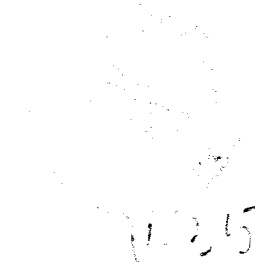


Zeszyty Naukowe Politechniki Białostockiej



MATEMATYKA FIZYKA CHEMIA

Zeszyt 20

Wydawnictwo Politechniki Białostockiej

Białystok 2001

Redaktorzy naukow
prof. dr hab. Zbigniew Bartosiewicz
prof. dr hab. Tadeusz Łukaszewicz

Recenzenci
dr hab. Piotr Grzeszczuk – 1, 2
prof. dr hab. Zbigniew Bartosiewicz – 3, 5, 6, 7, 8, 9
prof. dr hab. Grażyna Mirkowska-Salwicka – 4
prof. dr hab. Helena Puzanowska-Tarasiewicz – 10

Opracowanie redakcyjne
Jadwiga Żukowska

Skład komputerowy
Autorzy
Teresa Dawidowicz

ISSN 0208-8983

Wydawnictwo
Politechniki Białostockiej
Białystok 2001



Opracowanie i druk – Dział Wydawnictw i Poligrafii Politechniki Białostockiej
15-351 Białystok, ul. Wiejska 45C, tel. (85) 742 20 41 w. 184, 129

SPIS TREŚCI

Bartosiewicz Z., Mozyrska D.: <i>Observability of infinite-dimensional finitely presented discrete-time linear systems</i>	5
Filipowicz M., Puczyłowski E.R.: <i>On left ideals of associative rings</i>	15
Girejko E.: <i>On multidifferentials of multifunctions</i>	23
Kacprzak M.: <i>Weryfikacja temporalnych własności algorytmów niedeterministycznych</i>	37
Marchenko V., Zaczekiewicz Z.: <i>To the problem of observability for non-linear systems with time delay</i>	55
Pawłuszewicz E.: <i>Remarks on observability of Euler's discretization of control systems</i>	65
Valeev K., Janglajew K.: <i>On the dichotomy of a systems of higher order linear differential equations</i>	77
Wyrwas M.: <i>Multiobservers for nonlinear systems</i>	87
Żebrowska-Suchodolska D.: <i>Controllability of discrete-time infinite-dimensional linear systems</i>	99
Kucharski M., Łukaszewicz T., Karpowicz D., Lasocki Z.: <i>Elektrolit do wielowarstwowych urządzeń elektrochromowych</i>	107

SPIS STRESZCZEŃ

Bartosiewicz Z., Mozyrska D.: <i>Obserwowalność nieskończenie wymiarowych skończenie określonych układów liniowych z czasem dyskretnym</i>	14
Filipowicz M., Puczyłowski E.R.: <i>O lewych idealach pierścieni łącznych</i>	21
Girejko E.: <i>O multiróżniczkach multifunkcji</i>	35
Kacprzak M.: <i>Verification of temporal properties of nondeterministic algorithms</i>	53
Marchenko V., Zaczkiewicz Z.: <i>O problemie obserwowalności dla układów nieliniowych z opóźnieniem</i>	63
Pawłuszewicz E.: <i>Uwagi o obserwowalności dyskretyzacji Eulera układów sterowania</i>	76
Valeev K., Janglajew K.: <i>O dychotomii układu liniowych równań różniczkowych wyższego rzędu</i>	86
Wyrwas M.: <i>Multiobserwatory dla układów nieliniowych</i>	98
Żebrowska-Suchodolska D.: <i>Sterowalność nieskończenie wymiarowych układów liniowych z czasem dyskretnym</i>	106
Kucharski M., Łukaszewicz T., Karpowicz D., Lasocki Z.: <i>The electro-lite for electrochromic devices</i>	117

Zbigniew Bartosiewicz*

Dorota Mozyrska†

OBSERVABILITY OF INFINITE-DIMENSIONAL FINITELY PRESENTED DISCRETE-TIME LINEAR SYSTEMS

Abstract: Discrete-time infinite-dimensional linear dynamic systems, with one-dimensional output, are studied. They are described by infinite matrices that are row-finite. Observability of such systems is investigated. It may be described as the property that one can calculate the value of each state variable using only finitely many rows of observability matrix. Different conditions of observability of such systems are given.

Key-words: Infinite-dimensional system, linear discrete-time system, observability.

1 Introduction

We study here a particular type of discrete-time infinite-dimensional systems with output. They are defined on space $\mathbb{R}^{\mathbb{N}} := \prod_{i \in \mathbb{N}} \mathbb{R}$, the space of all infinite sequences of real numbers, and described by infinite matrices that are *row-finite*, i.e. rows have finitely many nonzero elements.

In [4, 5] there were presented infinite-dimensional nonlinear analytic systems with output. They were described by functions depending on a finite number of variables. The criterion of local observability of such systems was given.

One of the main reasons to study such systems is the concept of infinite dynamic extension of a control system. Such an extension of a continuous-time system appeared in works of Fliess [2], Pomet [7], Jakubczyk [3] and others.

*Supported by KBN under the Technical University of Białystok grant No. W/IMF/1/01

†Supported by KBN under the Technical University of Białystok grant No. W/IMF/2/99

By analogy, we may consider a dynamic extension of a discrete-time control systems. Let us consider a linear discrete-time control system with output

$$(\Sigma) : \begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k), \end{aligned} \quad (1.1)$$

where $k \in \mathbb{Z}$, $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}$ and $y(k) \in \mathbb{R}^p$. Then A, B, C, D are the matrices of dimensions: $n \times n$, $n \times 1$, $p \times n$, $p \times 1$. We extend the state variables by adding $u(k)$ and all its shifts. In this way we get the infinite dynamic extension of system Σ , denoted Σ_∞ . It may be written in the followig form:

$$(\Sigma_\infty) : \begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ u(k+1) &= u_1(k) \\ u_1(k+1) &= u_2(k) \\ &\vdots \\ u_k(k+1) &= u_{k+1}(k) \\ &\vdots \\ y(k) &= Cx(k) + Du(k). \end{aligned} \quad (1.2)$$

Let $\mathcal{U} = (u_0, u_1, u_2, \dots)$ be the infinite sequence of control $u = u_0$ and its shifts and $\mathcal{X} = (x, u_0, u_1, \dots) = (x, \mathcal{U})$. Then Σ_∞ is described by infinite matrices of the forms:

$$\tilde{A} = \begin{pmatrix} A & B & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & 1 & 0 & \dots \\ \mathbf{0} & \mathbf{0} & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \tilde{C} = (C \ D \ \mathbf{0} \ \dots).$$

Observe that $\tilde{A}$ and $\tilde{C}$ are row-finite. Let $\tilde{y}(k) = \tilde{C}\mathcal{X}(k)$ be the output of Σ_∞ . Then the output $\tilde{y}(k)$ corresponding to $\mathcal{X}(0)$ may be expressed in the following way:

$$\tilde{y}(k) = \tilde{C}\tilde{A}^k\mathcal{X}(0),$$

where $\tilde{A}^k = \tilde{A}^{k-1}\tilde{A}$ and the matrix $\tilde{C}\tilde{A}^k$ is row-finite too.

Another example where we may see such discrete-time infinite-dimensional systems is discretization of partial differential equations of parabolic type where infinitely many variables correspond to infinitely many discretization points of the state space (e.g. real line).

2 Linear finitely presented mappings on $\mathbb{R}^{\mathbb{N}}$

Let $\Pi_n : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^n$ denotes the projection on the first n coordinates, that is if $x = (x_1, x_2, \dots) \in \mathbb{R}^{\mathbb{N}}$ then $\Pi_n(x) = (x_1, \dots, x_n)$.

We say that a function $f : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$ is *finitely presented on $\mathbb{R}^{\mathbb{N}}$* if there exists $k \in \mathbb{N}$ and a function $\tilde{f} : \mathbb{R}^k \rightarrow \mathbb{R}$ such that $f = \tilde{f} \circ \Pi_k$. We say that a mapping $F : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ is *finitely presented on $\mathbb{R}^{\mathbb{N}}$* if every component of F is finitely presented on $\mathbb{R}^{\mathbb{N}}$.

Remark 2.1. Let $F, G : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ be finitely presented mappings on $\mathbb{R}^{\mathbb{N}}$. Then the composition $F \circ G$ is finitely presented too.

Let X be linear space. Then by X' we will denote its dual space, the space of all linear forms.

By $\mathbb{R}^{(\mathbb{N})}$ we denote the space of real sequences with finite number of nonzero elements. Then the dual space $(\mathbb{R}^{(\mathbb{N})})'$ consists of the linear mappings $x' : \mathbb{R}^{(\mathbb{N})} \rightarrow \mathbb{R}$. The following proposition is a well known fact from linear algebra.

Proposition 2.2. $(\mathbb{R}^{(\mathbb{N})})' \approx \mathbb{R}^{\mathbb{N}}$. $\square$

Corollary 2.3. Any linear map $\mathbb{R}^{(\mathbb{N})} \rightarrow \mathbb{R}^{(\mathbb{N})}$ is represented by a column-finite matrix (i.e. columns have only finitely many nonzero elements).

Proposition 2.4. Let a linear map $\bar{E} : \mathbb{R}^{(\mathbb{N})} \rightarrow \mathbb{R}^{(\mathbb{N})}$ be given by $\bar{E}(x) = Ex$ for $x \in \mathbb{R}^{(\mathbb{N})}$, where E is a column-finite matrix. Then the dual map $\bar{E}' : (\mathbb{R}^{(\mathbb{N})})' \rightarrow (\mathbb{R}^{(\mathbb{N})})'$ has the following form

$$\forall y' \in (\mathbb{R}^{(\mathbb{N})})' \quad \forall x \in \mathbb{R}^{(\mathbb{N})} \quad (\bar{E}'(y'))(x) = (E^T y)^T x,$$

where $y \in \mathbb{R}^{\mathbb{N}}$ and $\forall x \in \mathbb{R}^{(\mathbb{N})} \quad y'(x) = y^T x$.

Proof: Let $x \in \mathbb{R}^{(\mathbb{N})}$. Then

$$\bar{E}'(y')(x) = y'(\bar{E}(x)) = y'(Ex) = y^T(Ex) = y^T((E^T)^T x) = (E^T y)^T x.$$

In the last equality we use associativity of multiplication of row-finite matrices ([1]). $\square$

Proposition 2.5. $\bar{E} : \mathbb{R}^{(\mathbb{N})} \rightarrow \mathbb{R}^{(\mathbb{N})}$ given by $\bar{E}(x) = Ex$, where E is column-finite, is surjective iff there is a column-finite matrix E' such that: $EE' = I$.

Proof:

" $\Rightarrow$ "

If $\bar{E}$ is surjective then for all $e_i = (0, \dots, 1, 0, \dots)^T$, where 1 is on the i -th position, there is $E'_i \in \mathbb{R}^{(\mathbb{N})}$ such that $EE'_i = e_i$. Hence $E' = (E'_1, E'_2, \dots)$ is column-finite and $EE' = (e_1, e_2, \dots) = I$.

" $\Leftarrow$ "

Let $y \in \mathbb{R}^{(\mathbb{N})}$ and let $x = E'y$. Then $x \in \mathbb{R}^{(\mathbb{N})}$ and $Ex = E(E'y) = (EE')y = y$. Here we use associativity of multiplication of column-finite matrices. $\square$

Proposition 2.6. [4] Let X, Y be linear spaces and X', Y' their corresponding dual spaces. Let $F : X \rightarrow Y$ and $F' : Y' \rightarrow X'$.

F' is injective iff F is surjective.

Lemma 2.7. Let $\bar{D} : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$, where $\bar{D}(x) = Dx$ and D is row-finite. $\bar{D}$ is injective iff there is a row-finite matrix $'D$ such that $'DD = I$.

Proof:

" $\Rightarrow$ "

Let $E = D^T$. Observe that E is column-finite. If $\bar{D}$ is injective then, from Proposition 2.6, $\bar{E}$ is surjective. Hence from Proposition 2.5 there is a column-finite matrix E' such that $EE' = I$. Then $(E')^T E^T = I$ and we can take $'D = (E')^T$; $'D$ is row-finite.

" $\Leftarrow$ "

$Dx = 0 \Rightarrow 'D(Dx) = 0$. Then from associativity $(D'D)x = 0$, so $x = 0$. $\square$

Let D be a row-finite infinite matrix and $\bar{D} : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ be the (finitely presented linear) mapping $\bar{D}(x) = Dx$. By $D|_n$ we will denote the matrix consisting of first n rows from a matrix D . Then for $x \in \mathbb{R}^{\mathbb{N}}$ we have that $D|_n x = \Pi_n(\bar{D}(x))$.

Proposition 2.8. $\bar{D}$ is injective iff

$$\forall i \in \mathbb{N} \exists n_i \in \mathbb{N} : \forall x \in \mathbb{R}^{\mathbb{N}} (D|_{n_i} x = 0 \Rightarrow x_i = 0).$$

Proof: $\bar{D}$ is injective iff there is a row-finite $'D$ such that for $x \in \mathbb{R}^{\mathbb{N}} : 'D(Dx) = x$ iff $\forall i \in \mathbb{N} \exists n_i \in \mathbb{N} x_i = 'D_i(Dx) = (d'_1, d'_2, \dots, d'_{n_i}, 0, \dots)(Dx) = (d'_1, d'_2, \dots, d'_{n_i})(D|_{n_i} x)$. $\square$

Corollary 2.9. Let us consider an infinite system of linear equations:

$$f_k(x) = \sum_{n=1}^{\infty} a_{kn} x_n = 0, \quad (k \in \mathbb{N}), \quad (2.1)$$

where for every k only finitely many coefficients $a_{kn} \neq 0$. The system (2.1) has exactly one solution $x = (0, 0, \dots)$ if and only if for every $i \in \mathbb{N}$ there exists $m \in \mathbb{N}$ such that

$$x_i = \sum_{k=1}^m b_{ik} f_k(x)$$

for some $b_{ik} \in \mathbb{R}$.

It is easy to show the following necessary condition of injectivity.

Proposition 2.10. If $\bar{D}$ is injective then $\text{rank } D = \infty$, which means that D has nonzero minors of arbitrary order. $\square$

Example 2.11. Let us consider the following system:

$$\begin{aligned} f_1(x) &= x_1 + x_2 + x_3 = 0 \\ f_2(x) &= x_2 + x_3 = 0 \\ &\vdots \\ f_{2k-1}(x) &= x_{2k-1} + x_{2k} + x_{2k+1} = 0 \\ f_{2k}(x) &= x_{2k} + x_{2k+1} = 0 \\ &\vdots \end{aligned}$$

The above system has unique solution $x = 0$ and we can compute every x_k from finitely many functions f_i . For odd k we have that $x_k = f_k - f_{k+1}$ and for even k : $x_k = f_k - f_{k+1} + f_{k+2}$.

3 Discrete-time systems

Let us consider an infinite-dimensional finitely presented discrete-time system with one-dimensional output:

$$(\Sigma) : \begin{aligned} x(k+1) &= Ax(k) \\ y(k) &= Cx(k), \end{aligned} \quad (3.1)$$

where $k \in \mathbb{Z}$, $x(k) \in \mathbb{R}^{\mathbb{N}}$, $y(k) \in \mathbb{R}$ and A is a row-finite infinite matrix and C is an infinite row with only finitely many nonzero elements. For simplicity, we consider here one-dimensional output, but all the results may be formulated for systems with finite-dimensional output.

Let $(x(0), x(1), \dots), x(i) \in \mathbb{R}^{\mathbb{N}}$, be a trajectory of the dynamics of system Σ . Observe that it is uniquely defined by the initial condition $x(0)$. Then the output $y(k)$ corresponding to the initial condition $x(0)$ may be expressed in the following way:

$$y(k) = CA^k x(0),$$

where CA^k is an infinite row with only finitely many nonzero elements.

We say that $x, z \in \mathbb{R}^{\mathbb{N}}$ are *indistinguishable* (with respect to Σ) if for all $k \geq 0$:

$$CA^k x = CA^k z.$$

Otherwise x and z are *distinguishable*. We say that Σ is *observable* if every two points are distinguishable.

Let $D = (d_{ij}) = \begin{pmatrix} C \\ CA \\ \vdots \end{pmatrix}$ and $\bar{D} : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ be defined by $\bar{D}(x) = Dx$.

Remark 3.1. As in the finite-dimensional case, Σ is observable $\iff \bar{D}$ is injective.

Let C, A be the matrices of system Σ . Observe that for all $0 \leq k < \infty$ the rank of the matrix $\begin{pmatrix} C \\ \vdots \\ CA^k \end{pmatrix}$ is finite. Let E_i be the infinite row with 1 at

the i -th position and 0 at other positions and let $\bar{E}_i : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}$, for $x \in \mathbb{R}^{\mathbb{N}}$: $\bar{E}_i(x) = x_i$.

Theorem 3.2. *System Σ is observable iff $\forall i \in \mathbb{N} \exists k_i \geq 0$:*

$$\text{rank} \begin{pmatrix} C \\ \vdots \\ CA^{k_i} \end{pmatrix} = \text{rank} \begin{pmatrix} C \\ \vdots \\ CA^{k_i} \\ E_i \end{pmatrix}.$$

Proof: Let $D = \begin{pmatrix} C \\ CA \\ \vdots \end{pmatrix}$ and $\bar{D}(x) = Dx$. Then, from Proposition 2.8: $\bar{D}$ is injective iff $\forall i \in \mathbb{N} \exists D_i = (a_1, \dots, a_{n_i}, 0, \dots) \forall x \in \mathbb{R}^{\mathbb{N}} : x_i = D_i(Dx)$. This holds iff $\forall x \in \mathbb{R}^{\mathbb{N}} x_i = a_1 D_1 x + \dots + a_{n_i} D_{n_i} x$, where for every $j = 1, \dots, n_i$, D_j denotes the j -th row of matrix D . This is equivalent to the condition that $\forall i \in \mathbb{N} \exists a_1, \dots, a_{n_i} \in \mathbb{R} E_i = a_1 D_1 + \dots + a_{n_i} D_{n_i}$ which means that $\text{rank} \begin{pmatrix} D_{|n_i} \\ E_i \end{pmatrix} = \text{rank} (D_{|n_i})$, where $D_{|n_i} = \begin{pmatrix} D_1 \\ \vdots \\ D_{n_i} \end{pmatrix}$. Of course $D_{n_i} = CA^{n_i-1}$. $\square$

The following theorem gives a sufficient condition for observability.

Theorem 3.3. *If $\forall k \in \mathbb{N} \exists k' \geq k \exists i_1, \dots, i_{n_{k'}} \in \mathbb{N} : \forall i = i_1, \dots, i_{n_{k'}}, j > k'$*

$$d_{ij} = 0 \text{ and } \text{rank} (d_{ij})_{\substack{i = i_1, \dots, i_{n_{k'}} \\ j = 1, \dots, k'}} = k'$$

then Σ is observable.

Proof: Let $k \in \mathbb{N}$ and $k' \geq k$ be as in the assumption. Then

$$(d_{ij})_{\substack{i = i_1, \dots, i_{n_{k'}} \\ j = 1, \dots, k'}} = (\tilde{D}, \mathbf{0}, \dots),$$

where $\tilde{D}$ is a matrix of dimensions $k' \times k'$ and $\text{rank} \tilde{D} = k'$. Hence for all $x \in \mathbb{R}^{\mathbb{N}}$ we have that $(\tilde{D}, \mathbf{0}, \dots)x = \mathbf{0} \Rightarrow \forall j = 1, \dots, k' : x_j = 0$. Let take

$n_k = \max\{i_1, \dots, i_{n_k}\}$. Then $\forall x \in \mathbb{R}^{\mathbb{N}}$:

$$D|_{n_k} x = 0 \Rightarrow x_k = 0.$$

□

From the Proposition 2.10 we have the following:

Proposition 3.4. *If Σ is observable then $\text{rank } D = \infty$.*

□

4 Examples

Example 4.1. Let us consider the following infinite-dimensional system:

$$\begin{aligned} x_1(k+1) &= x_2(k) - x_1(k) \\ x_2(k+1) &= x_3(k) - x_2(k) \\ &\vdots \\ x_n(k+1) &= x_{n+1}(k) - x_n(k) \\ &\vdots \\ y(k) &= x_1(k). \end{aligned} \quad (4.1)$$

Observe that $D = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ -1 & 1 & 0 & 0 & 0 & \dots \\ 1 & -2 & 1 & 0 & 0 & \dots \\ -1 & 3 & -3 & 1 & 0 & \dots \\ \vdots & & & & & \ddots \end{pmatrix}$. And it is easy

to notice that the condition from theorem 3.3 is satisfied.

Example 4.2. Let us consider the system described by the matrices:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}, \text{ and } C = (0 \ 1 \ 0 \ \dots).$$

Then $\text{rank} \begin{pmatrix} C \\ CA \\ \dots \end{pmatrix} = \infty$ but Σ is not observable (e.g. $x_1 = (1, 0, \dots)$ and $x_2 = (0, 0, \dots)$ are indistinguishable).

Example 4.3. Let the system Σ be in the following form

$$\begin{aligned} x_{2n-1}(k+1) &= x_{2n}(k) - x_{2n+2}(k) - x_{2n+3}(k) \\ (\Sigma) : \quad x_{2n}(k+1) &= x_{2n+1}(k) + x_{2n+3}(k) + x_{2n+4}(k) + x_{2n+5}(k), \quad n \in \mathbb{N}. \\ y(k) &= x_1(k) + x_2(k) + x_3(k) \end{aligned}$$

Then the matrix $D = \begin{pmatrix} C \\ CA \\ \vdots \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \dots \\ 0 & 0 & 1 & 1 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & 1 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}$. From

Theorem 3.2 the system Σ is observable but the condition from Theorem 3.2 is not satisfied. Moreover the equation $Dx = 0$ is as in Example 2.11.

Acknowledgement. The authors wish to thank Karl Neeb for suggestions concerning the proof of Lemma 2.7.

References

- [1] Cooke R.G., *Infinite matrices and sequence spaces*, London 1950
- [2] Fliess M. et al. *On nonlinear controllability, infinite jets and prolongations*, in: Proceedings of European Control Conference ECC-97, Brussels, Belgium, July 1997.
- [3] Jakubczyk B., *Remarks on equivalence and linearization of nonlinear systems*, in: Proc. Nonlinear Control Systems Design Symposium, Bordeaux, France, 1992.
- [4] Komorowski J., *Od liczb zespolonych do tensorów, spinorów, algebr Liego i kwadryk* (in Polish), PWN, Warszawa, 1978.
- [5] Mozyrska D. and Bartosiewicz Z., *Families of germs in local observability of infinite-dimensional systems*, in: Proceedings of Conference Control and Self-Organisation in Nonlinear Systems, Białystok, Poland, February 2000.
- [6] Mozyrska D., Bartosiewicz Z., *Local observability of systems on $\mathbb{R}^\infty$* , in: Proceedings of MTNS'2000, Perpignan, France.

- [7] Pomet J.-B., *A differential geometric setting for dynamic equivalence and dynamic linearization*, in: Banach Center Publications 32 (1995), 319-339.

Authors' address: Instytut Matematyki i Fizyki, Politechnika Białostocka, Wiejska 45, 15-351 Białystok, Poland

OBSEKOWALNOŚĆ NIESKOŃCZENIE WYMIAROWYCH
SKOŃCZENIE OKREŚLONYCH
UKŁADÓW LINIOWYCH Z CZASEM DYSKRETNYM
Streszczenie

Badane są nieskończenie wymiarowe liniowe układy dynamiczne z dyskretnym czasem i z jednowymiarowym wyjściem. Układy te opisane są przez nieskończone macierze o „skończonych wierszach”. Zbadano obserwowalność takich układów. Własność ta została scharakteryzowana następująco: wartość każdej zmiennej stanu może być wyliczona na podstawie skończenie wielu wierszy macierzy obserwowalności. Pokazano kilka innych warunków na obserwowalność.

M. Filipowicz and E.R. Puczyłowski

ON LEFT IDEALS OF ASSOCIATIVE RINGS

Abstract: It is well known that the relation “ L is a left ideal of a ring R ” is not transitive. Several authors studied cases in which the transitive property does hold. In this paper we continue these studies.

It is well known that if A is a ring, B is a left ideal of A and C is a left ideal of B , then C need not be a left ideal of A . One can express this shortly saying that the concept of left ideal is not a transitive relation. The some concerns the concept of twosided ideal. However in some cases (e.g. for specific A , B or C) the transitive property does hold. Systematic studies of such cases were started by Sands in [2] and continued by Veldsman in [3]. In this paper we make some observations concerning the subject.

All rings in this paper are associative, but we do not assume that each ring has an identity element.

We use the notation $A \triangleleft B$, $A \triangleleft_l B$ to mean that A is ideal, a left ideal of the ring B .

We shall use some Morita contexts. A Morita context $\begin{pmatrix} R & V \\ W & S \end{pmatrix}$

consists of two rings R and S and two bimodules ${}_R V_S$ and ${}_S V_R$ together with mappings $V \times W \rightarrow R$ and $W \times V \rightarrow S$ (written multiplicatively) which induce bimodule homomorphism $V \otimes_S W \rightarrow R$ and $W \otimes_R V \rightarrow S$ and which satisfy the associativity conditions $(vw)v' = v(wv')$, $(wv)w' = w(vw')$ for $v, v' \in V$ and $w, w' \in W$. These conditions are equivalent to the requirement that the set $\begin{pmatrix} R & V \\ W & S \end{pmatrix}$ of all 2x2-matrices $\begin{pmatrix} r & v \\ w & s \end{pmatrix}$, $r \in R$, $v \in V$, $w \in W$, $s \in S$ forms a ring under the usual definitions of matrix addition and multiplication.

If X, Y, Z, T are subsets of R, V, W, S respectively, we shall denote by $\begin{pmatrix} X & Y \\ Z & T \end{pmatrix}$ the subset of $\begin{pmatrix} R & V \\ W & S \end{pmatrix}$ consisting of all matrices $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$ with $x \in X$, $y \in Y$, $z \in Z$ and $t \in T$.

Given a ring B , $E(B_B)$ denotes the ring of endomorphisms of the right B -module B .

1. In this section we make some remarks on the following Sands result.

Theorem 1. ([2], Theorem 3). *Let A be a left ideal of the ring B . Then the following conditions are equivalent*

- (i) $f \in E(B_B)$ implies $f(A) \subseteq A$;
- (ii) $B <_l C$ implies $A <_l C$, where C varies over all rings satisfying this condition.

Given a ring R , $r \in R$ and $L <_l R$, denote by $\varrho_{L,r}$ the map defined by $\varrho_{L,r}(x) = rx$, $x \in L$. Clearly $\varrho_{L,r} \in E(L_L)$. In case $L = R$, we denote $\varrho_{L,r}$ shortly by ϱ_r .

If C is a ring, $B <_l C$ and A is a subring of B such that $f(A) \subseteq A$ for every $f \in E(B_B)$, then for each $c \in C$, $\varrho_{B,c}(A) \subseteq A$. Consequently $CA \subseteq A$, so A is a left ideal of C .

This shows that (i) $\rightarrow$ (ii) in Theorem 1. The same proof of the implication was given in [2]. We apply different arguments to get the converse implication.

Given a subset X of a ring R , we denote by $l_R(X)$ the left annihilator of X in R , i.e., $l_r(X) = \{r \in R \mid rX = 0\}$. Instead of $l_R(R)$ we write just $l(R)$.

The following is in fact well known. We include its proof for completeness.

Proposition 2. *Suppose that R is a ring and $L <_l R$. The map $\Phi : R \rightarrow E(L)$ given by $\Phi(r) = \varrho_{L,r}$ is a ring homomorphism, $\text{Ker } \Phi = l_R(L)$ and $\Phi(L) <_l E(L_L)$.*

Proof. One easily checks that if $r_1, r_2 \in R$, then $\varrho_{L,r_1} + \varrho_{L,r_2} = \varrho_{L,r_1+r_2}$ and $\varrho_{L,r_1} \circ \varrho_{L,r_2} = \varrho_{L,r_1r_2}$, so indeed Φ is a ring homomorphism. Now $\text{Ker } \Phi = \{r \in R \mid \varrho_{L,r} = 0\} = \{r \in R \mid rL = 0\} = l_R(L)$. For every $f \in E(L_L)$ and each $l \in R$, $f \circ \varrho_{L,l} = \varrho_{L,f(l)}$, so indeed $\Phi(L) <_l E(L_L)$. $\square$

Corollary 3. *Every ring B such that $l(B) = 0$ can be embedded into $E(B_B)$ as a left ideal. Moreover if B is a left essential ideal of a ring C (i.e. for every nonzero ideal I of C , $B \cap I \neq 0$), then C is isomorphic to a subring of $E(B_B)$.*

Proof. It suffices to put in Proposition 2, $L = B$, $R = C$ and note that $l_C(B)$ is an ideal of C such that $l_C(B) \cap B = l(B)$. $\square$

Applying the first part of Corollary 3 one immediately gets (ii) $\Rightarrow$ (i) in case when $l(B) = 0$. Now we pass to the general case. Given a ring R , let

$$\begin{pmatrix} E(R_R) & R \\ 0 & 0 \end{pmatrix}$$

be the Morita context, where R is considered as a left $E(R_R)$ -module with respect to the addition of R and multiplication defined by $fr = f(r)$, $f \in E(R_R)$, $r \in R$.

Proof of (ii) $\Rightarrow$ (i) in Theorem 1. If $b, b_1 \in B$ and $f \in E(B_B)$, then

$$\begin{pmatrix} f & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \varrho_{b_1} & b_1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \varrho_{f(b_1)} & f(b_1) \\ 0 & 0 \end{pmatrix}.$$

This shows that $L = \left\{ \begin{pmatrix} \varrho_b & b \\ 0 & 0 \end{pmatrix} \mid b \in B \right\} <_l \begin{pmatrix} E(B_B) & B \\ 0 & 0 \end{pmatrix}$. One easily checks that the map $b \rightarrow \begin{pmatrix} \varrho_b & b \\ 0 & 0 \end{pmatrix}$ is a ring isomorphism of B onto L . Assuming (ii) we get that the image T of A under this isomorphism is a left ideal of $\begin{pmatrix} E(B_B) & B \\ 0 & 0 \end{pmatrix}$. Consequently for every $f \in E(B_B)$ and each $a \in A$,

$$\begin{pmatrix} f & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \varrho_a & a \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \varrho_{f(a)} & f(a) \\ 0 & 0 \end{pmatrix} \in T.$$

This implies that $f(a) \in A$, which is the condition (i). $\square$

Remark. Note that

$$I = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix} \triangleleft M = \begin{pmatrix} E(B_B) & B \\ 0 & 0 \end{pmatrix}$$

and $M/I \simeq E(B_B)$.

Moreover if $l(B) = 0$, then $\left\{ \begin{pmatrix} \varrho_b & b \\ 0 & 0 \end{pmatrix} \mid b \in B \right\} \cap I = 0$. These show that if $l(B) = 0$, then the above considered map gives an embedding of B into $E(B_B)$ as a left ideal.

To prove (ii) $\Rightarrow$ (i) Sands constructed a ring as follows. Let C consist of all pairs (b, λ) , $b \in B$, $\lambda \in E(B_B)$, where addition is defined in C pointwise and multiplication by $(b_1, \lambda_1)(b_2, \lambda_2) = (b_1b_2 + \lambda_1(b_2), b_1\lambda_2 + \lambda_1 \circ \lambda_2)$, where $(b_1\lambda_2)(x) = b_1\lambda_2(x)$ for all $x \in B$.

