Solvability of certain groups of time varying artificial neurons

Jan Chvalina*
Department of Mathematics
Faculty of Electrical Engineering and Communication
University of Technology
Czech Republic
chvalina@feec.vutbr.cz

Bedřich Smetana
Department of Quantitative Methods
Faculty of Military Leadership
University of Defence
Czech Republic
bedrich.smetana@unob.cz

Abstract. In the process of exploration structure of the most used artificial neural network-multilayer perceptron and functionality of artificial neuron, there were established structures of groups of artificial neurons. Currently the most interesting areas is the usage of time varying artificial neurons and their reflections in these algebraic structures. Using certain analogy with relations between structures based on certain groups of linear ordinary differential operators there is investigated access to new view point on these subjects. In this paper there is contained the solution of one classical problem-verification that the corresponding group of time is solvable.
Keywords: Neural network, linear ordinary differential operators, groups of neurons, solvable groups.

1. Introduction

Time dependence on neural networks is currently insufficiently explored. Yet, in biological systems this is an important aspect. Since forgetting or anticipation is as important as memorization. Most researchers have focused on memory as a time independent activity. However, if it was time independent as assumed in the connectionist models, healing, regrouping and creating new ideas in an unsupervised environment would not be possible. The biological nervous system's ability to process and produce spatio-temporal patterns of activity while interacting with real world environments is an attribute derived from the time dependent nature of neural spatial and temporal summation. These functional characteristics of neural activity, in conjunction with the principle of "learning" based on varying synaptic efficacy, have been incorporated into a "time dependent" neural network design. The behaviour of this network is responsive to the temporal relationship of stimulus "events" as well as the "spatial" intensity

∗. Corresponding author

pattern of the stimulation. In addition, this network has the potential of generating time varying patterns of activity. In this article time dependence of neural networks is examined when looking at the role of algebraic structures based on neural networks, with emphasis on the feed-forward neural network.

A neuron also called an artificial or formal neuron is the basic stone of the mathematical model of any neural network. Its design and functionality are derived from observation of a biological neuron that is basic building block of biological neural networks (systems) which includes the brain, spinal cord and peripheral ganglia. In case of the artificial neuron the information comes into the body of the artificial neuron via inputs that are weighted (i.e. each input can be individually multiplied with a weight). The body of the artificial neuron then sums the weighted inputs, bias and "processes" the sum with a transfer function. At the end the artificial neuron passes the processed information via outputs.

Recall that in the framework of Artificial neural networks a perceptron is a network of simple neurons called perceptrons. The basic concept of a single perceptron was introduced by Rosenblatt in the year 1958. The perceptron computes a single output from multiple real-valued inputs by forming a linear combination according to its input weights and then possibly putting the output through some nonlinear activation function. As usually mathematically this can be written as:

$$y = \varphi\left(\sum_{i=1}^{n} w_i x_i + b\right) = \varphi(\vec{w}^T \vec{x} + b),$$

where $\vec{w} = (w_1, \ldots, w_n)$ denotes the vector of weights, $\vec{x} = (x_1, \ldots, x_n)$ is the vector of inputs, b is the bias and φ is the activation function.

This model explains the function of a neural network, but there is an aspect of memory. The simplest way to insert memory into the model is to add registers which store consecutive values of the model input (Mozer, 1993). These registers can be implemented as tapped delay lines, which are fed with new value at each time increment, and can store up to k past values of the input. This corresponds to making the delay coordinate reconstruction from the time series. Traditional way of using neural networks in TSP is to convert the temporal sequence into concatenated vector via a tapped delay line, and to feed the resulting vector as an input to a network. Using the tapped delay line delay line memory the time dimension is converted into spatial representation.

In this reconstruction, neural network approximates the mapping between the input vectors and the desired output. Often the output is one-step prediction of the time series. In this function approximation, the contextual information between the consecutive input vectors is lost. Also, the time difference between consecutive samples is implicitly assumed to be equal in all input vectors. In many cases, this reconstruction can capture the essential information from the sequence, allowing to build models to predict it. Especially, if the series is stationary, the reconstruction is often successful. When the series is non-stationary, it can be difficult to determine a proper window length. Selecting the order of the delay coordinate embedding, or the length of the input vector, is usually carried out by using the model selection methods. However, the optimal window length for the model may vary in the time series, and thus the windowing should be adaptive. This architecture is known as Time delay neural network (TDNN). There are neural networks which employ the delay line memory include Time delay neural network (TDNN) (Waibel et al., 1989; Lang et al., 1990), Finite impulse response neural network (FIR NN) (Wan, 1993; Back and Tsoi, 1991), and Gamma memory neural network (de Vries and Principe, 1992). In the Gamma memory neural network, the memory is built from FIR filters, which include local feedback connections, thus making it a hybrid architecture, utilizing both delay lines and recurrent connections. For more information about artificial neural networks and artificial neurons there are sources to see like [1, 18, 20, 21, 22, 23, 24, 29, 33, 34, 36, 37, 38, 39, 41].

2. Linear differential operators of the n-th order and their groups

The substanding representative of the Brno school of differential equations František Neuman wrote in his paper [25]: "Algebraic, topological and geometrical tools together with the methods of the theory of dynamical systems and functional equations make possible to deal with problems concerning global properties of solutions by contrast to the previous local investigations and isolated results." Influence of mentioned ideas is a certain motivating factor of our investigations.

