POSSIBILITY THEORY I
THE MEASURE- AND INTEGRAL-THEORETIC GROUNDWORK

GERT DE COOMAN

Universiteit Gent, Vakgroep Elektrische Energietechniek
Technologiepark 9, B-9052 Zwijnaarde, Belgium.
E-mail: gert.decooman@rug.ac.be

Dedicated to Prof. dr. Etienne E. Kerre

In this paper, I provide the basis for a measure- and integral-theoretic formulation of possibility theory. It is shown that, using a general definition of possibility measures, and a generalization of Sugeno's fuzzy integral – the seminormed fuzzy integral, or possibility integral –, a unified and consistent account can be given of many of the possibilistic results extant in the literature. The striking formal analogy between this treatment of possibility theory, using possibility integrals, and Kolmogorov's measure-theoretic formulation of probability theory, using Lebesgue integrals, is explored and exploited. I introduce and study possibilistic and fuzzy variables as possibilistic counterparts of stochastic and real stochastic variables respectively, and develop the notion of a possibility distribution for these variables. The almost everywhere equality and dominance of fuzzy variables is defined and studied. The proof is given for a Radon-Nikodym-like theorem in possibility theory. Following the example set by the classical theory of integration, product possibility measures and multiple possibility integrals are introduced, and a Fubini-like theorem is proven. In this way, the groundwork is laid for a unifying measure- and integral-theoretic treatment of conditional possibility and possibilistic independence, discussed in more detail in Parts II and III of this series of three papers.

INDEX TERMS: Possibility measure, seminormed fuzzy integral, possibilistic variable, fuzzy variable, possibility distribution, Radon-Nikodym-like theorem, Fubini-like theorem.

1 INTRODUCTION

The theory of probability originated in the middle of the eighteenth century from considerations about games of chance, under the influence of, among others, Fermat, Pascal, Huyghens and James Bernoulli. Important mathematicians such as DeMoivre, Laplace, Gauss and Poisson further developed it, and it gradually became an impressive body of results about randomness and random events, with applications in statistics, demography, systems science, chemistry and physics, to name only a few important domains. As an inevitable consequence of this success and rapid growth, very soon the need was felt for a solid and formal foundation. At the same time, the theory was still lacking a consistent and unifying presentation of the diversity of its notions and results, gathered over decades and even centuries. The first formal foundation of probability theory was given by the Russian mathematician S. N. Bernstein. An alternative axiomatic foundation is due to yet another Russian scientist, A. N. Kolmogorov, who was the first to show that it is possible to combine probability theory with the measure theory to which Lebesgue had devoted much of his life. His position was that probability can be mathematically represented by a normalized measure. At the same time, using Lebesgue integration theory, he succeeded in formulating the theory of probability in a measure- and integral-theoretic language, and thus provided a unified and consistent account of the study of random events and chance [Kolmogorov, 1950].

On the other hand, many prominent scientists have defended the view that uncertainty and randomness are not one and the same thing. To put it more succinctly, uncertainty can be caused by more than chance alone. In 1978, Zadeh [1978a] observed in his seminal paper dealing with possibility theory, that people often convey information using simple affirmative sentences in natural language of the type 'subject-verb-predicate', where the predicate tells us something about the subject, or, to formulate it in a more mathematical way, imposes a restriction on the values that the subject may assume in an appropriate universe of discourse. The information transmitted or represented in this way, could be called *'linguistic information'*. In most cases in natural language, however, the information contained in such a sentence is not sufficient to unequivocally determine the subject, simply because the predicate involved is *imprecise* or *vague*. To give an example, the proposition 'John's age is between 20 and 30 years' gives us some information about how old John is, but does not completely determine his age, since the predicate 'between 20 and 30 years' is imprecise. Similarly, when we say 'Mary is tall', we give some information about Mary's stature, but since 'tall' is a vague predicate, this information is not sufficient to completely determine it.

In both cases, although the sentence involved contains information, it still leaves us uncertain as to the precise value which the subject of the sentence assumes in its universe of discourse. We have thus stumbled upon a kind of uncertainty, henceforth called '*linguistic uncertainty*', which cannot be directly attributed to randomness or chance, and which is nevertheless always present in our normal, everyday communication.

Zadeh has advanced the thesis that linguistic information, or dually, linguistic uncertainty, has nothing to do with probability and cannot be represented by probability measures. In order to provide a mathematical representation for this uncertainty, he introduced possibility measures, fuzzy variables and their possibility distributions, product possibility measures, and the notion of noninteractivity[1] [Zadeh, 1978a]. He thus laid the foundation of a *theory of possibility*, a collection of notions and results concerning possibility measures, and, in Zadeh's

[1] Noninteractivity is a special case of what I shall possibilistic independence in Part III of this series.

view, dealing with the mathematical description of linguistic uncertainty. Zadeh's work has been further refined and extended by a number of scientists. Without aiming at completeness, I want to mention Dubois and Prade and various co-workers, who among many other things introduced the dual notion of a *necessity measure*, studied the relationships between possibility and necessity measures and other representations of uncertainty [Dubois and Prade, 1985, 1988], and more recently presented a study of possibilistic independence [Benferhat *et al.*, 1994, 1995], [Dubois *et al.*, 1994a], of conditioning in a possibilistic framework [Dubois and Prade, 1990] and of possibilistic logic [Dubois *et al.*, 1994b]; Klir and various colleagues, who introduced the notion of a measure of uncertainty for possibility measures [Higashi and Klir, 1982], [Klir and Folger, 1988], studied ways of converting possibility measures into probability measures and *vice versa* [Klir, 1990], [Klir and Parviz, 1992], studied the relationship between possibility theory and the Dempster-Shafer theory of evidence and explored the relationship between possibility theory and modal logic [Klir and Harmanec, 1995]; and Hisdal [1978], Nguyen [1978] and Ramer [1989], who also contributed to the subject of conditional possibility.

It should be stressed that there are other interpretations of possibility measures than the one given by Zadeh. It has for instance been noted that possibility measures are special types of belief measures [Shafer, 1976] and upper probabilities [Walley, 1991], and that they can therefore be interpreted within the framework of the Dempster-Shafer theory of evidence and of imprecise probabilities, respectively. Giles [1982] made an attempt to interpret possibility measures as some kind of upper probabilities. Dubois and Prade have also studied interpretations of possibility theory in relation with likelihood functions [Dubois *et al.*, 1993a], and with epistemic states [Dubois *et al.*, 1993b]. It goes without saying that most, if not all, of the abstract mathematical results in this paper remain valid under any interpretation that might be given to the specific possibility measures that will be defined and used further on.

This is the first of a series of three papers dealing with the measure- and integral-theoretic treatment of possibility theory. With this series, it is my aim to accomplish to some extent in possibility theory what Kolmogorov has succeeded in doing for probability theory. In this first paper, using a very general definition of a possibility measure and an appropriate generalization of Sugeno's fuzzy integral – the seminormed fuzzy integral –, I give a consistent and unifying account of possibility theory, clarifying and resolving in the process some inconsistencies and difficulties still extant in the literature. At the same time, I show that possibility theory can be developed along the same general lines as Kolmogorov's theory of probability[2], and reveal the deep symmetry between both accounts of uncertainty. Finally, I present this material as mathematical evidence, confirming my suspicion that possibility measures and (seminormed) fuzzy integrals are a perfect match, in very much the same way as classical measures and Lebesgue integrals are. In the second and third parts of this series, I use the results derived here to work out a measure- and integral-theoretic treatment of conditional possibility and possibilistic independence. The results published in this series are a condensation of part of my doctoral dissertation [De Cooman, 1993], written in Dutch, and a significant extension of a short overview paper [De Cooman, 1995a].

In section 2, I have collected the basic mathematical notions, necessary for the proper understanding of the rest of the material. Possibility measures and seminormed fuzzy integrals are briefly introduced in section 3, together with a number of their properties that will be used later in the paper, and in the other papers in this series. A formal treatment of possibilistic variables,

[2]It should be stressed that in this series of papers, whenever I talk about probability theory, I mean Kolmogorov's measure- and integral-theoretic account of it.

the possibilistic counterparts of stochastic variables, is proposed in section 4. The special case of what I call fuzzy variables, the formal counterparts of the real stochastic variables in probability theory, is treated in more detail in section 5. Section 6 deals with the almost everywhere equality and dominance of these fuzzy variables. A Radon-Nikodym-like theorem for seminormed fuzzy integrals and possibility measures is the subject of section 7. An integral-theoretic approach to product possibility measures, and the product and chain integrals associated with them, follows in section 8, which also contains a possibilistic counterpart of the well-known theorem of Fubini. In section 9 I briefly look back at what has been accomplished, and pave the way for the treatment of conditional possibility and possibilistic independence in the second and third paper of this series.

2 PRELIMINARY DEFINITIONS

Let us start this discussion with a few preliminary definitions and notational conventions, valid in the rest of this paper, unless explicitly stated to the contrary. We denote by X an arbitrary *universe of discourse*, i.e., a nonempty set. The universes considered hereafter will be implicitly assumed to contain at least two elements.

By $(L, \leq)$ we mean a complete lattice that is arbitrary but *fixed throughout the whole text*. The smallest element of $(L, \leq)$ is denoted by 0_L and the greatest element by 1_L. We also assume that $0_L \neq 1_L$. The meet of $(L, \leq)$ is denoted by $\frown$, the join by $\smile$.

2.1 Triangular Seminorms and Norms

A *triangular seminorm* or, shortly, *t-seminorm* P on the complete lattice $(L, \leq)$ is a binary operator on L that satisfies the following two conditions:

(i) boundary conditions:
$(\forall \lambda \in L)(P(1_L, \lambda) = P(\lambda, 1_L) = \lambda)$;

(ii) isotonicity:
$(\forall (\lambda_1, \lambda_2, \mu_1, \mu_2) \in L^4)(\lambda_1 \leq \lambda_2 \text{ and } \mu_1 \leq \mu_2 \Rightarrow P(\lambda_1, \mu_1) \leq P(\lambda_2, \mu_2))$.

A *triangular norm* or, shortly, *t-norm* T on $(L, \leq)$ is a t-seminorm on $(L, \leq)$ that is furthermore associative and commutative. Of course, $\frown$ is a triangular norm on $(L, \leq)$, and more specifically, the only one that is idempotent. For a more involved account of triangular seminorms and norms defined on complete lattices, and more in general, on bounded partially ordered sets, I refer to [De Cooman and Kerre, 1994], where the reader will find more details about the notions discussed in this subsection.

A t-seminorm P on $(L, \leq)$ is called *completely distributive* w.r.t. supremum iff for any λ in L and any family $\{\mu_j \mid j \in J\}$ of elements of L: $P(\lambda, \sup_{j \in J} \mu_j) = \sup_{j \in J} P(\lambda, \mu_j)$ and $P(\sup_{j \in J} \mu_j, \lambda) = \sup_{j \in J} P(\mu_j, \lambda)$. In this case, the structure $(L, \leq, P)$ is called a *complete lattice with t-seminorm*. Similarly, whenever the t-norm T on $(L, \leq)$ is completely distributive w.r.t. supremum, we call the structure $(L, \leq, T)$ a *complete lattice with t-norm*. Of course, a complete lattice with t-norm $(L, \leq, \frown)$ is a complete Brouwerian lattice [Birkhoff, 1967].

Let λ and μ be elements of L and let P be a t-seminorm on $(L, \leq)$. An element α of L is called a *left-inverse* for P of λ w.r.t. μ iff $P(\alpha, \lambda) = \mu$ and a *right-inverse* for P of λ w.r.t. μ iff $P(\lambda, \alpha) = \mu$. For a t-norm T on $(L, \leq)$, left-inverses and right-inverses coincide due to the commutativity of T, and are simply called *inverses*. The following, rather interesting properties are easily proven:

$$\begin{aligned} \mu \not\leq \lambda \Rightarrow & \text{ there are no left- and right-inverses for } P \text{ of } \lambda \text{ w.r.t. } \mu \\ \mu \not\leq \lambda \Rightarrow & \text{ there are no inverses for } T \text{ of } \lambda \text{ w.r.t. } \mu \end{aligned}$$

In this light, a t-seminorm P on $(L, \leq)$ is called *weakly left-invertible* iff for any λ and μ in L with $\mu \leq \lambda$, there exists a left-inverse for P of λ w.r.t. μ. The definition of weak right-invertibility is completely similar. A t-norm T on $(L, \leq)$ is called *weakly invertible* iff for any λ and μ in L with $\mu \leq \lambda$, there exists an inverse for T of λ w.r.t. μ.

In the rest of this subsection, we assume that $(L, \leq, P)$ is a complete lattice with t-seminorm, and that $(L, \leq, T)$ is a complete lattice with t-norm. It has been shown previously [De Cooman and Kerre, 1994] that for these structures, there exists an important relation between the notion of a weak (left- and right-)inverse and the order-theoretic notion of *residuation* [Birkhoff, 1967]. In order to explain this, let us again consider arbitrary elements λ and μ of L. The *left-residual* $\lambda \lhd_P \mu$ for P of λ by μ is defined as

$$\lambda \lhd_P \mu = \sup \{ \nu \mid \nu \in L \text{ and } P(\nu, \mu) \leq \lambda \}.$$

For the right-residual $\mu \rhd_P \lambda$ for P of λ by μ a similar definition can be given. For the t-norm T the notions of left- and right-residual coincide and are simply called *residual*. The residual $\lambda \triangle_T \mu$ for T of λ by μ is defined as $\lambda \triangle_T \mu = \lambda \lhd_T \mu = \mu \rhd_T \lambda$. The connection between residuals and inverses is made explicit by the following propositions [De Cooman and Kerre, 1994].

Proposition 2.1. *Let λ and μ be elements of L.*

(i) If the equation $P(\nu, \lambda) = \mu$ in ν admits a solution, then $\mu \lhd_P \lambda$ is the greatest solution w.r.t. the order relation $\leq$, and if the equation $P(\lambda, \nu) = \mu$ in ν admits a solution, then $\lambda \rhd_P \mu$ is the greatest solution w.r.t. the order relation $\leq$.

(ii) If the equation $T(\nu, \lambda) = \mu$ in ν admits a solution, then $\mu \triangle_T \lambda$ is the greatest solution w.r.t. the order relation $\leq$.

Proposition 2.2. *(i) P is weakly left-invertible iff $(\forall(\lambda, \mu) \in L^2)(\lambda \geq \mu \Rightarrow P(\mu \lhd_P \lambda, \lambda) = \mu)$ and weakly right-invertible iff $(\forall(\lambda, \mu) \in L^2)(\lambda \geq \mu \Rightarrow P(\lambda, \lambda \rhd_P \mu) = \mu)$.*

(ii) T is weakly invertible iff $(\forall(\lambda, \mu) \in L^2)(\lambda \geq \mu \Rightarrow T(\mu \triangle_T \lambda, \lambda) = \mu)$.

In order to clarify these results, let us look at a number of familiar and important examples.

Example 2.3. If we consider the real unit interval $[0, 1]$, ordered by the natural order of real numbers, $([0, 1], \leq)$ is a complete chain, with meet min, the well-known minimum operator. Interestingly, $([0, 1], \leq)$ is a complete Brouwerian chain, or equivalently, $([0, 1], \leq, \min)$ is a complete chain with t-norm. It is easily verified that for any x and y in $[0, 1]$:

$$x \triangle_{\min} y = \sup \{ z \mid z \in [0, 1] \text{ and } \min(z, y) \leq x \} = \begin{cases} 1 & ; \quad y \leq x \\ x & ; \quad y > x \end{cases}$$

and $\min(x \triangle_{\min} y, y) = \min(x, y)$, which implies that min is weakly invertible.

