Fuzzy Transforms of Higher Order Approximate Derivatives: A Theorem

Irina Perfilieva[1] and Vladik Kreinovich[2]
[1]University of Ostrava
Inst. for Research and Applications of Fuzzy Modeling
Ostrava, Czech Republic, Irina.Perfilieva@osu.cz
[2]Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA, vladik@utep.edu

Abstract

In many practical applications, it is useful to represent a function $f(x)$ by its *fuzzy transform*, i.e., by the "average" values

$$F_i = \frac{\int f(x) \cdot A_i(x)\, dx}{\int A_i(x)\, dx}$$

over different elements of a *fuzzy partition* $A_1(x), \ldots, A_n(x)$ (for which $A_i(x) \geq 0$ and $\sum\limits_{i=1}^{n} A_i(x) = 1$). It is known that when we increase the number n of the partition elements $A_i(x)$, the resulting approximation get closer and closer to the original function: for each value x_0, the values F_i corresponding to the function $A_i(x)$ for which $A_i(x_0)$ tend to $f(x_0)$.

In some applications, if we approximate the function $f(x)$ on each element $A_i(x)$ not by a *constant* but by a *polynomial* (i.e., use a fuzzy transform of a *higher order*), we get an even better approximation to $f(x)$.

In this paper, we show that such fuzzy transforms of higher order (and even sometimes the original fuzzy transforms) not only approximate the function $f(x)$ itself, they also approximate its derivative(s). For example, we have $F_i'(x_0) \to f'(x_0)$.
Keywords: fuzzy transform, higher-order fuzzy transform, derivatives, universal approximation property

1 Fuzzy Transforms: A Brief Introduction

Need for fuzzy techniques: reminder. In many application areas, a significant part of our knowledge is in the form of human expertise, an expertise that most experts can only describe by using imprecise ("fuzzy") words from a natural language.

For example, a skilled driver cannot explain her skills by explicitly stating for how many angles one should turn the wheel to change the lane at a given speed, but this driver can express her rules by saying, say, that if we are traveling at a high speed, we should turn the wheel a little bit.

It is desirable to incorporate this "fuzzy" expertise into automatic control systems. To be able to do that, we must transform this knowledge into precise terms, terms understandable to a computer. Fuzzy techniques (see, e.g., [2, 3]) are the techniques have been specifically designed to formalize such "fuzzy" knowledge.

Fuzzy techniques: the main idea. In fuzzy techniques, every imprecise statement A (like "small") is represented by a *membership function*, i.e., a function that assigns, to every real number x, a value $A(x)$ which describes to what "degree" the property A is satisfied for this value x.

This membership function is also called a *fuzzy set.*

Fuzzy control: the main idea. Expert knowledge about control is usually formulated in the following terms: an expert selects words $A_1, \ldots, A_n$ like "small", "medium", "large", etc. describing his knowledge about the input x, and formulates (in similar imprecise terms) what actions to undertake under these n different assumptions about the input. Thus, we have rules of the following type:

- if the input x satisfies the property A_1 is true, then the control u should satisfy the property B_1;
- if the input x satisfies the property A_2 is true, then the control u should satisfy the property B_2;
- ...
- if the input x satisfies the property A_n is true, then the control u should satisfy the property B_n.

The corresponding fuzzy control methodology transforms these rules into a precise control strategy $u(x)$ [2, 3].

Typical selection of membership functions in fuzzy control. We want to select the rules in such a way that the corresponding terms cover all possible inputs x from a given range, i.e., for which, for every x, we have $\sum\limits_{i=1}^{n} A_i(x) = 1$. A collection of fuzzy sets with this property is called a *fuzzy partition.*

In the non-fuzzy case, when we talk about a partition, we usually mean that each element belongs to only one class. In the fuzzy case, transition is gradual, so when we move, e.g., from small to medium, then the degree of smallness gradually decreases from 1 (for a real small object) to 0 (for a truly medium object). Thus, it is reasonable to require that if we want to get good results on the domain $[\underline{x}, \overline{x}]$ of values x, then we should select values

$$x_0 < \underline{x} = x_1 < x_2 < \ldots < x_{n-1} < x_n = \overline{x} < x_{n+1}$$

for which each function $A_i(x)$:

- increases from 0 to 1 when $x_{i-1} \le x \le x_i$,
- attains the value 1 at $x = x_i$,
- decrease from 1 to 0 for $x_i \le x \le x_{i+1}$, and
- is equal to 0 outside the interval $[x_{i-1}, x_{i+1}]$.

Comment. Please notice that

- in our definition, the first and the last functions $A_1(x)$ and $A_n(x)$ are also defined slightly outside the domain of interest $[\underline{x}, \overline{x}]$ while
- in [4], these first and last functions are only considered inside the domain $[\underline{x}, \overline{x}]$.

This difference does not change the approximation property inside the domain (and our proofs), but it makes definitions and proofs much easier – because we no longer have to consider the first and the last functions separately.

Triangular membership functions and uniform partitions. In principle, we can use different functions $A_i(x)$. From the computational viewpoint, the simplest function $A_i(x)$ is a *triangular* function which is linear on both intervals $[x_{i-1}, x_i]$ and $[x_i, x_{i+1}]$:

- for $x \in [x_{i-1}, x_i]$, we have $A_i(x) = \dfrac{x - x_{i-1}}{x_i - x_{i-1}}$;
- for $x \in [x_i, x_{i+1}]$, we have $A_i(x) = 1 - \dfrac{x - x_i}{x_{i+1} - x_i}$.

