1 Introduction

Inaccuracies are omnipresent in our every day life ranging from our daily language to sketchy decision criteria to errors of measurement. Already Oskar Morgenstern[2] pointed out that all economic decisions are characterized by the fact that both quantitative and non-quantitative information influence the act of decision-making (cf. Morgenstern 1965, p. 1). Even emotions and mainly intuition can be of central importance in the process (cf. Holtfort 2011, p. 507).

Intuition can be regarded in different ways, one of which is considering it as knowledge gained by experience. Players knowingly and unknowingly gather information and keep them in mind. That is why they make decisions "from the gut" instead of employing a process of weighing up different alternatives (cf. Holtfort 2011, p. 508). Also economic theory deals with contents which are e.g. obtained from personal experiences or other "non-constructional" ways (cf. Morgenstern 1965, p. 88). Hence, subjective judgment can be an important factor.

Another matter that needs to be raised concerning impreciseness are linguistic inaccuracies which are present in our daily language. People speak of something being "likely" or "unlikely", that a "large" claim occurred, that an event will happen in the "near" future, etc. Usually, the terms are not precisely mathematically described and implemented in the decision-making process or mathematical models but remain vague.

An actuary at an insurance company is faced with similar situations. On the one hand, his/her calculations and the underlying models need to be as precise as possible and should correspond to the given data whether it is in the field of pricing, claims reserving, product controlling or other. At the same time, he/she should form an opinion or assessment to what extent his/her calculations express a realistic view of (future) reality.

[2] Oskar Morgenstern, * 1902, † 1977, was an Austrian-US-American economist.

Adjustments due to personal opinions are necessary or common in various situations: Circumstances are thinkable in which the given data is not sufficient or does not reflect a realistic view. Whether the derived consequences are too optimistic or pessimistic – in comparison to personal experience – adjustments in one or the other direction are possible.
Moreover, personal assessment can be found in the context of risk classification. "Determining the significant factors [...] however, may involve some subjective judgment" (cf. Outreville 1988, p. 150). Risks are often clustered in risk classes. If a risk can be seen as part of one class or another, an actuary will decide to which class the risk belongs. The presence of subjective judgment can also be seen in connection to pricing within the use of expert systems e.g. when historical data and/or information is not available (cf. Culp 2006, p. 164). In claims reserving subjective judgment is also present when estimating loss reserves on a single case basis (cf. Lemaire 1988, p. 396). In this particular case, actuaries often do not employ mathematical models but set up reserves according to the information they have at hand as well as their personal opinion and experience.

In practice, the actuary tends to adjust particular data (or parts of it) solely based on his/her intuition. This also applies in the field of claims reserving on which the present work will focus. This subdiscipline of insurance mathematics provides (stochastic) methods which are methodically advanced and diverse. The literature on claims reserving deals with various models which differ in a lot of facets: whether they are purely computational or make use of a stochastic framework, employ credibility theory or the Kalman filter, refer to paid or incurred data, etc. In addition, regulatory frameworks set more legal requirements in order to determine the reserve.

Nevertheless, the methods presented in the literature do not generally comprise a formalized approach to consider these subjective judgments. From a methodological point of view, the theory of fuzzy set seems to be appropriate. Fuzzy sets represent an approach to describe vagueness and imprecision. Particularly, it aims to depict limitations on knowledge and inaccuracies in context with numerical descriptions of problems. Thereby, it emanates from subjective judgments and not from global assumptions. Since its introduction in the seminal paper of Lotfi A. Zadeh (cf. Zadeh 1965) it underwent a continuous further development. It provides means to compute with vague expressions and a whole "computing tool box" has been developed. It took about 20 years until the first applications in insurance were discussed and another 20 years till the theory found its way into claims reserving.

In view of the above, the present work pursues two main goals. On the one hand, it aims to give an overview of existing methodical approaches for depicting subjective judgment in claims reserving. In this context, all necessary basics of the two branches of research, i.e. claims reserving and fuzzy set theory, are summarized. On the other hand, the present work will show in what way subjective judgment can be implemented in the two most popular claims reserving methods, i.e. the chain-ladder (CL) (cf. Section 4.3) and the Bornhuetter Ferguson (BF) (cf. Section 4.4) method. Therefore, three new approaches – two based on CL and one based on BF – are developed in the scope of this dissertation.