It turns out that this ring is isomorphic to $M = \begin{pmatrix} E(B_B) & B \\ 0 & 0 \end{pmatrix}$.

Proposition 4. *The map*

$$(b, \lambda) \longrightarrow \begin{pmatrix} \lambda + \varrho_b & b \\ 0 & 0 \end{pmatrix}$$

is an isomorphism of C onto M .

Proof. It is clear that the map is additive as well as that it is a bijection. Take (b_1, λ_1) and (b_2, λ_2) in C . We have

$$\begin{aligned} & \begin{pmatrix} \lambda_1 + \varrho_{b_1} & b_1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \lambda_2 + \varrho_{b_2} & b_2 \\ 0 & 0 \end{pmatrix} = \\ & \begin{pmatrix} (\lambda_1 + \varrho_{b_1}) \circ (\lambda_2 + \varrho_{b_2}) & \lambda_1(b_2) + \varrho_{b_1}(b_2) \\ 0 & 0 \end{pmatrix} = \\ & \begin{pmatrix} \lambda_1 \circ \lambda_2 + b_1\lambda_2 + \varrho_{\lambda_1}(b_2) + \varrho_{b_1}b_2 & \lambda_1(b_2) + b_1b_2 \\ 0 & 0 \end{pmatrix} = \\ & \begin{pmatrix} b_1\lambda_2 + \lambda_1 \circ \lambda_2 + \varrho_{b_1}b_2 + \lambda_1(b_2) & \lambda_1(b_2) + b_1b_2 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

This shows that the map is multiplicative. The result follows. $\square$

2. In [1] a twosided ideal I of a ring A was called a $*$ -ideal if for every ring B such that $A \triangleleft B$ also $I \triangleleft B$. It was proved there (Theorem 1) that a ring contains no nontrivial $*$ -ideals if and only if is either a simple idempotent ring or an algebra with zero multiplication over a field.

A left ideal L of a ring A will be called a left $*$ -ideal if for every ring B such that $A <_l B$ also $L <_l B$. In this section we shall characterize rings containing no nontrivial left $*$ -ideals.

Theorem 5. *A ring R contains no nontrivial left $*$ -ideals if and only if R is an algebra with zero multiplication over a field or R is isomorphic to the Morita context*

$$\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix}$$

where Δ is a division ring and V is a right Δ -module.

Proof. Suppose that R contains no nontrivial left $*$ -ideals. Note that R^2 is a left $*$ -ideal of R , so $R^2 = R$ or $R^2 = 0$. Suppose first that $R^2 = R$ (and $R \neq 0$). For every $r \in R$, Rr is a left $*$ -ideal of R , so for every $r \in R$, $Rr = 0$ or $Rr = R$. Let $V = \{r \in R \mid Rr = 0\}$. Since $R^2 \neq 0$, we have that $R \neq V$. Obviously for each $r \in R \setminus V$, $Rr = R$. Pick $a \in R \setminus V$. Then for some $e \in R$, $ea = a$. Observe that $L = \{xe - x \mid x \in R\}$ is a left $*$ -ideal of R . Hence $L = 0$ or R . Note however that $La = 0$, which together with the assumption $a \notin V$ force that $L = 0$. Hence e is a right identity of R (i.e., $xe = x$ for every $x \in R$). One can easily check that $V = \{x - ex \mid x \in R\}$ and R is the direct sum of right R -modules eR and V . Hence $eR \simeq R/V$. It is clear that the image of e in R/V is an identity element and since for each $r \in R \setminus V$, $Rr = R$, we have that every nonzero element of R/V is invertible from the left. It is a standard well known fact that this implies that R/V , and consequently $\Delta = eR$, is a division ring. Obviously V is in a natural way a right Δ -module. Every element $r \in R$ can be uniquely presented in the form $r = r' + v$, where $r' \in eR$ and $v \in V$, so we can define $f(r) = \begin{pmatrix} r' & 0 \\ v & 0 \end{pmatrix}$. Obviously f is an additive isomorphism. If $r_1 = r'_1 + v_1$, $r'_1 \in eR$, $v_1 \in V$, then $rr_1 = r'r'_1 + vr'_1 + r'v_1 + vv_1$. However $Rv = 0$, so $rr_1 = r'r'_1 + vr'_1$. Clearly $r'r'_1 \in Re$ and $vr'_1 \in V$. Thus $f(rr_1) = \begin{pmatrix} r'r'_1 & 0 \\ vr'_1 & 0 \end{pmatrix}$. On the other hand

$$f(r)f(r_1) = \begin{pmatrix} r' & 0 \\ v & 0 \end{pmatrix} \begin{pmatrix} r'_1 & 0 \\ v'_1 & 0 \end{pmatrix} = \begin{pmatrix} r'r'_1 & 0 \\ vr'_1 & 0 \end{pmatrix}.$$

Hence f is a ring isomorphism.

If a left ideal L of $\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix}$, where Δ is a division ring and V is a

right Δ -module, contains an element $\begin{pmatrix} a & 0 \\ v & 0 \end{pmatrix}$ with $a \neq 0$, then $\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix} = \begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ v & 0 \end{pmatrix} \subseteq L$.

Hence $L = \begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix}$. Consequently every proper left ideal of $\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix}$ is contained in $\begin{pmatrix} 0 & 0 \\ V & 0 \end{pmatrix}$. Let $V^* = \text{Hom}(V_\Delta, \Delta)$.

Note that V^* is a $\Delta - E$ -bimodule, where $E = \text{End}(V_\Delta, V_\Delta)$ with respect the usual addition and multiplication given by $(a\varphi)(v) = a\varphi(v)$, $\varphi f = \varphi \circ f$ for all $a \in \Delta$, $\varphi \in V^*$, $v \in V$ and $f \in E$. Defining $\varphi v = \varphi(v)$ and $(v\varphi)(v_1) = v\varphi(v_1)$ for $\varphi \in V^*$, $f \in E$ and $v, v_1 \in V$ we get that

$$M = \begin{pmatrix} \Delta & V^* \\ V & E \end{pmatrix}$$

is a Morita context. Note that $\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix} <_l M$. If L is a nontrivial left ideal of $\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix}$, then, as we have seen, $L \subseteq \begin{pmatrix} 0 & 0 \\ V & 0 \end{pmatrix}$. Take any $0 \neq \begin{pmatrix} 0 & 0 \\ v & 0 \end{pmatrix} \in L$.

There exists $\varphi \in V^*$ such that $\varphi(v) \neq 0$. Consequently $\begin{pmatrix} 0 & \varphi \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ v & 0 \end{pmatrix} = \begin{pmatrix} \varphi(v) & 0 \\ 0 & 0 \end{pmatrix} \notin L$. This shows that L is not a left $*$ -ideal of $\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix}$.

In case $R^2 = 0$, the result can be obtained similarly as Theorem 1 in [1]. $\square$

Remark. As we have seen in the proof of Theorem 5 if R is a ring containing no nontrivial left $*$ -ideals, then each principal left ideal Rr of R is equal to 0 or R . Suppose that R is a ring with nonzero multiplication containing no nontrivial principal left ideals. Let $V = \{x \in R \mid Rx = 0\}$. Clearly V is a proper ideal of R and R/V contains no nontrivial left ideals. Hence R/V is a division ring. Since $V^2 = 0$, the identity of R/V can be lifted to an idempotent e of R . Clearly $Re \neq 0$, so $Re = R$. Now similar arguments as in the proof of Theorem 5 show that R is isomorphic to $\begin{pmatrix} \Delta & 0 \\ V & 0 \end{pmatrix}$, where $\Delta = R/V$. These show that the class of rings with nonzero multiplication containing no nontrivial left

$*$ -ideal coincides with the class of rings with nonzero multiplication containing no nontrivial principal left ideals.

References

- [1] E.R. Puczyłowski, On unequivocal rings, Acta Math. Hungar. 36 (1980), 57-62.
- [2] A.D. Sands, On ideals in over-rings, Publ. Math. 35 (1988), 273-279.
- [3] S. Veldsman, Extensions and ideals of rings, Publ. Math. 38 (1991), 297-309.

Instytut Matematyki i Fizyki, Politechnika Białostocka, Wiejska 45, 15-351 Białystok

Instytut Matematyki, Uniwersytet Warszawski, Banacha 2, 02-097 Warszawa.

O LEWYCH IDEALACH PIERŚCIENI ŁĄCZNYCH

Streszczenie

Wiadomo, że relacja „ L jest lewym ideałem pierścienia R ” nie jest przechodnia. Kilku autorów badało przypadki, w których przechodniość ma miejsce. W pracy tej kontynuujemy te badania.

Ewa Girejko*

ON MULTIDIFFERENTIALS OF MULTIFUNCTIONS

Abstract: In this article some properties of multidifferentials of set-valued maps (multifunctions) are studied. The functions considered here are mostly those that are not differentiable in a classical sense. Existence of minimal multidifferentials has been proved.

Key words: multifunction, regular multifunction, multidifferential, minimal multidifferential.

1 Basic concepts

In the papers [Su1, Su2, Su3] H.J.Sussmann introduced multidifferentials of multifunctions and showed some of their basic properties: the Chain Rule, Open Mapping and Transversal Intersection Theorems. In order to develop results obtained in those papers we have to answer some simple, but important questions concerning multidifferentials: Is any overset of a multidifferential also a multidifferential? Is any intersection of multidifferentials also a multidifferential? Which of the multidifferentials are the best mathematical tools? These and some more questions are addressed by the following propositions and theorem. They extend the basic issues concerning multidifferentials.

Let X, Y be finite-dimensional real linear spaces and let $\bar{x} \in X, \bar{y} \in Y$ and $C \subset X$. We will call a nonempty subset C of X a *cone* if $(\forall_{x \in C})(\forall_{r \in \mathbb{R}})(r \geq 0 \Rightarrow r \cdot x \in C)$.

We use $REG(K, M)$ to denote the set of all regular set-valued maps from K to M and $Gr(F)$ to denote the *graph* of F (where $Gr(F) = \{(x, y) : y \in F(x)\}$) and $dist(q, Gr(F)) = \inf \{d(q, g) : g \in Gr(F)\}$.

*Supported by KBN under the Technical University of Białystok grant No. W/IMF/1/01

We shall be working in finite-dimensional real linear spaces X and Y whose open unit balls are denoted B_X, B_Y (resp.); closed unit balls are denoted $\bar{B}_X, \bar{B}_Y$.

Let us recall the basic definitions of Sussmann and Clarke.

Definition 1. Let K be a compact metric space and M be a metric space. We say that a set-valued map $F : K \rightarrow 2^M$ is *regular* if there exists a sequence $\{F_j\}_{j=1}^\infty$ of continuous single-valued maps $F_j : K \rightarrow M$ with the property that

$$\lim_{j \rightarrow \infty} \sup \{ \text{dist}(q, Gr(F)) : q \in Gr(F_j) \} = 0.$$

□

Definition 2. Suppose $F : X \rightarrow 2^Y$ is a multifunction, $(\bar{x}, \bar{y}) \in X \times Y$, where X, Y are finite-dimensional real linear spaces and D is a nonempty compact subset of $\mathcal{L}(X, Y)$ – the set of linear maps from X to Y . Let C be a closed convex cone in X . We say that D is a *multidifferential of F at $(\bar{x}, \bar{y})$ in the direction of C* and write $D \in MD_C(F; \bar{x}, \bar{y})$, if for every neighborhood D' of D in $\mathcal{L}(X, Y)$ there exists a pair (ρ, θ) such that $\rho > 0$, θ is a function on the interval $]0, \rho]$ with values in $[0, \infty[$ and such that $\lim_{s \downarrow 0} \theta(s) = 0$, and

(MD) for every $\epsilon > 0$ there exist f, H such that

(MD.1) $f \in REG(\bar{x} + (C \cap \rho \bar{B}_X), Y)$

(MD.2) $Gr(f) \subseteq Gr(F)$

(MD.3) $H \in C^1(\bar{x} + (C \cap \rho \bar{B}_X), Y)$

(MD.4) $H(\bar{x}) = \bar{y}$

(MD.5) $EH(x) \in D'$ for all $x \in \bar{x} + (C \cap \rho \bar{B}_X)$

(MD.6) the inequality

$\sup \{ \|y - H(x)\| : y \in f(x) \} \leq \epsilon + \theta(\|x - \bar{x}\|) \|x - \bar{x}\|$ holds whenever $x \in \bar{x} + (C \cap \rho \bar{B}_X)$

□

A finite union, intersection and relation of inclusion apply to multidifferentials just as they apply to algebra of sets.

Definition 3. Let f be Lipschitz near a given point, and let v be a vector in X . The *generalized directional derivative* of f at x in the direction v , denoted $f^\circ(x; v)$, is defined as follows:

$$f^\circ(x; v) = \limsup_{y \rightarrow x} \sup_{t \downarrow 0} \frac{f(y + tv) - f(y)}{t},$$

where y is a vector in X and t is a positive scalar.

Definition 4. Let $\zeta : X \rightarrow \mathbb{R}$ be bounded linear functional, which belongs to the dual space X^* of continuous linear functionals on X . Then the *generalized gradient* of f at x , denoted $\partial f(x)$, is the subset of X^* given by

$$\{ \zeta \in X^* : f^\circ(x; v) \geq \langle \zeta, v \rangle \text{ for all } v \text{ in } X \}.$$

Here are some basic properties of generalized gradients (we assume that f is Lipschitz near x):

(a) $\partial f(x)$ is a nonempty, convex, compact subset of X^*

(b) for every v in X , one has

$$f^\circ(x; v) = \max \{ \langle \zeta, v \rangle : \zeta \in \partial f(x) \}$$

(c) $\zeta \in \partial f(x)$ iff $f^\circ(x; v) \geq \langle \zeta, v \rangle$ for all v in X .

Example 1. Consider $f : \mathbb{R} \rightarrow \mathbb{R}$ such that $f(x) = x^2 \sin \frac{1}{x}$. Then f is differentiable at 0, Lipschitz near 0 and $f'(0) = 0$. We will show that $\partial f(0) = [-1, 1]$. Let's take $v \geq 0$, then we can calculate:

$$\begin{aligned} f^\circ(0; v) &= \limsup_{y \rightarrow 0} \sup_{t \downarrow 0} \frac{(y + tv)^2 \sin \frac{1}{y+tv} - y^2 \sin \frac{1}{y}}{t} \\ &= \limsup_{y \rightarrow 0} \sup_{tv \downarrow 0} \frac{v((y + tv)^2 \sin \frac{1}{y+tv} - y^2 \sin \frac{1}{y})}{tv} \\ &= \limsup_{y \rightarrow 0} (2vy \sin \frac{1}{y} - v \cos \frac{1}{y}). \end{aligned}$$

Fix $y_k = \frac{1}{\pi + 2k\pi}$; then $y_k \rightarrow 0$ for $k \rightarrow 0$. Thus

$$f^\circ(0; v) = \limsup_{k \rightarrow \infty} [2v \sin(\pi + 2k\pi) - v \cos(\pi + 2k\pi)] = v.$$

By analogy for $v < 0$ we'll get $f^\circ(0; v) = -v$, thus

$$f^\circ(0; v) = \begin{cases} v & \text{for } v \geq 0 \\ -v & \text{for } v < 0 \end{cases}$$

i.e. $f^\circ(0; v) = |v|$. Thus $\partial f(0)$ consists of those ζ satisfying $|v| \geq \zeta v$ for all v ; that is, $\partial f(0) = [-1, 1]$. We can also write, that $\partial f(0) \in MD(f; 0, f(0))$ for $f(0) = 0$ (see [Su1]). But for Lipschitz maps usually exist many multidifferentials smaller then $\partial f(x)$, like in our case: $\{f'(0)\} \in MD(f; 0, f(0))$ ([Su1]). $\square$

Theorem 1. [Su1] (The chain rule for multidifferentials) Let X_1, X_2, X_3 be finite-dimensional real linear spaces, and let $\bar{x}_i \in X_i$, $i = 1, 2, 3$. For $i = 1, 2$, let $F_i : X_i \rightarrow 2^{X_{i+1}}$ be a set-valued map from X_i to X_{i+1} , let C_i be a closed convex cone in X_i , and let D_i be a compact subset of $\mathcal{L}(X_i, X_{i+1})$ such that $D_i \in MD_{C_i}(F_i; \bar{x}_i, \bar{x}_{i+1})$. Let S be a linear subset of X_2 spanned by C_2 , and let Π be a linear projection from X_2 onto S . Assume that

$$F_1(\bar{x}_1 + C_1) \subseteq \bar{x}_2 + C_2.$$

Then

$$D_2 \circ \Pi \circ D_1 \in MD_{C_1}(F_2 \circ F_1; \bar{x}_1, \bar{x}_3),$$

where

$$D_2 \circ \Pi \circ D_1 = \{d_2 \circ \Pi \circ d_1 : d_2 \in D_2, d_1 \in D_1\}.$$

$\square$

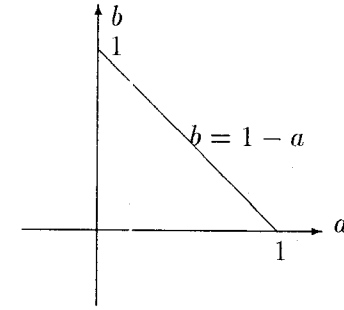
Example 2. (An application of the chain rule)

Consider $f_1 : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $f_2 : \mathbb{R} \rightarrow \mathbb{R}$ such that $f_1(x_1, x_2) = \max\{x_i; i = 1, 2\}$, $f_2(x) = |x|$. Then both f_1 and f_2 are continuous at 0 but they aren't differentiable at 0 in the classical sense. We have

$$[-1, 1] \in MD(f_2; 0, f_2(0)).$$

One can see [Cl] that Clarke's generalized gradient for f_1 at 0 is (if we identify $(\mathbb{R}^2)^* = \mathbb{R}^2$):

$$\begin{aligned} \partial f_1(0) &= \{\zeta \in \mathbb{R}^2 : \zeta_i \geq 0 \wedge \sum \zeta_i = 1, i = 1, 2\} \\ &= \{(a, b) \in \mathbb{R}^2 : a + b = 1; a, b \geq 0\}. \end{aligned}$$



From basic properties of multidifferentials [Su1] we know that $\partial f_1(0) \in MD(f_1; 0, f_1(0))$ (we have $f_i(0) = 0$, $i = 1, 2$).

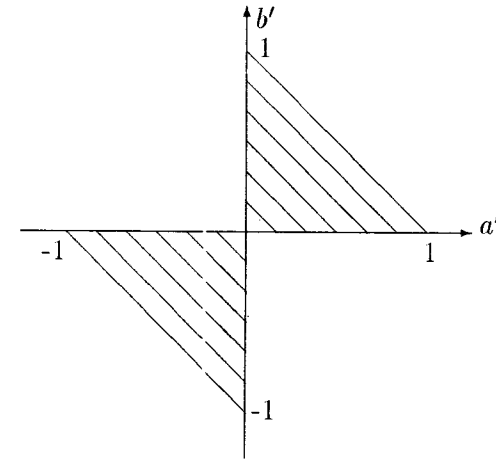
Let $D_1 = \partial f_1(0)$, $D_2 = [-1, 1]$. Then, by Theorem 1,

$$D_2 \circ D_1 \in MD(f_2 \circ f_1; 0, 0).$$

and

$$\begin{aligned} D_2 \circ D_1 &= [-1, 1] \cdot \{(a, b) \in \mathbb{R}^2 : a + b = 1; a, b \geq 0\} \\ &= \{(c \cdot a, -c \cdot a + c) : c \in [-1, 1], a \geq 0\} \\ &= \{(a', b') \in \mathbb{R}^2 : -1 \leq a' + b' \leq 1, -1 \leq a' \leq 1, -1 \leq b' \leq 1\}. \end{aligned}$$

where we identify linear maps with elements of $\mathbb{R}^2$ or $\mathbb{R}$:



Then $D_2 \circ D_1$ is one of the multidifferentials of the function $f_3 = f_2 \circ f_1$, where $f_3(x_1, x_2) = f_2(f_1(x_1, x_2)) = \max\{x_i; i = 1, 2\}$ at the point $(0, f_3(0))$ i.e.

$$D_2 \circ D_1 \subset MD(f_3; 0, f_3(0))$$

for $f_3(0) = 0$.

Remark 1. In the example above we can see, that neither $f_2 \circ f_1$ nor $D_2 \circ D_1$ are convex, so we can't apply the Chain Rule for Clarke's generalized gradients (see [Cl]).

2 Properties of multidifferentials

Proposition 1. Let $F : X \rightarrow 2^Y$ be a multifunction where X, Y are finite-dimensional real linear spaces and $\bar{x} \in X$, $\bar{y} \in Y$. Let C_1, C_2 be convex cones in a space X such that

$$C_1 \subseteq C_2.$$

Then

$$MD_{C_2}(F; \bar{x}, \bar{y}) \subseteq MD_{C_1}(F; \bar{x}, \bar{y}).$$

PROOF: Let's take any $D \in MD_{C_2}(F; \bar{x}, \bar{y})$. We will show that then $D \in MD_{C_1}(F; \bar{x}, \bar{y})$.

Assume, without loss of generality, that $\bar{x} = 0$ and $\bar{y} = 0$. Then for every neighborhood D' of D there exist ρ and θ such that for any ϵ there exist f and H such that

$$f \in REG(C_2 \cap \rho B_{\mathbb{R}}, \mathbb{R})$$

$$G(f) \subseteq G(F)$$

$$H \in C^1((C_2 \cap \rho B_{\mathbb{R}}, \mathbb{R}))$$

$$DH(x) \in D'$$

$$H(0) = 0$$

and

$$\sup\{\|y - H(x)\| : y \in f(x)\} \leq \epsilon + \theta(\|x - \bar{x}\|) \|x - \bar{x}\|$$

for $x \in C_2 \cap \rho B_{\mathbb{R}}$.

And now if we replace the neighborhood $C_2 \cap \rho B_{\mathbb{R}}$ of 0 by the smaller one i.e. $C_1 \cap \rho B_{\mathbb{R}}$ then the functions f and H will still keep their properties on this smaller set. Then we get $D \in MD_{C_1}(F; \bar{x}, \bar{y})$. □

Example 3. Consider $F : \mathbb{R} \rightarrow 2^{\mathbb{R}}$ - a multifunction and $C_1 = [0, +\infty]$, $C_2 = \mathbb{R}$. Then we have $C_1 \subset C_2$. Let F be multidifferentiable at $(0, 0)$ in the directions of C_1 and C_2 . Then

$$MD_{C_2}(F; 0, 0) \subset MD_{C_1}(F; 0, 0).$$

Proposition 2. Let $F : \mathbb{R} \rightarrow \mathbb{R}$ be continuously differentiable on a neighborhood of $\bar{x}$. Then $F'(\bar{x})$ is a multidifferential of F at $\bar{x}$, i.e. $\{F'(\bar{x})\} \in MD(F; \bar{x}, \bar{y})$ where $\bar{y} = F(\bar{x})$.

PROOF: Let D' be a neighborhood of $F'(x)$. Fix ρ equal to the radius of the neighborhood of $\bar{x}$, $\theta(s) \equiv 0$ and $f := H := F$. Then $f \in REG(\bar{x} + \rho B_{\mathbb{R}}, \mathbb{R})$, $G(f) = G(F)$, $H \in C^1((\bar{x} + \rho B_{\mathbb{R}}, \mathbb{R}))$, $DH(x) \in D'$, $H(\bar{x}) = \bar{y}$ and

$$\sup\{\|y - H(x)\| : y = f(x)\} \leq \epsilon + \theta(\|x - \bar{x}\|) \|x - \bar{x}\|$$

for $x \in \bar{x} + \rho B_{\mathbb{R}}$. And the proof is complete. □

Corollary 1. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a function continuously (Gateaux) differentiable on a neighborhood of $\bar{x}$ (i.e. on this neighborhood of $\bar{x}$ the Gateaux derivative DF exists and is continuous as a mapping from $\mathbb{R}^n$ to $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$). Then $DF(\bar{x}) \in MD(F; \bar{x}, \bar{y})$, where $\bar{y} \in F(\bar{x})$.

PROOF: If F is continuously differentiable in Gateaux sense on a neighborhood of $\bar{x}$, then F is Lipschitz near $\bar{x}$ and $\{DF(x)\} \in \partial F(x)$ for every $x \in \bar{x} + \epsilon B$. The required conclusion is now obvious. □

Example 4. Consider $F(x) = |x|$. We will show that

$$[-1, 1] \in MD(F; 0, 0).$$

Let D' be a neighborhood of $[-1, 1]$. Define $\rho = 1$ and $\theta \equiv 0$. Consider the function $H_\delta(x) = |x|^{1+\delta}$ for $\delta > 0$. Then F is regular on a neighborhood of 0 (because is single-valued and continuous near 0) and H_δ is differentiable for every $\delta > 0$. Now we will show that for such ρ, θ the estimation from Definition 2 holds. We have $H'_\delta(x) = (1+\delta)|x|^\delta \text{sgn}(x)$ so $|H'_\delta(x)| \leq (1+\delta)$ for $x \in B_\mathbb{R}$. Now take $\epsilon > 0$ and define $f := F$ and $H := H_\delta$ for some $\delta < \epsilon$ such that $1+\delta \in D'$. Then $H'(x) \in D'$ for $x \in B_\mathbb{R}$. We have to estimate the following

$$||x| - |x|^{1+\delta}|.$$

The maximum of this function is obtained at

$$x_\delta = \left(\frac{1}{1+\delta}\right)^{\frac{1}{\delta}} \rightarrow_{\delta \rightarrow 0} \frac{1}{e}.$$

So then we get

$$\begin{aligned} \sup ||x| - |x|^{1+\delta}| &= \left(\left(\frac{1}{1+\delta}\right)^{\frac{1}{\delta}} - \left[\left(\frac{1}{1+\delta}\right)^{\frac{1}{\delta}}\right]^{1+\delta}\right) \\ &= \left(\frac{1}{1+\delta}\right)^{\frac{1}{\delta}} \left(1 - \frac{1}{1+\delta}\right) \\ &= \left(\frac{1}{1+\delta}\right)^{\frac{1}{\delta}} \frac{\delta}{1+\delta} < \delta < \epsilon. \end{aligned}$$

And this is what we were supposed to obtain.

Example 5. Consider two single-valued functions $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$. The first one is given by $f_1(x) = x + 3x^2 \sin \frac{1}{x}$. Then f_1 is continuous at 0, differentiable at 0 with $f'_1(0) = 1$ and Lipschitz near 0. We will show that $\{1\} \in MD(f_1; 0, f_1(0))$, where $f_1(0) = 0$.

For a neighborhood D' of $\{1\}$ let $\rho = 1$, $\theta(s) = 3s$. Then $\rho > 0$ and $\lim_{s \rightarrow 0} \theta(s) = 0$. For such (ρ, θ) let $\epsilon > 0$. Take $f = f_1$ and $H(x) = x$. Then

$$f \in REG(B_\mathbb{R}, \mathbb{R})$$

$$\begin{aligned} G(j) &= G(f_1) \\ H &\in C^1(B_\mathbb{R}, \mathbb{R}) \\ H(0) &= 0 \\ DH(x) &= 1 \in D'. \end{aligned}$$

Moreover

$$\sup\{||y - H(x)|| : y = f(x)\} = \sup ||x + 3x^2 \sin \frac{1}{x} - x|| \leq 3x^2 + \epsilon$$

for $x \in B_\mathbb{R}$.

The second function is $f_2(x) = x + |x|^{\frac{3}{2}} \sin \frac{1}{x}$ for $x \neq 0$. This is continuous function at 0, differentiable at 0 with $f'(0) = 1$ but it is not Lipschitz near 0. And again we will show that $\{1\} \in MD(f_2; 0, f_2(0))$, where $f_2(0) = 0$.

For every neighbourhood D' of $\{1\}$ let $\rho = 1$ and $\theta(s) = \sqrt{s}$. Let $\epsilon > 0$ and put $f := f_2$, $H(x) = x$. Then

$$\begin{aligned} f &\in REG(B_\mathbb{R}, \mathbb{R}) \\ G(j) &= G(f_2) \\ H &\in C^1(B_\mathbb{R}, \mathbb{R}) \\ H(0) &= 0 \\ DH(x) &= 1 \in D' \end{aligned}$$

and

$$\sup\{||y - H(x)|| : y = f(x)\} = \sup ||x + |x|^{\frac{3}{2}} \sin \frac{1}{x} - x|| \leq |x|^{\frac{3}{2}} + \epsilon$$

for $x \in B_\mathbb{R}$.

Definition 5. A map $f : X \rightarrow Y$ is called a selection of multifunction $F : X \rightarrow 2^Y$ if for any point $x \in X$ the inclusion $f(x) \in F(x)$ is fulfilled. $\square$

Proposition 3. Let $F : X \rightarrow 2^Y$ be a multifunction having a selection $f \in C^1(X, Y)$. Then $f'(x_0) \in MD(F; x_0, y_0)$, where $y_0 = f(x_0)$.

PROOF: Assume without loss of generality that $x_0 = 0, y_0 = 0$. It is sufficient to show that for every neighbourhood D of $f'(x_0)$ there exist ρ (resp. small) and θ (arbitrary) that for every $\epsilon > 0$ there exist $\tilde{f}$ and H satisfying the definition of multidifferential. Let $H = \tilde{f} = f$. Then one obtains:

$$\begin{aligned} Gr(\tilde{f}) &\subseteq Gr(F) \\ \tilde{f} &\in REG(\rho B_X, Y) \end{aligned}$$

and

$$\begin{aligned} H(0) &= 0 \\ DH(x) &= f'(x) \in D \text{ for } x \in \rho B_X \\ \text{and } H &\in C^1(\rho B_X, Y). \end{aligned}$$

And then the following estimation is trivial

$$\sup \| \tilde{f} - H \| \leq \epsilon + \theta(\|x\|) \|x\|$$

for every ϵ , and $x \in \rho B_X$. It is clear that other conditions of the definition of multidifferential are satisfied and the proof is complete. $\square$

Proposition 4. *A single-valued function f is continuous on a neighbourhood of a point $\bar{x}$ if it is multidifferentiable at the point $(\bar{x}, f(\bar{x}))$.*

PROOF: Let function f be multidifferentiable at $\bar{x}$. Then there exists a single-valued $\tilde{f} \in REG(\bar{x} + \rho B_{\mathbb{R}}, \mathbb{R})$ such that $Gr(\tilde{f}) = Gr(f)$. But if $\tilde{f} \in REG(K, X)$ and $\tilde{f}$ is single-valued, then $\tilde{f} \in C^0(X, Y)$. Thus f is continuous on $x + \rho B_{\mathbb{R}}$. $\square$

Remark 2. For a single-valued function $f : \mathbb{R} \rightarrow \mathbb{R}$ that is multidifferentiable at the point $(\bar{x}, f(\bar{x}))$ the following holds

$$\exists_{\epsilon > 0} \forall_{x \in (\bar{x} - \epsilon, \bar{x} + \epsilon)} : f(\bar{x}) = \tilde{f}(\bar{x}),$$

where $\tilde{f}$ is a regular function from the definition of multidifferential.

Proposition 5. *Let $F : X \rightarrow 2^Y$ be a multifunction, where X, Y are finite-dimensional real linear spaces and $\bar{x} \in X, \bar{y} \in Y$. Suppose that $D \in MD(F; \bar{x}, \bar{y})$. Then any compact overset of D is a multidifferential of F at the point $(\bar{x}, \bar{y})$.*

PROOF: Take a compact set $\tilde{D} \in \mathcal{L}(X, Y)$ such that $\tilde{D} \supset D$. Then every neighborhood D' of $\tilde{D}$ will be also a neighborhood of D . Now take ρ and θ corresponding to D' as a neighborhood of D . They will work for $\tilde{D}$ as well. $\square$

Corollary 2. The union of multidifferentials of the function F at the point $(\bar{x}, \bar{y})$ is also a multidifferential at the same point.

