n \quad \text{with} \quad x_i = 1 \Leftrightarrow i \in S.$$

Geometrically, the characteristic vectors are the 2^n vertices of the hypercube $[0,1]^n$. The complementary characteristic vector of $S \subseteq N_n$ is defined by

$$\overline{e}_S^{(n)} := e_{N_n \setminus S}^{(n)}.$$

Moreover, we set $\theta_S^{(n)} := M^{(n)}(e_S^{(n)})$ and $\overline{\theta}_S^{(n)} := M^{(n)}(\overline{e}_S^{(n)})$ for all $S \subseteq N_n$.

For any $\pi \in \Pi_n$, we define

$$\mathcal{O}_\pi := \{x \in \mathbb{R}^n \mid x_{\pi(1)} \leq \cdots \leq x_{\pi(n)}\}.$$

The simplex of $[0,1]^n$ associated to π is then defined by

$$\mathcal{B}_\pi := \mathcal{B}_\pi \cap [0,1]^n,$$

that is the convex hull of

$$\{e^{(n)}_{\{\pi(i),\ldots,\pi(n)\}}\}_{i=1}^{n+1}.$$

Given a vector $(x_1,\ldots,x_n)$ and a permutation $\pi \in \Pi_n$, the notation $[x_1,\ldots,x_n]_\pi$ means $x_{\pi(1)},\ldots,x_{\pi(n)}$, that is, the permutation π of the indices.

Given a vector $(x_1,\ldots,x_n)$, let $(\cdot)$ denote the permutation of N_n which arranges all the elements $x_1,\ldots,x_n$ by increasing values: that is, $x_{(1)} \leq \ldots \leq x_{(n)}$.

The so-called *median* of an odd number of values $x_1,\ldots,x_{2k-1}$ is simply defined by

$$\text{median}(x_1,\ldots,x_{2k-1}) := x_{(k)}.$$

For any real interval E, $\Phi(E)$ denotes the automorphism group of E, that is the group of all strictly increasing bijections $\phi : E \to E$. $\Phi'(E)$ denotes the set of all strictly increasing functions $\phi : E \to E$.

Binary relations

A binary relation on a finite set is
- a *total preorder* (weak order, linear quasi-order) if it is strongly complete and transitive,
- a *partial preorder* (quasi-order) if it is reflexive and transitive.

Simplifications

When there is no fear of ambiguity, the superscript (n) is omitted in $M^{(n)}$, $e_S^{(n)}$, etc. It is used only to stress the dependency on the number of terms in the aggregation. Moreover, N often stands for N_n and x for the vector $(x_1,\ldots,x_n)$.

For any $x \in E^n$, the notation $\phi(x)$ means $(\phi(x_1),\ldots,\phi(x_n))$.

If no confusion can arise, we often use subscripts for arguments of set functions, e.g. writing μ_S, a_S instead of $\mu(S), a(S)$. However, we do not apply this rule for games, since v_T represents the unanimity game for T.

We often omit braces for singletons, e.g. writing $S \cup i$ instead of $S \cup \{i\}$, and μ_i instead of $\mu_{\{i\}}$. Also, for pairs, triples, we often write ij, ijk instead of $\{i,j\}$, $\{i,j,k\}$, as for example $S \cup ijk$.

Glossary of aggregation properties

(A)	Associativity, p. 24
(AD)	Autodistributivity, p. 28
(Add)	Additivity, p. 22
(B)	Bisymmetry, p. 28
(BOM)	Bisymmetry for orderable matrices, p. 30
(CM, CM')	Comparison meaningfulness for ordinal values, p. 20
(CMIS)	Comparison meaningfulness for ordinal values with independent scales, p. 20
(Co)	Continuity, p. 12
(CoAdd)	Comonotonic additivity, p. 23
(CoMax)	Comonotonic maxitivity, p. 23
(CoMin)	Comonotonic minitivity, p. 23
(Comp)	Compensativeness, p. 13
(Conj)	Conjunctiveness, p. 13
(D)	Decomposability, p. 25
(Disj)	Disjunctiveness, p. 13
(Ext)	Extension, p. 162
(GB)	General bisymmetry, p. 28
(GBOM)	General bisymmetry for orderable matrices, p. 30
(Id)	Idempotence, p. 13
(III)	Independent interval scales for the independent variables and interval scale for the dependent variable, p. 18
(In)	Increasingness, p. 12
(IRR)	Independent ratio scales for the independent variables and ratio scale for the dependent variable, p. 17
(ISUII)	Independent interval scales with same unit for the independent variables and interval scale for the dependent variable, p. 17
(ISZII)	Independent interval scales with same zero for the independent variables and interval scale for the dependent variable, p. 17
(LM)	Linearity w.r.t. the fuzzy measure, p. 161
(Max)	Maxitivity, p. 22
(Min)	Minitivity, p. 22
(OS, OS')	Ordinal stability, p. 19
(Rec)	Reciprocal property, p. 18
(SD)	Strong decomposability, p. 26
(Sep)	Separability, p. 31
(SId)	Self-identity, p. 30
(SII)	Same interval scales for the independent variables and interval scale for the dependent variable, p. 17