Alternatively, we may consider the algebraic product operator $\times$ on $[0,1]$, which clearly is a triangular norm on $([0,1],\leq)$ that is furthermore completely distributive w.r.t. supremum. In other words, the structure $([0,1],\leq,\times)$ is a complete chain with t-norm, and for any x and y in $[0,1]$, we find that

$$x\triangle_{\times}y = \sup\{\, z \mid z\in[0,1] \text{ and } z\times y\leq x\,\} = \begin{cases} 1 & ; \quad y\leq x \\ x/y & ; \quad y>x \end{cases}$$

and $(x\triangle_{\times}y)\times y = \min(x,y)$, which implies that $\times$ is also weakly invertible.

2.2 Ample Fields

An *ample field* $\mathcal{R}$ on the universe X is a set of subsets of X that is closed under *arbitrary* unions and intersections, and under complementation in X. The ample fields we shall consider, are assumed to be *proper*, i.e., $\{\emptyset, X\}\subset\mathcal{R}$. A special ample field on X, and at the same time the largest, is the *power set* $\wp(X)$ of X, i.e. the set of all subsets of X. In this sense, ample fields can be considered as immediate generalizations of this power set. For a discussion of this subject, I refer to [De Cooman and Kerre, 1993], [Wang, 1982] and [Wang and Klir, 1992].

The *atom* of $\mathcal{R}$ containing the element x of X is denoted by $[\,x\,]_{\mathcal{R}}$ and is defined by:

$$[\,x\,]_{\mathcal{R}} = \bigcap\{\, A \mid x\in A \text{ and } A\in\mathcal{R}\,\}.$$

It is easily proven that the set of all atoms of $\mathcal{R}$ constitutes a partition of X. Remark that $[\,x\,]_{\wp(X)} = \{x\}$, for any x in X. Therefore, atoms can be interpreted as generalizations of singletons. In this light, we also have that for any x in X and A in $\mathcal{R}$, $x\in A \Leftrightarrow [\,x\,]_{\mathcal{R}}\subseteq A$.

A subset E of X is called *$\mathcal{R}$-measurable* iff $E\in\mathcal{R}$. If no confusion can arise, we also simply call E measurable. A measurable set is also called an *event*. Furthermore, it is interesting to note that

$$E\in\mathcal{R} \Leftrightarrow E = \bigcup_{x\in E}[\,x\,]_{\mathcal{R}}. \tag{1}$$

Consider an arbitrary subset $\mathcal{E}$ of $\wp(X)$. Since the intersection of any family of ample fields is again an ample field, we know that

$$\tau(\mathcal{E}) = \bigcap\{\,\mathcal{R} \mid \mathcal{R} \text{ is an ample field on } X \text{ and } \mathcal{E}\subseteq\mathcal{R}\,\}$$

is an ample field on X, called the *ample field generated by* $\mathcal{E}$. τ can be regarded as an operator on $\wp(\wp(X))$, and is as such a closure operator on $\wp(X)$ [Davey and Priestley, 1990], i.e., for any subsets $\mathcal{E}$, $\mathcal{E}_1$ and $\mathcal{E}_2$ of $\wp(X)$, $\mathcal{E}\subseteq\tau(\mathcal{E})$, $\tau(\tau(\mathcal{E}))=\tau(\mathcal{E})$ and $\mathcal{E}_1\subseteq\mathcal{E}_2 \Rightarrow \tau(\mathcal{E}_1)\subseteq\tau(\mathcal{E}_2)$. Furthermore, $\tau(\mathcal{E})=\mathcal{E}$ iff $\mathcal{E}$ is an ample field on X. This notion can be used to introduce product ample fields. If we consider the universes X_1 and X_2 provided with the respective ample fields $\mathcal{R}_1$ and $\mathcal{R}_2$, then the *product ample field* of $\mathcal{R}_1$ and $\mathcal{R}_2$ is the ample field on $X_1\times X_2$ defined as

$$\mathcal{R}_1\times\mathcal{R}_2 = \tau(\{\,A_1\times A_2 \mid A_1\in\mathcal{R}_1 \text{ and } A_2\in\mathcal{R}_2\,\}). \tag{2}$$

Interestingly, for the atoms of $\mathcal{R}_1\times\mathcal{R}_2$, we have that

$$(\forall(x_1,x_2)\in X_1\times X_2)([\,(x_1,x_2)\,]_{\mathcal{R}_1\times\mathcal{R}_2} = [\,x_1\,]_{\mathcal{R}_1}\times[\,x_2\,]_{\mathcal{R}_2}), \tag{3}$$

which confirms the interpretation of an atom as a generalization of a singleton. The generalization of these results for products of more than two ample fields is immediate.

Finally, given an ample field $\mathcal{R}$ on a universe X, we may associate with it a $\wp(X) - \wp(X)$-mapping $p_{\mathcal{R}}$ – actually, the closure operator on X associated with the closure system $\mathcal{R}$ –, defined by, for any A in $\wp(X)$:

$$p_{\mathcal{R}}(A) = \bigcup_{x \in A} [\, x\,]_{\mathcal{R}}. \tag{4}$$

2.3 Fuzzy Sets and Fuzzy Variables

With any subset A of a universe X, we can associate its *characteristic* $X - L$ *mapping* χ_A, defined by

$$\chi_A(x) = \begin{cases} 1_L & ; \quad x \in A \\ 0_L & ; \quad x \in \text{co}A. \end{cases}$$

In accordance with the terminology introduced by Goguen [1967], an arbitrary $X - L$ mapping will be called a $(L, \leq)$-*fuzzy set* (or simply fuzzy set) in X. It is an obvious generalization of a characteristic $X - L$-mapping. The set of the $(L, \leq)$-fuzzy sets in X is denoted by $\mathcal{F}_{(L,\leq)}(X)$. We shall also make use of the partial order relation $\sqsubseteq$ on $\mathcal{F}_{(L,\leq)}(X)$, defined as follows: for any h_1 and h_2 in $\mathcal{F}_{(L,\leq)}(X)$,

$$h_1 \sqsubseteq h_2 \Leftrightarrow (\forall x \in X)(h_1(x) \leq h_2(x)).$$

For any λ in L, $\underline{\lambda}$ denotes the constant $X - \{\lambda\}$ mapping.

A $X - L$ mapping h is called $\mathcal{R}$-*measurable* iff it is constant on the atoms of $\mathcal{R}$. The origin of this definition of measurability is discussed in [De Cooman and Kerre, 1993], where it is also shown that it is an obvious and natural extension of the normal measurability condition imposed on real mappings in classical measure and probability theory. A $\mathcal{R}$-measurable $X - L$ mapping – or $(L, \leq)$-fuzzy set in X – is also called a $(L, \leq)$-*fuzzy variable* in $(X, \mathcal{R})$. Whenever we want to omit reference to the structures $(L, \leq)$ and $(X, \mathcal{R})$, we simply speak of fuzzy variables. A fuzzy variable can therefore be considered as a 'fuzzification' of a measurable set, and will sometimes be called a *fuzzy event*. Indeed, a subset E of X is $\mathcal{R}$-measurable iff its characteristic $X - L$-mapping is. The set of the $(L, \leq)$-fuzzy variables in $(X, \mathcal{R})$ is denoted by $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$.

If $n \in \mathbb{N}^*$, i.e., n is a strictly positive natural number, the $X - L^n$-mapping $(h_1, \ldots, h_n)$ is a n-*dimensional fuzzy variable* in $(X, \mathcal{R})$ iff each of its component mappings h_k, $k \in \{1, \ldots, n\}$, is a fuzzy variable in $(X, \mathcal{R})$. It is obvious that a fuzzy variable can always be considered as a special case of a multidimensional fuzzy variable. A more detailed account of the fuzzy variables introduced here, together with a deeper discussion of their meaning and interpretation, is deferred to section 5.

3 THE POSSIBILITY INTEGRAL AND ITS MEANING

Before addressing the details of the measure- and integral-theoretic aspects of possibility theory in the following section, let me define here the basic vocabulary, and explain what I mean by a possibility measure and a possibility integral.

3.1 Possibility Measures

Let $\mathcal{R}$ be an ample field on the universe X. A $(L,\leq)$-*possibility measure* Π on $(X,\mathcal{R})$ is a complete join-morphism between the complete lattices $(\mathcal{R},\subseteq)$ and $(L,\leq)$. This means by definition that Π satisfies the following requirement: for any family $\{A_j \mid j\in J\}$ of elements of $\mathcal{R}$

$$\Pi(\bigcup_{j\in J} A_j) = \sup_{j\in J}\Pi(A_j).$$

This definition immediately implies that $\Pi(\emptyset)=0_L$. For any A in $\mathcal{R}$, $\Pi(A)$ is called the $(L,\leq)$-*possibility of A*. The structure $(X,\mathcal{R},\Pi)$ is called a $(L,\leq)$-*possibility space*. Π is called *normal* iff $\Pi(X)=1_L$. Whenever we do not want to mention the complete lattice $(L,\leq)$ explicitly, we simply speak of possibility, possibility measures and possibility spaces.

A $\mathcal{R}$-measurable $X-L$-mapping π such that for any A in $\mathcal{R}$, $\Pi(A)=\sup_{x\in A}\pi(x)$, is called a *distribution* of Π. Such a distribution is *unique*, and satisfies $(\forall x\in X)(\pi(x)=\Pi([\,x\,]_{\mathcal{R}}))$.

$(L,\leq)$-possibility measures are generalizations towards more general domains and codomains of Zadeh's possibility measures [Zadeh, 1978a], Wang's fuzzy contactabilities [Wang, 1982], the possibility measures studied by Wang and Klir [1992], and the possibility measures introduced in [De Cooman *et al.*, 1992]. For a more detailed discussion of these generalizations, I refer to [De Cooman, 1993], [De Cooman and Kerre, 1993] and [De Cooman *et al.*, 1992]. Their introduction can be justified as follows. Using a complete lattice as a codomain allows us to model potential *incomparability* of possibilities, and for instance to associate possibility measures with the $(L,\leq)$-fuzzy sets introduced by Goguen [1967], in order to represent more general forms of linguistic uncertainty [De Cooman, 1995d], [Zadeh, 1978a]. Why possibility is defined here on ample fields, rather than on power sets, needs a justification that is more involved. Actually, we could call a mapping from an arbitrary subset $\mathcal{E}$ of $\wp(X)$ to L a $(L,\leq)$-possibility measure if it is *extendable* to a $(L,\leq)$-possibility measure with domain $\wp(X)$. It is shown elsewhere [Boyen *et al.*, 1995] that this extendability is equivalent with the extendability to a $(L,\leq)$-possibility measure with domain $\tau(\mathcal{E})$, the smallest ample field that includes $\mathcal{E}$. Therefore, ample fields arise naturally as domains of possibility measures. Moreover, generally speaking, if we want to define a possibility measure on $\wp(X)$, we have to be more specific than if we want to define one on any other ample field $\mathcal{R}$ on X, since, for instance, the atoms of $\wp(X)$ constitute a refinement of the atoms of $\mathcal{R}$. So, at least in principle, we want to be able to be as nonspecific as possible, and introduce possibility measures on ample fields and not just on power sets.

3.2 Possibility Integrals

In this subsection, we assume that P is a t-seminorm on $(L,\leq)$, such that $(L,\leq,P)$ is a complete lattice with t-seminorm. Furthermore $(X,\mathcal{R},\Pi)$ is a $(L,\leq)$-possibility space, and the distribution of Π is denoted by π.

In a recent article about possibility and necessity integrals [De Cooman and Kerre, 1995], I argued that a generalization of Sugeno's fuzzy integral, the seminormed fuzzy integral, is ideally suited for combination with $(L,\leq)$-possibility measures. In the rest of this paper and also in Parts II and III of this series, I want to explore this idea, and show that these seminormed fuzzy integrals can be used to give a consistent and unifying account of possibility theory. Let me, without going into detail, repeat here the most important points of the argumentation in the above-mentioned article, and explicitly write down the formulas that we shall need further on. For an explicit definition of the seminormed fuzzy integral, and a detailed account of its properties, I refer to [De Cooman and Kerre, 1995].

For a start, whenever the above-mentioned integral is associated with a possibility measure, we call it a *possibility integral*. It turns out that if we combine a seminormed fuzzy integral with a possibility measure, there exists a very simple formula for its calculation. Indeed, for an arbitrary $\mathcal{R}$-measurable set A and an arbitrary $\mathcal{R}$-measurable $X-L$-mapping h, the $(L,\leq,P)$-*possibility integral* of h on A (w.r.t. Π) can be written as:

$$(P)\,⨍_A h \mathrm{d}\Pi = \sup_{x\in A} P(h(x),\pi(x)). \tag{5}$$

Moreover, if μ is an element of L, we find for the constant mapping $\underline{\mu}$ that

$$(P)\,⨍_A \underline{\mu}\mathrm{d}\Pi = P(\mu,\Pi(A)). \tag{6}$$

As a special case, we can integrate the constant mapping $\underline{1_L}$ over A, which leads to

$$(P)\,⨍_A \mathrm{d}\Pi = (P)\,⨍_A \underline{1_L}\mathrm{d}\Pi = \Pi(A). \tag{7}$$

Also useful is the following relation:

$$(P)\,⨍_A h\mathrm{d}\Pi = (P)\,⨍_X (h \frown \chi_A)\mathrm{d}\Pi, \tag{8}$$

where, for any x in X, $(h \frown \chi_A)(x) = h(x) \frown \chi_A(x)$. As a corollary of this, we find for the characteristic $X-L$-mapping of the $\mathcal{R}$-measurable set A that

$$(P)\,⨍_X \chi_A\mathrm{d}\Pi = (P)\,⨍_A \underline{1_L}\mathrm{d}\Pi = \Pi(A). \tag{9}$$

If we define the mapping $\Pi_P\colon \mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X) \to L$ by

$$\Pi_P(h) = (P)\,⨍_X h\mathrm{d}\Pi = \sup_{x\in X} P(h(x),\pi(x)), \tag{10}$$

Eq. (9) tells us that $\Pi_P(\chi_A) = \Pi(A)$, which means that the mapping Π_P can be considered as an '*extension*' of the mapping Π from $\mathcal{R}$ to $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$. Furthermore, for any family $\{h_j \mid j \in J\}$ of fuzzy variables in $(X,\mathcal{R})$:

$$(P)\,⨍_X (\sup_{j\in J} h_j)\mathrm{d}\Pi = \sup_{j\in J}(P)\,⨍_X h_j\mathrm{d}\Pi, \tag{11}$$

which means that Π_P is a complete join-morphism between the complete lattices $(\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X),\sqsubseteq)$ and $(L,\leq)$, and therefore behaves like a possibility measure. The $(L,\leq,P)$-possibility integral allows us in other words to 'extend' the domain of the possibility measure Π from $\mathcal{R}$ to $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$. Π_P is called the *P-extension* of Π, and for any h in $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$, $\Pi_P(h)$ is called the $(L,\leq,P)$-*possibility*, or also (generalized) possibility, of h.

