INTEGRAL OBSERVABILITY OPERATORS OF NONLINEAR DYNAMICAL SYSTEMS

YURY V. ZAIKA

In terms of functional dependence, the description of observable functions in nonlinear dynamical systems, which are analytic with respect to phase variables, is obtained. For processing of measurements, integral operators are used, which provide certain noise stability of operation of phase state reconstruction. The analogue of the duality theory known for linear problems of observation and control is developed. Computing schemes for nonlinear observability problem are proposed.

1. Introduction. In the domain $U \subseteq \mathbb{R}^n$ consider the observation system

$$\dot{x} = f(x), \quad y = g(x), \quad f : U \longrightarrow \mathbb{R}^n, \, g : U \longrightarrow \mathbb{R}^m, \tag{1.1}$$

which models the law of motion and the available information about the phase state. We assume that the vector functions f and g are smooth and, for the main results of the paper, real analytic in U. We are given the observation time interval $[0, T]$ and the domain of all possible states $U_T = \{x(T)\} \subseteq U$, for which the solutions $x(\cdot; x, T)$ of the vector differential equation in (1.1) (where $x(T; x, T) = x \in U_T$) can be extended to $[0, T]$. The observation problem is to determine the phase vector $x = x(T)$ for the realization $y(\cdot; x, T) = g(x(\cdot; x, T)) : [0, T] \to \mathbb{R}^m$. The notation $y(\cdot; x, T)$ means that the known vector function $y(\cdot)$ at $[0, T]$ is identically determined by the unknown state x at time T. We assume that this problem is to be solved systematically. Thus we are interested in calculation for any possible realization $y(\cdot)$ of appropriate x from the domain U_T. It is possible to formulate the problem also in terms of unknown initial data $x^0 = x(0) \in U_0$. But usually the most interesting is a phase state at the end of the observation time. In case only the values $y(t)$, $t \in \Theta$ ($\Theta = [0, \vartheta]$, $\vartheta < T$), are used, then it is the prediction problem. Once we know $x(T)$ (or x^0), the motion can be reconstructed uniquely.

It is rather difficult to determine observability of the pair (f, g) directly by the conformity $x \mapsto y(\cdot; x, T)$ as inversion of the mapping to the space of vector functions is needed. Thus, here it is usually better to study the so-called observability mapping $H : x \mapsto y(\cdot; x, T) \mapsto z \in \mathbb{R}^p$, calculating the values

of p functionals on $y(\cdot)$. If H is injective on the set $V \subseteq U_T$ $(H(x) \leftrightarrow x \in V)$, then (f,g) is observable in V and the vector $x = x(T) \in V$ is uniquely identified by $z = H(x)$. The values of z are obtained from the processing of the measurements $y(\cdot)$. By definition, observability of the pair (f,g) in $V \subseteq U_T$ on $[0,T]$ means that one-to-one correspondence holds: $y(\cdot; x, T) \leftrightarrow x \in V$. The most frequently used methods of constructing H are the following:

(a) $x \mapsto (y'(t_1), \ldots, y'(t_\ell))'$, $t_i \in [0,T]$, $\ell m = p$;

(b) $x \mapsto (y'(T), \dot{y}'(T), \ldots, y^{(\ell-1)'}(T))'$;

(c) $x \mapsto (\langle k_1, y \rangle, \ldots, \langle k_p, y \rangle)'$, $\langle k, y \rangle = \int_0^T k'(\tau) y(\tau) d\tau$.

Some sufficient conditions of mapping injectivity in finite-dimensional spaces can be found, for instance, in [11, 12]. The following question is essential: in what class of pairs (f,g) the observability problem can be reduced to the problem of solvability of finite system of equations with unknown $x = x(T)$, obtained using (a), (b), and (c)?

To simplify the notation assume, without loss of generality, that $m = 1$.

Analytic observation theory was, in particular, developed in Starkov's work (see [13], which contains further bibliography). While determining $x(T)$ for a polynomial system (1.1) (a pair (f,g)), it is sufficient to calculate a finite number of derivatives $y^{(i)}(t_*)$, $t_* \in [0,T]$ (see [6]). But a necessary number of them may become arbitrarily large. For a stationary observable real-analytic pair (f,g), without loss of information about unknown $x(T)$, it suffices to have $2n + 1$ measurements $y(t_j)$ instead of $y(\cdot)$ (see [8]). Times t_j are fixed and do not depend on the realization $y(\cdot)$. But in the general case, a set of "successful" observation programs $\{t_1, \ldots, t_{2n+1}\}$ is not open in $[0,T]^{2n+1}$. Errors in fixing t_j can imply a loss of observability. Discrete observation programs that are stable to perturbations are considered in [14].

In case the measurements are considerably influenced by noises, it is preferable to use integral processing of $y(t)$. The fundamentals of the appropriate mathematics for linear case can be found, for instance, in [9]. Let $f = Fx$ and $g = Gx$. Here F and G are $n \times n$ and $m \times n$ matrices, respectively. If in the dual system

$$\dot{V}(t) = -F'V(t) + G'k(t), \quad V(0) = 0, \tag{1.2}$$

a control $k(\cdot)$ is built according to the condition $V(T) = h$, then according to the information $y(\cdot)$ it is possible to calculate a projection of unknown $x(T)$ to vector $h : h'x(T) = \langle k, y \rangle_{L_2}$ for all $x(T) \in \mathbb{R}^n$. A set of all $h \in \mathbb{R}^n$ for which, according to any possible vector function $y(\cdot)$, $h'x(T)$ is uniquely reconstructed is described by a set of attainability $\mathcal{D}_T = \{V(T)\}$. This approach, known as a duality principle, is extended to a nonlinear case in [7]. Construction of an operator for reconstruction of values of the given function $\varphi : U_T \to \mathbb{R}^1$ in the form

$$\varphi(x(T)) = \int_0^T k(\tau, y(\tau)) d\tau \quad \forall x(T) \in U_T, \tag{1.3}$$

is reduced to the control problem: choose $k(\cdot,\cdot)$ in the duality system

$$v_t(t,x) + v_x(t,x) \cdot f(x) = k(t,g(x)), \quad v(0,x) = 0, \tag{1.4}$$

for which $v(T,x) = \varphi(x)$, $x \in U_T$. The details will be specified later. In case $(f,g) = (F,G)$ and $k(t,y) = k'(t)y$, we obtain $v(t,x) = V'(t)x$. Here $V(t)$ meets (1.2). Appearance of the partial differential equation is natural since the nonlinear observability problem for the domain of phase space is, in essence, distributed. The most important here is the fact that (1.4) is linear with respect to the pair (k,v) and it is possible to use control theory and numerical methods for solving a linear boundary value problems.

2. Observability by a finite number of projections. Firstly, we consider the linear integral operators (1.3): $\varphi(x(T)) = \langle k,y \rangle_{L_2}$, $L_2 = L_2[0,T]$, $m = 1$. In applications, admissible weight functions $k(\cdot)$ for processing $y(\cdot)$ are considered piecewise continuous. Functionals $y(\cdot) \mapsto \langle k,y \rangle$ and values $\langle k,y \rangle$ will be called projections. Here some questions arise. From the numerical point of view, it is important to have a finite-dimensional presentation of $y(\cdot)$. Is it possible to find $k_1(\cdot),\ldots,k_p(\cdot)$ such that reducing the amount of information $y(\cdot)$ to values of a finite number of functionals $J_i(y(\cdot)) = \langle k_i,y \rangle_{L_2}$ does not imply loss of information about the unknown $x(T)$ in the sense of one-to-one correspondence $y(\cdot) \leftrightarrow (J_1(y(\cdot)),\ldots,J_p(y(\cdot)))$, $x(T) \in U_T$? In case of the positive answer, the problem of "memorizing" $y(\cdot)$ is reduced to integration in real time while measurements $y(t)$ are obtained, which is rather easily technically implemented. Another form of the question: is it possible that the pair (f,g) is observable (a mapping $x(T) \mapsto y(\cdot)$ is injective) while it is impossible to reconstruct uniquely $x(T)$ according to the finite number of projections $\langle k_i,y \rangle$? Here $k_i(\cdot)$ are fixed, where $1 \le i \le p$, the same is for all possible $y(\cdot)$ ($x(T) \in U_T$). If these sets $k_i(\cdot)$ exist, then how should minimal p be possibly chosen? A sequence $x(T) \mapsto y(\cdot) \mapsto \langle k,y \rangle$ defines a function $\varphi(x(T)) = \langle k,y \rangle$. How should analytic description of $\varphi(\cdot)$ be given? In a certain sense a vice-versa question is important. Usually only some of the phase coordinates are measured and it is necessary to reconstruct the others or, more generally, the values of given functions $\varphi(x(T))$. How should $k(\cdot,\cdot)$ be chosen for given $\varphi(\cdot)$ to satisfy (1.3), at least to some necessary precision? In this paper, we give some results in the case of analyticity with respect to phase variables.