(SIn)	Strict increasingness, p. 12
(SMax)	Stability for maximum with a constant vector, p. 20
(SMaxB)	Stability for maximum between Boolean and constant vectors, p. 21
(SMin)	Stability for minimum with a constant vector, p. 20
(SMinB)	Stability for minimum between Boolean and constant vectors, p. 21
(SPL)	Stability for the admissible positive linear transformations, p. 16
(SRR)	Same ratio scales for the independent variables and ratio scale for the dependent variable, p. 17
(SSi)	Stability for the admissible similarity transformations, p. 16
(SSN)	Stability for the standard negation, p. 18
(STr)	Stability for the admissible translations, p. 16
(Sy)	Symmetry, p. 11
(UIn)	Unanimous increasingness, p. 12
(WId)	Weak idempotence, p. 13

Bibliography

[1] M. Abramowitz and I.A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, (Dover Publications, New York, 7th printing, 1970).

[2] J. Aczél, On mean values, *Bulletin of the American Math. Society* **54** (1948) 392–400.

[3] J. Aczél, Sur les opérations définies pour des nombres réels, *Bull. Soc. Math. de France* **76** (1948) 59–64.

[4] J. Aczél, *Lectures on functional equations and applications*, (Academic Press, New York, 1966).

[5] J. Aczél, On weighted synthesis of judgements, *Aequationes Mathematicae* **27** (1984) 288–307.

[6] J. Aczél and C. Alsina, Synthesizing judgments: a functional equations approach, *Math. Modelling* **9** (1987) 311–320.

[7] J. Aczél and J. Dhombres, *Functional equations in several variables with applications to mathematics, information theory and to the natural and social sciences*, (Cambridge Univ. Press, Cambridge, 1989).

[8] J. Aczél and F.S. Roberts, On the possible merging functions, *Math. Social Sciences* **17** (1989) 205–243.

[9] J. Aczél, F.S. Roberts and Z. Rosenbaum, On scientific laws without dimensional constants, *Journal of Math. Analysis and Appl.* **119** (1986) 389–416.

[10] J. Aczél and T.L. Saaty, Procedures for synthesizing ratio judgements, *J. Math. Psychol.* **27** (1983) 93–102.

[11] C. Alsina, G. Mayor, M.S. Tomás and J. Torrens, A characterization of a class of aggregation functions, *Fuzzy Sets and Systems* **53** (1993) 33–38.

[12] C. Alsina, E. Trillas and L. Valverde, On some logical connectives for fuzzy sets theory, *J. Math. Anal. Appl.* **93** (1) (1983) 15–26.

[13] T. Arnould and A. Ralescu, From "and" to "or", *Lecture Notes in Comput. Sci.*, No. 682, Berlin, 1993, pp. 119–128.

[14] K.J. Arrow, Social choice and individual values, (John Wiley & Sons, New York, 1951).

[15] J.F. Banzhaf, Weighted voting doesn't work: A mathematical analysis, *Rutgers Law Review* **19** (1965) 317–343.

[16] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear programming, theory and algorithms*, (Wiley, New York, 2nd edition, 1993.)

[17] R.E. Bellman and L.A. Zadeh, Decision-making in a fuzzy environment, *Management Science* 17:4 (1970) 141–164.

[18] G. Bemporad, Sul principio della media aritmetica, *Atti della Accademia Nazionale dei Lincei* (6) **3** (1926) 87–91.

[19] C. Berge, *Principes de combinatoire*, (Dunod, Paris, 1968.) English translation: *Principles of combinatorics*, (Academic Press, New York, 1971.)

[20] D.P. Bertsekas, *Nonlinear programming*, (Athena Scientific, 1995.)

[21] J.M. Blin, Fuzzy relations in group decision theory, *Journal of Cybernetics* **4**, 2 (1974) 17–22.

[22] R.G. Brown, *Smoothing, forecasting and prediction of discrete time series* (Prentice-Hall, Englewood Cliffs, NJ, 1963).

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

[24] A. Charnes, B. Golany, M. Keane, J. Rousseau, Extremal principle solutions of games in characteristic function form: Core, Chebychev and Shapley value generalizations. Report CCS 526. The Center for Cybernetic Studies, University of Texas, 1985.

[25] A. Chateauneuf and J.Y. Jaffray, Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion, *Mathematical Social Sciences* **17** (1989) 263–283.