4 POSSIBILISTIC VARIABLES

In this section, I give a short description of what I mean by a *possibilistic variable*. I also introduce *possibility distributions* and *possibility distribution functions* for these variables. In order to keep this survey as tidy as possible, I have collected a number of preliminary definitions and results in subsection 4.1. The proper introduction of possibilistic variables and their distributions is given in subsection 4.2. A few related additional results about the calculation of possibility integrals by a *transformation of variables* can be found in subsection 4.3.

4.1 Preliminary Definitions

In probability theory, and in measure theory in general, it is possible to transfer a (probability) measure from one universe to another, using a mapping between these universes [Burrill, 1972]. This procedure is commonly called the *transformation of a measure by a mapping*. It is not difficult to show that something completely similar can be done with possibility measures.

Indeed, consider a $(L,\leq)$-possibility space $(X_1,\mathcal{R}_1,\Pi_1)$. Let furthermore X_2 be a universe and f a X_1-X_2-mapping. We use f and Π_1 to introduce the following set and mapping:

$$\mathcal{R}_1{}^{(f)} = \{\, B \mid B \in \wp(X_2) \text{ and } f^{-1}(B) \in \mathcal{R}_1 \,\} \text{ and } \Pi_1{}^{(f)} \colon \mathcal{R}_1{}^{(f)} \to L \colon B \mapsto \Pi_1(f^{-1}(B)),$$

where $f^{-1}(B) = \{\, x_1 \mid x_1 \in X_1 \text{ and } f(x_1) \in B \,\}$ is the inverse image of B under the mapping f. Of course, $\mathcal{R}_1{}^{(f)}$ is an ample field on X_2 and $\Pi_1{}^{(f)}$ is a $(L,\leq)$-possibility measure on $(X_2,\mathcal{R}_1{}^{(f)})$, which will henceforth be called the *transformed* $(L,\leq)$*-possibility measure on* $(X_2,\mathcal{R}_1{}^{(f)})$ *of* Π_1 *by* f. Furthermore, if Π_1 is normal, then $\Pi_1{}^{(f)}$ is normal, and *vice versa*. The distribution $\pi_1{}^{(f)}$ of $\Pi_1{}^{(f)}$ satisfies, for any x_2 in X_2:

$$\pi_1{}^{(f)}(x_2) = \sup_{f(x_1)\in[x_2]_{\mathcal{R}_1{}^{(f)}}} \pi_1(x_1), \tag{12}$$

where π_1 is the distribution of Π_1.

This course of reasoning can be taken yet a step further. Let me first remind the reader of a very general definition of measurability of mappings [Jacobs, 1978].

Definition 4.1. Let X_1 and X_2 be universes. Let $\mathcal{A}_1 \subseteq \wp(X_1)$ and $\mathcal{A}_2 \subseteq \wp(X_2)$, and let f be a X_1-X_2-mapping. Then f is called $\mathcal{A}_1-\mathcal{A}_2$-measurable iff $f^{-1}(\mathcal{A}_2) \subseteq \mathcal{A}_1$, or equivalently, $(\forall B \in \mathcal{A}_2)(f^{-1}(B) \in \mathcal{A}_1)$.

This definition plays a fairly important part in the transformation of possibility measures. Indeed, as before, let $(X_1,\mathcal{R}_1,\Pi_1)$ be a $(L,\leq)$-possibility space, let X_2 be a universe and let f be a X_1-X_2-mapping. Then f is obviously $\mathcal{R}_1-\mathcal{R}_1{}^{(f)}$-measurable. Let furthermore $\mathcal{R}_2$ be an ample field on X_2. If f is $\mathcal{R}_1-\mathcal{R}_2$-measurable, then it follows immediately that $\mathcal{R}_2 \subseteq \mathcal{R}_1^{(f)}$. We conclude that $\mathcal{R}_1{}^{(f)}$ is the greatest ample field $\mathcal{R}_2$ on X_2 w.r.t. set inclusion, such that f is still $\mathcal{R}_1-\mathcal{R}_2$-measurable.

We are now ready to generalize the transformation of possibility measures. Consider an ample field $\mathcal{R}_2$ on the universe X_2, and assume that the X_1-X_2-mapping f is $\mathcal{R}_1-\mathcal{R}_2$-measurable. We know that $\Pi_1{}^{(f)}$ is a $(L,\leq)$-possibility measure on $(X_2,\mathcal{R}_1{}^{(f)})$, and that $\mathcal{R}_2 \subseteq \mathcal{R}_1{}^{(f)}$, so that we may consider the restriction $\Pi_1{}^{(f)}|\mathcal{R}_2$. This is of course a $(L,\leq)$-possibility measure on

$(X_2, \mathcal{R}_2)$. We shall call ${\Pi_1}^{(f)}|\mathcal{R}_2$ the *transformed* $(L, \leq)$*-possibility measure on* $(X_2, \mathcal{R}_2)$ *of* Π_1 *by* f. Furthermore, for any B in $\mathcal{R}_2$

$$({\Pi_1}^{(f)}|\mathcal{R}_2)(B) = \Pi_1(f^{-1}(B)), \tag{13}$$

and the distribution π_2 of ${\Pi_1}^{(f)}|\mathcal{R}_2$ satisfies, for any x_2 in X_2

$$\pi_2(x_2) = \sup_{f(x_1) \in [x_2]_{\mathcal{R}_2}} \pi_1(x_1). \tag{14}$$

Clearly, ${\mathcal{R}_1}^{(f)}$ is the greatest ample field w.r.t. set inclusion on which a transformed $(L, \leq)$-possibility measure of Π_1 by f can be defined.

4.2 Possibilistic Variables

A variable may be informally defined as an *abstract object that can assume values in a certain universe*. This notion of a variable is used for instance by Zadeh [1978a], Hisdal [1978] and Nguyen [1978] in the context of fuzzy set theory and possibility theory. In probability theory, this notion appears in the guise of a *stochastic variable* [Burrill, 1972]. In this section, I intend to show how the notion of a *possibilistic variable* can be introduced, and in doing so I provide a generalization and formalization of Zadeh's more intuitive notion of a fuzzy variable.

Let us consider a universe X and a variable ξ in X of the type informally described above. In principle, ξ can assume any value in X. However, it is possible to *restrict* the values which ξ may take in X, for instance by stating that ξ can only take values in a subset D of X. By imposing such a restriction, the uncertainty about the value that ξ assumes is actually reduced. We thus get *more information* about the value which ξ assumes in X.

Such information need not always take the form of a subset D of X. Let us for instance consider the following case: the information about the value ξ takes in X is given in the form of a *probability measure* [Doob, 1953]. More explicitly, we consider a σ-field $\mathcal{S}$ of subsets of X and a probability measure Pr on $(X, \mathcal{S})$. We also assume that for any A in $\mathcal{S}$, $\Pr(A)$ is the probability that ξ takes a value in the subset A of X. It is clear that, by doing so, we impose a restriction on the values that ξ can take in X. This new restriction is more 'elastic' than the one specified by the subset D, and takes the form of a probability measure Pr. In this sense, it could be said that the information we have about the value of ξ in X is *probabilistic*.

Up to now, we have not yet given any formal mathematical definition of the variables we are talking about. It is nevertheless perfectly possible to do this, starting with the more intuitive model described above. In probability theory, this is normally accomplished as follows: first of all, a *basic space* Ω is considered, provided with a σ-field $\mathcal{S}_\Omega$ of subsets of Ω. Then a probability measure $\Pr_\Omega$ is defined on $(\Omega, \mathcal{S}_\Omega)$. A $\mathcal{S}_\Omega - \mathcal{S}$-measurable $\Omega - X$-mapping ξ is called a *stochastic variable* in $(X, \mathcal{S})$. The probability that this stochastic variable takes a value in the element A of $\mathcal{S}$ is then given by $\Pr_\Omega(\xi^{-1}(A))$. The universe X that this variable ξ takes its values in, is called a *sample space*. The relation with the more intuitive treatment given above should be clear. As an example, the above-mentioned probability measure Pr is determined by $\Pr(A) = \Pr_\Omega(\xi^{-1}(A))$, for any A in $\mathcal{S}$.

On the other hand, if the information we have is given by a *possibility measure*, a completely analogous approach can be used. In the intuitive picture of a variable ξ in X, we consider an ample field $\mathcal{R}$ of subsets of X. Furthermore, we assume that on $(X, \mathcal{R})$ there is defined a

$(L, \leq)$-possibility measure Π, in such a way that for any A in $\mathcal{R}$, $\Pi(A)$ is the $(L, \leq)$-possibility that the variable ξ takes a value in the set A. This effectively imposes a restriction on the values that ξ can take in X. Our information about ξ in X is in this case of a *possibilistic nature*.

We may use these more or less intuitive notions, together with the discussion of stochastic variables given above, to give a *formal definition of possibilistic variables*. We consider a *basic space* Ω, provided with an ample field $\mathcal{R}_\Omega$ of subsets of Ω. We also assume the existence of a $(L, \leq)$-possibility measure Π_Ω on $(\Omega, \mathcal{R}_\Omega)$, with distribution π_Ω.

Definition 4.2. A $\Omega - X$-mapping ξ that is $\mathcal{R}_\Omega - \mathcal{R}$-measurable, is called a possibilistic variable in $(X, \mathcal{R})$.

As mentioned before, this formalizes Zadeh's notion of a fuzzy variable. I prefer to use the name '*possibilistic* variable', however, because this notion will turn out to be of central importance in this treatise on possibility theory, and is only indirectly related with fuzzy sets.

Using Π_Ω, the $(L, \leq)$-possibility that the possibilistic variable ξ takes a value in the element A of $\mathcal{R}$ can be expressed as $\Pi_\Omega(\xi^{-1}(A))$. The relation with the more intuitive treatment given above, should once again be clear. In particular, the $(L, \leq)$-possibility measure Π in the intuitive picture can of course be determined by

$$(\forall A \in \mathcal{R})(\Pi(A) = \Pi_\Omega(\xi^{-1}(A))), \tag{15}$$

or equivalently, Π is the transformed $(L, \leq)$-possibility measure of Π_Ω on $(X, \mathcal{R})$ by the mapping ξ. The distribution π of Π is then given by

$$\pi(x) = \sup_{\xi(\omega) \in [x]_\mathcal{R}} \pi_\Omega(\omega), \tag{16}$$

for any x in X.

It should come as no surprise that possibilistic variables play a part in possibility theory that is to a high extent comparable to the one played in probability theory by stochastic variables. For one thing, they allow us to formally study the notions of conditional possibility and possibilistic independence. A detailed account of how this can be done, is given in Parts II and III of this series.

Drawing our inspiration from the treatment of stochastic variables in probability theory [Burrill, 1972], and from Eqs. (15) and (16) above, we introduce possibility distributions and possibility distribution functions for possibilistic variables.

Definition 4.3. Let ξ be a possibilistic variable in $(X, \mathcal{R})$. The $(L, \leq)$-possibility measure $\Pi_\xi = {\Pi_\Omega}^{(\xi)} | \mathcal{R}$ is called the possibility distribution (measure) of ξ. Its distribution $\pi_\xi \colon X \to L$ is called the possibility distribution function of ξ, and satisfies, for any x in X:

$$\pi_\xi(x) = \sup_{\xi(\omega) \in [x]_\mathcal{R}} \pi_\Omega(\omega).$$

Furthermore, for any A in $\mathcal{R}$, $\Pi_\xi(A) = \sup_{x \in A} \pi_\xi(A)$.

Remark that Π_ξ is normal iff Π_Ω is. I follow Zadeh's nomenclature [Zadeh, 1978a] in calling π_ξ the 'possibility distribution function' of the possibilistic variable ξ. On the other hand, I want to reserve the name 'possibility distribution of a possibilistic variable' for the possibility measure Π_ξ. It should also be noted that there exists an important conceptual difference between the *distribution* of a possibility measure, and the *possibility distribution* (measure) of a possibilistic variable[3].

4.3 Transformation of Variables

Let us now turn to a few theorems whose counterparts in classical integration theory fall into the category 'transformation of variables'. In what follows, we denote by $(X_1, \mathcal{R}_1, \Pi_1)$ and $(X_2, \mathcal{R}_2, \Pi_2)$ $(L, \leq)$-possibility spaces. f is a $X_1 - X_2$-mapping. Furthermore, ${\Pi_1}^{(f)}$ is the transformed $(L, \leq)$-possibility measure on $(X_2, {\mathcal{R}_1}^{(f)})$ of Π_1 by f. Whenever in the sequel f is assumed to be $\mathcal{R}_1 - \mathcal{R}_2$-measurable[4], we denote by ${\Pi_1}^{(f)}|\mathcal{R}_2$ the transformed $(L, \leq)$-possibility measure on $(X_2, \mathcal{R}_2)$ of Π_1 by f. The distributions of Π_1, Π_2 and ${\Pi_1}^{(f)}$ are denoted by respectively π_1, π_2 and ${\pi_1}^{(f)}$. For any subset A of X_1, $f(A) = \{ f(x_1) \mid x_1 \in X_2 \}$ is the *direct image* of A under the mapping f.

In Theorem 4.4 and Corollary 4.5 we encounter formulas that are very similar to what in classical integration theory is known as the *integral transport formula* [Burrill, 1972, Theorem 7-8A]. They tell us in precisely what way we must change the integrand and the domain of integration, when we transform a possibility integral associated with the possibility measure Π_1 on $(X_1, \mathcal{R}_1)$ into a possibility integral associated with the transformed possibility measures ${\Pi_1}^{(f)}$ on $(X_2, {\mathcal{R}_1}^{(f)})$ (Theorem 4.4) and ${\Pi_1}^{(f)}|\mathcal{R}_2$ on $(X_2, \mathcal{R}_2)$ (Corollary 4.5). These integral transport formulas will be of use in the next section, and in the discussion of conditional possibility, given in Part II.

Theorem 4.4. *Let P be a t-seminorm on $(L, \leq)$, such that $(L, \leq, P)$ is a complete lattice with t-seminorm. Let h be a $(L, \leq)$-fuzzy variable in $(X_2, {\mathcal{R}_1}^{(f)})$. Then for any E in ${\mathcal{R}_1}^{(f)}$:*

$$(P)\int_E h \mathrm{d}{\Pi_1}^{(f)} = (P)\int_{f^{-1}(E)} (h \circ f) \mathrm{d}\Pi_1.$$

Proof. Consider an arbitrary element E of ${\mathcal{R}_1}^{(f)}$. Since h is ${\mathcal{R}_1}^{(f)}$-measurable, we have on the one hand that, taking into account Eqs. (5) and (12),

$$\begin{aligned}
(P)\int_E h \mathrm{d}{\Pi_1}^{(f)} &= \sup_{y \in E} P(h(y), {\pi_1}^{(f)}(y)) \\
&= \sup_{y \in E} P(h(y), \sup_{x \in f^{-1}([y]_{{\mathcal{R}_1}^{(f)}})} \pi_1(x)) \\
&= \sup_{y \in E} \sup_{f(x) \in [y]_{{\mathcal{R}_1}^{(f)}}} P(h(y), \pi_1(x)) \\
&= \sup_{y \in E} \sup_{x \in f^{-1}([y]_{{\mathcal{R}_1}^{(f)}})} P(h(f(x)), \pi_1(x)).
\end{aligned}$$

[3]Of course, any distribution can be trivially considered as the possibility distribution function of an identical permutation.