As already mentioned, subjective judgment can arise in the field of claims reserving. This is addressed in Chapter 5 of which an earlier version has been published in *Insurance: Mathematics & Economics* (cf. Heberle and Thomas 2016). The fundamental idea of the fuzzy chain-ladder (FCL) model is to model the observation that actuaries might adjust previously calculated development factors according to their subjective judgment in the context of the CL method using fuzzy numbers and their corresponding arithmetic. To the best of our knowledge, the CL method has not been the center of studies within fuzzy applications in spite of its popularity. The method aims to model the development factors with the help of triangular shaped fuzzy numbers with equal spreads as they are easy to implement. Moreover, the derived reserve has been fuzzy as well in earlier models, i.e. not a specific value has been derived but a range of possible values. Since a reserve is a figure in the balance sheet of an insurance company it needs to be specified as a crisp number. Therefore, our method will apply a defuzzification method and also will make an attempt to quantify the uncertainty of the prediction.

This idea shall be further pursued in the second approach (see Chapter 6) by applying a fuzzy regression method which assumes a functional relation described with fuzzy coefficients to the CL method in its linear model representation. In fact, a regression "tube" will be derived in which all data points lie. Methods of fuzzy regression partially provide possibilities to subjectively assess the scattering of the data around the regression line which describes a functional relationship. This fact shall be additionally taken into consideration, such that actuaries can subjectively assess the informative value of the data. For this approach we will model the fuzzy coefficients with triangular shaped fuzzy numbers of different spreads. The aim is to reduce the prediction uncertainty – depending on the data – in this model compared to the FCL model.

An additional idea is pursued in the third approach (see Chapter 7): The considered BF method utilizes a priori information which is given either by an external source (market statistics, expert knowledge, organizational data, etc.) or internally. In case a priori information is available the BF method is a popular alternative to the CL method. Depending on its origin the information can be afflicted with vagueness. The fuzzy Bornhuetter Ferguson (FBF) model aims to map this with the help of fuzzy numbers. Development factors as well as the a priori information shall be modeled with fuzzy numbers. Again, the goal is to derive predictions for the reserves.

The structure of the work is as follows: Subsequent to the general introduction the theory of fuzzy sets is introduced in Chapter 2. The basic concepts are presented and all necessary theory for the later chapters is provided. Moreover, a distinction between the theory of fuzzy sets and probability theory is drawn. The comparison is followed by a literature survey of applications of fuzzy methods in insurance in Chapter 3. It is classified by the field of application: underwriting, risk classification, pricing and claims reserving. The latter point is only mentioned briefly since it is discussed in detail in Chapter 4. Here the problem of claims reserving as well as an overview of reserving methods is addressed. Subsequently, the FCL method which makes use of fuzzy numbers and their corresponding arithmetic is presented in Chapter 5. The model is motivated and examined followed by a numerical example. As a further utilization a model of fuzzy regression is applied to the CL method in Chapter 6, in particular the representation as a sequence of linear models. The description of the methodology is followed by an example for which the same data base as before is used in order to draw a comparison. The FBF method is introduced in Chapter 7 and additionally takes into account a priori information. Within this model framework reserves and their corresponding uncertainty are deduced. A short summary is given and questions for further research are raised in Chapter 8.

2 | Fuzzy Theory

In reality we often face situations in which we cannot decide precisely whether an object belongs to a certain class or not. We might consider classes like “all tall men” or “all damages much higher than 10,000 €”. In these cases it is not always possible to distinguish between objects which are members of those classes and those which are not. A class like e.g. “all good students” is not a set in the mathematical sense as the word “good” is not described in a mathematically precise manner. A first approach to even consider those sets has been introduced by Zadeh (cf. Zadeh 1965).
The aim of this chapter is to give an introduction to the theory of fuzzy sets, fuzzy numbers and fuzzy regression. Firstly, we give a motivation for the introduction of this theory and outline the historic development in Section 2.1.1. Subsequently, a presentation of fuzzy sets according to Zadeh is given. In Section 2.2 fuzzy numbers as a special case of fuzzy sets are introduced and the corresponding arithmetic is defined. Finally, the basic concepts of fuzzy regression are presented in Section 2.3.