DEFINITION 2.1. A function $\varphi : U_T \to \mathbb{R}^1$ is said to be observable in a set $M \subseteq U_T$ if there exists a functional Λ such that $\varphi(x) = \Lambda(y(\cdot; x, T))$, $x \in M$.

Given $y(\cdot)$, the observability of φ in M implies that values of $\varphi(x)$ on an a priori unknown $x = x(T)$ can be uniquely reconstructed if an additional inclusion $x \in M$ is sure. The observability of (f,g) is the same as the observability of all coordinates $\varphi(x) = x_i$ in U_T. If the observability of φ in M is studied

and φ is given only on the subset $\widetilde{M}$ ($M \subseteq \widetilde{M} \subseteq U_T$), then it is defined on $U_T \setminus \widetilde{M}$ arbitrarily. Let $\Phi(M)$ be the set of all functions φ observable in M. Obviously, $\Phi(\widetilde{M}) \subseteq \Phi(M)$ when $M \subseteq \widetilde{M}$.

DEFINITION 2.2. A basis of the set $\Phi(M)$ of all functions observable in M is a finite set of elements $\varphi_i \in \Phi(M)$, $1 \le i \le p$, such that a functional dependence holds:

$$\varphi(x) = F_\varphi(\varphi_1(x),\ldots,\varphi_p(x)) \quad \forall \varphi \in \Phi(M), \ \forall x \in M. \tag{2.1}$$

The set $\Phi(M)$ is a nonlinear (functional) hull of the basis observable functions. After calculating the values $\varphi_1(x),\ldots,\varphi_p(x)$, it becomes impossible to obtain more information about the unknown $x = x(T)$ from $y(\cdot;x,T)$ ($x \in M$). Observability of (f,g) in $M \subseteq U_T$ means one-to-one correspondence $(\varphi_1(x), \ldots,\varphi_p(x)) \leftrightarrow x \in M$. If one of the bases has the last property, then all the others do (in case they exist).

Indeed, let Λ_i be the functionals corresponding to basis functions $\varphi_i \in \Phi(M)$, and $\{k_i, \ i \ge 1\}$ is a full system in $L_2[0,T]$, that is, $\{\langle \phi, k_i \rangle, \ i \ge 1\} \leftrightarrow \phi(\cdot) \in L_2$. Then

$$\psi_i \in \Phi(U_T) \subseteq \Phi(M), \qquad \psi_i(x) = \langle k_i, y(\cdot;x,T) \rangle, \quad x \in U_T. \tag{2.2}$$

By the definition of a basis, $\psi_i(x) = F_i(\varphi_1(x),\ldots,\varphi_p(x))$ for all $i \ge 1$ and for all $x \in M$. Thus, from $\varphi_i(x) = \Lambda_i(y(\cdot;x,T))$, $1 \le i \le p$, the values $\psi_i(x)$, $i \ge 1$, are defined uniquely. The fullness of the system $\{k_i, \ i \ge 1\}$ in L_2 implies $(\varphi_1(x),\ldots,\varphi_p(x)) \leftrightarrow y(\cdot;x,T)$ and instead of functions $y(\cdot;x,T)$ on $[0,T]$ it is possible to use p-dimensional vectors $(\varphi_1(x),\ldots,\varphi_p(x))$, $x \in M$. These $\varphi_i(x)$ form a basis $\Phi(\widetilde{M})$ for all $\widetilde{M} \subseteq M$. Obviously, the functions $\psi(x) = \langle k, y(\cdot;x,T) \rangle$ are observable in any subset U_T, that is, $\psi \in \Phi(M)$ for all $M \subseteq U_T$. If necessary, it is possible to extend a class of possible $k(\cdot)$ to L_2.

THEOREM 2.3. *Let the pair (f,g) be real analytic in U, that is, $f,g \in C^\omega(U, \mathbb{R}^n)$. Then for any set M with compact closure in U_T from any full system in L_2 of admissible weight functions $\{k_i, \ i \ge 1\}$, it is possible to choose $k_{i_\nu}(\cdot)$ such that $\varphi_\nu : U_T \to \mathbb{R}^1$ (where $\varphi_\nu(x) = \langle k_{i_\nu}, y(\cdot;x,T) \rangle$, $1 \le \nu \le p$) form a finite basis $\Phi(M)$.*

PROOF. In the domain $U_T \times U_T$ consider the analytic functions

$$\begin{aligned} \Delta\psi_i(x^1,x^2) &= \psi_i(x^1) - \psi_i(x^2) \\ &= \langle k_i, y(\cdot;x^1,T) - y(\cdot;x^2,T) \rangle, \quad x^j \in U_T. \end{aligned} \tag{2.3}$$

One can assume that $\Delta\psi_i$ are defined and analytic in $W = U_T^c \times U_T^c \subseteq \mathbb{C}^{2n}$, where the domain U_T^c is a sufficiently small neighborhood of U_T in $\mathbb{C}^n$. This extension

can be given by the formula

$$\Delta\psi_i(z^1, z^2) = \langle k_i, y(\cdot; z^1, T) - y(\cdot; z^2, T)\rangle, \quad z^j \in U_T^c. \tag{2.4}$$

The sense of notation $y(\cdot; z, T)$, $z \in U_T^c$, is preserved for the solutions $\dot{x} = f(x)$ can be considered with complex Cauchy conditions $x(T) = z \in U_T^c \subseteq \mathbb{C}^n$. Such solutions can be extended to $[0, T]$ for a sufficiently small neighborhood of U_T.

Let Z_i be a set of zeros of the function $\Delta\psi_i$ in the domain W. Then common zeros of $\Delta\psi_i$ form an analytic subset $Z = \cap_{j=1}^{\infty} Z_j$ of W and numbers $i_1, \ldots, i_p$ exist such that

$$Z \cap (M \times M) = \left(\cap_{\nu=1}^{p} Z_{i_\nu}\right) \cap (M \times M). \tag{2.5}$$

Here we use the terminology and the statement of [3, page 53]. Let Ω be a domain in $\mathbb{C}^n$. The set $A \subset \Omega$ is called a (complex) analytic subset of Ω if for any point $a \in \Omega$, there exist its neighborhood U and holomorphic (analytic) functions $f_1, \ldots, f_N$ such that $A \cap U = \{z \in U \mid f_1(z) = \cdots = f_N(z) = 0\}$. In other words, A is locally a set of common zeros of finite sets of holomorphic functions. If $\{A_\alpha\}_{\alpha \in I}$ is any set of analytic subsets Ω, then $A = \cap_{\alpha \in I} A_\alpha$ is also an analytic subset in Ω, and for any K with compact closure in Ω (i.e., K is bounded and its closure $\mathrm{cl}\, K \subset \Omega$), there exists a finite subset $J \subset I$ such that $A \cap K = (\cap_{\alpha \in J} A_\alpha) \cap K$.

From $\Delta\psi_{i_\nu}(x^1, x^2) = 0$ for $1 \leq \nu \leq p$, $x^j \in M$, it follows that $\Delta\psi_i(x^1, x^2) = 0$, $i \geq 1$, and due to the fullness of $\{k_i, i \geq 1\}$ we obtain $y(\cdot; x^1, T) = y(\cdot; x^2, T)$. This implies that, for all $x \in M$,

$$(\varphi_1(x), \ldots, \varphi_p(x)) = (\langle k_{i_1}, y\rangle, \ldots, \langle k_{i_p}, y\rangle) \longleftrightarrow y(\cdot; x, T),$$
$$\varphi(x) = \Lambda(y(\cdot; x, T)) = F_\varphi(\varphi_1(x), \ldots, \varphi_p(x)) \quad \forall \varphi \in \Phi(M). \tag{2.6}$$

According to Definition 2.2, functions φ_i form a finite basis of $\Phi(M)$. □

REMARK 2.4. The demand of fullness of $\{k_i\}$ in L_2 can be weakened by demanding that $\{\langle k_i, y(\cdot; x, T)\rangle, i \geq 1\} \leftrightarrow y(\cdot; x, T)$, $x \in U_T$ (i.e., fullness on $Y = \{y(\cdot)\}$). It is possible to take into consideration also the limitations of realization $k_i(\cdot)$ (e.g., $|k_i(t)| \leq \bar{k}, \ldots$). The number of basis projections p depends not only on f, g, and M, but also on a choice of the system $\{k_i\}$.