PROOF: Follows from Proposition 1. $\square$

Theorem 2. *Let $D_k, k \in \mathbb{N}$, be multidifferentials of F at $(\bar{x}, \bar{y})$ i.e. $D_k \in MD(F; \bar{x}, \bar{y})$ for $k = 1, 2, \dots$ where $F : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^m}$ is a multifunction. Assume that*

$$D_1 \supset D_2 \supset \dots$$

and let $\bigcap_k D_k = D$ ($D \neq \emptyset$ and D is compact as implied by description of D_k). Then D is a multidifferential of F at $(\bar{x}, \bar{y})$ i.e.

$$D \in MD(F; \bar{x}, \bar{y}).$$

PROOF: We will show that for any neighborhood D' of D conditions of the definition of multidifferential are satisfied. Since for every D_k conditions of the definition are satisfied, it is sufficient to show that there exists k such that $D_k \subset D'$. We will apply the following theorem on compact sets:

Theorem 3. ([En])

Let U be an open subset of a topological space X . If the family $\{F_s\}_{s \in S}$ of closed subsets of the space X includes at least one compact set and $\bigcap_{s \in S} F_s \subset U$, then there exists a finite set $\{s_1, \dots, s_n\} \subset S$ such that $F_{s_1} \cap F_{s_2} \cap \dots \cap F_{s_n} \subset U$.

It then follows from the above theorem that there exists a finite set

$$\{k_1, \dots, k_n\} \subset K$$

such that

$$D_{k_1} \cap D_{k_2} \cap \dots \cap D_{k_n} \subset D'.$$

In our case this is a family of descending sets and we get

$$D_{k_1} \cap D_{k_2} \cap \dots \cap D_{k_n} = D_{k_n}, \quad k_n \in K$$

thus

$$\exists_{k=k_n} D \subset D_k \subset D'.$$

And the proof is complete. $\square$

Corollary 3. If the set of multidifferentials of a multifunction F at a point $(\bar{x}, \bar{y})$ is not empty then there exists in this set at least one minimal multidifferential D of multifunction F at this point in the sense of inclusion of sets.

PROOF: From Kuratowski-Zorn Lemma, a family of sets with descending sequences having a lower bound possesses a minimal element. In our case such sequences of multidifferentials (D_k) have a lower bound – their intersection D . Thus there exists for the family of multidifferentials a minimal element D , which we call a minimal multidifferential. $\square$

We use $mMD(F; \bar{x}, \bar{y})$ to denote the set of all minimal multidifferentials of multifunction F at the point $(\bar{x}, \bar{y})$, for $\bar{y} \in F(\bar{x})$.

Example 6. Consider $f(x) = x^2 \sin \frac{1}{x}$. Then f is differentiable at 0 in a classical sense and $f'(0) = 0$, thus we can write $\{0\} \in mMD(f; 0, f(0))$, where $f(0) = 0$.

References

- [Cl] F. Clarke, *Optimization and Nonsmooth Analysis*, SIAM, Philadelphia 1990.
- [En] R. Engelking, *General Topology*, Polish Scientific Publ., Warsaw 1977.
- [Ge] B. D. Gel'man, *On topological dimension of a set of solutions of functional inclusions*, Differential Inclusions and Optimal Control, vol.2, Lecture Notes in Nonlinear Analysis (1998), 163-178.
- [Su1] H. J. Sussmann, *Multidifferential Calculus: Chain Rule, Open Mapping and Transversal Intersection Theorems*, Optimal Control: Theory, Algorithms and Applications, University of Florida, Gainesville, Feb. 27,28 and March 1, 1997.
- [Su2] H. J. Sussmann, *Recent nonsmooth version of the maximum principle*, a talk at AMS Summer Research Institute in Boulder, Colorado, July 7 and 8, 1997.

- [Su3] H. J. Sussmann, *Transversality conditions and a strong maximum principle for systems of differential inclusions*, The 37th IEEE Conference on Decision and Control, Aug. 10. 1998.
- [Su4] H. J. Sussmann, *New theories of set-valued differentials and new versions of the maximum principle of optimal control theory*, in *Nonlinear Control in the Year 2000*, A. Isidori, F. Lamnabhi-Lagarigue and W. Respondek Eds., Springer-Verlag, London. 2000, pp.487-526.

Instytut Matematyki i Fizyki, Politechnika Białostocka, Wiejska 45, 15-351 Białystok

O MULTIRÓŻNICZKACH MULTIFUNKCJI

Streszczenie

W artykule tym badane są pewne własności multiróżniczek funkcji wielowartościowych (multifunkcji). Rozpatrywane funkcje nie są zwykle różniczkowalne w klasycznym sensie. Pokazano istnienie minimalnych multiróżniczek.

Magdalena Kacprzak

WERYFIKACJA TEMPORALNYCH WŁASNOŚCI ALGORYTMÓW NIEDETERMINISTYCZNYCH

Streszczenie: W pracy zaprezentowany został system dowodowy dla logiki temporalnej czasu rozgałęzionego, nazwany Temporalną Logiką Procesów. Pozwala on na badanie statycznych i dynamicznych własności programów wyrażonych jako formuły temporalne rozważanego języka. Udowodniono mocne twierdzenie o pełności dla przedstawionego systemu formalnego.

1 Wstęp

Logiki powstały i rozwinęły się jako narzędzie podstawowe w procesie formalizacji i porządkowania rozumowania prowadzonego przez człowieka. Klasyczna logika pierwszego rzędu odgrywa istotną rolę w dziedzinach związanych ze sztuczną inteligencją. Jednak mimo swej ogólności i siły wyrazu nie zawsze jest ona odpowiednia do formalizacji zdań języka naturalnego. Mówienie o takich pojęciach jak czas, czy wiedza staje się przy użyciu tej logiki niezwykle złożone, a niejednokrotnie wręcz niemożliwe. Jednym ze sposobów radzenia sobie z tymi problemami jest rozważanie logik modalnych.

Logiki modalne mają długą tradycję, sięgającą starożytnej Grecji. Możliwością i koniecznością zajmował się już Arystoteles, później także megarejczycy, stoicy, scholastycy. Logiki modalne przyciągały uwagę filozofów, matematyków, a w ostatnich latach również informatyków. Okazało się bowiem, że modalności pozwalają na rozwiązywanie problemów związanych z rozumieniem natury obliczeń programów komputerowych i dowodzeniem ich poprawności.

Czas jest obecny w niemal wszystkich zjawiskach życia codziennego. Dlatego wszelkie systemy sztucznej inteligencji muszą sobie radzić z pojęciem

czasu. Powstało wiele systemów, w których o czasie mówi się przez pryzmat logik modalnych. Do najpoważniejszych zaliczyć można systemy wspomagające dowodzenie własności programów, poprawności protokołów komunikacyjnych, układów scalonych itp. Spośród logik modalnych największym zainteresowaniem cieszą się logiki temporalne. Jednym z podstawowych zastosowań tych logik jest specyfikacja i weryfikacja programów. Szczególnie owocnym zastosowaniem jest dziedzina obliczeń współbieżnych i rozproszonych.

Opisana w pracy Temporalna Logika Procesów (TLP) jest jedną z wersji zdaniowej logiki temporalnej czasu rozgałęzionego (Computation Tree Logic). Większość logik temporalnych traktuje program jako obiekt zewnętrzny, który nie jest elementem języka logiki. W zaprezentowanym w pracy systemie jest inaczej. Program jest składową wyrażenia, które opisuje jego własności podobnie jak w logice dynamicznej [4] i algorytmicznej [6].

Przedstawione w pracy wyniki są uogólnieniem i poszerzeniem relatywnie zupełnego systemu opisanego w [3]. Ponadto dla TLP podano nieskończoną aksjomatyzację oraz zaprezentowano dowód mocnego twierdzenia o pełności.

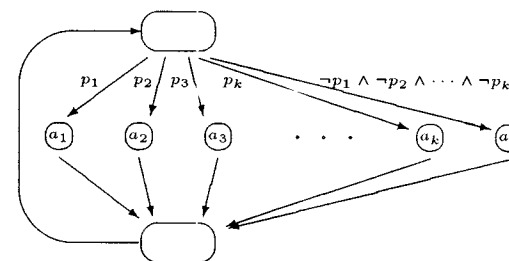
2 Składnia języka

Język TLP jest rozszerzeniem klasycznego rachunku zdań o formuły temporalne. Niech $L_o = \langle V_o, F_o \rangle$ oznacza język rachunku zdań, gdzie V_o - zbiór zmiennych zdaniowych, F_o - zbiór formuł języka L_o . Zdefiniujemy teraz zbiór procesów i zbiór formuł rozważanego w pracy języka.

Niech A_t będzie niepustym zbiorem deterministycznych akcji atomowych, a a^* jego wyróżnionym elementem.

Definicja 1. Zbiór procesów $\Pi(A_t)$ jest zbiorem wyrażen postaci:

◁ do
when $p_1 : a_1$;
⋮
when $p_k : a_k$;
otherwise a^* ;
▷ od
gdzie $p_i \in F_o$, $a_i, a^* \in A_t$, $i = 1, \dots, k$.



Rys. 1 Graficzna ilustracja procesu

Jeśli P jest procesem przedstawionym w definicji 1, to $A(P) = \{a_1, \dots, a_k, a^*\}$ jest zbiorem akcji atomowych procesu P , a zbiór $F(P) = \{p_1, \dots, p_k\}$ jest zbiorem formuł rachunku zdań, umożliwiającą wykonanie tych akcji. Intuicyjnie, dla każdego $i = 1, \dots, k$, wykonanie akcji a_i jest możliwe, gdy spełniona jest formuła p_i . Proces $P \in \Pi(A_t)$ będziemy w skrócie oznaczać $[p_i : a_i]_{i=1, \dots, k}$.

Definicja 2. F - zbiór formuł TLP jest to najmniejszy zbiór, taki że $F_o \subset F$ oraz jeśli $f, f_1, f_2 \in F$, $P \in \Pi(A_t)$, to $\neg f_1$, $f_1 \wedge f_2$, $f_1 \vee f_2$, $f_1 \rightarrow f_2$, $EXP f$, $AXP f$, $EGP f$, $AGP f$, $EUP[f_1, f_2]$, $AUP[f_1, f_2]$ są formułami TLP.

Formuły te definiują zarówno statyczne jak i dynamiczne własności procesów. Intuicyjne znaczenie operatorów temporalnych jest następujące: E - na pewnej ścieżce, A - na każdej ścieżce, X - w następnym stanie, G - w każdym stanie, U - "dopóki".

3 Semantyka języka

Intuicyjnie proces $P \in \Pi(A_t)$ w każdym pojedynczym kroku wykonuje jedną, możliwą do wykonania akcję ze zbioru $A(P)$. Akcja ta zostaje wybrana w sposób niedeterministyczny. Jeśli nie jest spełniona żadna z formuł $p_i \in F(P)$, to proces P wykonuje wyróżnioną akcję a^* . Zatem możemy powiedzieć, że P jest nieskończonym, jednopętlowym procesem z punktem kontroli na początku pętli.

Procesy i formuły TLP będą interpretowane w strukturze Kripkego K postaci $\langle S, R, v \rangle$, gdzie:

S - skończony zbiór stanów,

R - rodzina funkcji całkowitych określonych na S , $R = \{R_a\}_{a \in A_t}$, $R_a : S \rightarrow S$,

v - funkcja przyporządkowująca wartości logiczne zmiennym zdaniowym w stanach, $v : V_o \times S \rightarrow \{1, 0\}$.

Niech $P = [p_i : a_i]_{i=1, \dots, k}$. Wykonanie akcji atomowej $a_i \in A(P)$ jest możliwe w stanie $s \in S$ struktury K , jeśli w stanie tym spełniona jest formuła $p_i \in F(A)$. Formuła ta nazywana jest warunkiem wykonania akcji a_i . Akcję a^* nazywamy akcją pustą. Przypisuje ona każdemu stanowi s struktury K ten sam stan s .

Definicja 3. Niech $K = \langle S, R, v \rangle$ będzie strukturą Kripkego, a $P \in \Pi(A_t)$ procesem. Drzewo możliwych obliczeń procesu P w strukturze K przy zadanym stanie początkowym s_0 , ozn. $T(P, s_0)$, jest to etykietowane drzewo T takie, że:

- 1) wierzchołki drzewa T etykietowane są przez stany struktury K ,
- 2) krawędzie drzewa T etykietowane są przez akcje atomowe,
- 3) korzeń drzewa T jest etykietowany przez stan początkowy s_0 ,
- 4) jeśli w jest wierzchołkiem drzewa T etykietowanym przez stan s oraz $p_1, \dots, p_n$ są formułami procesu P spełnionymi w stanie s , to wierzchołek w ma dokładnie n bezpośrednich następników $w_1, \dots, w_n$ etykietowanych odpowiednio stanami $R_{a_1}, \dots, R_{a_n}$ oraz dla każdego $i = 1, \dots, n$ akcja a_i jest etykietą krawędzi (w, w_i) ,
- 5) jeśli w jest wierzchołkiem drzewa T etykietowanym przez stan s oraz nie istnieje formuła $p \in F(P)$ spełniona w stanie s , to wierzchołek w ma dokładnie jeden bezpośredni następnik w' etykietowany stanem s i krawędź (w, w') jest etykietowana akcją a^* .

Niech K będzie strukturą Kripkego. Relację spełniania $\models$ definiujemy następująco: dla każdego stanu s , procesu P , zmiennej zdaniowej q i dowolnych formuł $f, f_1, f_2 \in F$

$K, s \models q$ wttw, gdy $v(q, s) = 1$,

$K, s \models \neg f$ wttw, gdy $K, s \not\models f$,

$K, s \models f_1 \wedge f_2$ wttw, gdy $K, s \models f_1$ i $K, s \models f_2$,

$K, s \models f_1 \vee f_2$ wttw, gdy $K, s \models f_1$ lub $K, s \models f_2$,

$K, s \models f_1 \rightarrow f_2$ wttw, gdy $K, s \not\models f_1$ lub $K, s \models f_2$,

$K, s \models AXPf$ wttw, gdy w drzewie $T(P, s)$ każdy bezpośredni następnik stanu s spełnia formułę f ,

$K, s \models EXPf$ wttw, gdy w drzewie $T(P, s)$ istnieje bezpośredni następnik stanu s spełniający formułę f ,

$K, s \models AGPf$ wttw, gdy każdy stan na każdej ścieżce drzewa $T(P, s)$ spełnia formułę f ,

$K, s \models EGPf$ wttw, gdy istnieje ścieżka w drzewie $T(P, s)$ taka, że każdy jej stan spełnia formułę f ,

$K, s \models AUP[f_1, f_2]$ wttw, gdy na każdej ścieżce drzewa $T(P, s)$ istnieje stan spełniający formułę f_2 oraz wszystkie poprzedzające go stany na tej ścieżce spełniają formułę f_1 ,

$K, s \models EUP[f_1, f_2]$ wttw, gdy istnieje ścieżka w drzewie $T(P, s)$ i stan na tej ścieżce spełniający formułę f_2 oraz wszystkie poprzedzające go stany na tej ścieżce spełniają formułę f_1 .

Mówimy, że formuła f jest prawdziwa w strukturze K (ozn. $K \models f$) wttw, gdy $K, s \models f$ dla każdego $s \in S$.

Mówimy, że formuła f jest prawdziwa (ozn. $\models f$) wttw, gdy $K \models f$ dla każdej struktury K .

4 System dowodowy

System dowodowy składa się z następujących aksjomatów i reguł wnioskowania.

AKSJOMATY

Ax 0 Aksjomaty klasycznego rachunku zdań

Ax 1. $AXP(\neg f) \longleftrightarrow \neg EXP(f)$

Ax 2. $AXP(f_1 \rightarrow f_2) \longrightarrow (AXPf_1 \rightarrow AXPf_2)$

Ax 3. $AGP(\neg f) \longrightarrow \neg EGP(f)$

Ax 4. $EGPf \longleftrightarrow f \wedge EXP(EGPf)$

Ax 5. $AGPf \longleftrightarrow f \wedge AXP(AGPf)$

Ax 6. $AGP(f \rightarrow AXPf) \longrightarrow (f \rightarrow AGPf)$

Ax 7. $AUP[f_1, f_2] \longrightarrow AUP[true, f_2]$

Ax 8. $\neg EUP[true, \neg f] \longleftrightarrow AGPf$

Ax 9. $\neg AUP[f_1, f_2] \longleftrightarrow EGP(\neg f_2) \vee EUP[\neg f_2, \neg f_1 \wedge \neg f_2]$

Ax 10. $EUP[f_1, f_2] \longleftrightarrow f_2 \vee (f_1 \wedge EXP(EUP[f_1, f_2]))$

Ax 11. $AUP[f_1, f_2] \longleftrightarrow f_2 \vee (f_1 \wedge AXP(AUP[f_1, f_2]))$

Ax 12. $AGPf \longrightarrow EGPf$

Ax 13. $AXPf \longrightarrow EXPf$

Ax 14. $AUP[f_1, f_2] \longrightarrow EUP[f_1, f_2]$

Ax 15. $AX[p : a]f \longleftrightarrow EX[p : a]f$

Ax 16. $EX[p : a](f_1 \wedge f_2) \longleftrightarrow EX[p : a]f_1 \wedge EX[p : a]f_2$

Ax 17. $EXPf \longleftrightarrow (p_1 \wedge EX[true : a_1]f) \vee \dots \vee (p_k \wedge EX[true : a_k]f) \vee (f \wedge \neg p_1 \wedge \dots \wedge \neg p_k), \quad p_i \in F(P), a_i \in A(P), i = 1, \dots, k$

Ax 18. $AXPf \longleftrightarrow (\neg p_1 \vee AX[true : a_1]f) \wedge \dots \wedge (\neg p_k \vee AX[true : a_k]f) \wedge (p_1 \vee \dots \vee p_k \vee f), \quad p_i \in F(P), a_i \in A(P), i = 1, \dots, k$

Ax 19. $AXP \text{ true}$

W powyższych aksjomatach proces P ma postać $[p_i : a_i]_{i=1, \dots, k}$.

REGUŁY WNIOSKOWANIA

Wprowadzimy następujące oznaczenia:

- dla $i > 0$ $EXP^i(f_1, f_2) \stackrel{def}{=} f_1 \wedge EXP(f_1 \wedge (EXP(f_1 \wedge \dots EXP(f_2))))$, gdzie operator EXP występuje i -razy,
- dla $i = 0$ $EXP^0(f_1, f_2) \stackrel{def}{=} f_2$,
- jeśli w powyższych formułach $f_1 = f_2$, to będziemy w skrócie zapisywać $EXP^i f_1$.

Niech β będzie ciągiem akcji atomowych $b_1, \dots, b_n$, wtedy:

$$pref_\beta \stackrel{def}{=} EX[true : b_1]EX[true : b_2] \dots EX[true : b_n].$$

R 1. $\frac{\text{dla każdego } i \in \mathbb{N} \quad f_1 \rightarrow pref_\beta EXP^i f_2}{f_1 \rightarrow pref_\beta EGP f_2} \quad f_1, f_2 \in F$

R 2. $\frac{\text{dla każdego } i \in \mathbb{N} \quad pref_\beta EXP^i(f_1, f_2) \rightarrow f}{pref_\beta EUP[f_1, f_2] \rightarrow f} \quad f, f_1, f_2 \in F$

R 3. $\frac{f_1 \rightarrow f_2}{EXP f_1 \rightarrow EXP f_2} \quad f_1, f_2 \in F$

R 4. $\frac{f_1, f_1 \rightarrow f_2}{f_2}$

Definicja 4. Formalny dowód formuły f ze zbioru formuł Z jest to skończone drzewo etykietowane przez formuły, takie że :

- (1) korzeń jest etykietowany przez formułę f ,
- (2) etykietami wszystkich liści są aksjomaty lub formuły należące do zbioru Z ,
- (3) etykieta dowolnego wierzchołka w , niebędącego liściem, jest konkluzją reguły wnioskowania, której przesłanki są etykietami bezpośrednich poprzedników wierzchołka w .

Fakt, że formuła f ma formalny dowód ze zbioru Z będziemy oznaczać symbolem $Z \vdash f$ oraz $\vdash f$, jeśli Z jest zbiorem pustym.

Twierdzenie 1. Dla każdej formuły $f \in F$, jeśli $\vdash f$, to $\models f$.

Prosty dowód tego twierdzenia polega na wykazaniu, że:

- aksjomaty (1)-(19) są prawdziwe w każdej strukturze Kripkego,
- dla każdej reguły wnioskowania $R1 - R4$ i każdej struktury Kripkego K , jeżeli przesłanki reguły są prawdziwe w strukturze K , to wniosek również jest prawdziwy w tej strukturze.

Następujące twierdzenie jest wnioskiem wypływającym z twierdzenia 1.

Twierdzenie 2. Dla każdej formuły $f \in F$ i zbioru formuł Z , jeśli $Z \vdash f$, to $Z \models f$.

Poniższe twierdzenie jest bardzo użyteczne przy dowodzeniu formuł postaci $AUP[f_1, f_2]$.

Twierdzenie 3. Niech $Z \subset F$ - zbiór formuł, $f, f_1, f_2 \in F$, $P \in \Pi(A_t)$ - program postaci $[p_i : a_i]_{i=1, \dots, n}$, $\{\alpha_k\}_{k \in \mathbb{N}}$ - ciąg formuł. Jeśli

(1) $Z \vdash f \rightarrow f_2 \vee (f_1 \wedge \alpha_r)$ dla pewnego $r \in \mathbb{N}$,

(2) dla każdego $1 \leq t \leq r$ istnieje $1 \leq s < t$ takie, że

$Z \vdash f_1 \wedge \alpha_t \rightarrow AXP(f_2 \vee (f_1 \wedge \alpha_s))$,

(3) $Z \vdash f_1 \wedge \alpha_o \rightarrow AXP f_2$,

to $f \rightarrow AUP[f_1, f_2]$.

Dowód: Korzystając z (1), (2), reguły 3 i aksjomatu 1 otrzymujemy:
 $Z \vdash f \rightarrow f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge \alpha_t)))$ dla pewnego t .

Ponownie stosując (2), regułę 3 i aksjomat 1 mamy:

$Z \vdash f \rightarrow f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge \alpha_s)))))$ ($s < t$)

Jeszcze raz, po kilkakrotnym zastosowaniu (2), reguły 3 i aksjomatu 1 otrzymujemy:

$Z \vdash f \rightarrow f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2 \vee \dots (f_1 \wedge AXP(f_2 \vee (f_1 \wedge \alpha_o))))))))$ oraz z (3) $Z \vdash f \rightarrow f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2 \vee \dots (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2))))))))$.

Następnie, na mocy aksjomatu 11, $Z \vdash f_2 \rightarrow AUP[f_1, f_2]$.

Zatem $Z \vdash f \rightarrow f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2 \vee (f_1 \wedge AXP(f_2 \vee \dots AXP(f_2 \vee (f_1 \wedge AXP(AUP[f_1, f_2]))))))))$.

Ostatecznie po kilkakrotnym zastosowaniu aksjomatu 11 otrzymujemy:

$Z \vdash f \rightarrow f_2 \vee (f_1 \wedge AXP(AUP[f_1, f_2]))$ i $Z \vdash f \rightarrow AUP[f_1, f_2]$.

5 Pełność systemu dowodowego

Niech $T = \langle L, \vdash, A \rangle$ będzie teorią oparą na TLP, gdzie L - język TLP, $\vdash$ - operacja konsekwencji syntaktycznej, A - zbiór aksjomatów specyficznych.

Definicja 5. Struktura K jest modelem teorii $T = \langle L, \vdash, A \rangle$ wttw, gdy K jest modelem zbioru aksjomatów specyficznych A tej teorii.

Definicja 6. Formuła f jest twierdzeniem teorii T wttw, gdy f ma dowód ze zbioru A aksjomatów specyficznych.

Najpierw stworzymy algebrę Lindenbauma teorii T opartej na TLP. Niech $\sim$ oznacza relację równoważności, określoną w zbiorze formuł F , zdefiniowaną w następujący sposób: $f_1 \sim f_2$ wttw, gdy $A \vdash (f_1 \rightarrow f_2)$ i $A \vdash (f_2 \rightarrow f_1)$ dla każdego $f_1, f_2 \in F$. Niech $F/\sim$ oznacza zbiór wszystkich klas abstrakcji relacji $\sim$. Elementy zbioru $F/\sim$ będą oznaczane następująco: $\|f_1\|$, $\|f_2\|$, ... itp.

Algebra Lindenbauma teorii T jest to następująca struktura:
 $\mathcal{A} = (F/\sim; \cup, \cap, -, \mathbf{1}, \mathbf{0})$. Poniższe twierdzenie charakteryzuje tę algebrę.

Twierdzenie 4. (Rasiowa-Sikorski)

Algebra $\mathcal{A} = (F/\sim; \cup, \cap, -, \mathbf{1}, \mathbf{0})$ jest algebrą Boolea, gdzie

$\|f_1\| \cup \|f_2\| = \|f_1 \vee f_2\|$,

$\|f_1\| \cap \|f_2\| = \|f_1 \wedge f_2\|$,

$-\|f\| = \|\neg f\|$ oraz

f jest twierdzeniem teorii T wttw, gdy $\|f\| = \mathbf{1}$,

f nie jest twierdzeniem teorii T wttw, gdy $\|f\| \neq \mathbf{0}$.

Dowód twierdzenia 4 można znaleźć w [9].

Definiujemy relację częściowego porządku $\leq$ w $\mathcal{A}$ w następujący sposób:

$\|f_1\| \leq \|f_2\|$ wttw, gdy $A \vdash (f_1 \rightarrow f_2)$, $f_1, f_2 \in F$.

Lemat 1. Dla dowolnych formuł $f, f_1, f_2 \in F$, dowolnego ciągu akcji β i dowolnego procesu $P \in \Pi(A_t)$ zachodzą następujące równości:

a) $\|pref_\beta EGP f\| = \inf\{\|pref_\beta EXP^i f\|\}_{i \in \mathbb{N}}$

b) $\|pref_\beta EUP[f_1, f_2]\| = \sup\{\|pref_\beta EXP^i(f_1, f_2)\|\}_{i \in \mathbb{N}}$

Dowód:

a) Najpierw pokażemy, że $A \vdash EGP f \rightarrow EXP^i f$ dla każdego $i \in \mathbb{N}$. Przeprowadzimy dowód indukcyjny za względu na i .

1° Dla $i = 1$ z aksjomatu 4 i reguły 3 mamy $A \vdash EGP f \rightarrow f \wedge EXP f$.

2° Założenie indukcyjne: $A \vdash EGP f \rightarrow EXP^k f$ dla pewnego $k \in \mathbb{N}$.

Teza indukcyjna: $A \vdash EGP f \rightarrow EXP^{k+1} f$.

Dowód: z aksjomatu 4 oraz założenia indukcyjnego i reguły 3 otrzymujemy $A \vdash EGP f \rightarrow f \wedge EXP(EXP^k f)$, czyli $A \vdash EGP f \rightarrow f \wedge EXP(f \wedge EXP(f \wedge \dots EXP(f)))$, gdzie operator EXP występuje $k + 1$ razy. Zatem $A \vdash EGP f \rightarrow EXP^{k+1} f$.

Niech $A \vdash EGPf \rightarrow EXP^i f$ dla każdego $i \in \mathbb{N}$, zatem wykorzystując regułę 3 $A \vdash pref_\beta EGPf \rightarrow pref_\beta EXP^i f$, czyli $\| pref_\beta EGPf \| \leq \| pref_\beta EXP^i f \|$ dla każdego $i \in \mathbb{N}$. Zatem $\| pref_\beta EGPf \|$ jest dolnym ograniczeniem zbioru $\{\| pref_\beta EXP^i f \|\}_{i \in \mathbb{N}}$.

Następnie założmy, że istnieje formuła f_1 taka, że $\| f_1 \| \leq \| pref_\beta EXP^i f \|$ dla każdego $i \in \mathbb{N}$, czyli $A \vdash f_1 \rightarrow pref_\beta EXP^i f$. Korzystając z reguły 1, otrzymujemy $A \vdash f_1 \rightarrow pref_\beta EGPf$, czyli $\| f_1 \| \leq \| pref_\beta EGPf \|$. Ostatecznie $\| pref_\beta EGPf \|$ jest największym dolnym ograniczeniem zbioru $\{\| pref_\beta EXP^i f \|\}_{i \in \mathbb{N}}$.

b) Najpierw pokażemy, że $A \vdash EXP^i(f_1, f_2) \rightarrow EUP[f_1, f_2]$ dla każdego $i \in \mathbb{N}$. Przeprowadzimy dowód indukcyjny ze względu na i .

1° Dla $i = 1$ otrzymujemy $A \vdash EXP^1(f_1, f_2)$, czyli $A \vdash f_1 \wedge EXP(f_2)$. Na podstawie aksjomatu 10 i reguły 3 otrzymujemy $A \vdash EXP^1(f_1, f_2) \rightarrow f_1 \wedge EXP(EUP[f_1, f_2])$. Następnie z aksjomatu 10 $A \vdash EXP^1(f_1, f_2) \rightarrow EUP[f_1, f_2]$.

2° Założenie indukcyjne: $A \vdash EXP^k(f_1, f_2) \rightarrow EUP[f_1, f_2]$ dla pewnego $k \in \mathbb{N}$.

Teza indukcyjna: $A \vdash EXP^{k+1}(f_1, f_2) \rightarrow EUP[f_1, f_2]$.

Dowód: $A \vdash EXP^{k+1}(f_1, f_2)$, czyli $A \vdash f_1 \wedge EXP(f_1 \wedge \dots \wedge EXP(f_2))$, gdzie operator EXP występuje $k + 1$ razy. Zatem $A \vdash EXP^{k+1}(f_1, f_2) \rightarrow f_1 \wedge EXP(EXP^k(f_1, f_2))$ oraz na podstawie założenia indukcyjnego i reguły 3 $A \vdash EXP^{k+1}(f_1, f_2) \rightarrow f_1 \wedge EXP(EUP[f_1, f_2])$. Następnie z aksjomatu 10 $A \vdash EXP^{k+1}(f_1, f_2) \rightarrow EUP[f_1, f_2]$.

Niech $A \vdash EXP^i(f_1, f_2) \rightarrow EUP[f_1, f_2]$ dla każdego $i \in \mathbb{N}$, zatem wykorzystując regułę 3 mamy

$A \vdash pref_\beta EXP^i(f_1, f_2) \rightarrow pref_\beta EUP[f_1, f_2]$, czyli $\| pref_\beta EXP^i(f_1, f_2) \| \leq \| pref_\beta EUP[f_1, f_2] \|$ dla każdego $i \in \mathbb{N}$. Zatem $\| pref_\beta EUP[f_1, f_2] \|$ jest górnym ograniczeniem zbioru $\{\| pref_\beta EXP^i(f_1, f_2) \|\}_{i \in \mathbb{N}}$.

Następnie założmy, że istnieje formuła f taka, że $\| pref_\beta EXP^i(f_1, f_2) \| \leq \| f \|$ dla każdego $i \in \mathbb{N}$. Zatem $A \vdash pref_\beta EXP^i(f_1, f_2) \rightarrow f$. Stosując regułę 2, otrzymujemy $A \vdash pref_\beta EUP[f_1, f_2] \rightarrow f$, czyli $\| pref_\beta EUP[f_1, f_2] \| \leq \| f \|$. Ostatecznie $\| pref_\beta EUP[f_1, f_2] \|$ jest najmniejszym górnym ograniczeniem zbioru $\{\| pref_\beta EXP^i(f_1, f_2) \|\}_{i \in \mathbb{N}}$.