2.1 Fuzzy Sets

2.1.1 Historical Development

For a long time the scholars in philosophy taught that logic is a two-valued science. Therefore a proposition or statement is either e.g. “true” or “false”, “0” or “1”. Aristotle (384 BC – 322 BC), a student of Plato, was one of the founders of this school. Aristotle’s law of non-contradiction, the second law of “The Three Laws of Thought”, states that a proposition cannot be true and false at the same time. Additionally, his law of the excluded middle says that an object either possesses an attribute or its opposite.

Aristotle forbids that a proposition "in the middle" holds true, i.e. that an object only possesses an attribute to a certain extent (cf. Eberhart and Shi 2011, p. 284).

Scientists at the beginning of the twentieth century began to argue that not everything is two-valued. In mathematics a theory is developed for "perfect" objects. If objects fulfill these assumptions the proposition holds true, otherwise not. As objects in reality usually are not "perfect", the mathematical findings are – strictly speaking – not applicable. Therefore, Black proposes to introduce a symbolism for vagueness, i.e. for the case if requirements are not completely fulfilled (cf. Black 1937, pp. 427-429).

In one of his lectures Bertrand Russell pursues a similar idea. "I propose that all language is vague and that therefore my language is vague" (cf. Russell 1923, p. 84). Language is often a source of problems in modeling. As an example the following situation could be considered: If a person tells his friend to be at his house shortly after five p.m., this can be perceived by his friend in another way than it was actually meant. In one understanding this might be the time interval from 5:00 to 5:10 p.m. whereas another person could be of the opinion that this is the time interval from 5:00 to 5:15 p.m. In order to cope with the problem of imprecise language Black proposed the use of so-called *consistency-profiles* to picture this vagueness (cf. Black 1937, pp. 430ff.).

A first idea of a set theory which allows for a multi-valued logic came up from Menger. His *ensembles flous* are the French counterpart to fuzzy sets (cf. Menger 1951). In fact, the work of Lotfi A. Zadeh laid the foundation of the further research on fuzzy sets (cf. Zadeh 1965) and his theory is closer to the idea of Black's consistency-profiles (cf. Dubois and Prade 1980, p. 4).

After the publication of Zadeh's pioneer article a fast development in research took place. A survey about the early works can be found in Kaufmann (1975). For further literature see e.g. Dubois and Prade (1980), Bandemer and Gottwald (1993), Zimmermann (2001) or Dubois and Prade (2000). Nonetheless, the reactions to Zadeh's publication were divided. While the responses in the United States were quite cautious or there were harsh critics[3], scientists and technicians in Europe and Japan jumped onto the bandwagon. The use of fuzzy logic succeeded in Europe for the first time with the regulation of a steam raising unit in a power plant (cf. Altrock 1995, p. 7). Further applications in Japan were the regulation of a metro in which fuzzy technologies allowed for smooth

[3] For a collection of citations of critics see e.g. Altrock (1995, p. 7) and Dubois and Prade (2000, p. 3).

driveaways and slowdowns as well as implementations in camcorders, microwaves and washing machines (cf. Altrock 1995, pp. 8f.).

2.1.2 Basic Notations

The representation in this section basically relies on Bandemer and Gottwald (1993), Dubois and Prade (1980), Kruse et al. (1995) and Zimmermann (2001).
According to Cantor (1895) a set is “any collection M into a whole of definite, distinct objects m (which are called the “elements” of M) of our perception or of our thought” (cf. Cantor 1895, p. 481). In the following we will speak of a *crisp* set whenever we consider a set in the sense of Cantor. There are several ways to mathematically describe a set. Either every single element is listed or (for readability) the set is denoted with the help of descriptive characteristic traits, or the set is described with an indicator function (cf. Merz and Wüthrich 2013b, pp. 32ff.). As the objects are distinguishable in the case of crisp sets it is always possible to define an indicator function. Let Ω be a basic set and A a subset of Ω, i.e. $A \subset \Omega$. The function

$$\mathbf{1}_A : \ \Omega \longrightarrow \{0,1\}, \qquad \omega \mapsto \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

is called indicator function (or characteristic function) of A.

