Acta Applicandae Mathematicae **14** (1989), 191–237.

Local Invariants of Smooth Control Systems

A. A. AGRACHEV, R. V. GAMKRELIDZE, and A. V. SARYCHEV
V. A. Steklov Institute of Mathematics, Academy of Sciences of the U.S.S.R., ul. Vavilova 42, Moscow GPS 117966, U.S.S.R.

(Received in revised form: 12 July 1988)

Abstract. Methods are presented for locally studying smooth nonlinear control systems on the manifold **M**. The technique of chronological calculus is intensively exploited. The concept of chronological connection is introduced and is used when obtaining the invariant expressions in the form of Lie bracket polynomials for high-order variations of a nonlinear control system. The theorem on adduction of a family of smooth vector fields to the canonical form proved in Section 4 is then applied to the construction of a nilpotent polynomial approximation for a control system. Finally, the relation between the attainable sets of an original system and an approximating one is established; it implies some conclusions on the local controllability of these systems.

AMS subject classifications (1980). Primary 49B10, 49E15.

Key words. Nonlinear control systems, geometric/Lie algebraic methods, chronological calculus, nilpotent polynomial approximation, controllability, attainable sets.

1. In this paper we study a smooth control system of the form

$$\dot{q} = f(q) + g(q)u, \quad q \in M, \ u \in R, \ q(0) = q_0 \tag{1.1}$$

where M is a C^∞-manifold and $f(q)$, $g(q)$ are complete smooth vector fields.

For $t > 0$, we define a mapping $F_t\colon L_\infty[0, t] \to M$, which maps control $u(\cdot \in L_\infty[0, t]$ to the point $q(t) \in M$, where $\dot{q}(\tau) = f(q(\tau)) + g(q(\tau))u(\tau)$, $\tau \in [0, t]$, $q(0) = q_0$.

It is easy to show, that F_t is an infinitely differentiable mapping. The family of mappings F_τ, $\tau \in (0, +\infty)$ completely characterizes the control system.

The main goal of this paper is the description of some formalism which is 'inevitable' when methodically studying these families of mappings. Roughly speaking, the matter consists of eliminating from the Taylor expansions of F_t all terms which are unessential for the needs of control theory. Almost all definitions and results may be transferred without difficulties to the case of an arbitrary smooth nonlinear system $\dot{x} = f(x, u)$, $u \in R^r$. Corresponding extensions can be easily conjectured, but the formulations become more cumbersome. We intend to perform this in further publications.

When studying the control systems and/or mappings F_t, the using of operator notations of chronological calculus (see [1, 2]) is helpful. At first we will transfer these notations.

According to this calculus, we identify point $q \in M$ with a multiplicative functional $\varphi \to \varphi(q)$, $\forall \varphi \in C^\infty(M)$, defined on the algebra $C^\infty(M)$. Diffeomorphism P: $M \to M$ may be identified with a corresponding automorphism $\varphi(\cdot) \to \varphi(P(\cdot))$ of the algebra $C^\infty(M)$; the value of $P \in \operatorname{Diff} M$ at a point $q \in M$ is denoted as $q \circ P$, i.e., as a composition of an automorphism and multiplicative functional (which is a multiplicative functional again). The vector fields on M are arbitrary derivations of the algebra $C^\infty(M)$, i.e., R-linear mappings X: $C^\infty(M) \to C^\infty(M)$, which satisfy Leibnitz rule: $X(\varphi\psi) = (X\varphi)\psi + \varphi(X\psi)$. If Φ is a vector function, Φ: $M \to R^k$, $\Phi = (\varphi_1, \dots, \varphi_k)$, $\varphi_i \in C^\infty(M)$, $i = 1, \dots, k$, then we guess that X acts component-wise.

The Lie bracket $[X, Y] = X \circ Y - Y \circ X$ defines a Lie algebra structure on the set of vector fields. In what follows, this Lie algebra is denoted Der M. If, when using local coordinates $x = (x_1, \dots, x_n)$, the vector fields X and Y are presented in the form

$$X = \sum_{i=1}^n X_i \frac{\partial}{\partial x_i}, \qquad Y = \sum_{i=1}^n Y_i \frac{\partial}{\partial x_i},$$

then the Lie bracket

$$[X, Y] = \frac{\partial Y}{\partial x} X - \frac{\partial X}{\partial x} Y = \sum_{i=1}^n \left(\frac{\partial Y_i}{\partial x} X - \frac{\partial X_i}{\partial x} Y \right) \frac{\partial}{\partial x_i}.$$

The value of vector field X at a point $q \in M$ (being a tangent vector to the manifold M at q) is denoted $q \circ X$. T_qM is, as usual, the tangent space to M at q. We consider it as a space of R-linear functionals ξ on $C^\infty(M)$ satisfying, in addition, an equality

$$\xi(\varphi\psi) = (\xi\varphi)\psi(q) + \varphi(q)(\xi\psi).$$

If $P \in \operatorname{Diff} M$, we use $\operatorname{Ad} P$ to denote the inner automorphism of Der M: $\operatorname{Ad} PX = P \circ X \circ P^{-1}$, and $\mathbf{ad}\, Y$ (where $Y \in \operatorname{Der} M$) the inner derivation of Der M; $(\mathbf{ad}\, Y)X = [Y, X]$.

Let N be a C^∞-manifold and Φ be a diffeomorphism of M onto N; then Φ_* is a notation for the differential of Φ, Φ_*: $\operatorname{Der} M \to \operatorname{Der} M$, and Φ_{*q}: $T_qM \to T_{\Phi(q)}N$, $(q \in M)$ is the corresponding linear mapping of tangent spaces, though Φ_{*q} may be defined not only for diffeomorphisms but also for any smooth mapping Φ: $M \to N$. If $P \in \operatorname{Diff} M$, then we have $P_* = \operatorname{Ad} P^{-1}$.

The notations introduced above give exact information as to whether diffeomorphism P placed in some formula should be regarded as an operator or as a smooth mapping. It is determined in accordance with the left or right position of point q with respect to P.

We regard $C^\infty(M)$ as being provided with Whitney topology, i.e., the topology of uniform convergence of all the derivatives on compact sets. The Whitney topology can be set by means of seminorms $\|\circ\|_{s,K}$, when $s \geqslant 0$, $K \subseteq M$. This seminorm defines the topology of uniform convergence of all the derivatives up to the sth order on the compactum K. The seminorms $\|\circ\|_{s,K}$, in contrast to the topology they set, are not defined uniquely and may be chosen in many ways. In what follows, we assume that the choice is made and the seminorms are settled.

For matrix-valued functions on M we set

$$\|A\|_{s,K} = \sum_{j=1}^{t} \max_i \|a_j^i\|_{s,K};$$

specifically if Φ is a vector-function on M; $\Phi: M \to R^r$, then

$$\|\Phi\|_{s,K} = \max_{1 \leq i \leq r} \|\Phi_i\|_{s,K}.$$

For the vector fields on M, we define the seminorms

$$\forall X \in \operatorname{Der} M \|X\|_{s,K} = \sup\{\|X\varphi\|_{s,K} \mid \|\varphi\|_{s+1,K} = 1\}. \tag{1.2}$$

When the Whitney topology of $C^\infty(M)$ is set, we may define in various spaces of operators and functionals on $C^\infty(M)$, the topology of pointwise convergence. Further, we will often deal with one-parameter families of operators on $C^\infty(M)$ and the introduced topology of $C^\infty(M)$ gives the possibility of giving sense to the notions of continuity, measurability, differentiability, absolute continuity, etc., with respect to parameter. Indeed, if A_t is a family of linear operators $A_t: C^\infty(M) \to C^\infty(M)$ or linear functionals $A_t: C^\infty(M) \to R$ we say that A_t possesses property (*) with respect to t if $\forall \varphi \in C^\infty(M)$ function $A_t\varphi$ possesses the same property with respect to t (see [1, 2] for the details). The locally integrable (on t) families $X_t \in \operatorname{Der} M$ are called nonstationary vector fields on M and absolutely continuous (on t) families $P_t \in \operatorname{Diff} M$ – flows on M.

Nonstationary vector fields X_t, $t \in R$ provide an ordinary differential equation

$$\frac{dq}{dt} = q(t) \circ X_t \text{ on } M.$$

X_t is complete if every solution of this equation exists for all $t \in R$. A complete field X_t defines the flow P_t, $t \in R$ – the unique solution of the operator equation

$$\frac{dP_t}{dt} = P_t \circ X_t, \quad P_0 = \operatorname{Id}, \tag{1.3}$$

where Id is identical operator (identical diffeomorphism of M); we call this flow (see [1, 2]) the right chronological exponential in X_t and denote $\overrightarrow{\exp} \int_0^t X_\tau \, d\tau$.