It is also desirable to have a *uniform* partition, to avoid situations in which we have many rules about one part of the domain and very few rules about the other parts of the domain. This can be achieved if we select the values x_i equally spaced, i.e., $x_i = x_0 + i \cdot h_n$ for a step $h_n = \dfrac{\overline{x} - \underline{x}}{n-1}$, and select the functions $A_{n,1}(x)$, …, $A_{n,n}(x)$ which differ from each other only by shift, i.e., for which $A_{n,i}(x) = A_0\left(\dfrac{x - x_i}{h_n}\right)$ for some fixed function $A_0(x)$ defined on the interval $[-1, 1]$.

Such a partition is among the most frequently used in fuzzy control.

What if we apply similar ideas to a functional dependence. In some practical situations, we know the exact functional dependence $f(x)$ between the input quantity x and the auxiliary quantity y, but we do not know the exact value x. Instead, we only know whether the input x is, say, small, medium, or large.

More generally, we have the same fuzzy partition $A_1(x)$, ..., $A_n(x)$ (as in the case of fuzzy control), we know that the input x satisfies the property A_i, and we want to estimate the corresponding value y.

In this case, as the desired estimate, it is reasonable to take the "average" value of the function $f(x)$ over the fuzzy set $A_i(x)$, i.e., the value

$$F_i = \frac{\int f(x) \cdot A_i(x)\,dx}{\int A_i(x)\,dx}. \tag{1}$$

The corresponding set of values $F_1, \ldots, F_n$ is called the *fuzzy transform* of the function $f(x)$; see, e.g., [4] and references therein.

2 Fuzzy Transform: Formal Definition and Approximation Properties

Formal definition of a fuzzy transform. Thus, we arrive at the following definitions [4]. First, we define a fuzzy transform with respect to a single membership function:

Definition 1. *Let $f(x)$ and $A_i(x) \ge 0$ be functions. By a* fuzzy transform *of the function $f(x)$ with respect to the function $A_i(x)$, we mean a number*

$$F_i = \frac{\int f(x) \cdot A_i(x)\,dx}{\int A_i(x)\,dx}. \tag{2}$$

Now, we can define a fuzzy partition – as a sequence of membership functions satisfying certain properties – and a fuzzy transform with respect to a partition as a sequence of the corresponding fuzzy transforms.

Definition 2. *Let $[\underline{x}, \overline{x}]$ by an interval, let*

$$x_0 < \underline{x} = x_1 < x_2 < \ldots < x_{n-1} < x_n = \overline{x} < x_{n+1} \tag{3}$$

be a sequence of real numbers. By a fuzzy partition*, we mean a sequence of continuous functions $A_1(x), \ldots, A_n(x)$ for which each function $A_i(x)$:*

- *increases from 0 to 1 when $x_{i-1} \le x \le x_i$,*
- *attains the value 1 at $x = x_i$,*
- *decrease from 1 to 0 for $x_i \le x \le x_{i+1}$, and*
- *is equal to 0 outside the interval $[x_{i-1}, x_{i+1}]$.*

Definition 3. *Let $A_1(x), \ldots, A_n(x)$ be a fuzzy partition, and let $f(x)$ be a function. By a* fuzzy transform *of the function $f(x)$ with respect to the fuzzy partition $A_1(x), \ldots, A_n(x)$, we mean a sequence of values $F_1, \ldots, F_n$, where F_i is the fuzzy transform of $f(x)$ with respect to the function $A_i(x)$.*

Uniform partitions. For approximation purposes, it is useful to consider a special class of "uniform" partitions. To formulate this notion, we first need to define an auxiliary notion of the basic function:

Definition 4. *By a basic function $A_0(x)$, we mean a continuous function which:*

- *increases from 0 to 1 when $-1 \le x \le 0$,*
- *attains the value 1 at $x = 0$,*
- *decrease from 1 to 0 for $0 \le x \le 1$, and*
- *is equal to 0 outside the interval $[-1, 1]$.*

It is easy to prove the following result.

Proposition 1. *Let $[\underline{x}, \overline{x}]$ by an interval, let n be a positive integer, and let $A_0(x)$ be a basic function. Then, for the values*

$$x_{n,1} = \underline{x}, x_{n,2} = \underline{x} + h_n, \ldots, x_{n,i} = \underline{x} + (i-1) \cdot h_n, \ldots,$$
$$x_{n,n} = \underline{x} + (n-1) \cdot h_n = \overline{x}, \tag{4}$$

where $h_n \stackrel{\text{def}}{=} \dfrac{\overline{x} - \underline{x}}{n-1}$, the functions $A_{n,i}(x) = A_0\left(\dfrac{x - x_{n,i}}{h_n}\right)$ form a fuzzy partition.

Definition 5. *Let $A_0(x)$ be a basic function. The fuzzy partition formed by the functions*

$$A_{n,1}(x) = A_0\left(\frac{x - x_{n,1}}{h_n}\right), \ldots, A_{n,n}(x) = A_0\left(\frac{x - x_{n,n}}{h_n}\right)$$

is called a uniform *partition.*

Approximation properties of fuzzy transform. It is known (see [4] and references therein) that if fix a basic membership function $A_0(x)$ and a continuous function $f(x)$, and consider the uniform fuzzy partitions $A_{n,1}(x)$, ..., $A_{nn}(x)$ corresponding to different values n, then the fuzzy transform $F_{n,1}$, ..., $F_{n,n}$ tends to $f(x)$ as n increases, in the following precise sense: for every real number x_0, if we take, for every n, a value $i(n, x_0)$ for which $A_{n,i}(x_0) \ne 0$, then

$$F_{n,i(n,x_0)} \to f_i(x_0).$$

3 Fuzzy Transforms of Higher Order: Formulation of the Problem

Fuzzy transforms of higher order. To get a better approximation, it was proposed to approximate the function $f(x)$ on each fuzzy set $A_i(x)$ not by a *constant* $F_i(x)$, but by by a *polynomial*