As stated before, there are linguistic inaccuracies in our everyday life as well as in business or scientific contexts. Indicator functions do not allow for gradual memberships to a certain set. In order to cope with that impreciseness Zadeh has defined the concept of fuzzy sets according to which we can specify the grade of membership for all elements $x \in X$ to a fuzzy set $\tilde{A}$ (cf. Zadeh 1965, p. 339). The idea was to enlarge the image set of a characteristic function and to permit grades of membership in the interval $[0,1]$.

Definition 2.1 (Fuzzy set)
Let $X \neq \emptyset$ be a collection of objects. A **fuzzy set** $\tilde{A}$ in X is defined by

$$\tilde{A} = \Big\{(x, \mu_{\tilde{A}}(x)) \mid x \in X\Big\}.$$

where $\mu_{\tilde{A}} : X \longrightarrow [0,1]$ is called **membership function**. We refer to $\mu_{\tilde{A}}(x)$ as the **grade of membership** of an element x in X with respect to $\tilde{A}$. Furthermore, we consider $F\mathscr{P}(X)$ as the **fuzzy power set** of X, i.e. the set of all fuzzy sets in X.

The membership function μ specifies to which extent an element belongs to a fuzzy set $\tilde{A}$. Consequently, the closer the grade of membership to one, the higher the grade of membership of x in $\tilde{A}$. Usually, those elements are assigned grade of membership of one which definitely belong to a set. The membership function is not necessarily bounded by zero and one. Nevertheless, it is a preferable setting. In fact, a fuzzy set is defined by its membership function.

An example of a membership function is shown in Figure 2.1. Dubois and Prade (1980) even refer to this kind of representation as extended Venn diagram. A classical Venn diagram cannot be drawn for fuzzy sets but an extended form can be used to visually verify set theoretic operations (cf. Dubois and Prade 1980, p. 14).

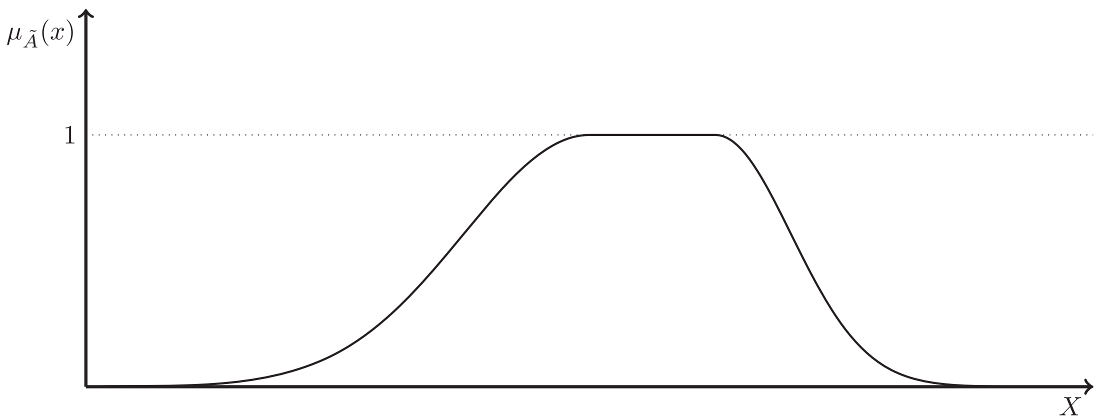


Figure 2.1: An example of a membership function of a fuzzy set $\tilde{A}$.

One criticism to fuzzy set theory often is the assignment of the grades of membership. In fact, they do not exist but are assigned to elements by an individual or a group. Therefore, it is a subjective assignment. There even might be fields of interest in which the mapping might be more controversial than in others. E.g. there might be more consensus in the field of “age” than in “smartness”.

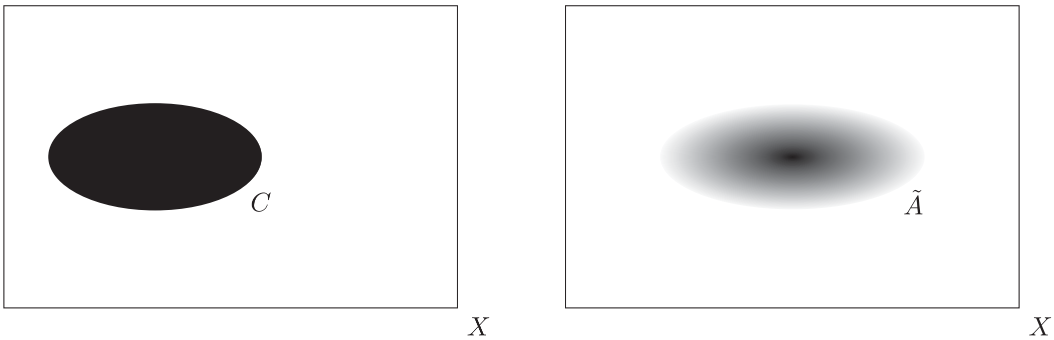
Remarks 2.2

a) The use of the interval $[0, 1]$ as the image set of the membership function allows for a convenient interpretation of the grade of membership.