Thus, we consider linear ordinary differential operators of the form:

$$L_n = \sum_{k=0}^{n} p_k(x) D^k,$$

where $D_k = \frac{d^k}{dx^k}$, $p_k(x)$ is a continuous function on some open interval $J \subset \mathbb{R}$, $k = 0, 1, \dots, n-1$, $p_n(x) \equiv 1$, i.e. $L_n(y) = 0$ which is a linear homogeneous ordinary differential equation of the form:

$$y^{(n)}(x) + \sum_{k=0}^{n} p_k(x) y^{(k)}(x) = 0.$$

By an ordered group we mean (as usually) a triad $(G, \cdot, \leq)$, where $(G, \cdot)$ is a group and $\leq$ is a reflexive, symmetrical and transitive binary relation on the set G such that for any triad $x, y, z \in G$ with the property $x \leq y$ also $x \cdot z \leq y \cdot z$, $z \cdot x \leq z \cdot y$ is satisfied. Further, $[a)_{\leq} = \{x \in G; a \leq x\}$ is the principal end generated by $a \in G$. To any element $a \in G$ there is assigned a pair of mappings $\lambda_a : G \to G, \rho_a : G \to G$, which are called a left translation, a right translation, respectively, determined by the element $a \in G$, i.e. $\lambda_a(x) = a \cdot x$, $\rho_a(x) = x \cdot a$.(Of course, in the case of a commutative group $(G, \cdot)$ we have $\lambda_a \equiv \rho_a$). Notice that a group with an ordering $(G, \cdot, \leq)$ is an ordered group if and only if all its left and right translations $\lambda_a, \rho_a, a \in G$ are order-preserving mappings, i.e. isotone selfmaps of the ordered set $(G, \leq)$.

The following lemma which is crucial for further constructions is proved in [8, 9, 10] (the Czech version is proved in [9], pp. 146, 147).

Lemma 2.1. *Let $(G, \cdot, \leq)$ be an ordered group. Define a hyperoperation $* : G \times G \to \mathcal{P}(G)^*$ by*

$$a * b = [a, b)_{\leq} (= \{x \in G; a \cdot b \leq x\}),$$

for all pairs of elements $a, b \in G$. Then $(G, *)$ is a hypergroup which is commutative if and only if the group $(G, \cdot)$ is commutative.

Application of the above lemma and many new results are obtained in papers of Michal Novák. See at least titles [26, 27, 28].

To present following results we use similar notation published in [9]. So there $\mathbb{R}$ stands for the set of all reals, $J \subset \mathbb{R}$ is an open interval (bounded or unbounded) of real numbers, $\mathbb{C}^k(J)$ is the ring (with respect to usual addition and multiplication of functions) of all real functions with continuous derivatives up to the order $k \geq 0$ including. We write $\mathbb{C}(J)$ instead of $\mathbb{C}^0(J)$. For a positive integer $n \geq 2$ we denote by $\mathbb{A}_n$ the set of all linear homogeneous differential equations of the n-th order with continuous real coefficients on J, i.e.

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_0(x)y = 0,$$

(cf. [8, 10], where $p_k \in \mathbb{C}(J)$, $k = 0, 1, \cdots, n-1$, $p_0(x) > 0$ for any $x \in J$ (this is not essential restriction). Denote $L(p_0, \cdots, p_{n-1}) : \mathbb{C}^n(J) \to \mathbb{C}^n(J)$ the above mentioned linear operator defined by

$$L(p_0, \cdots, p_{n-1})(y) = y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_0(x)y$$

and put

$$\mathbb{LA}_n(J) = \{L(p_0, \cdots, p_{n-1}); p_k \in \mathbb{C}(J),\ p_0 > 0\}.$$

Further $\mathbb{N}_0(n) = \{0, 1, \cdots, n-1\}$ and δ_{ij} stands for the Kronecker δ, $\overline{\delta_{ij}} = 1 - \delta_{ij}$. For any $m \in \mathbb{N}_0(n)$ we denote by $\mathbb{LA}_n(J)_m$ the set of all linear differential operators of the $n-$th order $L_0(p_0, \cdots, p_{n-1}) : \mathbb{C}^n(J) \to \mathbb{C}(J)$, where $p_k \in \mathbb{C}(J)$ for any $k \in \mathbb{N}_0(n)$, $p_m \in \mathbb{C}_+(J)$, (i.e. $p_m(x) > 0$ for each $x \in J$). Using the vector notation $\vec{p}(x) = (p_0(x), \cdots, p_{n-1}(x))$, $x \in J$ we can write $L_n(\vec{p}_0)y = y^{(n)} + (\vec{p}(x), (y, y', \cdots, y^{(n-1)}))$, (i.e. a scalar product).

We define a binary operation "$\circ_m$" and a binary relation "$\leq_m$ on the set $\mathbb{LA}_n(J)_m$ in this way:

For arbitrary pair $L(\vec{p}), L(\vec{q}) \in \mathbb{LA}_n(J)_m$, $\vec{p} = (p_0, \cdots, p_{n-1})$, $\vec{q} = (q_0, \cdots, q_{n-1})$ we put $L(\vec{p}) \circ_m L(\vec{q}) = L(\vec{u})$, $\vec{u} = (u_0, \cdots, u_{n-1})$, where

$$u_k(x) = p_m(x)q_k(x) + (1 - \delta_{km})p_k(x),\ x \in J$$

and $L(\vec{p}) \leq L(\vec{q})$ whenever $p_k(x) \leq q_k(x)$, $k \in \mathbb{N}_0(n)$, $p_m(x) = q_m(x)$, $x \in J$. Evidently, $(\mathbb{LA}_n(J)_m, \leq_m)$ is an ordered set.

In paper [11, 13] there is presented the sketch of the proof of the following lemma:

Lemma 2. *The triad* $(\mathbb{LA}_n(J)_m, \circ_m, \leq_m)$ *is an ordered (noncommutative) group.*

3. Group of artificial neurons

In what follows we will construct a group and hypergroup of artificial neurons using the above mentioned approach. So, recall the well-known mathematical description of a formal neuron:

Let $F : \mathbb{R} \to \mathbb{R}$ be a transfer function. Then the action of a neuron can be expressed as this model:

$$y(k) = F\left(\sum_{i=1}^{n} w_i(k)x_i(k) + b\right),$$

where $x_i(k)$ is input value in discrete time k where i goes from 0 to n, $w_i(k)$ is weight value in discrete time where i goes from 0 to m, b is bias, $y_i(k)$ is output value in discrete time k.

