Fuzzy Measures and Fuzzy Integrals

Toshiaki Murofushi and Michio Sugeno

Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, 4259-G3-47 Nagatsuta, Midori-ku, 226-8502 Yokohama, Japan

1. Introduction

This article gives a survey of the theory of fuzzy measures and fuzzy integrals. The measure is one of the most important concepts in mathematics and so is the integral with respect to the measure. They have many applications in engineering, and their main characteristic is the additivity. This characteristic is very effective and convenient, but often too inflexible or too rigid. As a solution to the rigidness problem the fuzzy measure was proposed [1]. It is an extension of the measure in the sense that the additivity of the measure is replaced with a weaker condition, the monotonicity. The non-additivity is the main characteristic of the fuzzy measure, so that it is also called a non-additive measure. 'Fuzzy integral' is a general term for integrals with respect to the fuzzy measure. There are many kinds of fuzzy integrals: the Choquet integral, the Šipoš integral, the Sugeno integral, the t-conorm integral, etc. We discuss mainly the Choquet integral among them.

We deal only with finite-valued fuzzy measures on finite sets and omit the discussion about infinite-valued fuzzy measures and fuzzy measures on infinite sets. One reason is that so far almost all practical applications have used only finite-valued fuzzy measures on finite sets. The other reason is that the theory of such fuzzy measures is much easier than the general one. Fuzzy measures and the Choquet integral on infinite sets are presented in detail in the monograph of Denneberg [2] (see also his article [3] in this volume).

This article is organized as follows. Section 2 discusses finite-valued measures and the ordinary integral on finite sets. From the mathematical point of view, measures on finite sets are meaningless; although such a measure is a set function, it is equivalent to a point function. Such measures are, however, important for the comparison with fuzzy measures. Infinite-valued measures on infinite sets are discussed in Appendix. Section 3 discusses fuzzy measures and the Choquet integral on finite sets; basic properties and examples are shown. Section 4 introduces various types of fuzzy measures: λ-fuzzy measures, possibility measures, and decomposable measures. Section 5 shows three types of fuzzy integrals: Šipoš integral, Sugeno integral, and t-conorm integral. Other kinds of fuzzy integral will be discussed in the articles by Benvenuti and Mesiar [4] and by Imaoka [5] in this volume.

We denote the set of real numbers by $\mathbf{R}$ and the set of non-negative real numbers by $\mathbf{R}_+$. All functions we deal with are real-valued. Throughout the article except the appendix, X is assumed to be a finite set. Its power set, which is the family of all subsets of X, is denoted by 2^X. We use the term 'family' for a set of sets.

2. Measures and Integral

2.1 Set functions

Definition 2.1. *A function ξ defined on a family of sets is called a* set function*. Let ξ be a set function defined on 2^X.*

(i) *The set function ξ is said to be* additive *if for every pair of disjoint subsets A and B of X*

$$\xi(A \cup B) \ = \ \xi(A) + \xi(B) \ .$$

(ii) *The set function ξ is said to be* monotone *if for every pair of subsets A and B of X such that $A \subset B$*

$$\xi(A) \ \leq \ \xi(B) \ .$$

(iii) *The set function ξ is said to be* normalized *if*

$$\min\{\xi(A) \mid A \subset X\} \ = \ 0 \quad \textit{and} \quad \max\{\xi(A) \mid A \subset X\} \ = \ 1 \ .$$

If ξ is additive, then $\xi(\emptyset) = 0$ since $\xi(\emptyset) = \xi(\emptyset) + \xi(\emptyset)$. A non-negative additive set function is monotone; if ξ is non-negative and additive, and if $A \subset B \subset X$, then $\xi(B) = \xi(A \cup (B \setminus A)) = \xi(A) + \xi(B \setminus A) \geq \xi(A)$, where $B \setminus A = \{x \mid x \in B, \ x \notin A\}$, since $\xi(B \setminus A) \geq 0$. Since X is a finite set, an additive set function ξ defined on 2^X can be represented as

$$\xi(A) \ = \ \sum_{x \in A} \xi(\{x\}) \qquad \text{for } A \subset X \ .$$

Definition 2.2. *Let ξ be a set function defined on 2^X and A a subset of X. The* restriction *ξ_A of ξ to A is defined as*

$$\xi_A(B) \ = \ \xi(A \cap B) \qquad \textit{for all } B \subset X \ .$$

A restriction ξ_A of ξ has the same properties as ξ; if ξ is additive (or monotone or non-negative) then so is ξ_A.

Definition 2.3. *For a set function ξ defined on 2^X such that $\xi(\emptyset) = 0$, its* conjugate *set function $\overline{\xi}$ is defined as*

$$\overline{\xi}(A) \ = \ \xi(X) - \xi(A^c) \qquad \textit{for all } A \subset X \ ,$$

where A^c is the complement of A.

By definition, $\overline{\xi}(\emptyset) = 0$. If $\xi(\emptyset) = 0$, then it follows that $\overline{\xi}(X) = \xi(X)$ and hence that $\overline{\overline{\xi}} = \xi$. If ξ is additive, then it is self-conjugate, i.e., $\overline{\xi} = \xi$. If ξ is monotone, then so is $\overline{\xi}$. When ξ is normalized and monotone, its conjugate $\overline{\xi}$ is also called the *dual* of ξ.

2.2 Measures

Definition 2.4. *A* measure *on X is a non-negative additive set function defined on 2^X. A normalized measure is called a* probability measure*. A* signed measure *on X is an additive set function defined on 2^X.*

A probability measure is a measure, and a measure is a signed measure (Fig. 2.1). A set function P is a probability measure iff (if and only if) it is a measure such that $P(X) = 1$.

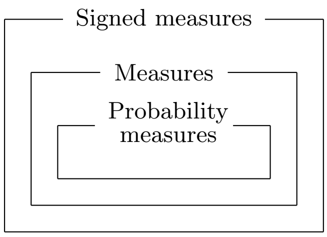


Fig.2.1. Families of probability measures, measures, and signed measures on X

A measure measures the size of sets. The number of elements in a set is a kind of measure of the size of sets.

Example 2.1. Let X be a finite set. The set function m_c defined as

$$m_c(A) \;=\; |A| \;,$$

where $|A|$ is the number of elements of A, is a measure on X, which is called the *counting measure* on X.

The volume and mass also can be regarded as sizes of sets. Although a size is usually non-negative, we can imagine a size taking a negative value. The quantity of electricity can be regarded as such a size.

Example 2.2. Let X be a finite set of objects (solid bodies).

(i) Let the volume of each object x be v_x cm^3. Then the set function $V : 2^X \to \mathbf{R}_+$ which measures the volume of each subset A of X,

$$V(A) \;=\; \sum_{x \in A} v_x \;,$$

is a measure on X.

(ii) Let the mass of each object x be m_x g. Then the set function $M : 2^X \to \mathbf{R}_+$ which measures the mass of each subset A of X,

$$M(A) \;=\; \sum_{x \in A} m_x \;,$$

is a measure on X.

(iii) Let each object x be electrified with q_x coulombs. Then the set function $Q : 2^X \to \mathbf{R}$ which measures the quantity of electricity of each subset A of X,

$$Q(A) \;=\; \sum_{x \in A} q_x \;,$$

is a signed measure on X.

The probability can be regarded as a size of sets.

Example 2.3. Consider the situation where one tosses a die and observes the number on the top face. Let $X = \{1, 2, 3, 4, 5, 6\}$, the set of possible outcomes. Then the set function $P : 2^X \to \mathbf{R}_+$ which measures the probability of each subset of X is a probability measure. If the die is unbiased, or fair, then $P(\{x\}) = 1/6$ for every $x \in X$.

The following is a special measure.

Example 2.4. Let x_0 be an element of X. The set function δ_{x_0} defined as

$$\delta_{x_0}(A) \;=\; \begin{cases} 1 & \text{if } x_0 \in A \,, \\ 0 & \text{if } x_0 \notin A \end{cases}$$

is a measure on X, which is called the *Dirac measure* on X focused on x_0.

Definition 2.5. *Let m be a signed measure on X. A subset N of X is called an m-*null *(or simply* null*) set if $m(M) = 0$ whenever $M \subset N$.*

Example 2.5. (Continued from Example 2.2 (iii)). A subset A of X is a Q-null set iff all the elements of A are not electrified at all.

Example 2.6. (Continued from Example 2.4.) Let $x_0 \in X$. A subset A of X is a δ_{x_0}-null set iff $x_0 \notin A$.

The following proposition shows properties of null sets.

Proposition 2.1. *Let m be a signed measure.*

(i) *The empty set is a null set.*
(ii) *A null set is of measure zero.*
(iii) *If m is non-negative, i.e., a measure, then a set of measure zero is a null set.*
(iv) *A subset of a null set is a null set.*
(v) *A union of null sets is a null set.*
(vi) *A set N is null*
$$\begin{aligned} &\iff \mu(A \cup M) = \mu(A) \quad \forall M \subset N,\, \forall A \subset X, \\ &\iff \mu(A \setminus M) = \mu(A) \quad \forall M \subset N,\, \forall A \subset X, \\ &\iff \mu(A \triangle M) = \mu(A) \quad \forall M \subset N,\, \forall A \subset X, \end{aligned}$$
where $A \triangle B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$.

Statement (vi) means that null sets are negligible.

Definition 2.6. *Let m be a signed measure on X, f a function on X, and F a non-null set. The* essential supremum *and* essential infimum *of f on F, denoted by* $\operatorname{ess\,sup}_{x\in F} f(x)$ *and* $\operatorname{ess\,inf}_{x\in F} f(x)$*, are defined as*

$$\operatorname*{ess\,sup}_{x\in F} f(x) \;=\; \min\{r \,|\, \text{the set } \{x \in F \,|\, f(x) > r\} \text{ is } m\text{-null}\,\}$$

and

$$\operatorname*{ess\,inf}_{x\in F} f(x) \;=\; \max\{r \,|\, \text{the set } \{x \in F \,|\, f(x) < r\} \text{ is } m\text{-null}\,\}\ ,$$

respectively.

The above definition is the one in the general case where X is not assumed to be a finite set. In the case of X being a finite set, the essential supremum and the essential infimum can be defined as

$$\begin{aligned}\operatorname*{ess\,sup}_{x\in F} f(x) &= \max\{f(x) \,|\, x \in F,\ \text{the singleton set } \{x\} \text{ is not null}\,\},\\ \operatorname*{ess\,inf}_{x\in F} f(x) &= \min\{f(x) \,|\, x \in F,\ \text{the singleton set } \{x\} \text{ is not null}\,\}\,.\end{aligned}$$

Then the essential supremum and the essential infimum may be called the essential maximum and the essential minimum, respectively.

Example 2.7. Let $X = \{x_1, x_2, x_3, x_4\}$, m be a signed measure on X such that $m(\{x_1\}) = 0$, $m(\{x_2\}) \neq 0$, $m(\{x_3\}) \neq 0$, $m(\{x_4\}) = 0$, and f a function on X such that $f(x_1) = 1$, $f(x_2) = 2$, $f(x_3) = 3$, $f(x_4) = 4$. Then, obviously x_1 gives the minimum of f, $\min_{x\in X} f(x) = 1$, and x_4 gives the maximum of f, $\max_{x\in X} f(x) = 4$. Since $\{x_1\}$ and $\{x_4\}$ are null sets, and since null sets are negligible, the minimum $f(x_1)$ and the maximum $f(x_4)$ are meaningless. On the other hand, since

$$\{x \in X \,|\, f(x) > r\} \;=\; \begin{cases} \emptyset & \text{if } 4 \le r\ ,\\ \{x_4\} & \text{if } 3 \le r < 4\ ,\\ \{x_3, x_4\} & \text{if } 2 \le r < 3\ ,\\ \{x_2, x_3, x_4\} & \text{if } 1 \le r < 2\ ,\\ \{x_1, x_2, x_3, x_4\} & \text{if } r < 1\ ,\end{cases}$$

$$\{x \in X \,|\, f(x) < r\} \;=\; \begin{cases} \emptyset & \text{if } r \le 1\ ,\\ \{x_1\} & \text{if } 1 < r \le 2\ ,\\ \{x_1, x_2\} & \text{if } 2 < r \le 3\ ,\\ \{x_1, x_2, x_3\} & \text{if } 3 < r \le 4\ ,\\ \{x_1, x_2, x_3, x_4\} & \text{if } 4 < r\ ,\end{cases}$$

and since $\{x_1\}$ and $\{x_4\}$ are null sets, it follows that

$$\operatorname*{ess\,sup}_{x\in X} f(x) \;=\; 3\ , \qquad \operatorname*{ess\,inf}_{x\in X} f(x) \;=\; 2\ .$$

Note that

$$\operatorname*{ess\,sup}_{x\in X} f(x) \;=\; \max_{x\in\{x_2,x_3\}} f(x)\ , \qquad \operatorname*{ess\,inf}_{x\in F} f(x) \;=\; \min_{x\in\{x_2,x_3\}} f(x)\ .$$

2.3 Integral

Definition 2.7. *Let m be a signed measure on X and f a function on X. The* integral $\int f(x)\,dm(x)$ *(or simply $\int f\,dm$) of f with respect to m is defined as*

$$\int f\,dm \;=\; \sum_{x\in X} f(x)\cdot m(\{x\}) \tag{2.1}$$

(Figs. 2.2 and 2.3). (Note that X is a finite set.)