The problem of searching a basis can be posed in algebraic terms. In the ring $C^\omega(W)$ of analytic functions in the domain W consider the ideal generated by the set $\{\Delta\psi_i, i \geq 1\}$. Its elements are finite linear combinations of functions $\Delta\psi_i$ with coefficients from $C^\omega(W)$. The finite basis of this ideal, if it exists, defines a number of basis projections $\langle k_i, y\rangle$ for $M = U_T$ (and thus for all $M \subseteq U_T$). In particular, from the fact that a ring of germs of analytic functions in a point is Noether, it follows (see [5, page 50]) that for a family of holomorphic functions $\mathcal{F}$ in a neighborhood $\mathbb{O}$ of a point a, it is possible

to choose a neighborhood $\widetilde{\mathbb{O}} \subseteq \mathbb{O}$ ($a \in \widetilde{\mathbb{O}}$) and a set $f_1, \ldots, f_N \in \mathscr{F}$ such that for any $f \in \mathscr{F}$, there exist holomorphic functions $h_1, \ldots, h_N$ in $\widetilde{\mathbb{O}}$ such that $f = h_1 f_1 + \cdots + h_N f_N$ in $\widetilde{\mathbb{O}}$. Thus, for all $\bar{x} \in U_T$, there exist a neighborhood

$$P_\varepsilon = \left\{ z \in \mathbb{C}^n \mid \|z - \bar{x}\| = \max_i |z_i - \bar{x}_i| < \varepsilon \right\} \subseteq U_T^c \quad (P_\varepsilon \cap \mathbb{R}^n \subseteq U_T) \qquad (2.7)$$

and a finite number of functions $\Delta \psi_{i_1}, \ldots, \Delta \psi_{i_q}$ such that

$$\Delta \psi_j(z^1, z^2) = \sum_{\nu=1}^{q} \alpha_{j\nu}(z^1, z^2) \Delta \psi_{i_\nu}(z^1, z^2), \quad j \ge 1, \qquad (2.8)$$

when $(z^1, z^2) \in P_\varepsilon \times P_\varepsilon$ and $\alpha_{j\nu} \in C^\omega(P_\varepsilon \times P_\varepsilon)$. Then

$$\begin{aligned} [\Delta \psi_{i_\nu}(x^1, x^2) = 0, \ 1 \le \nu \le q, \ x^j \in M = P_\varepsilon \cap U_T] \\ \Longrightarrow [\Delta \psi_i(x^1, x^2) = 0, \ i \ge 1] \\ \Longrightarrow y(\cdot; x^1, T) = y(\cdot; x^2, T). \end{aligned} \qquad (2.9)$$

The basis $\Phi(M)$ is formed by functions $\varphi_\nu(x) = \psi_{i_\nu}(x) = \langle k_{i_\nu}, y(\cdot; x, T) \rangle$:

$$(\varphi_1(x), \ldots, \varphi_q(x)) \longleftrightarrow y(\cdot; x, T), \quad x \in M = P_\varepsilon \cap U_T. \qquad (2.10)$$

If there are no a priori limitations on the structure of the weight functions $(k(\cdot) \in \{k_i\})$, then Theorem 2.3 can be strengthened as follows ($M = U_T$, $p = 2n + 1$).

THEOREM 2.5. *Let $\{k_i, \ i \ge 1\}$ be any full system of continuous functions on $[0, T]$ in L_2. There exists a family of sets of $2n + 1$ functions $\{r_i(\cdot)\}$, for which $\varphi_i(x) = \langle r_i, y(\cdot; x, T) \rangle$, $0 \le i \le 2n$, form a basis $\Phi(U_T)$. Each $r_j(\cdot)$ is represented by a uniformly convergent series on elements $\{k_i\}$ on $[0, T]$.*

PROOF. The following result from the theory of analytic sets (see [3, page 54]) will be used. Let $\{f_\alpha\}_{\alpha \in I}$ be any family of functions that is holomorphic in a given domain $\Omega \subseteq \mathbb{C}^n$. Then a set of their common zeros Z is an analytic subset in Ω, and sets of $n + 1$ functions $g_i \in C^\omega(\Omega)$ exist such that the common zeros of g_i also coincide with Z.

Here is a modification of this fact according to the concerned problem. Firstly, we will show how g_n should be built. Let Ω_i be connected components of Ω which do not belong to Z, and let $a_i \in \Omega_i \setminus Z$ be arbitrarily chosen points. For any i there exists a function f_{α_i} such that $f_{\alpha_i}(a_i) \ne 0$. We present Ω as a countable union $\bigcup K_j$ of the compact sets (where $K_j \subseteq K_{j+1}$ and for any compact set K there exists s such that $K \subset K_s$). Induction on j can help

to choose numbers c_j such that

$$
\begin{aligned}
&|c_j f_{\alpha_j}(z)| < 2^{-j} \quad \forall z \in K_j, \\
&\left| \sum_{k=1}^{j} c_k f_{\alpha_k}(a_i) \right| > 2^{-1} |c_i f_{\alpha_i}(a_i)| \quad \forall i \le j.
\end{aligned}
\tag{2.11}
$$

A series $\sum c_i f_{\alpha_i}$ converges uniformly on compacts on Ω to a holomorphic function, which will be named g_n. By construction, $g_n(a_i) \ne 0$ for all i, thus $\dim(Z_{g_n} \cap \Omega_i) < n$. Here Z_{g_n} are the zeros of g_n in Ω. Other $g_{n-1}, \dots, g_0$ have been constructed in [3] by induction in a similar way: $g_s|_Z \equiv 0$ and all irreducible components of dimension greater than or equal to s of analytic subset $Z_{g_n} \cap \dots \cap Z_{g_s}$ in Ω belong to Z. The set of common zeros $g_0, \dots, g_n$ coincides with Z.

This result should be applied to the set of analytic functions $\Delta\psi_i : W \to \mathbb{C}$ in the domain $\Omega = W = U_T^c \times U_T^c \subseteq \mathbb{C}^{2n}$, obtained using analytic extension of $\psi_i(x)$ from U_T to a sufficiently small neighborhood $U_T^c \supset U_T$ in $\mathbb{C}^n$: $\psi_i(z) = \langle k_i, y(\cdot; z, T) \rangle$, $z = x(T) \in U_T^c$. The only difference is that the coefficients c_j will be chosen from the following condition: for all $(z^1, z^2) \in K_j$,

$$
\begin{aligned}
|c_j \Delta\psi_{\alpha_j}(z^1, z^2)| &= |\langle c_j k_{\alpha_j}, y(\cdot; z^1, T) - y(\cdot; z^2, T) \rangle| \\
&\le \|c_j k_{\alpha_j}\|_C \cdot \|y(\cdot; z^1, T) - y(\cdot; z^2, T)\|_{L_1} \le 2^{-j}.
\end{aligned}
\tag{2.12}
$$

By doing this, we preserve the second inequality in (2.11) defining c_j. This correction provides convergence of not only the series $\sum c_i \Delta\psi_{\alpha_i}$ in W to an analytic function, but also the series $\sum c_i k_{\alpha_i}$ in $C[0, T]$. We use such constructions inductively and denote the sums of the series in $C[0, T]$ by $r_{2n}, \dots, r_0$. By construction, the set of common zeros of the functions

$$
q_i(z^1, z^2) = \langle r_i, y(\cdot; z^1, T) - y(\cdot; z^2, T) \rangle, \quad 0 \le i \le 2n,
\tag{2.13}
$$

in W coincides with $Z = \cap_{j=1}^{\infty} Z_j$ (where Z_j are zeros of $\Delta\psi_j$ in W). Due to fullness of $\{k_i, \ i \ge 1\}$, any unequal, on $[0, T]$, functions $y(\cdot; x^1, T) \ne y(\cdot; x^2, T)$, $x^j \in U_T$, have a different set of projections:

$$
\{\langle r_i, y(\cdot; x^1, T) \rangle\} \ne \{\langle r_i, y(\cdot; x^2, T) \rangle\}, \quad 0 \le i \le 2n.
\tag{2.14}
$$

From $y(\cdot; x, T) \leftrightarrow (\varphi_0(x), \dots, \varphi_{2n}(x))$, $x \in U_T$, $\varphi_i(x) = \langle r_i, y(\cdot; x, T) \rangle$, it follows that the set φ_i is the basis of $\Phi(U_T)$. There are infinitely many good sets of $\{r_j\}$ as $\{k_i\}$ and coefficients for r_j can be chosen rather arbitrarily. Discontinuous k_i are also possible if convergence of series r_j is studied in L_2. □

Observability of the pair (f, g) in $M \subseteq U_T$ (i.e., $y(\cdot; x, T) \leftrightarrow x \in M$) means that for a full system $\{k_i, \ i \ge 1\}$ in L_2 (or $Y = \{y(\cdot)\}$), the set of common zeros of the functions $\Delta\psi_i(x^1, x^2)$ in $M \times M$ coincides with the diagonal $\{(x, x) \mid x \in M\}$. For the basis $r_j(\cdot)$, the vector $x(T) \in M$ is uniquely determined by the set of

$2n+1$ projections $\mu_i = \langle r_i, y \rangle$. But even if (f, g) is not observable, to find a basis $\Phi(M)$ for given subdomain $M \subseteq U_T$ is important for applied purposes. For instance, to form a control $u(x(t))$ in control systems by the history of measurements on $[t - T, t]$, there is no need to reconstruct a full phase vector $x(t)$, but it is sufficient to observe directly the function $u(x)$. If any of the basis weight functions $k(\cdot)$ does not satisfy given limitations of realization $|k(t)| \leq \overline{k} = \text{const}$, then it is possible to replace it with $\alpha k(\cdot)$ with sufficiently small multiplier α. Such operations do not influence a basis, important here are "projections to directions."