$$F_i(x) = \sum_{j=0}^{d} F_{ij} \cdot (x - x_i)^j \tag{5}$$

of a given order d. The coefficients F_{ij} of this polynomial can be obtained from the condition that the mean square difference between $f(x)$ and $F_i(x)$ (weighted by $A_i(x)$) is the smallest possible:

$$\text{Minimize} \quad \int (f(x) - F_i(x))^2 \cdot A_i(x)\, dx. \tag{6}$$

Definition 6. *Let $f(x)$ and $A_i(x) \geq 0$ be functions, and let $d \geq 0$ be a natural number. By a* fuzzy transform of order d *of the function $f(x)$ with respect to the function $A_i(x)$, we mean a polynomial $F_i(x)$ of d-th order for which minimizes the expression $\int (f(x) - F_i(x))^2 \cdot A_i(x)\, dx$.*

Definition 7. *Let $A_1(x), \ldots, A_n(x)$ be a fuzzy partition, let $f(x)$ be a function, and let $d \geq 0$ be a natural number. By a* fuzzy transform of order d *of the function $f(x)$ with respect to the fuzzy partition $A_1(x), \ldots, A_n(x)$, we mean a sequence of polynomials $F_1(x), \ldots, F_n(x)$, where $F_i(x)$ is the fuzzy transform of order d of $f(x)$ with respect to the function $A_i(x)$.*

For $d = 0$, the optimization criterion (6) leads directly to the fuzzy transform (1). For $d > 0$, we can also produce an explicit solution to this optimization problem if we introduce the orthonormal basis $\varphi_{i,0}(x), \ldots, \varphi_{i,d}(x)$ on the class of all polynomials, i.e., polynomials $\varphi_{i,0}(x)$ of order 0, $\varphi_{i,1}(x)$ of order 1, …, and $\varphi_{i,d}(x)$ of order d for which

$$\int \varphi_{i,j}^2(x) \cdot A_i(x)\, dx = 1 \tag{7}$$

for all i and

$$\int \varphi_{i,j}(x) \cdot \varphi_{i,j'}(x) \cdot A_i(x)\, dx = 0 \tag{8}$$

for all $j \neq j'$.

In terms of this basis, the fuzzy transform of k-th order can be described as

$$F_i(x) = \sum_{j=0}^{d} f_{i,j} \cdot \varphi_{i,j}(x), \tag{9}$$

where

$$f_{i,j} \stackrel{\text{def}}{=} \int f(x) \cdot \varphi_{i,j}(x) \cdot A_i(x)\, dx. \tag{10}$$

What we do in this paper. In this paper, we prove that fuzzy transforms of the d-th order indeed approximate not only the original function $f(x)$ itself, but also its derivatives $f'(x)$, $f''(x)$, ..., $f^{(k)}(x)$, ..., $f^{(d)}(x)$.

4 Main Result: Approximation of Derivatives

Theorem 1. *Let:*

- *$d > 0$ be a natural number;*
- *$[\underline{x}, \overline{x}]$ be an interval;*
- *$A_0(x)$ be a basic function;*
- *for every n, $A_{n,1}(x)$, ..., $A_{n,n}(x)$ be a fuzzy partition generated by the basic function $A_0(x)$ on the interval $[\underline{x}, \overline{x}]$;*
- *for every $x_0 \in [\underline{x}, \overline{x}]$ and for every n, $i(n, x_0)$ will denote a value i for which $A_{n,i}(x_0) \neq 0$.*

Let:

- *$f(x)$ be a d times continuously differential function;*
- *$F_{n,1}(x)$, ..., $F_{n,n}(x)$ be the d-th order fuzzy transform of the function $f(x)$ with respect to the partition $A_{n,1}(x)$, ..., $A_{n,n}(x)$.*

Then, for every $k \le d$, as $n \to \infty$, we have

$$F^{(k)}_{n,i(n,x_0)}(x_0) \to f^{(k)}(x_0). \tag{11}$$

Comment. Specifically, we will prove that for every $\varepsilon > 0$, there exists an integer N such that for every x_0 and for every $n \ge N$, we have:

$$|F_{n,i(n,x_0)}(x_0) - f(x_0)| \le \varepsilon \cdot \frac{1}{n^d},$$

$$|F'_{n,i(n,x_0)}(x_0) - f'(x_0)| \le \varepsilon \cdot \frac{1}{n^{d-1}},$$

$$|F''_{n,i(n,x_0)}(x_0) - f''(x_0)| \le \varepsilon \cdot \frac{1}{n^{d-2}},$$

$$\ldots$$

$$|F^{(k)}_{n,i(n,x_0)}(x_0) - f^{(k)}(x_0)| \le \varepsilon \cdot \frac{1}{n^{d-k}}, \tag{12}$$

$$\dots$$
$$|F^{(d)}_{n,i(n,x_0)}(x_0) - f^{(d)}(x_0)| \le \varepsilon.$$

In terms of the "small o" notations, we have

$$F_{n,i(n,x_0)}(x_0) = f(x_0) + o\left(\frac{1}{n^d}\right),$$
$$\dots,$$
$$F^{(k)}_{n,i(n,x_0)}(x_0) = f^{(d)}(x_0) + o\left(\frac{1}{n^{d-k}}\right), \tag{13}$$
$$\dots,$$
$$F^{(d)}_{n,i(n,x_0)}(x_0) = f^{(d)}(x_0) + o\,(1)\,.$$

5 Proof of the Main Result

To prove this result, we introduce the following auxiliary definition.