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

[27] W. Cholewa, Aggregation of fuzzy opinions—an axiomatic approach, *Fuzzy Sets and Systems* **17** (1985) 249–258.

[28] G. Choquet, Theory of capacities, *Annales de l'Institut Fourier* **5** (1953) 131–295.

[29] A.C. Climescu, Sur l'équation fonctionnelle de l'associativité, *Bull. Inst. Politechn. Iasi* **1** (1946) 1–16.

[30] W.G. Cochran and G.M. Cox, *Experimental designs*, (Wiley, New York, 2nd edition, 1957).

[31] V. Cutello and J. Montero, Hierarchies of aggregation operators, working paper, 1993, unpublished.

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

[33] L.M. de Campos, M.T. Lamata and S. Moral, A unified approach to define fuzzy integrals, *Fuzzy Sets and Systems* **39** (1991) 75–90.

[34] B. de Finetti, Sul concetto di media, *Giornale dell'Istituto Italiano degli Attuari* (3) **2** (1931) 369–396.

[35] C. Dellacherie, Quelques commentaires sur les prolongements de capacités, *Lecture Notes in Mathematics*, Vol. 191, 1970.

[36] A.P. Denneberg, *Non-additive measure and integral*, (Kluwer Academic, Dordrecht, 1994).

[37] D. Denneberg and M. Grabisch, Shapley value and interaction index, Mathematik-Arbeitspapiere Nr. 46 (1996), Universität Bremen.

[38] J. Dombi, Basic concepts for a theory of evaluation: the aggregative operator, *European J. Oper. Res.* **10** (1981) 282–293.

[39] I. Dragan, New mathematical properties of the Banzhaf value, *European Journal of Operational Research* **95** (1996) 451–463.

[40] P. Dubey and L.S. Shapley, Mathematical properties of the Banzhaf power index, *Math. of Oper. Res.* **4** (1979) 99–131.

[41] D. Dubois, M. Grabisch, F. Modave and H. Prade, Relating decision under uncertainty and MCDM models. In, *AIII Fall Symposium on Frontiers in Soft Computing and Decisions Systems*, Boston, Ma, November 1997, 30–39.

[42] D. Dubois and J.-L. Koning, Social choice axioms for fuzzy set aggregation, *Fuzzy Sets and Systems* **43** (1991) 257–274.

[43] D. Dubois and H. Prade, *Fuzzy sets and systems* (Academic Press, New York, 1980).

[44] D. Dubois and H. Prade, Criteria aggregation and ranking of alternatives in the framework of fuzzy set theory, in: H.-J. Zimmermann, L.A. Zadeh and B. Gaines (Eds.), *Fuzzy Sets and Decision Making*, TIMS Studies in Management Sciences **20** (1984) 209–240.

[45] D. Dubois et H. Prade, *Théorie des possibilités. Application à la représentation des connaissances en Informatique*, (Masson, Paris, 1985.)

[46] D. Dubois and H. Prade, A review of fuzzy sets aggregation connectives, *Inform. Sci.* **36** (1985) 85–121.

[47] D. Dubois and H. Prade, Weighted minimum and maximum operators in fuzzy set theory, *Information Sciences* **39** (1986) 205–210.

[48] D. Dubois and H. Prade, A set-theoretic view of belief functions: logical operations and approximations by fuzzy sets, *Int. J. General Systems* **12** (1986) 193–226.

[49] D. Dubois and H. Prade, Aggregation of possibility measures, in: *Multiperson Decision Making Using Fuzzy Sets and Possibility Theory*, J. Kacprzyk, M. Fedrizzi (Eds), (Kluwer Academic. Publishers, Dordrecht, 1990).

[50] D. Dubois and H. Prade, Consonant approximations of belief functions, *Int. J. of Approximate Reasoning* **4** (1990) 419–449.

[51] D. Dubois and H. Prade, Possibility theory as a basis for qualitative decision theory. In *IJCAI'95*, pages 19–25, Montréal, Canada, August 1995.

[52] D. Dubois, H. Prade and R. Sabbadin, Qualitative decision theory with Sugeno integrals, *Proc. 14th Conference on Uncertainty in AI*, Madison, Wisconsin, Edited by G. Cooper and S. Moral, Morgan and Kaufmann, San Francisco, 121–128.

[53] D. Dubois, H. Prade and C. Testemale, Weighted fuzzy pattern-matching, *Fuzzy Sets and Systems* **28** (1988) 313–331.

[54] J.J. Dujmovic, Weighted conjunctive and disjunctive means and their application in system evaluation, *Univ. Beograd. Publ. Elektrotechn. Fak.* (1974) 147–158.

[55] J.J. Dujmovic, Extended continuous logic and the theory of complex criteria, *Univ. Beograd. Publ. Elektrotechn. Fak.* (1975) 197–216.