We discuss some generalizations. In case of multidimensional measurements $g : U \to \mathbb{R}^m$, $m \geq 1$, Theorems 2.3 and 2.5 remain mostly unchanged. They can be reformulated in terms of the predicting problem: given $y(\cdot)$ ($t \in \Theta$), find $x = x(T) \in U_T$ (or $\varphi(x(T))$). In stationary systems due to the uniqueness theorem for analytic functions $y|_{[0,\vartheta]} \leftrightarrow y|_{[0,T]}$, predictability is the same as observability (calculation details are not considered). In the nonstationary case $f = f(t, x)$, $g = g(t, x)$, when smoothness demands with respect to t can be significantly weakened, predictability implies observability, but not vice versa. For possible weight functions in the integral operators (1.3), one must define $k(t, y) = 0$, $t > \vartheta$. In the proofs of Theorems 2.3 and 2.5, actually, only real analyticity of the functions $\langle k, y(\cdot; x, T) \rangle$ with respect to Cauchy data $x = x(T) \in U_T$ was used. For convenience, let $\vartheta \leq T$, and then observability is a special case of predictability ($\vartheta = T$).

Now we consider the nonstationary observability system

$$\dot{x} = f(t, x), \qquad y = g(t, x), \tag{2.15}$$

posed in the domain $\Omega = (t_1, t_2) \times U$, $[0, T] \subset (t_1, t_2)$. Vector functions f and g are continuous in Ω and real analytic with respect to x in U for all fixed $t \in (t_1, t_2)$. Moreover, the following conditions hold:

$$f = f^c|_\Omega, \qquad g = g^c|_\Omega, \qquad f^c(t, \cdot) \in C^\omega(U^c), \qquad g^c(t, \cdot) \in C^\omega(U^c),$$
$$f^c \in C((t_1, t_2) \times U^c, \mathbb{C}^n), \qquad g^c \in C((t_1, t_2) \times U^c, \mathbb{C}^m). \tag{2.16}$$

A domain U^c is a neighborhood of U in $\mathbb{C}^n$. It guarantees existence and uniqueness of a Cauchy problem and analyticity with respect to initial data [4]. Suppose that for the subdomain $U_T \subseteq U$, the solutions $x(\cdot; x, T)$, $x = x(T) \in U_T$, can be extended to $[0, T]$. Then, for all $k(\cdot) \in L_2^m = L_2^m[0, T]$, the function $\psi(x) = \langle k, y(\cdot; x, T) \rangle$ is real analytic in U_T (and analytic in the domain U_T^c—a sufficiently small neighborhood of U_T in $\mathbb{C}^n$). The last property can serve as the initial assumption.

Let $\Phi_*(M)$ be a set of predictable functions in $M \subseteq U_T$:

$$\varphi \in \Phi_*(M) \iff \varphi(x) = \Lambda(y(\cdot; x, T)), \quad x \in M, \ y : \Theta \longrightarrow \mathbb{R}^m. \tag{2.17}$$

Definition 2.2 of a basis remains unchanged. In the argumentation after it let $y(\cdot) \in Y_* = \{y : \Theta \to \mathbb{R}^m \mid x(T) \in U_T\}$. A class of possible weight functions $k(\cdot)$ is piecewise continuous and equals zero on $(\vartheta, T]$.

THEOREM 2.6. *For the nonstationary pair (f, g) and any M with compact closure in U_T, it is possible to choose $k_{i_v}(\cdot)$ from any full system of admissible vector functions $\{k_i, \ i \geq 1\}$ in $L_2^m(\Theta)$ (on Y_*) such that the functions $\varphi_v(x) = \langle k_{i_v}, y(\cdot; x, T) \rangle$ (where $x \in U_T$, $1 \leq v \leq p$) form the basis of the set $\Phi_*(M)$. Without the restriction $k(\cdot) \in \{k_i\}$, it is possible to make $M = U_T$, $p = 2n + 1$.*

In Theorems 2.3, 2.5, and 2.6 it is possible to assume that $M = U_T$ if the domain U_T is known to be bounded and the solutions with $x(T) \in \hat{U}$ can be extended to the time interval $[0, T]$. Subdomain $\hat{U} \subseteq U$ contains the closure clU_T.

These results can be interpreted as a finite-dimensional presentation of Y (Y_*) in $\mathbb{R}^p$. Here one should note that the embedding operator has a special form (projections of $y(\cdot)$ to L_2^m), and local nonsingularity is not demanded (local observability in U_T by linear approximation). In other words, "Jacobian may be singular." In case (2.15), the output $y(\cdot)$ can be a nondifferentiable function of t.

3. Duality control problem. The results of Section 2 ensure that the integral operators (1.3) provide complete information. Instead of operator links $\varphi(\cdot)$ and $k(\cdot, \cdot)$ a convenient constructive equation is necessary. Here are some well-known results of a linear theory. They are presented to make the technique used later more easily understood.

3.1. Necessary results of the linear theory. Let $\dot{x}(t) = F(t)x(t)$ and $y(t) = G(t)x(t)$. The elements of the matrices F and G are continuous on $[0, T]$. The determination of $x(T)$ is the same as the calculation of projections $h'x(T)$ on a basis $\mathbb{R}^n$. Besides, the complete phase vector is not always necessary; for instance, some elements of it are measured. Thus, for a fixed vector $h \in \mathbb{R}^n$, we consider a problem of constructing an operator for reconstructing a projection: $h'x_T = I(y(\cdot; T, x_T))$. If the operator I exists, then a projection $\varphi(x) = h'x$ is called *observable* or *predictable* (when $y(t)$ is known on $\Theta = [0, \vartheta]$, $\vartheta < T$). We denote by $\mathcal{Y} = \{y(\cdot; T, x_T) \mid x_T \in \mathbb{R}^n\}$ a set of possible measurements:

$$y(t) = \Psi(t)x_T = G(t)\Phi(t)\Phi^{-1}(T)x_T, \qquad \mathcal{Y} \subset L_2^m[0, T], \quad \dim\mathcal{Y} \leq n. \quad (3.1)$$

Here $\dot{\Phi} = F\Phi$ and $\Phi(0) = E_n$. A canonical basis $\mathbb{R}^n$ is linearly mapped by $L : x_T \mapsto y(\cdot)$ to the columns of a matrix $\Psi(\cdot)$. If these columns are linearly independent on $[0, T]$ (on Θ), then $\dim\mathcal{Y} = \dim\mathcal{L}(\Psi(\cdot)) = n$ and a linear mapping L is invertible—a full observability (predictability) holds. A symbol $\mathcal{L}$ means a linear span of columns. A vector is uniquely defined by the projections on basis

elements:

$$y \longmapsto y(\cdot), \quad y = \int_0^T \Psi'(\tau)y(\tau)d\tau = \Gamma x_T, \quad \Gamma = \langle \Psi, \Psi \rangle. \qquad (3.2)$$

A functional which reconstructs a projection $h'x_T$ by $y(\cdot)$ exists only if $h \in \mathcal{L}(\Gamma)$. If $h'x_T = I(y(\cdot))$ holds for all $x_T \in \mathbb{R}^n$, then I is linear on $\mathcal{Y}$ and it can be represented as a scalar product $\langle k, y \rangle_{L_2}$. All what can be obtained about x_T from $y(\cdot)$ are functions of projections $h'x_T$ to columns of Gram matrix Γ. In particular, it is possible to replace a numerically ill-posed operation of differentiation $\dot{y}(s)$ by integral operator on $\mathcal{Y}$. If the operator $\varphi(x_T) = h'x_T = I(y(\cdot))$ exists, then it is better from the calculating point of view to construct it in the form

$$h'x_T = \int_0^T k'(\tau)y(\tau)d\tau \quad \forall x_T \in \mathbb{R}^n,$$
$$\Longrightarrow h = \Phi^{-1'}(T)\int_o^T \Phi'(\tau)G'(\tau)k(\tau)d\tau. \qquad (3.3)$$

Equation (3.3) is functional with respect to $k(\cdot)$. We reduce the problem to studying a "pointwise" equation. For that we define a vector function

$$V(t) = \Phi^{-1'}(t)\int_0^t \Phi'(\tau)G'(\tau)k(\tau)d\tau, \quad \vartheta < T,$$
$$\Longrightarrow k(\tau) = 0, \quad \tau \in (\vartheta, T]. \qquad (3.4)$$

The boundary conditions are

$$V(0) = 0, \qquad V(T) = h. \qquad (3.5)$$

To obtain an equation, we differentiate (3.4) noting that $\Phi^{-1}\Phi = E_n$ and thus $d\Phi^{-1}/dt = -\Phi^{-1}A$:

$$\dot{V}(t) = -F'(t)V(t) + G'(t)k(t). \qquad (3.6)$$

Equation (3.6) is understood as a duality control system. It is necessary to transfer a phase state V from zero to h by time T. In case of a predicting problem $k(\tau) = 0$, $\tau > \vartheta$. In terms of control theory, a set of all h, for which the projections $h'x_T$ are observable (predictable), coincides with a set of attainability $\mathcal{D}_T = \{V(T)\}$.