Definition 8. *Let $d > 0$ be a natural number, let $f(x)$ a d times continuously differentiable function, and let x_0 be a real number. By a d-*th order Taylor polynomial *of the function $f(x)$ at the point x_0, we mean the polynomial*

$$T_{f,x_0}(x) \stackrel{\text{def}}{=} f(x_0) + f'(x_0)\cdot(x-x_0) + \frac{1}{2!}\cdot f''(x_0)\cdot(x-x_0)^2 + \ldots +$$
$$\frac{1}{k!}\cdot f^{(k)}(x_0)\cdot(x-x_0)^k + \ldots + \frac{1}{d!}\cdot f^{(d)}(x_0)\cdot(x-x_0)^s. \tag{14}$$

Lemma 1. *Let $f(x)$ be a d times continuously differentiable function on the interval $[\underline{x},\overline{x}]$. Then for every real number $\varepsilon > 0$ there exists a value $\delta > 0$ such that for every x and x_0, if $|x - x_0| \le \delta$, then*

$$|f(x) - T_{f,x_0}(x)| \le \varepsilon\cdot(x-x_0)^d. \tag{15}$$

Comment. In terms of the "small o" notations, this lemma can be rewritten as

$$f(x) = T_f(x, x_0) + o((x-x_0)^d),$$

where it is understood that the convergence in o is uniform in x.

Proof of Lemma 1: idea. We will prove Lemma 1 by induction over the order d. In this proof, we will use standard facts and results from calculus; see, e.g., [1].

Proof of Lemma 1: base case. Let us start with the base case of $d = 1$. In this case, the function $f(x)$ is continuously differentiable, i.e.,

- the function $f(x)$ is differentiable, and
- its derivative $f'(x)$ is continuous on the interval $[\underline{x}, \overline{x}]$.

It is known that every continuous function on an interval – in particular, the derivative $f'(x)$ – is uniformly continuous on this interval. Thus, for every real number $\varepsilon > 0$, there exists a value $\delta > 0$ for which, for every two points x' and x'' from this interval for which $|x' - x''| \le \delta$, we have $|f'(x') - f'(x'')| \le \varepsilon$.

Let us prove that for $d = 1$, the desired inequality (15) holds for this same value δ.

Indeed, due to the Mean Value Theorem, for every x and x_0, we have

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(c(x, x_0)) \tag{16}$$

for some value $c(x, x_0)$ which located in between x and x_0. Since the value $c(x, x_0)$ is located in between x and x_0, its distance to x_0 cannot exceed the distance between x and x_0, i.e., $|c(x, x_0) - x_0| \le |x - x_0|$. So, if $|x - x_0| \le \delta$, then $|c(x, x_0) - x_0| \le \delta$, and thus, $|f'(c(x, x_0)) - f'(x_0)| \le \varepsilon$.

If we denote $s_1(x, x_0) \stackrel{\text{def}}{=} f'(c(x, x_0)) - f'(x_0)$, then we can say that

$$f'(c(x, x_0)) = f'(x_0) + s_1(x, x_0), \tag{17}$$

where $|s_1(x, x_0)| \le \varepsilon$.

Substituting the expression for $f'(c(x, x_0))$ from the formula (16) into the expression (17), we now have

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) + s_1(x, x_0). \tag{18}$$

Multiplying both sides of the equality (18) by $x - x_0$, we conclude that

$$f(x) - f(x_0) = f'(x_0) \cdot (x - x_0) + s_1(x, x_0) \cdot (x - x_0). \tag{19}$$

Moving $f(x_0)$ to the right-hand side, we get

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + s_1(x, x_0) \cdot (x - x_0). \tag{20}$$

By definition (14) of the Taylor polynomial, we thus get

$$f(x) = T_{f,x_0}(x) + s_1(x, x_0) \cdot (x - x_0), \tag{21}$$

where $|s(x, x_0)| \le \varepsilon$. Thus, for $d = 1$, we get the desired inequality

$$|f(x) - T_{f,x_0}(x)| \le \varepsilon \cdot |x - x_0|. \tag{22}$$

The base case is proven.

Proof of Lemma 1: induction step. Let us now assume that we have already proved Lemma 1 for functions which are d times continuously differentiable. Let us prove that in this case, a similar statement is true for functions which are $d+1$ times differentiable.

Indeed, if a function $f(x)$ is $d+1$ times continuously differentiable, then its first derivative $f'(x)$ is d times continuously differentiable. Thus, by the induction assumption, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $|z-x_0| \le \delta$, then

$$|f'(z) - T_{f',x_0}(z)| \le \varepsilon \cdot |z - x_0|^d, \tag{23}$$

where

$$T_{f',x_0}(z) = f'(x_0) + f''(x_0) \cdot (z - x_0) + \ldots + \frac{1}{j!} \cdot f^{(j+1)}(x_0) \cdot (z - x_0)^j + \ldots$$
$$+ \frac{1}{d!} \cdot f^{(d+1)}(x_0) \cdot (z - x_0)^d. \tag{24}$$

The expression (23) can be rewritten as

$$f'(z) = T_{f',x_0}(z) + s_d(z, x_0) \cdot (z - x_0)^d, \tag{25}$$

where $s_d(z, x_0) \stackrel{\text{def}}{=} \dfrac{f'(z) - T_{f',x_0}(z)}{(z - x_0)^d}$ satisfies the inequality

$$|s_d(z, x_0)| \le \varepsilon. \tag{26}$$

It is well known that once we know the derivative $f'(x)$ of a function $f(x)$, we can reconstruct the original function $f(x)$ as the integral of this derivative:

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(z)\, dz. \tag{27}$$

Substituting the formulas (25) and (24) into this integral expression, and explicitly integrating each terms $(z - x_0)^j$, we conclude that