In [1], the representation of the flow $P_t = \overrightarrow{\exp} \int_0^t X_\tau \, d\tau$ in the form of the so-called Volterra series is given

$$\begin{aligned} \overrightarrow{\exp} \int_0^t X_\tau \, d\tau &\approx \operatorname{Id} + \sum_{i=1}^{\infty} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \ldots \int_0^{\tau_{i-1}} d\tau_i (X_{\tau_i} \circ \cdots \circ X_{\tau_1}) \\ &= \operatorname{Id} + \int_0^t X_{\tau_1} \, d\tau_1 + \int_0^t \int_0^{\tau_1} (X_{\tau_2} \circ X_{\tau_1}) \, d\tau_1 \, d\tau_2 + \cdots \end{aligned} \tag{1.4}$$

In the real analytic case (to simplify the notation we will also suppose here that $M = R^n$), this series converges, when $\int_0^t \|\bar{X}_\tau\|_\sigma \, d\tau$ is sufficiently small. Here

$$\|\bar{X}_\tau\|_\sigma = \max_i \|\bar{X}_\tau^i\|_\sigma = \max_{1 \leq i \leq n} \sup_{z \in V_\sigma} |\bar{X}_\tau^i(z)|$$

and $\bar{X}_\tau$ is an analytic extension of the real analytic vector field X_τ to the complex neighborhood V_σ of $R^n \subseteq C^n$:

$$V_\sigma = \{(z_1, \ldots, z_n) \in C^n \mid |\mathrm{Im}\, z_i| \leq \sigma, i = \overline{1, n}\}$$

In the C^∞ case, the Volterra series is an asymptotic one for $\overrightarrow{\exp} \int_0^t X_\tau \, d\tau$: $\forall \varphi \in C^\infty(M)$

$$\left\| \left(\overrightarrow{\exp} \int_0^t X_\tau \, d\tau - \left(\mathrm{Id} + \sum_{i=1}^{m-1} \int_0^t d\tau_1 \ldots \int_0^{\tau_{i-1}} d\tau_1 (X_{\tau_i} \circ \cdots \circ X_{\tau_1}) \right) \right) \varphi \right\|_{s,K} \leq$$
$$\leq C_1 \, e^{C_2 \int_0^t \|X_\tau\|_{s,K'} d\tau} \left(\int_0^t \|X_\tau\|_{s+m-1,K'} \, d\tau \right)^m \|\varphi\|_{s+m,K'},$$

where K' is some neighborhood of compactum K (see [1] for details).

Let us consider a family of operators $\mathrm{Ad}\, P_t$ produced by the flow $P_t = \overrightarrow{\exp} \int_0^t X_\tau \, d\tau$. By the definition $\forall Y \in \mathrm{Der}\, M$ $\mathrm{Ad}\, P_t Y = P_t \circ Y \circ P_t^{-1}$. The differentiation of this expression with respect to t gives us (by virtue of (1.3))

$$\begin{aligned} \frac{d}{dt} \mathrm{Ad}\, P_t Y &= \frac{dP_t}{dt} \circ Y \circ P_t^{-1} + P_t \circ Y \circ \frac{dP_t^{-1}}{dt} \\ &= P_t \circ X_t \circ Y \circ P_t^{-1} - P_t \circ Y \circ X_t \circ P_t^{-1} \\ &= P_t \circ [X_t, Y] \circ P_t^{-1} = \mathrm{Ad}\, P_t \circ \mathrm{ad}\, X_t Y, \end{aligned}$$

whence $d/dt\, \mathrm{Ad}\, P_t = \mathrm{Ad}\, P_t \circ \mathrm{ad}\, X_t$, i.e., $\mathrm{Ad}\, P_t$ satisfies the differential equation similar to (1.3). So it is reasonable to set $\mathrm{Ad}\, P_t = \exp \int_0^t \mathrm{ad}\, X_\tau \, d\tau$, and then we obtain a representation

$$\begin{aligned} \mathrm{Ad}\left(\overrightarrow{\exp} \int_0^t X_\tau \, d\tau \right) &\approx \exp \int_0^t \mathrm{ad}\, X_\tau \, d\tau \\ &\approx \mathrm{Id} + \sum_{i=1}^{\infty} \int_0^t d\tau_i \ldots \int_0^{\tau_{i-1}} d\tau_i (\mathrm{ad}\, X_{\tau_i} \circ \cdots \circ \mathrm{ad}\, X_{\tau_1}). \end{aligned} \tag{1.5}$$

In the C^∞ case, the series at the right-hand side of (1.5) is asymptotic for $\mathrm{Ad}(\overrightarrow{\exp} \int_0^t X_\tau \, d\tau) : \forall Z \in \mathrm{Der}\, M$

$$\left\| \mathrm{Ad}\left(\overrightarrow{\exp} \int_0^t X_\tau \, d\tau \right) Z - \left(Z + \sum_{i=1}^{m-1} \int_0^t d\tau_1 \ldots \int_0^{\tau_{i-1}} d\tau_i (\mathrm{ad}\, X_{\tau_i} \circ \cdots \circ \mathrm{ad}\, X_{\tau_1}) Z \right) \right\|_{s,K}$$
$$\leq C_3 \, e^{C_4 \int_0^t \|X_\tau\|_{s+1,K'} d\tau} \|Z\|_{s+m,K'} \left(\int_0^t \|X_\tau\|_{s+m,K'} \, d\tau \right)^m.$$

When X_t is stationary (i.e., doesn't depend on t) $X_t \equiv X$, we use the traditional notation: $\overrightarrow{\exp} \int_0^t X_\tau \, d\tau = e^{tX}$.

For complete nonstationary vector fields X_t and $X_t + Y_t$, we will derive now the 'generalized variation of constants formula' presenting the flow $\overrightarrow{\exp} \int_0^t (X_\tau + Y_\tau) \, d\tau$ as a composition of flows

$$\overrightarrow{\exp} \int_0^t (X_\tau + Y_\tau) \, d\tau = C_t \circ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau, \tag{1.6}$$

i.e., as a perturbation of flow $\overrightarrow{\exp} \int_0^t X_\tau \, d\tau$ by a flow C_t. We will call them a nonperturbed flow and a perturbation one, correspondingly. In turn, we will look for C_t of a kind $C_t = \overrightarrow{\exp} \int_0^t Z_\tau \, d\tau$. Substituting the expression for C_t in (1.6) and differentiating the identity with respect to t, we get

$$\overrightarrow{\exp} \int_0^t (X_\tau + Y_\tau) \, d\tau \circ (X_t + Y_t) = C_t \circ Z_t \circ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau + C_t \circ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau \circ X_t$$

or by virtue of (1.6)

$$C_t \circ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau \circ Y_t = C_t \circ Z_t \circ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau.$$

The left action of the operator C_t^{-1} and the right action of $(\exp \int_0^t X_\tau \, d\tau)^{-1}$ on the both sides of the last equality, give us

$$\begin{aligned} Z_t &= \left(\overrightarrow{\exp} \int_0^t X_\tau \, d\tau \right) \circ Y_t \circ \left(\overrightarrow{\exp} \int_0^t X_\tau \, d\tau \right)^{-1} \\ &= \mathrm{Ad} \left(\overrightarrow{\exp} \int_0^t X_\tau \, d_\tau \right)^{-1} Y_t = \overrightarrow{\exp} \int_0^t \mathrm{ad}\, X_\tau \, d\tau Y_t. \end{aligned}$$

Hence

$$C_t = \overrightarrow{\exp} \int_0^t \overrightarrow{\exp} \int_0^\tau \mathrm{ad}\, X_\theta \, d\theta Y_\tau \, d\tau$$

and we get the 'variation of constants formula'

$$\overrightarrow{\exp} \int_0^t (X_\tau + Y_\tau) \, d\tau = \overrightarrow{\exp} \int_0^t \overrightarrow{\exp} \int_0^\tau \mathrm{ad}\, X_\theta \, d\theta Y_\tau \, d\tau \circ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau. \tag{1.7}$$

The 'variation of constants formula' is helpful when operating with complicated perturbations of nonstationary nonlinear systems of differential equations. For example, let us consider the family of flows

$$Q_t(\epsilon) = e^{\epsilon Z_0} \circ \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau) \, d\tau,$$

smoothly depending on parameter $\epsilon \in R$. Here, Z_0 is a stationary, and X_τ, Y_τ a nonstationary vector field, ϵY_τ is a perturbation of the right-hand side of the differential equation $\dot{x} = X_t(x)$, and $e^{\epsilon z_0}$ is a perturbation of the initial condition.

Calculate at first $\partial/\partial\epsilon \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau$. By virtue of (1.7), we get

$$
\begin{aligned}
&\overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau \\
&\quad = \overrightarrow{\exp} \int_0^t \mathrm{Ad}\Big(\overrightarrow{\exp} \int_0^\tau X_\theta\,d\theta\Big)\epsilon Y_\tau\,d\tau \circ \overrightarrow{\exp} \int_0^t X_\tau\,d\tau \\
&\quad = \Big(\mathrm{Id} + \epsilon \int_0^t \mathrm{Ad}\Big(\overrightarrow{\exp} \int_0^\tau X_\theta\,d\theta\Big) Y_\tau\,d\tau\Big) + O(\epsilon^2).
\end{aligned}
$$

Hence,

$$
\frac{\partial}{\partial\epsilon}\bigg|_{\epsilon=0} \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau = \int_0^t \mathrm{Ad}\Big(\overrightarrow{\exp} \int_0^\tau X_\theta\,d\theta\Big) Y_\tau\,d\tau \circ \overrightarrow{\exp} \int_0^t X_\tau\,d\tau.
$$

Similarly, for an arbitrary ϵ,

$$
\begin{aligned}
&\frac{\partial}{\partial\epsilon} \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau \\
&\quad = \int_0^t \mathrm{Ad}\Big(\overrightarrow{\exp} \int_0^\tau (X_\theta + \epsilon Y_\theta)\,d\theta\Big) Y_\tau\,d\tau \circ \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau.
\end{aligned}
$$