REMARK 3.1. Define an operator $\mathcal{A} : \mathbb{R}^n \to \mathcal{Y}$ as

$$\mathcal{A}x_T = y(\cdot; T, x_T) \Longrightarrow \langle k, y \rangle_{L_2} = \langle k, \mathcal{A}x_T \rangle_{L_2} = \langle h, x_T \rangle_{\mathbb{R}^n}, \quad h = V(T). \quad (3.7)$$

As embedding $y(\cdot) \in \mathcal{Y}$ holds, it is sufficient to consider only $k(\cdot) \in \mathcal{Y}$. By definition, $V(T) = \mathcal{A}^* k$. Here $\mathcal{A}^* : \mathcal{Y} \to \mathbb{R}^n$ is an adjoint operator. It can be said that $\mathcal{A} : \mathbb{R}^n \to L_2^m$ and $\mathcal{A}^* : L_2^m \to \mathbb{R}^n$. To determine $\mathcal{A}^* k$, it is necessary to integrate a dual system (3.6) on $[0, T]$ or use the integral representation (3.3). Image $\mathcal{A}^* L_2^m = \mathcal{A}^* \mathcal{Y}$ coincides with $\mathcal{L}(\Gamma)$ (with a linear span of the columns of matrix $\Gamma = \langle \Psi, \Psi \rangle$). The matrix γ is a matrix of the operator $\mathcal{A}^* \mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$: $h = \mathcal{A}^* \mathcal{A} x_T = \Gamma x_T$.

3.2. Nonlinear case. In the domain $\Omega = (t_1, t_2) \times U$ we consider the control system (2.15) and firstly weaken demands: $f, g, f_x, g_x \in C(\Omega)$. As the interval $[0, T]$ and $U_T = \{x(T)\}$ are fixed, all further constructions can be made on sheafs

$$
\begin{aligned}
W &= \{(t, x) \mid t \in [0, T],\ x \in x(t; U_T, T)\}, \\
W_g &= \{(t, y(t)) \mid t \in [0, T]\}.
\end{aligned}
\tag{3.8}
$$

In the domain $Q \supset W_g$ we choose a weight function for measurements processing $k(\cdot, \cdot) \colon Q = Q(k) \subset \mathbb{R}^{m+1}$, $k, k_y \in C(Q)$. Then the function

$$
v(t, x) = \int_0^t k(\tau, y(\tau; x, t)) d\tau
\tag{3.9}
$$

will be in class $C^1(W)$ because the solutions of differential equations are smooth with respect to the initial data $(t_0, x_0) = (t, x) \in W$.

The notation $v \in C^1(W)$ means that $v(\cdot, \cdot)$ can be extended to some domain $\widetilde{W} \supset W$ and $v \in C^1(\widetilde{W})$. It is possible to choose $\widetilde{W}$ as a union of integral curves $(t, x(t))$ corresponding to nonextendable solutions $x(\cdot; x(T), T)$ $(x(T) \in U_T)$, while $(t, y(t))$ are in Q. Thus even if somewhere below the values of t are limited by the observation interval $[0, T]$, derivative symbols v_t in the points $(0, x)$, (T, x) should not cause problems. They could be considered as one-side derivatives.

THEOREM 3.2. *The function v is a unique smooth solution of the linear equation*

$$
v_t(t, x) + v_x(t, x) \cdot f(t, x) = k(t, g(t, x)), \quad (t, x) \in W,
\tag{3.10}
$$

with zero initial conditions $v(0, x) = 0$, $x \in U_0 = x(0; U_T, T)$.

PROOF. Consider any point $(t, x) \in W$ and the solution $x(\tau)$ with initial data $x(t) = x$ ($\tau \in (-\varepsilon, t + \varepsilon)$, $\varepsilon = \varepsilon(t, x) > 0$). Calculate a time derivative of v on the solution (here D is a derivative symbol)

$$
L_f v(t, x) = D_\tau v(\tau, x(\tau))\big|_{\tau = t} = v_t(t, x) + v_x(t, x) \cdot f(t, x).
\tag{3.11}
$$

On the other hand (having in mind a specific representation of (3.9)),

$$
\begin{aligned}
v(\tau, x(\tau)) &= \int_0^\tau k(s, y(s; x(\tau), \tau)) ds \\
&= \int_0^\tau k(s, y(s; x(0), 0)) ds \\
&= \int_0^\tau k(s, y(s; x(t), t)) ds,
\end{aligned}
\tag{3.12}
$$
$$
D_\tau v(\tau, x(\tau))\big|_{\tau=t} = k(t, y(t; x, t)) = k(t, g(t, x)).
$$

UNIQUENESS. For the difference $\bar{v}$ of two solutions, we obtain equation (3.10) with zero right-hand side (first integral equation). Therefore, $\bar{v}(t, x(t)) \equiv$ const. The set W consists of integral curves on $[0, T]$, $\bar{v}(0, \cdot) = 0$. Thus $\bar{v}(t, x) = 0$, $(t, x) \in W$. $\qquad\square$

The meaning of the function (3.9) and equation (3.10) is the following. If $t = T$ in definition (3.9), then the integral operator is obtained in the right-hand side of (1.3). If the values of $\varphi(x(T))$ should be reconstructed, then a condition $v(T, x) = \varphi(x)$, $x \in U_T$, is added to the zero initial data for (3.10). Thus (3.10) can be understood as an equation of phase state transfer $v(t, \cdot)$ from zero to φ by time T. If $k(\cdot, \cdot)$ (where $k, k_y \in C(Q)$) brings solution to problem (1.3), then for $v(\cdot, \cdot)$, equation (3.10) holds together with boundary conditions $v(0, x) = 0$, $x \in U_0 = x(0; U_T, T)$, $v(T, x) = \varphi(x)$, $x \in U_T$. Inversely, if a choice of the function $k(\cdot, \cdot)$ brings the solution of boundary value problem, then, substituting x by any solution $x(t; x(T), T)$, $x(T) \in U_T$, in (3.10) and integrating the obtained identity on t on $[0, T]$ (at the left-hand side there is $\dot{v}(t, x(t))$), we obtain (1.3). Thus, the problem of constructing integral operators (1.3) for reconstructing the values of $\varphi(x(T))$ is the same as a boundary value problem $v(0, \cdot) = 0$, $v(T, \cdot) = \varphi$. Equation (3.10) is linear with respect to a pair (k, v). Here, an analogy is appropriate with methods of Lyapunov functions and Bellman's dynamical programming.

REMARK 3.3. Equation (3.10) can be sufficiently considered only on the set W, consisting of a union of all possible integral curves $(x(T) \in U_T)$. But with rather adequate modeling the extension problems occur rarely; the solutions $x(\cdot; x, t)$ with initial data $(t, x) \in [0, T] \times \tilde{U}$ $(U_T \subseteq \tilde{U} \subseteq U)$ are extendable to $[0, t]$ and possible phase curves with $x(T) \in U_T$ do not leave a known domain $\tilde{U}$. Then due to definition (3.9), $v \in C^1([0, T] \times \tilde{U})$ and (3.10) can be considered on a set of the form of direct product $[0, T] \times \tilde{U}$, which is more convenient. The definitional domain Q of $k(\cdot, \cdot)$ must contain $\{(t, g(t, x)) \mid t \in [0, T], \ x \in \tilde{U}\}$. For the predicting problem $k(\cdot, \cdot)$ (where $k, k_y \in C(Q)$), we reduce $k(t, \cdot) = 0$, $t > \vartheta$. Moreover, for linear $k(t, y) = k'(t) y$, a finite number of discontinuities of the first kind of the vector function $k(t)$ are possible. Continuity of $v(\cdot, \cdot)$ in (3.9) will remain, and (3.10) can be sufficiently considered outside the finite number of sections $t = \vartheta$, $t = t_j$. Similar specifications for possible $k(\cdot, \cdot)$ will be omitted.

We consider in more detail an interpretation of (3.10) as the linear control system. For that, it is convenient to consider the operator form

$$\dot{V}(t) = -A(t)V(t) + B(t)K(t), \quad V(0) = 0, \tag{3.13}$$

$$V(t) = v(t, \cdot) : x(t; U_T, T) \longrightarrow \mathbb{R}^1, \quad \dot{V}(t) = v_t(t, \cdot), \tag{3.14}$$

$$A(t)V(t) = v_x(t, \cdot)f(t, \cdot), \quad K(t) = k(t, \cdot), \quad B(t)K(t) = k(t, g(t, \cdot)). \tag{3.15}$$

If there are no extension problems (see Remark 3.3), then (3.13) is a linear control system in a "standard form" in a phase state $C^1(\tilde{U})$. Otherwise, the definitional domain of a phase vector $v(t, \cdot)$ (as a function of x) can be changed when $t \in [0, T]$. To satisfy (1.3) choosing $K(\cdot)$, one should solve the problem $V(T) = \varphi$ $(x \in U_T, x \in \tilde{U})$. Thus, we are interested in finding a set of attainability from zero point $\mathcal{D}_T = \{V(T) = v(T, \cdot)\} \subseteq C^1(U_T)$. By construction, $\mathcal{D}_T \subset \Phi(U_T)$. The restriction $|k(t, y)| \le \overline{k}$ is not considered. In applications, the phase orbits are usually situated in a limited domain and, if necessary, it is possible to use αk with a small constant α, dividing later the obtained result of integration in (1.3) by α.