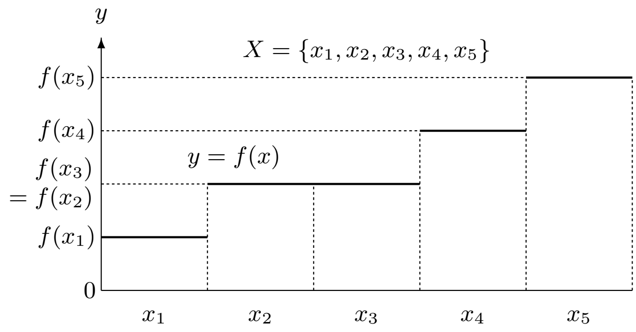


Fig. 2.2. Graph of f

Let $A \subset X$. The integral $\int_A f(x)\,dm(x)$ (or simply $\int_A f\,dm$) over A is defined as

$$\int_A f\,dm \;=\; \int f 1_A\,dm\,,$$

where 1_A is the indicator *(or* characteristic function*) of A;*

$$1_A(x) \;=\; \begin{cases} 1 & x\in A\,, \\ 0 & x\notin A\,. \end{cases}$$

Obviously

$$\int_X f\,dm \;=\; \int f\,dm$$

and

$$\int_A f\,dm \;=\; \sum_{x\in A} f(x)\cdot m(\{x\}) \;=\; \int f\,dm_A\,.$$

Example 2.8. (Continued from Example 2.1.) Let m_c be the counting measure on X. Then for every $A \subset X$ and every function f on X,

$$\int_A f\,dm_c \;=\; \sum_{x\in A} f(x)\,.$$

Example 2.9. (Continued from Example 2.2 (i), (ii)). Let the density of each $x \in X$ be $f(x)$ g/cm^3. The integral of f with respect to V is equal to the total mass of X. Moreover, for every $A \subset X$

$$M(A) = \int_A f\,dV\,.$$

The integral of f with respect to a probability measure P is called the *expectation* or *expected value* of f, and it is denoted as $\mathrm{E}(f;P)$, $\mathrm{E}_P(f)$, or simply $\mathrm{E}(f)$.

Example 2.10. (Continued from Example 2.3). Consider the situation where one tosses a die and, if the number on the top face is x, then the person gets $f(x)$ dollars. Then the expectation of the money the person will get is given by

$$\mathrm{E}(f;P) = \int f\,dP = \sum_{i=1}^{6} f(i)\cdot P(\{i\})\,.$$

Example 2.11. (Continued from Example 2.4.) Let $x_0 \in X$ and δ_{x_0} be the Dirac measure focused on x_0. Then for every function f on X,

$$\int f\,d\delta_{x_0} = f(x_0)\,.$$

The integral has the following properties.

$$\int f\,dm = \textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4} + \textcircled{5}$$

$\textcircled{1} = f(x_1)\cdot m(\{x_1\})$ $\quad$ $\textcircled{4} = f(x_4)\cdot m(\{x_4\})$

$\textcircled{2} = f(x_2)\cdot m(\{x_2\})$ $\quad$ $\textcircled{5} = f(x_5)\cdot m(\{x_5\})$

$\textcircled{3} = f(x_3)\cdot m(\{x_3\})$

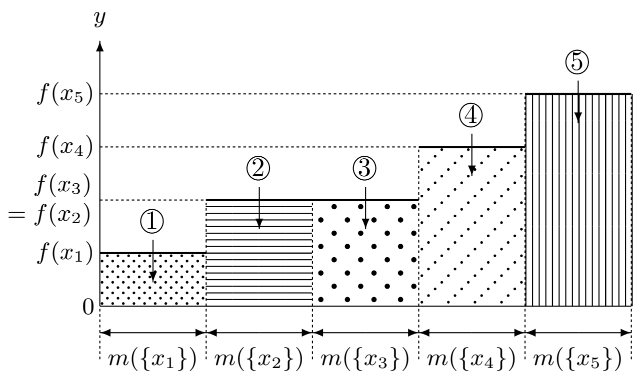


Fig. 2.3. The integral of f

Proposition 2.2. *Let m be a signed measure on X. Let a, b, a_1, a_1, $\ldots$, a_n be real numbers, f and g functions on X, and A, A_1, A_2, $\ldots$, A_n subsets of X.*

(i)
$$\int (af + bg)\, dm \;=\; a \int f\, dm + b \int g\, dm\,.$$

(ii)
$$\int \sum_{i=1}^{n} a_i 1_{A_i}\, dm \;=\; \sum_{i=1}^{n} a_i \cdot m(A_i)\;;$$

especially
$$\int 1_A\, dm \;=\; m(A)\,.$$

(iii) *If m is a measure and $f \leq g$, then*
$$\int f\, dm \;\leq\; \int g\, dm\,.$$

(iv) *If N is a null set, and if $f(x) = g(x)$ for all $x \notin N$, then*
$$\int f\, dm \;=\; \int g\, dm\,.$$

Concerning (ii) of the above proposition, note that every function f on X can be represented as
$$f \;=\; \sum_{i=1}^{n} a_i 1_{A_i}\;;$$
for instance,
$$f \;=\; \sum_{x \in X} f(x) 1_{\{x\}} \tag{2.2}$$
$$=\; \sum_{i=1}^{n} a_i 1_{\{x|f(x)=a_i\}} \tag{2.3}$$
$$=\; \sum_{i=1}^{n} (a_i - a_{i-1})\, 1_{\{x|f(x)\geq a_i\}}\;, \tag{2.4}$$
where $\{a_1, a_2, \ldots, a_n\}$ is the range $\{f(x)\,|\,x \in X\}$ of f, $a_1 \leq a_2 \leq \cdots \leq a_n$, and $a_0 = 0$.

For a function f represented as (2.2), by Proposition 2.2 (ii) its integral is given by (2.1) as in Fig. 2.3. For (2.3) the integral is given by
$$\int f\, dm \;=\; \sum_{i=1}^{n} a_i \cdot m(\{x|f(x) = a_i\}) \tag{2.5}$$

as in Fig. 2.4. For (2.4), the integral is given by

$$\int f\,dm \;=\; \sum_{i=1}^{n} (a_i - a_{i-1}) \cdot m(\{x|f(x) \geq a_i\}) \tag{2.6}$$

as in Fig. 2.5. The right hand sides of (2.1), (2.5), and (2.6) are the same value since the (signed) measure m is additive.

$$\int f\,dm \;=\; ① + ② + ③ + ④$$

$$① = a_1 \cdot m(\{x|f(x) = a_1\}) \qquad ③ = a_3 \cdot m(\{x|f(x) = a_3\})$$

$$② = a_2 \cdot m(\{x|f(x) = a_2\}) \qquad ④ = a_4 \cdot m(\{x|f(x) = a_4\})$$

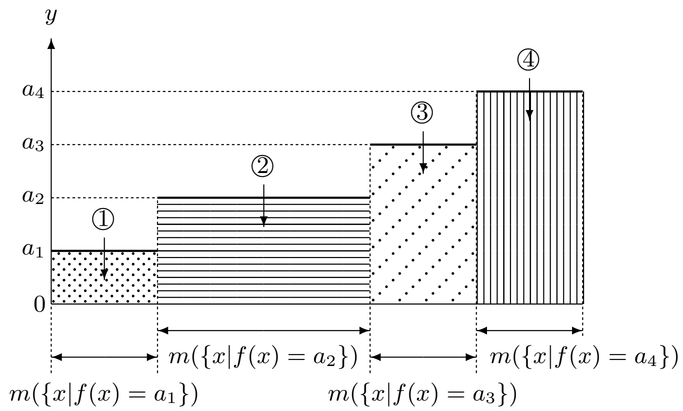


Fig. 2.4. The integral of f

3. Fuzzy Measures and the Choquet Integral

The contents of this section, except the last part, are based on [6, 7].

3.1 Fuzzy measures

Definition 3.1. *A (monotonic)* fuzzy measure *on X is a monotone set function defined on 2^X which vanishes at the empty set. A* non-monotonic *(or* signed*) fuzzy measure is a set function defined on 2^X which vanishes at the empty set.*

Obviously a fuzzy measure is a particular case of non-monotonic fuzzy measure. A fuzzy measure is non-negative since $\mu(A) \geq \mu(\emptyset) = 0$ for every

$$\int f\,d\mu \;=\; ① + ② + ③ + ④$$

$$
\begin{aligned}
① &= (a_1 - a_0) \cdot \mu(\{x|f(x) \geq a_1\}) \\
② &= (a_2 - a_1) \cdot \mu(\{x|f(x) \geq a_2\}) \\
③ &= (a_3 - a_2) \cdot \mu(\{x|f(x) \geq a_3\}) \\
④ &= (a_4 - a_3) \cdot \mu(\{x|f(x) \geq a_4\})
\end{aligned}
$$

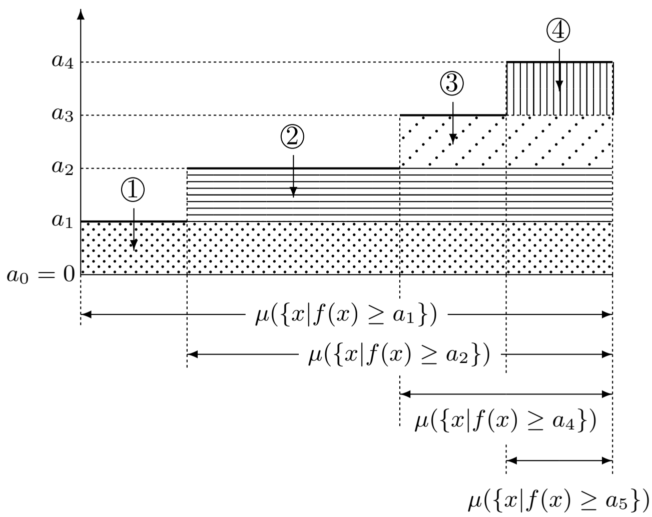


Fig. 2.5. The integral of f

$A \subset X$. An additive non-monotonic fuzzy measure is a signed measure, and an additive fuzzy measure is a measure since it is non-negative. A signed measure is a non-monotonic fuzzy measure since it vanishes at the empty set, and a measure is a fuzzy measure since it is monotone; therefore the (resp. non-monotonic) fuzzy measure is an extension of the (resp. signed) measure (Fig. 3.1).

Note that a fuzzy measure is not necessarily a measure. The difference between a fuzzy measure and a measure (or a non-monotonic fuzzy measure and a signed measure) is that the former is not necessarily additive. The main characteristic of a (non-monotonic) fuzzy measure is the non-additivity, so that a (non-monotonic) fuzzy measure is also called a non-additive measure.

We give concrete examples of monotonic and non-monotonic fuzzy measures.

Example 3.1. Let X be the set of all workers in a workshop, and suppose they produce the same products. For each $A \subset X$, we consider the situation that the members of group A work in the workshop. Each group may have various ways to work: various combinations of joint work and divided work.

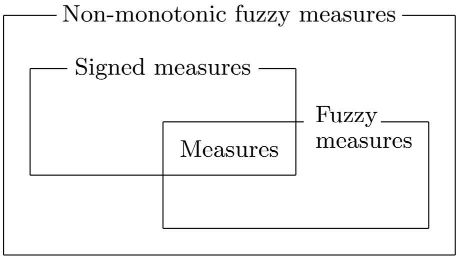


Fig. 3.1. Families of set functions on a finite set X

But suppose that every group works in the most efficient way. Let $\mu(A)$ be the number of the products made by group A in one hour. Then the set function $\mu : 2^X \to \mathbf{R}_+$ is monotone and vanishes at the empty set, and therefore it is a fuzzy measure.