[4]In classical integration theory an analogous mapping is called a *measurable transformation*.

Since $E \in \mathcal{R}_1{}^{(f)}$, we find, taking into account Eq. (1) and the properties of the inverse image of a mapping,

$$
\begin{aligned}
&= \sup_{x \in f^{-1}(E)} P((h \circ f)(x), \pi_1(x)) \\
&= (P)\int_{f^{-1}(E)} (h \circ f)\mathrm{d}\Pi_1,
\end{aligned}
$$

again taking into account Eq. (5), since it easily verified that $h \circ f$ is $\mathcal{R}_1$-measurable. □

Corollary 4.5. *Let P be a t-seminorm on $(L, \leq)$, such that $(L, \leq, P)$ is a complete lattice with t-seminorm. Let the $X_1 - X_2$-mapping f be $\mathcal{R}_1 - \mathcal{R}_2$-measurable. Let furthermore h be a $(L, \leq)$-fuzzy variable in $(X_2, \mathcal{R}_2)$. Then for any E in $\mathcal{R}_2$:*

$$(P)\int_E h\mathrm{d}(\Pi_1{}^{(f)}|\mathcal{R}_2) = (P)\int_{f^{-1}(E)} (h \circ f)\mathrm{d}\Pi_1.$$

Proof. Consider an arbitrary E in $\mathcal{R}_2$. Since f is $\mathcal{R}_1 - \mathcal{R}_2$-measurable, we know that $\mathcal{R}_2 \subseteq \mathcal{R}_1{}^{(f)}$. Therefore, h is $\mathcal{R}_1{}^{(f)}$-measurable and $E \in \mathcal{R}_1{}^{(f)}$. This implies that we may use Theorem 4.4. Also taking into account Eqs. (1) and (5), and the $\mathcal{R}_2$- and $\mathcal{R}_1{}^{(f)}$-measurability of h, we find that

$$
\begin{aligned}
(P)\int_{f^{-1}(E)} (h \circ f)\mathrm{d}\Pi_1 &= (P)\int_E h\mathrm{d}\Pi_1{}^{(f)} \\
&= \sup_{x_2 \in E} P(h(x_2), \Pi_1{}^{(f)}([\, x_2 \,]_{\mathcal{R}_1{}^{(f)}})) \\
&= \sup_{y_2 \in E} \sup_{x_2 \in [\, y_2 \,]_{\mathcal{R}_2}} P(h(x_2), \Pi_1{}^{(f)}([\, x_2 \,]_{\mathcal{R}_1{}^{(f)}})). \\
&= \sup_{y_2 \in E} \sup_{x_2 \in [\, y_2 \,]_{\mathcal{R}_2}} P(h(y_2), \Pi_1{}^{(f)}([\, x_2 \,]_{\mathcal{R}_1{}^{(f)}})) \\
&= \sup_{y_2 \in E} P(h(y_2), \sup_{x_2 \in [\, y_2 \,]_{\mathcal{R}_2}} \Pi_1{}^{(f)}([\, x_2 \,]_{\mathcal{R}_1{}^{(f)}})) \\
&= \sup_{y_2 \in E} P(h(y_2), \Pi_1{}^{(f)}([\, y_2 \,]_{\mathcal{R}_2})) \\
&= (P)\int_E h\mathrm{d}(\Pi_1{}^{(f)}|\mathcal{R}_2). \quad \square
\end{aligned}
$$

The following rather remarkable result illustrates that the analogy between classical integration theory and possibility integral theory can be carried very far. Theorem 4.6 is indeed a perfect analogon of the *theorem of the Jacobian* for Lebesgue integrals [Burrill, 1972, Theorem 7-8B]. It should be noted that the mapping g in this theorem plays the role of the *Jacobian* of the 'measurable transformation' f. The t-norm T in this theorem has a similar function as the product operator in the classical theory.

Theorem 4.6. *Let T be a t-norm on $(L, \leq)$, such that $(L, \leq, T)$ is a complete lattice with t-norm. Let the $X_1 - X_2$-mapping f be $\mathcal{R}_1 - \mathcal{R}_2$-measurable. If*

(i) $(\forall E \in \mathcal{R}_2)(\Pi_1^{(f)}(E) \leq \Pi_2(E))$,

(ii) T is weakly invertible,

then there exists a g in $\mathcal{G}^{\mathcal{R}_2}_{(L,\leq)}(X_2)$, such that for any h in $\mathcal{G}^{\mathcal{R}_2}_{(L,\leq)}(X_2)$ and for any E in $\mathcal{R}_2$:

$$(T)\int_{f^{-1}(E)}(h\circ f)\mathrm{d}\Pi_1 = (T)\int_E T\circ(h,g)\mathrm{d}\Pi_2.$$

Proof. Consider an arbitrary h in $\mathcal{G}^{\mathcal{R}_2}_{(L,\leq)}(X_2)$ and any E in $\mathcal{R}_2$. Since h is $\mathcal{R}_2$-measurable, we may write, taking into account Corollary 4.5 and Eq. (5),

$$(T)\int_{f^{-1}(E)}(h\circ f)\mathrm{d}\Pi_1 = (T)\int_E h\mathrm{d}({\Pi_1}^{(f)}|\mathcal{R}_2) = \sup_{y\in E} T(h(y), {\Pi_1}^{(f)}([\,y\,]_{\mathcal{R}_2})),$$

On the other hand, we immediately find for any g in $\mathcal{G}^{\mathcal{R}_2}_{(L,\leq)}(X_2)$ that, taking into account $T\circ(h,g)\in\mathcal{G}^{\mathcal{R}_2}_{(L,\leq)}(X_2)$, Eq. (5) and the associativity of the t-norm T,

$$(T)\int_E T\circ(h,g)\mathrm{d}\Pi_2 = \sup_{y\in E} T(T(h(y),g(y)),\pi_2(y)) = \sup_{y\in E} T(h(y),T(g(y),\pi_2(y))).$$

It therefore remains to be proven that there exists a $\mathcal{R}_2$-measurable X_2-L-mapping g such that

$$(\forall x_2\in X_2)({\Pi_1}^{(f)}([\,x_2\,]_{\mathcal{R}_2}) = T(g(x_2),\pi_2(x_2)). \qquad (17)$$

Taking into account assumption (i), we have in particular that ${\Pi_1}^{(f)}([\,x_2\,]_{\mathcal{R}_2})\leq\pi_2(x_2)$, for any x_2 in X_2. Assumption (ii) and Proposition 2.2 now guarantee that the X_2-L-mapping g, defined as $g(x_2)={\Pi_1}^{(f)}([\,x_2\,]_{\mathcal{R}_2})\triangle_T\pi_2(x_2)$, for any x_2 in X_2, satisfies Eq. (17). This g is at the same time clearly $\mathcal{R}_2$-measurable. □

5 FUZZY VARIABLES

In section 2 I introduced fuzzy variables as measurable *fuzzy* sets, i.e., fuzzifications of measurable sets (or events). Further on, I shall show that these fuzzy variables are special instances of the possibilistic *variables*, introduced and studied in the previous section. This, in my opinion, justifies my using the name '*fuzzy variables*'.

As mentioned before, it should on the other hand be noted that Zadeh, in his seminal paper about possibility theory [Zadeh, 1978a], uses the name 'fuzzy variable' more generally for what I have been calling here a possibilistic variable[5]. Since, however, these variables are the possibilistic counterparts of the stochastic variables in probability theory, I prefer to call them *possibilistic*, and want to reserve the qualification 'fuzzy' for those variables *that are also fuzzy sets.*

In probability theory, a very prominent class of variables are the *real* stochastic variables, for which a large body of results have been derived [Burrill, 1972]: using Lebesgue integration theory, it is possible for these real variables to introduce mean values or expectations, higher order moments, characteristic functions, etc. If we take a closer look, we see that a deeper reason

[5]Following Zadeh, Nahmias [1978] defines fuzzy variables as real-valued mappings. Wang [1982] gives a similar definition for his fuzzy variables, but also imposes a measurability condition. Ralescu [1982] defines fuzzy variables as real stochastic variables. It should be noted that Nahmias and Wang introduce their fuzzy variables in a possibilistic context, although both of them use a different name (scale, fuzzy contactability) for the possibility measures they work with. Ralescu does not work with possibility measures, but with Sugeno's fuzzy measures.

for this diversity of results is the following: real stochastic variables assume values in the set of the reals, and so do the probability measures used to describe their behaviour. This makes it possible for those variables to appear in the integrand of the Lebesgue integral associated with these measures.

If we want to look for a possibilistic formal analogon of real stochastic variables, it is clear that we must look for those possibilistic variables that take values in essentially the same set as the possibility measures used to describe their behaviour. Only then shall we be able to make these variables appear in the integrand of possibility integrals associated with these possibility measures, and closely follow the analogy with probability theory. That is the most important reason why I single out fuzzy variables here, why I have given them a special name, and dedicate this section to their study.

In what follows, we denote by $(X, \mathcal{R}, \Pi)$ a $(L, \leq)$-possibility space; the distribution of Π is denoted by π. Furthermore, Ω is a universe, $\mathcal{S}_\Omega$ a σ-field on Ω and Pr_Ω a probability measure on $(\Omega, \mathcal{S}_\Omega)$. For a more detailed treatment of the probability theoretic material that follows, I refer to [Burrill, 1972] and [Doob, 1953].

For a start, let us find out what the formal analogy between real stochastic variables and fuzzy variables really consists in. Consider the universe Ω, provided with the σ-field $\mathcal{S}_\Omega$ and the probability measure Pr_Ω as a *basic space*. Furthermore, consider a real stochastic variable ξ, i.e., ξ is a $\mathcal{S}_\Omega - \sigma(T_\mathbb{R})$-measurable $\Omega - \mathbb{R}$-mapping, or equivalently, a stochastic variable in $(\mathbb{R}, \sigma(T_\mathbb{R}))$. $T_\mathbb{R}$ is of course the metric topology on the reals, and $\sigma(T_\mathbb{R})$ the σ-algebra generated by this topology, also called the Borel algebra on that set. This real stochastic variable ξ has a *probability distribution function* F_ξ, defined as:

$$F_\xi \colon \mathbb{R} \to [0,1] \colon x \mapsto \mathrm{Pr}_\Omega(\xi^{-1}(]-\infty, x])).$$

It can be shown that with this F_ξ there is associated a *unique* probability measure Pr_ξ on $(\mathbb{R}, \sigma(T_\mathbb{R}))$, satisfying $(\forall x \in \mathbb{R})(\mathrm{Pr}_\xi(]-\infty, x]) = F_\xi(x))$. The *completion* of this probability measure is precisely the *Lebesgue-Stieltjes measure* m_{F_ξ} *induced by* F_ξ (or by ξ). It can also be shown that for any element B of $\sigma(T_\mathbb{R})$:

$$\mathrm{Pr}_\xi(B) = \int_B \mathrm{d}F_\xi, \tag{18}$$

where the integral on the right hand side of the equation is the Lebesgue-Stieltjes integral associated with the distribution function F_ξ. Furthermore, it can be proven that the probability measure Pr_ξ is precisely the transformed probability measure on $(\mathbb{R}, \sigma(T_\mathbb{R}))$ of Pr_Ω by ξ, which really means that $(\forall B \in \sigma(T_\mathbb{R}))(\mathrm{Pr}_\xi(B) = \mathrm{Pr}_\Omega(\xi^{-1}(B)))$. There are other ways of determining the probability measure Pr_ξ, which are however *equivalent* with the specification of the probability distribution function F_ξ . When the real stochastic variable ξ is *discrete*, it is possible to define a $\mathbb{R} - [0,1]$-mapping f_ξ by

$$(\forall x \in \mathbb{R}) \left(f_\xi(x) = \mathrm{Pr}_\Omega(\xi^{-1}\{x\}) = F_\xi(x) - \lim_{t \overset{<}{\to} x} F_\xi(t) \right).$$

f_ξ is called the *frequency function* of ξ. When ξ is a *continuous* real stochastic variable, there exists a $\mathbb{R} - \mathbb{R}$-mapping f_ξ satisfying

$$(\forall x \in \mathbb{R}) \left(F_\xi(x) = \int_{-\infty}^{x} f_\xi(t) \mathrm{d}t \right)$$

which is called the *(probability) density function* of ξ.

The ultimate aim of these functions is of course to convert the probabilistic information, embodied in $(\Omega, \mathcal{S}_\Omega, \Pr_\Omega)$, into the probabilistic information $(\mathbb{R}, \sigma(T_\mathbb{R}), \Pr_\xi)$ about the values that the real stochastic variable ξ can take in $\mathbb{R}$. The latter form of information is the starting point for further derivations, such as the calculation of the mean, the higher order moments and the characteristic function of the real stochastic variable ξ.

For fuzzy variables, it is possible to follow an analogous course of reasoning. Let us consider a universe X, provided with an ample field $\mathcal{R}$ and the $(L, \leq)$-possibility measure Π as a *basic space*. It is easily verified that a $(L, \leq)$-fuzzy variable h in $(X, \mathcal{R})$ is a $\mathcal{R} - \wp(L)$-measurable $X - L$-mapping, which can therefore be formally considered as a *possibilistic variable in* $(L, \wp(L))$. The set L is in this case interpreted as a sample space, and is provided with the ample field of measurable subsets of[6] $\wp(L)$. Drawing our inspiration from the probabilistic discussion above, we can now try and convert the possibilistic information, contained in the $(L, \leq)$-possibility space $(X, \mathcal{R}, \Pi)$, into possibilistic information about the value that h assumes in L.

In probability theory, using distribution functions and Lebesgue-Stieltjes integrals to calculate transformed probability measures is a standard procedure, and in many cases by far the most convenient way of obtaining these measures. On the other hand, it turns out that in possibility theory a more direct approach is ideally suited: we shall use the plain, brute-force transformation of possibility measures as the appropriate way of transmitting possibilistic information. In Proposition 5.1 and Definition 5.2 it is explained in more detail how this comes about. In order to keep this discussion as general as possible, we shall work from the outset with multidimensional fuzzy variables, of which the ordinary, one-dimensional fuzzy variables are a special case. The proof of Proposition 5.1 is straightforward, and can therefore be omitted.

Proposition 5.1. *Let n be an element of $\mathbb{N}^*$ and let $h = (h_1, \ldots, h_n)$ be a n-dimensional $(L, \leq)$-fuzzy variable in $(X, \mathcal{R})$. Since h is a $X - L^n$-mapping, we can use h to transform Π into a $(L, \leq)$-possibility measure $\Pi^{(h)}$ on $(L^n, \mathcal{R}^{(h)})$. Furthermore, $\mathcal{R}^{(h)} = \wp(L^n)$, and the $(L, \leq)$-possibility measure $\Pi^{(h)}$ on $(L^n, \wp(L^n))$ has as its distribution the $L^n - L$-mapping $\pi^{(h)}$, defined by $\pi^{(h)}(\lambda) = \Pi(h^{-1}(\{\lambda\}))$, $\lambda \in L^n$. For $\Pi^{(h)}$ itself, we have that $\Pi^{(h)}(B) = \Pi(h^{-1}(B))$, $B \in \wp(L^n)$. $\Pi^{(h)}$ is normal iff Π is.*

Definition 5.2. Let n be an element of $\mathbb{N}^*$ and let $h = (h_1, \ldots, h_n)$ be a n-dimensional $(L, \leq)$-fuzzy variable in $(X, \mathcal{R})$. The $L^n - L$-mapping $\gamma_h = \pi^{(h)}$ is called the possibility distribution function of h. The $\wp(L^n) - L^n$-mapping Γ_h is the unique $(L, \leq)$-possibility measure on $(L^n, \wp(L^n))$ that has γ_h as its distribution, i.e. $\Gamma_h = \Pi^{(h)}$. We call Γ_h the possibility distribution (measure) of h. Remark that Γ_h is normal iff Π is.