[56] J.J. Dujmovic, Evaluation, comparison, and optimization of hybrid computers using the theory of complex criteria, in *Simulation of Systems* (L. Dekker, Ed., North-Holland, Amsterdam, 1976), pp. 553–566.

[57] H. Dyckhoff and W. Pedrycz, Generalized means as model of compensative connectives, *Fuzzy Sets and Systems* **14** (1984) 143–154.

[58] P.H. Edelman, A note on voting, *Mathematical Social Sciences* **34** (1997) 37–50.

[59] W.M. Faucett, Compact semigroups irreducibly connected between two idempotents, *Proc. Amer. Math. Soc.* **6** (1955) 741–747.

[60] D.P. Filev and R.R. Yager, Fuzzy logic controllers with flexible structures, *Proc. Second Int. Conf. on Fuzzy Sets and Neural Networks*, Iizuka, 1992, pp. 317–320.

[61] D.P. Filev and R.R. Yager, On the issue of obtaining OWA operator weights, *Fuzzy Sets and Systems* **94** (1998) 157–169.

[62] P.C. Fishburn, *Utility theory for decision making*, (New York, Wiley, 1970).

[63] J.C. Fodor, A remark on constructing t-norms, *Fuzzy Sets and Systems* **41** (1991) 195–199.

[64] J.C. Fodor, Strict preference relations based on weak t-norms, *Fuzzy Sets and Systems* **43** (1991) 327–336.

[65] J.C. Fodor, An extension of Fung-Fu's theorem, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 4, No. 3 (1996) 235–243.

[66] J.C. Fodor and S. Ovchinnikov, On aggregation of T-transitive fuzzy binary relations, *Fuzzy Sets and Systems* **72** (1995) 135–145.

[67] J.C. Fodor and J.-L. Marichal, On nonstrict means, *Aequationes Mathematicae* **54** (1997) 308–327.

[68] J.C. Fodor, J.-L. Marichal and M. Roubens, Characterization of the ordered weighted averaging operators, *IEEE Trans. Fuzzy Syst.*, vol. 3, no. 2 (1995) 236–240.

[69] J.C. Fodor and M. Roubens, Aggregation and scoring procedures in multicriteria decision making methods, *Proc. 1st FUZZ'IEEE Cong.*, San Diego, CA (1992) 1261–1267.

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

[71] J.C. Fodor and M. Roubens, Characterization of weighted maximum and some related operations, *Information Sciences* **84** (1995) 173–180.

[72] J.C. Fodor and M. Roubens, On meaningfulness of means, *Journal of Computational and Applied Mathematics* **64** (1995) 103–115.

[73] L. Fuchs, On mean systems, *Acta Math. Acad. Sci. Hungar.* **1** (1950) 303–320.

[74] S. Fujishige, Characterization of subdifferentials of submodular functions and its relation to Lovász extension of submodular functions. Report No. 82241-OR, Inst. für Ökonometrie und Operations Research, Bonn, 1982.

[75] 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*, eds. L.A. Zadeh, K.S. Fu, K. Tanaka and M. Shimura, (Academic Press, New York–San Francisco–London, 1975) 227–256.

[76] I. Gilboa and D. Schmeidler, Canonical representation of set functions, *Mathematics of Operations Research*, Vol. 20, No. 1 (1995) 197–212.

[77] M. Grabisch, On the use of fuzzy integral as a fuzzy connective, *Proceedings of the Second IEEE International Conference on Fuzzy Systems*, San Francisco (1993) 213–218.

[78] M. Grabisch, On equivalence classes of fuzzy connectives : the case of fuzzy integrals, *IEEE Trans. Fuzzy Systems* **3** (1) (1995) 96–109.

[79] M. Grabisch, Fuzzy integral in multicriteria decision making, *Fuzzy Sets and Systems*, **69** (1995) 279–298.

[80] M. Grabisch, The application of fuzzy integrals in multicriteria decision making, *European Journal of Operational Research*, **89** (1996) 445-456.

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

[82] M. Grabisch, Alternative expressions of the discrete choquet integral, *7th IFSA World Congress*, Prague, Czech Republic, June 1997.

[83] M. Grabisch, Alternative representations of discrete fuzzy measures for decision making, *Int. J. of Uncertainty, Fuzziness, and Knowledge Based Systems* **5** (1997) 587–607.

[84] M. Grabisch, Fuzzy integral as a flexible and interpretable tool of aggregation. *Aggregation of evidence under fuzziness*, B. Bouchon-Meunier (ed), (Physica Verlag, 1998), 51–72.

[85] M. Grabisch, On lower and upper approximation of fuzzy measures by k-order additive measures, *8th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'98)*, Paris, 1998.