The fuzzy measure μ is not necessarily additive. Let A and B be disjoint subsets of X, and consider the productivity of the coupled group $A \cup B$. If A and B work separately, then $\mu(A \cup B) = \mu(A) + \mu(B)$. But, since they generally interact on each other, the equality may not always hold. The effective cooperation of members of $A \cup B$ yields the inequality $\mu(A \cup B) > \mu(A) + \mu(B)$. On the other hand, the incompatibility between A's operation and B's, i.e., the impossibility of separate working, yields the opposite inequality $\mu(A \cup B) < \mu(A) + \mu(B)$. For example, the incompatibility is caused by limited space and/or insufficient equipments; sufficient space together with sufficient equipments makes separate working possible.

In the above example, the assumption "every group works in the most efficient way" brings the monotonicity of μ. Let A and B be disjoint subsets of X. If groups A and B are on bad terms with each other, and if they work together against their will, then their productivity may fall below that of either group; $\mu(A \cup B) < \mu(A)$ and/or $\mu(A \cup B) < \mu(B)$. The monotonicity assumption works to such a case. The most efficient way of working is to turn some troublemakers out of the workshop; in the worst case all members of one group are turned out. Then the monotonicity is recovered; $\mu(A \cup B) \geq \mu(A)$ and $\mu(A \cup B) \geq \mu(B)$. If the monotonicity assumption is removed, we can obtain a non-monotonic version:

Example 3.2. Instead of the monotonicity assumption we assume that, for each group A, all members of A must work together in the workshop. Let $\nu(A)$ be the number of the products made by A in one hour. Obviously the set function $\nu : 2^X \to \mathbf{R}_+$ is a non-monotonic fuzzy measure. As mentioned above, ν is not necessarily monotonic.

In the following example, we try to formalize the normalized fuzzy measures as a measure of certainty.

Example 3.3. Let x_0 be the right answer to a specific question, and X be the set of possible answers. Examples of questions and their answers are as follows:

(a) 'What the number on the top face will come out if the dice is thrown?'
x_0 : the number on the top face.
$X = \{1, 2, 3, 4, 5, 6\}$.

(b) 'Who is the criminal among the suspects a, b, c?'
x_0 : the criminal.
$X = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$.

Let us make assumptions on x_0 and X as follows:

Exhaustiveness Assumption:
The set X contains all the possible answers to the question.

Exclusiveness Assumption:
All the elements of X are mutually exclusive, or in other words, two distinct answers cannot become right at the same time.

As a consequence, there is one and only one right answer x_0 in X.

Now let us try to represent our knowledge or judgment about the certainty of the proposition '$x_0 \in E$' by giving a real number $P(E)$. We can express the degree of certainty of the proposition '$x_0 \in E$' by $P(E)$ instead of $P(x_0 \in E)$ because the logical combinations among propositions of the form '$x_0 \in E$' can be represented in the same form. Indeed,

$$x_0 \in E \text{ or } x_0 \in F \iff x_0 \in E \cup F \ ,$$

by Exhaustiveness Assumption

$$x_0 \in E \text{ and } x_0 \in F \iff x_0 \in E \cap F \ ,$$

and by Exhaustiveness and Exclusiveness Assumptions

$$\text{not } x_0 \in E \iff x_0 \in E^c \ .$$

We shall set the following three conditions concerning the numerical representation of certainty, which are independent of the above-mentioned two assumptions.

Monotonicity Condition:
If the proposition '$x_0 \in F$' is equally or more certain than the proposition '$x_0 \in E$', then $P(E) \leq P(F)$.

Lower Bound Condition:
If certainly $x_0 \notin E$, then $P(E) = 0$.

Upper Bound Condition:
If certainly $x_0 \in E$, then $P(E) = 1$.

Since '$x_0 \in E$' implies '$x_0 \in F$' whenever $E \subset F$, it follows from Monotonicity Condition that

$$E \subset F \implies P(E) \leq P(F) . \tag{3.1}$$

Since $x_0 \notin \emptyset$, it follows from Lower Bound Condition that

$$P(\emptyset) \; = \; 0 . \tag{3.2}$$

Since $x_0 \in X$ by Exhaustiveness Assumption, it follows from Upper Bound Condition that

$$P(X) \; = \; 1 . \tag{3.3}$$

From mathematical standpoint, we will adopt (3.1)–(3.3) as the axioms of the measure of certainty (or the normalized fuzzy measure).

If x_0 is known, our knowledge must be expressed by the Dirac measure δ_{x_0}. Obviously the probability measure satisfies (3.1)–(3.3) and can be regarded as a measure of certainty. In the next section, we introduce another measure of certainty named a possibility measure.

Now we discuss null set with respect to fuzzy measures.

Definition 3.2. *Let μ be a non-monotonic fuzzy measure on X. A set $N \subset X$ is called a μ-*null *set (or simply* null *set) if*

$$\mu(A \cup M) \; = \; \mu(A) \quad \forall M \subset N \; , \; \forall A \subset X \; .$$

Example 3.4. (Continued from Example 3.1). Assume that a worker x_0 can neither produce any products by oneself nor help any other workers. Then evidently the worker x_0 is suitable to be called incompetent or null. This situation can be expressed by

$$\mu(A \cup \{x_0\}) \; = \; \mu(A) \quad \forall A \subset X \; ,$$

and hence $\{x_0\}$ is a null set.

The following proposition shows properties of null sets. The statements (v) and (vi) show that the null set defined above is an extension of that in the ordinary measure theory.

Proposition 3.1. *Let μ be a non-monotonic fuzzy measure.*

(i) *The empty set is a null set.*
(ii) *A null set is of measure zero.*
(iii) *A set N is null* $\iff \mu(A \setminus M) \; = \; \mu(A) \quad \forall M \subset N, \forall A \subset X,$
$\iff \mu(A \triangle M) \; = \; \mu(A) \quad \forall M \subset N, \forall A \subset X.$
(iv) *If μ is monotone, i.e., a fuzzy measure, then a necessary and sufficient condition for $N \subset X$ to be a null set is that*

$$\mu(A \cup N) \; = \; \mu(A) \quad \forall A \subset X \; .$$

(v) *If μ is additive, then the necessary and sufficient condition for $N \subset X$ to be a null set is that $\mu(M) = 0$ whenever $M \subset N$.*

(vi) *If μ is additive and non-negative, then the necessary and sufficient condition for $N \subset X$ to be a null set is that $\mu(N) = 0$.*

(vii) *A subset of a null set is a null set.*

(viii) *A union of null sets is a null set.*

3.2 The Choquet integral

In this subsection we introduce the Choquet integral. It is an extension of the ordinary integral and the most natural fuzzy integral.

Since a fuzzy measure is generally non-additive, the right hand sides of (2.1), (2.5), and (2.6) are generally different from each other. The right hand side of (2.6) is the most appropriate to the integration with respect to (non-monotonic) fuzzy measures, and this is the Choquet integral.

Definition 3.3. *Let μ be a non-monotonic fuzzy measure on X and f a function on X with range $\{a_1, a_2, \ldots, a_n\}$ where $a_1 \le a_2 \le \cdots \le a_n$. The* Choquet integral $(\mathrm{C})\int f(x)\,d\mu(x)$ *(or simply* $(\mathrm{C})\int f\,d\mu$*) of f with respect to μ is defined as*

$$(\mathrm{C})\int f\,d\mu \;=\; \sum_{i=1}^{n} (a_i - a_{i-1}) \cdot \mu(\{x | f(x) \ge a_i\}) \,,$$

where $a_0 = 0$.

Fig. 2.3 shows the Choquet integral with $a_1 > 0$. If $a_1 < 0$, it is shown as in Fig. 3.2. From the arguments before the definition it follows that the Choquet integral with respect to a signed measure coincides with the ordinary integral, that is, the Choquet integral is an extension of the ordinary integral.

We give concrete examples of the Choquet integral.

Example 3.5. (Continued from Example 3.1). Let $X = \{x_1, x_2, \ldots, x_n\}$. One day each worker x_i works $f(x_i)$ hours from the opening hour. Without loss of generality, we can assume that $f(x_1) \le f(x_2) \le \cdots \le f(x_n)$. Then we have for $i \ge 2$,

$$f(x_i) - f(x_{i-1}) \;\ge\; 0$$

and

$$f(x_i) \;=\; f(x_1) + [f(x_2) - f(x_1)] + [f(x_3) - f(x_2)] + \cdots + [f(x_i) - f(x_{i-1})] \,.$$

Now let us aggregate the working hours of all the workers in the following way. First the group X with n workers works $f(x_1)$ hours, next the group $X \setminus \{x_1\} = \{x_2, x_3, \ldots, x_n\}$ works $f(x_2) - f(x_1)$ hours, then the group $X \setminus \{x_1, x_2\} = \{x_3, x_4, \ldots, x_n\}$ works $f(x_3) - f(x_2)$ hours, ..., lastly worker x_n works $f(x_n) - f(x_{n-1})$ hours. Therefore, since a group $A \subset X$ produces the

$$(\mathrm{C})\int f\,d\mu \;=\; \textcircled{1}+\textcircled{2}+\textcircled{3}+\textcircled{4}$$

$$\begin{aligned}
\textcircled{1} &= (a_1-a_0)\cdot\mu(\{x|f(x)\geq a_1\})\\
\textcircled{2} &= (a_2-a_1)\cdot\mu(\{x|f(x)\geq a_2\})\\
\textcircled{3} &= (a_3-a_2)\cdot\mu(\{x|f(x)\geq a_3\})\\
\textcircled{4} &= (a_4-a_3)\cdot\mu(\{x|f(x)\geq a_4\})
\end{aligned}$$

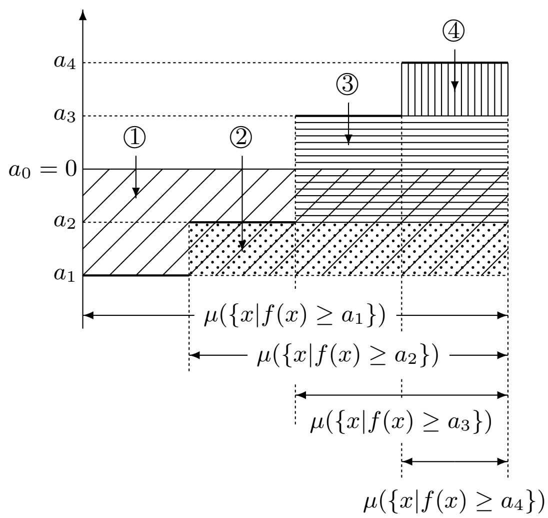


Fig. 3.2. The Choquet integral of f

amount $\mu(A)$ in one hour, the total number of the products produced by the workers is expressed by

$$\begin{aligned}
& f(x_1)\cdot\mu(X)\\
& +[f(x_2)-f(x_1)]\cdot\mu(X\setminus\{x_1\})\\
& +[f(x_3)-f(x_2)]\cdot\mu(X\setminus\{x_1,x_2\})\\
& +\cdots\\
& +[f(x_n)-f(x_{n-1})]\cdot\mu(\{x_n\})\\
=\; & \sum_{i=1}^{n}[f(x_i)-f(x_{i-1})]\cdot\mu(\{x_i,x_{i+1},\ldots,x_n\})\;,
\end{aligned}$$

where $f(x_0)=0$. This is nothing but the Choquet integral of f with respect to μ.

Example 3.6. (Continued from Example 3.2). In the same situation as in the previous example. The total number of the products produced by the workers is expressed by the Choquet integral $(\mathrm{C})\int f\,d\nu$.

Note 3.1. There are two different definitions of $(\mathrm{C})\int_A f\,d\mu$; one is

$$(\mathrm{C})\int_A f\,d\mu \;=\; (\mathrm{C})\int f 1_A\,d\mu \tag{3.4}$$

and the other is

$$(\mathrm{C})\int_A f\,d\mu \;=\; (\mathrm{C})\int f\,d\mu_A\ , \tag{3.5}$$

where μ_A is the restriction of μ to A (Definition 2.2). If f is non-negative, the right hand sides of (3.4) and (3.5) coincide with one another. If f takes a non-negative value, however, they are generally different.

The Choquet integral has the following properties.

Proposition 3.2. *Let f and g be functions on X and A a subset of X.*

(i)

$$(\mathrm{C})\int 1_A\,d\mu \;=\; \mu(A)\ .$$

(ii) *If μ is a fuzzy measure and $f \le g$, then*

$$(\mathrm{C})\int f\,d\mu \;\le\; (\mathrm{C})\int g\,d\mu\ .$$

(iii) *If a is a non-negative real number and b is a real number, then*

$$(\mathrm{C})\int (af+b)\,d\mu \;=\; a\cdot(\mathrm{C})\int f\,d\mu + b\cdot\mu(X)\ .$$

(iv)

$$(\mathrm{C})\int (-f)\,d\mu \;=\; -(\mathrm{C})\int f\,d\overline{\mu}\ .$$

(v)

$$(\mathrm{C})\int (-f)\,d\mu \;=\; -(\mathrm{C})\int f\,d\mu \qquad \textit{for all functions } f \textit{ on } X$$

iff $\overline{\mu} = \mu$.