Of course, fuzzy variables are special possibilistic variables, and the possibility distributions and distribution functions of the former are special instances of those of the latter.

It should be noted that the definition of a possibility distribution function and that of its probabilistic counterpart are somewhat different to the letter: in the definition of a probability distribution function, the real intervals $]-\infty, x]$, $x \in \mathbb{R}$, play a prominent part; in the definition of a possibility distribution function, this role is played by the atoms $\{\lambda\}$, $\lambda \in L^n$, of the ample field $\wp(L^n)$. This means that possibility distribution functions bear a closer formal resemblance to the *density* and *frequency functions* in probability theory. Despite this difference, both distribution

[6] Remark also that $\wp(L) = \tau(T_L)$, where T_L is the order topology on the complete lattice $(L, \leq)$ [Birkhoff, 1967]. This makes the analogy between real stochastic variables and fuzzy variables even more complete.

functions are very similar *in spirit*: the probability distribution function F_ξ uniquely determines the probability measure $\Pr_\xi$; the possibility distribution function γ_h characterizes the $(L,\leq)$-possibility measure Γ_h. The name '*possibility distribution function*' for the $L^n - L$-mapping γ_h is not just plucked out of the air: this function (mapping) tells us how the possibility is distributed over the elements of L^n. A second reason for my terminology is that Zadeh[7] [1978a] uses it in a comparable context (see also Definition 4.3). A third reason is that the terms 'density' and 'frequency' do not belong in a theory of possibility. 'Frequency' is a statistical term, and is therefore tightly connected with the probabilistic paradigm. The term 'density' has originated in mechanics: probability densities draw their inspiration from mass densities, which are essentially mass-volume ratios, and therefore alien to this theory of possibility.

Of course, when we want to look for analogous theorems between probability and possibility theory, these will primarily be found between possibility distribution functions on the one hand, and density and distribution functions on the other hand, due to the formal resemblance between these notions. To a lesser extent, we can of course also expect similarities between results about possibility and probability distribution functions. To give a few examples, let us start from the definition above and deduce a number of interesting results, which are clearly possibilistic counterparts of well-known theorems in probability theory. See, for instance, Eq. (18) and [Burrill, 1972, Theorem 11-1A, section 11-1 and Theorem 11-2A with corollary] respectively. The proofs of Propositions 5.3 and 5.4, and of Corollary 5.6 are immediate, and therefore omitted.

Proposition 5.3. *Let P be a t-seminorm on $(L,\leq)$ such that $(L,\leq,P)$ is a complete lattice with t-seminorm. Let furthermore n be an element of $\mathbb{N}^*$ and let $h=(h_1,\dots,h_n)$ be a n-dimensional $(L,\leq)$-fuzzy variable in $(X,\mathcal{R})$. Then for any B in $\wp(L^n)$:*

$$\Pi(h^{-1}(B)) = (P)\int_B \mathrm{d}\Gamma_h.$$

Proposition 5.4. *Let P be a t-seminorm on $(L,\leq)$ such that $(L,\leq,P)$ is a complete lattice with t-seminorm. Let m and n be elements of $\mathbb{N}^*$. Let furthermore $h=(h_1,\dots,h_n)$ be a n-dimensional $(L,\leq)$-fuzzy variable in $(X,\mathcal{R})$ and let $g=(g_1,\dots,g_m)$ be a L^n-L^m-mapping. Then $g\circ h$, or $g(h)$, is a m-dimensional $(L,\leq)$-fuzzy variable in $(X,\mathcal{R})$, and for any λ in L^m and $B\subseteq L^m$:*

$$\gamma_{g(h)}(\lambda) = (P)\int_{g^{-1}(\{\lambda\})} \mathrm{d}\Gamma_h \quad \text{and} \quad \Gamma_{g(h)}(B) = (P)\int_{g^{-1}(B)} \mathrm{d}\Gamma_h.$$

Theorem 5.5. *Let P be a t-seminorm on $(L,\leq)$, such that $(L,\leq,P)$ is a complete lattice with t-seminorm. Let m and n be elements of $\mathbb{N}^*$. Let furthermore $h=(h_1,\dots,h_n)$ be a n-dimensional $(L,\leq)$-fuzzy variable in $(X,\mathcal{R})$ and let $g=(g_1,\dots,g_m)$ be a L^n-L^m-mapping. Then*

$$\Pi_P(g(h)) = (P)\int_{L^n} g\,\mathrm{d}\Gamma_h.$$

Proof. For a start, it should be noted that the formula above is a shorthand for

$$(\forall \ell\in\{1,\dots,m\})\left(\Pi_P(g_\ell(h)) = (P)\int_{L^n} g_\ell\,\mathrm{d}\Gamma_h\right).$$

[7] Nahmias [1978] gives a similar definition for his fuzzy variables, but uses the term '*membership function*' instead of 'possibility distribution function'. Wang [1982] on the other hand generalizes Nahmias definition of a fuzzy variable, but uses still other names: *fuzzy density* and *fuzzy distribution*.

Let therefore ℓ be an arbitrary natural number with $1 \leq \ell \leq m$. Then it must be shown that

$$(P)\int_X (g_\ell \circ h)\mathrm{d}\Pi = (P)\int_{L^n} g_\ell \mathrm{d}\Gamma_h.$$

This equality immediately follows from Theorem 4.4, by taking into account the following correspondences: $X_1 \to X$, $X_2 \to L^n$, $f \to h$, $\mathcal{R}_1 \to \mathcal{R}$, $\mathcal{R}_1{}^{(f)} \to \mathcal{R}^{(h)} = \wp(L^n)$, $\Pi_1 \to \Pi$, $\Pi_1{}^{(f)} \to \Pi^{(h)} = \Gamma_h$, $h \to g_\ell$, $E \to L^n$ and $f^{-1}(E) \to h^{-1}(L^n) = X$. □

Corollary 5.6. *Let P be a t-seminorm on $(L, \leq)$, such that $(L, \leq, P)$ is a complete lattice with t-norm. Let furthermore h be an arbitrary $(L, \leq)$-fuzzy variable in $(X, \mathcal{R})$. Then*

$$\Pi_P(h) = (P)\int_L \mathrm{id}_L \mathrm{d}\Gamma_h = \sup_{\lambda \in L} P(\lambda, \gamma_h(\lambda)),$$

where id_L *is the identical permutation of* L.

Let me conclude this section with a short discussion of the interpretation and the significance of these results. Although the formal analogy between real stochastic variables and fuzzy variables is apparent, there is a notable difference between their respective interpretations. Even though a fuzzy variable can be formally considered as a variable that takes values in L, we have been interpreting it as a measurable fuzzy set, the 'fuzzification' of the notion of event, or measurable set. Whereas the Lebesgue integral in probability theory is used to define the mean value or expectation of a real stochastic variable, here the possibility integral is used to extend the notion of possibility, and define the possibility of fuzzy events, as is explained in section 3. Similarly, as I shall presently argue, the results in this section have an interpretation that differs from their formal counterparts in probability theory.

In Zadeh's fuzzy set theory, it is possible to change the meaning of a $([0,1], \leq)$-fuzzy set[8] by taking its composition with transformations of the real unit interval $[0,1]$, called *linguistic hedges* [Zadeh, 1975a, 1975b, 1976, 1978b]. We can illustrate Zadeh's course of reasoning more or less as follows. Let h_{tall} be the $([0,1], \leq)$-fuzzy set in $\mathbb{R}$, associated with the property 'tall'. Also consider the following transformations of $[0,1]$:

$$g_{very}\colon [0,1] \to [0,1]\colon x \mapsto x^2$$
$$g_{more\ or\ less}\colon [0,1] \to [0,1]\colon x \mapsto \sqrt{x}$$
$$g_{not}\colon [0,1] \to [0,1]\colon x \mapsto 1-x.$$

Then, still according to Zadeh, $g_{very} \circ h_{\text{tall}}$ is the $([0,1], \leq)$-fuzzy set in $\mathbb{R}$ associated with the property '*very* tall'; $g_{more\ or\ less} \circ h_{\text{tall}}$ is the $([0,1], \leq)$-fuzzy set in $\mathbb{R}$ associated with the property '*more or less* tall' and $g_{not} \circ h_{\text{tall}}$ is the $([0,1], \leq)$-fuzzy set in $\mathbb{R}$ associated with the property '*not* tall'. The $L^n - L^m$-mapping g in the results in this section can be considered as a *generalization* of these linguistic hedges. The mapping g enables us to convert n $(L, \leq)$-fuzzy sets into m new $(L, \leq)$-fuzzy sets. Possible choices for g are for instance [De Cooman, 1993], [De Cooman and Kerre, 1994]:

- $m = n = 1$;
  g is a negation operator on $(L, \leq)$, which can be used to define a pointwise complement operator for $(L, \leq)$-fuzzy sets.

[8] Remark that I consistently use the terminology introduced in section 2. In the language of this paper, the *membership function* of a Zadeh fuzzy set is a $([0,1], \leq)$-fuzzy set.

- $m = 1, n = 2$;
  g is a t-norm on $(L, \leq)$, which can be used to introduce a pointwise intersection operator for $(L, \leq)$-fuzzy sets.

- $m = 1, n = 2$;
  g is a t-conorm on $(L, \leq)$, which can be used to introduce a pointwise union operator for $(L, \leq)$-fuzzy sets.

Whereas its probabilistic counterpart is used for the calculation of the moments and the characteristic functions of real stochastic variables, starting from their probability distribution function, Theorem 5.5 enables us to calculate the possibility of a combination of fuzzy variables, using their possibility distribution function. As a special case, Corollary 5.6 can be used to calculate the possibility of a fuzzy variable from its possibility distribution function. Proposition 5.4 tells us how the possibility distribution function of a combination of fuzzy variables can be expressed in terms of the possibility distribution functions of these fuzzy variables themselves.

In summary, these results tell us that if we want to work with fuzzy variables, their possibility distributions functions (or measures) contain all the information we need, i.e., it is possible to perform all the calculations with these functions, without having to go back to the possibility measure Π defined on $(X, \mathcal{R})$.

6 THE NOTION 'ALMOST EVERYWHERE'

In this section, I introduce the notions of *almost everywhere equality* and *almost everywhere dominance* of fuzzy variables. Similar notions for real-valued functions play an important part in the classical theory of measure and integration [Burrill, 1972]. It will appear from the propositions below that in order to make these notions as useful in possibility theory as they are in classical measure theory, we *cannot* adopt definitions that are immediate extensions of the classical ones. As it turns out, we need more specific definitions, of which such immediate extensions turn out to be generalizations. The notions discussed here are of crucial importance for the discussion of conditional possibility and possibilistic independence, given in Parts II and III.

In what follows, we consider a $(L, \leq)$-possibility space $(X, \mathcal{R}, \Pi)$ and denote the distribution of Π by π.

Definition 6.1. Let P be a t-seminorm on $(L, \leq)$. For any E in $\mathcal{R}$, we define the following binary relations on $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$. For h_1 and h_2 in $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$:

(i) $h_1 \overset{(\Pi,P,E)}{\leq} h_2 \Leftrightarrow (\forall x \in E)(P(h_1(x), \pi(x)) \leq P(h_2(x), \pi(x)))$.
When $h_1 \overset{(\Pi,P,E)}{\leq} h_2$, we say that h_2 (Π, P)-dominates h_1 almost everywhere on E.

(ii) $h_1 \overset{(\Pi,P,E)}{=} h_2 \Leftrightarrow (\forall x \in E)(P(h_1(x), \pi(x)) = P(h_2(x), \pi(x)))$.
When $h_1 \overset{(\Pi,P,E)}{=} h_2$, we say that h_1 and h_2 are (Π, P)-equal almost everywhere on E.

When $h_1 \overset{(\Pi,P,X)}{\leq} h_2$, we also write $h_1 \overset{(\Pi,P)}{\leq} h_2$ and say that h_2 (Π, P)-dominates h_1 almost everywhere. When $h_1 \overset{(\Pi,P,X)}{=} h_2$, we also write $h_1 \overset{(\Pi,P)}{=} h_2$ and say that h_1 and h_2 are almost everywhere (Π, P)-equal, or equivalently, that they are (Π, P)-equivalent.

On any complete lattice $(L, \leq)$, there always exist at least two t-seminorms [De Cooman and Kerre, 1994]. This means that this definition is generally applicable, and is never vacuous. Just like their classical counterparts, the relations introduced in Definition 6.1 satisfy some interesting properties. These are gathered in Proposition 6.2, the proof of which is trivial, and therefore omitted.

Proposition 6.2. *Let P be a t-seminorm on $(L, \leq)$. For any E in $\mathcal{R}$:*

(i) $\overset{(\Pi,P,E)}{\leq}$ *is a partial preorder relation on* $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$*;*

(ii) $\overset{(\Pi,P,E)}{=}$ *is an equivalence relation on* $\mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)$*;*

(iii) $\overset{(\Pi,P,E)}{=}$ *is the equivalence relation associated with* $\overset{(\Pi,P,E)}{\leq}$*, i.e.,*

$$(\forall (h_1, h_2) \in \mathcal{G}^{\mathcal{R}}_{(L,\leq)}(X)^2)((h_1 \overset{(\Pi,P,E)}{\leq} h_2 \text{ and } h_2 \overset{(\Pi,P,E)}{\leq} h_1) \Leftrightarrow h_1 \overset{(\Pi,P,E)}{=} h_2).$$

In the following propositions we show that some well-known integral-theoretic results from classical measure theory have analogous counterparts in possibility theory. Among these, Proposition 6.4 is important, because it provides a characterization of the notions of almost everywhere dominance and almost everywhere equality in terms of possibility integrals. It is precisely this result which led me to consider Definition 6.1 instead of a definition that immediately extends the classical approach. Indeed, for such a straightforward extension no characterization in terms of possibility integrals has been found. The proof of Proposition 6.3 is trivial, taking into account Eq. (5).

Proposition 6.3. *Let P be a t-seminorm on $(L, \leq)$, such that $(L, \leq, P)$ is a complete lattice with t-seminorm. Let E be a $\mathcal{R}$-measurable set and let h_1 and h_2 be $(L, \leq)$-fuzzy variables in $(X, \mathcal{R})$.*

(i) $h_1 \overset{(\Pi,P,E)}{\leq} h_2 \Rightarrow (P)\!\!\displaystyle\int\!\!\!\!\!-_E h_1 \mathrm{d}\Pi \leq (P)\!\!\int\!\!\!\!\!-_E h_2 \mathrm{d}\Pi.$

(ii) $h_1 \overset{(\Pi,P,E)}{=} h_2 \Rightarrow (P)\!\!\displaystyle\int\!\!\!\!\!-_E h_1 \mathrm{d}\Pi = (P)\!\!\int\!\!\!\!\!-_E h_2 \mathrm{d}\Pi.$

As mentioned before, Proposition 6.4 gives an alternative characterization of the relations introduced in Definition 6.1. It also gives a characterization of the (Π, P)-equivalence of fuzzy variables that will play an important role in the discussion of conditional possibility in Part II. The proof of Corollary 6.5 is immediate, and can therefore be omitted.