So differentiating $Q_t(\epsilon)$ with respect to ϵ, we get

$$
\begin{aligned}
&\frac{\partial}{\partial\epsilon} Q_t(\epsilon) \\
&\quad = Z_0 \circ e^{\epsilon Z_0} \circ \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau + \\
&\qquad + \int_0^t e^{\epsilon Z_0}\,\mathrm{Ad}\Big(\overrightarrow{\exp} \int_0^\tau (X_\theta + \epsilon Y_\theta)\,d\theta\Big) Y_\tau\,d\tau \circ \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau \\
&\quad = Z_0 \circ Q_t(\epsilon) + \int_0^t e^{\epsilon Z_0} \circ \overrightarrow{\exp} \int_0^\tau (X_\theta + \epsilon Y_\theta)\,d\theta \circ \\
&\qquad \circ Y_\tau \circ \Big(\overrightarrow{\exp} \int_0^\tau (X_\theta + \epsilon Y_\theta)\,d\theta\Big)^{-1} \circ e^{-\epsilon Z_0}\,d\tau \circ e^{\epsilon Z_0} \circ \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\,d\tau \\
&\quad = \Big(Z_0 + \int_0^t \mathrm{Ad}\,Q_\tau(\epsilon)\,Y_\tau\,d\tau\Big) \circ Q_t(\epsilon).
\end{aligned}
$$

If $Z_t(\epsilon) = Z_0 + \int_0^t \mathrm{Ad}\,Q_\tau(\epsilon)\,Y_\tau\,d\tau$, then we get

$$
\frac{\partial Q_t}{\partial\epsilon} = Z_t(\epsilon) \circ Q_t(\epsilon). \tag{1.8}
$$

In turn

$$\frac{\partial Z_t(\epsilon)}{\partial \epsilon}$$
$$= \int_0^t \frac{\partial}{\partial \epsilon} \operatorname{Ad} Q_\tau(\epsilon) Y_\tau \, d\tau$$
$$= \int_0^t \frac{\partial}{\partial \epsilon} (Q_\tau(\epsilon) \circ Y_\tau \circ Q_\tau^{-1}(\epsilon)) \, d\tau$$
$$= \int_0^t [Z_\tau(\epsilon), \operatorname{Ad} Q_\tau(\epsilon) Y_\tau] \, d\tau$$
$$= \int_0^t \left[Z_\tau(\epsilon), \frac{\partial}{\partial \tau} Z_\tau(\epsilon) \right] d\tau. \tag{1.8}$$

In what follows up to the end, two admissions are adopted:

(1) All the vector fields (nonstationary or usual ones) are supposed to be complete. The sufficient conditions of completeness are well known, so we need not dwell on it.

(2) We operate only with the diffeomorphisms which can be included in some flow P_t, $t \in R$ on M, $P_0 = \mathrm{Id}$. In other words, the notation Diff M denotes below not the whole group of the diffeomorphisms of M, but its linear-connected component, which contains Id.

One more technical detail: for the family of vectors $l_\tau (\tau \in [0, t])$ of some vector space, let us define

$$\operatorname{vraispan}\{l_\tau \mid \tau \in [0, t]\} = \bigcap_{\substack{i \subset [0,t] \\ \operatorname{mes} I = 0}} \operatorname{Span}\{l_\tau \mid \tau \in [0, t] \not\supset I\};$$

$$\operatorname{Span} \phi = 0.$$

2. Returning to the control system (1.1) and the family of mappings F_t, $t \in R$, we may now put

$$F_T(u(\cdot)) = q_0 \circ \overrightarrow{\exp} \int_0^t (f + gu(\tau)) \, d\tau.$$

From now on, we shall study the mapping F_t locally near the fixed admissible control $\tilde{u}(\cdot)$. Note that this problem differs from the one of the ascertaining of small time local controlability; it seems that the first problem is more complicated than the second one.

It is convenient to study, in place of F_t, the equivalent mapping

$$G_t \colon u(\cdot) \to F_t(u(\cdot)) \circ \left(\overrightarrow{\exp} \int_0^t (f + g\tilde{u}(\tau)) \, d\tau \right)^{-1}.$$

Denote

$$h_\tau = \left(\operatorname{Ad} \overrightarrow{\exp} \int_0^\tau (f + g\tilde{u}(\theta))\, d\theta\, g = \overrightarrow{\exp} \int_0^\tau \operatorname{ad}(f + g\tilde{u}(\theta))\, d\theta\, g\right).$$

Employing (1.7), we get

$$G_t(u(\cdot)) = q_0 \circ \overrightarrow{\exp} \int_0^t h_\tau u(\tau)\, d\tau.$$

So instead of control system (1.1), we may consider the control system

$$\dot{q} = h_\tau(q) u, \quad q(0) = q_0, \tag{2.1}$$

with the right side equal to zero, when $u = 0$.

The study of the local properties of attainable sets and the conditions of the local extremality of the given control $u(\cdot)$ for some functional (i.e., the main problems of control theory), are reduced, generally speaking, to descriptions of the images of G_t restricted (maybe) on some special subsets of the space of admissible controls $L_\infty[0, t]$ – the ones being contained in some small neighborhood of zero in $L_\infty[0, t]$.

For any smooth mapping, its differential and Hessian are the simplest local invariants. If $G: A \to M$ is an arbitrary smooth mapping of Banach manifold to $M(G(a_0) = q_0)$, then a differential G' is a linear mapping of the tangent space $T_{a_0}A$ to the tangent space $T_{q_0}M$. Choosing some local coordinates $\alpha: A \to \bar{A}$, $\mu: M \to R^n$, $\alpha(a_0) = 0$, $\mu(q_0) = 0$, we get the smooth mapping $\mu \circ G \circ \alpha^{-1}: \bar{A} \to R^n$. (Here $\bar{A}$ is a Banach space, modeling manifold A locally.) In these local coordinates, the differential of G comes to the differential of $\mu \circ G \circ \alpha^{-1}$ at zero of $\bar{A}$, which linearly maps $\bar{A}$ to R^n.

As for the Hessian, let us note at first that the second derivative of mapping $\mu \circ G \circ \alpha^{-1}$ at zero is some quadratic (or corresponding symmetric bilinear) mapping of $\bar{A}$ to R^n. It is obvious that this bilinear mapping essentially depends on the choice of coordinate systems. For example, if a_0 is a regular point for G, i.e., the differential $G'(a_0): T_{a_0}A \to T_{q_0}M$ is surjective, then the implicit function theorem implies the possibility of choosing such local coordinates that the mapping $\mu \circ G \circ \alpha^{-1}$ becomes linear, hence the Hessian equals zero. However, if we restrict the bilinear mapping defined above to the kernel of the differential $\mu' \circ G' \circ \alpha'^{-1}$ and, in addition, factorize its values modulo the image of the same differential, then it turns out to be a correctly defined (and invariant!) bilinear symmetric mapping

$$G'': \operatorname{Ker} G' \times \operatorname{Ker} G' \to \operatorname{coker} G'.$$

We will call G'' the Hessian of G at a point a_0.

Let us return to the control system (2.1) and the mappings G_t. Let Φ be some coordinate mapping of the neighborhood $V_{q_0} \supset q_0$ to R^n, $\Phi\ (q_0) = 0$. Using the Volterra-series expansion of the right chronological exponential, we get

$$\Phi(G_t(u(\circ))) = q_0 \circ \overrightarrow{\exp} \int_0^t h_\tau u(\tau)\,\mathrm{d}\tau \circ \Phi = \int_0^t (q_0 \circ h_\tau) u(\tau)\,\mathrm{d}\tau \circ \Phi +$$

$$+ q_0 \circ \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 (h_{\tau_2} u(\tau_2) \circ h_{\tau_1} u(\tau_1)) \circ \Phi + 0\left(\int_0^t |u(\tau)|\,\mathrm{d}\tau\right)^2.$$

Evidently, the differential G_t': $L_\infty[0, t] \to T_{q_0}M$ of the mapping G_t at zero is

$$G_t' u(\circ) = \int_0^t (q_0 \circ h_\tau) u(\tau)\,\mathrm{d}\tau,$$

and its image im G_t' coincides with the varispan$\{q_0 \circ h_\tau \mid \tau \in [0, t]\}$. We denote it E_t^1.

Consider now the quadratic term

$$q_0 \circ \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2\, h\tau_2 u(\tau_2) \circ h\tau_1 u(\tau_1)$$

of Volterra series, and restrict this quadratic form to ker G_t', i.e., the set of $u(\cdot)$ satisfying the equality

$$\int_0^t (q_0 \circ h_\tau) u(\tau)\,\mathrm{d}\tau = 0. \tag{2.2}$$

The integration by parts gives us

$$\begin{aligned} &q_0 \circ \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2\, h\tau_2 u(\tau_2) \circ h\tau_1 u(\tau_1) \\ &= q_0 \circ \int_0^t \left(\left(\int_0^{\tau_1} h\tau_2 u(\tau_2)\,\mathrm{d}\tau_2\right)\right) \circ \mathrm{d}\left(\int_0^{\tau_1} h\tau_2 u(\tau_2)\,\mathrm{d}\tau_2\right) \\ &= q_0 \circ \int_0^t h\tau_2 u(\tau_2)\,\mathrm{d}\tau_2 \circ \int_0^t h\tau_2 u(\tau_2)\,\mathrm{d}\tau_2 - \\ &\quad - q_0 \circ \int_0^t h\tau_1 u(\tau_1) \circ \int_0^{\tau_2} h\tau_2 u(\tau_2)\,\mathrm{d}\tau_2\,\mathrm{d}\tau_1. \end{aligned} \tag{2.3}$$

The first term on the right-hand side of (2.3) vanishes by virtue of (2.2); hence

$$\begin{aligned} &- q_0 \circ \int_0^t \mathrm{d}\tau_1 \int_0^{\tau_1} \mathrm{d}\tau_2 h_{\tau_2} u(\tau_2) \circ h_{\tau_1} u(\tau_1) \\ &= \tfrac{1}{2} \int_0^t q_0 \circ \left(\int_0^{\tau_1} h_{\tau_2} u(\tau_2)\,\mathrm{d}\tau_2 \circ h_{\tau_1} u(\tau_1) - h_{\tau_1} u(\tau_1) \circ \int_0^{\tau_1} h_{\tau_2} u(\tau_2)\,\mathrm{d}\tau_2\right) \mathrm{d}\tau_1 \\ &= \tfrac{1}{2} \int_0^t q_0 \circ \left[\int_0^{\tau_1} h_{\tau_2} u(\tau_2)\,\mathrm{d}\tau_2,\, h_{\tau_1} u(\tau_1)\right] \mathrm{d}\tau_1. \end{aligned}$$