(vi)

$$(\mathrm{C})\int f\,d\mu \;=\; (\mathrm{C})\int f^+\,d\mu - (\mathrm{C})\int f^-\,d\overline{\mu}\ ,$$

where $f^+(x) = \max\{f(x), 0\}$ *and* $f^-(x) = \max\{-f(x), 0\}$.

(vii) *If a is a real number, then*

$$(\mathrm{C})\int f\,d(a\cdot\mu) \;=\; a\cdot(\mathrm{C})\int f\,d\mu\ .$$

(viii) *If μ and ν are fuzzy measures on X such that $\mu \le \nu$ and $\mu(X) = \nu(X)$, then for all function f on X*

$$(\mathrm{C})\int f\,d\mu \;\le\; (\mathrm{C})\int f\,d\nu\;.$$

(ix) *If N is a null set, and if $f(x) = g(x)$ for all $x \notin N$, then*

$$(\mathrm{C})\int f\,d\mu \;=\; (\mathrm{C})\int g\,d\mu\;.$$

In the rest of this section, we show that the Choquet integral has a great expressive power; it can represent several important quantities [8]. First, we show that it can represent maximum, minimum, essential supremum, and essential infimum.

Definition 3.4. *A* 0-1 fuzzy measure *is a fuzzy measure whose range is* $\{0,1\}$.

Proposition 3.3. *Let μ be a 0-1 fuzzy measure. Then for every function f on X*

$$(\mathrm{C})\int f\,d\mu \;=\; \max_{A:\mu(A)=1}\,\min_{x\in A} f(x)\;.$$

Definition 3.5. *Let F be a non-empty subset of X. A set function* Pos_F *defined as*

$$\mathrm{Pos}_F(A) \;=\; \begin{cases} 1 & \text{if } A\cap F \neq \emptyset\;, \\ 0 & \text{if } A\cap F = \emptyset \end{cases}$$

is called the 0-1 possibility measure *focused on F. A set function* Nec_F *defined as*

$$\mathrm{Nec}_F(A) \;=\; \begin{cases} 1 & \text{if } F\subset A\;, \\ 0 & \text{if } F\not\subset A \end{cases}$$

is called the 0-1 necessity measure *focused on F.*

Obviously, 0-1 possibility measure and 0-1 necessity measure are 0-1 fuzzy measures, and Pos_F and Nec_F are conjugate, or dual. If $x_0 \in X$, then $\mathrm{Pos}_{\{x_0\}} = \mathrm{Nec}_{\{x_0\}} = \delta_{x_0}$, the Dirac measure focused on x_0. For every normalized fuzzy measure μ on X, it holds that $\mathrm{Nec}_X \le \mu \le \mathrm{Pos}_X$.

Proposition 3.4. *Let F be a non-empty subset of X and f a function on X.*

(i)

$$(\mathrm{C})\int f\,d\mathrm{Pos}_F \;=\; \max_{x\in F} f(x)\;.$$

(ii)

$$(\mathrm{C})\int f\,d\mathrm{Nec}_F \;=\; \min_{x\in F} f(x)\;.$$

(iii) *For every normalized fuzzy measure μ on X*

$$\min_{x\in X} f(x) \;\leq\; \text{(C)}\int f\, d\mu \;\leq\; \max_{x\in X} f(x)\ .$$

The Choquet integral as an expectation has good properties: Proposition 3.2 (iii) and Proposition 3.4 (iii).

Definition 3.6. *Let m be a signed measure and F a non-null subset of X. A set function* $\mathrm{ess\,Pos}_F$ *defined as*

$$\mathrm{ess\,Pos}_F(A) \;=\; \begin{cases} 1 & \textit{if } A\cap F \textit{ is not } m\textit{-null}\,, \\ 0 & \textit{if } A\cap F \textit{ is } m\textit{-null} \end{cases}$$

is called the 0-1 essential possibility measure *focused on F. A set function* $\mathrm{ess\,Nec}_F$ *defined as*

$$\mathrm{ess\,Nec}_F(A) \;=\; \begin{cases} 1 & \textit{if } F\setminus A \textit{ is } m\textit{-null}\,, \\ 0 & \textit{if } F\setminus A \textit{ is not } m\textit{-null} \end{cases}$$

is called the 0-1 essential necessity measure *focused on F.*

Proposition 3.5. *Let m be a signed measure, F a non-null subset of X, and f a function on X. Then*

$$\begin{aligned} \text{(C)}\int f\, d(\mathrm{ess\,Pos}_F) &\;=\; \operatorname*{ess\,sup}_{x\in F} f(x)\ , \\ \text{(C)}\int f\, d(\mathrm{ess\,Nec}_F) &\;=\; \operatorname*{ess\,inf}_{x\in F} f(x)\ . \end{aligned}$$

Since X is assumed to be a finite set, 0-1 essential possibility measures and 0-1 essential necessity measures are represented as 0-1 possibility measures and 0-1 necessity measures, respectively. Indeed, for every non-null set F, if we define

$$\mathrm{ess}\,F \;=\; \{x\in F \,|\, \text{the singleton set } \{x\} \text{ is not null}\,\}\ ,$$

then

$$\mathrm{ess\,Pos}_F \;=\; \mathrm{Pos}_{\mathrm{ess}\,F}\ , \qquad \mathrm{ess\,Nec}_F \;=\; \mathrm{Nec}_{\mathrm{ess}\,F}\ .$$

When X is a infinite set, 0-1 essential possibility (or 0-1 essential necessity) measures are not always represented as 0-1 possibility (or 0-1 necessity) measures.

The Choquet integral can represent the OWA operators, or the L-estimators. The definition of OWA operators is as follows.

Definition 3.7. *[9] Let* $w_1, w_2, \ldots, w_n$ *be non-negative numbers such that* $\sum_{i=1}^n w_i = 1$ *and* $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$*. The* ordered weighted averaging (OWA) *operator with weight* $\boldsymbol{w}$ *is a function* F *from* $[0,1]^n$ *into* $[0,1]$ *defined as*

$$F(r_1, r_2, \ldots, r_n) \; = \; \sum_{i=1}^{n} w_i \cdot r_{(i)} \; ,$$

where (i) *is a permutation satisfying* $r_{(1)} \leq r_{(2)} \leq \cdots \leq r_{(n)}$*.*

The OWA operator can be extended to a function from $\mathbf{R}^n$ into $\mathbf{R}$. In statistics such a function is called an *L-estimator*, which is a linear combination of order statistics. The L-estimator, or the extended OWA operator, is the arithmetic mean when $w_i = 1/n$ $(i = 1, 2, \ldots, n)$, it is the median when

$$w_i \; = \; \begin{cases} 1 & \text{if } n \text{ is odd and } i = (n+1)/2 \; , \\ 1/2 & \text{if } n \text{ is even and } i = n/2 \text{ or } (n/2)+1 \; , \\ 0 & \text{otherwise} \; , \end{cases}$$

it is the α-trimmed mean when

$$w_i \; = \; \begin{cases} 1/(n - 2[n\alpha]) & \text{if } [n\alpha] < i \leq n - [n\alpha] \; , \\ 0 & \text{otherwise} \; , \end{cases}$$

it is the α-Winsorized mean when

$$w_i \; = \; \begin{cases} 1/n & \text{if } [n\alpha] < i \leq n - [n\alpha] \; , \\ [n\alpha]/n & \text{if } i = [n\alpha] \text{ or } n - [n\alpha] + 1 \; , \\ 0 & \text{otherwise} \; , \end{cases}$$

and it is the k-th minimum, or the $(n-k+1)$-th maximum, when

$$w_i \; = \; \begin{cases} 1 & \text{if } i = k \; , \\ 0 & \text{otherwise} \; . \end{cases}$$

Note that [ ] is the Gauss symbol, i.e., $[r]$ stands for the greatest integer not exceeding r.

We can treat any $\boldsymbol{r} = (r_1, r_2, \ldots, r_n) \in \mathbf{R}^n$ as the function $f_{\boldsymbol{r}}$ on $X = \{1, 2, \ldots, n\}$ such that $f_{\boldsymbol{r}}(i) = r_i$ $(i = 1, 2, \ldots, n)$. Then the Choquet integral induces a function from $\mathbf{R}^n$ into $\mathbf{R}$ such that

$$\boldsymbol{r} = (r_1, r_2, \ldots, r_n) \;\; \mapsto \;\; \text{(C)}\int f_{\boldsymbol{r}} \, d\mu \; .$$

Proposition 3.6. *The L-estimator, or the extended OWA, with weight* $\boldsymbol{w}$ *is represented as the Choquet integral* $\text{(C)}\int f_{\boldsymbol{r}} \, d\mu$*, where*

$$\mu(A) \; = \; 1 - \sum_{i=1}^{n-|A|} w_i \; .$$

The Choquet integral with respect to a normalized fuzzy measure μ *induces an L-estimator, or an extended OWA, iff the value* $\mu(A)$ *depends only on* $|A|$*, i.e., there are* $\{\mu_i \,|\, i = 1, 2, \ldots, n\} \subset [0,1]$ *such that* $\mu(A) = \mu_i$ *whenever* $|A| = i$*; the weight* $\boldsymbol{w}$ *is given by* $w_i = \mu_{n-i+1} - \mu_{n-i}$*.*

The above proposition implies that the Choquet integral can represent the arithmetic mean, median, trimmed mean, Winsorized mean, and the k-th minimum.

4. Various Fuzzy Measures

As shown in the previous section, fuzzy measures (and the Choquet integral) have great powers of description. If $|X| = n$, then a measure has n parameters (or $n-1$ parameters when it is normalized) while a fuzzy measure has $2^n - 1$ parameters (or $2^n - 2$ parameters when it is normalized). This fact brings great powers of description to a fuzzy measure; however, it also brings a problem of complexity.

One solution to this problem is a constraint like the additivity of measures. Sugeno [1] introduced the λ-fuzzy measure g_λ, the normalized fuzzy measure with the λ-additivity; it has $n-1$ parameters. The possibility measure proposed by Zadeh [10], which is also a normalized fuzzy measure, has $n-1$ parameters. The decomposable measure proposed by Weber [11], which is an extension of the λ-fuzzy measure and the possibility measure, also has $n-1$ parameters. In this section we discuss these fuzzy measures.

Another solution to the problem is the k-additivity, which is discussed in the article by Grabisch [12] in this volume. The k-additivity is equivalent to the concept of inclusion-exclusion covering, which is discussed in the article by Fujimoto and Murofushi [13] in this volume.

4.1 λ-Fuzzy measures

Definition 4.1. *Let $\lambda \in (-1, \infty)$. A normalized set function g_λ defined on 2^X is called a* λ-fuzzy measure *on X if for every pair of disjoint subsets A and B of X*

$$g_\lambda(A \cup B) \;=\; g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B) \,.$$

Obviously, if $\lambda = 0$, then a λ-fuzzy measure is a normalized additive measure, i.e., a probability measure. A Dirac measure is a λ-fuzzy measure for all $\lambda > -1$.

A λ-fuzzy measure g_λ is a fuzzy measure. Since $\lambda > -1$ and g_λ is normalized, $1 + \lambda g_\lambda(A) > 0$ for all $A \subset X$. Therefore, since $g_\lambda(\emptyset) = g_\lambda(\emptyset) + g_\lambda(\emptyset) + \lambda g_\lambda(\emptyset) g_\lambda(\emptyset)$, it follows that $g_\lambda(\emptyset)[1 + \lambda g_\lambda(\emptyset)] = 0$, and hence that $g_\lambda(\emptyset) = 0$. On the other hand, if $A \subset B \subset X$, then

$$\begin{aligned} g_\lambda(B) &= g_\lambda(A) + g_\lambda(B \setminus A) + \lambda g_\lambda(A) g_\lambda(B \setminus A) \\ &= g_\lambda(A) + g_\lambda(B \setminus A)[1 + \lambda g_\lambda(A)] \\ &\geq g_\lambda(A) \end{aligned}$$

since $g_\lambda(B \setminus A) \geq 0$ and $1 + \lambda g_\lambda(A) > 0$.