Z powyższego twierdzenia wynika, że algebra Lindenbauma może być rozważana jako algebra Boolea z przeliczalnym zbiorem Q następujących nieskończonych operacji:

(1) $\inf\{\| pref_\beta EXP^i f \|\}_{i \in \mathbb{N}}$

(2) $\sup\{\| pref_\beta EXP^i(f_1, f_2) \|\}_{i \in \mathbb{N}}$

dla dowolnego procesu P , dowolnego ciągu akcji atomowych β oraz formuł f, f_1, f_2 .

Definicja 7. Q -filtrem D nazywamy maksymalny, właściwy filtr taki, że:

1) jeśli $\| pref_\beta EUP[f_1, f_2] \| \in D$, to istnieje $i \in \mathbb{N}$ takie, że

$\| pref_\beta EXP^i(f_1, f_2) \| \in D$,

2) jeśli $\| pref_\beta EGPf \| \notin D$, to istnieje $i \in \mathbb{N}$ takie, że $\| pref_\beta EXP^i f \| \notin D$.

Lemat 2. (Rasiowa-Sikorski)

Jeśli w algebrze Boolea zbiór Q nieskończonych operacji jest przeliczalny, to każdy niezerowy element tej algebry jest zawarty w pewnym Q -filtrze.

Dowód lematu 2 można znaleźć w [9].

Definicja 8. Niech ρ oznacza funkcję, która każdej formule przyporządkowuje liczbę porządkową w następujący sposób:

$\rho(p) = 1$ jeśli $p \in V_o$,

$\rho(\neg f) = \rho(f) + 1$,

$\rho(f_1 \wedge f_2) = \max(\rho(f_1), \rho(f_2)) + 1$,

$\rho(f_1 \vee f_2) = \max(\rho(f_1), \rho(f_2)) + 1$,

$\rho(f_1 \rightarrow f_2) = \max(\rho(f_1), \rho(f_2)) + 1$,

$\rho(EXP f) = \rho(f) + 1$,

$\rho(AXP f) = \rho(f) + 1$,

$\rho(EGPf) = \omega \times \rho(f) + 1$,

$\rho(AGPf) = \omega \times \rho(f) + 1$,

$\rho(EUP[f_1, f_2]) = \omega \times \max(\rho(f_1), \rho(f_2)) + 1$,

$\rho(AUP[f_1, f_2]) = \omega \times \max(\rho(f_1), \rho(f_2)) + 1$.

Definicja 9. Niech $<$ oznacza relację określoną w zbiorze formuł F w następujący sposób: $f_1 < f_2 \stackrel{\text{def}}{\iff} \rho(f_1) < \rho(f_2)$, gdzie $<$ jest porządkiem określonym na liczbach porządkowych.

Lemat 3. Relacja $\prec$ spełnia następujące warunki:

$f_1 \prec \neg f_1$,
 $f_1 \prec f_1 \wedge f_2$, $f_2 \prec f_1 \wedge f_2$,
 $f_1 \prec f_1 \vee f_2$, $f_2 \prec f_1 \vee f_2$,
 $f_1 \prec f_1 \rightarrow f_2$, $f_2 \prec f_1 \rightarrow f_2$,
 $f_1 \prec EXP f_1$, $f_1 \prec AXP f_1$,
dla każdego $i \in \mathbb{N}$, $EXP^i f_1 \prec EGP f_1$,
dla każdego $i \in \mathbb{N}$, $EXP^i(f_1, f_2) \prec EUP[f_1, f_2]$.

Lemat 4. Relacja $\prec$ jest relacją porządku w zbiorze formuł F .

Lemat 5. Relacja $\prec$ jest dobrze ufundowana, tzn. dla każdego zbioru formuł Z istnieje formuła f będąca najmniejszym elementem zbioru Z ze względu na relację $\prec$.

Definicja 10. Kanoniczna struktura teorii T jest to następująca trójka

$$M = \langle QF, \rightarrow, w \rangle$$

gdzie: QF jest rodziną wszystkich Q -filtrów w algebrze Lindenbaum teorii T ,
 $\rightarrow$ jest rodziną funkcji $\{\overset{a}{\rightarrow}\}_{a \in A}$,
 $\overset{a}{\rightarrow}$ jest funkcją, która każdemu Q -filtrowi D przyporządkowuje Q -filtr D' taki, że $\| EX[true : a]f \| \in D \rightarrow \| f \| \in D'$ dla każdego $f \in F$,
 w jest wartościowaniem takim, że $w(p, D) = 1$ wttw, gdy $\| p \| \in D$, dla każdego $p \in V_o$, $D \in QF$.

Lemat 6. Jeśli D jest Q -filtrem i $\| EXP f \| \in D$, to istnieje Q -filtr D' i akcja atomowa $a \in A(P)$ takie, że $D \xrightarrow{a} D'$ i $\| f \| \in D'$.

Dowód:

Niech D będzie Q -filtrem i $\| EXP f \| \in D$. Załóżmy, że P jest procesem mającym postać $[p_i : a_i]_{i=1, \dots, k}$. Na podstawie aksjomatu 17 mamy:
 $EXP f \leftrightarrow (p_1 \wedge EX[true : a_1]f) \vee \dots \vee (p_k \wedge EX[true : a_k]f) \vee (f \wedge \neg p_1 \wedge \dots \wedge \neg p_k)$. Jeśli $\| EXP f \| \in D$, to $\| (p_1 \wedge EX[true : a_1]f) \vee \dots \vee (p_k \wedge EX[true : a_k]f) \vee (f \wedge \neg p_1 \wedge \dots \wedge \neg p_k) \| \in D$.

D jest Q -filtrem, czyli filtrem maksymalnym i pierwszym. Zatem na podstawie definicji filtru pierwszego i twierdzenia 4, istnieje $1 \leq i \leq k$ takie, że

$\| p_i \| \in D$ i $\| EX[true : a_i]f \| \in D$ lub $\| f \wedge \neg p_1 \wedge \dots \wedge \neg p_k \| \in D$.

Jeśli zachodzi drugi przypadek: $\| f \wedge \neg p_1 \wedge \dots \wedge \neg p_k \| \in D$, to $\| f \| \in D$ i dla każdego $i \leq k$ $\| p_i \| \notin D$. Zatem możemy przyjąć $D' = D$ i $a = a^*$.

Rozważmy przypadek pierwszy: $\| p_i \| \in D$ i $\| EX[true : a_i]f \| \in D$ dla pewnego $1 \leq i \leq k$.

Niech $D' = \{ \alpha \mid \| EX[true : a_i]\alpha \| \in D \}$. Zauważmy, że $\| f \| \in D'$. Zatem D' nie jest zbiorem pustym.

Najpierw pokażemy, że D' jest filtrem. Niech $\| f_1 \|, \| f_2 \| \in D'$. Zatem $\| EX[true : a_i]f_1 \| \in D$ i $\| EX[true : a_i]f_2 \| \in D$ (z definicji zbioru D'). Ponieważ D jest filtrem, to $\| EX[true : a_i]f_1 \| \cap \| EX[true : a_i]f_2 \| \in D$. Następnie możemy przeprowadzić następujące rozumowanie:

$\| EX[true : a_i]f_1 \| \cap \| EX[true : a_i]f_2 \| \in D \Leftrightarrow$ (z twierdzenia 4)
 $\| EX[true : a_i]f_1 \wedge EX[true : a_i]f_2 \| \in D \Leftrightarrow$ (z aksjomatu 16) $\| EX[true : a_i](f_1 \wedge f_2) \| \in D \Leftrightarrow$ (z definicji zbioru D') $\| f_1 \wedge f_2 \| \in D' \Leftrightarrow$ (z twierdzenia 4) $\| f_1 \| \cap \| f_2 \| \in D'$.

Teraz wykażemy, że D' jest filtrem właściwym. Załóżmy, że tak nie jest, czyli że $\| false \| \in D'$.

Zatem $\| false \| \in D' \Leftrightarrow$ (z definicji zbioru D') $\| EX[true : a_i]false \| \in D \Leftrightarrow$ (z faktu, że D jest maksymalnym filtrem) $\| \neg EX[true : a_i]false \| \notin D \Leftrightarrow$ (z aksjomatu 1) $\| AX[true : a_i]true \| \notin D$ i otrzymujemy sprzeczność z aksjomatem 19. Ostatecznie D' jest filtrem właściwym.

Następnie pokażemy, że D' jest filtrem pierwszym, czyli maksymalnym. Załóżmy, że dla dowolnych formuł f_1 i f_2 , $\| f_1 \| \cup \| f_2 \| \in D'$, czyli $\| f_1 \vee f_2 \| \in D'$. Zatem $\| EX[true : a_i](f_1 \vee f_2) \| \in D$ i na podstawie następującego faktu: $EXP(f_1 \vee f_2) \rightarrow EXP f_1 \vee EXP f_2$ mamy $\| EX[true : a_i]f_1 \vee EX[true : a_i]f_2 \| \in D$. Ponieważ D jest Q -filtrem, to $\| EX[true : a_i]f_1 \| \in D$ lub $\| EX[true : a_i]f_2 \| \in D$. Dlatego $\| f_1 \| \in D'$ lub $\| f_2 \| \in D'$, czyli że D' jest filtrem pierwszym.

Na zakończenie dowodu, należy jeszcze sprawdzić, czy zbiór D' spełnia warunki Q -filtru. Niech $P_1 \in \Pi(A_t)$, $f_1, f_2 \in F$, β będzie ciągiem akcji $b_1, b_2, \dots, b_n$

i β' będzie ciągiem akcji $a_i, b_1, \dots, b_n$.

Załóżmy, że $\| \text{pref}_\beta \text{EUP}_1[f_1, f_2] \| \in D'$. Korzystając z definicji zbioru D' otrzymujemy $\| \text{EX}[true : a_i] \text{pref}_\beta \text{EUP}_1[f_1, f_2] \| \in D$. Dlatego

$\| \text{pref}_{\beta'} \text{EUP}_1[f_1, f_2] \| \in D$. Ponieważ D jest Q -filtrem, to istnieje $i \in \mathbb{N}$ takie, że $\| \text{pref}_{\beta'} \text{EXP}_1^i(f_1, f_2) \| \in D$ i $\| \text{EX}[true : a_i] \text{pref}_\beta \text{EXP}_1^i(f_1, f_2) \| \in D$. Zatem $\| \text{pref}_\beta \text{EXP}_1^i(f_1, f_2) \| \in D'$.

Załóżmy, że $\| \text{pref}_\beta \text{EGP}_1 f_1 \| \notin D'$. Korzystając z definicji zbioru D' otrzymujemy: $\| \text{EX}[true : a_i] \text{pref}_\beta \text{EGP}_1 f_1 \| \notin D$. Dlatego $\| \text{pref}_{\beta'} \text{EGP}_1 f_1 \| \notin D$. Ponieważ D jest Q -filtrem, to istnieje $i \in \mathbb{N}$ takie, że $\| \text{pref}_{\beta'} \text{EXP}_1^i f_1 \| \notin D$ i $\| \text{EXP}[true : a_i] \text{pref}_\beta \text{EXP}_1^i f_1 \| \notin D$. Zatem $\| \text{pref}_\beta \text{EXP}_1^i f_1 \| \notin D'$, czyli D' jest Q -filtrem.

Ostatecznie z definicji funkcji $\xrightarrow{a_i}$ i definicji zbioru D' otrzymujemy: $D \xrightarrow{a_i} D'$.

Lemat 7. Dla każdego $f \in F$ i $D \in QF$ zachodzi: $M, D \models f$ wttw, gdy $\| f \| \in D$.

Dowód:

Przeprowadzimy dowód indukcyjny ze względu na strukturę formuły f .

1°. Niech $f = p$, $p \in V_o$. Zatem $M, D \models p$ wttw, gdy $\| p \| \in D$, co wynika z definicji wartościowania w .

2° Założenie indukcyjne: Niech $f \in F$. Dla każdego $g \in F$ takiego, że $g < f$ zachodzi: $M, D \models g$ wttw, gdy $\| g \| \in D$.

Teza indukcyjna: $M, D \models f$ wttw, gdy $\| f \| \in D$.

Dowód:

1. Załóżmy, że $f = \neg f_1$ i $\| f \| \in D$. Zatem $\| \neg f_1 \| \in D$. Ponieważ D jest filtrem maksymalnym, to $\| f_1 \| \notin D$. Korzystając z założenia indukcyjnego otrzymujemy $\| f_1 \| \notin D$ wttw, gdy $M, D \not\models f_1$. Zatem $M, D \models \neg f_1$, czyli $M, D \models f$.

2. Załóżmy, że $f = f_1 \wedge f_2$ i $\| f \| \in D$. Zatem $\| f_1 \wedge f_2 \| \in D$. Na podstawie definicji filtru $\| f_1 \| \in D$ i $\| f_2 \| \in D$. Korzystając z założenia indukcyjnego otrzymujemy $M, D \models f_1$ i $M, D \models f_2$. Zatem $M, D \models f_1 \wedge f_2$, czyli $M, D \models f$.

3. Załóżmy, że $f = f_1 \vee f_2$ i $\| f \| \in D$. Zatem $\| f_1 \vee f_2 \| \in D$. Na podstawie definicji Q -filtru $\| f_1 \| \in D$ lub $\| f_2 \| \in D$. Korzystając z założenia indukcyjnego otrzymujemy $M, D \models f_1$ lub $M, D \models f_2$. Zatem $M, D \models f_1 \vee f_2$, czyli $M, D \models f$.

4. Załóżmy, że $f = f_1 \rightarrow f_2$ i $\| f \| \in D$. Zatem $\| \neg f_1 \vee f_2 \| \in D$. Na podstawie definicji Q -filtru $\| f_1 \| \notin D$ lub $\| f_2 \| \in D$. Korzystając z założenia induk-

cyjnego otrzymujemy $M, D \not\models f_1$ lub $M, D \models f_2$. Zatem $M, D \models \neg f_1 \vee f_2$ i $M, D \models f_1 \rightarrow f_2$. Zatem $M, D \models f$.

5. Załóżmy, że $f = \text{EXP} f_1$ i $\| f \| \in D$. Zatem $\| \text{EXP} f_1 \| \in D$. Na podstawie lematu 6 istnieje Q -filtr D' i akcja atomowa $a \in A(P)$ takie, że $D \xrightarrow{a} D'$ oraz $\| f_1 \| \in D'$. Korzystając z założenia indukcyjnego otrzymujemy $\| f_1 \| \in D'$ wttw, gdy $M, D' \models f_1$. Zatem (z definicji semantyki i modelu kanonicznego M) $M, D \models \text{EXP} f_1$, czyli $M, D \models f$.

Załóżmy, że $M, D \models \text{EXP} f_1$. Z definicji semantyki i struktury M istnieje Q -filtr D' i akcja atomowa $a \in A(P)$ takie, że $D \xrightarrow{a} D'$ i $M, D' \models f_1$. Korzystając z założenia indukcyjnego otrzymujemy $\| f_1 \| \in D'$ oraz z definicji funkcji $\xrightarrow{a}$ mamy $\| \text{EXP} f_1 \| \in D$.

6. Załóżmy, że $f = \text{EGP} f_1$. Zatem $\| \text{EGP} f_1 \| \in D \Leftrightarrow$ (na podstawie definicji Q -filtru i lematu 1) $\| \text{EXP}^i f_1 \| \in D$ dla każdego $i \in \mathbb{N} \Leftrightarrow$ (na podstawie założenia indukcyjnego) $M, D \models \text{EXP}^i f_1$ dla każdego $i \in \mathbb{N} \Leftrightarrow M, D \models \text{EGP} f_1 \Leftrightarrow M, D \models f$.

7. Załóżmy, że $f = \text{EUP}[f_1, f_2]$. Zatem $\| \text{EUP}[f_1, f_2] \| \in D \Leftrightarrow$ (z lematu 1 i definicji Q -filtru) $\| \text{EXP}^i(f_1, f_2) \| \in D$ dla pewnego $i \in \mathbb{N} \Leftrightarrow$ (na podstawie założenia indukcyjnego) $M, D \models \text{EXP}^i(f_1, f_2)$ dla pewnego $i \in \mathbb{N} \Leftrightarrow M, D \models \text{EUP}[f_1, f_2] \Leftrightarrow M, D \models f$.

Dowód dla formuł postaci $\text{AXP} f$, $\text{AGP} f$, $\text{AUP}[f_1, f_2]$ wynika z następujących równoważności:

$\text{AXP}(f) \longleftrightarrow \neg \text{EXP}(\neg f)$ (aksjomat 1),

$\text{AGP} f \longleftrightarrow \neg \text{EUP}[true, \neg f]$ (aksjomat 8),

$\text{AUP}[f_1, f_2] \longleftrightarrow \neg \text{EGP}(\neg f_2) \wedge \neg \text{EUP}[\neg f_2, \neg f_1 \wedge \neg f_2]$ (aksjomat 9).

Twierdzenie 5. Dla każdej formuły f niesprzecznej teorii $T = \langle L, C, A \rangle$ następujące warunki są równoważne:

(1) formuła f jest twierdzeniem teorii T ,

(2) formuła f jest prawdziwa w każdym modelu teorii T .

Dowód:

Implikacja (1) $\rightarrow$ (2) wynika z twierdzenia 2.

Udowodnimy implikację odwrotną (2) $\rightarrow$ (1).

Załóżmy, że $f \in F$, f jest prawdziwa w każdym modelu teorii T i f nie jest twierdzeniem teorii T . Stosując twierdzenie 4 otrzymujemy $\| \neg f \| \neq 0$. Zatem, na podstawie lematu 2, istnieje Q -filtr D taki, że $\| \neg f \| \in D$. Korzystając z lematu 7 otrzymujemy $M, D \models \neg f$, czyli $M, D \not\models f$. Ostatecznie M nie jest

modelem formuły f , a więc formuła f nie jest prawdziwa w każdym modelu teorii T . Otrzymaliśmy sprzeczność, która kończy dowód.

6 Podsumowanie

Zaprezentowana w pracy logika może być stosowana do weryfikacji temporalnych własności algorytmów niedeterministycznych. W literaturze, można znaleźć różne temporalne logiki czasu rozgałęzionego, np. [2], [8]. TLP wyróżnia się tym, że proces jest składową języka i występuje w formułach. W przyszłości zostanie podjęta próba stworzenia podobnego systemu dla programów współbieżnych i rozproszonych.

Bibliografia

- [1] L. Bolc, K. Dziewicki, P. Rychlik, A. Szalas, Wnioskowanie w logikach nieklasycznych. Podstawy teoretyczne, Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1995
- [2] E. A. Emerson, J. Y. Halpern, Decision Procedures and Expressiveness in the Temporal Logic of Branching Time, *Journal of Computer and System Sciences* 30, 1985, 1-24
- [3] L. Fix, O. Grumberg, Verification of Temporal Properties, *Journal of Logic and Computation*, vol. 6 No. 3, 1996, 343-361
- [4] D. Harel, Dynamic Logic, *Handbook of Philosophical Logic*, vol. 2, D. Reidel Pub. Co., 1984, 497-607
- [5] F. Kröger, A Generalized Nexttime Operator in Temporal Logic, *Journal of Computer and System Sciences* 29, 1984, 80-98
- [6] G. Mirkowska, A. Salwicki, *Algorithmic Logic*, PWN, Polish Scientific Publishers, Warszawa, 1987
- [7] W. Penczek, A Temporal Logic for Event Structures, *Fundamenta Informaticae* XI, 1988, 297-326
- [8] W. Penczek, Axiomatizations of Temporal Logic on Trace Systems, *Fundamenta Informaticae* 25, 1996, 183-200
- [9] H. Rasiowa, R. Sikorski, *The Mathematics of Metamathematics*, PWN, Warszawa, 1970

- [10] A. Szalas, U. Petermann, On Temporal Logic for Distributed Systems and Its Application to Processes Communicating by Interrupts, *Fundamenta Informaticae* 12, 1989, 191-204

VERIFICATION OF TEMPORAL PROPERTIES OF NONDETERMINISTIC ALGORITHMS

Abstract

This paper provides one of the versions of propositional, branching time temporal logic, called Process Temporal Logic. It allows to verify static and dynamic properties of programs, which are expressed as temporal formulas. PTL system is presented with the infinitary axiomatic system and the proof of the completeness theorem. The majority of temporal logics treats a program as a separate object and not as a language element. The system shown in this article is different. Program is a part of expression which describes its quality just as in dynamic and algorithmic logic.

Vladimir Marchenko and Zbigniew Zaczekiewicz*

TO THE PROBLEM OF OBSERVABILITY FOR NONLINEAR SYSTEMS WITH TIME DELAY

Abstract: Several kinds of observability notions are introduced for nonlinear time-delay systems and comparative analysis of such notions is presented. Sufficient conditions for observability are given. Special attention is paid to the two dimensional system case. The results obtained are illustrated by examples.

1 Introduction

The paper deals with the problem of observability for nonlinear dynamical systems with time delay. In the no time delay case the problem was intensively investigated by many authors (see [1]-[6] and their references). In the case of time-delay systems we transform the system under consideration to some "canonical" form and obtain some sufficient conditions of observability for the original time delay system by analyzing the problem of observability for such a canonical system.

2 Preliminaries

Let us consider an output system Σ of the form:

$$(\Sigma) : \begin{cases} \dot{x}(t) = f(t, x(t), x(t-h)) \\ y(t) = g(t, x(t), x(t-h)) \end{cases} \quad \text{for } t > t_0 \quad (2.1)$$

with initial conditions

$$x(\tau) = \phi(\tau), \quad \tau \in [t_0 - h, t_0], \quad x(t_0^+) = \phi_0 \in \mathbb{R}^n \quad (2.2)$$

*Supported by KBN under the Technical University of Białystok grant No. W/IMF/1/01

where h is a constant delay, $h > 0$; $x(t) \in \mathbb{R}^n$, $y(t) \in \mathbb{R}^m$, n -vector-function f and initial data ϕ and ϕ_0 are such that the corresponding solution $x(t)$ of the system exists, is unique and continuous for $t > t_0$. We also assume that the m -vector-function g is continuous in its arguments.

By the symbol $x(t, t_0, \phi, \phi_0)$ we denote the solution of system Σ corresponding to the initial conditions (2.2) and evaluated at time $t > t_0$.

Similarly, the outputs corresponding to two different solutions $x(t) = x(t, t_0, \phi, \phi_0)$ and $\tilde{x}(t) = x(t, t_0, \tilde{\phi}, \tilde{\phi}_0)$ are denoted by the symbols $y(t, t_0, \phi, \phi_0)$ and $\tilde{y}(t, t_0, \tilde{\phi}, \tilde{\phi}_0)$ or briefly $y(t)$ and $\tilde{y}(t)$, $t > t_0$, respectively.

Definition 1. Let s be a real non-negative number. System Σ is said to be:

1) s -observable in (τ_1, τ_2) (by measurement in the interval $(\tau_1, \tau_2) \subset (t_0, +\infty)$) if

$$y(t) = \tilde{y}(t) \text{ for } t \in [\tau_1, \tau_2] \Rightarrow x(t) = \tilde{x}(t) \text{ for } t \geq t_0 + s;$$

2) weakly observable in (τ_1, τ_2) if

$y(t) = \tilde{y}(t)$ for $t \in (\tau_1, \tau_2) \Rightarrow$ there exists a time moment $t_1 > t_0$ such that $x(t) = \tilde{x}(t)$ for all $t > t_1$;

3) superstrongly observable in $[\tau_1, \tau_2]$ if

$$y(t, t_0, \phi, \phi_0) = \tilde{y}(t, t_0, \tilde{\phi}, \tilde{\phi}_0) \text{ for } t \in (\tau_1, \tau_2) \Rightarrow \phi = \tilde{\phi} \text{ and } \phi_0 = \tilde{\phi}_0;$$

4) linearly s -observable in (τ_1, τ_2) if for any n -vector p_0 and for any integrable in the interval $[t_0 + s, t_0 + s + h]$ n -vector-function $p(\cdot)$ there exists an integrable in $[\tau_1, \tau_2]$ m -vector-function $\nu(\cdot)$ such that the following condition

$$\int_{\tau_1}^{\tau_2} \nu^T y(t, t_0, \phi, \phi_0) dt = p_0^T x(t_0 + s + h + 0, t_0, \phi, \phi_0) + \int_{t_0 + s}^{t_0 + s + h} p^T(t) x(t, t_0, \phi, \phi_0) dt$$

is valid for all admissible initial data ϕ and ϕ_0 (here $()^T$ = transposition);

5) observable if it is s -observable in (τ_1, τ_2) with $s = 0$ and $(\tau_1, \tau_2) = (t_0, +\infty)$. It means that every solution of the system can be distinguished by the output measurement, i.e.

$$y(t) = \tilde{y}(t) \text{ for } t > t_0 \Rightarrow x(t) = \tilde{x}(t) \text{ for } t > t_0.$$

Consider the following special case of system (2.1):

$$\begin{aligned} (\Sigma_1) : \quad & \dot{x}(t) = A(t, x(t), x(t-h))x(t) + A_1(t, x(t), x(t-h))x(t-h) \\ & y(t) = C(t, x(t), x(t-h))x(t), \quad t > t_0, \end{aligned} \quad (2.3)$$

where elements of $(n \times n)$ -matrices A , A_1 and $(1 \times n)$ -matrix C are n times continuously differentiable with respect to argument t along all the solutions $x(t)$, $t > t_0 - h$, of the system.

Introduce a determining equation for system (2.3):

$Q_{k+1}(x)(t) = Q_k(x)(t)A(t, x(t), x(t-h)) + \frac{d}{dt}Q_k(x)(t)$ for $k = 0, 1, \dots, n-1$ with initial conditions of the form $Q_0(x)(t) = C(t, x(t), x(t-h))$ where $x = x(t)$, $t > t_0$, is a solution of system (2.3).

3 Main results

We can state

Proposition 2. The following implications are valid:

a) "linear s -observability in (τ_1, τ_2) " $\Rightarrow$ " s -observability in (τ_1, τ_2) " $\Rightarrow$ "weak observability in (τ_1, τ_2) ";

b) "superstrong observability in $(t_0, +\infty)$ " $\Rightarrow$ "observability" $\Rightarrow$ " s -observability in $(t_0, +\infty)$ for $s \geq 0$ " $\Rightarrow$ "weak observability in $(t_0, +\infty)$."

Theorem 3. If along all the solutions $x(t)$, $t > t_0 - h$, of system (2.3) the following conditions are satisfied:

$$i) \det W(x)(t) \stackrel{\text{def}}{=} \det \begin{bmatrix} Q_0(x)(t) \\ Q_1(x)(t) \\ \vdots \\ Q_{n-1}(x)(t) \end{bmatrix} \neq 0, \quad \forall t > t_0;$$

ii) mapping $z(t) = W(x)(t)x(t)$, $t > t_0$, is superstrongly injective in the sense that

$$z(t) = \tilde{z}(t), \quad t > t_0, \Rightarrow x(t) = \tilde{x}(t) \text{ for } t > t_0 - h;$$

$$iii) Q_k(x)(t)A_1(t, x(t), x(t-h)) = 0 \text{ for } t > t_0 \text{ and } k = 0, 1, \dots, n-1;$$

then system Σ_1 is superstrongly observable in $(t_0, +\infty)$.

Proof. Let us define $x(\tau) = 0$ for $\tau \leq t_0 - h$ and consider the following transformation $z(t) = W(t, x)x(t)$. Then we have

$$\begin{aligned} \dot{z}(t) &= \frac{d}{dt}(W(x)(t)x(t)) = \frac{dW(x)(t)}{dt}x(t) + W(x)(t)\dot{x}(t) \stackrel{(2.3)}{=} \\ &= \left(\frac{dW(x)(t)}{dt} + W(x)(t)A(t, x(t), x(t-h)) \right) x(t) + \\ &+ \begin{bmatrix} Q_0(x)(t) \\ \vdots \\ Q_{n-1}(x)(t) \end{bmatrix} A_1(t, x(t), x(t-h)) x(t-h) = \\ &= \begin{bmatrix} Q_1(x)(t) \\ \vdots \\ Q_n(x)(t) \end{bmatrix} W^{-1}(x)(t)z(t) + \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \\ \beta_1 & \beta_2 & \dots & \beta_N \end{bmatrix} z(t-h) \end{aligned}$$

If follows from here that

$$\begin{aligned} \dot{z}(t) &= \begin{bmatrix} 0 & 1 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \vdots & 0 & 0 \\ \dots & \dots & \dots & \dots & \ddots & \dots & \dots \\ 0 & 0 & 0 & 0 & \vdots & 0 & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \dots & \alpha_{n-1} & \alpha_n \end{bmatrix} z(t) + \\ &+ \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \\ \beta_1 & \beta_2 & \dots & \beta_n \end{bmatrix} z(t-h) \end{aligned} \quad (3.1)$$

$$y(t) = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} z(t) \quad (3.2)$$

where $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$ are some non-zero in general functions of $t, x(t)$ and $x(t-h)$. It is not difficult to directly check that system (3.1) with output (3.2) is observable. Thus, if we assume that $y(t) = \tilde{y}(t)$ for $t > t_0$ then we have that $z(t) = \tilde{z}(t)$ for $t > t_0$, i.e. $W(x)(t)x(t) = W(\tilde{x})(t)\tilde{x}(t)$ for $t > t_0$,

and by assumption (ii) we conclude that $x(t) = \tilde{x}(t)$ for $t > t_0 - h$. Hence the system is superstrongly observable in $(t_0, +\infty)$ that finishes the proof.

Remark 4. Examples show that the property of superstrong observability is stronger than the well-known Kalman's type of observability. For example, a necessary condition for the superstrong observability of the following linear stationary system

$$\dot{x}(t) = Ax(t) + A_1x(t-h), \quad y(t) = Cx(t), \quad t > 0.$$

is $\det A_1 \neq 0$ that is not realized if we consider a system with no delay: $A_1 = 0$.

Corollary 5. Let the assumptions i) and iii) of Theorem 3 be valid and additionally ii) a mapping $z(t) = W(x)(t)x(t)$, $t > t_0$, is injective in sense that

$$z(t) = \tilde{z}(t) \text{ for } t > t_0 \Rightarrow x(t) = \tilde{x}(t), \quad t > t_0,$$

then a system (2.3) is observable.

The proof is very similar to the proof of Theorem 3.

Remark 6. By analogy with Theorem 3 and Corollary 5 we can give sufficient conditions of s -observability in $(t_0, +\infty)$ of system (2.3).

Let us consider the following stationary two dimensional system (2.3) with output:

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \end{bmatrix} \begin{bmatrix} x_1(t-h) \\ x_2(t-h) \end{bmatrix} \\ y(t) &= \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \quad t > t_0 \end{aligned} \quad (3.3)$$

where system matrix coefficients $a_{11}, a_{12}, a_{21}, a_{22}, a_{11}^1, a_{12}^1, a_{21}^1, a_{22}^1$ are Lipschitzian functions of $x_1(t), x_2(t), x_1(t-h), x_2(t-h)$ and c_1, c_2 are real numbers, $c_1^2 + c_2^2 \neq 0$.

Proposition 7. If along all the solutions $x_1(t), x_2(t)$, $t > t_0 - h$, of system (3.3) the following conditions hold:

$$i) \begin{bmatrix} c_1 & c_2 \end{bmatrix} \cdot \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} -c_2 \\ c_1 \end{bmatrix} \neq 0 \text{ for } t > t_0,$$

ii) $\begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_1 & -c_2 \\ c_2 & c_1 \end{bmatrix} \stackrel{\text{def}}{=} f(c_1 x_1(t) + c_2 x_2(t))$ is a (1×2) -matrix-function depending on the sum $c_1 x_1(t) + c_2 x_2(t)$ only,

iii) $\begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \end{bmatrix} = 0$ for $t > t_0$

then system (3.3) is observable.