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + \frac{1}{2!} \cdot f''(x_0) \cdot (z - x_0)^2 + \ldots +$$
$$\frac{1}{(j+1)!} \cdot f^{(j+1)}(x_0) \cdot (z - x_0)^{j+1} + \ldots + \frac{1}{(d+1)!} \cdot f^{(d+1)}(x_0) \cdot (z - x_0)^{d+1} +$$
$$S(x, x_0), \tag{28}$$

where

$$S(x, x_0) \stackrel{\text{def}}{=} \int_{x_0}^{x} s_d(z, x_0) \cdot (z - x_0)^d\, dz. \tag{29}$$

All the terms in the right-hand side of (28) – with the exception of the term $S(x, x_0)$ – form the Taylor polynomial $T_{f,x_0}(x)$ of order $(d+1)$ corresponding to the function $f(x)$. Thus, the expression (28) can be rewritten as

$$f(x) = T_{f,x_0}(x) + S(x, x_0). \tag{30}$$

To complete our proof, we need to estimate the term $S(x, x_0)$. From (26), we conclude that

$$|S(x, x_0)| = \left| \int_{x_0}^{x} s_d(z, x_0) \cdot (z - x_0)^d \, dz \right| \le \int_{x_0}^{x} |s_d(z, x_0)| \cdot |z - x_0|^d \, dz \le$$

$$\varepsilon \cdot \int_{x_0}^{x} |s_d(z, x_0)| \cdot |z - x_0|^d \, dz = \varepsilon \cdot \frac{1}{d+1} \cdot |x - x_0|^{d+1}. \tag{31}$$

Since $d \ge 0$ and therefore $\dfrac{1}{d+1} \le 1$, we conclude that

$$|S(x, x_0)| \le \varepsilon \cdot |x - x_0|^{d+1}, \tag{32}$$

Therefore, we can rewrite $S(x, x_0)$ as

$$S(x, x_0) = s_{d+1}(x, x_0) \cdot (x - x_0)^{d+1}, \tag{33}$$

where $s_{d+1}(x, x_0) \stackrel{\text{def}}{=} \dfrac{S(x, x_0)}{(x - x_0)^{d+1}}$ satisfies the inequality $|s_{d+1}(x, x_0)| \le \varepsilon$.

The induction step is proven, and so is the lemma.

Resuming the proof of the theorem itself. Now that the lemma is proven, let us prove the theorem itself.

Let a function $f(x)$ be continuously differentiable on the interval $[\underline{x}, \overline{x}]$, and let x_0 be a value on this interval. According to Lemma 1, this means that for every $\varepsilon > 0$, there exists a value $\delta > 0$ such that for all the values x in the δ-vicinity of x_0, we have

$$f(x) = T_{f,x_0}(x) + s_d(x, x_0) \cdot (x - x_0)^d, \tag{34}$$

where $|s_k(x, x_0)| \le \varepsilon$. Since $|x - x_0| \le \delta$, we thus have

$$f(x) = T_{f,x_0}(x) + r_{x_0}(x), \tag{35}$$

where the term

$$r_{x_0}(x) \stackrel{\text{def}}{=} s_d(x, x_0) \cdot (x - x_0)^d \tag{36}$$

satisfies the inequality

$$|r_{x_0}(x)| \le \varepsilon \cdot |x - x_0|^d. \tag{37}$$

For each x_0, it is sufficient to consider values x which are close to x_0. By definition, the basic function $A_0(x)$ is only different from 0 for $x \in [-1, 1]$. For every n, each membership function $A_{n,i}(x) = A_0\left(\dfrac{x - x_{n,i}}{h_n}\right)$ is only different from 0 when $|x - x_{n,i}| \le h_n$. We have defined $i(n, x_0)$ as the value for which $A_{n,i}(x_0) \ne 0$. Thus, we have

$$|x_0 - x_{n,i}| \le h_n. \tag{38}$$

We are interested in the fuzzy transform $F_{n,i}(x)$ corresponding to this index $i = i(n, x_0)$. For this i, the computation of the fuzzy transform only involves values x for which $A_{n,i}(x) \neq 0$. For all the values x, we have $|x - x_{n,i}| \leq h_n$. Thus, due to (38), we get $|x - x_0| \leq |x - x_{n,i}| + |x_0 - x_{n,i}| \leq 2 \cdot h_n$, i.e.,

$$|x - x_0| \leq 2 \cdot h_n. \tag{39}$$

Here, $h_n = \dfrac{\overline{x} - \underline{x}}{n-1}$ is of order $\dfrac{1}{n}$. The value h_n decreases with n and tends to 0 as $n \to \infty$. Thus, for every $\varepsilon > 0$, once n gets so large that $2 \cdot h_n \leq \delta$, the inequality (37) becomes valid. So, for all x for which $A_{n,i}(x) \neq 0$, we have

$$|r_{x_0}(x)| \leq \varepsilon \cdot 2^k \cdot h_n^k. \tag{40}$$

Since $h_n \sim n^{-1}$, for the remainder $r_{x_0}(x)$, we thus get an upper bound of order n^{-k}.

The desired fuzzy transform as a sum of two fuzzy transforms. According to the formulas (9) and (10), the fuzzy transform of a sum of two functions is equal to the sum of their fuzzy transforms. Thus, the desired fuzzy transform $F_{n,i}(x)$ of the function $f(x)$ is equal to the sum of the following two fuzzy transforms:

- the fuzzy transform of the Taylor polynomial $T_{f,x_0}(x)$ of the d-th order, and
- the fuzzy transform $R_{n,i}(x)$ of the remainder function $r_{x_0}(x)$.