For every $\lambda > -1$, define

$$\psi_\lambda(r) = \begin{cases} \log_{(1+\lambda)}(1+\lambda r) & \text{if } \lambda \neq 0 \,, \\ r & \text{if } \lambda = 0 \,. \end{cases} \tag{4.1}$$

Then for every λ-fuzzy measure g_λ on X, the set function $\psi_\lambda \circ g_\lambda$ is a probability measure on X. Indeed, $(\psi_\lambda \circ g_\lambda)(X) = \psi_\lambda(g_\lambda(X)) = \psi_\lambda(1) = 1$ and, if $\lambda \neq 0$, then for every pair of disjoint subsets A and B of X

$$\begin{aligned} (\psi_\lambda \circ g_\lambda)(A \cup B) &= \psi_\lambda(g_\lambda(A \cup B)) \\ &= \psi_\lambda[g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B)] \\ &= \log_{(1+\lambda)}\{1 + \lambda[g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B)]\} \\ &= \log_{(1+\lambda)}[(1+\lambda g_\lambda(A))(1+\lambda g_\lambda(B))] \\ &= \log_{(1+\lambda)}(1+\lambda g_\lambda(A)) + \log_{(1+\lambda)}(1+\lambda g_\lambda(B)) \\ &= (\psi_\lambda \circ g_\lambda)(A) + (\psi_\lambda \circ g_\lambda)(B) \,. \end{aligned}$$

Since there exists the inverse of ψ_λ,

$$\psi_\lambda^{-1}(r) = \begin{cases} \dfrac{1}{\lambda}[(1+\lambda)^r - 1] & \text{if } \lambda \neq 0 \,, \\ r & \text{if } \lambda = 0 \,, \end{cases}$$

for every pair of disjoint subsets A and B of X

$$g_\lambda(A \cup B) = \psi_\lambda^{-1}[\psi_\lambda(g_\lambda(A)) + \psi_\lambda(g_\lambda(B))] \,,$$

and hence for every finite sequence of mutually disjoint subsets A_1, A_2, $\ldots$, A_n of X,

$$g_\lambda\left(\bigcup_{i=1}^n A_i\right) = \psi_\lambda^{-1}\left[\sum_{i=1}^n \psi_\lambda(g_\lambda(A_i))\right] \,,$$

that is,

$$g_\lambda\left(\bigcup_{i=1}^n A_i\right) = \begin{cases} \dfrac{1}{\lambda}\left(\prod_{i=1}^n [1 + \lambda g_\lambda(A_i)] - 1\right) & \text{if } \lambda \neq 0 \,, \\ \sum_{i=1}^n g_\lambda(A_i) & \text{if } \lambda = 0 \,; \end{cases}$$

especially, since X is a finite set, for every subset A of X

$$g_\lambda(A) = \begin{cases} \dfrac{1}{\lambda}\left(\prod_{x \in A} [1 + \lambda g_\lambda(\{x\})] - 1\right) & \text{if } \lambda \neq 0 \,, \\ \sum_{x \in A} g_\lambda(\{x\}) & \text{if } \lambda = 0 \,. \end{cases}$$

Let P be a probability measure on X. Then the set function $\psi_\lambda^{-1} \circ P$ is a λ-fuzzy measure. For every function f on X, we define its λ-*expectation* $\mathrm{E}_\lambda(f;P)$ (or simply $\mathrm{E}_\lambda(f)$) as

$$\mathrm{E}_\lambda(f;P) \;=\; (\mathrm{C})\int f\, d(\psi_\lambda^{-1} \circ P)\ .$$

The λ-expectation has the following properties:

(i) If $\lambda \le \lambda'$, then

$$\mathrm{E}_\lambda(f;P) \;\ge\; \mathrm{E}_{\lambda'}(f;P)\ .$$

(ii)

$$\lim_{\lambda\to -1} \mathrm{E}_\lambda(f;P) \;=\; \operatorname*{ess\,sup}_{x\in X} f(x)\ .$$

(iii)

$$\mathrm{E}_0(f;P) \;=\; \mathrm{E}(f;P)\ ,$$

where the right hand side is the ordinary expectation of f.

(iv)

$$\lim_{\lambda\to \infty} \mathrm{E}_\lambda(f;P) \;=\; \operatorname*{ess\,inf}_{x\in X} f(x)\ .$$

Now consider a decision-making problem as follows. When a decision maker takes an alternative f and a state of nature $x \in X$ occurs, the monetary asset position of the decision maker is given by $f(x)$; that is, the alternative f is a function from the set X of states of nature into the set $\mathbf{R}$ of real numbers. Assume there is a probability measure P on X and the utility of the alternative f is given by the λ-expectation $\mathrm{E}_\lambda(f)$ with respect to a fixed value of λ; that is, $\mathrm{E}_\lambda(f) \ge \mathrm{E}_\lambda(g)$ iff f is indifferent or preferred to g.

In this decision-making problem, the value of λ is interpreted as a measure of risk aversion. The sign of λ indicates whether the decision maker is risk averse or risk prone, in other words, whether $\mathrm{E}(f)$ is preferred to f or not, where $\mathrm{E}(f)$ is regarded as the alternative by which the utility $\mathrm{E}(f)$ is certainly obtained regardless of states of nature, that is, $[\mathrm{E}(f)](x) = \mathrm{E}(f)$ for all $x \in X$. In contrast, f is an alternative the outcomes of which are generally affected by states of nature; f is an alternative with risk while $\mathrm{E}(f)$ is an alternative without risk.

First, consider the case of λ being negative. Since $\mathrm{E}_\lambda(f)$ is monotone non-increasing with respect to λ, it holds that $\mathrm{E}_\lambda(f) \ge \mathrm{E}_0(f)$. In addition, since $\mathrm{E}_0(f) = \mathrm{E}(f)$ and $\mathrm{E}(f) = \mathrm{E}_\lambda(\mathrm{E}(f))$, consequently it follows that $\mathrm{E}_\lambda(f) \ge \mathrm{E}_\lambda(\mathrm{E}(f))$, and therefore that f is indifferent or preferred to $\mathrm{E}(f)$. In other words, the decision maker is risk prone. As λ approaches -1, the difference between $\mathrm{E}_\lambda(f)$ and $\mathrm{E}(f)$ enlarges, that is, the risk proneness increases. In contrast to this, if λ is positive, then the decision maker is risk averse since $\mathrm{E}_\lambda(f) \le \mathrm{E}(f)$, and the risk aversion increases as λ increases. When $\lambda = 0$, the decision maker is risk neutral (Fig. 4.1).

The region of λ representing risk proneness is $(-1, 0)$ and that representing risk aversion is $(0, \infty)$. It therefore appears that the latter region is much

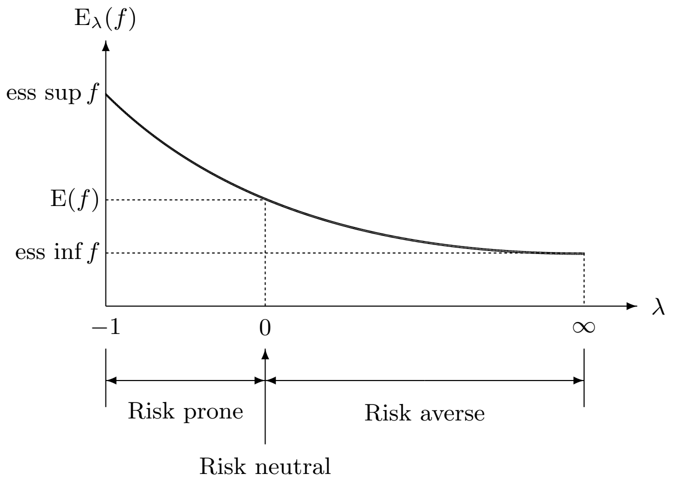


Fig. 4.1. Interpretation of parameter λ

wider and two regions are asymmetric. These two regions, however, can be regarded as symmetrical by the correspondence

$$\lambda \;\mapsto\; \overline{\lambda} \;=\; -\frac{\lambda}{1+\lambda}\ .$$

It is easily verified that the two λ-fuzzy measures $\psi_\lambda^{-1} \circ P$ and $\psi_{\overline{\lambda}}^{-1} \circ P$ are dual, i.e.,

$$(\psi_{\overline{\lambda}}^{-1} \circ P)(A) \;=\; 1 - (\psi_\lambda^{-1} \circ P)(A^c)\ ,$$

and therefore it follows from Proposition **3.2.** (iv) that

$$\mathrm{E}_\lambda(f) \;=\; -\mathrm{E}_{\overline{\lambda}}(-f)\ ;$$

this equality implies that λ and $\overline{\lambda}$ represent the mutually symmetrical attitudes toward risk.

4.2 Possibility measures and necessity measures

Definition 4.2. *A function* $\pi : X \to [0,1]$ *satisfying* $\max\{\pi(x) \,|\, x \in X\} = 1$ *is called a* possibility distribution. *A set function* Pos *is called a* possibility measure *on* X *if there exists a possibility distribution* π *on* X *such that for every* $A \subset X$

$$\mathrm{Pos}(A) \;=\; \max\{\pi(x) \,|\, x \in A\}\ .$$

A set function Nec *is called a* necessity measure *on* X *if there exists a possibility measure* Pos *on* X *such that for every* $A \subset X$

$$\mathrm{Nec}(A) \;=\; 1 - \mathrm{Pos}(A^c)\ .$$

The above equation is a mathematical expression of the statement that 'A is necessary' is equivalent to 'not-A is impossible'. A necessity measure Nec can be represented by the possibility distribution π of Pos as

$$\mathrm{Nec}(A) \;=\; 1 - \max\{\pi(x) \,|\, x \notin A\} \;.$$

For every $A \subset X$, $\mathrm{Pos}(A) = 1$ or $\mathrm{Pos}(A^c) = 1$, and $\mathrm{Nec}(A) = 0$ or $\mathrm{Nec}(A^c) = 0$. In addition, $\mathrm{Nec}(A) \leq \mathrm{Pos}(A)$ for all $A \subset X$.

The following example is due to Zadeh [10].

Example 4.1. Consider the question 'How many eggs did Hans eat for breakfast?' and let the set of possible answers be $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$. We may associate a possibility distribution π with the right answer x_0 by interpreting $\pi(x)$ as the degree of ease with which Hans can eat x eggs; for assessing the degree we may use some explicit or implicit criterion. We may associate a probability distribution p with x_0 by interpreting $p(x)$ as the probability of Hans eating x eggs for breakfast; the probability may be a subjective one or the relative frequency. They might be shown in Table 4.1.

Table 4.1. The possibility and probability distributions

x	1	2	3	4	5	6	7	8
$\pi(x)$	1	1	1	1	0.8	0.6	0.4	0.2
$p(x)$	0.1	0.8	0.1	0	0	0	0	0

In the example it holds that $p(x) \leq \pi(x)$ for all $x \in X$ and moreover that $\mathrm{Pr}(A) \leq \mathrm{Pos}(A)$ for all $A \subset X$, where $\mathrm{Pr}(A) = \sum_{x \in A} p(x)$ and $\mathrm{Pos}(A) = \max\{\pi(x) \,|\, x \in A\}$. This inequality indicates the *possibility/probability consistency principle*: a high degree of probability implies a high degree of possibility while the converse does not always hold, and a low degree of possibility implies a low degree of probability while the converse does not always hold. Note that, if $\mathrm{Pr}(A) \leq \mathrm{Pos}(A)$ for all $A \subset X$, then it holds that $\mathrm{Nec}(A) \leq \mathrm{Pr}(A) \leq \mathrm{Pos}(A)$ for all $A \subset X$.

A set function Pos is a possibility measure on X iff for every pair of subsets A and B of X

$$\mathrm{Pos}(A \cup B) \;=\; \mathrm{Pos}(A) \vee \mathrm{Pos}(B) \;;$$

a set function Nec is a necessity measure on X iff for every pair of subsets A and B of X

$$\mathrm{Nec}(A \cap B) \;=\; \mathrm{Nec}(A) \wedge \mathrm{Nec}(B) \;,$$

where $\vee$ and $\wedge$ mean max and min, respectively.

Obviously, 0-1 possibility measures and 0-1 necessity measures are possibility measures and necessity measures, respectively. The possibility distribution function of the 0-1 possibility (or 0-1 necessity) measure focused on F is the indicator 1_F of F.