The expression

$$\tfrac{1}{2} \int_0^t q_0 \circ \left[\int_0^{\tau_1} h_{\tau_2} u(\tau_2)\,\mathrm{d}\tau_2,\, h\tau_1 u(\tau_1)\right] \mathrm{d}\tau_1$$

is (for given $u(\cdot)$) a vector from the tangent space Tq_0M and, hence, is defined invariantly. After factorization, the image of this quadratic mapping modulo im G'_t, we conclude that the Hessian of G_t at zero is

$$G''_t u(\cdot) = q_0 \circ \int_0^t \left[\int_0^\tau h_\theta u(\theta)\, d\theta,\, h_\tau u(\tau) \right] d\tau + E_t^1, \tag{2.4}$$

when $u(\cdot) \in \operatorname{Ker} G'_t$, i.e., satisfies (2.2). Corresponding symmetric bilinear mapping

$$G''_t(u(\cdot), v(\cdot)) = q_0 \circ \int_0^t \left[\int_0^\tau h_\theta u(\theta)\, d\theta,\, h_\tau v(\tau) \right] d\tau + E_t^1 \tag{2.5}$$

will be denoted by the same symbol G''_t (the sharpening will be done if necessary).

We denote

$$E_t^1 = \operatorname{vraispan}\{q_0 \circ h_\tau,\ \tau \in [0, t]\} = \operatorname{Im} G'_t,$$

$$E_t^2 = E_t^1 + \operatorname{vraispan}\{q_0 \circ [h\tau_1, h\tau_2],\ \tau_1, \tau_2 \in [0, t]\} = \operatorname{Span} \operatorname{Im} G''_t,$$

$$E_t^1 \subseteq E_t^2 \subseteq T_{q_0}M.$$

By setting $h_\tau^1 = q_0 \circ h_\tau \in E_t^1$, we may consider h_τ^1 as a constant vector field on the vector space E_t^1, evidently $[h_{\tau_1}^1, h_{\tau_2}^1] = 0 \forall \tau_1, \tau_2$. The control system $\dot{x} = h_\tau^1 u$, $x(0) = 0$ being defined on E_t^1 is actually a linearization of system (2.1). In fact, the mapping $G_t^1 : L_\infty[0, t] \to E_t^1$, where $G_t^1(u(\cdot)) = 0 \circ \overrightarrow{\exp} \int_0^t h_\tau^1 u(\tau)\, d\tau$ coincides with G'_t.

When $E_t^1 = T_{q_0}M$, then (in accordance with the implicit function theorem) the mapping G_t coincides in some proper coordinates with G_t^1. So, if the case, the control system (2.1) is completely characterized (locally) by its linearization, otherwise ($E_t^1 \neq T_{q_0}M$) we need to take into consideration the high-order approximations.

Let $\mathscr{E}_t^1 = \{X \in \operatorname{Der} M \mid q_0 \circ X \in E_t^1\}$ be the set of all fields on M, whose values at q_0 lie in E_t^1, and

$$\mathfrak{N}(\mathscr{E}_t^1) = \{X \in \operatorname{Der} M \mid q_0 \circ [X, \mathscr{E}_t^1] \in E_t^1\}$$

be a normalizer of the subspace $\mathscr{E}_t^1$ included in the Lie algebra Der M.

The foregoing formula (2.5) gives the direct expression for the bilinear mapping G''_t in terms of the vector fields $h_\tau \in \mathscr{E}_t^1$, $\tau \in [0, t]$. By definition, the vector fields $\int_0^\tau h_\theta u(\theta)\, d\theta$, $h_\tau v(\tau)$ belong to $\mathscr{E}_t^1$ for almost every $\tau \in [0, t]$. Then, for any nonstationary perturbations Z_τ^1, $Z_\tau^2 \in \mathfrak{N}(\mathscr{E}_t^1) \subset \mathscr{E}_t^1 \subset \operatorname{Der} M$, there is carried out an equality

$$q_0 \circ \int_0^t \left[\int_0^\tau (h_\theta + Z_\theta^1) u(\theta)\, d\theta,\, (h_\tau + Z_\tau^2) v(\tau) \right] d\tau$$

$$= q_0 \circ \int_0^t \left[\int_0^\tau h_\theta u(\theta)\, d\theta,\, h_\tau v(\tau) \right] d\tau \pmod{E_t^1}.$$

Hence, G_t'' depends, in fact, only on an image of h_τ under canonical factorization $\mathscr{E}_t^1 \to \mathscr{E}_t^1/\mathfrak{N}(\mathscr{E}_t^1)$. Suppose that

$$\dim E_t^1 = k_1 \quad \text{and} \quad (x, y), \quad x = (x_1, \ldots, x_{k_1}), \quad y = (y_1, \ldots, y_{k_2})$$

are local coordinates on M such that $(x, y)(q_0) = 0$, and

$$(x, y)_*|_{q_0} E_t^1 = \{(\xi, \eta) \mid \xi \in \mathbf{R}^{k_1}, \eta \in \mathbb{R}^{k_2}, \eta = 0\}.$$

In this case

$$\mathscr{E}_t^1 = \quad a(x, y)\frac{\partial}{\partial x} + b(x, y)\frac{\partial}{\partial y} \Big| \, b(0, 0) = 0\}.$$

Here

$$a(x, y) = (a_1(x, y_1), \ldots, a_{k_1}(x, y), \qquad b(x, y) = (b_1(x, y), \ldots, b_{k_2}(x, y))$$

are row vectors, and

$$\frac{\partial}{\partial x} = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_{k_1}}\right)^T, \qquad \frac{\partial}{\partial y} = \left(\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_{k_2}}\right)^T$$

are column vectors. Let us prove that

$$\mathfrak{N}(\mathscr{E}_t^1) = \left\{a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} \; a(0, 0) = 0, \; b(0, 0) = 0, \frac{\partial b}{\partial x}(0, 0) = 0\right\}.$$

The inclusion

$$0 \circ \left[a_1\frac{\partial}{\partial x} + b_1\frac{\partial}{\partial y}, a_2\frac{\partial}{\partial x} + b_2\frac{\partial}{\partial y}\right] \in E_t^1$$

provided that

$$b_1(0, 0) = b_2(0, 0) = 0, \qquad a_2(0, 0) = 0, \frac{\partial b_2}{\partial x}(0, 0) = 0$$

can be verified directly. On the other hand, if

$$Z = a(x, y)\frac{\partial}{\partial x} + b(x, y)\frac{\partial}{\partial y} \quad \text{and} \quad a(0, 0) \neq 0 \quad \text{or} \quad b(0, 0)\, W \neq 0,$$

then the value at zero of the Lie bracket $[Z, X]$, where

$$X = (a(0, 0)x + b(0, 0)y)\frac{\partial}{\partial y_1} \in \mathscr{E}_t^1$$

does not lie in E_t^1; its coordinate η_1 is equal to

$$\langle a(0, 0), a^T(0, 0)\rangle + \langle b(0, 0), b^T(0, 0)\rangle \neq 0.$$

If $\partial b_i/\partial x_j(0, 0) \neq 0$, then the value at zero of the Lie bracket $[Z, Y]$, where

$Y=\partial/\partial x_j\in\mathscr{E}_t^1$, does not lie in E_t^1 (coordinate η_i of this vector is equal $\partial b_i/\partial x_j(0,0)\neq 0$).

So the factorization $\mathscr{E}_t^1\to\mathscr{E}_t^1/\mathfrak{N}(\mathscr{E}_t^1)$ takes the form

$$\left(a(x,y)\frac{\partial}{\partial x}+b(x,y)\frac{\partial}{\partial y}\right)\to\left(a(0,0)\frac{\partial}{\partial x}+\frac{\partial b}{\partial x}(0,0)x\frac{\partial}{\partial y}\right).$$

Suppose now that

$$h_\tau=a_\tau(x,y)\frac{\partial}{\partial x}+b_\tau(x,y)\frac{\partial}{\partial y}.$$

Then

$$\left(\frac{\partial b_\tau}{\partial x}(0,0)x\right)^t\in(x,y)_*|_{q_0}E_t^2.$$

The result of action of the linear mapping $(x,y)_{*\,q_0}^{-1}$ on the field

$$a_\tau(0,0)\frac{\partial}{\partial x}+\frac{\partial b_\tau}{\partial x}(0,0)x\frac{\partial}{\partial y},$$

is a vector field on E_t^2 denoted h_τ^2, $\tau\in[0,t]$. Moreover,

$$[h_{\tau_1}^2,[h_{\tau_2}^2,h_{\tau_3}^2]]=0\forall\tau_1,\tau_2,\tau_3\in[0,t],$$

i.e., the Lie algebra generated by the vector fields h_τ^2 is nilpotent of length 2. It can be easily seen that the dimension of this Lie algebra does not exceed $k_1(n-k_1)+n$. In the simplest case $M=R^2$, $k_1=1$ we get the Heisenberg algebra generated by $\partial/\partial x$, $\partial/\partial y$, $x\,\partial/\partial y$. The mapping G_t^2: $L_\infty[0,t]\to E_t^2$ defined as

$$G_t^2(u(\cdot))=0\circ\overrightarrow{\exp}\int_0^t h_\tau^2u(\tau)\,d\tau,$$

takes, when using the coordinates (x,y), the following form

$$(x,y)_*|_{q_0}G_t^2(u(\cdot))=\int_0^t a^T(0,0)u(\tau)\,d\tau+$$
$$+\int_0^t\int_0^\tau\frac{\partial b_\tau}{\partial x}(0,0)a_\theta^\tau(0,0)u(\theta)\,d\theta u(\tau)\,d\tau.$$