Proof. Using the transformation $\begin{bmatrix} \omega_1(t) \\ \omega_2(t) \end{bmatrix} = \omega(t) = \begin{bmatrix} c_1 & c_2 \\ -c_2 & c_1 \end{bmatrix} x(t)$,
 $\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = x(t) = \frac{1}{\Delta} \begin{bmatrix} c_1 & -c_2 \\ c_2 & c_1 \end{bmatrix} \omega(t)$, $t > t_0$, where $\Delta = c_1^2 + c_2^2$, we have

$$\begin{aligned} \dot{\omega}(t) &= \begin{bmatrix} c_1 & c_2 \\ -c_2 & c_1 \end{bmatrix} \left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} x(t) + \begin{bmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \end{bmatrix} x(t-h) \right) \stackrel{(iii),(ii)}{=} \\ &= \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \omega(t) + \begin{bmatrix} 0 & 0 \\ \beta_{21} & \beta_{22} \end{bmatrix} \omega(t-h) \end{aligned} \quad (3.4)$$

with the output

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) = x_1(t), \quad t > t_0,$$

where elements α_{11} and α_{12} are functions of $\omega_1(t)$ only and

$$\alpha_{12} = \frac{1}{\Delta} \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} -c_2 \\ c_1 \end{bmatrix} \neq 0.$$

Check the assumptions of Theorem 3:

$$Q_0(\omega)(t) = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad Q_1(\omega)(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \end{bmatrix}$$

$$(i) \quad \det \begin{bmatrix} Q_0(\omega)(t) \\ Q_1(\omega)(t) \end{bmatrix} = \det \begin{bmatrix} 1 & 0 \\ \alpha_{11} & \alpha_{12} \end{bmatrix} = \alpha_{12} \neq 0$$

(it is satisfied);

$$(iii) \quad Q_0(\omega)(t) \begin{bmatrix} 0 & 0 \\ \beta_{21} & \beta_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} \quad (\text{it is satisfied});$$

$$(ii) \quad W(\omega)(t) \omega(t) = \begin{bmatrix} 1 & 0 \\ \alpha_{11} & \alpha_{12} \end{bmatrix} \omega(t). \quad \text{Assuming } W(\tilde{\omega})(t) \tilde{\omega}(t) = W(\omega)(t) \omega(t)$$

for $t > t_0$ we have $\begin{bmatrix} \tilde{\omega}_1(t) \\ \alpha_{11}(\tilde{\omega}_1(t)) \tilde{\omega}_1(t) + \alpha_{12}(\tilde{\omega}_1(t)) \tilde{\omega}_2(t) \end{bmatrix} =$
 $= \begin{bmatrix} \omega_1(t) \\ \alpha_{11}(\omega_1(t)) \omega_1(t) + \alpha_{12}(\omega_1(t)) \omega_2(t) \end{bmatrix}$ for $t > t_0$. It follows from here

that $\tilde{\omega}_1(t) = \omega_1(t)$ and $\tilde{\omega}_2(t) = \omega_2(t)$ for $t > t_0$.

Hence, the mapping $z(t) = W(\omega)(t) \omega(t)$, $t > t_0$, is injective. By Theorem 3 system (3.4) is observable. Therefore system (3.3) is also observable.

Remark 8. It is clear that the most difficult problem to apply Theorem 3 is checking the assumption (ii). Below we give an effective way for checking such an assumption in a scalar case.

Proposition 9. For a continuous function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ the following conditions are equivalent:

- 1) φ is injective;
- 2) φ is a monotonic decreasing or increasing function;
- 3) φ has no local extreme value.

Proof. 1) $\Rightarrow$ 2) If we assume that φ is not a monotonic function we can find different points v_1, v_2, v_3 where $v_1 < v_2 < v_3$ such that $f(v_1) \leq f(v_2)$ and $f(v_2) \geq f(v_3)$ or $f(v_1) \geq f(v_2)$ and $f(v_2) \leq f(v_3)$. In both cases by the Darboux theorem the function takes the same value at least for two different points that contradicts the statement 1). Then 1) implies 2).

It is obvious that 2) $\Rightarrow$ 3). Let us prove 3) $\Rightarrow$ 1). If a continuous function φ has no a local extreme value then it is strongly monotonic, i.e. for any $v_1 \neq v_2$ we have $\varphi(v_1) > \varphi(v_2)$ or $\varphi(v_1) < \varphi(v_2)$ that implies $\varphi(v_1) \neq \varphi(v_2)$ and φ is injective. The proof is complete.

In the general multidimensional case the problem of characterization of the property for a mapping to be injective is much more complicated.

4 Examples

Example 10. Consider the following system

$$\begin{aligned} \dot{x}_1 &= -6x_1 - 2x_1x_2 + 6x_1^3 + 11x_2x_1^2 + 4x_2 - 2x_2^2 - 2x_1x_2^2 + 2x_2^3 \\ &\quad + 6x_2x_1(t-h) + 6x_1x_2x_1(t-h) + 2x_2x_2(t-h) + 2x_1x_2x_2(t-h), \\ \dot{x}_2 &= -x_1 + 3x_1x_2 - 3x_2x_1^2 + x_1x_2^2 + x_2 + x_2^2 + x_2^3 - 3x_1(t-h) - 3x_2^2x_1(t-h) - \\ &\quad - x_2(t-h) - x_2^2x_2(t-h), \quad y(t) = x_1x_2^2 + x_1 + x_2^2. \end{aligned}$$

where

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix}, \quad x_1 = x_1(t), \quad x_2 = x_2(t),$$

$$A = \begin{bmatrix} -6 - 2x_2 + 6x_1^2 + 8x_1x_2 & 4 - 2x_2 + 3x_1^2 - 2x_1x_2 + 2x_2^2 \\ -1 + 3x_2 - 3x_1x_2 - x_2^2 & 1 + x_2 + 2x_1x_2 + x_2^2 \end{bmatrix},$$

$$A_1 = \begin{bmatrix} 6x_2 + 6x_1x_2 & 2x_2 + 2x_1x_2 \\ -3 - 3x_2^2 & -1 - x_2^2 \end{bmatrix}, \quad C = \begin{bmatrix} x_2^2 + 1 & x_2 \end{bmatrix},$$

then

$$W(x)(t) =$$

$$\begin{bmatrix} x_2^2 + 1 & x_2 \\ 6x_1^2 + (6x_2^3 + 6x_2)x_1 - 4x_2^3 - 6 - 4x_2 & (3 + 7x_2^2)x_1^2 + 6x_1x_2^2 + 8x_2^2 + 2x_2^4 + 4 \end{bmatrix},$$

$$\text{and } W(x)(t)x(t) = \begin{bmatrix} (x_2^2 + 1)x_1 + x_2^2 \\ 6x_1^3 + (13x_2^3 + 9x_2)x_1^2 + (2x_2^3 - 6 - 4x_2)x_1 + 4x_2 + 8x_2^3 + 2x_2^5 \end{bmatrix}$$

is injective and

$\det W(x)(t) = (10x_2^2 + 7x_2^4 + 3 - 6x_2)x_1^2 + 14x_2^4 + 4 + 2x_2^6 + 16x_2^2 + 6x_2$ is always positive. It is clear that all the assumptions of Corollary 5 are true. Thus the system is observable.

Example 11. As an example of unobservable system let us consider the following one

$$\dot{x}(t) = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} x(t) + \begin{bmatrix} A_{11}^1 & 0 \\ A_{21}^1 & A_{22}^1 \end{bmatrix} x(t-h), \quad y = [c_1, 0]x(t)$$

where $A_{11} = A_{11}(x_1, x_1(t-h))$, $A_{21} = A_{21}(x_1, x_2, x_1(t-h), x_2(t-h))$,

$A_{22} = A_{22}(x_1, x_2, x_1(t-h), x_2(t-h))$, $A_{11}^1 = A_{11}^1(x_1, x_1(t-h))$,

$A_{21}^1 = A_{21}^1(x_1, x_2, x_1(t-h), x_2(t-h))$,

$A_{22}^1 = A_{22}^1(x_1, x_2, x_1(t-h), x_2(t-h))$, $c_1 = c_1(x_1, x_1(t-h))$. It is clear that for

the solution $x(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\tilde{x}(t) = \begin{bmatrix} 0 \\ f(t) \end{bmatrix}$ with $f(t)$ satisfying the equation

$\dot{f}(t) = A_{22}f(t) + A_{22}^1f(t)$ we have $y(t) \equiv \tilde{y}(t)$ for $t \geq 0$. Thus the system is

not observable. It is not difficult to check that $W(x)(t) = \begin{bmatrix} W_{11} & 0 \\ W_{21} & 0 \end{bmatrix}$ and the assumptions of Corollary 5 are not valid.

References

- [1] Z. Bartosiewicz *Local observability of nonlinear systems*, Systems & Control Letters 25 (1995), 295-298.

- [2] Z. Bartosiewicz, M. Wyrwas, *On multiobservers for nonlinear systems*, in: Progress in Simulation, Modelling, Analysis and Synthesis of Modern Electrical and Electronic Devices and Systems, World Scientific, 1999.
- [3] Z. Bartosiewicz *Lokalna obserwowalność układów z czasem ciągłym i czasem dyskretnym*, Zeszyty Naukowe Politechniki Białostockiej, Matematyka Fizyka Chemia, 19 (1999)
- [4] Z. Bartosiewicz *Universes in nonlinear control*, Proceedings of the Conference Mathtools'99, Sankt Petersburg, 1999
- [5] D. Mozyrska, Z. Bartosiewicz *Local observability of systems on $\mathbb{R}^\infty$* , Proceedings of Fourteen International Symposium of Mathematical Theory of Networks and Systems MTNS 2000, Perpignan, France, 19-25 June 2000
- [6] Z. Bartosiewicz *Computational complexity of observability problems*, in: Control and Self-Organization in Nonlinear Systems. Proceedings of the First International Conference, 15-18 February 2000, Białystok, Poland, Eds. Z. Bartosiewicz, M. Marczak, E. Pawluszewicz, Białystok 2000
- [7] J. K. Hale, S. M. V. Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag New York, Inc.(1993).
- [8] E. B. Lee, A. Olbrot, *Observability and related structural results for linear hereditary systems.*, Int. J. Control vol 34 no 6 (1981), 1061-1078.
- [9] W. Rudin, *Podstawy analizy matematycznej*, Wydawnictwo Naukowe PWN SA (1998).

O PROBLEMIE OBSERWOWALNOŚCI DLA UKŁADÓW NIELINIOWYCH Z OPÓŹNIENIEM

Streszczenie

Wprowadzono kilka pojęć obserwowalności dla układów nieliniowych z opóźnieniem i przeprowadzono analizę porównawczą tych pojęć. Podano warunki wystarczające na obserwowalność. Szczególną uwagę zwrócono na układy dwuwymiarowe. Uzyskane wyniki zilustrowano przykładami.

Ewa Pawłuszewicz

REMARKS ON OBSERVABILITY OF EULER'S DISCRETIZATION OF CONTROL SYSTEMS *

Abstract: In the paper relations between observability, single-experiment observability and finite-state determinability of continuous-time system Σ_c and discrete-time system obtained by Euler's discretization of Σ_c are studied. All presented results are local, i.e. they hold in some neighborhood of a given point from state-space.

Key words: nonlinear system, Euler's discretization, distinguishability, observability

1 Introduction

The necessity to study the problem of observability of continuous-time and discrete-time systems is well understood from as well theoretical as practical point of view. The classical definition of distinguishability, that is a background of idea of observability, roughly says that two distinct states of the given control system are distinguishable if there exists a control that gives rise to distinct outputs [2, 1, 3]. Then the control system is observable if for every pair of different states there exists control that distinguishes them. It was proved in [5, 7, 8] that for analytic systems there exists an analytic universal control, i.e. control that distinguishes every pair of states that can be distinguished by some control.

The aim of this paper is to study relations between observability of continuous-time system Σ_c and discrete-time system Σ_d , obtained from Σ_c by Euler's discretization. The background for our considerations are definitions given by E.D.Sontag [5, 6].

The paper is organized as follows: in Section 2 there is given basic notation. Section 3 presents ideas of distinguishability of events and states. It is proved

*Supported by KBN under the Technical University of Białystok grant No. W/IMF/3/99

that for any pair of states there exists a small discretization step in Euler's method such that if two different events are distinguishable by some control in the discrete-time system then they are distinguishable by similar control in the continuous-time case and vice versa. In Section 4 there are considered relations between observability, single-experiment observability, finite-state determinability of continuous-time system and its Euler's discretization. The observability and single-experiment observability are far from equivalent, at least for finite automata. For linear systems they are equivalent. For discrete-time systems single-experiment observability is rather restrictive. We proved that continuous-time system Σ_c is single-experiment observable if and only if discrete-time system Σ_d is. Single-experiment observability implies final-state determinability. It was proved [5] that a continuous-time system is observable if and only if it is final-state observable. We prove that continuous-time system Σ_c is final-state observable if and only if discrete-time system Σ_d is observable. All presented results are local, i.e. hold in some neighborhood of a certain point from state-space of the given system.

2 Preliminaries

Let us consider a nonlinear continuous-time (i.e. $t \in \mathbb{R}$) system defined on $\mathbb{R}^n$ with controls $u \in \mathbb{R}^m$ and outputs $y \in \mathbb{R}^p$:

$$\Sigma_C : \begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ y(t) = h(x(t)) \end{cases} \quad (1)$$

By $\mathcal{U}^{[\sigma, \tau]}$ we will denote the set of all controls defined on interval $[\sigma, \tau] \subset \mathbb{R}$ for which system Σ_C has a unique solution for every initial condition $x(\sigma) = x_0$. A map $\phi_C : \{(\tau, \sigma, x, \omega) : \tau, \sigma \in \mathcal{T} \subset \mathbb{R}, \sigma \leq \tau, x \in \mathbb{R}^n, \omega \in \mathcal{U}^{[\sigma, \tau]}\} \rightarrow \mathbb{R}^n$ is called the *transition map* of the Σ_C if the following properties hold:

- (*nontriviality*) for each $x \in \mathbb{R}^n$ there is at least one pair $\sigma < \tau$ in $\mathcal{T}$ and some $\omega \in \mathcal{U}^{[\sigma, \tau]}$ such that ω is admissible for x .
- (*restriction*) if $\omega \in \mathcal{U}^{[\sigma, \tau]}$ is admissible for x , then for each $\tau \in [\sigma, \mu)$ the restriction $\omega_1 := \omega|_{[\sigma, \tau]}$ is also admissible for x and the restriction $\omega_2 := \omega|_{[\tau, \mu]}$ admissible for $\phi(\tau, \sigma, x, \omega_1)$.
- (*semigroup*) if $\sigma, \tau, \mu \in \mathcal{T}$ are such that $\sigma < \tau < \mu$, if $\omega_1 := \omega|_{[\sigma, \tau]}$ and $\omega_2 := \omega|_{[\tau, \mu]}$, and if x is a state so that $\phi_C(\tau, \sigma, x, \omega_1) = x_1$ and $\phi_C(\mu, \tau, x_1, \omega_2) = x_2$.

then $\omega = \omega_1 \omega_2$ is also admissible for x and $\phi_C(\mu, \sigma, x, \omega) = x_2$.

We read $\phi_C(\tau, \sigma, x, \omega)$ as "the state at time τ resulting from starting at time σ and applying the control ω ".

By *inner trajectory* of this system we will mean a pair of functions $(\xi(\cdot), \omega(\cdot))$ where $\xi(\cdot) \in \mathbb{R}^n$, $\omega(\cdot) \in \mathbb{R}^m$ and $\xi(\tau) := \phi_C(\tau, \sigma, \xi(\sigma), \omega|_{[\sigma, \tau]})$ for each $\sigma, \tau \in \mathcal{T}$, $\sigma < \tau$. An *external trajectory* of system Σ_C is any pair of functions $(\zeta(\cdot), \omega(\cdot))$, $\zeta \in \mathbb{R}^p$, $\omega \in \mathbb{R}^m$, such that $\zeta(\tau) := h(\phi_C(\tau, \sigma, \xi(\sigma), \omega|_{[\sigma, \tau]}))$. For fixed σ and τ we will denote: $\lambda_x^{\sigma, \tau}(\omega) = h(\phi_C(\tau, \sigma, x, \omega|_{[\sigma, \tau]}))$. It represents the value of the last output of the system Σ_C initialized at the point x at time τ when applying the control $\omega \in \mathcal{U}^{[\sigma, \tau]}$.

Let us consider a nonlinear discrete-time (i.e. $t \in \mathbb{Z}$) system

$$\Sigma_D : \begin{cases} x(t+1) = g(x(t), u(t)) \\ y(t) = h_1(x(t)) \end{cases} \quad (2)$$

defined on $\mathbb{R}^n$ with controls $u \in \mathbb{R}^m$ and outputs $y \in \mathbb{R}^p$. Transition map of Σ_D is the map $\phi_D : \{(\tau, \sigma, x, \omega) : \tau, \sigma \in \mathcal{T} \cap \mathbb{Z}, \mathcal{T} \in \mathbb{R}, \sigma \leq \tau, x \in \mathbb{R}^n, \omega \in \mathcal{U}^{[\sigma, \tau]}\} \rightarrow \mathbb{R}^n$ that satisfies conditions nontriviality, restriction and semigroup defined in the same way as in the continuous-time case.

By the *inner trajectory* of system Σ_D we will mean a pair of functions $(\xi(\cdot), \omega(\cdot))$, $\xi(\cdot) \in \mathbb{R}^n$, $\omega(\cdot) \in \mathbb{R}^m$ and $\xi(\tau) = \phi_D(\tau, \sigma, \xi(\sigma), \omega|_{[\sigma, \tau] \cap \mathbb{Z}})$ for each $\sigma, \tau \in \mathcal{T} \cap \mathbb{Z}$, $\mathcal{T} \in \mathbb{R}$, $\sigma < \tau$. An *external trajectory* of the system Σ_D is a pair of functions $(\zeta(\cdot), \omega(\cdot))$, $\zeta \in \mathbb{R}^p$, $\omega \in \mathbb{R}^m$, such that $\zeta(\tau) := h(\xi(\tau)) = \phi_D(\tau, \sigma, \xi(\sigma), \omega|_{[\sigma, \tau] \cap \mathbb{Z}})$. By $\mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$ we will denote the set of all controls defined on $[\sigma, \tau] \cap \mathbb{Z}$ for which the system Σ_D has a unique solution for every initial condition $x(\sigma) = x_0$. The set $\mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$ can be identified with the set of all sequences $\omega(\sigma), \dots, \omega(\tau-1)$ consisting of elements of $\mathcal{U}$.

Let us consider an input-output continuous-time control system defined on $\mathbb{R}^n$ with $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$:

$$\Sigma_c : \begin{cases} \dot{x}(t) = f(x(t), u(t)) \\ y(t) = h(x(t)) \end{cases} \quad (3)$$

We will assume that:

- (A) the set $\mathcal{U}$ consists of all controls that are piecewise constant functions
- (B) system Σ_c is complete, i.e every control is admissible for every state.

Using Euler's method, on the base of the form of system Σ_c , let us construct an input-output discrete-time system also defined in $\mathbb{R}^n$ with $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$:

$$\Sigma_d: \begin{aligned} x(t+1) &= x(t) + \alpha f(x(t), u(t)) \\ y(t) &= h(x(t)) \end{aligned} \quad (4)$$

where α is any positive real number. System Σ_d is complete.

3 Indistinguishability

Let us consider a continuous-time system Σ_c and let $x, z \in \mathbb{R}^n$, $\sigma, \tau \in \mathcal{T} \subset \mathbb{R}$, $\sigma \leq \tau$. We say that [5]:

- the control $\omega \in \mathcal{U}^{[\sigma, \tau]}$ distinguishes between the events (x, σ) and (z, σ) if it is admissible for both and $\lambda_x^{\sigma, \tau}(\omega) \neq \lambda_z^{\sigma, \tau}(\omega)$.
- the events (x, σ) and (z, σ) are distinguishable on the interval $[\sigma, \tau]$ or in time at most $T = \tau - \sigma$ if there is some $t \in [\sigma, \tau]$ and some control $\omega \in \mathcal{U}^{[\sigma, t]}$ that distinguishes them. If there exists at least one such interval, then (x, σ) and (z, σ) are distinguishable.
- the states x, z are distinguishable by a given control (respectively distinguishable on interval $[\sigma, \tau]$ or distinguishable) if there is same $\sigma \in \mathcal{T}$ so that the respective property holds for events (x, σ) and (z, σ) .

Let us consider a discrete-time system Σ_D and let $x, z \in \mathbb{R}^n$, $\sigma, \tau \in \mathcal{T} \subset \mathbb{Z}$, $\sigma \leq \tau$. We say that [5, 6]:

- the control $\omega \in \mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$ distinguishes between the events (x, σ) and (z, σ) if it is admissible for both and $\lambda_x^{\sigma, \tau}(\omega) \neq \lambda_z^{\sigma, \tau}(\omega)$.
- the events (x, σ) and (z, σ) are distinguishable on the set $[\sigma, \tau] \cap \mathbb{Z}$ or in time at most $T = [\tau - \sigma]^\dagger$ if there is some $t \in [\sigma, \tau] \cap \mathbb{Z}$ and some control $\omega \in \mathcal{U}^{[\sigma, t] \cap \mathbb{Z}}$ that distinguishes them. If there exists at least one such interval, then (x, σ) and (z, σ) are distinguishable.
- the states x, z are distinguishable by a given control (respectively distinguishable on set $[\sigma, \tau] \cap \mathbb{Z}$ or distinguishable) if there is same $\sigma \in \mathcal{T}$ so that the respective property holds for events (x, σ) and (z, σ) .

[†] $[c]$ means the integral part of real number c

Let us consider system Σ_c given by (3) and system Σ_d given by (4). Let us assume that:

(C) the function f is of class C^1 with respect to x and u .

Lemma 3.1. Let $\sigma, \tau \in \mathcal{T} \subset \mathbb{Z}$, $\sigma \leq \tau$. For any $x, z \in \mathbb{R}^n$ there exists a small fixed $\alpha \in \mathbb{R}_+$ such that if events (x, σ) and (z, σ) are distinguishable by control $\omega \in \mathcal{U}^{[\sigma, \tau]}$ for the system Σ_d then they are distinguishable by control $\bar{\omega} \in \mathcal{U}^{[\sigma, \sigma + \alpha(\tau - \sigma)]}$ for the system Σ_c .

Proof: The events (x, σ) and (z, σ) are distinguishable by control $\omega \in \mathcal{U}^{[\sigma, \tau]}$ for the system Σ_d , so $\lambda_x^{\sigma, \tau}(\omega) \neq \lambda_z^{\sigma, \tau}(\omega)$ or $h(\phi_d(\tau, \sigma, x, \omega)) \neq h(\phi_d(\tau, \sigma, z, \omega))$. This condition is fulfilled only if $\phi_d(\tau, \sigma, x, \omega) \neq \phi_d(\tau, \sigma, z, \omega)$. It means that inner trajectories of system Σ_d starting at points x and z are different. One can notice that each of these trajectories is the set of points that are the ends of Euler's broken line. Let us joint these points by rectilinear segments. Because Euler's broken lines approximate the inner trajectory of continuous-time system Σ_c hence $\phi_c(\bar{\tau}, \sigma, x, \bar{\omega}) \neq \phi_c(\bar{\tau}, \sigma, z, \bar{\omega})$ where

$$\bar{\tau} := \sigma + \alpha(\tau - \sigma) \quad (5)$$

and for fixed α control $\bar{\omega}$ is defined as follows: let l be any positive integer such that $\sigma =: t_1, \dots, t_l := \tau - 1$. Let us put $\omega_1 = \omega(\sigma), \dots, \omega_{l-1} = \omega(\tau - 1)$ for $\omega \in \mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$. Then

$$\bar{\omega}(t) := \omega_i \quad (6)$$

for $t \in [t_i, t_{i+1})$ for $i = 1, \dots, l - 1$. Then $h(\phi_c(\bar{\tau}, \sigma, x, \bar{\omega})) \neq h(\phi_c(\bar{\tau}, \sigma, z, \bar{\omega}))$. Hence, the events (x, σ) and (z, σ) are distinguishable by control $\bar{\omega} \in \mathcal{U}^{[\sigma, \bar{\tau}]}$ for system Σ_c . $\square$

Proposition 3.2. Let $\sigma, \tau \in \mathbb{Z}$, $\sigma < \tau$.

- a) For any $x, z \in \mathbb{R}^n$ there exists a small positive real α such that if events (x, σ) and (z, σ) are distinguishable on the set $[\sigma, \tau] \cap \mathbb{Z}$ for the system Σ_d then they are distinguishable on the interval $[\sigma, \sigma + \alpha(\tau - \sigma)]$ for system Σ_c .
- b) If events (x, σ) and (z, σ) are distinguishable in time at most $T = [\tau - \sigma]$ for the system Σ_d then they are distinguishable in time at most $\bar{T} = \alpha(\tau - \sigma)$ for system Σ_c .
- c) If events (x, σ) and (z, σ) are distinguishable for system Σ_d that they are also distinguishable for the system Σ_c .

Proof: a) Distinguishability of events (x, σ) and (z, σ) on the set $[\sigma, \tau] \cap \mathbb{Z}$ for system Σ_d means that there is some $t \in [\sigma, \tau] \cap \mathbb{Z}$ and control $\omega \in \mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$ that distinguishes these events. Let us take $\bar{t}$ such that $\bar{t} \leq \bar{\tau}$ and $\bar{\tau}$ is given by (5). Then $\bar{\omega} \in \mathcal{U}^{[\sigma, \bar{\tau}]}$ can be taken as in (6), but now $\sigma =: t_1, \dots, t_l := \bar{t} - 1$. Hence there exist some $\bar{t} \in [\sigma, \sigma + \alpha(\tau - \sigma)]$ and $\omega \in \mathcal{U}^{[\sigma, \tau]}$ that events (x, σ) , (z, σ) are distinguishable on the interval $[\sigma, \sigma + \alpha(\tau - \sigma)]$ for system Σ_c .

b) Distinguishability of events (x, σ) and (z, σ) in time at most $T = [\tau - \sigma]$ for system Σ_d means distinguishability of these events on the interval $[\sigma, \sigma + \alpha(\tau - \sigma)]$ for the system Σ_c . Hence (x, σ) and (z, σ) are distinguishable in time at most $\bar{T} = \alpha(\tau - \sigma)$ for system Σ_c .

c) This fact follows from the proof of previous facts and definitions. $\square$

Proposition 3.3. *Let $\sigma, \tau \in \mathbb{Z}$, $\sigma < \tau$. If the states x and z are distinguishable by a control $\omega \in \mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$ or just distinguishable for the system Σ_d then these states are distinguishable by a control $\bar{\omega} \in \mathcal{U}^{[\sigma, \tau]}$ given by (6) or just distinguishable for the system Σ_c .*

Proof follows from Lemma 3.1, Proposition 3.2 and from the definition of distinguishability of states x and z . $\square$

Lemma 3.4. *Let $\sigma, \tau \in \mathbb{R}$, $\sigma \leq \tau$. For any $x, z \in \mathbb{R}^n$ there exists a fixed $\alpha \in \mathbb{R}_+$ such that if events (x, σ) and (z, σ) are distinguishable by control $\omega \in \mathcal{U}^{[\sigma, \tau]}$ for the system Σ_c then they are distinguishable by control $\bar{\omega} = \omega|_{\mathbb{Z}} \in \mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$ for the system Σ_d .*

Proof: Distinguishability of the events (x, σ) and (z, σ) by control $\omega \in \mathcal{U}^{[\sigma, \tau]}$ for the system Σ_c implies that $\lambda_x^{\sigma, \tau}(\omega) \neq \lambda_z^{\sigma, \tau}(\omega)$ or equivalently

$$h(\phi_c(\tau, \sigma, x, \omega)) \neq h(\phi_c(\tau, \sigma, z, \omega)). \quad (7)$$

The last condition is fulfilled when $\phi_c(\tau, \sigma, x, \omega) \neq \phi_c(\tau, \sigma, z, \omega)$.

Let us consider two different inner trajectories of the system Σ_c : one starting at the point x and the second one - starting at z . Let us define positive integer α - a step of Euler's method - as

$$\alpha := \frac{\tau - \sigma}{[\tau - \sigma] + 1} \quad (8)$$

where $[c]$ denotes an integral part of a real number c . If (7) is fulfilled, then in neighbourhoods (of a radius at least α) of the last points of the inner trajectories initialized respectively at x and z (reached in the time τ) the Euler's broken lines are different. So inner trajectories of system Σ_d given by (4) initialized respectively at points x and z (for α), that consist of the set of points being the ends of Euler's broken lines, are different. Let $\bar{\omega} := \omega|_{\mathbb{Z}}$. Then $h(\phi_d(\tau, \sigma, x, \bar{\omega})) \neq h(\phi_d(\tau, \sigma, z, \bar{\omega}))$. So, events (x, σ) and (z, σ) are distinguishable by control $\bar{\omega} \in \mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$ for the system Σ_d . $\square$

Proposition 3.5. *Let $\sigma, \tau \in \mathbb{R}$, $\sigma < \tau$.*

a) *For any $x, z \in \mathbb{R}^n$ there exists a fixed $\alpha \in \mathbb{R}_+$ such that if events (x, σ) and (z, σ) are distinguishable on interval $[\sigma, \tau]$ for system Σ_c then they are distinguishable on the set $[\sigma, \tau] \cap \mathbb{Z}$ for system Σ_d .*

b) *If events (x, σ) and (z, σ) are distinguishable in time at most $T = \tau - \sigma$ for system Σ_c then they are distinguishable in time at most $\bar{T} = [\tau - \sigma]$ for system Σ_d .*

c) *If events (x, σ) and (z, σ) are distinguishable for system Σ_c that they are also distinguishable for system Σ_d .*

Proof follows from the Lemma 3.4. $\square$

Proposition 3.6. *Let $\sigma, \tau \in \mathbb{R}$, $\sigma < \tau$. If the states x and z are distinguishable by control $\omega \in \mathcal{U}^{[\sigma, \tau]}$ or just distinguishable for system Σ_c then there exists $\alpha \in \mathbb{R}_+$ such that these states are distinguishable by control $\bar{\omega} = \omega|_{\mathbb{Z}} \in \mathcal{U}^{[\sigma, \tau] \cap \mathbb{Z}}$, or just distinguishable for system Σ_d given by (4).*

Proof follows from Lemma 3.4 and Proposition 3.5. $\square$

4 Observability

The system Σ_C given by (1) is ([5]):

- *observable on the interval $[\sigma, \tau]$* if for every pair of distinct states x and z the events (x, σ) and (z, σ) are distinguishable on this interval.
- *observable in time T* if any two distinct states are distinguishable in time (at most) T .
- *observable* if any two distinct states are distinguishable.

Moreover, system Σ_C is:

- *locally observable on interval* $[\sigma, \tau]$ (respectively *locally observable in time at most* T) if there exists an open neighborhood $\mathcal{O}_{\tilde{x}}$ of a point $\tilde{x} \in \mathbb{R}^n$ such that Σ_C is observable on this interval (respectively observable in time at most T) for any pair of distinct states from $\mathcal{O}_{\tilde{x}}$.
- *locally observable* if there exists an open neighborhood $\mathcal{O}_{\tilde{x}}$ of a point $\tilde{x} \in \mathbb{R}^n$ such that any two distinct states from $\mathcal{O}_{\tilde{x}}$ are distinguishable.