4.3 t-Conorms and Decomposable measures

4.3.1 t-Conorms

Definition 4.3. *A binary operation on the unit interval* $[0,1]$ *is called a* t-conorm *if it satisfies the following conditions:*

(i) $r \perp 0 \;=\; 0 \perp r \;=\; r,$
(ii) $r \perp s \;=\; s \perp r,$
(iii) $(r \perp s) \perp t \;=\; s \perp (r \perp t),$
(iv) *if* $r \leq u$ *and* $s \leq v$*, then* $r \perp s \;\leq\; u \perp v.$

Example 4.2. The following are t-conorms, which are frequently used in the fuzzy set theory.

logical sum:

$$r \vee s \;=\; \max\{r, s\}\ .$$

bounded sum:

$$r \oplus s \;=\; \min\{r+s, 1\}\ .$$

algebraic sum:

$$r \dot{+} s \;=\; r + s - rs\ .$$

drastic sum:

$$r \,\dot{\vee}\, s \;=\; \begin{cases} 1 & \text{if } r > 0 \ \&\ s > 0\ , \\ r & \text{if } s = 0\ , \\ s & \text{if } r = 0\ . \end{cases}$$

For every t-conorm $\perp$ the following holds:

$$r \vee s \;\leq\; r \perp s \;\leq\; r \,\dot{\vee}\, s \qquad \forall r,\, s \in [0,1]\ .$$

For every $r \in [0,1]$

$$r \perp 1 \;=\; 1 \perp r \;=\; 1$$

since $1 \geq r \perp 1 = 1 \perp r \geq 1 \perp 0 = 1$.

A t-conorm $\perp$ is said to be *Archimedean* if for every pair of real numbers r and s for which $0 < r < s < 1$ there is a positive integer n such that

$$s \;<\; \mathop{\perp}_{i=1}^{n} r\ ,$$

where

$$\mathop{\perp}_{i=1}^{n} r_i \;=\; r_1 \perp r_2 \perp \cdots \perp r_n\ .$$

The bounded sum, the algebraic sum, and the drastic sum are Archimedean, and the logical sum is non-Archimedean. If a t-conorm $\perp$ is Archimedean, then $r < r \perp r \quad \forall r \in (0,1)$. If a t-conorm $\perp$ is continuous, and if $r < r \perp r$ $\forall r \in (0,1)$, then it is Archimedean.

A t-conorm $\perp$ is said to be *nilpotent* if for every $r \in (0,1)$ there is a positive integer n such that

$$\mathop{\perp}_{i=1}^{n} r \;=\; 1 \,.$$

The bounded sum and the drastic sum are nilpotent, and neither the logical sum nor the algebraic sum is nilpotent.

The following theorem is well-known:

Theorem 4.1. *A binary operation $\perp$ is a continuous Archimedean t-conorm iff there exists a continuous strictly increasing function $\psi : [0,1] \to [0,\infty]$ such that*

$$r \perp s \;=\; \psi^{(-1)}(\psi(r)+\psi(s)) \qquad \forall r,\, s \in [0,1] \,, \tag{4.2}$$

where $\psi^{(-1)}$ is the pseudo-inverse *of ψ defined as*

$$\psi^{(-1)}(r) \;=\; \begin{cases} \psi^{-1}(r) & \text{if } r \le \psi(1)\,, \\ 1 & \text{if } r > \psi(1)\,, \end{cases}$$

and for every real number r

$$r + \infty \;=\; \infty + r \;=\; \infty + \infty \;=\; \infty \,.$$

A function ψ satisfying (4.2) is called an (*additive*) *generator* of $\perp$. If ψ is a generator of $\perp$, then so is $a\psi$ for $a > 0$; furthermore, if φ is another generator of $\perp$, then there is a positive number a such that $\varphi = a\psi$. If $\perp$ has a generator ψ, then the necessary and sufficient condition for $\perp$ to be nilpotent is that $\psi(1) < \infty$. By the above theorem, if ψ is a generator of $\perp$, then it holds that

$$\mathop{\perp}_{i=1}^{n} r_i \;=\; \psi^{(-1)}\left(\sum_{i=1}^{n} \psi(r_i)\right) \,.$$

The logical sum has no generator since it is non-Archimedean, the bounded sum has a generator $\psi(r) = r \;\forall r \in [0,1]$, the algebraic sum has a generator $\psi(r) = -\log(1-r) \;\forall r \in [0,1]$, and the drastic sum has no generator since it is not continuous.

A *parameterized family* of t-conorms is a family of t-conorms generated by a parameterized generator. For example, the λ-sums,

$$r \oplus_\lambda s \;=\; \min\{r + s + \lambda rs,\, 1\} \,,$$

are the parameterized family with the parameterized generator ψ_λ defined as in (4.1), where λ is the parameter with range $(-1,\infty)$. The λ-sum $\oplus_\lambda$ becomes the bounded sum when $\lambda = 0$, the algebraic sum when $\lambda \to -1$, and the drastic sum when $\lambda \to \infty$.

4.3.2 Decomposable measures

Definition 4.4. *Let $\perp$ be a t-conorm. A set function m defined on 2^X is called a* $\perp$-decomposable measure *(or simply* decomposable measure*) on X if $m(\emptyset) = 0$, $m(X) = 1$, and for every pair of disjoint subsets A and B of X*

$$m(A \cup B) \;=\; m(A) \perp m(B) \;.$$

Obviously a λ-fuzzy measure is a $\oplus_\lambda$-decomposable measure, a probability measure is a $\oplus$-decomposable measure, and a possibility measure is a $\vee$-decomposable measure. A 0-1 possibility measure is a $\perp$-decomposable measure for any t-conorm $\perp$.

For every t-conorm $\perp$, a $\perp$-decomposable measure is a normalized fuzzy measure; the monotonicity follows from the non-decreasingness of $\perp$ (Definition 4.3 (iv)).

Let m be a $\perp$-decomposable measure on X. Since X is a finite set, for every subset A of X

$$m(A) \;=\; \mathop{\perp}_{x \in A} m(\{x\}) \;.$$

If $\perp$ is not nilpotent, then there is at least one element $x \in X$ such that $m(\{x\}) = 1$ since $\perp_{x \in X} m(\{x\}) = m(X) = 1$.

Let $\perp$ be a t-conorm with a generator ψ. A $\perp$-decomposable measure m such that $\psi \circ m$ is a measure, which may take ∞ as its value, is useful (see subsection 5.3). Note that a set function m is a λ-fuzzy measure iff it is a $\oplus_\lambda$-decomposable measure such that $\psi_\lambda \circ m$ is a probability measure.

5. Various Fuzzy integrals

This section discusses the Šipoš integral, the Sugeno integral, and t-conorm integral. They and the Choquet integral are the principal fuzzy integrals, and some other fuzzy integrals are modifications of them.

5.1 Šipoš integral

Definition 5.1. *Let μ be a non-monotonic fuzzy measure on X, and f a function on X with range $\{a_1, a_2, \ldots, a_m, b_1, b_2, \ldots, b_n\}$ where $b_n \le b_{n-1} \le \cdots \le b_1 \le 0 \le a_1 \le \cdots \le a_m$. The* Šipoš integral *(Š)$\int f(x)\, d\mu(x)$ (or simply (Š)$\int f\, d\mu$) of f with respect to μ is defined as*

$$\begin{aligned}(\check{S})\int f\, d\mu \;&=\; \sum_{i=1}^{m} (a_i - a_{i-1}) \cdot \mu(\{x | f(x) \ge a_i\}) \\ &\quad + \sum_{i=1}^{n} (b_i - b_{i-1}) \cdot \mu(\{x | f(x) \le b_i\}) \;,\end{aligned}$$

where $a_0 = b_0 = 0$ (Fig. 5.1).

$$(\check{\mathrm{S}})\int f\,d\mu \;=\; ① + ② + ③ + ④$$

$$\begin{aligned}
① &= (a_1 - a_0)\cdot\mu(\{x|f(x)\geq a_1\})\\
② &= (a_2 - a_1)\cdot\mu(\{x|f(x)\geq a_2\})\\
③ &= (b_1 - b_0)\cdot\mu(\{x|f(x)\leq b_1\})\\
④ &= (b_2 - b_1)\cdot\mu(\{x|f(x)\leq b_2\})
\end{aligned}$$

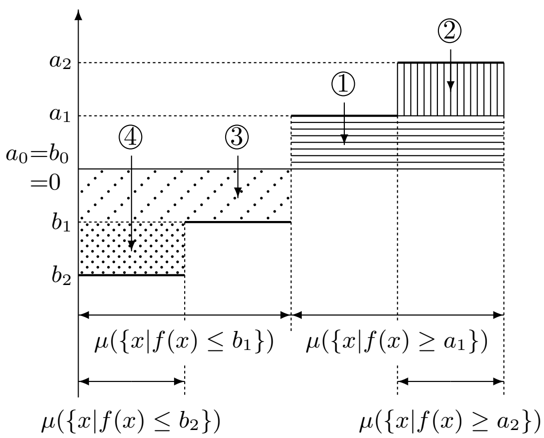


Fig. 5.1. The Šipoš integral of f

We can define $(\check{\mathrm{S}})\int_A f\,d\mu$ as

$$(\check{\mathrm{S}})\int_A f\,d\mu \;=\; (\check{\mathrm{S}})\int f1_A\,d\mu \;=\; (\check{\mathrm{S}})\int f\,d\mu_A$$

since the last two integrals are equal to one another for every function f on X.

The Šipoš integral has the following properties.

Proposition 5.1. *Let a be a non-negative real number, f and g functions on X, and A a subset of X.*

(i) *If f is non-negative, then*

$$(\check{\mathrm{S}})\int f\,d\mu \;=\; (\mathrm{C})\int f\,d\mu\;.$$

Especially

$$(\check{\mathrm{S}})\int 1_A\,d\mu \;=\; \mu(A)\;.$$

(ii) *If μ is a fuzzy measure and $f \leq g$, then*

$$(\check{\mathrm{S}})\int f\,d\mu \;\leq\; (\check{\mathrm{S}})\int g\,d\mu\,.$$

(iii)

$$(\check{\mathrm{S}})\int af\,d\mu \;=\; a\cdot(\check{\mathrm{S}})\int f\,d\mu\,.$$

Especially

$$(\check{\mathrm{S}})\int(-f)\,d\mu \;=\; -(\check{\mathrm{S}})\int f\,d\mu\,.$$

(iv)

$$(\check{\mathrm{S}})\int f\,d\mu \;=\; (\check{\mathrm{S}})\int f^+\,d\mu - (\check{\mathrm{S}})\int f^-\,d\mu\,,$$

where $f^+(x) = \max\{f(x), 0\}$ *and* $f^-(x) = \max\{-f(x), 0\}$.

(v)

$$(\check{\mathrm{S}})\int f\,d(a\cdot\mu) \;=\; a\cdot(\check{\mathrm{S}})\int f\,d\mu\,.$$

(vi) *If N is a null set, and if $f(x) = g(x)$ for all $x \notin N$, then*

$$(\check{\mathrm{S}})\int f\,d\mu \;=\; (\check{\mathrm{S}})\int g\,d\mu\,.$$

5.2 Sugeno integral

The Sugeno integral is defined only for functions whose range is included in $[0,1]$ and normalized fuzzy measures.

Definition 5.2. *Let μ be a normalized fuzzy measure on X and f a function on X with range $\{a_1, a_2, \ldots, a_n\}$ where $0 \leq a_1 \leq a_2 \leq \cdots \leq a_n \leq 1$. The* Sugeno integral $⨍ f(x)\circ\mu(x)$ *(or simply $⨍ f\circ\mu$) of f with respect to μ is defined as*

$$⨍ f\circ\mu \;=\; \bigvee_{i=1}^{n}\left[a_i \wedge \mu(\{x|f(x) \geq a_i\})\right]\,.$$

The Sugeno integral has the following properties.