Notice that in some very special cases the transfer function F can be also linear. Transfer function defines the properties of artificial neuron and can be any mathematical function. Usually it is chosen on the basis of problem that artificial neuron (artificial neural network) needs to solve and in most cases it is taken (as mentioned above) from the following set of functions: step function, linear function and non-linear (sigmoid) function.

In what follows we will consider a certain generalization of classical artificial neurons mentioned above consisting in such a way that inputs x_i and weight w_i will be functions of an argument t belonging into a linearly ordered (tempus) set T with the least element 0. As the index set we use the set $\mathbb{C}(J)$ of all continuous functions defined on an open interval $J \subset \mathbb{R}$. So, denote by W the set of all non-negative functions $w : T \to \mathbb{R}$ forming a subsemiring of the ring of all real functions of one real variable $x : \mathbb{R} \to \mathbb{R}$. Denote by $Ne(\vec{w}_r) = Ne(w_{r1}, \ldots, w_{rn})$ for $r \in \mathbb{C}(J)$, $n \in \mathbb{N}$ the mapping

$$y_r(t) = \sum_{k=1}^{n} w_{r,k}(t)x_{r,k}(t) + b_r,$$

which will be called the artificial neuron with the bias $b_r \in \mathbb{R}$. By $\mathbb{AN}(T)$ we denote the collection of all such artificial neurons.

Neurons are usually denoted by capital letters X, Y or X_i, Y_i, nevertheless we use also notion $Ne(\vec{w})$, where $\vec{w} = (w_1, \ldots, w_n)$ is the vector of weights.

We suppose - for the sake of simplicity - that transfer functions (activation functions) φ, σ(or f) are the same for all neurons from the collection $\mathbb{AN}(T)$ or the role of this function plays the identity function $f(y) = y$.

Now, similarly as in the case of the collection of linear differential operators above, we will construct a group and hypergroup of artificial neurons.

Denote by δ_{ij} the so called Kronecker delta, $i, j \in \mathbb{N}$, i.e. $\delta_{ii} = \delta_{jj} = 1$ and $\delta_{ij} = 0$, whenever $i \neq j$.

Suppose $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}(T)$, $r, s \in \mathbb{C}(J)$, $\vec{w}_r = (w_{r,1}, \ldots, w_{r,n})$, $\vec{w}_s = (w_{s,1}, \ldots, w_{s,n}), n \in \mathbb{N}$.

Let $m \in \mathbb{N}$, $1 \leq m \leq n$ be such an integer that $w_{r,m} > 0$. We define

$$Ne(\vec{w}_r) \cdot_m Ne(\vec{w}_s) = Ne(\vec{w}_u),$$

where

$$\vec{w}_u = (w_{u,1}, \ldots, w_{u,n}) = (w_{u,1}(t), \ldots, w_{u,n}(t)),$$

$$\vec{w}_{u,k}(t) = w_{r,m}(t) w_{s,k}(t) + (1 - \delta_{m,k}) w_{r,k}(t), t \in T$$

and, of course, the neuron $Ne(\vec{w}_u)$ is defined as the mapping:

$$y_u(t) = \sum_{k=1}^{n} w_k(t) x_k(t) + b_u,\ t \in T,\ b_u = b_r b_s.$$

Further for a pair $Ne(\vec{w}_r)$, $Ne(\vec{w}_s)$ of neurons from $\mathbb{AN}(T)$ we put $Ne(\vec{w}_r) \leq_m Ne(\vec{w}_s), \vec{w}_r{=}(w_{r,1}(t), \ldots, w_{r,n}(t)), \vec{w}_s{=}(w_{s,1}(t), \ldots, w_{s,n}(t))$ if $w_{r,k}(t) \leq w_{s,k}(t)$, $k \in \mathbb{N}$, $k \neq m$ and $w_{r,m}(t) = w_{s,m}(t)$, $t \in T$ and with the same bias. Evidently $(\mathbb{AN}(T), \leq_m)$ is an ordered set. A relationship (compatibility) of the binary operation "$\cdot$" and the ordering $\leq_m$ on $\mathbb{AN}(T)$ is given by this assertion analogical to the above one.

4. Solvable groups

Lemma 4.1. *The triad* $(\mathbb{AN}(T), \cdot_m, \leq_m)$ *(algebraic structure with an ordering) is a non-commutative ordered group.*

The sketch of the proof is contained in [13].

Denoting

$$\mathbb{AN}(T)_m = \{Ne(\vec{w}); \vec{w} = (w_1, \ldots, w_n),\ w_k \in \mathbb{C}(T),\ k = 1, \ldots, n,\ w_m(t) \equiv 1\},$$

we get the following assertion:

Proposition 4.1. *Let* $T = [0, t_0) \subset \mathbb{R}$, $t_0 \in \mathbb{R} \cup \{\infty\}$. *Then for any positive integer* $n \in \mathbb{N}$, $n \geq 2$ *and for any integer* m *such that* $1 \leq m \leq n$ *the semigroup* $(\mathbb{AN}_1(T)_m, \cdot_m)$ *is an invariant subgroup of the group* $(\mathbb{AN}(T)_m, \cdot_m)$.

The sketch of the proof is contained in [13].