The system Σ_D given by (2) is ([5, 6])

- *observable on the set* $[\sigma, \tau] \cap \mathbb{Z}$ if for every pair of distinct states x and z the events (x, σ) and (z, σ) are distinguishable on this set.
- *observable in time* T if any two distinct states are distinguishable in time (at most) T .
- *observable* if any two distinct states are distinguishable.
- *locally observable on the set* $[\sigma, \tau] \cap \mathbb{Z}$ (respectively *locally observable in time at most* T) if there exists an open neighborhood $\mathcal{O}_{\tilde{x}}$ of a point $\tilde{x} \in \mathbb{R}^n$ such that Σ_D is observable (respectively observable in time at most T) on this set for any pair of distinct states from $\mathcal{O}_{\tilde{x}}$.
- *locally observable* if there exists an open neighborhood $\mathcal{O}_{\tilde{x}}$ of a point $\tilde{x} \in \mathbb{R}^n$ such that any two distinct states from $\mathcal{O}_{\tilde{x}}$ are distinguishable.

Proposition 4.1. *Let $\tilde{x}$ a point be fixed and $\sigma, \tau \in \mathbb{Z}$, $\sigma < \tau$. There exists $\alpha \in \mathbb{R}_+$ such that if system Σ_d is locally observable on the set $[\sigma, \tau] \cap \mathbb{Z}$ then the system Σ_c is locally observable on the interval $[\sigma, \sigma + \alpha(\tau - \sigma)]$.*

Proof: Let us take a neighborhood $\mathcal{O}_{\tilde{x}}$ of the point $\tilde{x} \in \mathbb{R}^n$, that has a compact closure. Let $x, z \in \mathcal{O}_{\tilde{x}}$. Then, one can choose a common α for both x and z (from continuity of α on a compact set).

Local observability of system Σ_d on the set $[\sigma, \tau] \cap \mathbb{Z}$ implies distinguishability of events (x, σ) and (z, σ) on the set $[\sigma, \tau] \cap \mathbb{Z}$ for every pair of distinct states x, z from neighborhood $\mathcal{O}_{\tilde{x}}$ of point $\tilde{x}$. Then the Proposition 3.2 implies that these events are distinguishable on interval $[\sigma, \sigma + \alpha(\tau - \sigma)]$ for the system Σ_c . So, the system Σ_c is locally observable on $[\sigma, \sigma + \alpha(\tau - \sigma)]$. $\square$

Proposition 4.2. *Let $\sigma, \tau \in \mathbb{R}$, $\sigma < \tau$. If the system Σ_c is locally observable on the interval $[\sigma, \tau]$ then there exists $\alpha \in \mathbb{R}_+$ such that the system Σ_d is also locally observable on this interval.*

Proof: If system Σ_c is locally observable on the interval $[\sigma, \tau]$ then for every pair of distinct states x and z , $x, z \in \mathcal{O}_{\tilde{x}}$, $\tilde{x} \in \mathbb{R}^n$, the events (x, σ) and (z, σ) are distinguishable on this interval. Then Proposition 3.5 implies that there exists $\alpha \in \mathbb{R}_+$ such that these events are also distinguishable on this interval for the system Σ_d . Hence Σ_d is locally observable on $[\sigma, \tau]$. $\square$

Proposition 4.3. a) *If system Σ_c is locally observable in time at most T then there exists $\alpha \in \mathbb{R}_+$ such that system Σ_d is locally observable in time at most $\bar{T} = [T]$.*

b) *If there exists $\alpha \in \mathbb{R}_+$ such that system Σ_d is locally observable in time at most T then system Σ_c is locally observable in time $\bar{T} = \alpha[T]$.*

Proof follows from Propositions 3.2 and 3.5. $\square$

Proposition 4.4. *System Σ_c is locally observable if and only if there exists a positive real α such that system Σ_d is locally observable.*

Proof: " $\Rightarrow$ " Let $\sigma, \tau \in \mathbb{R}$, $\sigma < \tau$. Local observability of system Σ_c implies distinguishability of any two distinct states in $\mathcal{O}_{\tilde{x}}$, $\tilde{x} \in \mathbb{R}^n$. Then, by Proposition 3.6, there exists a real positive number α such that these two distinct states are also distinguishable for the system Σ_d . Hence Σ_d is observable.

" $\Leftarrow$ " Let $\sigma, \tau \in \mathbb{Z}$, $\sigma < \tau$. Observability of system Σ_d means that any two distinct states from $\mathcal{O}_{\tilde{x}}$, $\tilde{x}$ - any point from $\mathbb{R}^n$, are distinguishable. Proposition 3.3 implies that these states are also distinguishable for the system Σ_c . Hence Σ_c is observable. $\square$

Above there were considered the relations between local observability of continuous-time system Σ_c and local observability of it's Euler's discretization, i.e. local observability of system Σ_d given by (4). But besides classical observability of control systems there can be studied also (classical) single-experiment observability or final-state determinability [5, 6, 4]. System Σ_c (or Σ_D) is:

- *single-experiment observable* if there exists control ω that distinguishes every pair of states
- *final-state determinable* if there exists a control ω such that for each pair of states x, z either

$$\lambda_x^{\sigma, \tau}(\omega) \neq \lambda_z^{\sigma, \tau}(\omega) \quad \text{or} \quad \phi_C(\tau, \sigma, x, \omega) = \phi_C(\tau, \sigma, z, \omega) \quad (9)$$

respectively for system Σ_D :

$$\lambda_x^{\sigma,\tau}(\omega) \neq \lambda_z^{\sigma,\tau}(\omega) \quad \text{or} \quad \phi_D(\tau, \sigma, x, \omega) = \phi_D(\tau, \sigma, z, \omega). \quad (10)$$

Moreover, we say that system Σ_C or Σ_D is:

- *locally single-experiment observable* if it is single-experiment observable for every pair of states from $\mathcal{O}_{\tilde{x}}$ where $\tilde{x}$ is any point from $\mathbb{R}^n$
- *locally final-state determinable* if there exists a control ω such that for each pair of states $x, z \in \mathcal{O}_{\tilde{x}}$ holds (9) or (10).

In continuous-time analytic case single-experiment observability implies observability [5]. In discrete-time analytic case observability implies final-state observability. From these facts and proposition 4.4 follows that local observability of analytic system Σ_c given by (3) implies local final-state determinability of it's Euler's discretization Σ_d given by (4).

Proposition 4.5. *Analytic system Σ_c is locally final-state determinable if and only if there exists $\alpha \in \mathbb{R}_+$ such that system Σ_d (with analytic function f) is locally final-state determinable.*

Proof: Let us assume that $\omega \in \mathcal{U}^{[\sigma,\tau]}$ and $x, z \in \mathcal{O}_{\tilde{x}}$ for some $\tilde{x} \in \mathbb{R}^n$.
 $\Rightarrow$ Lemma 3.4 implies that from the first condition of definition of finite-state determinability i.e. $\lambda_x^{\sigma,\tau}(\omega) \neq \lambda_z^{\sigma,\tau}(\omega)$ for system Σ_c follows that $\lambda_x^{\sigma,\tau}(\omega) \neq \lambda_z^{\sigma,\tau}(\omega)$ for system Σ_d .

Let us assume that $\phi_c(\tau, \sigma, x, \omega) = \phi_c(\tau, \sigma, z, \omega)$. It means that inner trajectories of system Σ_c starting at points x and z , respectively, coincide. From proof of lemma 3.4 follows that if the second condition of (9) holds then in some neighbourhood of radius at least α of the last points of trajectories starting respectively at x and z , the Euler's broken lines coincide. Then inner trajectories of system Σ_d starting at points x and z also coincide. Hence $\phi_d(\tau, \sigma, x, \omega) = \phi_d(\tau, \sigma, z, \omega)$.

$\Leftarrow$ Similar as in previous point, lemma 3.1 implies that inequality $\lambda_x^{\sigma,\tau}(\omega) \neq \lambda_z^{\sigma,\tau}(\omega)$ for system Σ_d follows that $\lambda_x^{\sigma,\tau}(\omega) \neq \lambda_z^{\sigma,\tau}(\omega)$ for system Σ_c .

Now, let us assume that $\phi_d(\tau, \sigma, x, \omega) = \phi_d(\tau, \sigma, z, \omega)$. Then inner trajectories of system Σ_d starting at points x and z coincide. So, Euler's broken lines that approximate inner trajectories of system Σ_c also coincide. Hence $\phi_c(\bar{\tau}, \sigma, x, \bar{\omega}) = \phi_c(\bar{\tau}, \sigma, z, \bar{\omega})$ where $\bar{\tau}$ and $\bar{\omega}$ are, respectively, given by (5) and

(6). $\square$.

Because in continuous-time case single-experiment observability coincides with final-state determinability [5], then local single-experiment observability of system Σ_c is equivalent to local final-state determinability of system Σ_d for some small $\alpha \in \mathbb{R}_+$. Moreover, local observability of Σ_c in analytic case implies local final-state determinability of system Σ_d for some small $\alpha \in \mathbb{R}_+$.

References

- [1] Z.Bartosiewicz, Local observability of control systems, Control & systems Letters, 1995
- [2] R.Hermann, A.Krener, Nonlinear controllability and observability, IEEE Trans.AC-22, 1977
- [3] Isidori A.: Nonlinear Control Systems, Springer, 1996
- [4] U.Kotta, On decomposition of discrete-time nonlinear control systems, to appear
- [5] E.D.Sontag, Mathematical Control Theory, Springer-Verlag 1990
- [6] E.D.Sontag, On the observability of polynomial systems. I: Finite-time problems, SIAM J.Control and Optimization, No1, 1979
- [7] E.D.Sontag, Y.Wang, Orders of IÓ Equations and Uniformly Universal Inputs, Proc. of the 33th Conf. on Decision and Control, Dec.1994
- [8] H.J.Sussmann, Single-Input Observability of Continuous-Time Systems, Math. Systems Theory 12, 1979

Instytut Matematyki i Fizyki, Politechnika Białostocka, ul.Wiejska 45a, 15-351 Białystok

UWAGI O OBSERWOWALNOŚCI DYSKRETYZACJI EULERA UKŁADÓW STEROWANIA

Streszczenie

W pracy są przedstawione relacje pomiędzy różnymi rodzajami obserwowalności (np. obserwowalność, obserwowalność w sensie odróżnialności w stanie końcowym, obserwowalność typu „single-experiment observability”) nieliniowych układów sterowania z czasem ciągłym Σ_c i układów otrzymanych poprzez dyskretyzację Eulera Σ_e .

K. Valeev

K. Janglajew

ON THE DICHOTOMY OF A SYSTEM OF HIGHER ORDER LINEAR DIFFERENTIAL EQUATIONS

Abstract: In this paper we develop a method for the decomposition of solutions of a system of higher order linear differential equations under an exponential dichotomy. This method is based on properties of the parametric transfer function and on properties of the two-sided Laplace transform.

Key-words: Dichotomy, the Greer matrix, the transfer function, the Laplace transform.

1 Introduction.

The notion of dichotomy for a linear system of differential equations has gained prominence since the appearance of two fundamental books: [4]. and [2]. These were followed by the important book of Coppel [1]. Several papers on the subject appeared afterward, see [3]. [5].

We consider the following system of m linear differential equations with variable coefficients:

$$\sum_{k=0}^n A_k(t) \frac{d^k Y(t)}{dt^k} = \Phi(t), \dim Y(t) = m, \quad (1)$$

where $t \in (-\infty, \infty)$, $\det A_n(t) \neq 0$.

For the convenience of the reader we repeat some known material without proofs, thus making our exposition self-contained.

If for any continuous vector-function $\Phi(t)$ such that

$$\|\Phi(t)\| \leq M, (-\infty < t < +\infty)$$

there exists a unique solution of System (1) of the form

$$Y(t) = \int_{-\infty}^{\infty} G(t, \tau) \Phi(\tau) d\tau, \quad (2)$$

satisfying

$$\|Y(t)\| \leq M_1 = \text{const}, (-\infty < t < \infty),$$

then the matrix $G(t, \tau)$ is called the Green matrix.

In the particular case when

$$\Phi(t) = E\delta(t - \tau),$$

where $\delta(t)$ is the Dirac function, from (2) we get

$$Y(t) = \int_{-\infty}^{\infty} G(t, s) \delta(s - \tau) ds = G(t, \tau),$$

Therefore the Green matrix is a discontinuous solution of the matrix differential equation

$$\sum_{k=0}^n A_k(t) \frac{\partial^k G(t, \tau)}{\partial t^k} = E\delta(t - \tau) \quad (3)$$

and this solution satisfies the following conditions (see [7]):

$$\frac{\partial^k G(\tau + 0, \tau)}{\partial t^k} = \frac{\partial^k G(\tau - 0, \tau)}{\partial t^k}, (k = 0, 1, \dots, n-2);$$

$$\frac{\partial^{n-1} G(\tau + 0, \tau)}{\partial t^{n-1}} - \frac{\partial^{n-1} G(\tau - 0, \tau)}{\partial t^{n-1}} = A_n^{-1}(\tau);$$

$$\int_{-\infty}^{\infty} \|G(t, \tau)\| d\tau < \infty \quad (4)$$

It should be noted that for a system of linear differential equations with constant coefficients

$$\sum_{k=0}^n A_k \frac{d^k Y(t)}{dt^k} = \Phi(t), \det A_n \neq 0$$

the Green matrix is determined by the formula

$$G(t, \tau) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} L^{-1}(p) e^{p(t-\tau)} dp$$

where $L(p) = \sum_{k=0}^n A_k p^k$.

In a series papers I.Z.Shtokolo and L.A.Zadeh [6] [8] developed the method of the transfer function, which may be used to find the Green matrix. The procedure is to find a solution of the matrix differential equation

$$\sum_{k=0}^n A_k(t) \frac{\partial^k}{\partial t^k} (e^{pt} W(t, p)) = E e^{pt}, \quad \text{Re } p \leq \theta \quad (5)$$

The matrix $W(t, p)$ is called a transfer matrix.

From (2) we obtain

$$W(t, p) = \int_{-\infty}^{\infty} G(t, \tau) e^{-p(t-\tau)} d\tau = \int_{-\infty}^{\infty} G(t, t - \tau) e^{-p\tau} d\tau \quad (6)$$

Thus the transfer matrix is expressed in terms of the Green matrix. The inverse transformation leads to

$$G(t, t - \tau) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} W(t, p) e^{p\tau} dp; \quad (7)$$

$$G(t, \tau) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} W(t, p) e^{p(t-\tau)} dp.$$

Thus it is possible to use the transfer matrix to search for the Green matrix. The transfer matrix satisfies an equation of the form (5), i.e.,

$$\sum_{k=0}^n A_k(t) (p + D)^k W(t, p) = E, \quad D := \frac{\partial}{\partial t} \quad (8)$$

or the equivalent equation

$$\sum_{k=0}^n \frac{1}{k!} \frac{\partial^k L(t, p)}{\partial p^k} \frac{\partial^k W(t, p)}{\partial t^k} = E$$

$$L(t, p) := \sum_{k=0}^n A_k(t) p^k, \quad (9)$$

which was given by L.A.Zadeh.

2 Finding the Green matrix.

To seek approximations to the Green matrix and the transfer matrix, we may use expansions in powers of a small parameter.

Let us consider the perturbed system of differential equations

$$(L(D) + \varepsilon B(t, D)) Y(t, \varepsilon) = 0, \quad (10)$$

where the differential operators have the form

$$L(D) = \sum_{k=0}^n A_k D^k, \quad \det A_n \neq 0, \quad B(t, D) = \sum_{k=0}^{n-1} B_k(t) D^k \quad (11)$$

From now on, D denotes $\frac{\partial}{\partial t}$.

We assume that the matrix $B_k(t)$ be k times differentiable with respect to t and satisfy for $f \in (-\infty, \infty)$ the boundedness conditions

$$\left\| \frac{d^l B_k(t)}{dt^l} \right\| \leq b_{lk} = \text{const}, \quad (l = 0, 1, \dots, k; \quad k = 0, 1, \dots, n-1)$$

The Green matrix of System (10) will be denoted by $G(t, \tau, \varepsilon)$. Assuming that the matrix $L^{-1}(p)$ does not have poles on the line $\text{Re } p=0$, we obtain for $\varepsilon = 0$

$$G(t, \tau, 0) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} L^{-1}(p) e^{p(t-\tau)} dp.$$

From (2) and (3) we get the matrix integral equation

$$G(t, \tau, \varepsilon) = G(t, \tau, 0) - \varepsilon \int_{-\infty}^{\infty} G(t, s, 0) B(s, D) G(s, \tau, \varepsilon) ds.$$

Integrating the last equation by parts eliminates the derivative of an unknown-matrix and leads to the equation

$$G(t, \tau, \varepsilon) = G(t, \tau, 0) + \varepsilon \int_{-\infty}^{\infty} R(t, s) G(s, \tau, \varepsilon) ds. \quad (12)$$

where the matrix kernel $R(t, s)$ is determined by the formula

$$R(t, s) = \sum_{k=0}^{n-1} (-1)^{k+1} \frac{\partial^k}{\partial t^k} (G(t, s, 0) B_k(s))$$

The appropriate integral equation for the parametric transfer matrix

$$W(t, p, \varepsilon) = \int_{-\infty}^{\infty} G(t, \tau, \varepsilon) e^{-p(t-\tau)} d\tau$$

has the form

$$W(t, p, \varepsilon) = L^{-1}(p) + \varepsilon \int_{-\infty}^{\infty} R(t, \tau) e^{-p(t-\tau)} W(\tau, p, \varepsilon) d\tau \quad (13)$$

A solution of the systems of matrix integral equations (12), (13) may be found by using a method of successive approximations which converges under conditions

$$|\varepsilon| \int_{-\infty}^{\infty} \|R(t, \tau)\| e^{-\operatorname{Re} p(t-\tau)} d\tau < 1, \quad |\operatorname{Re} p| < \theta \quad (14)$$

Under this condition, the exponential dichotomy of solutions is preserved and consequently, also possibility of a decomposition of solutions.

3 Dichotomy

Suppose the Green matrix has been found. We show that it is possible to decompose a solution set of a homogeneous system of the type

$$\sum_{k=0}^n A_k(t) \frac{d^k Y(t)}{dt^k} = 0 \quad (15)$$

by using the Green matrix.

A solution $Y(t, \tau)$ of System (15) with initial conditions, at $t = \tau$

$$\left. \frac{d^k Y(t)}{dt^k} \right|_{t=\tau} = Y_k(\tau), \quad (k = 0, 1, \dots, n-1) \quad (16)$$

may be represented as a solution of the inhomogeneous system with the right side depending on Dirac's generalized function and its derivatives

$$\sum_{k=0}^n A_k(t) \frac{d^k Y(t)}{dt^k} = \sum_{k=0}^{n-1} \left(\sum_{l=0}^{n-k-1} A_{l+k+1}(t) Y_l(\tau) \right) \frac{d^k \delta(t-\tau)}{dt^k}$$

A solution of System (15) with conditions (16) may be written by using (2)

$$Y(t, \tau) = \int_{-\infty}^t G(t, s) Q(s, \tau) ds + \int_t^{\infty} G(t, s) Q(s, \tau) ds \quad (17)$$

where

$$Q(s, \tau) := \sum_{k=0}^{n-1} \left(\sum_{l=0}^{n-k-1} A_{l+k+1}(s) Y_l(\tau) \right) \frac{d^k \delta(s-\tau)}{ds^k}$$

For $t > \tau$ we have a particular solution of System (15)

$$Y(t, \tau) = S(t, \tau),$$

$$S(t, \tau) := \sum_{k=0}^{n-1} (-1)^{k-1} \left(\frac{\partial^k}{\partial \tau^k} \sum_{l=0}^{n-k-1} G(t, \tau) A_{l+k+1}(\tau) \right) Y_l(\tau)$$

For $t = \tau + 0$, we get

$$Y(\tau + 0, \tau) = S(\tau + 0, \tau)$$

4 Conclusions

If $S(\tau + 0, \tau) \equiv 0$, then the solution set of System (15) does not contain solutions tending to zero exponentially for $t \rightarrow +\infty$.

Similarly, if $S(\tau - 0, \tau) \equiv 0$, then the solution set does not contain solutions tending to zero exponentially for $t \rightarrow -\infty$. Therefore the equations

$$S(\tau + 0, \tau) \equiv 0, \quad S(\tau - 0, \tau) \equiv 0 \quad (18)$$

bring about the exponential dichotomy of solutions of System (15).

If the system of m equations (15) of order n is written in the form of a system of (nm) first order equations, then conditions (18) may be obtained

by using the standard Green matrix [2] of order $(nm \times nm)$ and by employing projectors. This leads to rather lengthy computations.

Example. Let us consider the third order linear differential equation

$$\frac{d^3 y}{dt^3} + 2 \frac{d^2 y}{dt^2} - \frac{dy}{dt} - 2y = 0 \quad (19)$$

The transfer function may be written in the form

$$L(p) = \frac{1}{p^3 + 2p^2 - p - 2} = \frac{1}{6(p-1)} - \frac{1}{2(p+1)} + \frac{1}{3(p+2)}$$

The Green function has the form

$$G(t, \tau) = -\frac{1}{2}e^{-(t-\tau)} + \frac{1}{3}e^{-2(t-\tau)} \quad (t > \tau)$$

$$G(t, \tau) = -\frac{1}{6}e^{t-\tau}, \quad (t < \tau)$$

A solution of Equation (19) with the initial conditions

$$y(0) = y_0, \quad \frac{dy(0)}{dt} = y_1, \quad \frac{d^2 y(0)}{dt^2} = y_2$$

may be obtained as a solution of the inhomogeneous equation, with a generalized right-hand side

$$y''' + 2y'' - y' - 2y = \delta''(t)y_0 + \delta'(t)(y_1 + 2y_0) + \delta(t)(y_2 + 2y_1 - y_0)$$

Conditions (18) for $\tau = 0$ take the form

$$y_2 - y_1 - 2y_0 = 0, \quad y_2 - y_0 = 0 \quad (20)$$

$$y_2 + 3y_1 + 2y_0 = 0 \quad (21)$$

If we denote by

$$y(t) = c_1 e^{-t} + c_2 e^{-2t} + c_3 e^t \quad (22)$$

the general solution of Equation(19) then we have

$$y_0 = c_1 + c_2 + c_3, \quad y_1 = -c_1 - 2c_2 + c_3, \quad y_2 = c_1 + 4c_2 + c_3,$$

Conditions (20) lead to $c_1 = 0, c_2 = 0$, that is in the general solution terms tending to zero as $t \rightarrow +\infty$, are absent.

The condition (21) leads to $c_3 = 0$ i.e. in the general solution the term tending to zero as $t \rightarrow -\infty$, is absent.

References

- [1] W. A. Coppel, Dichotomies in Stability Theory, Lecture Notes in Mathematics, No. 629, Springer-Verlag, New York-Berlin-Heidelberg, 1978.
- [2] Ju. L. Daleckii and M. G. Krein, Stability of Solutions of Differential Equations in Banach Spaces, American Mathematical Society Translations of Mathematical Monographs, Vol. 43, Providence, 1974.
- [3] S. Elaydi and K. Janglajew, Dichotomy and trichotomy of difference equations, J. Difference Equations and Appl, Vol. 3, No.5- 6 (1998), 417-448.
- [4] J. L. Massera and J. J. Schäffer, Linear Differential Equations and Function Spaces, Academic Press, New York, 1966.
- [5] K. J. Palmer, Exponential Dichotomies, the Shadowing Lemma and Transversal Homoclinic points, Dynamics reported, vol.1, John Wiley & Sons and G. Teubner (1988).
- [6] I. Shtokalo, Operational Methods and their Development in Theory of Linear Differential Equations with Variable Coefficients, AN USSR, Kiev, 1961. (in Russian).
- [7] K. G. Valeev, Linear Differential Equations (Decomposition of Solutions), Znaniye, Kiev, 1978.(in Russian).

- [8] L.Zadeh, C.A. Desoer, Linear System Theory - the State Space Approach. New York, McGraw-Hill, 1963.

Kiev National University of Economics, Ukraine.
Institute of Mathematics, University of Białystok, Akademicka 2, 15-267 Białystok, Poland, e-mail: jang@math.uwb.edu.pl

O DYCHOTOMII UKŁADU LINIOWYCH RÓWNAŃ RÓŻNICZKOWYCH WYŻSZEGO RZĘDU

Streszczenie

W artykule tym rozwijamy metodę dekompozycji rozwiązań układu liniowych równań różniczkowych wyższego rzędu ze względu na dychotomię wykładniczą. Metoda ta bazuje na własnościach parametrycznej transmitancji i dwustronnej transformaty Laplace'a.

Małgorzata Wyrwas*

MULTIOBSERVERS FOR NONLINEAR SYSTEMS

Abstract: In this paper we give the construction of some family of multiobservers for locally observable systems and prove that the dynamics of the multiobserver system gives an estimation of the successive derivatives of the output based upon the knowledge of the output $y = h(x)$. The multiobserver's output is a multifunction which gives simultaneously finitely many estimations of the state of the original system.

Let us consider a nonlinear analytic system Σ on $\Omega \subset \mathbb{R}^n$:

$$\dot{x} = f(x) \quad (1)$$

$$y = h(x) \quad (2)$$

For simplicity we will assume that the output y is scalar. The vector field f and the function h are assumed to be real analytic.

Let us recall that two points x_1 and x_2 in $\mathbb{R}^n$ are *indistinguishable* if

$$h(x(t, x_1)) = h(x(t, x_2))$$

for every $t \geq 0$ such that both sides of the equation are defined. Here $x(t, x_0)$ denotes the trajectory of f starting at $x(0) = x_0$, evaluated at time t . Otherwise x_1 and x_2 are called *distinguishable*. It is known that x_1 and x_2 are indistinguishable iff $\varphi(x_1) = \varphi(x_2)$ for every φ of the form $L_f^k h$, where L_f is the Lie derivative with respect to the vector field f . These functions generate the *observation algebra* $H(\Sigma)$ of the system Σ . Indistinguishability is an equivalence relation so the state space can be decomposed into disjoint indistinguishability classes.

*Supported by KBN under the Technical University of Białystok grant No. W/IMF/1/01

We say that Σ is *observable* if any two distinct points are distinguishable.

The system Σ is *locally observable* at x_0 if there is a neighborhood U of x_0 such that for every $x \in U$, x and x_0 are distinguishable. Σ is *locally observable* if it is locally observable at every point. Necessary and sufficient conditions for local observability were developed in [1, 2].

Assume Σ is locally observable in a compact set $\Omega \subset \mathbb{R}^n$. Then any indistinguishability class in Ω is finite and there is a common bound on the number of elements in a class.

Proposition 1. *If the system Σ is locally observable on compact set Ω then $\exists N \in \mathbb{N} \forall k \geq N \exists \varphi_k : \mathbb{R}^N \rightarrow \mathbb{R}$, φ_k – continuous function such that*

$$L_f^k h = \varphi_k(h, \dots, L_f^{N-1} h).$$

Proof. We assume that the system Σ is locally observable on Ω . Then the set

$$[x] := \{\tilde{x} \in \Omega : x \text{ and } \tilde{x} \text{ are indistinguishable}\}$$

is a discrete set, as indistinguishable points are isolated and there are only finitely many of them in each compact subset of the state space. The criterion of local observability given in [2, 3] is based on the sequence $(L_f^i h)_{i \geq 0}$. Let I_x denote the ideal generated by the germs of the functions $(L_f^i h - L_f^i h(x))_{i \geq 0}$. Then

$$\forall x \in \Omega \exists N_x \langle h - h(x), \dots, L_f^{N_x-1} h - L_f^{N_x-1} h(x) \rangle = I_x$$

because the ring of analytic function germs is Noetherian.

Let $N = \max_{x \in \Omega} N_x$. So let us assume that the first N functions

$$h, L_f h, \dots, L_f^{N-1} h$$

determine the indistinguishability relation. It means that if

$$Y = (h(x), \dots, L_f^{N-1} h(x)),$$

then the indistinguishability class of x , denoted by $[x]$, consists of those $\tilde{x}$ that $(h(\tilde{x}), \dots, L_f^{N-1} h(\tilde{x})) = Y$.

For $k \geq N$ the function $L_f^k h$ doesn't distinguish points of the set $[x] = \{\tilde{x} \in \Omega : \Phi_N(x) = \Phi_N(\tilde{x})\} \in \Omega/\sim$, where $\Phi_N : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^N$ i $\Phi_N(x) = (h(x), \dots, L_f^{N-1} h(x))$. The function $L_f^k h : \mathbb{R}^n \rightarrow \mathbb{R}$ is analytic for every $k \geq N$.

Denote by $\tilde{\Omega}$ the quotient space of Ω with respect to the indistinguishability relation ($\tilde{\Omega} = \Omega/\sim$). Let $y_i = L_f^i h$, $i \geq 0$ and the map $\tilde{\Phi}_N : \tilde{\Omega} \rightarrow \mathbb{R}^N$ be given by $\tilde{\Phi}_N([x]) := \Phi_N(x)$. We impose the quotient topology on $\tilde{\Omega}$. Then $\tilde{\Omega}$ is compact and the map $\tilde{\Phi}_N$ is continuous and injective. Then $\Psi := \tilde{\Phi}_N^{-1}$ is a continuous map (homeomorphism) from $\tilde{\Phi}_N(\tilde{\Omega})$ to $\tilde{\Omega}$. The map Ψ assigns to the extended output $Y = (y_0, \dots, y_{N-1})$ the class of indistinguishable states that produced Y . For every $x \in \Omega$ and $k \geq N$ we have

$$L_f^k h(x) = L_f^k h([x]) = L_f^k h(\Psi(y_0, \dots, y_{N-1})) = (L_f^k h \circ \Psi)(y_0, \dots, y_{N-1})$$

Let $\varphi_k = L_f^k h \circ \Psi : \mathbb{R}^N \rightarrow \mathbb{R}$. We get then

$$L_f^k h = \varphi_k(h, L_f h, \dots, L_f^{N-1} h)$$

and φ_k is a continuous function. $\square$

When the continuous function φ_N which we get from Proposition 1 is Lipschitz on compact sets, then we can construct a multiobserver for locally observable systems (see [4]). This construction is based on the idea of Jouan and Gauthier [5].

The system $\mathcal{S}$

$$\dot{z} := F(z, y) \quad (3)$$

$$\dot{z} = g(z) \quad (4)$$

is called a *multiobserver of the system Σ on the compact subset $\Omega \subset \mathbb{R}^n$* if:

- (1) the input y is the output of the system Σ ,
- (2) g is a multivalued map whose values are finite subsets of $\mathbb{R}^n$,
- (3) $\lim_{t \rightarrow +\infty} d(x(t); \hat{x}(t)) = 0$,

where $d(x_0; A) := \min_{x \in A} |x_0 - x|$ is a distance from x_0 to A and $|\cdot|$ is the Euclidian norm. We assume that F and g are continuous in appropriate topologies.

Equation (3) is called the *dynamical* part of the multiobserver $\mathcal{S}$, while (4) is the *static* part. The multiobserver $\mathcal{S}$ will be *continuous* (*smooth*) if F and g are continuous (*smooth*).

The multiobserver $\mathcal{S}$ of the system Σ (on the compact subset $\Omega \subset \mathbb{R}^n$) is a system whose input is the output of the system Σ and whose output is a multivalued map which estimates the whole class of states that are indistinguishable from the current state of the system Σ . The state of the multiobserver $\mathcal{S}$ approximates the time derivatives of the output y of Σ based upon the knowledge of the output $y = h(x)$. To achieve this we assume that higher order derivatives of the output may be expressed as continuous functions of a finite number of its first derivatives and the function is Lipschitz on compact sets. If the system is restricted to a compact subset of the state space and is locally observable, the values of the multiobserver's output are finite sets.