[86] M. Grabisch, Alternative representations of OWA operators. In: *The Ordered Weighted Averaging Operators: Theory and Applications*, R.R. Yager and J. Kacprzyk (eds), (Kluwer Academic Publisher, 1997), pages 73–85.

[87] M. Grabisch, J.-L. Marichal and M. Roubens, Equivalent representations of a set function with applications to game theory and multicriteria decision making, Preprint 9801, GEMME, Faculty of Economics, University of Liège, Belgium, 1998; also in: Preprint 98.002, Institute of Mathematics, University of Liège, Belgium, 1998.

[88] M. Grabisch, T. Murofushi and M. Sugeno, Fuzzy measure of fuzzy events defined by fuzzy integrals, *Fuzzy Sets and Systems* **50** (1992) 293–313.

[89] M. Grabisch, H.T. Nguyen and E.A. Walker, *Fundamentals of uncertainty calculi with applications to fuzzy inference*, (Kluwer Academic, Dordrecht, 1995).

[90] M. Grabisch and M. Roubens, Equivalent representations of a set function with application to decision making, in: *6th IEEE Int. Conf. on Fuzzy Systems*, Barcelona, Spain, July 1996.

[91] M. Grabisch and M. Roubens, An axiomatic approach of interaction in multicriteria decision making, *5th Eur. Congr. on Intelligent Techniques and Soft Computing (EUFIT'97)*, Aachen, Germany, september 1997.

[92] M. Grabisch and M. Roubens, An axiomatic approach to the concept of interaction among players in cooperative games, *Int. Journal of Game Theory*, submitted.

[93] M. Grabisch and M. Roubens, Probabilistic interactions among players of a cooperative game, working paper.

[94] M. Grabisch and M. Roubens, Application of the Choquet integral in multicriteria decision making, in: *Studies in Fuzziness*, M. Grabisch, T. Murofushi and M. Sugeno (eds), Physica Verlag, to appear.

[95] P.L. Hammer and R. Holzman, Approximation of pseudo-Boolean functions; applications to game theory, *ZOR - Methods and Models of Operations Research* **36** (1992) 3–21.

[96] P.L. Hammer and S. Rudeanu, *Boolean methods in operations research and related areas*, (Springer, Berlin, 1968).

[97] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities*, (Cambridge Univ. Press, Cambridge, 1952).

[98] J.C. Harsanyi, A simplified bargaining model for the n-person cooperative game, *International Economic Review* **4** (1963) 194–220.

[99] S. Hart and A. Mas-Colell, The potential of the Shapley value. In A.E. Roth, editor, *The Shapley Value. Essays in Honor of Lloyd S. Shapley*, pages 127–137. (Cambridge University Press, 1988).

[100] S. Hart and A. Mas-Colell, Potential, value and consistency, *Econometrica* **57**(3) (1989) 589–614.

[101] J.B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization-algorithms I*, (Springer-Verlag, Berlin-Heidelberg-New York, 1991).

[102] J. Horváth, Sur le rapport entre les systèmes de postulats caractérisant les valeurs moyennes quasi arithmétiques symétriques, *C. R. Acad. Sci. Paris*, **225** (1947) 1256–1257.

[103] J. Kacprzyk and R.R. Yager (eds.), *The ordered weighted averaging operators: theory and applications* (Kluwer Academic Publishers: Norwell, MA, 1997).

[104] A. Kandel and W.J. Byatt, Fuzzy sets, fuzzy algebra, and fuzzy statistics, *Proc. IEEE* **66** (1978) 1619–1639.

[105] R.L. Keeney and H. Raiffa, *Decision with multiple objectives.* (Wiley, New York, 1976.)

[106] S.-R. Kim, On the possible scientific laws, *Mathematical Social Sciences* **20** (1990) 19–36.

[107] A.N. Kolmogoroff, Sur la notion de la moyenne, *Atti delle Reale Accademia Nazionale dei Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez.* (6) **12** (1930) 388–391.

[108] D.H. Krantz, R.D. Luce, P. Suppes and A. Tversky, *Foundations of measurement*, volume I: *Additive and polynomial representations*, (Academic Press, San Diego, 1971).

[109] E. Lehrer, An axiomatization of the Banzhaf value, *Int. J. of Game Theory* **17** (1988) 89–99.

[110] C.H. Ling, Representation of associative functions, *Publ. Math. Debrecen*, **12** (1965) 189–212.

[111] L. Lovász, Submodular function and convexity. In: *Mathematical programming. The state of the art.* Bonn 1982, Eds. A. Bachem, M. Grötschel, B. Korte. (Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1983), 235–257.

[112] R.D. Luce, On the possible psychophysical laws, *Psych. Rev.* **66** (1959) 81–95.