Now recall some basic concepts from the theory of group series which will be needed in the following considerations. For more details see e.g. [4, 12, 13, 16, 17, 19]. A sequence of subgroups $H_i (i = 0, 1, ..., n)$ of a group G such that

$$1 = H_0 \lhd H_1 \lhd \cdots \lhd H_{n-1} \lhd H_n = G \tag{1}$$

is said to be a subnormal series of the group G, the corresponding factor-groups $H_i/H_{i-1} (i = 1, 2, ..., n)$ are termed factors of the series $(\mathcal{H})$.

Definition 4.1 ([4, 11, 12, 19])**.** *If G is a group then its subnormal series*

$$1 = H_0 \lhd H_1 \lhd \cdots \lhd H_{n-1} \lhd H_n = G$$

is said to be solvable, if all factors $H_i/H_{i-1} (i = 1, 2, ..., n)$ are commutative thus they are abelian groups. If a group G possesses at least one solvable subnormal series then G is called a solvable group.

Remark. Evidently, an arbitrary abelian group is solvable, however there exist solvable groups which are not abelian. In the literature one can find well-known results concerning solvable groups. Let us mention some of them: First of all it is famous Feit-Thompson Theorem or odd order Theorem which states that every finite group of odd order is solvable. It was proved by Walter Feit and John Griggs Thompson (1962, 1963)-[16, 17]. Further, every finite group of order $|G| \leq 100$ except 60 is solvable. (For this information and for other details both authors are very obliged to Professor Václav Havel.) Any p-group G, i.e. $|G| = p^n$ for some prime number p is solvable. If $K \lhd G$, then the group G is solvable if and only if both groups K and G/K are solvable. A group is solvable if and only if it possesses a subnormal series all factors of which are solvable. In particular a direct product of a finite number of solvable groups is a solvable group. Of course, any finite group G of the degree 5 and higher is not solvable. Other important information can be found in the literature [4, 12, 13, 16, 17, 19]. Recall that a subgroup $G^{'}$ of a group G generated by the set of commutators

$$[a, b] = a^{-1}b^{-1}ab$$

of all pairs $[a, b] \in G \times G$ is called the commutant of the group G. The commutant $G^{''}$ of the group $G^{'}$ is called the second commutant of the group G. The usual notation is $G^{'} = G^{(1)}$, $G^{''} = G^{(2)}$, etc. In general $G^{(n+1)} = (G^{(n)})^{'}$. The commutant $G^{(n)}$ is also termed as the n-th derivative of the group G and the chain of commutants of the group G

$$G = G^{(0)} \supset G^{(1)} \supset G^{(2)} \supset \dots$$

is also called a derived chain of the group G.

Theorem 4.1 ([11], [30], Theorem 5., p.117, 2)**.** *Let G be a group. The following conditions are equivalent:*

(i) *The group G is solvable.*

(ii) *There exists a positive integer n such that $G^{(n)} = 1$.*

(iii) *There exists a subnormal series of the group G, which is solvable.*

Now we can use reverse approach. For reasons that will be apparent from the next interpretation, we will now introduce the following designation:

Example. We will present a concrete example. Let's name the structure $[\mathbb{C}(J)]^3 = \mathbb{C}(J)\times\mathbb{C}(J)\times\mathbb{C}(J)$ as a functional cube, then denote group $([\mathbb{C}(J)]^3_2, \circ_2)$, with elements $\vec{p}(x) \in [\mathbb{C}(J)]^3$, $\vec{p}(x) = (p_0(x), p_1(x), p_2(x))$,
$\vec{q}(x) = (q_0(x), q_1(x), q_2(x))$, then

$$[\mathbb{C}(J)_2]^3 = \{\vec{p}(x) = (p_0(x), p_1(x), p_2(x)); p_k \in \mathbb{C}(J), p_2(x) \neq 0, x \in J\},$$

Define a binary operation $\circ_2 : [\mathbb{C}(J)_2]^3 \times [\mathbb{C}(J)_2]^3 \to [\mathbb{C}(J)_2]^3$ by:

$$\vec{p}(x) \circ_2 \vec{q}(x) = (p_0(x), p_1(x), p_2(x)) \circ_2 (q_0(x), q_1(x), q_2(x))$$

$$= \vec{w}(p_2(x)q_0(x) + p_0(x), p_2(x)q_1(x) + p_1(x), p_2(x)q_2(x)),$$

$$\text{then } (\vec{p}(x))^{-1} = (p_0(x), p_1(x), p_2(x))^{-1} = \left(-\frac{p_0(x)}{p_2(x)}, -\frac{p_1(x)}{p_2(x)}, \frac{1}{p_2(x)}\right) = \vec{v}(x).$$

These formulation with analogy of the operators have sufficiently general form showing the way to solving of the above mentioned problem in quite generality. Nevertheless, solving of that problem (i.e. the question of solvability of the group $(\mathbb{C}(J)_n, \circ_m)$ seems to be open up to now.

By the above considerations we have that $[\mathbb{C}(J)_2, \circ_2]^3$ is a non-commutative group with the neural element $\vec{e}(x) = (0, 0, 1)$ (here $\vec{e}(x) \circ_2 \vec{p}(x) = \vec{p}(x)$ for any element $\vec{p}(x) \in \mathbb{C}(J)_2]^3$. The inverse element to $\vec{p}(x) = (p_0(x), p_1(x), p_2(x))$ is given above. It is obvious that

$$\vec{p}(x) \circ_2 (\vec{p}(x))^{-1} = (p_0(x), p_1(x), p_2(x)) \circ_2 \left(-\frac{p_0(x)}{p_2(x)}, -\frac{p_1(x)}{p_2(x)}, \frac{1}{p_2(x)}\right)$$
$$= \left(-\frac{p_2(x)p_0(x)}{p_2(x)} + p_0(x), -\frac{p_2(x)p_1(x)}{p_2(x)} + p_1(x), \frac{p_2(x)}{p_2(x)}\right) = (0, 0, 1) = \vec{e}(x).$$

and similarly we get the neutral element from the product of the above vectors from the functional cube in the opposite order. Denote

$$[\mathbb{C}_{2C}(J)_2]^3 = \{(p_0(x), p_1(x), r); p_0, p_1 \in \mathbb{C}(J)\}, r \in R, r \neq 0$$

and

$$[\mathbb{C}_{11}(J)_2]^3 = \{(p_0(x), p_1(x), 1); p_0, p_1 \in \mathbb{C}(J)\}.$$