The construction of multiobservers for locally observable systems can be divided into the following steps:

- (1) We compute the sequence $(L_f^i h)_{i \geq 0}$ and choose the first N function

$$y_0 = h, y_1 = L_f h, \dots, y_{N-1} = L_f^{N-1} h,$$

that determine the indistinguishability relation.

- (2) We compute higher derivatives of the output as a continuous function of lower derivatives, in particular

$$L_f^N h = \varphi_N(y_0, y_1, \dots, y_{N-1}).$$

- (3) Let us consider the map $\Phi_N : \Omega \rightarrow \mathbb{R}^N$,

$$\Phi_N(x) = (h(x), \dots, L_f^{N-1} h(x)).$$

- (4) We consider x and $\tilde{x}$ – the indistinguishable states (denote: $x \sim \tilde{x}$) and define

$$\tilde{\Omega} := \Omega / \sim,$$

$$\tilde{\Phi}_N([x]) := \Phi_N(x), \tilde{\Phi}_N : \tilde{\Omega} \rightarrow \mathbb{R}^N,$$

$\tilde{\Phi}_N$ is continuous and injective, so there is a continuous map

$$\Psi := \tilde{\Phi}_N^{-1} : \tilde{\Phi}_N(\tilde{\Omega}) \rightarrow \tilde{\Omega}.$$

- (5) We extend Ψ to a continuous multifunction g on entire $\mathbb{R}^N$,

$$g : \mathbb{R}^N \rightarrow \tilde{\Omega}.$$

- (6) We find φ a continuous extension of the function φ_N (we would like φ to be Lipschitz, in order to have exponential estimation of the derivatives of the outputs). φ appears in the dynamics of the multiobserver, so its dynamics is only continuous.

- (7) The system

$$\mathcal{S}_{\Sigma, \theta, \Omega} : \begin{cases} \dot{Z} = (A - K_\theta C)Z + K_\theta Cy + b\varphi(Z) \\ \hat{x} = g(Z) \end{cases}$$

is a continuous multiobserver of Σ , where $Z \in \mathbb{R}^N$,

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}_{N \times N}, \quad b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{N \times 1}, \quad \Delta_\theta = \begin{bmatrix} \theta & 0 & 0 & \dots & 0 \\ 0 & \theta^2 & 0 & \dots & 0 \\ 0 & 0 & \theta^3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \theta^N \end{bmatrix}_{N \times N},$$

$$\theta > 1, C = [1 \ 0 \dots 0]_{1 \times N}, K_\theta = \Delta_\theta K,$$

K is such that $(A - KC)$ is Hurwitz and φ comes from (6) and g comes from (5).

An example of a multiobserver is presented in [4].

The problem appears when the function φ_N is not Lipschitz on some compact subset and in this case we can't find a continuous extension of φ_N which would be Lipschitz. Therefore we want to improve it so that to receive a Lipschitz function on each compact set and in this way we would like to be able to construct multiobservers not only for locally observable systems, for which the function φ_N is Lipschitz, but also for each locally observable system. Let us consider the following example:

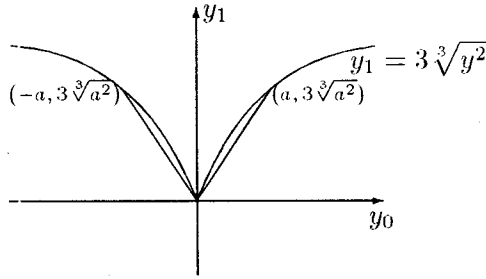
Example 2. Let Σ be the system

$$\Sigma : \begin{cases} \dot{x} = 1 \\ y = x^3 \end{cases}$$

The system Σ is locally (and globally) observable and the function

$$y_1 = \varphi_1(y_0) = 3\sqrt[3]{y_0^2}$$

is continuous but it is not Lipschitz in a neighbourhood of a point $y_0 = 0$.



Because the function φ_1 is not Lipschitz on the compact set $[-a, a]$ ($a > 0$), we are going to improve it by defining the following function

$$\tilde{\varphi}(y) := \begin{cases} \frac{3}{\sqrt[3]{a}}|y|, & \text{if } |y| \leq a \\ \varphi_1(y), & \text{if } |y| > a \end{cases}$$

The function φ_1 is only a continuous function, but $\tilde{\varphi}$ is continuous and Lipschitz so

$$\exists L = \frac{3}{\sqrt[3]{a}} \in \mathbb{R} \forall z_1, z_2 \in \mathbb{R} |\tilde{\varphi}(z_1) - \tilde{\varphi}(z_2)| \leq L|z_1 - z_2|.$$

Moreover

$$\forall z \in \mathbb{R} \tilde{\varphi}(z) \leq \varphi_1(z).$$

Let us consider the system $\mathcal{S}_{\Sigma, \theta, a}$

$$\dot{z} = -\theta kz + \theta ky + \tilde{\varphi}(z) \quad (5)$$

$$\hat{x} = \sqrt[3]{z}. \quad (6)$$

If a is small enough, the system $\mathcal{S}_{\Sigma, \theta, a}$ will be a multiobserver of Σ and the solution of the equation (5) will approximate the output y while the equation (6) gives a (single-valued) continuous estimation of $x(t)$.

Now we show that an error, which we make when the solution of (5) approximates to the output y , can be estimated by some constant which is small enough.

Set $\epsilon(t) = z(t) - y(t)$. ϵ satisfies the following equation

$$\frac{d}{dt}\epsilon(t) = -\theta k\epsilon(t) + \tilde{\varphi}(z(t)) - 3\sqrt[3]{y^2(t)}.$$

Because

$$0 \leq \varphi(y) - \tilde{\varphi}(y) \leq \frac{4}{9}\sqrt[3]{a^2}$$

so

$$\begin{aligned} \frac{1}{2} \frac{d}{dt}\epsilon^2(t) &= \epsilon(t) \cdot \dot{\epsilon}(t) = -\theta k\epsilon^2(t) + \epsilon(t)(\tilde{\varphi}(z(t)) - 3\sqrt[3]{y^2(t)}) \\ &= -\theta k\epsilon^2(t) + \epsilon(t)(\tilde{\varphi}(z(t)) - \tilde{\varphi}(y(t)) + \tilde{\varphi}(y(t)) - 3\sqrt[3]{y^2(t)}) \\ &\leq -\theta k\epsilon^2(t) + |\epsilon(t)|(|\tilde{\varphi}(z(t)) - \tilde{\varphi}(y(t))| + L|\epsilon(t)|) \\ &\leq (-\theta k + L)\epsilon^2(t) + \frac{4}{9}\sqrt[3]{a^2}|\epsilon(t)|, \end{aligned}$$

where $L = \frac{3}{\sqrt[3]{a}}$. Hence:

$$\frac{1}{2} \frac{d}{dt}\epsilon^2(t) \leq \left(-\theta k + \frac{3}{\sqrt[3]{a}}\right)\epsilon^2(t) + \frac{4}{9}\sqrt[3]{a^2}|\epsilon(t)|.$$

Let us substitute $v(t)$ ($v(t) > 0$) for $|\epsilon(t)|$, then

$$v \frac{dv}{dt} \leq \left(-\theta k + \frac{3}{\sqrt[3]{a}}\right)v^2 + \frac{4}{9}\sqrt[3]{a^2}v.$$

The following trajectory

$$v(t) = \left(|\epsilon(0)| - \frac{4}{9} \cdot \frac{a}{\theta k \sqrt[3]{a} - 3}\right) \cdot e^{(-\theta k + \frac{3}{\sqrt[3]{a}})t} + \frac{4}{9} \cdot \frac{a}{\theta k \sqrt[3]{a} - 3}$$

satisfies the equation

$$\frac{dv}{dt} = \left(-\theta k + \frac{3}{\sqrt[3]{a}}\right)v(t) + \frac{4}{9}\sqrt[3]{a^2}.$$

Therefore

$$|\epsilon(t)| \leq e^{(-\theta k + \frac{3}{\sqrt[3]{a}})t} \cdot \left(|\epsilon(0)| - \frac{4}{9} \cdot \frac{a}{\theta k \sqrt[3]{a} - 3}\right) + \frac{4}{9} \cdot \frac{a}{\theta k \sqrt[3]{a} - 3}.$$

Hence we can estimate the error as follows

$$\lim_{t \rightarrow +\infty} |\epsilon(t)| \leq \frac{4}{9} \cdot \frac{a}{\theta k \sqrt[3]{a-3}}$$

(because we can choose θ, k big enough). If a approaches zero, then $\lim_{t \rightarrow +\infty} |\epsilon(t)|$ will approach zero too.

In this case we get some family of (multi)observers which depends on a . The solution of (5) approximates the output y with an accuracy of some constant depended on a , while (6) gives a (single-valued) continuous estimation of $x(t)$.

On the basis of the above-mentioned example we want to construct some family of multiobservers for each locally observable system.

Let Σ be a locally observable system described by the equations (1) and (2) and $\Omega \subset \mathbb{R}^n$ be a compact subset of $\mathbb{R}^n$. Assume that

Assumption: There exists a family of continuous and Lipschitz functions $\tilde{\varphi}$ (with a Lipschitz constant L) such that

$$\forall \underline{y} \in \mathbb{R}^N \quad |\tilde{\varphi}(\underline{y}) - \varphi_N(\underline{y})| \leq A,$$

where $\sup_{\underline{y} \in \mathbb{R}^N} |\tilde{\varphi}(\underline{y}) - \varphi_N(\underline{y})| = A$ and $\frac{A}{L}$ goes to zero as $A \rightarrow 0$.

Let us take one function from above-mentioned family and then consider a candidate for multiobserver system $\mathcal{S}_{\Sigma, \theta, \Omega, A}$ for Σ :

$$\begin{aligned} \dot{Z} &= (\mathbf{A} - \mathbb{K}_\theta \mathbf{C})Z + \mathbb{K}_\theta \mathbf{C}y + \mathbf{b}\tilde{\varphi}(Z) \\ \hat{x} &= g(Z) \end{aligned}$$

where $Z \in \mathbb{R}^N$ and $\mathbf{A}, \mathbf{b}, \mathbf{C}, \mathbb{K}_\theta$ are defined above in point (7) of the multiobserver's construction, and g is defined in (5) of it.

We have the following theorem:

Theorem 3. For each $Z(0) \in \mathbb{R}^N$, $x(0) \in \Omega$ and some constants $\alpha \geq 0$, $s \geq 0$ the solution $Z(t)$ of $\mathcal{S}_{\Sigma, \theta, \Omega, A}$ satisfies the following inequality

$$\begin{aligned} \|Z(t) - \Phi_N(x(t))\| &\leq \theta^N \left(\frac{\sqrt{\alpha s} \|\epsilon(0)\|}{\theta} - \frac{sA}{\theta^N \sqrt{\alpha(\frac{\theta}{2} - sL)}} \right) \cdot e^{-\alpha(\frac{\theta}{2} - sL)t} + \\ &+ \frac{sA}{\sqrt{\alpha(\frac{\theta}{2} - sL)}}, \end{aligned}$$

where $\theta \geq 1$, A and L are constants. A and L are such constants that the above-mentioned assumption is true. This inequality holds for all t such that $\{x(s) : 0 \leq s \leq t\} \subset \Omega$ ($x(t)$ is the solution at time t of Σ starting from $x(0)$).

Proof. Take S , a positive definite solution of the Lyapunov matrix equation

$$(\mathbf{A} - \mathbb{K}\mathbf{C})^T S - S(\mathbf{A} - \mathbb{K}\mathbf{C}) = -Id.$$

Such an S does exist by the standard linear theory (see [6]). Let $x(0) \in \Omega$ and $Z(0) \in \mathbb{R}^N$.

We notice that

$$\frac{d}{dt} \Phi_N(x(t)) = \mathbf{A} \cdot \Phi_N(x(t)) + \mathbf{b}\varphi_N(\Phi_N(x(t))),$$

as long as $x(t) \in \Omega$.

Set $\Phi_N(x(t)) = H(t)$ and $\epsilon(t) = Z(t) - H(t)$. Then ϵ satisfies

$$\frac{d}{dt} \epsilon(t) = (\mathbf{A} - \mathbb{K}_\theta \mathbf{C})Z(t) + \mathbb{K}_\theta y(t) + \mathbf{b}\tilde{\varphi}(Z(t)) - \mathbf{A}H(t) - \mathbf{b}\varphi_N(H(t)).$$

Because $\mathbb{K}_\theta y(t) = \Delta_\theta \mathbb{K}y(t) = \mathbb{K}_\theta \mathbf{C}H(t)$ so

$$\frac{d}{dt} \epsilon(t) = (\mathbf{A} - \mathbb{K}_\theta \mathbf{C})\epsilon(t) + \mathbf{b}[\tilde{\varphi}(Z(t)) - \varphi_N(H(t))].$$

Set $\epsilon(t) = \Delta_\theta \tilde{\epsilon}(t)$, $Z(t) = \Delta_\theta \tilde{Z}(t)$, $H(t) = \Delta_\theta \tilde{H}(t)$. Then we obtain

$$\frac{d}{dt} \tilde{\epsilon}(t) = \theta(\mathbf{A} - \mathbb{K}_\theta \mathbf{C})\tilde{\epsilon}(t) + \frac{1}{\theta^N} \mathbf{b}[\tilde{\varphi}(\Delta_\theta \tilde{Z}(t)) - \varphi_N(\Delta_\theta \tilde{H}(t))].$$

Hence

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} (\tilde{\epsilon}^T S \tilde{\epsilon}) &= -\frac{\theta}{2} \tilde{\epsilon}^T \tilde{\epsilon} + \tilde{\epsilon}^T S \frac{1}{\theta^N} \mathbf{b}[\tilde{\varphi}(\Delta_\theta \tilde{Z}(t)) - \tilde{\varphi}(\Delta_\theta \tilde{H}(t))] + \tilde{\varphi}(\Delta_\theta \tilde{H}(t)) - \varphi_N(\Delta_\theta \tilde{H}(t)) \\ &\leq -\frac{\theta}{2} \|\tilde{\epsilon}\|^2 + \tilde{\epsilon}^T S \frac{1}{\theta^N} \mathbf{b}[\tilde{\varphi}(\Delta_\theta \tilde{Z}(t)) - \tilde{\varphi}(\Delta_\theta \tilde{H}(t))] + \tilde{\epsilon}^T S \frac{1}{\theta^N} \mathbf{b}[\tilde{\varphi}(\Delta_\theta \tilde{H}(t)) - \varphi_N(\Delta_\theta \tilde{H}(t))] \\ &\leq -\frac{\theta}{2} \|\tilde{\epsilon}\|^2 + \|S\|L\|\tilde{\epsilon}\|^2 + \|S\| \frac{1}{\theta^N} A \|\tilde{\epsilon}\| \\ &\leq -\left(\frac{\theta}{2} - \|S\|L\right) \alpha \tilde{\epsilon}^T S \tilde{\epsilon} + \|S\| \frac{1}{\theta^N} A \sqrt{\alpha} \sqrt{\tilde{\epsilon}^T S \tilde{\epsilon}} \end{aligned}$$

for some $\alpha \geq 0$.

We have to solve the inequality

$$\frac{1}{2} \frac{d}{dt} (\tilde{\epsilon}^T S \tilde{\epsilon}) \leq -\left(\frac{\theta}{2} - \|S\|L\right) \alpha \tilde{\epsilon}^T S \tilde{\epsilon} + \|S\| \frac{1}{\theta^N} A \sqrt{\alpha} \sqrt{\tilde{\epsilon}^T S \tilde{\epsilon}}.$$

Let $\sqrt{\tilde{\epsilon}^T S \tilde{\epsilon}(t)} := v(t)$ ($v(t) \geq 0$) then we obtain

$$v \frac{dv}{dt} \leq - \left(\frac{\theta}{2} - \|S\|L \right) \alpha v^2 + \|S\| \frac{1}{\theta^N} A \sqrt{\alpha} v.$$

If $v(t) = 0$, then $\|\epsilon(t)\| = 0$.

We can assume that $v(t) > 0$ then we have

$$\frac{dv}{dt} \leq - \left(\frac{\theta}{2} - \|S\|L \right) \alpha v + \|S\| \frac{1}{\theta^N} A \sqrt{\alpha}.$$

The trajectory

$$v(t) = \left(\sqrt{\tilde{\epsilon}^T S \tilde{\epsilon}(0)} - \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right) \cdot e^{-\alpha(\frac{\theta}{2} - \|S\|L)t} + \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)}$$

satisfies the equation

$$\frac{dv}{dt} = - \left(\frac{\theta}{2} - \|S\|L \right) \alpha v + \|S\| \frac{1}{\theta^N} A \sqrt{\alpha}.$$

Because $\tilde{\epsilon}^T S \tilde{\epsilon}(t) = v^2(t)$ so

$$\begin{aligned} \tilde{\epsilon}^T S \tilde{\epsilon} \leq & e^{-2\alpha(\frac{\theta}{2} - \|S\|L)t} \cdot \left(\sqrt{\tilde{\epsilon}^T S \tilde{\epsilon}(0)} - \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right)^2 + \\ & + e^{-\alpha(\frac{\theta}{2} - \|S\|L)t} \cdot \frac{2\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \cdot \left(\sqrt{\tilde{\epsilon}^T S \tilde{\epsilon}(0)} - \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right) + \\ & + \frac{\|S\|^2 A^2}{\theta^{2N} \alpha (\frac{\theta}{2} - \|S\|L)} \end{aligned}$$

and

$$\begin{aligned} \|\tilde{\epsilon}\|^2 \leq & e^{-2\alpha(\frac{\theta}{2} - \|S\|L)t} \cdot \left(\sqrt{\alpha\|S\|\tilde{\epsilon}^T \tilde{\epsilon}(0)} - \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right)^2 + \\ & + e^{-\alpha(\frac{\theta}{2} - \|S\|L)t} \cdot \frac{2\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \cdot \left(\sqrt{\alpha\|S\|\tilde{\epsilon}^T \tilde{\epsilon}(0)} - \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right) + \\ & + \frac{\|S\|^2 A^2}{\theta^{2N} \alpha (\frac{\theta}{2} - \|S\|L)}. \end{aligned}$$

For $\theta \geq 1$:

$$\|\epsilon\|^2 = \tilde{\epsilon}^T \cdot \Delta_\theta \cdot \tilde{\epsilon} \leq \theta^{2N} \|\tilde{\epsilon}\|,$$

Hence

$$\|\epsilon\|^2 \leq \theta^{2N} \left[\left(\sqrt{\alpha\|S\|\tilde{\epsilon}^T(0)(\Delta_\theta)^{-2}\tilde{\epsilon}(0)} - \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right) \cdot e^{-\alpha(\frac{\theta}{2} - \|S\|L)t} + \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right]^2.$$

Then

$$\|\epsilon\| \leq \theta^N \left(\frac{\sqrt{\alpha\|S\|\|\epsilon(0)\|}}{\theta} - \frac{\|S\|A}{\theta^N \sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)} \right) \cdot e^{-\alpha(\frac{\theta}{2} - \|S\|L)t} + \frac{\|S\|A}{\sqrt{\alpha}(\frac{\theta}{2} - \|S\|L)}.$$

If A goes to zero and t goes to infinity, then $\epsilon(t)$ will approach zero. $\square$

References

- [1] Z. BARTOSIEWICZ, *Local Observability of Nonlinear Systems*, System & Control Letters, 25 (1995), pp. 295–298.
- [2] Z. BARTOSIEWICZ, *Real Analytic Geometry and Local Observability*, in Differential Geometry and Control, Proceedings of Symposia in Pure Mathematics, G. Ferreyra, R. Gardner, H. Hermes, and H. Sussmann, eds., American Mathematical Society. 1998, pp. 65–72.
- [3] Z. BARTOSIEWICZ AND D. MOZYRSKA, *Algebraic Criteria for Stable Local Observability of Analytic Systems on $\mathbb{R}^n$* , in Proceedings of European Control Conference ECC-97, Brussels, July 1997.
- [4] Z. BARTOSIEWICZ AND M. WYRWAS, *On Multiobservers for Nonlinear Systems*, in Progress in Simulation, Modelling, Analysis and Synthesis of Modern Electrical and Electronic Devices and Systems, World Scientific, 1999.
- [5] P. JOUAN AND J.-P. GAUTHIER, *Finite Singularities of Nonlinear Systems. Output Stabilization, Observability and Observers*, J. Dynamical and Control Systems, 2 (1996), pp. 255–288.
- [6] T. KACZOREK, *Teoria sterowania i systemów*, Wydawnictwo Naukowe PWN, wydanie trzecie ed., 1999

Instytut Matematyki i Fizyki, Politechnika Białostocka, Wiejska 45, 15-351 Białystok

MULTIOBSERWATORY DLA UKŁADÓW NIELINIOWYCH

Streszczenie

W artykule podana jest konstrukcja pewnej rodziny multiobserwatorów dla układów lokalnie obserwowalnych. Udowodniono, że dynamika multiobserwatora estymuje kolejne pochodne wyjścia bazując na znanym $y = h(x)$. Wyjściem multiobserwatora jest multifunkcja, która dostarcza jednocześnie skończenie wiele estymat stanu układu oryginalnego.

Dorota Żebrowska-Suchodolska

CONTROLLABILITY OF DISCRETE-TIME INFINITE-DIMENSIONAL LINEAR SYSTEMS

Abstract: Discrete-time infinite-dimensional linear control systems, with one-dimensional output, are studied. They are described by infinite matrices with finite columns i.e. columns with finitely many nonzero elements. Conditions of controllability of such systems are given.

Key-words: Infinite-dimensional system, linear discrete-time system, controllability, reachability.

1 Introduction

We consider discrete-time infinite-dimensional control systems with output. They are defined on space $\mathbb{R}^{(\mathbb{N})} = \{x : \mathbb{N} \rightarrow \mathbb{R}, \exists n(x) \in \mathbb{N} \forall i > n(x): x(i) = 0\}$ and described by infinite matrices with finite columns, i.e. columns with finitely many nonzero elements.

Reachability and controllability were described for a finite-dimensional space $\mathbb{R}^n$ in many articles. Reachability and controllability in this article is described for infinite dimensional space $\mathbb{R}^{(\mathbb{N})}$.

In article [2] were presented conditions for observability of discrete-time infinite-dimensional systems with output. They were described by linear functions depending on a finite number of variables, represented by matrices with finite rows. In this paper it is assumed that matrices of a system have finite columns. This will be needed for definition of controllability.

Let us consider a discrete-time infinite-dimensional system with one-dimensional output:

$$(\Sigma) : \begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k), \end{aligned} \quad (1.1)$$

where $k \in \mathbb{N}$, $x(k) \in \mathbb{R}^{(\mathbb{N})}$, $y(k) \in \mathbb{R}$, $u(k) \in \mathbb{R}$ with initial state $x(0) = x^0 = \begin{pmatrix} x_1^0 \\ x_2^0 \\ \vdots \end{pmatrix}$.

We assume that matrices A and B have finite columns i.e. columns with finitely many nonzero elements. Matrices A and B are identified with operators $A : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ and $B : \mathbb{R} \rightarrow \mathbb{R}^{\mathbb{N}}$. System (1.1) may be written in form:

$$x_i(k+1) = \sum_{j=1}^{\infty} a_{ij} x_j(k) + b_i u(k), \quad (k \in \mathbb{N}, i \in \mathbb{N}), \quad (1.2)$$

where for every j only finitely many coefficients $a_{ij} \neq 0$.

2 Solution of the system

A solution of the system (1.1) is a sequence of discrete-time functions $x = (x_1(k), x_2(k), \dots)$, $x_i : \mathbb{N} \cup \{0\} \rightarrow \mathbb{R}$ which substituted to the equations change them to identity for some u and satisfy initial conditions $x(0) = x^0$. Such solution at time k is denoted by $x(k, x^0, u)$ and output denoted by $y(k, x^0, u) = Cx(k, x^0, u)$. It has the form:

$$x(k+1, x^0, u) = A^{k+1} x^0 + \sum_{j=0}^k A^{k-j} B u(j) \quad (k \in \mathbb{N} \cup \{0\}), \quad (2.1)$$

or otherwise

$$x(k+1, x^0, u) = A^{k+1} x^0 + \sum_{j=0}^k A^j B u(k-j) \quad (k \in \mathbb{N}, i \in \mathbb{N}). \quad (2.2)$$

Definition 2.1. The function $f * g = \sum_{s=0}^k f(k-s)g(s)$ is called the convolution of the functions f and g ; $f, g : \mathbb{N} \cup \{0\} \rightarrow \mathbb{R}$.

3 Controllability

Let $\mathbb{R}^{(\mathbb{N})}$ be the space of finite columns (columns with finitely many nonzero elements) and $\mathbb{R}^{\mathbb{N}}$ the space of all infinite columns. Then by $(\mathbb{R}^{(\mathbb{N})})'$ we denote the dual space to $\mathbb{R}^{(\mathbb{N})}$. The space $(\mathbb{R}^{(\mathbb{N})})'$ is isomorphic to $\mathbb{R}^{\mathbb{N}}$.

The solution of system (2.2) with $x(0) = 0$ may be written for any k in the form:

$$x(k+1) = \sum_{j=0}^k A^j B u(k-j) = (B, AB, \dots) \cdot \begin{pmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(0) \end{pmatrix} = (B, AB, \dots) \cdot \begin{pmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(0) \\ 0 \\ 0 \\ \vdots \end{pmatrix} = E \cdot w = (E * u)(k), \quad E = (B, AB, \dots), \quad w = \begin{pmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(0) \\ 0 \\ 0 \\ \vdots \end{pmatrix},$$

$$u = \begin{pmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \end{pmatrix}.$$

Let $\bar{E} : \mathbb{R}^{(\mathbb{N})} \rightarrow \mathbb{R}^{(\mathbb{N})}$ and $\bar{E}(w) = Ew$, where w is as above and $k \in \mathbb{N}$ is arbitrary.

Proposition 3.1. The set of all $E \cdot w$ is a linear subspace of $\mathbb{R}^{(\mathbb{N})}$, where

$$w \in \mathbb{R}^{(\mathbb{N})}, \quad \bar{E} : \mathbb{R}^{(\mathbb{N})} \rightarrow \mathbb{R}^{(\mathbb{N})}, \quad E = (B, AB, \dots), \quad w = \begin{pmatrix} u(k) \\ u(k-1) \\ \vdots \\ u(0) \\ 0 \\ 0 \\ \vdots \end{pmatrix}, \quad k \in \mathbb{N}.$$

Definition 3.1. $\text{Im} \bar{E}$ is the reachability set with zero.

We denote the reachability set by $R(0)$.

Definition 3.2. The system is controllable when the reachable set $R(0) = \mathbb{R}^{(\mathbb{N})}$.

Theorem 3.2. System Σ is controllable $\Leftrightarrow$ for $v \in \mathbb{R}^{\mathbb{N}}$ $v^T \cdot (B, AB, \dots) = 0 \Rightarrow v = 0 \Leftrightarrow \forall k \in \mathbb{N} v^T A^k B = 0 \Rightarrow v = 0$ where $v \in \mathbb{R}^{\mathbb{N}}$. $\square$

Proof. $R(0) = \mathbb{R}^{\mathbb{N}}$ (with definition (3.2)) $\Leftrightarrow v \perp R(0) \Rightarrow v = 0$ for $v \in \mathbb{R}^{\mathbb{N}}$ $\Leftrightarrow v^T \cdot \text{im } \bar{E} = 0 \Rightarrow v = 0 \Leftrightarrow ((\forall w \in \mathbb{R}^{\mathbb{N}} : v^T E w = 0) \Rightarrow v = 0)$. $\square$

4 Controllability and observability

Let us consider a discrete-time infinite-dimensional system with one-dimensional output:

$$(\Sigma) : \begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k), \end{cases} \quad (4.1)$$

where $k \in \mathbb{N}$, $x(k) \in \mathbb{R}^{\mathbb{N}}$, $y(k) \in \mathbb{R}$. (A and B have finite columns) and

$$(\Sigma^T) : \begin{cases} x(k+1) &= A^T x(k) + C^T u(k) \\ y(k) &= B^T x(k), \end{cases} \quad (4.2)$$

where $k \in \mathbb{N}$, $x(k) \in \mathbb{R}^{\mathbb{N}}$, $y(k) \in \mathbb{R}$ (A and B have finite rows)

$$y(k, x^0, u) = Cx(k, x^0, u).$$

Definition 4.1. We say that $x_1, x_2 \in \mathbb{R}^{\mathbb{N}}$ are indistinguishable (with respect to Σ) if for all $k \geq 0$:

$$y(k, x^1, u) = y(k, x^2, u)$$

for all k and u . Otherwise x_1 and x_2 are distinguishable.

We say that Σ is observable if every two distinct points are distinguishable.

Let $D = (d_{ij}) = \begin{pmatrix} C \\ CA \\ \vdots \end{pmatrix}$. Then the following theorem gives a necessary

condition for observability.

Theorem 4.1. [2] If Σ is observable then $\text{rank } D = \infty$. $\square$

Theorem 4.2. [2] If $\forall k \in \mathbb{N} \exists k' \geq k \exists i_1, \dots, i_{n_{k'}} \in \mathbb{N} : \forall i = i_1, \dots, i_{n_{k'}}, j > k'$

$$d_{ij} = 0 \text{ and } \text{rank}(d_{ij})_{\substack{i = i_1, \dots, i_{n_{k'}} \\ j = 1, \dots, k'}} = k'$$

then Σ is observable. $\square$

Theorem (4.2) is a sufficient condition for observability.

Theorem 4.3. [2] System Σ is observable $\Leftrightarrow Dx = 0 \Rightarrow x = 0$, where $x \in \mathbb{R}^{\mathbb{N}}$,

$$D = \begin{pmatrix} C \\ CA \\ \vdots \end{pmatrix}$$

Theorem 4.4. System Σ^T is observable $\Leftrightarrow \Sigma$ is controllable.

Proof. Σ^T is observable $\Leftrightarrow \begin{pmatrix} B^T \\ B^T A^T \\ \vdots \end{pmatrix} x = 0, x \in \mathbb{R}^{\mathbb{N}} \Rightarrow x = 0 \Leftrightarrow \begin{pmatrix} B^T \\ (AB)^T \\ \vdots \end{pmatrix} x = 0$, where $x \in \mathbb{R}^{\mathbb{N}} \Rightarrow x = 0 \Leftrightarrow x^T (B, AB, \dots) = 0 \Rightarrow x = 0 \Leftrightarrow \Sigma$ is controllable. $\square$

We may use the condition for observability to write down a similar condition for controllability.

Proposition 4.5. If Σ is controllable then $\text{rank}(B, AB, \dots) = \infty$. $\square$

Proposition 4.6. If $\forall j \in \mathbb{N} \exists n_i \geq j \exists k_1, \dots, k_{n_i} : \forall j = k_1, \dots, k_{n_i}, i > n_i$

$$e_{ij} = 0 \text{ and } \text{rank}(e_{ij})_{\substack{i = 1, \dots, n_i \\ j = k_1, \dots, k_{n_i}}} = n_i$$

then Σ is controllable. $\square$

Remark 4.7. $x^T E = 0 \Rightarrow x = 0$ where $(B, AB, \dots) = 0$ (E has finite columns) $\Leftrightarrow \exists E' : EE' = I$ (E' has finite columns).

Proof 4.8. $\exists E' : EE' = I$ (E' has finite columns) $\Leftrightarrow \forall i \in \mathbb{N} \exists n_i \in \mathbb{N} : x^T E_{n_i} = 0 \Rightarrow x_i = 0 \Leftrightarrow \forall i \in \mathbb{N} \exists n_i \in \mathbb{N} : \text{rank } E_{n_i} = n_i$, where E_{n_i} denotes the first n_i rows of E .