The control of systems with infinite-dimensional phase space is a popular subject in general system theory and functional analysis. In this particular case there is the following specificity. It is unnecessary to study full controllability (or ε-controllability) of the infinite-dimensional system (3.13). To observe a pair (f, g), it is sufficient to have $w_i : U_T \to \mathbb{R}^1$ in the set of attainability $\mathcal{D}_T$ such that $(w_1(x), \ldots, w_p(x)) \leftrightarrow x \in U_T$, $x = H(w_1(x), \ldots, w_p(x))$. If it is necessary to determine only the values of the given function $\varphi : U_T \to \mathbb{R}^1$ (where $\varphi(x) = \Lambda(y(\cdot; x, T))$), then it is sufficient to ascertain either inclusion $\varphi \in \mathcal{D}_T$ or dependence $\varphi = H_\varphi(w_1, \ldots, w_p)$ in a domain U_T and use the operator $\varphi(x(T)) = H_\varphi(\mu_1, \ldots, \mu_p)$. Here μ_i are calculated as integrals of functions $k_i(\tau, y(\tau))$ on $\tau \in [0, T]$.

In a linear case $f = F(t)x$, $y = G(t)x$, and $k = k'(t)y$, the observable projections $h'x(T)$ (where $U_T = \mathbb{R}^n$) are described by the set $\{h = V(T)\}$ due to (3.6), that is, the linear span $\mathcal{L}$ of basis $V_i(T)$, $1 \le i \le p$, $p \le n$. Controllability (3.6) means that $\mathcal{L} = \mathbb{R}^n$, that is, $\{V_i'(T)x, \ 1 \le i \le n\} \leftrightarrow x \in U_T(\mathbb{R}^n)$, $p = n$. For (3.13) as a dual system to (f, g), we are interested in "nonlinear projections" $\varphi(x(T)) = v(T, x(T))$ and a "fullness" of not linear, but functional hull of sets of elements of the attainability set. As a linear space, the set $\mathcal{D}_T$ is infinite dimension-al with the exception of singular cases (e.g., if f is linear, g is polynomial). It can be proved that even if linear $k(t, y) = k'(t)y$ are considered, $\mathcal{D}_T$ is finite dimensional only in the case when the linear hull $\mathcal{L}\{y(\cdot; x, T) \mid x \in U_T\}$ is finite dimensional (and then $\dim \mathcal{D}_T = \dim \mathcal{L}$). These facts approve using a functional hull for a nonlinear observation problem (yet linear control problem (3.13)) instead of a linear one.

DEFINITION 3.4. A basis in $M \subseteq U_T$ of the set of attainability of the dual control system $\mathcal{D}_T = \{V(T) = v(T, \cdot) : U_T \to \mathbb{R}\}$ is a finite set $w_i \in \mathcal{D}_T$ for which

$w(x) = H_w(w_1(x),\ldots,w_p(x))$ for all $x \in M$ and for all $w \in \mathscr{D}_T$. The system
(3.10) (or (3.13)) is considered controllable in a subset $M \subseteq U_T$ if a basis w_i
in M exists and $(w_1(x),\ldots,w_p(x)) \leftrightarrow x \in M$ (i.e., a functional hull of basis w_i
coincides with a whole space of functions of $x \in M$).

Controllability in $M \subseteq U_T$ does not depend on a basis $\mathscr{D}_T$ in M. For the full
system $\{k_i,\ i \geq 1\}$ in $L_2^m[0,T]$ of piecewise continuous vector functions, we
define elements $v_i(T,\cdot) \in \mathscr{D}_T$, $v_i(T,x) = \langle k_i, y(\cdot;x,T) \rangle$. "Fourier coefficients"
$\langle k_i, y \rangle$ are uniquely determined on $w_\nu(x)$ in M: $v_i(T,x) = H_i(w_1(x),\ldots,$
$w_p(x))$. Thus, $(w_1(x),\ldots,w_p(x)) \leftrightarrow y(\cdot;x,T)$, $x \in M$.

Thus we come to generalization of duality in control and observation the-
ory to a nonlinear case. We consider at once a nonstationary case (2.15) and
a predicting problem (observation problem is corresponded by $\vartheta = T$). Con-
sider the suggestions pointed out above about real analyticity with respect to
phase variables in (2.15). Then, for all M with a compact closure in U_T, it is
possible to select elements $k_{i_\nu}(\cdot)$, $1 \leq \nu \leq p$, from any full system in $L_2^m(\Theta)$
of piecewise continuous vector functions $\{k_i,\ i \geq 1\}$ such that the functions
$w_\nu(x) = \langle k_{i_\nu}, y(\cdot;x,T) \rangle = v_{i_\nu}(T,x)$, $x \in U_T$, form a basis in M of attainability
set $\mathscr{D}_T$. Provided that the conditions mentioned after Theorem 2.6 hold, this
is also true for $M = U_T$. If $\vartheta < T$, then $k_{i_\nu}(t) = 0$, $k(t,\cdot) = 0$, $t > \vartheta$ is always
assumed. Without an a priori restriction $k(\cdot) \in \{k_i,\ i \geq 1\}$ on the time interval
$\Theta = [0,\vartheta]$, it is possible to choose a family of sets $\{k_j(\cdot)\}$ (where $1 \leq j \leq p$,
$p = 2n + 1$) from admissible $k(\cdot)$ such that the corresponding $w_j = v_j(T,\cdot)$
form a basis of $\mathscr{D}_T$ in $M = U_T$. In a brief form, the following result is obtained.

Theorem 3.5. *The set $\Phi_*(M)$ of predictable functions in M is described as
a functional hull $\mathscr{H}(M) = \{H(w_1,\ldots,w_p)\}$ of some basis w_i in M of the attain-
ability set $\mathscr{D}_T$. In particular, only linear weight functions $k(t,y) = k'(t)y$ can
be considered and assume that $M = U_T$, $p = 2n + 1$. Contraction of elements
$\mathscr{H}(U_T)$ to M form $\Phi_*(M)$. A pair (f,g) is predictable (observable when $\vartheta = T$)
in $M \subseteq U_T$ if and only if a dual system (3.10) is controllable in M.*

In a stationary linear case, the set of attainability $\mathscr{D}_T$ is conveniently de-
scribed as a linear hull of columns of the controllability matrix $\mathscr{K} = (G',\ldots,$
$F'^{n-1}G')$. For the infinite-dimensional dual system (3.13), things are more com-
plex. Let, for simplicity, $f = f(x)$, $g = g(x)$, $m = 1$, $\vartheta = T$, $k(t,y) = k(t)y$.
Consider consistent derivatives $L_f^0 g(x) = g(x)$, $L_f^{i+1} g(x) = (L_f^i g(x))_x \cdot f(x)$,
$x \in U_T$ ($x \in U$). Derivatives of the output $y^{(i)}(t)$ equal $L_f^i g(x(t))$. Theoret-
ically, it is convenient to study the system of n equations $L_f^i g(x) = y^{(i)}(T)$
from the point of view of solvability with respect to $x \in U_T$. Use of injection
criteria for mappings from $M \subseteq \mathbb{R}^n$ to $\mathbb{R}^n$ gives sufficient conditions of observ-
ability in M. But consequent differentiation is hardly possible in practice. In
this sense, operators (1.3) are correct; each integration is done independently
of another and the measurements are smoothed. In applications, usually the

components of the vector functions f and g are superpositions of elementary functions—then $L_f^i g$ are as well. In operator terms (3.13) the derivatives $L_f^i g = A^i B$ (where $A = D_x(\cdot)f$, $B = g$, $BK(t) = k(t)g$) are analogues of the columns of controllability matrix.

REMARK 3.6. Using Taylor's formula, with sufficiently small T, we obtain the representation of the elements of $\mathcal{D}_T$ as a series: $L_i = L_f^i g$, $v(T,x) = c_0 + c_1 L_1(x) + c_2 L_2(x) + \cdots$, $c_i = \langle k, (\tau - T)^i \rangle / i!$. Here arises a moment power problem. Thus in a general case, basis elements $A^i B = L_f^i g$ (columns of controllability matrix) do not themselves belong to $\mathcal{D}_T$. Differentiation of the output cannot be replaced by integral operators. The finite decomposition $v(T,x)$ on $L_i(x)$ is possible, but coefficients will be functions of x. A set $\mathcal{D}_T$ depends on T, and a restriction $k(t) = 0$, $t > \vartheta$, for predicting problem, reduces $\mathcal{D}_T$ (unlike the linear model).