b) Definition 2.1 ensures that every crisp set $A \in \mathscr{P}(X)$, where $\mathscr{P}(X)$ denotes the power set of X, can be considered as a fuzzy set since $\{0,1\} \subseteq [0,1]$.

c) Even though the membership function might be seen as a fuzzy analogon to a density function in probability theory there are remarkable differences which support that those two concepts should not be mixed up. Normally, the area enclosed by the membership function and the abscissae, i.e. the axis referring to the elements of the basic set X, is not necessarily one. The interpretation differs as the value of a membership function $\mu_{\tilde{A}}(x)$ expresses the grade of membership of a certain element x to a fuzzy set $\tilde{A}$. In contrast, a density function specifies probabilities of subsets of a basic set Ω where $(\Omega, \mathcal{A}, P)$ is a probability space.

The difference between a crisp and a fuzzy set is also visualized in Figure 2.2. While it is easy to verify whether an element $x \in X$ belongs to the crisp set $C \subset X$ in Figure 2.2a an element x of a fuzzy set $\tilde{A}$ cannot be identified clearly in all cases in Figure 2.2b. The lighter the color in Figure 2.2b, the lower is the grade of membership. This comes along with Zadeh's interpretation who pointed out that fuzziness stands for diffuse, i.e. fuzzy, boundaries (cf. Kruse et al. 1995, p. vi).



(a) A crisp set $C \in \mathscr{P}(X)$. (b) A fuzzy set $\tilde{A} \in F\mathscr{P}(X)$.

Figure 2.2: Illustration of a crisp and a fuzzy subset of a basic set X.

In the following some characteristics of fuzzy sets are presented in order to describe them.

Definition 2.3 (Height of a fuzzy set)
Let $\tilde{A}$ be a fuzzy set. We call

$$\mathrm{hgt}(\tilde{A}) := \sup_{x \in X} \mu_{\tilde{A}}(x)$$

height of a fuzzy set $\tilde{A}$.

Definition 2.3 states that the height of a fuzzy set $\tilde{A}$ specifies the least upper bound of the membership function $\mu_{\tilde{A}}$. Consequently, it indicates the largest grade of membership of a fuzzy set. If the height is one, normal fuzzy sets are being considered which are defined in the following Definition 2.4.

Definition 2.4 (Normal fuzzy set and core)
Let $\tilde{A}$ be a fuzzy set. $\tilde{A}$ is called **normal** if $\mathrm{hgt}(\tilde{A}) = 1$. Moreover, the crisp set of elements having a degree of membership of one is refered to as the **core** of $\tilde{A}$, i.e. $\mathrm{core}(\tilde{A}) = \{x \in X \mid \mu_{\tilde{A}}(x) = 1\}$.

Remarks 2.5

a) Fuzzy sets $\tilde{A}$ of which the basic set is nonempty with a height strictly between zero and one, i.e. $0 < \mathrm{hgt}(\tilde{A}) < 1$, are called **unnormal**.

b) A nonempty fuzzy set $\tilde{A}$ can always be normalized by division of $\mu_{\tilde{A}}(x)$ by $\sup_{x \in X} \mu_{\tilde{A}}(x)$ for all $x \in X$.

c) Definition 2.4 states that a fuzzy set $\tilde{A}$ is normal if the core is nonempty.

For convenience sake we will always consider normal fuzzy sets if not explicitly stated otherwise. Referring to Definition 2.1, the concept of fuzzy sets and its associated membership function is a generalization of a set and its indicator function in the classical sense (cf. Zimmermann 2001, p. 14). While considering fuzzy sets it is possible that a set also comprises elements with grade of membership zero. Consequently, one is interested in those elements with a grade of membership unequal to zero.