Then, we obtain

Theorem 4.2. *Let $J \subseteq \mathbb{R}$ be an open interval. The grupoids* $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$, $([\mathbb{C}_{2C}(J)_2]^3, \circ_2)$ *are subgroups of the group* $([\mathbb{C}(J)_2]^3, \circ_2)$ *- the first from them is commutative - and we have*

$$([\mathbb{C}_{21}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}_{2C}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}(J)_2]^3, \circ_2).$$

Proof of Theorem 4.2. Evidently

$$(0,0,1) \in [\mathbb{C}_{21}(J)_2]^3 \subseteq [\mathbb{C}_{2C}(J)_2]^3.$$

If $(p_0(x), p_1(x), r)$, $(q_0(x), q_1(x), s) \in [\mathbb{C}_{2C}(J)_2, \circ_2]^3$ (with $r, s \in \mathbb{R}, r \neq 0 \neq s$) are vectors of continuous functions, then we have:

$$(p_0(x), p_1(x), r) \circ_2 (q_0(x), q_1(x), s)^{-1} = (p_0(x), p_1(x), r) \circ_2 \left(-\frac{q_0(x)}{s}, -\frac{q_1(x)}{s}, \frac{1}{s}\right)$$

$$= \left(p_0(x) - \frac{r}{s}q_0(x), p_1(x) - \frac{r}{s}q_1(x), \frac{r}{s}\right) \in [\mathbb{C}_{2C}(J)_2]^3.$$

Similarly

$$(p_0(x), p_1(x), 1) \circ_2 (q_0(x), q_1(x), 1) \in [\mathbb{C}_{2C}(J)_2]^3.$$

Thus $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$, $([\mathbb{C}_{2C}(J)_2]^3, \circ_2)$ are subgroups of the group $([\mathbb{C}(J)_2]^3, \circ_2)$. Further,

$$(p_0(x), p_1(x), 1) \circ_2 (q_0(x), q_1(x), 1) = (p_0(x) + q_0(x), p_1(x) + q_1(x), 1)$$

$$= (q_0(x), q_1(x), 1) \circ_2 (p_0(x), p_1(x), 1),$$

for arbitrary vectors of continuous functions $(p_0(x), p_1(x), 1)$, $(q_0(x), q_1(x), 1) \in [\mathbb{C}_{21}(J)_2]^3$, hence the group $([\mathbb{C}_{11}(J)_2]^3, \circ_2)$ is commutative.

Now, for any vectors $(p_0(x), p_1(x), p_2(x)) \in [\mathbb{C}(J)_2]^3$ and $(q_0(x), q_1(x), r) \in [\mathbb{C}_{2C}(J)_2]^3$ we obtain:

$$\begin{aligned}
&(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 (q_0(x), q_1(x), r) \circ_2 (p_0(x), p_1(x), p_2(x)) \\
&= \left(-\frac{p_0(x)}{p_2(x)}, -\frac{p_1(x)}{p_2(x)}, \frac{1}{p_2(x)}\right) \circ_2 (rp_0(x) + q_0(x), rp_1(x) + q_1(x), rp_2(x)) \\
&= (\varphi_0(x), \varphi_1(x), r) \text{ where } \varphi_k(x) = r(p_k(x) + q_k(x)),\ k = 0, 1.
\end{aligned}$$

Since $(\varphi_0(x), \varphi_1(x), r) \in [\mathbb{C}_{2C}(J)_2]^3$, we have

$$(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 [\mathbb{C}_{2C}(J)_2]^3 \circ_2 (p_0(x), p_1(x), p_2(x)) \subseteq [\mathbb{C}_{2C}(J)_2]^3,$$

consequently $([\mathbb{C}_{2C}(J)_2]^3, \circ_2)$ is a normal subgroup of the group $([\mathbb{C}(J)_2]^3, \circ_2))$.

In a similar way we obtain that

$$(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 [\mathbb{C}_{21}(J)_2]^3 \circ_2 (p_0(x), p_1(x), p_2(x)) \subseteq [\mathbb{C}_{21}(J)_2]^3$$

for an arbitrary vector $(p_0(x), p_1(x), r) \in [\mathbb{C}_{2C}(J)_2]^3$. Consequently

$$([\mathbb{C}_{21}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}_{2C}(J)_2]^3, \circ_2) \lhd ([\mathbb{C}(J)_2]^3, \circ_2).$$

□

Theorem 4.3. *Let $J \subseteq \mathbb{R}$ be an open interval. The group $([\mathbb{C}(J)_2]^3, \circ_2)$ of all vectors of continuos functions is solvable.*