Proposition 6.4. *Let P be a t-seminorm on $(L, \leq)$, such that $(L, \leq, P)$ is a complete lattice with t-seminorm. Let E be a $\mathcal{R}$-measurable set and let h_1 and h_2 be $(L, \leq)$-fuzzy variables in $(X, \mathcal{R})$.*

(i) $h_1 \overset{(\Pi,P,E)}{\leq} h_2$ *iff for any A in $\mathcal{R}$*

$$A \subseteq E \Rightarrow (P)\!\!\int\!\!\!\!\!-_A h_1 \mathrm{d}\Pi \leq (P)\!\!\int\!\!\!\!\!-_A h_2 \mathrm{d}\Pi. \tag{19}$$

(ii) $h_1 \overset{(\Pi,P,E)}{=} h_2$ *iff for any* A *in* $\mathcal{R}$

$$A \subseteq E \Rightarrow (P)\!\!\int_A h_1 \mathrm{d}\Pi = (P)\!\!\int_A h_2 \mathrm{d}\Pi. \tag{20}$$

(iii) h_1 *and* h_2 *are* (Π, P)*-equivalent iff for any* A *in* $\mathcal{R}$

$$(P)\!\!\int_A h_1 \mathrm{d}\Pi = (P)\!\!\int_A h_2 \mathrm{d}\Pi. \tag{21}$$

Proof. It suffices to prove (i), since (ii) and (iii) immediately follow from (i), taking into account Proposition 6.2(iii). If $E = \emptyset$, the equivalence is trivially satisfied. Let us therefore assume that $E \neq \emptyset$. Assume on the one hand that $h_1 \overset{(\Pi,P,E)}{\leq} h_2$. It immediately follows that for any A in $\mathcal{R}$, with $A \subseteq E$, $\sup_{x\in A} P(h_1(x), \pi(x)) \leq \sup_{x\in A} P(h_2(x), \pi(x))$, whence (19), taking into account Eq. (5).

Assume on the other hand that (19) holds for any A in $\mathcal{R}$. Let x be an arbitrary element of E and let in (19) A be equal to $[\,x\,]_{\mathcal{R}}$. Since, by definition, $[\,x\,]_{\mathcal{R}} \subseteq E$, it follows from the assumption that

$$(P)\!\!\int_{[\,x\,]_{\mathcal{R}}} h_1 \mathrm{d}\Pi \leq (P)\!\!\int_{[\,x\,]_{\mathcal{R}}} h_2 \mathrm{d}\Pi.$$

Using Eq. (5) and the $\mathcal{R}$-measurability of π, h_1 and h_2, we deduce from this inequality that $P(h_1(x), \pi(x)) \leq P(h_2(x), \pi(x))$. We may therefore conclude that $h_1 \overset{(\Pi,P,E)}{\leq} h_2$. □

Corollary 6.5. *Let* P *be a t-seminorm on* $(L, \leq)$*, such that* $(L, \leq, P)$ *is a complete lattice with t-seminorm. Let* h *be a* $(L, \leq)$*-fuzzy variable in* $(X, \mathcal{R})$ *and* E *an element of* $\mathcal{R}$*. Then*

$$h \overset{(\Pi,P,E)}{=} \underline{0_L} \Leftrightarrow (P)\!\!\int_E h \mathrm{d}\Pi = 0_L \quad \textit{and} \quad h \overset{(\Pi,P,E)}{=} \underline{1_L} \Leftrightarrow (\forall A \in \mathcal{R}) \Big(A \subseteq E \Rightarrow (P)\!\!\int_A h \mathrm{d}\Pi = \Pi(A) \Big).$$

In the rest of this subsection, I intend to further explore the relation between my definition of almost everywhere equality and dominance, and the immediate extensions of the definitions in classical measure theory. Consider two $(L, \leq)$-fuzzy variables h_1 and h_2 in $(X, \mathcal{R})$ and an arbitrary element E of $\mathcal{R}$. Also consider the set $A = \{\, x \mid x \in E \text{ and } h_1(x) \not\leq h_2(x) \,\}$. It is easily proven that this set is $\mathcal{R}$-measurable. If, for instance, we had extended the classical definition of almost everywhere dominance towards fuzzy variables, we would have found as a defining condition for the almost everywhere dominance on E of h_1 by h_2:

$$\Pi(\{\, x \mid x \in E \text{ and } h_1(x) \not\leq h_2(x) \,\}) = 0_L. \tag{22}$$

Let us find out how this relates to my definition. First, assume that $E \neq \emptyset$. Consider an arbitrary x in E. Either x belongs to A, and then of course $\pi(x) = 0_L$, or x belongs to $E \setminus A$, and then $h_1(x) \leq h_2(x)$. In either case, we find that $P(h_1(x), \pi(x)) \leq P(h_2(x), \pi(x))$, whence $h_1 \overset{(\Pi,P,E)}{\leq} h_2$. If $E = \emptyset$, the same conclusion is trivially reached. Thus it appears that my definition is implied by an immediate extension of the classical definition. However, it is easily verified that the reverse implication is not necessarily valid. We conclude from this that an immediate extension of the classical definition of almost everywhere dominance would lead to a definition that is *less specific* than mine. Of course, for almost everywhere equality, similar results can be obtained.

Let us now take one further step and assume that the t-seminorm P on $(L, \leq)$ satisfies the following property:

$$(\forall(\lambda_1, \lambda_2) \in L^2)(\forall \mu \in L)(P(\lambda_1, \mu) \leq P(\lambda_2, \mu) \Rightarrow (\mu = 0_L \text{ or } \lambda_1 \leq \lambda_2)). \tag{23}$$

In this case, P is called *strongly resolving on the left*. This interesting potential property of t-seminorms is studied in more detail in [De Cooman and Kerre, 1994]. Furthermore, assume that $h_1 \overset{(\Pi,P,E)}{\leq} h_2$. Eq. (23) then immediately implies that $\Pi(\{\, x \mid x \in E \text{ en } h_1(x) \not\leq h_2(x) \,\}) = 0_L$. We conclude that in this particular case, my definition coincides with the immediate extensions of the classical definitions. It is furthermore interesting to note that the product operator $\times$ on the unit interval $[0, 1]$ is a t-(semi)norm on $([0, 1], \leq)$ which is in particular strongly resolving on the left. We are thus led to the interesting conclusion that the reason why we have to use definitions which are more specific than the classical ones, is that we are using operators P which are more general than the product operator used in classical measure theory.

7 A RADON-NIKODYM-LIKE THEOREM

In this section I prove a result that is an analogon of the famous theorem of Radon-Nikodym in classical measure theory [Burrill, 1972]. Let us denote by P a triangular seminorm on $(L, \leq)$, such that $(L, \leq, P)$ is a complete lattice with t-seminorm.

As a classical first step towards this Radon-Nikodym-like theorem, the following proposition shows that it is always possible to construct a new possibility measure using a fuzzy variable and a possibility integral associated with another possibility measure.

Proposition 7.1. *Let $(X, \mathcal{R}, \Pi)$ be a $(L, \leq)$-possibility space. Let h be a $(L, \leq)$-fuzzy variable in $(X, \mathcal{R})$. The $\mathcal{R} - L$ mapping Φ, defined as*

$$(\forall A \in \mathcal{R}) \left(\Phi(A) = (P) \int\!\!\!\!-_A h \mathrm{d}\Pi \right),$$

is a $(L, \leq)$-possibility measure on $(X, \mathcal{R})$.

Proof. We denote by π the distribution of Π. Consider an arbitrary family $\{A_j \mid j \in J\}$ of elements of $\mathcal{R}$. Then, taking into account Eq. (5) and the associativity of supremum in the complete lattice $(L, \leq)$:

$$\Phi(\bigcup_{j \in J} A_j) = \sup_{x \in \bigcup_{j \in J} A_j} P(h(x), \pi(x)) = \sup_{j \in J} \sup_{x \in A_j} P(h(x), \pi(x)) = \sup_{j \in J} \Phi(A_j). \quad \square$$

This result leads us to the following important question. Consider two $(L, \leq)$-possibility measures Π and Φ on $(X, \mathcal{R})$; can we find a $(L, \leq)$-fuzzy variable h in $(X, \mathcal{R})$, such that for any E in $\mathcal{R}$:

$$\Phi(E) = (P) \int\!\!\!\!-_E h \mathrm{d}\Pi?$$

In the theorem below, I give a sufficient condition in order that this would indeed be the case.

Theorem 7.2 (Radon-Nikodym). *Let Π and Φ be two $(L, \leq)$-possibility measures on $(X, \mathcal{R})$ with respective distributions π and φ. If* *(i)* $(\forall E \in \mathcal{R})(\Phi(E) \leq \Pi(E))$, *or equivalently,* $\varphi \sqsubseteq \pi$,

(ii) P *is weakly invertible,*

there exists a $(L, \leq)$*-fuzzy variable* h *in* $(X, \mathcal{R})$ *such that*

$$(\forall E \in \mathcal{R}) \left(\Phi(E) = (P) \int_E h \mathrm{d}\Pi \right). \tag{24}$$

This $(L, \leq)$*-fuzzy variable* h *is furthermore unique in the sense of* (Π, P)*-equivalence.*

Proof. Let us give a proof by construction. For any $(L, \leq)$-fuzzy variable h in $(X, \mathcal{R})$, we find, using Eq. (5), that (24) is equivalent with

$$(\forall x \in X)(\varphi(x) = P(h(x), \pi(x))). \tag{25}$$

Taking into account the assumptions and Proposition 2.2, we know that the $X - L$ mapping h, defined by $(\forall x \in X)(h(x) = \varphi(x) \lhd_P \pi(x))$, satisfies (25). Since the distributions φ and π are by definition $\mathcal{R}$-measurable, it immediately follows that h is also $\mathcal{R}$-measurable and is therefore a solution for the problem considered. Taking into account Proposition 6.4(iii), this solution is unique in the sense of (Π, P)-equivalence, which of course means that every solution of the problem considered must be (Π, P)-equivalent with h. Or, to formulate it in yet another way, the set of solutions of Eq. (24) is the equivalence class of the equivalence relation $\overset{(\Pi,P)}{=}$ that contains h. □

Theorem 7.2 is obviously analogous to the theorem of Radon and Nikodym. Eq. (24) can be considered as an integral equation in the fuzzy variable h. Theorem 7.2 provides us with sufficient conditions for the existence of solutions of this equation. Integral equations of the type (24) occur very frequently in the measure- and integral-theoretic account of possibility theory, and more particularly in the treatment of conditional possibility, as will be shown in Part II. The theorem above, together with the method of solving (24), implicit in its proof, will turn out to be a significant aid in dealing with such equations.

To conclude this section, let me point out that it is interesting to compare the sufficient conditions in Theorem 7.2 with the one that appears in the classical version of the Radon-Nikodym theorem. It is easily seen that condition (i) is a stronger variant of the classical condition of *absolute continuity* [Burrill, 1972]. Condition (ii), however, has no classical counterpart. It should nevertheless be noted that in this theory the operator P plays the same role as the product operator $\times$ does in the classical theory of measure and integration, and that the t-norm $\times$ on $([0, 1], \leq)$ turns out to be weakly invertible [De Cooman and Kerre, 1994]. From this we may conclude that (ii) is an additional condition, which occurs in this theory because we work with operators that are more general than the product operator.

8 PRODUCT MEASURES AND MULTIPLE INTEGRALS

In what follows, the structures $(X_1, \mathcal{R}_1, \Pi_1)$ and $(X_2, \mathcal{R}_2, \Pi_2)$ denote $(L, \leq)$-possibility spaces. The distributions of Π_1 and Π_2 are denoted by respectively π_1 and π_2. By $\mathcal{R}_1 \times \mathcal{R}_2$ we mean the product ample field of $\mathcal{R}_1$ and $\mathcal{R}_2$, and not the Cartesian product of the sets $\mathcal{R}_1$ and $\mathcal{R}_2$. Furthermore, T denotes a t-norm on the complete lattice $(L, \leq)$, such that $(L, \leq, T)$ is a complete lattice with t-norm.

In section 2, I have already indicated how the structures $(X_1, \mathcal{R}_1)$ and $(X_2, \mathcal{R}_2)$ can be combined into a new structure $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$. The ample fields $\mathcal{R}_1$ and $\mathcal{R}_2$ are the building blocks for the construction of the product ample field $\mathcal{R}_1 \times \mathcal{R}_2$ on the product universe $X_1 \times X_2$. The questions I want to answer in this section are the following. Can we also, using as building blocks the possibility measures Π_1 on $(X_1, \mathcal{R}_1)$ and Π_2 on $(X_2, \mathcal{R}_2)$, construct a possibility measure on $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$? And furthermore, can the $(L, \leq, T)$-possibility integral, associated with this possibility measure, be calculated in terms of the $(L, \leq, T)$-possibility integrals associated with Π_1 and Π_2? The first of these problems is treated in subsection 8.1, where I show how, using the notion of a $(L, \leq, T)$-possibility integral, the possibility measures Π_1 and Π_2 can be combined into their T-product. In subsection 8.2 I deal with the second problem. In doing so, I prove a result that is analogous to Fubini's theorem in the classical theory of measure and integration. The reader will undoubtedly notice that the analogy between our discussion and the treatment in the classical case is not restricted to Fubini's theorem alone. Indeed, the material in this section is developed along the same general lines as the treatment of product measures, chain and double integrals in classical measure theory [Burrill, 1972].

8.1 Product Possibility Measures

Let us attempt a step-by-step construction of a $(L, \leq)$-possibility measure on the structure $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$, using the possibility measures Π_1 and Π_2. For a start, we devise a recipe which allows us to associate with every $\mathcal{R}_1 \times \mathcal{R}_2$-measurable set an element of L, having $(L, \leq, T)$-possibility integrals as its most important ingredient. This approach results in the definition of a particular $\mathcal{R}_1 \times \mathcal{R}_2 - L$-mapping. From the study of its properties, it will appear that this mapping is a $(L, \leq)$-possibility measure on $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$.

Consider an arbitrary element E of $\mathcal{R}_1 \times \mathcal{R}_2$. We use the following notations, for any (x_1, x_2) in $X_1 \times X_2$: $x_1E = \{ y_2 \mid (x_1, y_2) \in E \}$ and $Ex_2 = \{ y_1 \mid (y_1, x_2) \in E \}$. The following measurability properties are easily proven [Wang, 1982]: for any x_1 in X_1 and x_2 in X_2, $x_1E \in \mathcal{R}_2$ and $Ex_2 \in \mathcal{R}_1$. This allows us to meaningfully associate with E the following mappings, which are, in a sense, symmetrical counterparts: $g_E \colon X_1 \to L \colon x_1 \mapsto \Pi_2(x_1E)$ and $h_E \colon X_2 \to L \colon x_2 \mapsto \Pi_1(Ex_2)$. For these mappings, we now prove the following auxiliary property.