Definition 2.6 (Support of a fuzzy set)
Let $\tilde{A}$ be a fuzzy set. The crisp set

$$\mathrm{supp}(\tilde{A}) := \{x \in X \mid \mu_{\tilde{A}}(x) > 0\}$$

is called **support** of $\tilde{A}$.

This issue can be generalized by the concept of an α-level set. So far only a representation with the help of a membership function has been considered. There exists an alternative representation via α-level sets, $\alpha \in \mathbb{R}$. We consider $X = \mathbb{R}$. Then, they denote unions of intervals $[a, b] \subseteq X$ in which the elements are assigned a grade of membership of at least α. This issue is sometimes referred to as *horizontal representation* of a fuzzy set (cf. Kruse et al. 1995, p. 16).[4] As we are mostly considering fuzzy sets on the basic set $X = \mathbb{R}$ we are restricting ourselves to that case if not explicitly stated otherwise.

Definition 2.7 (α-level set)
Let $\tilde{A}$ be a fuzzy set and $\alpha \in [0, 1]$. The set

$$\tilde{A}_\alpha := \{x \in X \mid \mu_{\tilde{A}}(x) \geq \alpha\}$$

is called **α-level set** or **α-level (cut)**. The set $\tilde{A}'_\alpha := \{x \in X \mid \mu_{\tilde{A}}(x) > \alpha\}$ is called **strong α-level set** or **strong α-level (cut)**.

The concept of an strong α-level set and support of a fuzzy set is visualized in Figure 2.3. All elements $x \in X$ having a membership grade exceeding the threshold α belong to the strong α-level set. A strong α-level set is a subset of the support where for a choice of $\alpha = 0$ the support is given, i.e. $\tilde{A}'_0 = \text{supp}(\tilde{A})$. In fact, this concept allows us to form a family of crisp sets out of a fuzzy set if we build an α-level set for every $\alpha \in [0, 1]$. The notation α-level cut yields from the fact that the prescription for composing α-level sets divides up an originally fuzzy set into several crisp sets.

Moreover, every fuzzy set can be characterized by its α-cuts due to the following Proposition 2.8.

Proposition 2.8
Let $\tilde{A}$ be a fuzzy set over a basic set X and $\mu_{\tilde{A}}$ denotes the corresponding membership function. We then have

$$\mu_{\tilde{A}}(x) = \sup_{\alpha \in [0,1]} \left\{ \min \left(\alpha, \mathbf{1}_{\tilde{A}_\alpha}(x) \right) \right\}.$$

[4] Kruse et al. denote the representation via membership functions analogously as vertical representation.

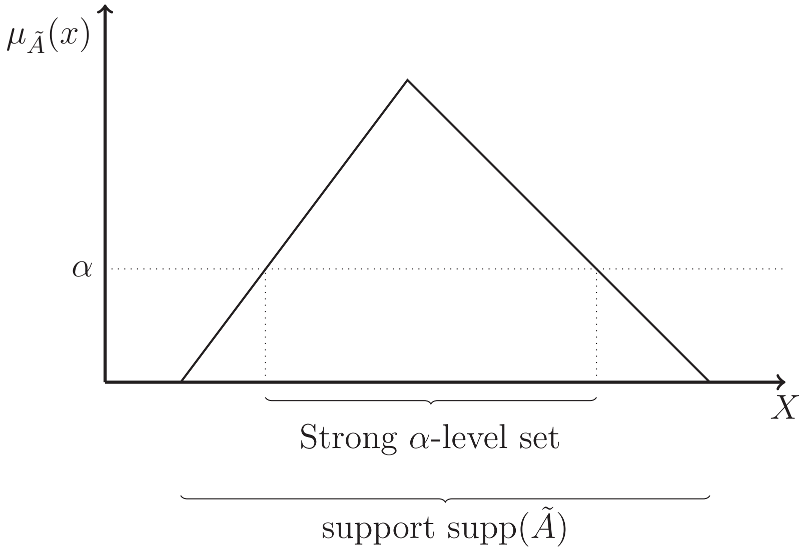


Figure 2.3: Strong α-level set $\tilde{A}'_\alpha$ and support supp$(\tilde{A})$ of a fuzzy set $\tilde{A}$.