3.3. Ideal observability. We briefly cover the observability problem of the perturbed systems. Consider a dual system (3.10) on the set $[0,T] \times \widetilde{U}$. The solutions $x(\cdot;x,t)$ with initial data $(t,x) \in [0,T] \times \widetilde{U}$ can be extended to $[0,t]$, and possible phase curves ($x(T) \in U_T \subseteq \widetilde{U}$) do not leave the known domain $\widetilde{U}$ (Remark 3.3). Now let the movement equations in the model (2.15) be perturbed:

$$\dot{x} = f(t,x) + \sum_{i=1}^{r} \xi_i(t) h^i(t,x), \quad h^i, h_x^i \in C(\Omega). \tag{3.16}$$

Functions $\xi_i(t)$ are unknown and piecewise continuous, $|\xi_i(t)| \leq \bar{\xi} = \text{const}$. Assume, choosing a possible function $k(\cdot,\cdot)$ in (3.10), that problem (1.3) is solved ($v(T,x) = \varphi(x)$, $x \in U_T(\widetilde{U})$) and additionally that the following holds:

$$v_x(t,x) \cdot h^i(t,x) = 0, \quad (t,x) \in [0,T] \times \widetilde{U}, \ 1 \leq i \leq r. \tag{3.17}$$

Then in the right-hand side of the dual system (3.10), it is possible to replace f by the right-hand side of (3.16). The formal result obtained will be named $(3.10)_\xi$. Consider any perturbed solution $x(\cdot;x(T),T,\xi)$ with initial data $x(T) \in U_T$, defined on $[0,T]$ and with a phase orbit in the set $\widetilde{U}$. Substitute it instead of x in $(3.10)_\xi$ and integrate this identity on $t \in [0,T]$:

$$\varphi(x(T)) = \int_0^T k(\tau, y(\tau; x(T), T, \xi)) d\tau, \quad x(T) \in U_T, \ x(\tau) \in \widetilde{U}. \tag{3.18}$$

In analogy with a linear case, such operator will be called an ideal observability operator; a weight function $k(\cdot,\cdot)$ does not depend on the realization of perturbations $\xi = (\xi_1, \ldots, \xi_r)'$. They are taken into consideration indirectly by

means of measurements $y(t) = g(t,x(t))$. It is possible to estimate the component $\varphi(x(T))$ of the phase state when perturbations are unknown and with given uncertainty $x(T) \in U_T$. In the conditions of extendability on t this can be done periodically.

In the operator form (3.13), condition (3.17) has a form of linear phase restrictions: $P(t)V(t) = 0 : v_x(t,\cdot)H(t,\cdot) = 0$, $H = (h^1,\ldots,h^r)$.

In case the directions of perturbations h^i are fixed, we obtain the problem of controlling not only a finite state $v(T,\cdot) = \varphi$ but also a gradient $v_x(t,\cdot)$. Provided that PV is small in the appropriate norm, one obtains integral observability operator that is stable to perturbations. Conversely, it is possible to solve unperturbed problem and then condition (3.17) gives the description of all invariant directions h^i, which is an important property of the pair (φ,k).

4. Approximation schemes. Firstly, we consider a power series technique. In the domain $\Omega = (t_1,t_2)\times U$ consider the observation model (2.15) with posed suppositions of analyticity on phase variables. Series will be constructed with respect to deviations from some basis movement, which is considered zero (this can be obtained by change of variables): $f(t,0) = 0$, $g(t,0) = 0$, $0 \in U_T \subseteq U$. Firstly, consider only linear admissible weight functions $k(t,y) = k'(t)y$. In some neighborhood Q of zero when $t \in [0,T]$, the functions $v(t,x)$, $f(t,x)$, and $g(t,x)$ are expanded in power series of x. The coefficients are continuous on $[0,T]$, smoothness of coefficients for $v(t,x)$ may vanish only in discontinuities of the first kind of $k(t)$. Convergence is uniform with respect to $t \in [0,T]$. Equating homogeneous polynomials of x of the same degree (upper index $p \geq 1$) at both parts in the dual system (3.10), we get

$$v_t^{(p)}(t,x) + \sum_{i=1}^{p} v_x^{(i)}(t,x) \cdot f^{(p-i+1)}(t,x) = k'(t)g^{(p)}(t,x), \qquad (4.1)$$

where $v^{(p)}(0,x) = 0$, $x \in Q$, $t \in [0,T]$. Homogeneous polynomial $w^{(p)}(\cdot)$ of degree p is corresponded by a unique symmetrical p-linear form $w^{(p)}(\cdot,\ldots,\cdot)$ from the condition $w^{(p)}(x) \equiv w^{(p)}(x,\ldots,x)$, $x \in \mathbb{R}^n$ [1, 2]. For instance, polynomial $v^{(2)}(x) = x'Px$ is corresponded by the bilinear form $v^{(2)}(x,z) = v^{(2)}(z,x) = x'(P+P')z/2$. In terms of symmetrical polylinear forms (t is considered as a parameter),

$$\begin{aligned} v_t^{(p)}(t,x,\ldots,x) + \sum_{i=1}^{p} \sum_{j=1}^{i} v^{(i)}(t,x,\ldots,f^{(p-i+1)}(t,x,\ldots,x),\ldots,x) \\ = k'(t)g^{(p)}(t,x,\ldots,x). \end{aligned} \qquad (4.2)$$

Due to symmetry, it is possible to equate the coefficients at the same lexicographically ordered monomials $x_{i_1} \cdot \cdots \cdot x_{i_p}$:

$$\dot{V}^{(p)} + \sum_{i=1}^{p} \left(\sum_{j=1}^{i} E \otimes \cdots \otimes F^{(p-i+1)'} \otimes \cdots \otimes E \right) \cdot V^{(i)} = G^{(p)'} k. \qquad (4.3)$$

The index j means an ordinal of $F^{(p-i+1)'}(t)$ in a sequence of i products $\otimes$. A symbol $\otimes$ means direct (tensor) matrix product [10]. In $A \otimes B$, instead of elements a_{ij} of a matrix A there are blocks $a_{ij}B$. The set of indexes $(i_1,\ldots,i_p)$ is lexicographically before $(j_1,\ldots,j_p)$ if the first difference of nonzero differences $j_1 - i_1,\ldots,j_p - i_p$ is greater than zero. The name E is for unit matrix $n \times n$, $X^{(1)} = x$, $X^{(s)} = x \otimes \cdots \otimes x$,

$$\begin{aligned}
V^{(s)'}(t)X^{(s)} &= v^{(s)}(t,x,\ldots,x) = v^{(s)}(t,x),\\
G^{(s)}(t)X^{(s)} &= g^{(s)}(t,x,\ldots,x) = g^{(s)}(t,x),\\
F^{(s)}(t)X^{(s)} &= f^{(s)}(t,x,\ldots,x) = f^{(s)}(t,x).
\end{aligned} \qquad (4.4)$$

For a united vector $V = (V^{(1)'},\ldots,V^{(p)'},\ldots)'$, one obtains

$$\dot{V}(t) = -\mathscr{F}'(t)V(t) + \mathscr{G}'(t)k(t), \quad V(0) = 0. \qquad (4.5)$$

Here $\mathscr{G} = (G^{(1)}, G^{(2)},\ldots)$, the first n rows of upper block-triangular matrix $\mathscr{F}(t)$ are equal to $(F^{(1)}, F^{(2)},\ldots)$ and next n^2 rows are equal to

$$\left(\mathbf{0}, F^{(1)} \otimes E + E \otimes F^{(1)},\ldots,F^{(p)} \otimes E + E \otimes F^{(p)},\ldots \right) \qquad (4.6)$$

and so on.

Note that in such notation (f,g) will be written as

$$\dot{X} = \mathscr{F}(t)X, \quad y = \mathscr{G}(t)X, \qquad X = (x', X^{(2)'},\ldots)'. \qquad (4.7)$$

Multiplying (4.5) by X in a scalar sense, we obtain a dual control system: $v(t,x) = V'(t)X$. Here is a complete analogy with (1.1), (1.2) in the linear case, but in the obtained "coordinate" representation (4.5), (4.7), the matrices $\mathscr{F}$ and $\mathscr{G}$ are infinite. In a stationary case (when observing in a neighborhood of equilibrium), blocks $\mathscr{F}'^i \mathscr{G}'$ of a controllability matrix are the coefficients of the consequent derivatives $L_f^i g(x) = \mathscr{G}\mathscr{F}^i X$.