Proposition 5.2. *Let μ and ν be normalized fuzzy measures, f and g functions from X into $[0,1]$, A a subset of X, and $a \in [0,1]$.*

(i)

$$⨍ 1_A \circ\mu \;=\; \mu(A)\,.$$

(ii) *If $f \leq g$, then*

$$⨍ f\circ\mu \;\leq\; ⨍ g\circ\mu\,.$$

(iii)
$$\fint (a \vee f) \circ \mu \;=\; a \vee \left(\fint f \circ \mu \right) .$$

(iv)
$$\left| \fint f \circ \mu \;-\; (\mathrm{C})\int f \, d\mu \right| \;\leq\; \frac{1}{4} .$$

(v) *If μ is a 0-1 fuzzy measure, then*
$$\fint f \circ \mu \;=\; (\mathrm{C})\int f \, d\mu .$$

(vi) *If $\mu \leq \nu$, then*
$$\fint f \circ \mu \;\leq\; \fint f \circ \nu .$$

(vii)
$$\fint f \circ (\mu \vee \nu) \;=\; \left(\fint f \circ \mu \right) \vee \left(\fint f \circ \nu \right) .$$

(viii) *If N is a null set, and if $f(x) = g(x)$ for all $x \notin N$, then*
$$\fint f \circ \mu \;=\; \fint g \circ \mu .$$

From statements (v) and (vi) and Proposition 3.4 it follows that for every normalized fuzzy measure μ on X and for every function f from X into $[0, 1]$
$$\min_{x \in X} f(x) \;\leq\; \fint f \circ \mu \;\leq\; \max_{x \in X} f(x) .$$

5.3 t-Conorm integral

Definition 5.3. *A* t-conorm system *is a quadruplet* $(\triangle, \perp, \mathrm{II}, \diamond)$ *consisting of continuous t-conorms* $\triangle$, $\perp$, II, *and an operation* $\diamond : [0, 1] \times [0, 1] \to [0, 1]$ *such that*

TS1: $\diamond$ *is continuous on* $(0, 1]^2$,
TS2: $a \diamond r = 0 \iff a = 0$ *or* $r = 0$,
TS3: if $r \perp s < 1$,
$$a \diamond (r \perp s) \;=\; (a \diamond r) \,\mathrm{II}\, (a \diamond s) ,$$
TS4: if $a \triangle b < 1$,
$$(a \triangle b) \diamond r \;=\; (a \diamond r) \,\mathrm{II}\, (b \diamond r) .$$

We denote generators of $\triangle$, $\perp$, and II, if exist, by φ, ψ, and ζ, respectively. Many t-conorm systems are trivial and/or useless; for example, $a \diamond r = \text{const.}$ for all a, $r \in (0, 1]$. The non-trivial t-conorm systems are following ones:

– Max type:
The t-conorms $\triangle$, $\perp$, and $\amalg$ are the max operator $\vee$.

– Archimedean type:
The t-conorms $\triangle$, $\perp$, and $\amalg$ have their generators φ, ψ, and ζ, respectively, and for all $a \in [0,1]$ and $r \in [0,1]$

$$a \diamond r \;=\; \zeta^{(-1)}(\varphi(a) \cdot \psi(r)) \;.$$

We call the triplet $(\varphi,\, \psi,\, \zeta)$ the *generator* of $(\triangle,\, \perp,\, \amalg,\, \diamond)$.

Definition 5.4. *A $\perp$-decomposable measure m is said to be* normal *iff* (i) *or* (ii) *holds:*

(i) $\perp = \vee$,
(ii) $\perp$ *has a generator ψ and $\psi \circ m$ is a measure.*

Obviously, λ-fuzzy measures and possibility measures are normal.

Definition 5.5. *Let $\mathcal{F} = (\triangle, \perp, \amalg, \diamond)$ be a t-conorm systems, m a normal $\perp$-decomposable measure on X, and f a function from X into* $[0,1]$. *The $\mathcal{F}$-*integral *(or* t-conorm integral*)* $(\mathcal{F}) \int f(x)\, dm(x)$ *(or simply* $(\mathcal{F}) \int f\, dm$*) of f with respect to m is defined as*

$$(\mathcal{F}) \int f\, dm \;=\; \coprod_{x \in X} f(x) \diamond m(\{x\}) \;. \tag{5.1}$$

If the t-conorm system $\mathcal{F} = (\triangle, \perp, \amalg, \diamond)$ is Archimedean, i.e., it has a generator (φ, ψ, ζ), then the t-conorm integral is expressed as

$$(\mathcal{F}) \int f\, dm \;=\; \zeta^{(-1)} \left(\int \varphi(f)\, d(\psi \circ m) \right) \;,$$

where the integral of the right hand side is the ordinary integral.

Definition 5.6. *Let $\triangle$ be a t-conorm. Define a binary operation $-_\triangle$ on* $[0,1]$ *as*

$$a -_\triangle b \;=\; \inf\{c \,|\, b \triangle c \geq a\} \;.$$

If $\triangle = \vee$, then

$$a -_\triangle b \;=\; \begin{cases} a & \text{if } a \geq b\,, \\ 0 & \text{if } a < b\,; \end{cases}$$

and, if $\triangle$ has a generator φ, then

$$a -_\triangle b \;=\; \varphi^{(-1)}(0 \vee [\varphi(a) - \varphi(b)]) \;.$$

Obviously, $a -_\triangle 0 = a$. If $\triangle$ is continuous, then $(a -_\triangle b) \triangle b = a$ whenever $a \geq b$.

Definition 5.7. *Let $\mathcal{F} = (\triangle, \perp, \amalg, \diamond)$ be a non-trivial t-conorm systems, μ a normalized fuzzy measure on X, and f a function on X with range $\{a_1, a_2, \ldots, a_n\}$ where $a_1 \leq a_2 \leq \cdots \leq a_n$. The $\mathcal{F}$-*fuzzy integral *(or* t-conorm fuzzy integral*) $(\mathcal{F}) \int f(x)\, d\mu(x)$ (or simply $(\mathcal{F}) \int f\, d\mu$) of f with respect to μ is defined as*

$$(\mathcal{F}) \int f\, d\mu \;=\; \coprod_{i=1}^{n} (a_i -_{\triangle} a_{i-1}) \diamond \mu(\{x | f(x) \geq a_i\})\;, \tag{5.2}$$

where $a_0 = 0$.

If μ is a normal $\perp$-decomposable measure, then the $(\triangle, \perp, \amalg, \diamond)$-fuzzy integral with respect to μ coincides with the $(\triangle, \perp, \amalg, \diamond)$-integral with respect to μ. The $(\vee, \vee, \vee, \wedge)$-fuzzy integral is the Sugeno integral. If $(\triangle, \perp, \amalg, \diamond)$ has a generator (φ, ψ, ζ), then $(\triangle, \perp, \amalg, \diamond)$-fuzzy integral can be represented as

$$(\mathcal{F}) \int f\, d\mu \;=\; \zeta^{(-1)} \left((\mathrm{C}) \int \varphi(f)\, d(\psi \circ m) \right)\;,$$

where the integral of the right hand side is the Choquet integral.

References

1. Sugeno, M. (1974): Theory of fuzzy integrals and its applications, Doctoral Thesis, Tokyo Institute of Technology
2. Denneberg, D. (1997): Non-additive measure and integral, 2nd ed., Kluwer Academic
3. Denneberg, D. (1999): Non-additive measure and integral, basic concepts and their role for applications, in: M. Grabisch, T. Murofushi, and M. Sugeno, Eds., Fuzzy Measures and Integrals: Theory and Applications, Springer-Verlag
4. Benvenuti, B. and Mesiar, R. (1999): Integrals with respect to general fuzzy measure, in: M. Grabisch, T. Murofushi, and M. Sugeno, Eds., Fuzzy Measures and Integrals: Theory and Applications, Springer-Verlag
5. Imaoka, H. (1999): Generalized fuzzy integral, in: M. Grabisch, T. Murofushi, and M. Sugeno, Eds., Fuzzy Measures and Integrals: Theory and Applications, Springer-Verlag
6. Murofushi, T., Sugeno, M. (1991): A theory of fuzzy measures: representations, the Choquet integral, and null sets, J. Math. Anal. Appl. **159**, 532–549
7. Murofushi, T., Sugeno, M. , Machida, M. (1994): Non-monotonic fuzzy measures and the Choquet integral, Fuzzy Sets and Systems **64**, 73–86
8. Murofushi, T., Sugeno, M. (1993): Some quantities represented by the Choquet integral, Fuzzy Sets and Systems, **56**, 229–235
9. Yager, R.R. (1986): On ordered weighted averaging aggregation operators in multicriteria decision making, IEEE trans. Systems, Man & Cybern. **18**, 183–190
10. Zadeh, L.A. (1978): Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems **1**, 3–28
11. Weber, S. (1984): $\perp$-Decomposable measures and integrals for Archimedean t-conorms $\perp$, J. Math. Anal. Appl. **101**, 114–138

12. Grabishc, M. (1999): The interaction and Möbius representations of fuzzy measures on finite spaces, k-additive measures: a survey, in: M. Grabisch, T. Murofushi, and M. Sugeno, Eds., Fuzzy Measures and Integrals: Theory and Applications, Springer-Verlag
13. Fujimoto, K. and Murofushi, T. (1999): Hierarchical decomposition of the Choquet integral, in: M. Grabisch, T. Murofushi, and M. Sugeno, Eds., Fuzzy Measures and Integrals: Theory and Applications, Springer-Verlag

A. Measures and Integral

In the appendix, we discuss measures on infinite sets and the integral on infinite sets. Arguments on signed measures on infinite sets are omitted.

A.1 Measures

By an *extended real number* we mean a real number, the positive infinity $+\infty$ (or simply ∞), or the negative infinity $-\infty$. The following relations among $\pm\infty$ and real numbers r hold:

$$\begin{aligned}
(\pm\infty)+(\pm\infty) \;=\; r+(\pm\infty) \;=\; (\pm\infty)+r \;&=\; \pm\infty\ , \\
r(\pm\infty) \;=\; (\pm\infty)r \;&=\; \begin{cases} \pm\infty & \text{if } r>0\ , \\ 0 & \text{if } r=0\ , \\ \mp\infty & \text{if } r<0\ , \end{cases} \\
(\pm\infty)(\pm\infty) \;&=\; +\infty\ , \\
(\pm\infty)(\mp\infty) \;&=\; -\infty\ , \\
r/(\pm\infty) \;&=\; 0\ , \\
-\infty \;<\; r \;&<\; +\infty\ .
\end{aligned}$$

Neither $(\pm\infty)+(\mp\infty)$ nor $(\pm\infty)/(\pm\infty)$ can be defined. We denote the set of extended real numbers by $\overline{\mathbf{R}}$ and the set of non-negative extended real numbers by $\overline{\mathbf{R}}_+$,

Definition A.1. *Let μ be an extended real-valued set function defined on a family $\mathcal{X}$ of subsets of a set X.*

(i) *The set function μ is said to be* σ-additive *if*

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) \;=\; \sum_{n=1}^{\infty}\mu(A_n)$$

whenever $\{A_n\}$ is a disjoint sequence of sets in $\mathcal{X}$ (i.e., $\{A_n\}\subset\mathcal{X}$ and $A_i\cap A_j=\emptyset$ for $i\neq j$) and $\bigcup_{n=1}^{\infty}A_n\in\mathcal{X}$.

(ii) *The set function* μ *is said to be* finite *if* $-\infty < \mu(A) < \infty$ *for all* $A \in \mathcal{X}$.
(iii) *The set function* μ *is said to be* continuous from below *if* $\lim_{n\to\infty} m(A_n) = m(A)$ *whenever* $\{A_n\} \subset \mathcal{X}$, *and* $A_n \uparrow A \in \mathcal{X}$, *where* $A_n \uparrow A$ *means that* $\{A_n\}$ *is a non-decreasing sequence and* $\bigcup_{n=1}^{\infty} A_n = A$.
(iv) *The set function* μ *is said to be* continuous from above *if* $\lim_{n\to\infty} m(A_n) = m(A)$ *whenever* $\{A_n\} \subset \mathcal{X}$, $A_n \downarrow A \in \mathcal{X}$, *and* $|m(A_k)| < \infty$ *for some* k. *where* $A_n \downarrow A$ *means that* $\{A_n\}$ *is a non-increasing sequence and* $\bigcap_{n=1}^{\infty} A_n = A$.

The additivity and monotonicity of μ are defined as in Definition 2.1 on condition that all sets in Definition 2.1 are in $\mathcal{X}$.

If $\emptyset \in \mathcal{X}$ and $\mu(\emptyset) = 0$, then σ-additivity of μ implies additivity. If $\emptyset \in \mathcal{X}$ and $\mu(\emptyset) = 0$, and if $\mathcal{X}$ is a finite family, e.g., X is a finite set, then additivity of μ is equivalent to σ-additivity. If $\emptyset \in \mathcal{X}$ and μ is finite and additive, then $\mu(\emptyset) = 0$ (cf. the remark below Definition 2.1).

Definition A.2. *A family* $\mathcal{X}$ *of subsets of* X *is called a* σ-algebra *if* $\mathcal{X}$ *satisfies the following conditions:*

(i) $\emptyset \in \mathcal{X}$ *and* $X \in \mathcal{X}$.
(ii) *If* $A \in \mathcal{X}$, *then* $A^c \in \mathcal{X}$.
(iii) *If* $\{A_n\} \subset \mathcal{X}$, *then* $\bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} A_n \in \mathcal{X}$.

The pair $(X, \mathcal{X})$ *of a set* X *and a* σ*-algebra* $\mathcal{X}$ *of subsets of* X *is called a* measurable space. *When* $(X, \mathcal{X})$ *is a measurable space, a set* A *in* $\mathcal{X}$ *is said to be* $\mathcal{X}$-measurable *(or simply* measurable*).*

Obviously the power set 2^X is the largest σ-algebra and $\{\emptyset, X\}$ is the smallest one.