[113] R.D. Luce, D.H. Krantz, P. Suppes and A. Tversky, *Foundations of measurement*, volume III, (Academic Press, New York, 1990).

[114] R.D. Luce and H. Raiffa, Games and decisions: introduction and critical survey, (John Wiley & Sons, New York, 1967).

[115] J.-L. Marichal, On Sugeno integral as an aggregation function, *Fuzzy Sets and Systems*, to appear.

[116] J.-L. Marichal, On the associativity functional equation, *Fuzzy Sets and Systems*, to appear.

[117] J.-L. Marichal, On an axiomatization of the quasi-arithmetic mean values without the symmetry axiom, *Aequationes Mathematicae*, submitted.

[118] J.-L. Marichal and P. Mathonet, On comparison meaningfulness of aggregation functions, *Journal of Mathematical Psychology*, submitted.

[119] J.-L. Marichal and P. Mathonet, A characterization of the ordered weighted averaging functions based on the ordered bisymmetry property, *IEEE Trans. Fuzzy Syst.*, in press.

[120] J.-L. Marichal, P. Mathonet and E. Tousset, Characterization of some aggregation functions stable for positive linear transformations, *Fuzzy Sets and Systems*, in press.

[121] J.-L. Marichal and M. Roubens, Characterization of some stable aggregation functions, in: *Proc. Intern. Conf. on Industrial Engineering and Production Management*, Mons, Belgium (1993) 187–196.

[122] J.-L. Marichal and M. Roubens, The chaining interaction index among players in cooperative games, in: N. Meskens and M. Roubens (eds.), *Advances in Decision Theory*, Kluwer, 1998, in press.

[123] P. Miranda and M. Grabisch, Optimization issues for fuzzy measures, *8th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'98)*, Paris, 1998.

[124] F. Modave, D. Dubois, M. Grabisch and H. Prade, A Choquet integral representation in multicriteria decision making, *AAAI Symposium*, Boston, November 7–9, 1997.

[125] F. Modave and M. Grabisch, Preference representation by the Choquet integral: the commensurability hypothesis, *8th Int. Conf. on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'98)*, Paris, 1998.

[126] F.J. Montero, A note on Fung-Fu's theorem, *Fuzzy Sets and Systems* **17** (1985) 259–269.

[127] F.J. Montero, Aggregation of fuzzy opinions in a nonhomogeneous group, *Fuzzy Sets and Systems* **25** (1988) 15–20.

[128] P.S. Mostert and A.L. Shields, On the structure of semigroups on a compact manifold with boundary, *Ann. of Math.* **65** (1957) 117–143.

[129] T. Murofushi, A technique for reading fuzzy measures (I): the Shapley value with respect to a fuzzy measure, in: *2nd Fuzzy Workshop*, pp. 39–48, Nagaoka, Japan, October 1992. In Japanese.

[130] T. Murofushi and S. Soneda, Techniques for reading fuzzy measures (III): interaction index, in: *9th Fuzzy System Symposium*, pp. 693–696, Sapporo, Japan, May 1993. In Japanese.

[131] T. Murofushi and M. Sugeno, An interpretation of fuzzy measure and the Choquet integral as an integral with respect to a fuzzy measure, *Fuzzy Sets and Systems* **29** (1989) 201–227.

[132] T. Murofushi and M. Sugeno, A theory of fuzzy measures. Representation, the Choquet integral and null sets, *Journal of Mathematical Analysis and Applications* 159/2 (1991) 532–549.

[133] T. Murofushi and M. Sugeno, Fuzzy t-conorm integrals with respect to fuzzy measures: generalization of Sugeno integral and Choquet integral, *Fuzzy Sets and Systems* **42** (1991) 57–71.

[134] T. Murofushi and M. Sugeno, Non-additivity of fuzzy measures representing preferential dependence. In *2nd Int. Conf. on Fuzzy Systems and Neural Networks*, Iizuka, Japan, July 1992, pp. 617–620.

[135] T. Murofushi and M. Sugeno, Some quantities represented by the Choquet integral, *Fuzzy Sets & Systems* **56** (1993) 229–235.

[136] M. Nagumo, Über eine klasse der mittelwerte, *Japanese Journal of Mathematics*, **7** (1930) 71–79.

[137] A.S. Nowak, On the axiomatization of the Banzhaf value without the additivity axiom, *Int. J. of Game Theory* **26** (1997) 137–141.

[138] M. O'Hagan, Aggregating template rule antecedents in real-time expert systems with fuzzy set logic, *Proc. 22nd Ann. IEEE Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA (1988) 681–689.

[139] M. O'Hagan, Using maximum entropy-ordered weighted averaging to construct a fuzzy neuron, *Proceedings 24th Annual IEEE Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA (1990) 618–623.