Proof of Theorem 4.3. Denote $G = ([\mathbb{C}(J)_2]^3, \circ_2)$. The first derivate $G^{'}$, i.e. the first commutant of the group G is its subgroup generated by all commutators of the form

$$(p_0(x), p_1(x), p_2(x))^{-1} \circ_2 (q_0(x), q_1(x), q_2(x))^{-1}$$
$$\circ_2 (p_0(x), p_1(x), p_2(x)) \circ_2 (q_0(x), q_1(x), q_2(x)),$$

where $(p_0(x), p_1(x), p_2(x))$, $(q_0(x), q_1(x), q_2(x)) \in [\mathbb{C}(J)_2]^3$. Since the above commutator

$$\begin{aligned}
&(\vec{p}(x))^{-1} \circ_2 (\vec{q}(x))^{-1} \circ_2 \vec{p}(x) \circ_2 \vec{q}(x) \\
&= \left(-\frac{p_0(x)q_2(x) + p_0(x)}{p_2(x)q_2(x)}, -\frac{p_1(x)q_2(x) + q_1(x)}{p_2(x)q_2(x)}, \frac{1}{p_2(x)q_2(x)} \right) \\
&\circ_2 (p_2(x)q_0(x) + p_0(x), p_2(x)q_1(x) + p_1(x), p_2(x)q_2(x)) \\
&= \left(\frac{q_0(x)(p_2(x) - 1) + p_0(x)(1 - q_2(x))}{p_2(x)q_2(x)}, \frac{q_1(x)(p_2(x) - 1) + p_1(x)(1 - q_2(x))}{p_2(x)q_2(x)}, 1 \right)
\end{aligned}$$

belongs into the group $([\mathbb{C}_{21}(J)_2]^3, \circ_2)$, we have that $G^{'}$ is a subgroup of the last group above.

Consider an arbitrary pair of vectors $(u_0(x), u_1(x), 1)$, $(v_0(x), v_1(x), 1) \in ([\mathbb{C}_{21}(J)_2]^3, \circ_2)$. The second derivative of G, i.e. the second commutant $G^{''}$ of the group $G = ([\mathbb{C}(J)_2]^3, \circ_2)$ is its subgroup generated by the set of all commutators

$$(u_0(x), u_1(x), 1)^{-1} \circ_2 (v_0(x), v_1(x), 1)^{-1} \circ_2 (u_0(x), u_1(x), 1) \circ_2 (v_0(x), v_1(x), 1)$$

$$= (-u_0(x), -u_1(x), 1) \circ_2 (-v_0(x), -v_1(x), 1) \circ_2 (u_0(x) + v_0(x), u_1(x) + v_1(x), 1) =$$

$$(-(u_0(x) + v_0(x)), -(u_1(x) + v_1(x)), 1) \circ_2 (u_0(x) + v_0(x), u_1(x) + v_1(x), 1) =$$

$$= (0, 0, 1)$$

hence $G^{''} = \{(0, 0, 1)\}$ which is trivial subgroup (formed by the unit only) of the group $G = ([\mathbb{C}(J)_2]^3, \circ_2)$. Consequently

$$\{(0, 0, 1)\} = G^{''} \subset G^{'} \subset G^{(0)} = G,$$

therefore the group $([\mathbb{C}(J)_2]^3, \circ_2)$. is solvable.

Notice that to decide whether the group $G^{'}$ is equal to $([\mathbb{C}_{11}(J)_2]^3, \circ_2)$ or not needs some other calculations. This consists in the question whether an arbitrary continuous function $f \in \mathbb{C}$ can be expressed in the form of coefficients of the operator

$$L^{-1}(\vec{p}(x)) \circ_2 L^{-1}(\vec{q}(x)) \circ_2 L(\vec{p}(x)) \circ_2 L(\vec{q}(x)).$$

□

Theorem 4.4. *The above defined group* $(\mathbb{AN}(T), \bullet_m)$ *of artificial neurons is solvable.*

Proof of Theorem 4.4. Let us denote $G = (\mathbb{AN}(T), \bullet_m)$. The first derivative G' of the group G, which means the first commutant of the group G is its subgroup generated by all commutators of the form

$$Ne^{-1}(\vec{w}_r) \bullet_m Ne^{-1}(\vec{w}_s) \bullet_m Ne(\vec{w}_r) \bullet_m Ne(\vec{w}_s),$$

where $Ne(\vec{w}_r)$, $Ne(\vec{w}_s) \in \mathbb{AN}(T)$. Since the mentioned commutator is of the form

$$\begin{aligned}
&Ne\left(-\frac{w_{r,1}}{w_{r,m}}, \dots, \frac{1}{w_{r,m}}, \dots, -\frac{w_{r,n}}{w_{r,m}}\right) \bullet_m Ne\left(-\frac{w_{s,1}}{w_{s,m}}, \dots, \frac{1}{w_{s,m}}, \dots, -\frac{w_{s,n}}{w_{s,m}}\right) \bullet_m \\
&Ne\left(w_{r,m}w_{s,1} + w_{r,1}, \dots, w_{r,m}w_{s,m}, \dots, w_{r,m}w_{s,n} + w_{r,n}\right) \\
&= Ne\left(-\frac{w_{s,m}w_{r,1} + w_{s,1}}{w_{r,m}w_{s,m}}, \dots, \frac{1}{w_{r,m}w_{s,m}}, \dots, -\frac{w_{s,m}w_{r,n} + w_{s,n}}{w_{r,m}w_{s,m}}\right) \bullet_m \\
&Ne\left(w_{r,m}w_{s,1} + w_{r,1}, \dots, w_{r,m}w_{s,m}, \dots, w_{r,m}w_{s,n} + w_{r,n}\right) \\
&= Ne\left(\frac{(w_{r,m} - 1)w_{s,1} + (1 - w_{s,m})w_{r,1}}{w_{r,m}w_{s,m}}, \dots, 1, \right. \\
&\qquad\qquad \left. \dots, \frac{(w_{r,m} - 1)w_{s,n} + (1 - w_{s,m})w_{r,n}}{w_{r,m}w_{s,m}}\right) \\
&\in \mathbb{AN}_1(T)_m.
\end{aligned}$$