Fuzzy transform of the Taylor polynomial: result. By Definition 6, the fuzzy transform of d-th order of the function with respect to a membership function $A_{n,i}(x)$ is defined as a polynomial $F_{n,i}(x)$ that minimizes the squared difference between a polynomial of d-order and the given function. When the given function is already a d-th order polynomial, this difference can be made equal to the smallest possible value 0 if we select this same given polynomial as $F_{n,i}(x)$.

Thus, for a polynomial of d-th order – in particular, for the Taylor polynomial $T_{f,x_0}(x)$ – its fuzzy transform is simply equal to this same polynomial.

Fuzzy transform of the Taylor polynomial: conclusion. Since we know the fuzzy transform of the Taylor polynomial, to find the fuzzy transform of the function $f(x)$, we must now estimate the estimate the fuzzy transform of the remainder term $r_{x_0}(x)$. Since, as we have mentioned, this fuzzy transform is one of the two components of the desired fuzzy transform $F_{n,i}(x)$, we thus conclude that

$$F_{n,i}(x) = T_{f,x_0}(x) + R_{n,i}(x). \tag{41}$$

We want to prove that the values $F_{n,i}(x_0)$ and $F^{(k)}_{n,i}(x_0)$ $(1 \leq k \leq d)$ of the function $F_{n,i}(x)$ and of its first d derivatives at the point x_0 is close to the corresponding values $f(x_0)$ and $f^{(k)}(x_0)$.

By differentiating both sides of the equality (41) one or several times and substituting $x = x_0$, we conclude that for every $j \le k$, we have

$$F^{(k)}_{n,i}(x_0) = T^{(k)}_{f,x_0}(x_0) + R^{(k)}_{n,i}(x_0). \tag{42}$$

Let us first analyze the first term in this sum: the derivatives of the Taylor polynomial.

Derivatives of the Taylor polynomial. By definition of the Taylor polynomial $T_{f,x_0}(x)$ (Definition 8), for this polynomial, we have

$$T_{f,x_0}(x_0) = f(x_0), T'_{f,x_0}(x_0) = f'(x_0), \ldots, T^{(d)}_{f,x_0}(x_0) = f^{(d)}(x_0), \ldots,$$

$$T^{(k)}_{f,x_0}(x_0) = f^{(k)}(x_0). \tag{43}$$

So, the formula (41) takes the form

$$F^{(k)}_{n,i}(x_0) = f^{(k)}(x_0) + R^{(k)}_{n,i}(x_0). \tag{44}$$

Thus, to show that the differences between the values of the original derivatives $f^{(k)}(x_0)$ and of their fuzzy transform approximations $F^{(k)}_{n,i}(x_0)$ are indeed small, we need to show that the values and the derivatives $R^{(k)}_{n,i}(x_0)$ of the fuzzy transform $R_{n,i}(x)$ of the remainder term $r_{x_0}(x)$ are also small.

Towards estimating fuzzy transform of the remainder term: we need to find the orthonormal basis. We are interested in the fuzzy transform of the remainder term $r_{x_0}(x)$ with respect to the membership function $A_{n,i}(x) \stackrel{\text{def}}{=} A_0\left(\dfrac{x - x_i}{h_n}\right)$, where $i = i(n, x_0)$. According to the formulas (9) and (10), this fuzzy transform has the form

$$R_{n,i}(x) = \sum_{j=0}^{d} f_{n,i,j} \cdot \varphi_{n,i,j}(x), \tag{45}$$

where

$$f_{n,i,j} \stackrel{\text{def}}{=} \int r_{x_0}(x) \cdot \varphi_{n,i,j}(x) \cdot A_{n,i}(x)\, dx \tag{46}$$

and the functions $\varphi_{n,i,j}(x)$ are the orthonormal basis functions corresponding to the membership function $A_{n,i}(x)$. Thus, to describe the desired fuzzy transform $R_{n,i}(x)$, it is useful to find the corresponding orthonormal basis functions $\varphi_{n,i,j}(x)$.

Finding the orthonomal basis. The membership function $A_{n,i}(x)$ is obtained from the basic function $A_0(x)$ by a linear change of variables: $A_{n,i}(x) = A_0(y_{n,i}(x))$, where

$$y_{n,i}(x) \stackrel{\text{def}}{=} \frac{x - x_{n,i}}{h_n}. \tag{47}$$

It is therefore reasonable to try to describe the basis functions $\varphi_{n,i,j}(x)$ corresponding to the membership function $A_{n,x}(x)$ in terms of the orthonormal basis corresponding to the basic function $A_0(x)$.

Let $\varphi_0(x)$, $\varphi_1(x)$, ..., $\varphi_d(x)$ be the orthonormal polynomials of orders 0, 1, ..., d that correspond to the basic function $A_0(x)$, i.e., for which

$$\int \varphi_j^2(x) \cdot A_0(x)\, dx = 1, \tag{48}$$

and

$$\int \varphi_j(x) \cdot \varphi_{j'}(x) \cdot A_0(x)\, dx = 1 \tag{49}$$

for all $j \neq j'$.