Nonlinear functions $k(\cdot,\cdot)$ will be limited to being real analytic, $k(t,0) = 0$. It is sufficient to be continuous and analytic with respect to y in a neighborhood of zero (where continuity with respect to (t,y) is preserved in $(t_1,t_2) \times P$ for $(t_1,t_2) \supset [0,T]$ and in a neighborhood P of a zero in $\mathbb{C}^m$). Then after equating

homogeneous polynomials and transfering to symmetrical forms, one obtains

$$
\begin{aligned}
\dot{V}^{(p)}(t) &+ \left(\sum_{j=1}^{p} E \otimes \cdots \otimes F^{(1)'}(t) \otimes \cdots \otimes E \right) \cdot V^{(p)}(t) \\
&+ \sum_{q=1}^{p-1} \left(\sum_{j=1}^{q} E \otimes \cdots \otimes F^{(p-q+1)'} \otimes \cdots \otimes E \right) \cdot V^{(q)}(t) \\
&- G^{(p)'} \cdot k^{(1)}(t) - \left(G^{(2)'} \otimes G^{(1)'} + G^{(1)'} \otimes G^{(2)'} \right) \cdot K^{(2)}(t) - \cdots \\
&- \sum G^{(i_1)'} \otimes \cdots \otimes G^{(i_{p-1})'} \cdot K^{(p-1)} \\
&= G^{(1)'} \otimes \cdots \otimes G^{(1)'} \cdot K^{(p)}.
\end{aligned}
\tag{4.8}
$$

Here $K^{(s)'}(t) Y^{(s)} = k^{(s)}(t, y)$, $Y^{(s)} = y \otimes \cdots \otimes y$, $i_1 + \cdots + i_{p-1} = p$. Thus in (4.5), there will be a vector $K(t) = (k^{(1)'}, K^{(2)'}, \dots)'$ instead of $k(t)$. The first m rows of $\mathcal{G}$ are equal to $(G^{(1)}, G^{(2)}, \dots)$, next m^2 rows are equal to

$$
(\mathbf{0}, G^{(1)} \otimes G^{(1)}, G^{(2)} \otimes G^{(1)} + G^{(1)} \otimes G^{(2)}, \dots), \dots, V'(t) X = v(t, x). \tag{4.9}
$$

For approximate calculations, it is possible to approximate $\varphi(x)$ $(\varphi(0) = 0)$ by the polynomial $\varphi(x) \approx \varphi^{(1)}(x) + \cdots + \varphi^{(r)}(x) = W_r' X_r$, $X_r = (x', \dots, X^{(r)'})'$, and solve a finite-dimensional linear problem $V_r(T) = (V^{(1)'}, \dots, V^{(r)'})' \approx W_r$ choosing $K_r(t) = (k^{(1)'}, \dots, K^{(r)'})'$. Block-triangular structure of matrices in a dual system simplifies the problem and allows writing a subsystem for $V_r(t)$, which will not include $V^{(i)}$, $K^{(i)}$, $i > r$. This property consequently allows solving control problems in the form $V^{(j)}(T) \approx W^{(j)}$. Finally,

$$
\varphi(x(T)) \approx \int_0^T k_r(t, y(t)) dt, \qquad k_r(t, y) = K_r'(t) Y_r, \tag{4.10}
$$

k_r is a polynomial of degree r of y. Due to the growth of dimension, large r are not used. In applications, when linear approximation is singular, it is possible to consider $r = 2$. In the problem of ideal observability, there will be additional restrictions $H_j' V(t) = 0$, where H_j are constructed using h^j in the same way as a matrix $\mathcal{F}$ is using f.

Instead of power functions, it is possible to use other functions, taking into consideration specificity of nonlinearity of f and g. Analyticity is not necessary. Choose basis $\psi_1(x), \dots, \psi_N(x)$ (where $x \in \tilde{U}$, see Remark 3.3). Choose functions h_j, $1 \le j \le r$, such that $\varphi(x)$ and $h_j(t, g(t, x))$ are rather precisely approximated in a basis when $x \in \tilde{U}$:

$$
\varphi(x) \approx \sum_{\nu=1}^{N} d_\nu \psi_\nu(x), \qquad h_j(t, g(t, x)) \approx \sum_{\nu=1}^{N} b_{j\nu}(t) \psi_\nu(x). \tag{4.11}
$$

In the same way, approximate

$$\mathcal{A}\psi_\nu = \psi_{\nu x}(x) \cdot f(t,x) \approx \sum_{\nu=1}^{N} a_{\nu\mu}(t)\psi_\mu(x). \tag{4.12}$$

Construct the functions k, v in the form

$$k(t,y) \approx \sum_{j=1}^{r} k_j(t)h_j(y), \qquad v(t,x) \approx \sum_{\nu=1}^{N} v_\nu(t)\psi_\nu(x). \tag{4.13}$$

After substituting the above expressions in the dual system (3.10) and equating the coefficients at ψ_ν, we obtain a finite-dimensional two-point problem

$$\dot{V}(t) = -A'V(t) + B'k(t), \qquad V(0) = 0, \qquad V(T) \approx d,$$
$$d = \{d_j\}, \qquad V = \{v_j\}, \qquad k = \{k_j\}, \qquad A = \{a_{ij}\}, \qquad B = \{b_{ij}\}. \tag{4.14}$$

Finally,

$$\varphi(x(T)) \approx \int_0^T \sum_{j=1}^{r} k_j(t)h_j(t,y(t))dt. \tag{4.15}$$

For the predicting problem $k_j(t) = 0$, $t > \vartheta$ should be considered.

The problem can be understood from the point of view of the theory of approximate solving of linear boundary value problems. Select any smooth function

$$v(t,x) : v(0,x) = 0, \quad v(T,x) = \varphi(x), \quad x \in \tilde{U} \ (v(t,x) = t\varphi(x)/T). \tag{4.16}$$

Add a sum $\alpha_1 v_1(t,x) + \cdots + \alpha_N v_N(t,x)$, $v_i(0,x) = v_i(T,x) = 0$ (e.g., $v_i(t,x) = t(t-T)\vartheta_i(t)\eta_i(x)$). In the same way, $k(t,y) = \beta_1 k_1(t,y) + \cdots + \beta_M k_M(t,y)$. Substituting these expressions in the distributed dual equation (3.10), we obtain a discrepancy $R(t,x;\alpha_1,\ldots,\beta_M)$. It should be minimized on parameters in the appropriate norm in $[0,T] \times \tilde{U}$. In the ideal observation problem, it is necessary to consider also a discrepancy in phase restrictions. Linearity with respect to a pair (k,v) allows using a large amount of direct methods which are developed for the problems of mathematical physics.

Usually, only a part of phase variables is measured: $y_i = x_i$, $i \le m$. Formally, this can be achieved by changing or adding variables. Even for weaker suppositions of smoothness, in the right-hand side of the dual system (3.10) there is a function of $t, x_1, \ldots, x_m$, and it is possible to consider the problem

$$\mathcal{L}v = 0, \qquad v(0,\cdot) = 0, \qquad v(T,\cdot) = \varphi, \qquad \mathcal{L}v = \left\{ (v_t + v_x \cdot f)_{x_i} \right\}_{i=m+1}^{n}. \tag{4.17}$$

As a weight function $k(t, y_1, \ldots, y_m)$ in the integral observation operator (1.3) there will be an expression $L_f v$.

Note that the initial problem is nonlinear inverse, but finally we used direct methods of solving a linear equation though being distributed (due to constructing observation operators for the domain of a phase space).

References

[1]   H. Cartan, *Calcul Différentiel*, Hermann, Paris, 1967 (French).
[2]   ———, *Formes différentielles. Applications élémentaires au calcul des variations et à la théorie des courbes et des surfaces*, Hermann, Paris, 1967 (French).
[3]   E. M. Chirka, *Complex Analytic Sets*, Nauka, Moscow, 1985 (Russian).
[4]   E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.
[5]   M. Hervé, *Several Complex Variables. Local Theory*, Oxford University Press, London, 1963.
[6]   Y. Inouye, *On the observability of autonomous nonlinear systems*, J. Math. Anal. Appl. **60** (1977), no. 1, 236–247.
[7]   N. E. Kirin, *Estimation and Control Methods in Dynamical Systems*, Izdatel' stvo Sankt-Peterburgskogo Universiteta, St. Petersburg, 1993 (Russian).
[8]   K. V. Kozerenko, *On the number of readings*, Dokl. Akad. Nauk SSSR **296** (1987), no. 5, 1069–1071 (Russian).
[9]   N. N. Krasovskiĭ, *Theory of Control of Motion: Linear Systems*, Izdat. Nauka, Moscow, 1968 (Russian).
[10] P. Lancaster, *Theory of Matrices*, Academic Press, New York, 1969.
[11] H. Nikaidô, *Convex Structures and Economic Theory*, Mathematics in Science and Engineering, vol. 51, Academic Press, New York, 1968.
[12] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
[13] K. E. Starkov, *Number characteristics of observability for nonlinear continuous-time control systems*, IMA J. Math. Control Inform. **17** (2000), no. 4, 425–437.
[14] Yu. V. Zaika, *Stable discrete observation programs in analytic dynamical systems*, Math. Notes **66** (1999), no. 2, 153–159.

Yury V. Zaika: Institute of Applied Mathematical Research, Karelian Research Center, Russian Academy of Sciences, Petrozavodsk 185610, Russia
E-mail address: zaika@krc.karelia.ru