Lemma 8.1. *g_E is $\mathcal{R}_1$-measurable and h_E is $\mathcal{R}_2$-measurable.*

Proof. Let us for instance prove that g_E is $\mathcal{R}_1$-measurable. Consider an arbitrary x_1 in X_1 and an arbitrary y_1 in $[\, x_1 \,]_{\mathcal{R}_1}$. For any y_2 in X_2, we have, taking into account Eq. (3) and $[\, x_1 \,]_{\mathcal{R}_1} = [\, y_1 \,]_{\mathcal{R}_1}$, that $[\, (x_1, y_2) \,]_{\mathcal{R}_1 \times \mathcal{R}_2} = [\, (y_1, y_2) \,]_{\mathcal{R}_1 \times \mathcal{R}_2}$, whence

$$y_2 \in x_1E \Leftrightarrow [\, (x_1, y_2) \,]_{\mathcal{R}_1 \times \mathcal{R}_2} \subseteq E \Leftrightarrow [\, (y_1, y_2) \,]_{\mathcal{R}_1 \times \mathcal{R}_2} \subseteq E \Leftrightarrow y_2 \in y_1E.$$

We conclude that $x_1E = y_1E$ whence $\Pi_2(x_1E) = \Pi_2(y_1E)$ and therefore also $g_E(x_1) = g_E(y_1)$. This implies that g_E is $\mathcal{R}_1$-measurable. □

Using the $(L, \leq, T)$-possibility integrals associated with Π_1 and Π_2, we now define the mappings $\rho_1 \colon \mathcal{R}_1 \times \mathcal{R}_2 \to L$ and $\rho_2 \colon \mathcal{R}_1 \times \mathcal{R}_2 \to L$, which can also be considered as symmetrical counterparts:

$$\rho_1(E) = (T) \int_{X_1} g_E \mathrm{d}\Pi_1 = (T) \int_{X_1} \Pi_2(\cdot E) \mathrm{d}\Pi_1$$

and

$$\rho_2(E) = (T)\int_{X_2} h_E \mathrm{d}\Pi_2 = (T)\int_{X_2} \Pi_1(E\cdot)\mathrm{d}\Pi_2,$$

for any E in $\mathcal{R}_1 \times \mathcal{R}_2$. In Theorem 8.2, we show that these mappings constitute one and the same possibility measure on $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$.

Theorem 8.2. *(i)* $(\forall E \in \mathcal{R}_1 \times \mathcal{R}_2)(\rho_1(E) = \rho_2(E))$.

(ii) Let us write $\rho = \rho_1 = \rho_2$ *and define the* $X_1 \times X_2 - L$*-mapping* π *as*

$$(\forall (x_1, x_2) \in X_1 \times X_2)(\pi(x_1, x_2) = T(\pi_1(x_1), \pi_2(x_2)))$$

Then ρ *is the only* $(L, \leq)$*-possibility measure on* $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$ *with distribution* π.

(iii) Furthermore, ρ *satisfies*

$$(\forall A_1 \in \mathcal{R}_1)(\forall A_2 \in \mathcal{R}_2)(\rho(A_1 \times A_2) = T(\Pi_1(A_1), \Pi_2(A_2))) \tag{26}$$

and this determines ρ *uniquely, i.e., there is only one* $(L, \leq)$*-possibility measure on* $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$ *for which* (26) *holds.*

Proof. Let us first prove (i). Consider an arbitrary E in $\mathcal{R}_1 \times \mathcal{R}_2$. Taking into account Lemma 8.1, g_E is $\mathcal{R}_1$-measurable. Eq. (5) and the definition of g_E therefore allow us to write that

$$\begin{aligned}
\rho_1(E) &= \sup_{x_1 \in X_1} T(g_E(x_1), \pi_1(x_1)) \\
&= \sup_{x_1 \in X_1} T(\sup_{x_2 \in x_1 E} \pi_2(x_2), \pi_1(x_1)) \\
&= \sup_{x_1 \in X_1} \sup_{x_2 \in x_1 E} T(\pi_2(x_2), \pi_1(x_1)) \\
&= \sup_{(x_1, x_2) \in E} T(\pi_2(x_2), \pi_1(x_1)).
\end{aligned}$$

In an analogous way, we find that

$$\rho_2(E) = \sup_{(x_1, x_2) \in E} T(\pi_1(x_1), \pi_2(x_2)).$$

It now follows from the commutativity of the t-norm T that $\rho_1(E) = \rho_2(E)$.

The proof of (ii) is now immediate. Indeed, from the proof of (i) and the definition of ρ it follows for any E in $\mathcal{R}_1 \times \mathcal{R}_2$ that

$$\rho(E) = \sup_{(x_1, x_2) \in E} T(\pi_1(x_1), \pi_2(x_2)) = \sup_{(x_1, x_2) \in E} \pi(x_1, x_2).$$

Using this formula, it is easily proven that ρ is a $(L, \leq)$-possibility measure on $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$. It is clear that π is $\mathcal{R}_1 \times \mathcal{R}_2$-measurable, and that it is the unique distribution of ρ. Of course, ρ is the only $(L, \leq)$-possibility measure on $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$ that has π as its distribution.

To conclude this proof, let us show that (iii) holds. On the one hand, consider arbitrary A_1 in $\mathcal{R}_1$ and A_2 in $\mathcal{R}_2$. Since by definition $A_1 \times A_2 \in \mathcal{R}_1 \times \mathcal{R}_2$, it follows from (ii) that

$$\begin{aligned}
\rho(A_1 \times A_2) &= \sup_{(x_1,x_2)\in A_1\times A_2} T(\pi_1(x_1), \pi_2(x_2)) \\
&= \sup_{x_1\in A_1} \sup_{x_2\in A_2} T(\pi_1(x_1), \pi_2(x_2)) \\
&= T(\sup_{x_1\in A_1} \pi_1(x_1), \sup_{x_2\in A_2} \pi_2(x_2)) \\
&= T(\Pi_1(A_1), \Pi_2(A_2)).
\end{aligned}$$

On the other hand, let κ be a $(L, \leq)$-possibility measure on $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$ satisfying

$$(\forall A_1 \in \mathcal{R}_1)(\forall A_2 \in \mathcal{R}_2)(\kappa(A_1 \times A_2) = T(\Pi_1(A_1), \Pi_2(A_2))).$$

We already know that there exists at least one such κ. In particular, such a κ by definition satisfies $(\forall(x_1, x_2) \in X_1 \times X_2)(\kappa([\,(x_1, x_2)\,]_{\mathcal{R}_1\times\mathcal{R}_2}) = T(\pi_1(x_1), \pi_2(x_2)))$, which means that κ has π as its distribution. We conclude from (ii) that $\kappa = \rho$. □

The possibility measure ρ in this theorem can be constructed in a very simple way, using the possibility measures Π_1 and Π_2. In Definition 8.3 it is given a a special name, which is more or less obvious, given the analogy with the standard introduction of classical product measures. Let me also point out that the t-norm T in this treatment plays the same role as the algebraic product operator $\times$ in the classical theory.

Definition 8.3. The $(L, \leq)$-possibility measure ρ is called the T-product possibility measure – or, shortly, the T-product measure or the T-product – of Π_1 and Π_2, and will from now on be denoted by $\Pi_1 \times_T \Pi_2$.

The product possibility measures introduced here are generalizations of the product possibility measures found in the literature for $(L, \leq) = ([0, 1], \leq)$ and $T = \min$ [Wang, 1982], [Zadeh, 1978a]. To my knowledge, however, it is noted here for the first time that fuzzy (and, more particularly, possibility) integrals can be used to define such product measures, and this using a methodology that is analogous to the one used in the Lebesgue theory of measure and integration.

So far, we have concentrated on the definition of the T-product of *two* possibility measures. There is no reason, however, why we should not be able to extend this course of reasoning to the product of more than two possibility measures. In a fairly trivial and completely analogous way, this extension leads to the definition of the T-product $\Pi_1 \times_T \cdots \times_T \Pi_n$ on the structure $(X_1 \times \cdots \times X_n, \mathcal{R}_1 \times \cdots \times \mathcal{R}_n)$ of an n-tuple, $n \in \mathbb{N}^*$, of $(L, \leq)$-possibility measures Π_k on $(X_k, \mathcal{R}_k)$, $k = 1, \ldots, n$. This extension is left implicit, however.

8.2 Multiple Possibility Integrals

Let us consider an arbitrary $(L, \leq)$-fuzzy variable h in $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$ and an arbitrary element E of $\mathcal{R}_1 \times \mathcal{R}_2$. In the previous subsection, we have introduced the $(L, \leq)$-possibility measure $\Pi_1 \times_T \Pi_2$ on $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$. We are therefore able to integrate h over E, using the possibility integral associated with the T-product measure $\Pi_1 \times_T \Pi_2$:

$$(T)\!\!\int_E h\mathrm{d}(\Pi_1 \times_T \Pi_2).$$

There are, however, other ways of integrating the fuzzy variable h over E, which involve the partial mappings of h. Taking into account Eq. (3), we find that $h(x_1, \cdot)$ is $\mathcal{R}_2$-measurable, for any x_1 in X_1; and that $h(\cdot, x_2)$ is $\mathcal{R}_1$-measurable, for any x_2 in X_2.

Now, consider an arbitrary x_1 in X_1. Since it is easily shown that $x_1E \in \mathcal{R}_2$, we may write the following, also taking into account Eq. (5):

$$(T)\!\!\!\!\!\!\!\!\;\;\;\; ⨍_{x_1E} h(x_1, \cdot)\mathrm{d}\Pi_2 = \sup_{x_2 \in x_1E} T(h(x_1, \cdot)(x_2), \pi_2(x_2)) = \sup_{x_2 \in x_1E} T(h(x_1, x_2), \pi_2(x_2)).$$

A similar integral can be considered for every choice of x_1. This course of reasoning therefore leads to the introduction of a mapping

$$h_1^E \colon X_1 \to L \colon x_1 \mapsto (T)⨍_{x_1E} h(x_1, \cdot)\mathrm{d}\Pi_2.$$

In a completely symmetrical way, the mapping

$$h_2^E \colon X_2 \to L \colon x_2 \mapsto (T)⨍_{Ex_2} h(\cdot, x_2)\mathrm{d}\Pi_1$$

is defined. It is fairly obvious that the integration process can be completed by considering the integrals

$$(T)⨍_{X_1} h_1^E \mathrm{d}\Pi_1 \text{ and } (T)⨍_{X_2} h_2^E \mathrm{d}\Pi_2.$$

Informally stated, this means that we integrate h over E by first performing an integration in one possibility space and then integrating the partial result in the other possibility space. In Definition 8.4 we summarize this course of reasoning.

Definition 8.4. Let h be an $X_1 \times X_2 - L$-mapping that is $\mathcal{R}_1 \times \mathcal{R}_2$-measurable, and let E be a $\mathcal{R}_1 \times \mathcal{R}_2$-measurable set. We can associate with h and E the mappings h_1^E and h_2^E. We call the integrals $(T)⨍⨍_E h\mathrm{d}\Pi_1\mathrm{d}\Pi_2 = (T)⨍_{X_2} h_2^E\mathrm{d}\Pi_2$ and $(T)⨍⨍_E h\mathrm{d}\Pi_2\mathrm{d}\Pi_1 = (T)⨍_{X_1} h_1^E\mathrm{d}\Pi_1$ chain $(L, \leq, T)$-possibility integrals of h on E. Furthermore, $(T)⨍_E h\mathrm{d}(\Pi_1 \times_T \Pi_2)$ will be called the double $(L, \leq, T)$-possibility integral of h on E.

We have in all given three ways in which the fuzzy variable h can be integrated over the set E. We proceed to show in Theorem 8.6 that they all lead to the same result. In the classical theory of measure and integration, an analogous result is known as Fubini's theorem.

Lemma 8.5. *Let h be a $(L, \leq)$-fuzzy variable in $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$. Then for every E in $\mathcal{R}_1 \times \mathcal{R}_2$ the $X_1 - L$-mapping h_1^E is a $(L, \leq)$-fuzzy variable in $(X_1, \mathcal{R}_1)$ and the $X_2 - L$-mapping h_2^E is a $(L, \leq)$-fuzzy variable in $(X_2, \mathcal{R}_2)$.*

Proof. We give the proof for h_1^E. The proof for h_2^E is completely analogous. Consider an arbitrary x_1 in X_1 and y_1 in $[\, x_1 \,]_{\mathcal{R}_1}$. Then, by definition,

$$h_1^E(y_1) = (T)⨍_{y_1E} h(y_1, \cdot)\mathrm{d}\Pi_2.$$

A course of reasoning analogous to that of Lemma 8.1 shows that $x_1E = y_1E$. Furthermore, consider an arbitrary x_2 in X_2. We know that $h(\cdot, x_2)$ is $\mathcal{R}_1$-measurable, whence $h(x_1, x_2) =$

$h(y_1, x_2)$. We may therefore conclude that the partial mappings $h(x_1, \cdot)$ and $h(y_1, \cdot)$ are identical. Consequently,

$$(T)\fint_{y_1 E} h(y_1, \cdot)\mathrm{d}\Pi_2 = (T)\fint_{x_1 E} h(x_1, \cdot)\mathrm{d}\Pi_2,$$

and therefore also $h_1^E(x_1) = h_1^E(y_1)$. □

Theorem 8.6 (Fubini). *Let h be a $(L, \leq)$-fuzzy variable in $(X_1 \times X_2, \mathcal{R}_1 \times \mathcal{R}_2)$. Then for any E in $\mathcal{R}_1 \times \mathcal{R}_2$:*

$$(T)\fint_E h\mathrm{d}(\Pi_1 \times_T \Pi_2) = (T)\fint\fint_E h\mathrm{d}\Pi_1\mathrm{d}\Pi_2 = (T)\fint\fint_E h\mathrm{d}\Pi_2\mathrm{d}\Pi_1.$$

Proof. Since h is $\mathcal{R}_1 \times \mathcal{R}_2$-measurable, we have on the one hand that

$$(T)\fint_E h\mathrm{d}(\Pi_1 \times_T \Pi_2) = \sup_{(x_1,x_2)\in E} T(h(x_1, x_2), T(\pi_1(x_1), \pi_2(x_2))).$$

On the other hand, we have by definition that, also taking into account Eq. (5), Lemma 8.5 and the commutativity and associativity of T,

$$\begin{aligned}
(T)\fint\fint_E h\mathrm{d}\Pi_2\mathrm{d}\Pi_1 &= (T)\fint_{X_1} h_1^E \mathrm{d}\Pi_1 \\
&= \sup_{x_1\in X_1} T(h_1^E(x_1), \pi_1(x_1)). \\
&= \sup_{x_1\in X_1} T(\sup_{x_2\in x_1 E} T(h(x_1, x_2), \pi_2(x_2)), \pi_1(x_1)) \\
&= \sup_{x_1\in X_1} \sup_{x_2\in x_1 E} T(T(h(x_1, x_2), \pi_2(x_2)), \pi_1(x_1)) \\
&= \sup_{(x_1,x_2)\in E} T(T(h(x_1, x_2), \pi_2(x_2)), \pi_1(x_1)) \\
&= \sup_{(x_1,x_2)\in E} T(h(x_1, x_2), T(\pi_1(x_1), \pi_2(x_2))).
\end{aligned}$$

This proves the first equality. The proof of the second equality is completely analogous. □

The course of reasoning followed in this subsection can be extended in a fairly trivial way to the case of more than two possibility spaces. This leads to the definition of multiple possibility integrals, for which an immediate generalization of Theorem 8.6 remains valid. As before, such an extension will be left implicit.