The differential and the Hessian of G_t^2 at zero correspondingly coincide with G_t' and G_t''. So the control system

$$\dot x=h_\tau^2(x)u,\quad x(0)=0,\ \tau\in[0,t]$$

is a second-order approximation to the system (2.1)* (and, hence, system (1.1) for

* The employment of local coordinates for the definition of this system is essential. The matter is that the image of the field h_τ under the factorization $\mathscr{E}_t^1\to\mathscr{E}_t^1/\mathfrak{N}(\mathscr{E}_t^1)$, which is an invariant object, actually characterizing the second-order approximation, does not transform (under coordinate transformation) as a vector field on E_t^2. See also Remark 1 at the end of Section 4.

the controls close to zero one, when $E_t^2 = Tq_0M$. In reality, the mapping G_t may be transformed to G_t^2 by means of smooth coordinate transformations in $L_\infty[0, t]$ and $\mathbf{M}$, *only* if some hard and fast additional restrictions are imposed ([4]). Nevertheless, a number of properties being important for the control theory, in particular the structure of attainable sets, are similar in many cases, even when the proper coordinate transformation does not exist. We put off the discussion of this problem to the end of the paper in order to concern ourself now with constructing high-order approximations, when $E_t^2 \neq Tq_0M$.

It was mentioned above that the Hessian of the mapping G_t being an 'invariant part' of the second derivation of G_t, is defined on a kernel of the first derivative while its values lie in the factor space of E_t^2 modulo the image of the first derivative. The elementary analysis of formula (2.3) shows that just the equality of the first derivative to zero makes it possible to express the second derivative in terms of Lie brackets of the vector fields h_τ, $\tau \in [0, t]$. At the same time, the factorization of values of the second derivative makes it possible to pass from h_τ to h_τ^2, i.e., from the Lie algebra Der M to a special finite-dimensional nilpotent Lie algebra of the length 2. When dealing with high-order derivatives, each of the two items mentioned above (the constructing of proper Lie bracket expressions and the transition from Der M to some special nilpotent Lie algebra) is based on its own general construction which falls outside the limits of control theory. Let us pass to the description of these constructions.

3. Chronological Connection

Let Z_0 be an usual, while X_t, Y_t, $t \in R$ are nonstationary vector fields. Let us denote

$$Q_t(\epsilon) = e^{\epsilon Z_0} \circ \overrightarrow{\exp} \int_0^t (X_\tau + \epsilon Y_\tau)\, d\tau, \tag{3.1}$$

as a family of flows on M which depends smoothly on $\epsilon \in R$. Let $t \in R$ be fixed. One of the results of our further calculations will give us an invariant 'bracket' formula for the tangent vector to the curve $\epsilon \to q_0 \circ Q_t(\epsilon)$ in M at point $q_0 \circ Q_t(0)$ in the case when arbitrary singularities take place.

We calculated above (see (1.8)–(1.9)) that

$$\frac{\partial}{\partial \epsilon} Q_t(\epsilon) = Z_t(\epsilon) \circ Q_t(\epsilon), \tag{3.2}$$

where

$$Z_t(\epsilon) = \int_0^t \operatorname{Ad} Q_\tau(\epsilon)\, Y_\tau\, d\tau + Z_0,$$

and

$$\frac{\partial}{\partial \epsilon} Z_t(\epsilon) = \int_0^t \left[Z_\tau(\epsilon), \quad \frac{\partial}{\partial \tau} Z_\tau(\epsilon) \right] \mathrm{d}\tau = (Z(\epsilon) * Z(\epsilon))_t. \tag{3.3}$$

Here '*' denotes the 'chronological product' (see [2]), putting in accordance with an arbitrary pair of absolutely continuous on t nonstationary vector fields A_t, B_t the field

$$(A * B)_t = \int_0^t \left[A_\tau, \frac{\mathrm{d}}{\mathrm{d}\tau} B_\tau \right] \mathrm{d}\tau.$$

We notice that the '*'-product is nonassociative, although the following relation, more weak than associativity, takes place:

$$A * (B * C) - B * (A * C) = (A * B - B * A) * C.$$

Thus, we get

$$\frac{\partial}{\partial \epsilon} Q = Z(\epsilon) \circ Q(\epsilon), \tag{3.4}$$

$$\frac{\partial}{\partial \epsilon} Z(\epsilon) = Z(\epsilon) * Z(\epsilon), \tag{3.5}$$

The given equations have clear geometric interpretations, but before its description, some small digression is necessary.

Invariant connections on Lie groups. Let G be a Lie group; for any $x \in G$ let $\rho(x)$: $G \to G$ denote right translation on G, $\rho(x)y = yx\ \forall x \in G$. Then $\rho(G)$ is imbedded in Diff G – a group of diffeomorphisms of G on itself. $\rho(G)$ is a subgroup of Diff G which acts freely and transitively on G. Thus, any Lie group may be realized as a Lie subgroup of a group of diffeomorphisms of a smooth manifold. This subgroup acts freely and transitively on this manifold.

Let M be an arbitrary smooth manifold and Diff $M \supset G$ a Lie group of diffeomorphisms of M (onto itself), $\mathfrak{G}$ is a Lie algebra of G. Then $\mathfrak{G}$ is a Lie subalgebra of Der M (a Lie algebra of smooth vector fields on M). If G is connected, then

$$G = \left\{ \overrightarrow{\exp} \int_0^s X_\tau \, \mathrm{d}\tau \,|\, X_\tau \in \mathfrak{G};\ \tau, s \in R \right\}.$$

Smooth vector field on G is a name for the smooth mapping $p \to p \circ X_p$, where $X_p \in \mathfrak{G}\ \forall p \in G$, or, equivalently $p \to Y_p \circ p$, where $Y_p \in \mathfrak{G}\ \forall p \in G$; obviously $Y_p = \mathrm{Ad}\, pX_p$. The vector field $p \circ X_p$ on G is left-invariant (invariant with respect to the transformation of the kind $p \to p_0 \circ p$, where $p_0 \in G$ is fixed) iff X_p does not depend on p, and right-invariant (invariant with respect to the transformation of the kind $p \to p \circ p_0$) iff $Y_p = \mathrm{Ad}\, pX_p$ does not depend on p.

Let $(X, Y) \to \nabla_X Y$ be an arbitrary bilinear mapping of the Lie algebra $\mathfrak{G}$ into itself. Then the correspondence

$$(p \circ X_p, p \circ Y_p) \rightarrow p \circ \left(\nabla_{X_p} Y_p + \frac{\partial Y_p}{\partial p} X_p\right)$$

defines a left-invariant linear connection on the Lie group G (a covariant derivative of the vector field $p \circ Y_p$ in the direction of $p \circ X_p$). The correspondence

$$(X_p \circ p,\ Y_p \circ p) \rightarrow \left(-\nabla_{X_p} Y_p + \frac{\partial Y_p}{\partial p} X_p\right) \circ p$$

defines a right-invariant linear connection on G. All the left (right)-invariant connections on G may be described in this manner, because each of them is completely defined when its values for arbitrary left (right)-invariant vector fields are given. Let us put

$$R_\nabla(X,\ Y) = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]},$$

$$T_\nabla(X,\ Y) = \nabla_X Y - \nabla_Y X - [X,\ Y], \forall X,\ Y \in \mathfrak{G}.$$

The curvature tensor of the corresponding left(right)-invariant connection has the form

$$(p \circ X_p, p \circ Y_p) \rightarrow p \circ R_\nabla(X_p,\ Y_p) \qquad ((X_p \circ p,\ Y_p \circ p) \rightarrow R_\nabla(X_p,\ Y_p) \circ p);$$

and the torsion tensor has the form

$$(p \circ X_p, p \circ Y_p) \rightarrow p \circ T_\nabla(X_p,\ Y_p) \qquad ((X_p \circ p,\ Y_p \circ p) \rightarrow -T_\nabla(X_p,\ Y_p) \circ p).$$

Let P_s be an arbitrary absolutely continuous curve in G, then the fields

$$X_s = P_s^{-1} \circ \frac{\mathrm{d}}{\mathrm{d}s} P_s \quad \text{and} \quad Y_s = \frac{\mathrm{d}}{\mathrm{d}s} P_s \circ P_s^{-1} \quad (X_s,\ Y_s \in \mathfrak{G})$$

are called the left or right angular velocities of the curve P_s. The term 'angular velocity' originates from the analogy with the rotation of a rigid body, i.e., with the case when $M = R^3$, $G = \mathrm{SO}(3)$ (see [5]).