Proof. Let $x \in X$ and $\alpha \in [0,1]$. It holds

$$\min\left(\alpha, \mathbf{1}_{\tilde{A}_\alpha}(x)\right) = \begin{cases} \alpha & \text{if } x \in \tilde{A}_\alpha \Leftrightarrow \mu_{\tilde{A}}(x) \geq \alpha \\ 0 & \text{if } x \notin \tilde{A}_\alpha \Leftrightarrow \mu_{\tilde{A}}(x) < \alpha \end{cases}.$$

We yield

$$\begin{aligned} \mu_{\tilde{A}}(x) &= \sup\{\alpha \mid \alpha \leq \mu_{\tilde{A}}(x)\} \\ &= \sup_{\alpha \in [0,1]} \left\{\min\left(\alpha, \mathbf{1}_{\tilde{A}_\alpha}(x)\right)\right\}. \end{aligned}$$

Another important concept in the theory of fuzzy sets is convexity. In crisp set theory convexity is defined with the help of the support whereas fuzzy set theory makes use of the membership function. This definition will be of use later on in order to define fuzzy numbers.

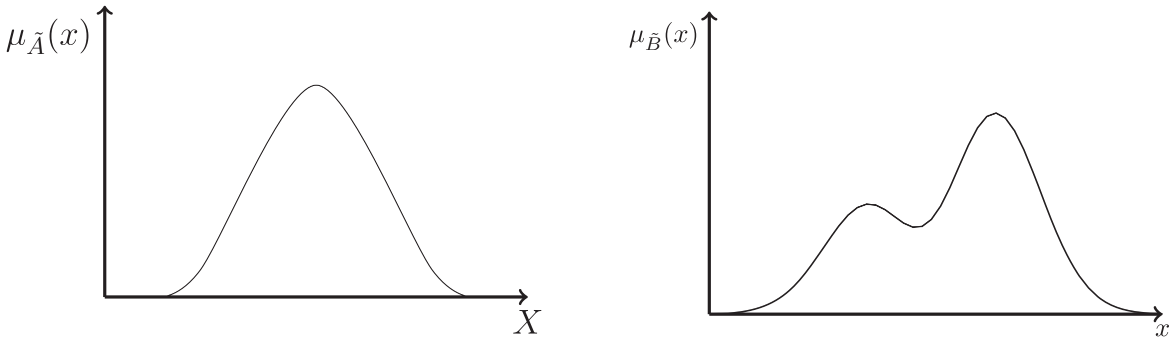
Definition 2.9 (Convex fuzzy set)
Let $\tilde{A}$ be a fuzzy set. A fuzzy set $\tilde{A}$ is called **convex** if

$$\mu_{\tilde{A}}(\lambda x_1 + (1-\lambda)x_2) \geq \min(\mu_{\tilde{A}}(x_1), \mu_{\tilde{A}}(x_2)), \qquad x_1, x_2 \in X, \lambda \in [0,1].$$

Remark 2.10

We also speak of a convex fuzzy set if all α-level sets are convex (cf. Zimmermann 2001, p. 15). As α-level sets are crisp sets here the definition of convexity for crisp sets applies.

An illustration of a convex and a non-convex fuzzy set is given in Figure 2.4. In plain English a fuzzy set is convex if the graph of its membership function does not possess any "valleys".



(a) A convex fuzzy set $\tilde{A}$. (b) A non-convex fuzzy set $\tilde{B}$.

Figure 2.4: Convex and non-convex fuzzy sets

Like for crisp sets set operations can be also defined for fuzzy sets. Generally they are stated as defined in Definition 2.11 (cf. Zimmermann 2001, pp. 16ff.).

Definition 2.11

Let $\tilde{A}$ and $\tilde{B}$ be fuzzy sets and $\mu_{\tilde{A}}$ and $\mu_{\tilde{B}}$ denote their corresponding membership functions. We then define the following:

a) Two fuzzy sets $\tilde{A}$ and $\tilde{B}$ are said to be **equal** if

$$\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x) \quad \text{for all } x \in X.$$

b) The membership function $\mu_{\tilde{A}\cap\tilde{B}}$ of the **intersection** $\tilde{A} \cap \tilde{B}$ is defined for every $x \in X$ by

$$\mu_{\tilde{A}\cap\tilde{B}}(x) := \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}.$$