We conclude this section with a result that can be informally described as 'the conversion of a double integral into a product of simple integrals'. Let us start by repeating here the definition of the well-known notion of *cylindric extension* [Kerre, 1991]. It is of course possible to provide this notion with a far more general definition than the one given here. Our Definition 8.7 is however sufficiently general for the use that will be made of this notion in this series of papers.

Definition 8.7. Let h_1 be a $X_1 - L$-mapping and let h_2 be a $X_2 - L$-mapping.

(i) The $X_1 \times X_2 - L$-mapping $\overline{h_1}$, defined by $\overline{h_1}(x_1, x_2) = h_1(x_1)$, for any (x_1, x_2) in $X_1 \times X_2$, is called the cylindric extension of h_1 to $X_1 \times X_2$.

(ii) The $X_1 \times X_2 - L$-mapping $\overline{h_2}$, defined by $\overline{h_2}(x_1, x_2) = h_2(x_2)$, for any (x_1, x_2) in $X_1 \times X_2$, is called the cylindric extension of h_2 to $X_1 \times X_2$.

Proposition 8.8. *Let h_1 and h_2 be $(L, \leq)$-fuzzy variables in $(X_1, \mathcal{R}_1)$ and $(X_2, \mathcal{R}_2)$ respectively. Let furthermore A_1 be an element of $\mathcal{R}_1$ and A_2 an element of $\mathcal{R}_2$.*

(i) $\overline{h_1}$, $\overline{h_2}$ *and* $T \circ (\overline{h_1}, \overline{h_2})$ *are* $\mathcal{R}_1 \times \mathcal{R}_2$*-measurable.*

(ii) $(T)\!\!\!\!\!-\!\!\!\!\int_{A_1 \times A_2} T \circ (\overline{h_1}, \overline{h_2}) \mathrm{d}(\Pi_1 \times_T \Pi_2) = T\left((T)\!\!\!\!\!-\!\!\!\!\int_{A_1} h_1 \mathrm{d}\Pi_1, (T)\!\!\!\!\!-\!\!\!\!\int_{A_2} h_2 \mathrm{d}\Pi_2\right).$

Proof. The proof of (i) is immediate. We give the proof of (ii). Taking into account (i), Eq. (5) and the commutativity and associativity of T, we have for any A_1 in $\mathcal{R}_1$ and A_2 in $\mathcal{R}_2$ that

$$\begin{aligned}(T)\!\!\!\!\!-\!\!\!\!\int_{A_1 \times A_2} T \circ (\overline{h_1}, \overline{h_2}) \mathrm{d}(\Pi_1 \times_T \Pi_2) &= \sup_{(x_1,x_2) \in A_1 \times A_2} T(T_{k=1}^2 \overline{h_k}(x_1, x_2), T_{k=1}^2 \pi_k(x_k)) \\ &= \sup_{x_1 \in A_1} \sup_{x_2 \in A_2} T(T(h_1(x_1), \pi_1(x_1)), T(h_2(x_2), \pi_2(x_2))) \\ &= T(\sup_{x_1 \in A_1} T(h_1(x_1), \pi_1(x_1)), \sup_{x_2 \in A_2} T(h_2(x_2), \pi_2(x_2))).\end{aligned}$$

Again taking into account Eq. (5), this completes the proof. □

This result is also a possibilistic counterpart of a well-known result in classical measure and integration theory [Burrill, 1972]. The role of the product operator is of course in possibility theory taken over by a t-norm T.

9 CONCLUSION

In the first paper of this series, I have laid the foundations for a measure- and integral-theoretic formulation of possibility theory. I have formalized the notion of a variable in possibility theory, and studied possibilistic and fuzzy variables together with their possibility distributions. Using the notion of a seminormed fuzzy integral, a variety of integral-theoretic notions and results have been discussed and derived, which show that these seminormed fuzzy integrals and possibility measures are a perfect match, in very much the same way as Lebesgue integrals and classical measures are. I explicitly mention possibility integrals and their properties; the (extended) possibility of fuzzy variables; the almost everywhere equality and dominance of fuzzy variables; product possibility measures, double possibility integrals and a Fubini-like theorem; how to work with possibility distributions to calculate integrals, other possibility distributions and the possibility of fuzzy variables; and finally the discussion of a special class of integral equations and a Radon-Nikodym-like theorem for their solution.

I have attempted to make this discussion of possibility theory very general, and some will probably say that it is more general than called for. My reason for doing so is that possibility theory is still a very young branch of mathematics, and I do not want to preclude *a priori* and wittingly any direction of subsequent research. That is why I am working with a complete lattice as a codomain for possibility measures, instead of the much more popular, but more restrictive real unit interval. In this way, it remains possible to consider incomparability of possibilities. Moreover, I want to leave room for the claim that possibility is an ordinal notion, and not a cardinal one. In the same spirit, I consider possibility measures defined on ample fields, and not just on power sets, as is generally the case. Finally, I use triangular seminorms and norms defined on complete lattices instead of the minimum operator on the unit interval. In an earlier paper [De Cooman and Kerre, 1995], I have shown that these operators are in a sense the most general operators using which possibility integrals may be defined. The reader will have noticed that I only use t-norms where their associativity and commutativity is necessary. In all other places, I only make use of t-seminorms. This is in keeping with my *maximum generality* approach.

This brings us to a very interesting open question. The triangular (semi)norms in this paper serve as possibilistic counterparts for the product operator in probability theory. Is this many-to-one relationship fundamental, or does there exist a unique t-(semi)norm P that is to possibility theory what the product operator is to probability theory? It should be mentioned here that at least some results [De Cooman, 1995b, 1995c] seem to indicate that the choice $P = \frown$ is of special importance in possibility theory, because it yields a number of interesting properties, formally analogous to probability theory, which cannot be obtained for other choices of P.

There is one obvious way in which the results in this paper might still be further generalized, namely by considering the definition of possibility measures on more general set structures than ample fields: plump fields, σ-fields, fields, etc. Actually, it might be asked why I have used ample fields from the beginning, and closely related to that question, why I do not define possibility measures as set mappings that preserve countable suprema, in analogy with classical measure theory. I feel that such a definition would be artificial in this context. Indeed, one of the reasons why in classical measure theory σ-fields and countable additivity are used, is that uncountable additivity is impossible. There seems to be nothing, on the other hand, that prevents us from considering uncountable 'supitivity'. Sticking to countable supitivity in this context is, to my knowledge, not really called for, however interesting the mathematical theory such a definition would generate. On the other hand, philosophical reasons might lead to the consideration of finitary definitions of possibility measures on fields of sets. This step might make matters more complicated, however, and is the subject of my current research.

In the introduction to this paper, I mentioned that possibility theory can be couched in a measure- and integral-theoretic language, and that the approach described here can be used to unify a number of existing results in possibility theory. I am now in a position to give at least a partial justification for my claim. Indeed, I have shown in section 3 that possibility integrals can be used to define the (extended) possibility of a measurable fuzzy set [Dubois and Prade, 1985, 1988], [Zadeh, 1978a], (indirectly) to formalize and generalize the notion of a possibilistic variable [Zadeh, 1978a] and its possibility distribution, and to introduce the notion of a product possibility measure [Zadeh, 1978a].

The material presented here is a necessary prelude to the more involved measure- and integral-theoretic treatment of conditional possibility and possibilistic independence, given in Parts II and III of this series of papers.

10 ACKNOWLEDGEMENT

Gert de Cooman is a Postdoctoral Fellow of the Belgian National Fund for Scientific Research (NFWO). He would like to thank the NFWO for supporting the research reported on in this paper.

11 REFERENCES

Benferhat, S., D. Dubois and H. Prade [1994], "Expressing independence in a possibilistic framework and its application to default reasoning." In: *Proceedings of ECAI 94 (11th European Conference on Artificial Intelligence)*, edited by A. Cohn, John Wiley & Sons, pp. 150–154.

Benferhat, S., D. Dubois and H. Prade [1995], "Possibilistic independence and plausible reasoning." In: *Foundations and Applications of Possibility Theory*, edited by G. de Cooman, D. Ruan and E. E. Kerre, World Scientific, Singapore, pp. 47–63.

Birkhoff, G. [1967], *Lattice Theory.* A.M.S. Colloquium Publications, Providence, Rhode Island.

Boyen, L., G. de Cooman and E. E. Kerre [1995], "On the extension of P-consistent mappings." In: *Foundations and Applications of Possibility Theory*, edited by G. de Cooman, D. Ruan and E. E. Kerre, World Scientific, Singapore, pp. 88–98.

Burrill, C. W. [1972], *Measure, Integration and Probability.* McGraw-Hill, New York.

Davey, B. A. and H. A. Priestley [1990], *Introduction to Lattices and Order.* Cambridge University Press, Cambridge.

De Cooman, G. [1993], *Evaluatieverzamelingen en -afbeeldingen – Een ordetheoretische benadering van vaagheid en onzekerheid [Evaluation Sets and Mappings—An Order-Theoretic Approach to Vagueness and Uncertainty].* PhD thesis, Universiteit Gent.

De Cooman, G. [1995a], "The formal analogy between possibility and probability theory." In: *Foundations and Applications of Possibility Theory*, edited by G. de Cooman, D. Ruan and E. E. Kerre, World Scientific, Singapore, pp. 71–87.

De Cooman, G. [1995b], "Reliability theory: a possibilistic perspective." *Fuzzy Sets and Systems*, submitted for publication.

De Cooman, G. [1995c], "Towards a possibilistic logic." In: *Fuzzy Set Theory and Advanced Mathematical Applications*, edited by D. Ruan, Kluwer Academic, Boston, pp. 89–133.

De Cooman, G. [1995d], "Generalized possibility and necessity measures on fields of sets." In: *Proceedings of the International ISCS Symposium on Fuzzy Logic*, edited by N. C. Steele, ICSC Academic Press, pp. A91–A98.

De Cooman, G. and E. E. Kerre [1993], "Ample fields." *Simon Stevin*, **67**, pp. 235–244.

De Cooman, G. and E. E. Kerre [1994], "Order norms on bounded partially ordered sets." *Journal of Fuzzy Mathematics*, **2**, pp. 281–310.

De Cooman, G. and E. E. Kerre [1995], "Possibility and necessity integrals." *Fuzzy Sets and Systems*, in print.

De Cooman, G., E. E. Kerre and F. Vanmassenhove [1992], "Possibility theory: An integral theoretic approach." *Fuzzy Sets and Systems*, **46**, pp. 287–300.

Doob, J. L. [1953], *Stochastic Processes.* John Wiley & Sons, New York.

Dubois, D., L. Fariñas del Cerro, A. Herzig and H. Prade [1994a], "An ordinal view of independence with application to plausible reasoning." In: *Uncertainty in Artificial Intelligence (Proceedings of the Tenth Conference (1994))*, pp. 195–203.

Dubois, D., J. Lang and H. Prade [1994b], "Possibilistic logic." In: *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 3, edited by D. M. Gabbay *et al.*, Clarendon Press, Oxford, pp. 439–513.

Dubois, D., S. Moral and H. Prade [1993a], *A semantics for possibility theory based on likelihoods.* Annual Report CEC-ESPRIT III BRA 6156 DRUMS II.

Dubois, D. and H. Prade [1985], *Théorie des possibilités.* Masson, Paris.

Dubois, D. and H. Prade [1988], *Possibility Theory – An Approach to Computerized Processing of Uncertainty.* Plenum Press, New York.

Dubois, D. and H. Prade [1990], "The logical view of conditioning and its application to possibility and evidence theories." *International Journal of Approximate Reasoning*, **4**, pp. 23–46.

Dubois, D., H. Prade and R. R. Yager [1993b], *Readings in Fuzzy Sets for Intelligent Systems.* Morgan Kaufman, San Mateo, CA.

Giles, R. [1982], "Foundations for a theory of possibility." In: *Fuzzy Information and Decision Processes*, edited by M. M. Gupta and E. Sanchez, North Holland, New York, pp. 183–195.

Goguen, J. A. [1967], "L-fuzzy sets." *Journal of Mathematical Analysis and Applications*, **18**, pp. 145–174.

Higashi, M. and G. J. Klir [1982], "Measures of uncertainty and information based on possibility distributions." *International Journal of General Systems*, **9**, pp. 43–58.

Hisdal, E. [1978], "Conditional possibilities independence and noninteraction." *Fuzzy Sets and Systems*, **1**, pp. 283–297.

Jacobs, K. [1978], *Measure and Integral.* Academic Press, New York.

Kerre, E. E. [1991], "Basic principles of fuzzy set theory for the representation and manipulation of imprecision and uncertainty." In: *Introduction to the Basic Principles of Fuzzy Set Theory and Some of Its Applications*, edited by E. E. Kerre, Communication & Cognition, Ghent, pp. 1–158.

Klir, G. J. [1990], "A principle of uncertainty and information invariance." *International Journal of General Systems*, **17**, pp. 249–275.

Klir, G. J. and T. A. Folger [1988], *Fuzzy Sets, Uncertainty and Information.* Prentice Hall, Englewood Cliffs, New Jersey.

Klir, G. J. and D. Harmanec [1995], "On some bridges to possibility theory." In: *Foundations and Applications of Possibility Theory*, edited by G. de Cooman, D. Ruan and E. E. Kerre, World Scientific, Singapore, pp. 3–19.

Klir, G. J. and B. Parviz [1992], "Probability-possibility transformations: a comparison." *International Journal of General Systems*, **21**, pp. 291–310.

Kolmogorov, A. N. [1950], *Foundations of Probability.* Chelsea Publishing, New York, translation.

Nahmias, S. [1978], "Fuzzy variables." *Fuzzy Sets and Systems*, **1**, pp. 97–110.

Nguyen, H. T. [1978], "On conditional possibility distributions." *Fuzzy Sets and Systems*, **1**, pp. 299–309.

Ralescu, D. [1982], "Toward a general theory of fuzzy variables." *Journal of Mathematical Analysis and Applications*, **86**, pp. 176–193.

Ramer, A. [1989], "Conditional possibility measures." *Cybernetics and Systems*, **20**, pp. 233–247.

Shafer, G. [1976], *A Mathematical Theory of Evidence.* Princeton University Press, Princeton.

Walley, P. [1991], *Statistical Reasoning with Imprecise Probabilities.* Chapman and Hall, London.

Wang, P.-Z. [1982], "Fuzzy contactability and fuzzy variables." *Fuzzy Sets and Systems*, **8**, pp. 81–92.

Wang, Z. and G. J. Klir [1992], *Fuzzy Measure Theory.* Plenum Press, New York.

Zadeh, L. A. [1975a], "The concept of a linguistic variable and its application to approximate reasoning I." *Information Science*, **8**, pp. 199–249.

Zadeh, L. A. [1975b], "The concept of a linguistic variable and its application to approximate reasoning II." *Information Science*, **8**, pp. 301–357.

Zadeh, L. A. [1976], "The concept of a linguistic variable and its application to approximate reasoning III." *Information Science*, **9**, pp. 43–80.

Zadeh, L. A. [1978a], "Fuzzy sets as a basis for a theory of possibility." *Fuzzy Sets and Systems*, **1**, pp. 3–28.

Zadeh, L. A. [1978b], "Pruf — a meaning representation language for natural languages." *International Journal of Man-Machine Studies*, **10**, pp. 395–460.