The parallel translation along the curve

$$P_s = P_0 \circ \overrightarrow{\exp} \int_0^s X_\tau \, \mathrm{d}\tau = \overleftarrow{\exp} \int_0^s Y_\tau \, \mathrm{d}\tau \circ P_0$$

by virtue of the left(right)-invariant connection defined by ∇, transforms an arbitrary vector $p_0 \circ Z(Z \circ p_0)$, being tangent to G at p_0, into the vector

$$P_s \circ \overleftarrow{\exp} \int_0^s -\nabla_{X_\tau} \, \mathrm{d}\tau \circ Z \quad \left(\overleftarrow{\exp} \int_0^s \nabla_{Y_\tau} \, \mathrm{d}\tau Z \circ P_s\right)$$

The curve P_s is a geodesic for the left(right)-invariant connection if it is smooth and

$$\frac{\partial}{\partial s} X_s + \nabla_{X_s} X_s = 0 \quad \left(\frac{\partial}{\partial s} Y_s = \nabla_{Y_s} Y_s\right).$$

In what follows, nothing hinders us to employ the concepts, which were used above when dealing with the Lie subgroup $G \subset \operatorname{Diff} M$, for the study of the whole group Diff M or, moreover, the group of flows on M (i.e., absolutely continuous curves on Diff M with the pointwise multiplication). The Lie algebra of the group Diff M is certainly Der M. Indeed, if P_ϵ ($\epsilon \geqslant 0$) is a smooth curve in Diff M ($P_0 = \mathrm{Id}$), then the direct computation shows that $\forall \epsilon\ P_\epsilon^{-1} \circ \partial P_\epsilon / \partial \epsilon$ is a derivation of the algebra $C^\infty(M)$, i.e., element of Der M. Hence, $P_\epsilon^{-1} \circ \partial P_\epsilon/\partial\epsilon|_{\epsilon=0} \in \operatorname{Der} M$, and so the tangent space to Diff M at 'point' Id is contained in Der M. On the other hand, any complete vector field X is a tangent vector at 'point' Id to the curve $e^{\epsilon X}$ in Diff M. If a vector field X is not complete, then we can consider a truncated field X-(complete) vector field which vanishes outside some compactum K' and coincides with X on some (chosen in advance) compactum K. Thus, if the dynamics we are interested in evolves on some compact subset of $M \times R$, then there is no need to take into consideration a completeness of the vector field we deal with. So we may regard Der M as a Lie algebra of Diff M.

Let us now consider Flow M – a Lie group of flows on M consisting of curves P_t in Diff M which are absolutely continuous with respect to t; the pointwise (on t) composition is a group operation: $(P_\bullet\ Q_\bullet)_t = P_t \circ Q_t$.

We will show that the Lie algebra of the Lie group Flow M is an algebra of all absolutely continuous curves in Der M with the pointwise (on t) commutator as an operation $[X_\bullet, Y_\bullet]_t = [X_t, Y_t]$. In fact, if $P_t(\epsilon)$ is a smooth (on ϵ) curve in the Lie group Flow M, such that $P_t(0) \stackrel{t}{\equiv} \mathrm{Id}$, then 'tangent vector' $\partial P_t/\partial\epsilon|_{\epsilon=0}$ is an absolutely continuous (on t) curve in Der M. Vice-versa, if X_t is an arbitrary curve in Der M absolutely continuous with respect to t, then there exists a vector field X_0 and locally integrable curve Y_t in Der M such that $X_t = X_0 + \int_0^t Y_\tau\, d\tau$. If $Y_\bullet$ is a complete nonstationary vector field, then consider in Flow M the curve $P_t(\epsilon) = e^{\epsilon X_0} \circ \exp \int_0^t Y_\tau\, d\tau$. Differentiating this flow with respect to X, we get

$$\left.\frac{\partial P_t}{\partial \epsilon}\right|_{\epsilon=0} = X_0 + \int_0^t Y_\tau\, d\tau = X_t,$$

i.e., X_t is a 'tangent vector' to Flow M at 'point' $P_t(0) \stackrel{t}{\equiv} \mathrm{Id}$.

Let us return to the family (3.1) of flows on M and the equations (3.2)–(3.5). It follows from them that the curve $\epsilon \to Q(\epsilon)$, lying in the group of flows, is geodesic for some right-invariant linear connection defined on this group. The connection cannot be re-established uniquely if we use only the equations for geodesics, since these equations contain only a covariant derivative of the angular velocity along itself. However, if in addition, we require the equality of the torsion tensor to zero, then the connection is re-established quite uniquely.

DEFINITION. The chronological connection is a right-invariant linear connection on the group of flows defined by bilinear mapping $(A_\bullet, B_\bullet) \to \nabla^c_A B$, where A_t, $B_t (t\in R)$ are arbitrary absolutely continuous curves in Der M and $\nabla^c_A B = A{*}\, B + \frac{1}{2}[A_0, B_0]$. It takes place $R\nabla^c(A.,B.) = \frac{1}{4}\,\mathrm{ad}[B_0, A_0]$, $T_{\nabla^c}(A.,B.) = 0$.

PROPOSITION 3.1. *Let $A_t (t\in R)$ be an absolutely continuous curve in Der M, $\mathring{A}_t = (\mathrm{d}/\mathrm{d}t)A_t$; $Q^0_t (t\in R)$ be a flow on M. Then there exists one and only one geodesic $\epsilon \to Q_*(\epsilon)$ of the connection ∇^c-meeting the conditions $Q_t(0) = Q^0_t$, $\partial/\partial\epsilon\, Q_t(0) = A_t \circ Q^0_t$, $t\in R$. This geodesic is defined by the formula*

$$Q_t(\epsilon) = \mathrm{e}^{\epsilon A_0} \circ \overrightarrow{\exp} \int_0^t \epsilon \mathring{A}_\tau \,\mathrm{d}\tau \circ Q^0_t .$$

It follows from the reasoning adduced above, that this formula really defines a geodesic of the connection ∇^c. The uniqueness, unlike the case of the finite-dimensional Lie group, is not trivial here but can be proved by means of one important 'composition property' of angular velocity, which we will now describe.

Let us denote $\Omega_t(\epsilon, A)$ the right angular velocity of the geodesic

$$Q_t(\epsilon) = \mathrm{e}^{\epsilon A_0} \circ \overrightarrow{\exp} \int_0^t \epsilon \mathring{A}_\tau \,\mathrm{d}\tau, \qquad \frac{\partial}{\partial\epsilon} Q_t(\epsilon) = \Omega_t(\epsilon, A) \circ Q_t(\epsilon);$$

we have $\Omega_t(\epsilon, A) = A_0 + \int_0^t \mathrm{Ad}\; Q_\tau(\epsilon)\mathring{A}_\tau d_\tau$. Direct calculation using formula (1.7) gives an identity

$$\Omega_t(\epsilon_1 + \epsilon_2; A.) = \Omega_t(\epsilon_1; \Omega(\epsilon_2; A.)). \tag{3.6}$$

Indeed,

$$\begin{aligned}
&\Omega_t(\epsilon_1 + \epsilon_2, A.)\\
&\quad = A_0 + \int_0^t \mathrm{Ad}\; Q_\tau(\epsilon_1 + \epsilon_2)\mathring{A}_\tau \,\mathrm{d}\tau\\
&\quad = A_0 + \int_0^t \mathrm{Ad}\,\mathrm{e}^{(\epsilon_1+\epsilon_2)A_0} \circ \mathrm{Ad}\Big(\overrightarrow{\exp} \int_0^\tau (\epsilon_1 + \epsilon_2)\mathring{A}_\theta \,\mathrm{d}\theta\Big)\mathring{A}_\tau \,\mathrm{d}\tau\\
&\quad = A_0 + \int_0^t \mathrm{Ad}\,\mathrm{e}^{\epsilon_1 A_0}\,\mathrm{Ad}\,\mathrm{e}^{\epsilon_2 A_0} \circ \mathrm{Ad}\,\overrightarrow{\exp} \int_0^\tau \mathrm{Ad}\,\overrightarrow{\exp} \int_0^\theta \epsilon_2 \mathring{A}_\xi \,\mathrm{d}\xi \mathring{A}_\theta \,\mathrm{d}\theta \circ\\
&\qquad \circ \mathrm{Ad}\,\overrightarrow{\exp} \int_0^\tau \epsilon_2 \mathring{A}_\theta \,\mathrm{d}\theta \mathring{A}_\tau \,\mathrm{d}\tau\\
&\quad = A_0 + \int_0^t \mathrm{Ad}\,\mathrm{e}^{\epsilon_1 A_0} \circ \Big(\mathrm{Ad}\,\mathrm{e}^{\epsilon_2 A_0} \circ \mathrm{Ad}\,\overrightarrow{\exp} \int_0^\tau \mathrm{Ad}\,\overrightarrow{\exp} \int_0^\theta \epsilon_2 \mathring{A}_\xi \,\mathrm{d}\xi \epsilon_1 \mathring{A}_\theta \,\mathrm{d}\theta \circ\\
&\qquad \circ \mathrm{Ad}\,\mathrm{e}^{-\epsilon_2 A_0}\Big) \circ \mathrm{Ad}\,\mathrm{e}^{\epsilon_2 A_0} \circ \mathrm{Ad}\,\overrightarrow{\exp} \int_0^\tau \epsilon_2 \mathring{A}_\theta \,\mathrm{d}\theta \mathring{A}_\tau \,\mathrm{d}\tau
\end{aligned}$$

$$= A_0 + \int_0^t \mathrm{Ad}\left(e^{\epsilon_1 A_0} \circ \overrightarrow{\exp} \int_0^\tau \epsilon_1 \left(\mathrm{Ad}\, e^{\epsilon_2 A_0} \circ \overrightarrow{\exp} \int_0^\theta \epsilon_2 \mathring{A}_\xi \, d\xi\right) \mathring{A}_\theta \, d\theta\right) \circ$$

$$\circ \mathrm{Ad}\left(e^{\epsilon_2 A_0} \circ \overrightarrow{\exp} \int_0^\tau \epsilon_2 \mathring{A}_\theta \, d\theta\right) \mathring{A}_\tau \, d\tau.$$