As we have mentioned, the membership function $A_{n,i}(x)$ is obtained from the basic function $A_0(x)$ by a linear change of variables $A_{n,i}(x) = A_0(y_{n,i}(x))$. To find the orthonormal functions corresponding to the membership function $A_{n,i}(x)$, let us therefore try to perform a similar linear change of variables in the functions $\varphi_j(x)$, i.e., let us consider the auxiliary functions

$$\psi_{n,i,j}(x) \stackrel{\text{def}}{=} \varphi_j(y_{n,i}(x)) = \varphi_j\left(\frac{x - x_{n,i}}{h_n}\right). \tag{50}$$

For the new variable $y = y_{n,i}(x)$, formulas (48) and (49) lead to

$$\int \varphi_j^2(y) \cdot A_0(y)\, dy = 1, \tag{51}$$

and

$$\int \varphi_j(y) \cdot \varphi_{j'}(y) \cdot A_0(y)\, dy = 1 \tag{52}$$

for all $j \neq j'$, i.e., to

$$\int \psi_{n,i,j}^2(x) \cdot A_{n,i}(x)\, dy = 1, \tag{53}$$

and

$$\int \psi_{n,i,j}(x) \cdot \psi_{n,i,j'}(x) \cdot A_{n,i}(x)\, dy = 1 \tag{54}$$

for all $j \neq j'$.

From (47), we get $dy = \dfrac{dx}{h_n}$ and therefore, the formula (53) takes the form

$$\int \psi_{n,i,j}^2(x) \cdot A_{n,i}(x)\, \frac{dx}{h_n} = 1, \tag{55}$$

i.e., the form

$$\int \psi^2_{n,i,j}(x) \cdot A_{n,i}(x)\, dx = h_n. \tag{56}$$

Therefore, the functions

$$\varphi_{n,i,j}(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{h_n}} \cdot \psi_{n,i,j}(x) \tag{57}$$

satisfy the desired orthonormality conditions

$$\int \varphi^2_{n,i,j}(x) \cdot A_{n,i}(x)\, dx = 1, \tag{58}$$

and

$$\int \varphi_{n,i,j}(x) \cdot \varphi_{n,i,j'}(x) \cdot A_{n,i}(x)\, dx = 1 \tag{59}$$

for all $j \neq j'$.

Substituting the formula (50) into the expression (57), we conclude that the functions $\varphi_{m,i,j}(x)$ from the desired orthonormal basis have the form

$$\varphi_{n,i,j}(x) = \frac{1}{\sqrt{h_n}} \cdot \varphi_j\left(\frac{x - x_{n,i}}{h_n}\right). \tag{60}$$

Estimating the coefficients $f_{n,i,j}$ at the basis functions $\varphi_{n,i,j}(x)$. Let us use the above expression (60) for the orthonormal basis functions to estimate the coefficients $f_{n,i,j}$ (as described by the formula (46)).

Substituting the formula (60) and the expression for $A_{n,i}(x)$ into the formula (46), we conclude that

$$f_{n,i,j} = \int r_{x_0}(x) \cdot \varphi_{n,i,j}(x) \cdot A_{n,i}(x)\, dx =$$

$$f_{n,i,j} = \int r_{x_0}(x) \cdot \frac{1}{\sqrt{h_n}} \cdot \varphi_j\left(\frac{x - x_{n,i}}{h_n}\right) \cdot A_0\left(\frac{x - x_{n,i}}{h_n}\right) dx. \tag{61}$$

To estimate this interval, let us find all the bounds on all the factors in the integrated expression.

- First, we know, from (40), that

$$|r_{x_0}(x)| \leq \varepsilon \cdot 2^d \cdot h_n^d. \tag{62}$$

- The second factor $\dfrac{1}{\sqrt{h_n}}$ is simply a constant.

- Let us estimate the third factor. Let C denote the largest possible value of the absolute values $|\varphi_0(x)|, \ldots, |\varphi_d(x)|$ of all the polynomials $\varphi_0(x), \ldots, \varphi_d(x)$ on the interval $[-1, 1]$. Then, we can conclude that $|\varphi_j(x)| \leq C$ for all j and for all C. In particular, for the third factor, we have:

$$\left|\varphi_j\left(\frac{x - x_{n,i}}{h_n}\right)\right| \leq C. \tag{63}$$

- Finally, by the definition of the basic function $A_0(x)$, we have $|A_0(x)| \le 1$ for all x. In particular, for the third factor, we have a similar bound:

$$\left|A_0\left(\frac{x - x_{n,i}}{h_n}\right)\right| \le 1. \tag{64}$$

By combining all these inequalities, we conclude that the integrated expression in (61) is bounded by

$$\left|r_{x_0}(x) \cdot \frac{1}{\sqrt{h_n}} \cdot \varphi_j\left(\frac{x - x_{n,i}}{h_n}\right) \cdot A_0\left(\frac{x - x_{n,i}}{h_n}\right)\right| =$$

$$|r_{x_0}(x)| \cdot \frac{1}{\sqrt{h_n}} \cdot \left|\varphi_j\left(\frac{x - x_{n,i}}{h_n}\right)\right| \cdot \left|A_0\left(\frac{x - x_{n,i}}{h_n}\right)\right| \le$$

$$\varepsilon \cdot 2^d \cdot h_n^d \cdot \frac{1}{\sqrt{h_n}} \cdot 1 \cdot C = (C \cdot 2^d) \cdot \varepsilon \cdot h_n^{d-1/2}. \tag{65}$$

The integration is bounded by the values x for which $|x - x_{n,i}| \le h_n$, i.e., by the values from the interval $[x_{n,i} - h_n, x_{n,i} + h_n]$ of width $2 \cdot h_n$. Thus,

$$|f_{n,i,j}| \le \int_{x_{n,i}-h_n}^{x_{n,i}+h_n} (C \cdot 2^d) \cdot \varepsilon \cdot h_n^{d-1/2}\, dx =$$

$$(2 \cdot h_n) \cdot [(C \cdot 2^d) \cdot \varepsilon \cdot h_n^{d-1/2}] = (C \cdot 2^{d+1}) \cdot \varepsilon \cdot h_n^{d+1/2}. \tag{66}$$