[140] A. Orlov, The connection between mean quantities and admissible transformations, *Mathematical Notes* **30** (1981) 774–778.

[141] S. Ovchinnikov, Means on ordered sets, *Mathematical Social Sciences* **32** (1996) 39–56.

[142] S. Ovchinnikov, An analytic characterization of some aggregation operators, *International Journal of Intelligent Systems* **13** (1998) 3-10.

[143] S. Ovchinnikov, Invariant functions on simple orders, *ORDER*, to appear.

[144] G. Owen, Multilinear extensions of games, *Management Sciences* **18** (1972) 64–79.

[145] G. Owen, Multilinear extensions of games. In A.E. Roth, editor, *The Shapley Value. Essays in Honor of Lloyd S. Shapley*, pages 139–151. (Cambridge University Press, 1988).

[146] E. Pap, Null-additive set functions, (Kluwer Academic, Dordrecht, 1995).

[147] J. Quiggin, A theory of anticipated utility, *Journal of Economic Behavior and Organization* **3** (1982) 323–343.

[148] A.L. Ralescu and D.A. Ralescu, Extensions of fuzzy aggregation, *Fuzzy Sets and Systems* **86** (1997) 321–330.

[149] D.A. Ralescu and M. Sugeno, Fuzzy integral representation, *Fuzzy Sets and Systems* **84** (1996) 127–133.

[150] U. Ricci, Confronti tra medie, *Giornale degli Economisti e Rivista di Statistica* **26** (1915) 38–66.

[151] F.S. Roberts, *Measurement theory with applications to decision-making, utility and the social sciences*, (Addison-Wesley Pub., Reading, MA, 1979).

[152] F.S. Roberts, Merging relative scores, *J. Math. Anal. & Appl.* **147** (1990) 30–52.

[153] F.S. Roberts, Limitations on conclusions using scales of measurement, in A. Barnett, S.M. Pollock, and M.H. Rothkopf (eds.), *Operations Research and the Public Sector*, (Elsevier, Amsterdam, 1994), pp. 621–671.

[154] G.C. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **2** (1964) 340–368.

[155] J.J. Rotman, *An introduction to the theory of groups*, (Springer-Verlag New York, fourth edition, 1995.)

[156] M. Roubens, Interaction between criteria and definition of weights in MCDA problems. In *44th Meeting of the European Working Group "Multicriteria Aid for Decisions"*, Brussels, Belgium, October 1996.

[157] B. Roy, Classement et choix en présence de points de vue multiples (la méthode ELECTRE), *Revue Française d'Informatique et de Recherche Opérationnelle* **8** (1968) 57–75.

[158] B. Roy, How outranking relations helps multiple criteria decision making, in: J.L. Cochrane and M. Zeleny (Eds.), *Multiple Criteria Decision Making* (University of South California Press, Columbia, 1973) 179–201.

[159] T. Saaty, Exploring the interface between hierarchies, multiple objectives and fuzzy sets, *Fuzzy Sets and Systems* **1** (1978) 25–30.

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

[161] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* **57** (1989) 571–587.

[162] B. Schweizer and A. Sklar, Associative functions and statistical triangle inequalities, *Publ. Math. Debrecen* **8** (1961) 169–186.

[163] B. Schweizer and A. Sklar, Associative functions and abstract semigroups, *Publ. Math. Debrecen* **10** (1963) 69–81.

[164] B. Schweizer and A. Sklar, *Probabilistic metric spaces*, (North-Holland, Amsterdam, 1983).

[165] D. Scott and P. Suppes, Foundational aspects of theories of measurement, *J. Symbolic Logic* **23** (1958) 113–128.

[166] U. Segal, Order indifference and rank-dependent probabilities, *Journal of Math. Economics* **22** (1993) 373–397.

[167] G. Shafer, *A mathematical theory of evidence.* (Princeton University Press, 1976).

[168] C.E. Shannon and W. Weaver, A mathematical theory of communication, (University of Illinois Press, Urbana, 1949).

[169] L.S. Shapley, A value for n-person games. In: H.W. Kuhn and A.W. Tucker (eds.), *Contributions to the Theory of Games*, Vol. II, Annals of Mathematics Studies, 28, (Princeton University Press, Princeton, NJ, 1953), 307–317.

[170] L.S. Shapley, Core of convex games, *Int. J. Game Theory* **1** (1971) 11–26.

[171] L.S. Shapley and M. Shubik, A method for evaluating the distribution of power in a committee system, *The American Political Science Review* **48** (1954) 787–793.

[172] W. Silvert, Symmetric summation: a class of operations on fuzzy sets, *IEEE Trans. on Systems, Man and Cybernetics* **9** (1979) 657–659.