(We used here the following evident identities:

$$\mathrm{Ad}\, P \circ \mathrm{Ad}\, Q = \mathrm{Ad}\, PQ,\ \mathrm{Ad}\, P \circ \mathrm{Ad}\, P^{-1} = \mathrm{Id},\ P \circ \overrightarrow{\exp} \int_0^t X_\tau \, d\tau \circ P^{-1}$$

$$= \overrightarrow{\exp} \int_0^t \mathrm{Ad}\, PX_\tau \, d\tau.\Big)$$

On the other hand

$$\Omega_t(\epsilon_2, A.) = A_0 + \int_0^t \mathrm{Ad}\left(e^{\epsilon_2 A_0} \circ \overrightarrow{\exp} \int_0^\tau \epsilon_2 \mathring{A}_\theta \, d\theta\right) \mathring{A}_\tau \, d\tau,$$

and

$$\Omega_t(\epsilon_1, \Omega.(\epsilon_2, A.)) = A_0 + \int_0^t \mathrm{Ad}\, Q_\tau(\epsilon_1) \circ \mathrm{Ad}\left(e^{\epsilon_2 A_0} \circ \overrightarrow{\exp} \int_0^\tau \epsilon_2 \mathring{A}_\theta \, d\theta\right) \mathring{A}_\tau \, d\tau,$$

where

$$Q_t(\epsilon_1) = e^{\epsilon_1 A_0} \circ \overrightarrow{\exp} \int_0^\tau \epsilon_1 \mathring{\Omega}_\theta(\epsilon_2, A.) \, d\theta$$

$$= e^{\epsilon_1 A_0} \circ \overrightarrow{\exp} \int_0^\tau \epsilon_1 \, \mathrm{Ad}\left(e^{\epsilon_2 A_0} \circ \overrightarrow{\exp} \int_0^\theta \epsilon_2 \mathring{A}_\xi \, d\xi\right) \mathring{A}_\theta \, d\theta,$$

whence the identity (3.6) follows.

Moreover, this important property turns out to be true when scalar ϵ is replaced by a scalar function of t. More accurately, let u: $R \to R$ be an absolutely continuous function. Let us denote

$$Q_t(u) = e^{u(0) A_0} \circ \overrightarrow{\exp} \int_0^t \dot{u}(\tau) \dot{A}_\tau \, d\tau,$$

$$\Omega_t(u, A.) = u(0) A_0 + \int_0^t \mathrm{Ad}\, Q_\tau(u) \mathring{A}_\tau \, d\tau:$$

then

$$\Omega_t(u+v, A.) = \Omega_t(u; \Omega(v, A.)), \quad \forall u, v, A.$$

The last identity is also a direct consequence of a variations formula. It expresses the fact that the correspondence $u \to \Omega(u)$ defines a (nonlinear) action of the additive group of absolutely continuous functions on the space of (absolutely continuous on t) nonstationary vector fields.

Let us return to the proof of the uniqueness of the solution of the Caushi problem

$$\frac{\partial}{\partial \epsilon} Z_t(\epsilon) = \int_0^t [Z_\tau(\epsilon), \partial/\partial\tau Z_\tau(\epsilon)]\, d\tau, \quad Z_t(0) = A_t\,. \tag{3.7}$$

In this case, the solution coincides with $\Omega_t(\epsilon, A.)$. By virtue of (3.6), it is sufficient to prove that $\forall \epsilon\ \Omega_t(-\epsilon, Z.(\epsilon)) \equiv A.$, when $Z_t(\epsilon)$ satisfies (3.7). Since, for $\epsilon = 0$, this equality is carried out trivially, then it remains to prove that

$$\frac{\partial}{\partial \epsilon} \Omega_t(-\epsilon, Z.(\epsilon)) \equiv 0,$$

or as far as $\Omega_0(-\epsilon, Z.(\epsilon)) \stackrel{\epsilon}{=} A_0$, then it is sufficient to prove the following identity

$$\frac{\partial}{\partial \epsilon}\frac{\partial}{\partial t} \Omega_t(-\epsilon, Z.(\epsilon)) \equiv 0.$$

Denote

$$\omega_t(\epsilon) = \frac{\partial}{\partial t} \Omega_t(-\epsilon, Z.(\epsilon)) = \mathrm{Ad}\, Q_t(-\epsilon)\dot{Z}_t(\epsilon) = e^{-\epsilon\, \mathrm{ad}\, A_0} \circ$$

$$\circ \overrightarrow{\exp} \int_0^t -\epsilon\, \mathrm{ad}\, \dot{Z}_\tau\, d\tau\, \dot{Z}_t(\epsilon);$$

$$\hat{\omega}_t(\epsilon) = e^{-\epsilon\, \mathrm{ad}\, A_0} \circ \overrightarrow{\exp} \int_0^t -\epsilon\, \mathrm{ad}\, \dot{Z}_\tau\, d\tau\, Z_t(\epsilon), \quad \hat{\omega}_0(\epsilon) = A_0.$$

The differentiation of the chronological exponentials gives us

$$\frac{\partial \omega_t}{\partial \epsilon} = \mathrm{ad}\, Y_t(\epsilon) \circ \mathrm{Ad}\, Q_t(-\epsilon)\dot{Z}_t(\epsilon) + \mathrm{Ad}\, Q_t(-\epsilon)\frac{\partial}{\partial \epsilon} \dot{Z}_t(\epsilon), \tag{3.8}$$

where

$$Y_t(\epsilon) = -A_0 - \int_0^t \mathrm{Ad}\, Q_\tau(-\epsilon)\Big(\dot{Z}_\tau(\epsilon) + \frac{\partial}{\partial \epsilon}\frac{\partial}{\partial \tau} Z_\tau(\epsilon)\Big)\, d\tau$$

$$= -A_0 - \int_0^t (\omega_\tau + \epsilon[\hat{\omega}_\tau, \omega_\tau])\, d\tau.$$

Substituting the expression for $Y_t(\epsilon)$ into (3.8), we get

$$\frac{\partial \omega_t(\epsilon)}{\partial \epsilon} = -[A_0, \omega_t] - \left[\left(\int_0^t (\omega_\tau + \epsilon[\hat{\omega}_\tau, \omega_\tau]\right) \mathrm{d}\tau, \omega_t\right] + [\hat{\omega}_t, \omega_t]$$

$$= \left[\hat{\omega}_t - A_0 - \int_0^t (\omega_\tau + \epsilon[\hat{\omega}_\tau, \omega_\tau]) \mathrm{d}\tau, \omega_t\right].$$

On the other hand, differentiating ω_t with respect to t, we get

$$\frac{\partial \omega(\epsilon)}{\partial t} = \omega_t + \epsilon[\hat{\omega}_t, \omega_t].$$

With regard for the equality $\hat{\omega}_0(\epsilon) = A_0$, it implies $\partial\omega_t/\partial\epsilon \stackrel{\epsilon}{=} 0$, so the uniqueness of the geodesic is proved.

Let us now return to Equations (3.4) and (3.5). Let

$$Z_t(\epsilon) = \sum_{i=1}^{\infty} \epsilon^{i-1} \xi_t^i$$

be a Taylor expansion in powers of ϵ for the family of vector fields $Z_t(\epsilon)$; $\xi_t^i \in \operatorname{Der} M$, $i = 1, 2, \ldots$. For a given $t \in R$, let k be the smallest integer, such that $q_0 \circ \xi_t^k \neq 0$, and let us put $q_t(\epsilon) = q_0 \circ Q_t(\epsilon)$. Equation (3.4), being rewritten for the curve $\epsilon \to q_t(\epsilon)$ in M, gives us an equality

$$\frac{\partial}{\partial \epsilon} q_t(\epsilon) = q_0 \circ Z_t(\epsilon) \circ Q_t(\epsilon) = q_0 \circ Q_t(\epsilon) \circ \operatorname{Ad} Q_t^{-1}(\epsilon) Z_t(\epsilon)$$

$$= q_t(\epsilon) \circ \operatorname{Ad} Q_t^{-1}(\epsilon) Z_t(\epsilon).$$

Hence

$$q_t(\epsilon) = q_t(\mathrm{o}) + \int_0^\epsilon q_t(\sigma) \operatorname{Ad} Q_t^{-1}(\sigma) Z_t(\sigma) \,\mathrm{d}\sigma,$$

which implies the relation

$$q_t(\epsilon) = q_t(0) + \frac{\epsilon^k}{k} q_t(0) \circ \operatorname{Ad} Q_t^{-1}(0) \xi_t^k + O(\epsilon^{k+1}). \tag{3.9}$$

Thus, the tangent vector to the curve $\epsilon \to q_t(\epsilon)$ at $\epsilon = 0$ can be expressed in terms of 'non-perturbed' flow $Q_t(0)$ and the field ξ_t^k. In turn, Equation (3.5) gives us the recurrent formulae for successive calculations of the field ξ_t^i, $i = 1, 2, \ldots$ in terms of $Q_\tau(0)$ and Y_τ. In particular, we have

$$\xi_t^1 = Z_0 + \int_0^t \operatorname{Ad} Q_\tau(0) Y_\tau;$$

$$\xi_t^2 = \int_0^t \left[\int_0^\tau \operatorname{Ad} Q_\theta(0) Y_\theta \,\mathrm{d}\theta, \operatorname{Ad} Q_\tau(0) Y_\tau\right] \mathrm{d}\tau + \left[Z_0, \int_0^t \operatorname{Ad} Q_\tau(0) Y_\tau \,\mathrm{d}\tau\right].$$

In order to get a compact notation of the expressions for $\xi_t^i (\forall i)$, we will put into consideration a special sequence of polynomials over a nonassociative variable. Let us consider a free nonassociative algebra $\mathfrak{A}(\lambda)$ with only one generator λ. Quotient algebra $\mathfrak{A}(\lambda)$ modulo an ideal, generated by the elements of the form

$$x(yz)-y(xz)-(xy-yx)z, \tag{3.10}$$

will be called a free chronological algebra with generator λ and will be denoted Λ. Algebra Λ inherits the natural graduation of $\mathfrak{A}(\lambda)$, $\deg\lambda=1$. (For details of the matter concerning chronological algebras, particularly the construction of a basis in a free chronological algebra, see [2].)