c) The membership function $\mu_{\tilde{A}\cup\tilde{B}}$ of the **union** $\tilde{A}\cup\tilde{B}$ is defined for every $x\in X$ by

$$\mu_{\tilde{A}\cup\tilde{B}}(x) := \max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}.$$

d) The membership function $\mu_{\tilde{A}^C}$ of the complement $\tilde{A}^C$ of a normalized fuzzy set $\tilde{A}$ for every $x\in X$ is defined by

$$\mu_{\tilde{A}^C}(x) = 1 - \mu_{\tilde{A}}(x).$$

Remark 2.12
The definition of the membership function of the union and intersection is not intuitively at first hand. Bellman and Giertz came up with a justification for the definition with the help of the min- and max-operators (cf. Bellman and Giertz 1973).

2.1.3 The Extension Principle

The aim of this section is to present one of the most applied propositions in fuzzy set theory, the extension principle, which was firstly stated by Zadeh in a rather simple form (cf. Zadeh 1965, p. 346). The form which is used nowadays has been firstly shown in Zadeh (1975a,b,c).

The objective has been to transfer familiar operations like e.g. addition and multiplication to fuzzy sets. These will be used in Section 2.2. The definition will be stated in the sense of a generalization. Therefore, it is intended that the well-known operations are equivalent in the case of crisp sets. With the help of Definition 2.13 we can allow for fuzzy sets in a mapping instead of crisp elements.
We follow the presentations by Hanss (2005, pp. 44f.) and Zimmermann (2001, pp. 55f.).

Definition and Proposition 2.13 (Extension principle)
Let $X_1,\dots,X_n,X$ be non-empty sets and

$$f:\ X_1\times\ldots\times X_n \longrightarrow X$$

be a mapping. Let $\tilde{A}_1\in F\mathscr{P}(X_1),\ldots\tilde{A}_n\in F\mathscr{P}(X_n)$ be fuzzy sets with corresponding membership functions

$$\mu_{\tilde{A}_1}:\ X_1 \longrightarrow [0,1], \ldots, \mu_{\tilde{A}_n}:\ X_n \longrightarrow [0,1].$$

We yield the membership function $\mu_{f(\tilde{A}_1,\ldots,\tilde{A}_n)} : X \longrightarrow [0,1]$ of a fuzzy set $f(\tilde{A}_1,\ldots,\tilde{A}_n) \in F\mathscr{P}(X)$ by

$$\mu_{f(\tilde{A}_1,\ldots,\tilde{A}_n)}(x) = \begin{cases} \sup\limits_{x=f(x_1,\ldots,x_n)} \min\left(\mu_{\tilde{A}_1}(x_1),\ldots,\mu_{\tilde{A}_n}(x_n)\right) & \text{if there exist } x_1,\ldots,x_n \\ & \text{with } x = f(x_1,\ldots,x_n) \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

and therefore a mapping

$$\tilde{f}:\ F\mathscr{P}(X_1) \times \ldots \times F\mathscr{P}(X_n) \longrightarrow F\mathscr{P}(X).$$

We say that $f:\ X_1 \times \ldots \times X_n \longrightarrow X$ has been extended to $\tilde{f}:\ F\mathscr{P}(X_1) \times \ldots \times F\mathscr{P}(X_n) \longrightarrow F\mathscr{P}(X)$ by means of the **extension principle**.

Remarks 2.14

a) Every n-tuple $(x_1,\ldots,x_n) \in X_1 \times \ldots \times X_n$ is mapped to an element $x \in X$ via the mapping f. First, the minimal grades of membership are determined component-wise. Then, the supremum of these grades of membership is taken as the grade of membership of the element x to the fuzzy set $f(\tilde{A}_1,\ldots,\tilde{A}_n)$.

b) If there exist no inverse images in (2.1) we use the common convention that $\sup\emptyset = 0$.

2.1.4 Relationship between Probability Theory and Fuzzy Set Theory

This section aims to highlight the differences between classical probability theory and fuzzy set theory. The prevailing discussion exists since the emergence of fuzzy set theory with the seminal paper of Lotfi A. Zadeh in 1965 (cf. Zadeh 1965). In the course of controversial discussions about the differences and similarities Kosko wonders: “Is uncertainty the same as randomness? If we are not sure about something, is it only up