Resulting bounds on the values of the fuzzy transform $R_{n,i}(x)$. Substituting the expression (60) for the functions $f_{n,i,j}(x)$ from the orthonomal basis into the formula (45), we conclude that

$$R_{n,i}(x) = \sum_{j=0}^{d} f_{n,i,j} \cdot \frac{1}{\sqrt{h_n}} \cdot \varphi_j\left(\frac{x - x_{n,i}}{h_n}\right). \tag{67}$$

Since the absolute value of the sum cannot exceed the sum of the absolute values, we get

$$|R_{n,i}(x)| \le \sum_{j=0}^{d} |f_{n,i,j}| \cdot \frac{1}{\sqrt{h_n}} \cdot \left|\varphi_j\left(\frac{x - x_{n,i}}{h_n}\right)\right|. \tag{68}$$

We know the bounds (66) on the absolute values $|f_{n,i,j}|$ of the coefficients, we have denoted by C the bound on the absolute value of $|\varphi_j(x)|$. So, we can get the bound for each term in the sum (68):

$$|f_{n,i,j}| \cdot \frac{1}{\sqrt{h_n}} \cdot \left|\varphi_j\left(\frac{x - x_{n,i}}{h_n}\right)\right| \le [(C \cdot 2^{d+1}) \cdot \varepsilon \cdot h_n^{d+1/2}] \cdot \frac{1}{\sqrt{h_n}} \cdot C = (C^2 \cdot 2^{d+1}) \cdot \varepsilon \cdot h_n^d. \tag{69}$$

To get the bound for the whole sum (68), we can now multiply this bound by the number $(d+1)$ of terms in this sum:

$$|R_{n,i}(x)| \le (d+1) \cdot (C^2 \cdot 2^{d+1}) \cdot \varepsilon \cdot h_n^d = (C^2 \cdot 2^{d+1} \cdot (d+1)) \cdot \varepsilon \cdot h_n^d. \quad (70)$$

In particular, for $x = x_0$, we have

$$|R_{n,i}(x_0)| = |f(x_0) - F_{n,i(n,x_0)}(x_0)| \le (C^2 \cdot 2^{d+1} \cdot (d+1)) \cdot \varepsilon \cdot h_n^d. \quad (71)$$

Since $h_n \sim \dfrac{1}{n}$, we get the desired asymptotic result

$$F_{n,i(n,x_0)}(x_0) = f(x_0) + o\left(\frac{1}{n^d}\right). \quad (72)$$

Resulting bounds on the derivatives of the fuzzy transform $R_{n,i}(x)$. Differentiating both sides of the formula (67), and taking into account that

$$\frac{d}{dx}\left(\frac{x - x_{n,i}}{h_n}\right) = \frac{1}{h_n}, \quad (73)$$

we conclude that

$$R'_{n,i}(x) = \sum_{j=0}^{d} f_{n,i,j} \cdot \frac{1}{\sqrt{h_n}} \cdot \frac{1}{h_n} \cdot \varphi'_j\left(\frac{x - x_{n,i}}{h_n}\right). \quad (74)$$

Similarly, for every $k \le d$, if we differentiate the formula (67) k times, we multiply by the same factor (73) each of these k times and thus, we get an expression

$$R^{(k)}_{n,i}(x) = \sum_{j=0}^{d} f_{n,i,j} \cdot \frac{1}{\sqrt{h_n}} \cdot \left(\frac{1}{h_n}\right)^k \cdot \varphi^{(k)}_j\left(\frac{x - x_{n,i}}{h_n}\right). \quad (75)$$

Let C_k denote the largest of the maxima of the absolute values $|\varphi^{(k)}_j(x)|$ of the k-th order derivatives of all $d+1$ polynomials $\varphi_j(x)$ on the interval $[-1,1]$. With this notation, we have

$$\left|\varphi^{(k)}_j\left(\frac{x - x_{n,i}}{h_n}\right)\right| \le C_k. \quad (76)$$

Since we already know the bounds (66) on $f_{n,i,j}$, so we can conclude that

$$|R^{(k)}_{n,i}(x)| \le (d+1) \cdot (C \cdot 2^{d+1}) \cdot \varepsilon \cdot h_n^{d+1/2} \cdot \frac{1}{\sqrt{h_n}} \cdot \left(\frac{1}{h_n}\right)^k \cdot C_k =$$

$$(C \cdot C_k \cdot 2^{d+1} \cdot (d+1)) \cdot \varepsilon \cdot h_n^{d-k}. \quad (77)$$

In particular, for $x = x_0$, we have

$$|R^{(k)}_{n,i}(x_0)| = |f^{(k)}(x_0) - F^{(k)}_{n,i(n,x_0)}(x_0)| \le (C \cdot C_k \cdot 2^{d+1} \cdot (d+1)) \cdot \varepsilon \cdot h_n^{d-k}. \quad (78)$$

Since $h_n \sim \frac{1}{n}$, we get the desired asymptotic result

$$F_{n,i(n,x_0)}^{(k)}(x_0) = f^{(k)}(x_0) + o\left(\frac{1}{n^{d-k}}\right). \tag{79}$$

The theorem is proven.
!!!

Acknowledgments.

This work was supported in part by the National Science Foundation grant HRD-0734825, by Grant 1 T36 GM078000-01 from the National Institutes of Health, and by Grant MSM 6198898701 from MŠMT of Czech Republic.

References

[1] T. Apostol, *Calculus*, Jon Wiley & Sons, 1967.

[2] G. Klir amd B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applicatins*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[3] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman & Hall/CRC Press, Boca Raton, Florida, 2006.

[4] I. Perfilieva, "Fuzzy transforms: a challenge to conventional transforms", *Advances in Imaging and Electron Physics*, 2007, Vol. 47, pp. 137–196.