We obtain that the subgroups of the group $(\mathbb{AN}(T), \bullet_m)$ generated by all commutators of the above described form is just the group $(\mathbb{AN}_1(T)_m, \bullet_m)$. Indeed, suppose that k is a positive integer, $1 \le k \le n$, $k \ne m$ and $f \in \mathbb{C}(T)$ is an arbitrary function then choosing $w_{r,m}(t) = 1$, $w_{s,m}(t) = t^2 + 1$ and $w_{r,k}(t) = -\frac{f(t)}{t^2}(t^2 + 1)$, we obtain

$$\begin{aligned}
&\frac{1}{w_{r,m}(t)w_{s,m}(t)}\left((w_{r,m}(t) - 1)w_{s,k}(t) + (1 - w_{s,m}(t)w_{r,k}(t)\right) \\
&\qquad = \frac{-t^2 w_{r,k}(t)}{t^2 + 1} = f(t),\ t \in T.
\end{aligned}$$

Thus an arbitrary continuous weight from the above used neuron can be expressed in the described form. Therefore $G' = (\mathbb{AN}_i(T)_m, \bullet_m)$. Furthermore consider an arbitrary pair of neurons

$$Ne(w_{u,1}, \dots, 1, \dots, w_{u,n}),\ Ne(w_{v,1}, \dots, 1, \dots, w_{v,n}) \in \mathbb{AN}_1(T)_m.$$

Under a consideration which is similar as the above one we have:

$$\begin{aligned}
&Ne^{-1}(w_{u,1}, \dots, 1, \dots, w_{u,n}) \bullet_m Ne^{-1}(w_{v,1}, \dots, 1, \dots, w_{v,n}) \bullet_m \\
&Ne(w_{u,1}, \dots, 1, \dots, w_{u,n}) \bullet_m Ne(w_{v,1}, \dots, 1, \dots, w_{v,n}) \\
&= Ne(-w_{u,1}, \dots, 1, \dots, -w_{u,n}) \bullet_m Ne(-w_{v,1}, \dots, 1, \dots, -w_{v,n}) \bullet_m
\end{aligned}$$

$$\begin{aligned}&Ne(w_{u,1}+w_{v,1},\dots,1,\dots,w_{u,n}+w_{v,n})\\&=Ne(-(w_{u,1}+w_{v,1}),\dots,1,\dots,-(w_{u,n}+w_{v,n}))\bullet_m\\&Ne(w_{u,1}+w_{v,1},\dots,1,\dots,w_{u,n}+w_{v,n})=Ne(0,\dots,0,1,0,\dots,0),\end{aligned}$$

thus the second derivative G'' is the trivial group, which contains only the identity.

Therefore, we have obtained as the result

$$\{Ne(0,\dots,0,1,0,\dots,0)\}=G''\subseteq G'=G^{(0)}=G=(\mathbb{AN}(T),\bullet_m),$$

hence the group $(\mathbb{AN}(T),\bullet_m)$ is solvable.

5. Conclusion

The treated group of artificial (formal) neurons is constructed on formal similarity with the group of ordinary linear n-th order operators which form left hand sides of corresponding linear homogeneous differential equations. In particular, using the theory of series of groups motivated by the Galois theory of solvability of algebraic equations and the modern theory of extensions of fields, we have obtained that the considered group of artificial neurons is solvable. It is to be noted that there exists the large and deep Galois theory of linear differential equations-contained e.g. in the book [31] and in other papers of authors of the just mentioned monograph, but results obtained in our contribution seems to be independent on the mentioned theory. Moreover, in paper [13] there is defined (so called) a differential time varying neuron, which can serve for construction of the solution space of certain differential equation. Investigation of its properties seems to be the task of a further analysis.

References

[1] S. Behnke, *Hierarchical neural networks for image interpretation*, Lecture Notes in Computer Science, vol. 2766, Springer, Heidelberg (2003), 245.

[2] J. Beránek, J. Chvalina, *Structural properties of certain groups of solution spaces of third order linear homogeneous differential equations*, In Acta Mathematica 15, Faculty of Natural Sciences, Constantine the Philosopher University, Nitra, 2012, 43-50.

[3] J. Beránek, J. Chvalina, *Invariant subgroups of groups of ordinary second-order linear differential operators*, Fac. Nat. Sci. Constantine the Philosopher Univ. Nitra, Acta Math., 13 (2010), 43-47.

[4] J. Beránek, J. Chvalina, *Solvability of groups of linear differential operators of the n-th order*, In: Acta Math. 16, Fac. Nat. Sci., Constantine the Plilosopher Univ. Nitra, 2013, 23-30.

[5] O. Borůvka, *Foundations of the theory of groupoids and groups*, translated from the German Grundlagen der Gruppoid-und Gruppen-theorie by Milada Borůvková, F. Šik and M. Sekanina, Berlin: Deutscher Verlag der Wissenschaften, 1974, 215.

[6] O. Borůvka, *Lineare differentialtransformationen 2. Ordnung*, VEB Verlag, Berlin, 1967, extended English version Linear Differential Transformations of the Second Order, English Univ. Press, London, 1973.

[7] S. Buchholz, *A theory of neural computation with Clifford-algebras*, Technical Report Number 0504, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik (2005), 1-135.

[8] J. Chvalina, *Join spaces of linear ordinary differential operators of the second order*, Colloquium on Differential and Difference Equations, CDDE 2002, Proceedings (M. Bartušek, O. Došlý, ed.) Folia Fac. Sci. Nat. Univ. Masarykianae Brunensis, Mathematica 13, MU Brno 2003, 77-86. Masaryk Univ. Brno 1995, 305 (in Czech).

[9] J. Chvalina, *Functional graphs, quasi-ordered sets and commutative hypergroups*, Masaryk Univ. Brno, 205 (1995), (in Czech).

[10] J. Chvalina, L. Chvalinová, *Solvability of certain groups of second-order linear differential operators*, In: Proceedings Aplimat 2011, 10th Internat. Conf. FME, Slovak Univ. of Technology in Bratislava, 2011, 113-119.