Let $\mathcal{X}$ be a σ-algebra and μ an extended real-valued set function defined on $\mathcal{X}$. Then the σ-additivity of μ implies the continuity from below. If μ is additive and continuous from below, then it is σ-additive and continuous from above.

Definition A.3. *Let* $(X, \mathcal{X})$ *be a measurable space. A* measure *on* $(X, \mathcal{X})$ *(or on* $\mathcal{X}$*) is an extended real-valued, non-negative,* σ*-additive set function defined on* $\mathcal{X}$ *which vanishes at the empty set. A measure* P *on* $(X, \mathcal{X})$ *is called a* probability measure *if* $P(X) = 1$. *The triplet* $(X, \mathcal{X}, m)$ *of a set* X, *a* σ*-algebra* $\mathcal{X}$ *of subsets of* X, *and a measure* μ *on* $\mathcal{X}$ *is called a* measure space.

The Lebesgue measure is one of the most important measures. The following are brief explanations of the Lebesgue measures on $\mathbf{R}$ and $\mathbf{R}^2$. Details are shown in elementary textbooks on measure theory.

Example A.1. Let $\mathcal{M}$ be the σ-algebra of Lebesgue measurable sets in $\mathbf{R}$. The family $\mathcal{M}$ is very large; intervals, open sets, closed sets, and Borel sets are all in $\mathcal{M}$. The Lebesgue measure λ on $\mathcal{M}$ measures the length of Lebesgue measurable sets; for example, if $a \le b$, then

$$\lambda((a,b)) \;=\; \lambda((a,b]) \;=\; \lambda([a,b)) \;=\; \lambda([a,b]) \;=\; b-a\ .$$

If $a \le b \le c$, then it follows that $(a,b] \cap (b,c] = \emptyset$, $(a,c] = (a,b] \cup (b,c]$, and $c-a = (b-a)+(c-b)$, and therefore the additivity holds (Fig. A.1):

$$\lambda((a,c]) \;=\; \lambda((a,b]) + \lambda((b,c])\ .$$

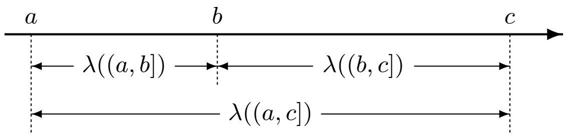


Fig. A.1. Additivity of length

Let $\{a_n\}$ and $\{b_n\}$ be sequences of real numbers such that $a_n \le b_n$ for all n, $a_n \downarrow a$, and $b_n \uparrow b$ (for example, $a_n = 1/n$, $b_n = 3-(1/n)$, $a=0$, and $b=3$). In addition, let $A_n = (a_n, b_n)$ for all n and $A=(a,b)$. Then $A_n \uparrow A$, $\lambda(A_n) = b_n - a_n$ for all n, and $\lambda(A) = b-a$. Since $a_n \downarrow a$ and $b_n \uparrow b$, it follows that $(b_n - a_n) \uparrow (b-a)$, and hence this is an example of the continuity from below: $\lambda(A_n) \uparrow \lambda(A)$ (Fig. A.2). Similarly, we can easily make an example of the continuity from above.

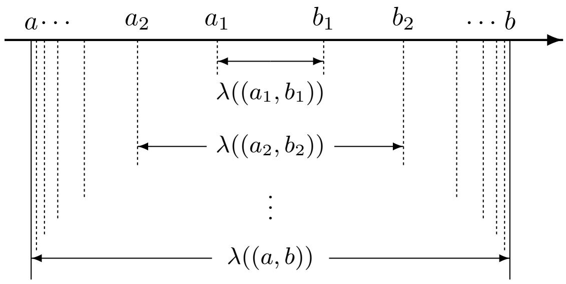


Fig. A.2. Continuity of length

Example A.2. Let $\mathcal{M}$ be the σ-algebra of Lebesgue measurable sets in $\mathbf{R}^2$. The Lebesgue measure λ on $\mathcal{M}$ measures the area of Lebesgue measurable sets; for example, if $a \le b$ and $c \le d$, then

$$\lambda((a,b] \times (c,d]) \;=\; (b-a)\cdot(d-c)$$

(Fig. A.3). We know that the area is additive (Fig. A.4), and that the area is continuous (Fig. A.5).

Definition A.4. *Let $(X, \mathcal{X}, m)$ be a measure space. A measurable set N is called a* null *set (or m-*null *set) if $m(N) = 0$.*

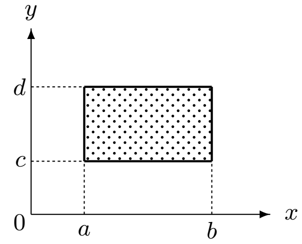


Fig. A.3. $(a, c] \times (c, d]$

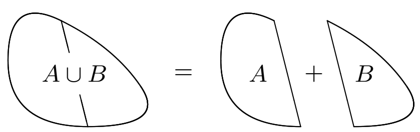
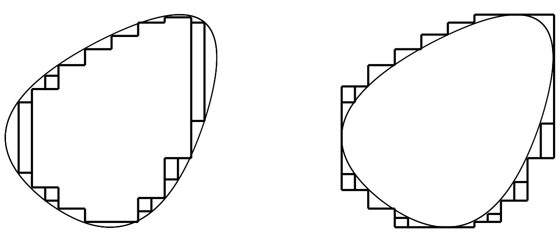


Fig. A.4. Additivity of area



(a) Continuity from below (b) Continuity from above

Fig. A.5. Continuity of area

A.2 Integral

Let $(X, \mathcal{X})$ be a measurable space.

Definition A.5. *An extended real-value function f defined on X is said to be* $\mathcal{X}$-measurable *(or simply* measurable*) if for every real number r, the set $\{x \mid f(x) > r\}$ is $\mathcal{X}$-measurable (Fig. A.6). A* simple function *is a measurable function whose range is a finite subset of* $\mathbf{R}$.

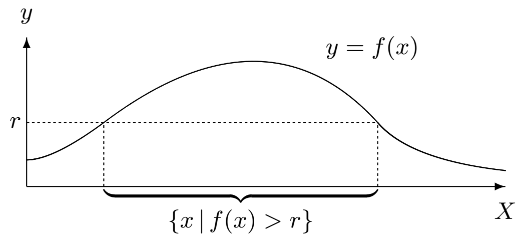


Fig. A.6. Measurability of function f

Note that in the above definition the set $\{f(x) > r\}$ can be replaced with any of $\{x \mid f(x) \geq r\}$, $\{x \mid f(x) < r\}$, and $\{x \mid f(x) \leq r\}$. If f and g are $\mathcal{X}$-measurable functions, then so are $af + bg$ (a and b are constants), $f \cdot g$, $\max\{f, g\}$, $\min\{f, g\}$, and $|f|$. If f is a simple function, then it is represented as

$$f \;=\; \sum_{i=1}^{n} a_i 1_{A_i} \;, \tag{A.1}$$

where $\{a_i\} \subset \mathbf{R}$ and $\{A_i\} \subset \mathcal{X}$.

Definition A.6. *Let $(X, \mathcal{X}, m)$ be a measure space. The* integral $\int f(x)\, dm(x)$ *(or simply $\int f\, dm$) of a measurable function f with respect to m is defined in steps as follows.*

(i) *For a non-negative simple function $f = \sum_{i=1}^{n} a_i 1_{A_i}$,*

$$\int f\, dm \;=\; \sum_{i=1}^{n} a_i m(A_i) \;.$$

(See Figs. 2.2–2.5.)

(ii) *For a non-negative measurable function f,*

$$\int f\, dm \;=\; \sup\left\{ \int g\, dm \,\middle|\, g \text{ is a simple function, } 0 \leq g \leq f \right\} \;.$$

(See Fig. A.7.)

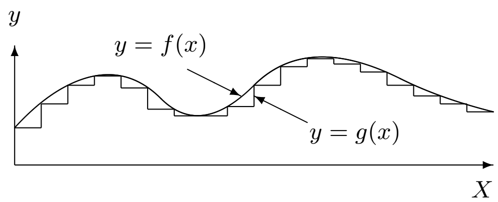


Fig. A.7. Simple function g such that $g \leq f$

(iii) *For a measurable function f,*

$$\int f\, dm \;=\; \int f^+\, dm - \int f^-\, dm \;,$$

where $f^+(x) = \max\{f(x), 0\}$, $f^-(x) = \max\{-f(x), 0\}$ (Fig. A.8), and at least one of $\int f^+ dm$ and $\int f^- dm$ is finite.

The integral $\int_A f(x)\, dm(x)$ over a measurable set A of f with respect to m is defined as in Definition 2.7. The same proposition as Proposition 2.2 holds on condition that all sets and functions are measurable.

The following example illustrates the meaning of the integral. However, it is not mathematically strict.

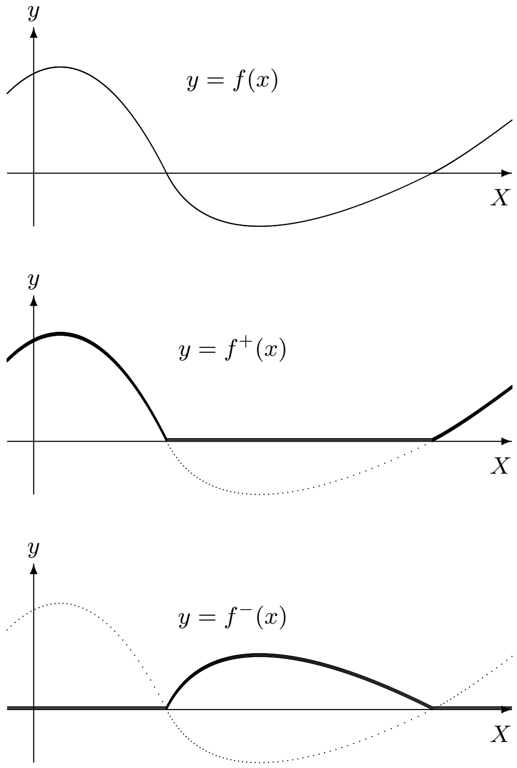


Fig. A.8. Function f, its positive part f^+, and its negative part f^-

Example A.3. Consider a wire with segmentally homogeneous density; the wire is segmented into A_1, A_2, $\ldots$, A_n, and the line density of each part A_i is d_i g/cm (Fig. A.9). The total mass w g of this wire is expressed as

$$w \;=\; \sum_{i=1}^{n} d_i \cdot \lambda(A_i) \;,$$

where λ is a measure which measures the length of parts of the wire in cm.

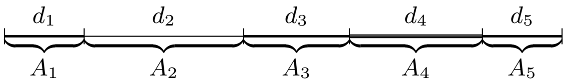


Fig. A.9. Segmentally homogeneous wire

Next consider another wire whose line density at position x is $f(x)$ g/cm. The line density $f(x)$ can be regarded as the limit of density of a part A containing x when A becomes infinitely short. (The line density of the segmentally homogeneous wire at position $x \in A_i$ is d_i g/cm.) The total mass w g of the wire is also expressed as

$$w = \int f \, d\lambda$$

and the mass w_A g of a part A is expressed as

$$w_A = \int_A f \, d\lambda \, .$$

Similar arguments can be applied to higher dimensions. In two dimension, objects are shifted from wires to metallic plates, the density $f(x)$ is shifted from line density (g/cm) to surface density (g/cm^2), and the quantity measured by λ is shifted from the length (cm) to the area (cm^2). The integral $\int f \, d\lambda$ represents the mass (g) of the plate. The images of segmentally homogeneous and heterogeneous metallic plates are shown in Figs. A.10 and A.11, respectively. In three dimensions, the object becomes a lump of metal, the integrand $f(x)$ is volume density (g/cm^3), and the measure λ measures the volume (cm^3). Regardless of dimensions, the integral $\int f \, d\lambda$ expresses the mass (g) of the object.

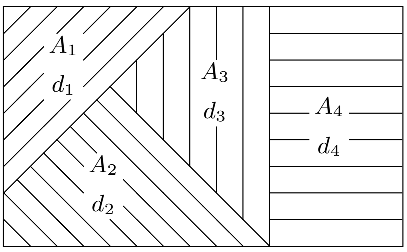


Fig. A.10. Segmentally homogeneous metallic plate

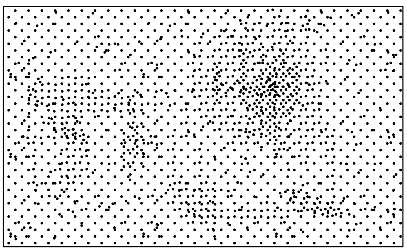


Fig. A.11. Heterogeneous metallic plate