5 Examples

Example 5.1. Let us consider the following infinite-dimensional system:

$$(\Sigma): \begin{aligned} x_1(k+1) &= u(k) \\ x_2(k+1) &= x_1(k) \\ &\vdots \end{aligned}$$

Observe that $E = (B, AB, \dots) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \\ \vdots & & & & & \ddots \end{pmatrix}$. And it is

easy to notice that the condition from Proposition (4.6) is satisfied and Σ is controllable.

Example 5.2. Let us consider the system described by the matrices: $A =$

$$\begin{pmatrix} -1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & -1 & 0 & \dots \\ 1 & -1 & 0 & 1 & \dots \\ 1 & -1 & 0 & 0 & \dots \\ 0 & -1 & 1 & -1 & \dots \\ 0 & 0 & 1 & -1 & \dots \\ 0 & 0 & 0 & -1 & \dots \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}, \text{ and } B = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ \vdots \end{pmatrix}.$$

$$\text{Then rank } E = \text{rank} (B, AB, \dots) = \text{rank} \begin{pmatrix} 1 & 1 & -1 & \dots \\ 1 & 1 & 1 & \dots \\ 1 & 1 & -1 & \dots \\ 0 & 0 & -4 & \dots \\ 0 & -2 & 0 & \dots \\ 0 & -1 & 1 & \dots \\ 0 & -1 & 3 & \dots \\ 0 & 0 & 2 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \infty \text{ but}$$

$$\Sigma \text{ is not controllable e.g. } x = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \text{ satisfies equations:}$$

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 - 2x_5 - x_6 - x_7 &= 0 \\ -x_1 + x_2 - x_3 - 4x_4 + x_6 + 4x_7 + 2x_8 &= 0 \\ &\vdots \end{aligned}$$

Example 5.3. Matrix E like in Example (5.1). The condition from Remark (4.7) is satisfied. There exists a matrix E' that $EE' = I$, where

$$E' = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & & & & \ddots \end{pmatrix}.$$

References

- [1] Kaczorek T., *Teoria sterowania i systemów*, PWN, Warszawa 1993.
- [2] Mozyrska D., Bartosiewicz Z., *Local observability of systems on $\mathbb{R}^\infty$* , in: Proceedings of MTNS'2000, Perpignan, France.

STEROWALNOŚĆ NIESKOŃCZENIE WYMIAROWYCH UKŁADÓW LINIOWYCH Z CZASEM DYSKRETNYM

Streszczenie

Zbadano nieskończenie wymiarowe liniowe układy sterowania z dyskretnym czasem i jednowymiarowym wyjściem. Opisane są one przez nieskończone macierze ze „skończonymi kolumnami”, tzn. kolumnami mającymi skończenie wiele niezerowych elementów. Warunki na sterowalność takich układów zostały podane.

Marian Kucharski, Tadeusz Łukaszewicz,
Danuta Karpowicz, Zdzisław Lasocki

ELEKTROLIT DO WIELOWARSTWOWYCH URZĄDZEŃ ELEKTROCHROMOWYCH

Streszczenie: W publikacji przedstawiono przegląd artykułów opisujących wyniki badań przeprowadzonych w ostatnich latach nad otrzymywaniem elektrolitu stałego (żelowego) do urządzeń elektrochromowych. Podstawowy warunek stawiany takim elektrolitom to wysokie przewodnictwo właściwe (minimum $10^{-4} \text{ S} \cdot \text{cm}^{-1}$), a także odpowiednia trwałość w czasie, wysoka transmitancja w zakresie promieniowania widzialnego (VIS) oraz odpowiednia przyczepność do podłoża. Elektrolit żelowy spełniający te wymagania powinien zawierać:

- przewodnik jonowy, którym powinna być sól litu np.: LiCF_3SO_3 – trifluorosulfonian litu, LiClO_4 – chloran(VII) litu lub $\text{LiN}(\text{CF}_3\text{SO}_2)_2$ – trifluorosulfonimidek litu.
- matrycę polimerową np.: polimetakrylan metylu (PMMA), poli(chlorek winylu) (PCV), poli(tlenek etylenu) (PEO) lub poliakrylonitryl (PAN) w odpowiedniej postaci.
- odpowiedni rozpuszczalnik lub ich mieszaninę z następujących: węgiel propylenowy (PC), węgiel etylenu (EC), γ -butyrolakton (γ -BL), odpowiedniej czystości.
- rozpuszczalnik do wstępnego przygotowania żelu matrycy, np. acetonitryl (AcCN) lub tetrahydrofuran (THF).

Wykonane badania własne pozwoliły otrzymać elektrolit żelowy o składzie: 15% PEO, 40% PC, 40% EC, 5% LiClO_4 , którego przewodnictwo wynosiło $8,4 \text{ mS} \cdot \text{cm}^{-1}$ w temperaturze pokojowej po 3 dobach żelowania. Elektrolit charakteryzował się trwałością w czasie, i wysoką transmitancją promieniowania w zakresie widzialnym, co daje możliwość jego zastosowania w urządzeniach elektrochromowych.

Słowa kluczowe: elektrolit żelowy (stały), urządzenie elektrochromowe, sól litu, matryca polimerowa, rozpuszczalnik.

Przegląd literaturowy

Jednym z najważniejszych elementów wielowarstwowego urządzenia elektrochromowego jest warstwa elektrolitu o odpowiednio dużym przewodnictwie właściwym (powyżej $10^{-4} \text{ S} \cdot \text{cm}^{-1}$). Dodatkowo elektrolit musi się charakteryzować trwałością w czasie, łatwością nanoszenia na warstwy elektrochromowe, dobrą przyczepnością (kleistością), odpornością na temperaturę, a właściwie zachowa-

niem odpowiednich właściwości w dużym zakresie temperatury (od -20 do 60°C). Trudności z otrzymaniem elektrolitu spełniającego te wymagania spowodowały, że w wielu ośrodkach naukowych zajmowano się tym zagadnieniem w ostatnich latach [1-12].

Elektrolit stały (żelowy) opisywany w literaturze zawiera jako podstawowy składnik przewodnik jonowy (jest to najczęściej sól litu), matrycę organiczną, którą stanowi odpowiednio spreparowane tworzywo sztuczne oraz odpowiednie rozpuszczalniki organiczne (bezwodne). Rozpuszczalniki te muszą umożliwiać dysocjację i odpowiednią ruchliwość jonów, a tym samym odpowiednie przewodnictwo.

W ostatnich latach badania prowadzone są w kierunku uzyskiwania jak największego przewodnictwa i trwałości, co daje szansę długiej pracy urządzenia elektrochromowego z dobrą pamięcią, to znaczy z identycznym (lub zbliżonym) zachowaniem, jak na początku, nawet po wielu tysiącach cykli pracy urządzenia.

Su L. i współpracownicy [1] przygotowując elektrolit stały zastosowali, jako przewodnik jonowy chloran(VII) litu (LiClO_4) oraz matrycę z poli(tlenku etylenu); podstawowym rozpuszczalnikiem był węglan propylenu. Zależność przewodnictwa tego elektrolitu od temperatury ma przebieg typu Arrheniusa. Badanie metodą dyfrakcji rentgenowskiej wykazało, że polimer (matryca) był amorficzny (o strukturze amorficznej). Z nachylenia prostej (zależność przewodnictwa σ [$\text{S}\cdot\text{cm}^{-1}$] od $1/T$ [K]) obliczono energię aktywacji E_a jonowego przewodnictwa elektrolitu, wykorzystując zależność $\sigma = \sigma_0 \exp(-E_a/kT)$, gdzie k to stała Boltzmana, T – temperatura [K]. Energia pseudoaktywacji dla elektrolitu żelowego wynosiła $31,7 \text{ kJ/mol}$, gdy tymczasem dla roztworu chloranu(VII) litu (LiClO_4) w węglanie propylenu wynosi $13,1 \text{ kJ/mol}$. Przewodnictwo elektrolitu na początku wynosiło $4,6 \cdot 10^{-4} \text{ S}\cdot\text{cm}^{-1}$, a po 3 dobach wynosiło $3,2 \cdot 10^{-4} \text{ S}\cdot\text{cm}^{-1}$ i nie zmieniło się po następnych dobach.

W innej publikacji Su L. i współpracownicy [2] przedstawiają wyniki badań nad elektrolitem stałym z matrycą z poli(chlorku winylu). Skład (% wagowy) elektrolitu był następujący:

15% PCV (poli(chlorek winylu)), 40% PC (węglan propylenu),

40% EC (węglan etylenu), 5% LiClO_4 (chloran(VII) litu).

Mieszanina stanowiąca elektrolit przed nanoszeniem rozpuszczona była w tetrahydrofuranie (THF). Wartość przewodnictwa elektrolitu stałego ustaliła się po 3 dobach i wynosiła około $2,3 \text{ mS}\cdot\text{cm}^{-1}$, a dla porównania przewodnictwo chloranu(VII) litu- LiClO_4 o stężeniu 1 mol/dm^3 w węglanie propylenu (PC) wynosi $8 \text{ mS}\cdot\text{cm}^{-1}$.

Stosunkowo wysokie przewodnictwo elektrolitu z matrycą PVC świadczy o tym, że ta matryca nie ma istotnego wpływu na rezystancję. Zastosowanie jako rozpuszczalnika mieszaniny EC i PC powoduje wysoką dysocjację soli litu, a równocześnie małą lepkość w układzie tych rozpuszczalników, co umożliwia dużą ruchliwość jonów.

Modyfikacja składu elektrolitu to znaczy zastosowanie matrycy z poli(chlorku winylu), oraz zastosowanie dodatkowo tetrahydrofuranu (THF), jako rozpuszczalnika do matrycy przed nałożeniem elektrolitu na warstwę elektrochromową, pozwoliło uzyskać elektrolit o przewodnictwie równym około $2 \text{ mS}\cdot\text{cm}^{-1}$ [3].

Zastosowanie matrycy z poli(tlenku etylenu) i dobranie optymalnego stężenia chlorku litu, a także układu rozpuszczalników (węglanu propylenu i węglanu etylenu w stosunku wagowym 1:2) pozwoliło otrzymać elektrolit żelowy o przewodnictwie właściwym też równym około $2 \text{ mS}\cdot\text{cm}^{-1}$ [4]. W pracy tej, opierając się na badaniach voltamperometrycznych w układzie trójelektrodowym wyznaczono współczynnik dyfuzji, który wynosił $D = 3 \cdot 10^{-10} \text{ cm}^2 \cdot \text{S}$, co jest wartością podobną dla tego typu elektrolitów [5].

Kolejny krok w poszukiwaniu elektrolitu spełniającego jak najlepiej wszystkie wymagania urządzenia elektrochromowego to zastosowanie polimetakrylanu metylu (PMMA), jako matrycy [6] w elektrolicie żelowym. Skład tego elektrolitu był następujący:

25% PMMA	Polimetakrylan metylu	} Mieszaninę elektrolitu rozpuszczono w tetrahydrofuranie (THF)
35% PC	Węglan propylenu	
35% EC	Węglan etylenu	
5% LiClO_4	Chloran(VII) litu	

Obliczono energię aktywacji dla tego elektrolitu ($E_{ao} = 16,9 \text{ kJ mol}^{-1}$). Dla porównania energia aktywacji dla 1 M roztworu LiClO_4 w węglanie propylenu wynosi $13,1 \text{ kJ mol}^{-1}$. Interesujące, że przewodnictwo elektrolitu po 3 dobach spadło tylko o $0,0014 \text{ S}\cdot\text{cm}^{-1}$, gdy równocześnie przewodnictwo roztworu LiClO_4 (1 mol/dm^3) w węglanie propylenu spadło o $0,008 \text{ S}\cdot\text{cm}^{-1}$.

Sekhon i współpracownicy [7,8] podejmują badania i dyskusję nad problemem stosowania rozpuszczalników w przygotowywaniu i właściwościach otrzymywanych elektrolitów stałych. W pracy [8] zaprezentowano wyniki badań nad elektrolitem złożonym z:

- matrycy polimerowej na bazie polimetakrylanu metylu (PMMA),
- rozpuszczalnika jednoskładnikowego lub mieszanego dwuskładnikowego z następujących: węglan propylenu (PC), węglan etylenu (EC), γ -butyrolaktonu (γ -BL) oraz
- przewodnika jonowego – trifluorosulfonianu litu (LiCF_3SO_3).

Sposób przygotowania wymienionego elektrolitu był następujący: sól rozpuszczono w rozpuszczalniku (lub mieszaninie rozpuszczalników), a następnie dodawano polimer w rozpuszczalniku i dokładnie mieszano. Wówczas następowało częściowe odparowanie rozpuszczalnika i po pewnym czasie elektrolit można było zastosować.

Wcześniejsze badania wykazały, że polimer (PMMA) nie uczestniczy w przewodnictwie omawianego elektrolitu.[9]. Z drugiej jednak strony polimetakrylan metylu wykazuje wysoką transmitancję w zakresie promieniowania widzialnego (VIS) i stąd możliwość jego zastosowania w urządzeniach elektrochromowych.

Agnitory i współpracownicy [10] otrzymali nowy elektrolit do „Smart Windows” o przewodnictwie $3 \cdot 10^{-3} \text{ S} \cdot \text{cm}^{-1}$. Badany elektrolit zawierał sól litu $\text{LiN}(\text{CF}_3\text{SO}_2)_2$ lub LiCF_3SO_3 , rozpuszczalnik węglan propylenu (lub mieszaninę węglan propylenu + węglan etylenu) oraz matrycę z PMMA. Po różnych próbach i badaniach zaproponowano elektrolit, w którym stężenie soli litu wynosiło 1 mol/dm^3 , a elektrolit zawierał 20% PMMA oraz PC (lub $\text{PC} + \text{EC} = 1 + 1, 1+3$). Badaniom poddano elektrolit o następującym składzie: PMMA – PC/EC/ γ -BL – LiCF_3SO_3 , gdzie: PMMA – polimetakrylan metylu (matryca),

PC – węglan propylenu,

EC – węglan etylenu,

LiCF_3SO_3 – trifluorosulfonian litu.

Ostatecznie przeprowadzono badanie właściwości elektrolitów, które zawierały rozpuszczalniki pojedyncze lub układy dwuskładnikowe w następujących wariantach:

PC, EC, γ -BL – jako jednoskładnikowe

PC + γ -BL, PC + EC, EC + γ BL – jako dwuskładnikowe

W wyniku przeprowadzonych badań stwierdzono, że najwyższe przewodnictwo wykazuje roztwór soli (LiCF_3SO_3) o stężeniu od 0,8 do $1,2 \text{ mol/dm}^3$. Powyżej stężenia od $1,2 \text{ mol/dm}^3$ przewodnictwo istotnie maleje (tworzą się pary jonowe, co obniża przewodnictwo).

Jeżeli chodzi o wpływ zastosowanych rozpuszczalników, to najniższe przewodnictwo przy określonym stężeniu soli wykazuje roztwór węglanu propylenu

(PC), a najwyższe roztwór w węglanie etylenu (EC). Kolejność przewodnictwa według stosowanych rozpuszczalników jest następująca:

$$\text{PC} < \text{PC} + \text{EC} < \text{PC} + \gamma\text{-BL} < \text{EC} + \gamma\text{-BL} < \gamma\text{-BL} < \text{EC}$$

Dodawanie matrycy z PMMA istotnie zmniejsza przewodnictwo co zbadano w zakresie od 0 do 30% wagowej zawartości PMMA w otrzymanym żelowym elektrolicie. Tłumaczy się to przede wszystkim wzrostem lepkości układu. Do zawartości 10% PMMA spadek przewodnictwa jest mniejszy. Badanie przewodnictwa w funkcji temperatury daje zależność typu Arrheniusa, ale charakterystyczną dla elektrolitów o wysokiej lepkości. Zdaniem autorów tego artykułu [8] takie cechy otrzymanego elektrolitu, jak:

- wysokie przewodnictwo $\sigma = 3 + 5 \text{ mS/cm}$, przy zawartości 10% PMMA,
- wysoka transmitancja w zakresie widzialnym,
- duża przyczepność (kleistość),

powodują, że ten elektrolit, jako elektrolit żelowy, bardzo dobrze nadaje się do urządzeń typu elektrochromowego.

W pracy [11] przedstawiono wyniki badań nad nowym elektrolitem dla urządzeń elektrochromowych, w którym przewodnikiem jonowym był trifluorosulfonimidek litu $\text{LiN}(\text{CF}_3\text{SO}_2)_2$, matrycą – polimetakrylan metylu (PMMA) oraz rozpuszczalniki γ -butyrolacton (γ -BL), węglan propylenu (PC) i węglan etylenu (EC). Rozpuszczalników używano pojedynczo lub jako układów dwuskładnikowych. Przewodnictwo maksymalne dla elektrolitu z γ -BL wynosiło $1,26 \cdot 10^{-2} \text{ S} \cdot \text{cm}^{-1}$, przy stężeniu związku soli litu, wynoszącym $1,2 \text{ mol/dm}^3$. Zastosowanie rozpuszczalników w układzie dwuskładnikowym (np. γ -BL + EC, γ -BL + PC) powodowało w rezultacie mniejsze przewodnictwo. Zwiększenie procentowej zawartości matrycy (PMMA) powodowało także zmniejszanie przewodnictwa. Ostatecznie zaproponowano następujący skład elektrolitu:

25% PMMA, $1,25 \text{ mol/dm}^3 \text{ LiN}(\text{CF}_3\text{SO}_2)_2$, γ -BL + EC = 1 + 1 lub (γ -BL lub EC)

Przewodnictwo takiego elektrolitu wynosiło około $10^{-3} \text{ S} \cdot \text{cm}^{-1}$, a w zakresie temperatury 10–85°C zaobserwowano mały wpływ temperatury na wartość zmianę przewodnictwa.

W innych badaniach Sekhon i współpracownicy [12] przedstawili wyniki badań nad elektrolitem w którego skład wchodziła matryca – poliakrylonitrylowa (PAN), elektrolit – trifluorosulfonian litu – LiCF_3SO_3 i rozpuszczalniki – węglan etylenu (EC) oraz węglan propylenu (PC). Wykonano badania dla różnej proporcji rozpuszczalników, różnego stężenia elektrolitu litowego, różnej zawartości matrycy. Ostatecznie zaproponowano następujący skład elektrolitu: stężenie soli wynosiło $1 \text{ mol/dm}^3 \text{ LiCF}_3\text{SO}_3$, 20% (wagowo) PAN, $\text{PC} + \text{EC} = 1 + 3$. Przewodnictwo

tego elektrolitu wynosiło około $2 \text{ mS} \cdot \text{cm}^{-1}$. Sposób przygotowania elektrolitu był następujący: sól litową rozpuszczono w odpowiednim (lub mieszaninie) rozpuszczalniku, następnie dodawano PAN i otrzymano po mieszanii elektrolit żelowy o odpowiedniej gęstości i lepkości.

Z przeprowadzonego przeglądu literaturowego wynika, że elektrolit stały (żelowy) zastosowany w wielowarstwowym urządzeniu elektrochromowym powinien zawierać poniżej wymienione składniki:

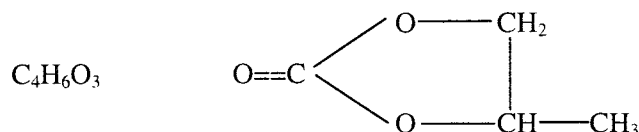
1. Przewodnik jonowy, którym powinna być sól litu np. LiCF_3SO_3 – trifluorosulfonian litu, LiClO_4 – chloran(VII) litu, lub $\text{LiN}(\text{CF}_3\text{SO}_2)_2$ – trifluorosulfoni-midek litu.
2. Marycę polimerową np.: polimetakrylan metylu (PMMA), poli(chlorek winylu) PCV, poli(tlenek etylenu) PEO lub poliakrylonitryl (PAN).
3. Odpowiedni rozpuszczalnik lub mieszaninę – węglan propylenu (PC), węglan etylenu (EC) lub γ -butyrolakton (γ -BL).
4. Rozpuszczalnik do przygotowania żelu matrycy np. acetonitryl lub tetrahydrofuran.

Elektrolit powinien charakteryzować się przewodnictwem właściwym minimum $10^{-4} \text{ S} \cdot \text{cm}^{-1}$, (a jeszcze lepiej, jeśli przewodnictwo wynosi powyżej $10^{-3} \text{ S} \cdot \text{cm}^{-1}$), odpowiednią trwałością, wysoką transmitancją w zakresie promieniowania widzialnego (VIS) oraz odpowiednią przyczepnością do podłoża.

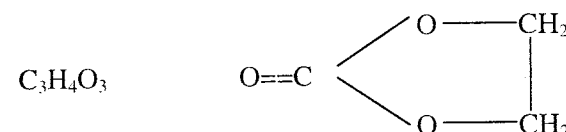
Przygotowane elektrolitu stałego

Do badań zastosowano następujące substancje:

1. PVC – $[-\text{CH}_2\text{CH}(\text{Cl})-]_m$ – poly(vinyl chloride), poli(chlorek winylu) firmy Aldrich
2. PEO – $[-\text{CH}_2\text{CH}_2\text{O}-]_n$ M_v ca 600.000
poly(ethylene oxide) poli(etyleny tlenek) firmy Aldrich
3. PC – propylene carbonate (węglan propylenu) firmy Fluka



4. EC – ethylene carbonate (węglan etylenu) firmy Fluka



5. LiClO_4 chloran(VII) litu firmy POCH-Gliwice
6. Acetonitryl CH_3CN czystości chromatograficznej (do HPLC)
7. LiCF_3SO_3 – trifluorosulfonian litu firmy Merck
8. γ -butyrolacton (γ -BL) firmy Merck
9. Polimetakrylan metylu (PMMA) firmy Aldrich

Przygotowane elektrolitu stałego z matrycą z PVC

Na podstawie danych literaturowych przygotowano elektrolit o następującym składzie:

15% PVC, 40% PC, 40% EC, 5% LiClO_4

Sposób przygotowania

1,5g PVC rozpuszczono w 10 cm^3 THF (intensywnie mieszano i rozcierano). Dodano 4 g EC i mieszając powoli wprowadzono 5 cm^3 roztworu LiClO_4 w PC o stężeniu 1 mol/dm^3 .

Przy próbie zastosowania elektrolitu pojawiły się jednak w tym wypadku trudności z uzyskaniem żelu jednorodnego, a dodatkowo elektrolit nie skleił płytek i powstały po pewnym czasie białe plamy (oddzielanie matrycy), co zmniejszyło transmitancję urządzenia elektrochromowego w zakresie promieniowania widzialnego (VIS). Zauważono niejednorodność elektrolitu w warstwie.

Elektrolit stały z PEO

Skład procentowy taki sam, jak poprzednio to znaczy:

15% PEO, 40% PC, 40% EC, 5% LiClO_4

Przygotowanie elektrolitu

1,5g PEO rozpuszczono w 5 cm^3 acetonitrylu (energicznie mieszając i rozcierając bagietką). Następnie dodawano powoli roztwór LiClO_4 w PC/EC, który przygotowano w następujący sposób: odważono 4g EC i 4g PC, zmieszano i rozpuszczo-

no, mieszając mieszadłem magnetycznym, po czym dodano 0,5g LiClO_4 , który rozpuszczono w mieszaninie EC/PC.

Tak otrzymany elektrolit wykazywał niskie przewodnictwo właściwe (poniżej $10^{-4} \text{ S} \cdot \text{cm}^{-1}$). Następnie zwiększono w dalszych próbach stężenie przewodnika jonowego do 10%, co odpowiada stężeniu około 1 mol/dm^3 .

Przygotowanie elektrolitu stałego o wyższym przewodnictwie właściwym

1. Przygotowanie matrycy z poli(tlenku etylenu) – PEO:
1,5g PEO rozpuszczono (w wysokiej zlewce poj. 50 cm^3) w 5 cm^3 acetonitrylu (CH_3CN). PEO wsypywano bardzo powoli i energicznie mieszano zawartość za pomocą odpowiedniej bagietki (lub z zastosowaniem mieszadła magnetycznego).
2. Przygotowanie roztworu elektrolitu:
Do zlewki o pojemności 50 cm^3 odważono na wadze analitycznej 5g węglanu etylenu (EC) i do tego powoli odważono 2,5g (około $2,1 \text{ cm}^3$) roztworu węglanu propylenu (PC). Następnie rozpuszczono EC w PC intensywnie mieszając przy użyciu mieszadła magnetycznego. W otrzymanym roztworze rozpuszczono 1g LiClO_4 cz.d.a. (roztwór zajmuje objętość $= 6 \text{ cm}^3$, co oznacza, że stężenie LiClO_4 wynosi $1,5 \text{ mol/dm}^3$ (w $\text{PC}+\text{EC}=1+2$)).
3. Przygotowanie elektrolitu stałego (żelowego)
Zmieszano oba roztwory, dodając powoli przygotowany roztwór elektrolitu (LiClO_4 w $\text{EC}+\text{PC}$) do matrycy z PEO, energicznie mieszając (rozcieńczając) za pomocą bagietki.

Ostateczny skład procentowy elektrolitu stałego:

15% PEO, 25% PC, 50% EC, 10% LiClO_4

Po 3 godzinach żel (elektrolit stały) nanoszono na warstwę elektrochromową WO_3 (która była naniesiona na warstwę przewodzącą ITO). Po równomiernym rozprowadzeniu odpowiednio przygotowanym przyrządem szklanym, przyciskano drugą płytką (z warstwą błękitu pruskiego – BP naniesionego też na warstwę ITO). Żelowanie trwało 3 doby w temperaturze pokojowej.

Przeprowadzono pomiary przewodnictwa właściwego z zastosowaniem konduktometru CC-551 firma Elmetron z czujnikiem CD-2, Nr 505 i stałą $k = 0,5 \text{ cm}^{-1}$.

Tabela 1.

Wyniki pomiarów przewodnictwa roztworu chloranu(VII) litu o stężeniu 1 mol/dm^3 w węglanie propylenu (PC)

Temperatura [°C]	δ [$\text{mS} \cdot \text{cm}^{-1}$]
23,8	5,02
26,0	5,08
27,0	5,16
28,2	5,21
29,0	5,23
30,0	5,27
32,5	5,33
40,5	5,46
50,0	5,66
60,0	6,16

Tabela 2.

Wyniki pomiarów przewodnictwa roztworu chloranu(VII) litu stężeniu 1 mol/dm^3 w mieszaninie węglanu propylenu (PC) i węglanu etylenu (EC) – (stosunek wagowy 1+2)

Temperatura [°C]	δ [$\text{mS} \cdot \text{cm}^{-1}$]
22,0	6,45
24,2	6,70
30,0	6,78
35,0	6,80
40,0	6,86
46,0	7,03
50,0	7,20
56,0	7,75
60,0	8,14

Przewodnictwo elektrolitu żelowego (15% PEO, 50% EC, 25% PC, 10% LiClO_4 , matryca przygotowana w acetonitrylu w temp. 25°C wynosiło:

po 30 min. 25°C – $9,84 \text{ mS} \cdot \text{cm}^{-1}$
po 24 godz. – $8,71 \text{ mS} \cdot \text{cm}^{-1}$
po 3 dobach – $8,40 \text{ mS} \cdot \text{cm}^{-1}$

Wysokie przewodnictwo (około $8,4 \text{ mS}\cdot\text{cm}^{-1}$), a także trwałość w czasie umożliwiają zastosowanie otrzymanego elektrolitu do wielowarstwowych układów elektrochromowych.

Literatura

1. **Su L., Fang J., Lu Z.:** „All-solid-state Electrochromic Window of Electrodeposited WO_3 and Prussian Blue with Poly(ethylene oxide) Gel Electrolyte” *Jpn.J.Appl.Phys.* pp. 5747-5750, **36**, (1997).
2. **Su L., Wang H., Lu Z.:** „All Solid-state Electrochromic Smart Window of Electrodeposited WO_3 and Prussian Blue Film with PVC Gel Electrolyte”, *Supramolecular Science*, pp.657-659, **5**, (1998).
3. **Su L., Xiao Z., Lu Z.:** „All Solid-state Electrochromic Window of Electrodeposited WO_3 and Prussian Blue Film with PVC Gel Electrolyte”, *Thin Solid Films*, pp. 285-289, **320**, (1998).
4. **Su L., Wang H., Lu Z.:** „All-Solid-State Electrochromic Window of Prussian Blue and Electrodeposited WO_3 Film with Poly(Ethylene Oxide) Gel Electrolyte” *Materials Chemistry and Physics* pp.266-270, **56** (1998).
5. **Dini D., Decker F., Masetti E.,** *J.Appl.Electrochem.* 647, **26** (1996).
6. **Su L., Xiao Z., Lu Z.:** „All Solid-state Electrochromic Device with PMMA Gel Electrolyte”, *Materials Chemistry and Physics*, pp.180-183, **52**, (1998).
7. **Sekhon S.S., Prodeep, Agnihotry S.A., in B.V.R. Chowdari, Lai K., Agnihotry S.A., Khare N., Sekhon S.S., Srivastava P.C., Chandra S.** (Eds.), *Solid State Ionics: Science and Technology*, World Scientific, p217, 1998.
8. **Sekhon S.S., Deepa , Agnihotry S.A.:** „Solvent Effect on Gel Electrolytes Containing Lithium Salts”, *Solid State Ionics*, pp.1189-1192, **136-137**, (2000)
9. **Bohnke O., Frand G., Rezrazi M., Rousselot C., Truche C.,** *Solid State Ionics*, pp. 97-105, **66** (1993).
10. **Agnihotry S.A., Pradeep, Sekhon S.S.:** „PMMA Based Gel Electrolyte for EC Smart Windows”, *Electrochimica Acta*, pp. 3121-3126, **44** (1999).
11. **Agnihotry S.A., Nidhi, Pradeep, Sekhon S.S.:** „ Li^+ Conducting Gel Electrolyte for Electrochromic Windows”, *Solid State Ionics* pp.573-576, **136-137** (2000).
12. **Sekhon S.S., Arora N., Agnihotry S.A.:** „PAN-Based Gel Electrolyte with Lithium Salts” *Solid State Ionic*, pp.1201-1204, **136-137** (2000).

THE ELECTROLITE FOR ELECTROCHROMIC DEVICES

Abstract

The paper is a review of investigations solid polymeric electrolytes and the results of the research a new gel electrolyte for electrochromic devices. Polymer gel electrolytes (polymer-salt-solvent system) for use in electrochromic devices application contained:

- Lithium salt : LiClO_4 - lithium chlorate(VII) , $\text{LiN}(\text{CF}_3\text{SO}_2)_2$ – lithium trifluorosulfonimide or LiCF_3SO_3 – lithium triflate,
- Based polymer: poly(ethylene oxide) – PEO, poli(vinyl chloride) – PVC, poliacylonitrile – PAN, polymethylmetacrylate -PMMA,
- Solvent (or mixed solvent): propylene carbonate (PC), ethylene carbonate (EC),
- γ -butyrolacton (γ -BL), tetrahydrofuran (THF).

Poly(ethylene oxide) PEO – based polymer electrolytes, containing lithium chlorate(VII) (LiClO_4) and using mixed solvent propylene carbonate (PC) + ethylene carbonate (EC) are presented the gel electrolyte prepared. The high conductivity attained by the gels ($8,4 \text{ mS}\cdot\text{cm}^{-1}$) at room temperature and thermally stable, make them very promising for electrochromic devices.

Keywords: Gel electrolyte; Electrochromic devices; Lithium salt; Based polymer; Solvent

Praca została wykonana w ramach realizacji projektu badawczego Nr 7 T08D 018 19 sponzorowanego przez Komitet Badań Naukowych.