[11] J. Chvalina, L. Chvalinová, *Modelling of join spaces by n- th order linear ordinary differential operators*, In: Fourth International Conference APLIMAT 2005. Bratislava: Slovak University of Technology, Bratislava, 2005, 279-284.

[12] J. Chvalina, L. Chvalinová, J. Moučka, *Solvability of a certain group of the third-order linear differential operators*, In Proc. XXX International Colloquium of the Management of Educational Process,Brno: Univ. of Defence Brno, May 17, 2012, 69-78.

[13] J. Chvalina, B. Smetana, *Groups and hypergroups of artificial neurons*, In Proceedings, 17th Conference on Aplied Mathematics Aplimat 2018, Bratislava: Slovak University of Technology in Bratislava, Faculty of Mechanical Engineering, 2018, 232-243.

[14] P. Corsini, *Prolegomena of hypergroup theory*, Aviani Editore Tricesimo, 1993.

[15] P. Corsini, V. Leoreanu, *Applications of hyperstructure theory,* Kluwer Academic Publishers, Dordrecht/Boston/London, 2003.

[16] W. Feit, J.G. Thomson, *A solvability criterion for finite groups and some consequences*, Proc. Nat. Acad. Sci. 48 (1962), 968-970.

[17] W. Feit, J.G. Thomson, *Solvability of groups of odd order*, Pacific J. Math., 13 (1963), 775-1029

[18] M. Hagan, H. Demuth, M. Beale, *Neural network design*, Boston, MA: PWS Publishing, 1996.

[19] M. Hall Jr., *The theory of groups*, The Macmillan Co., New York, N.Y., 434 (1959).

[20] V. Koudelka, Z. Raida, P. Tobola, *Simple electromagnetic modeling of small airplanes: neural network approach*, Radioengineering, 18 (2019), 38-41.

[21] T. Koskela, *Neural network methods in analysing and modelling time varying processes*, Helsinki University of Technology Laboratory of Computational Engineering Publications. Report B, 35, Department of Electrical and Communications (2003), 113.

[22] A. Krenker, J. Bešter, A. Kos, *Introduction to the artificial neural networks*, K. Suzuki (Ed.), Artificial Neural Networks-Methodological advances and biomedical applications, In: Tech, Rijeka (2011), 3-18.

[23] L. Merad, F.T. Bendimerad, S.M. Meriah, S.A. Djennas, *Neural networks for synthesis and optimization of antenna arrays*, Radioengineering, 16 (2007), 23-30.

[24] W. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bulletin of Mathematical Biophysics, 5 (1943), 115-133.

[25] F. Neuman, *Global properties of linear ordinary differential equations*, Academia Praha - Kluwer Academic Publishers Dordrecht/Boston/London 1991, 320.

[26] M. Novák, *n-ary hyperstructures constructed from binary quasi- ordered semigroups*, Analele Stiintifice Ale Universitatii Ovidius Constanta, Seria Matematica, 22 (2014), 147-168.

[27] M. Novák, *Some basic properties of EL-hyperstructures*, European Journal of Combinatorics, 34 (2013),446-459.

[28] M. Novák, *On EL-semihypergroups*, European Journal of Combinatorics, 4 (2015), 274-286.

[29] D. Pollock, M.B. Waldron, *Phase dependent output in a time varying neural net*, Proc. of Ann Conf. on EMBS, vol 11, part 6, 1989, 2054-2055.

[30] L. Procházka, L. Bican, T. Kepka, P. Němec, *Algebra*, Academia Praha, 1990.

[31] M. Put van der, M. Singer, *Galois theory of linear differential equations*, Grundlehren der Math. Wissenschaften, Vol. 328, Springer Verlag, Berlin-New York-Heidelberg, 2003.

[32] P. Račková, *Hypergroups of symmetric matrices*, 10th International Congress of Algebraic Hyperstructures and Applications, Proceedings of AHA 2008, University of Defence, Brno, (2009), 267-272.

[33] Z. Raida, Z. Lukeš, V. Otevřel, *Modeling broadband microwave structures by artificial neural networks*, Radioengineering, 13 (2004), 3-11.

[34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, *Dropout: a simple way to prevent neural networks from overfitting*, J. Machine Learning Res., 15 (2014), 1929-1958.

[35] D. Staněk, *Rough sets on state spaces of automata*, In Studen EEICT, Proceedings of the 22nd conference. Brno: BUT, FEEC, 2016, 753-757.

[36] J. Tučková, *Comparison of two approaches in the fundamental frequency control by neural nets*, Proceedings of the 6th Czech-German Workshop in Speech Processing, Praha, September, 1996, 37.

[37] J. Tučková, P. Boreš, *The neural network approach in fundamental frequency control*, Speech Processing: Forum Phoneticum (Ed.: Wodarz,H.-W.), Frankfurt am Main, Hector Verlag, 1997, 143-154.

[38] J. Tučková, V. Šebesta, *Data mining approach for prosody modelling by ANN in text-to-speech synthesis*, Proc. of the IASTED Inernational Conference on "Artificial Intelligence and Applications - AIA 2001", Ed. M.H. Hamza, ACTA Press Anaheim-Calgary-Zurich, 6th Czech-German Workshop in Speech Processing, 164-166, September 2001, Marbella, Spain.

[39] E. Volná, *Neural networks 1. 2*, ed. Ostrava: University of Ostrava, 86 (2008).

[40] M.W. Warner, *Semigroup, group quotient and homogeneous automata*, Inform. Control, 47 (1980), 59-66.

[41] M.B. Waldron, *Time varying neural networks*, Engineering in Medicine and Biology Society, 1988. Proceedings of the Annual International Conference of the IEEE.

Accepted: 25.09.2018