[173] I. Singer, Extensions of functions of 0-1 variables and applications to combinatorial optimization, *Numerical Functional Analysis and Optimization* 7(1) (1984-85) 23–62.

[174] H.J. Skala, Concerning ordered weighted averaging aggregation operators, *Statistical Papers* **32** (1991) 35–44.

[175] S.S. Stevens, Mathematics, measurement, and psychophysics, in: S.S. Stevens (ed.); *Handbook of Experimental Psychology*, (Wiley, New York, 1951) pp. 1–49.

[176] S.S. Stevens, Measurement, psychophysics, and utility, in: C.W. Churchman and P. Ratoosh (eds.), *Measurement: Definitions and Theories*, (Wiley, New York, 1959) pp. 18–63.

[177] M. Sugeno, Theory of fuzzy integrals and its applications, Ph.D. Thesis, Tokyo Institute of Technology, Tokyo, 1974.

[178] M. Sugeno, Fuzzy measures and fuzzy integrals: a survey, in: M.M. Gupta, G.N. Saridis and B.R. Gaines (Eds.), *Fuzzy Automata and Decision Processes*, (North-Holland, Amsterdam, 1977): 89–102.

[179] P. Suppes, D.H. Krantz, R.D. Luce and A. Tversky, *Foundations of measurement*, volume II, (Academic Press, San Diego, 1989).

[180] K. Tanaka and M. Sugeno, A study on subjective evaluations of color printing images, *Int. J. of Approximate Reasoning* **5** (1991) 213–222.

[181] E. Trillas, Sobre funciones de negación en la teoría de conjuntos difusos, *Stochastica* **3** (1979) 47–60.

[182] E. Trillas, S. Cubillo and J.L. Castro, Conjunction and disjunction on $([0,1],\leq)$, *Fuzzy Sets and Systems* **72** (1995) 155-165.

[183] A. Tversky and D. Kahneman, Advances in prospect theory: cumulative representation of uncertainty, *J. Risk Uncert.* **5** (1992) 297–323.

[184] B.L. van der Waerden, *Mathematical statistics*, (Springer-Verlag, Berlin, 1969).

[185] D. von Winterfeldt and W. Edwards, *Decision analysis and behavioral research*, (Cambridge University Press, 1986).

[186] P. Wakker, *Additive representations of preferences. A new foundation of decision analysis*, (Kluwer Academic Publishers, 1989).

[187] R.J. Weber, Probabilistic values for games. In A.E. Roth, editor, *The Shapley Value. Essays in Honor of Lloyd S. Shapley*, pages 101–119. (Cambridge University Press, 1988).

[188] S. Weber, A general concept of fuzzy connectives, negations and implications based on t-norms and t-conorms, *Fuzzy Sets and Systems* **11** (1983) 115–134.

[189] R.R. Yager, Entropy and specificity in a mathematical theory of evidence, *Int. J. Gen. Systems* **9** (1983) 249–260.

[190] R.R. Yager, On the aggregation of processing units in neural networks, *Proc. 1st IEEE Int. Conf. on Neural Networks*, San Diego, Vol. II, 1987, pp. 927–933.

[191] R.R. Yager, A note on weighted queries in information retrieval systems, *J. Amer. Soc. Information Sciences* **28** (1987) 23–24.

[192] R.R. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, *IEEE Trans. on Systems, Man and Cybernetics* **18** (1988) 183–190.

[193] R.R. Yager, Connectives and quantifiers in fuzzy sets, *Fuzzy Sets and Systems* **40** (1991) 39–75.

[194] R.R. Yager, OWA Neurons: A new class of fuzzy neurons, *Proc. Int. Joint on Neural Networks*, vol. 1, Baltimore, 1992, pp. 226–231.

[195] R.R. Yager, Fuzzy quotient operator, *Proc. Fourth Int. Conf. on Information Processing and Management of Uncertainty*, Palma de Majorca, 1992, pp. 317–322.

[196] R.R. Yager, MAM and MOM operators for aggregation, *Inform. Sci.* **69** (1993) 259–273.

[197] R.R. Yager, Aggregation operators and fuzzy systems modeling, *Fuzzy Sets and Systems* **67** (1994) 129–145.

[198] R.R. Yager and A. Rybalov, Noncommutative self-identity aggregation, *Fuzzy Sets and Systems* **85** (1997) 73–82.

[199] L.A. Zadeh, Fuzzy sets, *Inform. and Control* **8** (1965) 338–353.

[200] L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* **1** (1978) 3–28.

[201] H.-J. Zimmermann and P. Zysno, Latent connectives in human decision making, *Fuzzy Sets and Systems* **4** (1980) 37–51.

[202] H.-J. Zimmermann and P. Zysno, Decisions and evaluations by hierarchical aggregation of information, *Fuzzy Sets and Systems* 10 (1983) 243–260.