Remember that a derivation of an algebra $\mathfrak{A}$ is an R-linear mapping $\omega\colon \mathfrak{A}\to\mathfrak{A}$, satisfying the Leibnitz rule; $\forall a_1, a_2$ $\omega(a_1a_2)=(\omega a_1)a_2+a_1(\omega a_2)$. Denote Der $\mathfrak{A}$ as the set of derivations of algebra $\mathfrak{A}$.

Apparently, a derivation ω in a free nonassociative algebra $\mathfrak{A}(\lambda)$, with a single generator λ, may be correctly defined by putting down its action on a generator. It can be easily shown that any derivation $\omega\in \operatorname{Der}\mathfrak{A}(\lambda)$ gives rise to some derivation of the chronological algebra Λ. We will denote this derivation as $\hat{\omega}$. In order to prove this fact, we need only satisfy ourselves that the ideal generated by the elements of the kind (3.10) is invariant with respect to ω. Direct computation by Leibnitz rule shows that the result of applying ω to an element of the kind (3.10) is a sum of elements of the same kind. This fact proves the invariance.

Let us define the derivation δ by formula; $\delta\lambda=\lambda\lambda$. The polynomials we are interested in are of the kind

$$\delta_k(\lambda)=\delta^{k-1}\lambda. \tag{3.11}$$

In particular,

$$\begin{aligned}
&\delta_1(\lambda)=\lambda, \qquad \delta_2(\lambda)=\lambda^2, \qquad \delta_3(\lambda)=\lambda^2\lambda+\lambda\lambda^2,\\
&\delta_4(\lambda)=2\lambda^2\lambda^2+(\lambda^2\lambda+\lambda\lambda^2)\lambda+\lambda(\lambda^2\lambda+\lambda\lambda^2)\\
&\qquad\quad =\lambda(3\lambda^2\lambda+\lambda\lambda^2)+(3\lambda^2\lambda-\lambda\lambda^2)\lambda.
\end{aligned}$$

The correspondence $\Gamma\colon \lambda\to Z(\epsilon)$ defines the unique homomorphism of the free chronological algebra Λ in the algebra of derivations of $C^\infty(M)$ with '$*$' as the algebra multiplication. We shall prove that

$$\xi_t^k=\delta_k(Z.(0))=\delta_k\left(Z_0+\int_0^{\cdot}\operatorname{Ad}Q_\theta(0)\,Y_\theta\,d\theta\right), \tag{3.12}$$

where δ_k is a polynomial with respect to '$*$'-multiplication. Indeed, on the one hand $\partial/\partial\epsilon$ and δ are both the derivations (satisfying the Leibnitz rule), and on the other hand $\partial/\partial\epsilon Z=Z*Z$ and $\delta\lambda=\lambda\lambda$, i.e., $\partial/\partial\epsilon$ and δ act similarly on the corresponding generators $Z(\epsilon)$ and λ. Hence $(\partial^k/\partial\epsilon^k)Z(\epsilon)=\Gamma\delta^k(\lambda)$ and

$$\xi_t^k=\frac{\partial^k}{\partial\epsilon^k}\bigg|_{\epsilon=0}Z_t(\epsilon)=\delta_k(Z.(0))=\delta_k\left(Z_0+\int_0^{\cdot}\operatorname{Ad}Q_\theta(0)\,Y_\theta\,d\theta\right)_t.$$

In order to find the tangent vector $\epsilon \to q_t(\epsilon)$ (see (3.9)), it is necessary to act with the differential of the diffeomorphism $Q_t(0)$ on the tangent vector $q_0 \circ \xi_t^k$. It is not difficult to get the recurrent relation for the fields $\operatorname{Ad} Q_t^{-1}(0)\xi_t^i$ directly. Let us denote

$$R_t(\epsilon) = \operatorname{Ad} Q_t^{-1}(0) Z_t(\epsilon); \qquad \rho_t^i = \operatorname{Ad} Q_t^{-1}(0)\xi_t^i.$$

Using Equation (3.5), we get

$$\frac{\partial}{\partial \epsilon}\frac{\partial}{\partial t} Z_t = \left[Z_t, \frac{\partial}{\partial t} Z_t\right],$$

$$\frac{\partial}{\partial \epsilon}\left(\frac{\partial}{\partial t} + \operatorname{ad} X_t\right) R_t = \left[R_t, \left(\frac{\partial}{\partial t} + \operatorname{ad} X_t\right) R_t\right],$$

$$\left(\frac{\partial}{\partial t} + \operatorname{ad} X_t\right) R_t(0) = Y_t,\ R_0(0) = Z_0.$$

One can see that the equation for R_t differs from the equation for Z_t in the detail that, in this case the operator $(\partial/\partial t + \operatorname{ad} X_t)$ is employed in place of $\partial/\partial t$. The coefficients of the expansion of R_t into the powers of ϵ satisfy the equations

$$\left(\frac{\partial}{\partial t} + \operatorname{ad} X_t\right)\rho_t^1 = Y_t, \qquad \rho_1^0 = Z_0;$$

$$\left(\frac{\partial}{\partial t} + \operatorname{ad} X_t\right)\rho_t^2 = [\rho_t^1, Y_t], \qquad \rho_0^2 = 0, \quad \text{etc.}$$

Hence

$$\rho_t^1 = \operatorname{Ad} Q_t^{-1} Z_0 + \int_0^t \left(\operatorname{Ad} \overleftarrow{\exp} \int_\tau^t - X_\theta \, d\theta\right) Y_\tau \, d\tau,$$

$$\rho_t^2 = \int_0^t \operatorname{Ad}\left(\overleftarrow{\exp} \int_\tau^t - X_\theta \, d\theta\right)[\rho_\tau^1, Y_\tau] \, d\tau, \quad \text{etc.}$$

Let us introduce a more detailed consideration the formula for ξ_t^k, one representing the field ξ_t^k as a polynomial with respect to the chronological product '$*$'.

By virtue of (3.9), it is particularly important to calculate the value $q_0 \circ \xi_t^k$ provided that $q_0 \circ \xi_t^i = 0$ for $i < k$. Evidently this value will not vary if we add the arbitrary Lie brackets of $\xi_t^i (i < k)$ to the vector field ξ_t^k. At the same time

$$(A * B)_t - (B * A)_t = [A_t, B_t] - [A_0, B_0] \quad \forall A_\tau, B_\tau,$$

and if $B_0 = 0$, then $A * B - B * A = [A, B]$.

Therefore, the sequence of chronological polynomials we described above is not the only possibility. Actually, if we define the commutator in chronological algebra Λ as follows: $\forall x, y \in \Lambda [x, y] = xy - yx$, then by virtue of what is stated above, any sequence of polynomials $\sigma_1(\lambda), \ldots, \sigma_n(\lambda) \ldots$ such that $\sigma_i(\lambda)$ comes out from $\delta_i(\lambda)$ by adding some bracket polynomial on $\delta_1, \ldots, \delta_{i-1}$, is adequate.

The sequences of polynomials different from (3.12) are mentioned in [2, 3] (see, also [6]). Let us consider one more example (obtained by A. I. Tretijak) of such a sequence, consisting of chronological *monomials*.

EXAMPLE 3.1. Consider a sequence of monomials

$$\pi_1 = \lambda,\ \pi_k = \pi_{k-1}\lambda, \quad k = 2, 3, \ldots;\ \pi_k(\lambda) = \ldots((\lambda\lambda)\lambda)\ldots\lambda).$$

Then the following equality holds

$$\delta\pi_k = k\pi_{k+1} + [\lambda, \pi_k] \tag{3.13}$$

In fact, for $k = 1$ this equality is obviously fulfilled. Let it be fulfilled for $k = l - 1$, i.e. $\delta\pi_{l-1} = (l-1)\pi_l + [\lambda, \pi_{l-1}]$. Then

$$\begin{aligned}\delta\pi_l &= \delta(\pi_{l-1}\lambda) = (\delta\pi_{l-1})\lambda + \pi_{l-1}(\delta\lambda) = (l-1)\pi_l\lambda + [\lambda, \pi_{l-1}]\lambda + \pi_{l-1}(\lambda\lambda)\\ &= (l-1)\pi_{l+1} + [\lambda, \pi_{l-1}]\lambda + \pi_{l-1}(\lambda\lambda).\end{aligned}$$

Transforming the term $[\lambda, \pi_{l-1}]\lambda$ by means of the identity

$$x(yz) - y(xz) = (xy - yx)z = [x, y]z, \tag{3.14}$$

we get

$$\begin{aligned}\delta\pi_l &= (l-1)\pi_{l+1} + \lambda(\pi_{l-1}\lambda) = (l-1)\pi_{l+1} + (\pi_{l-1}\lambda)\lambda + [\lambda, \pi_{l-1}]\lambda\\ &= l\pi_{l+1} + [\lambda, \pi_l],\end{aligned}$$

so the step of induction is completed.

The equality (3.13) implies

$$\delta_k(\lambda) = (k-1)!\,\pi_k(\lambda) + \mu_k(\lambda), \tag{3.15}$$

where $\mu_k(\lambda)$ is some bracket polynomial on $\delta_1, \ldots, \delta_{k-1}$ (or $\pi_1, \ldots, \pi_{k-1}$). Indeed, if $k = 1$ then $\delta_1 = \pi_1 = \lambda$, and if the equality (3.15) holds for $k = l - 1$, then

$$\delta_l(\lambda) = \delta(\delta_{l-1}(\lambda)) = (l-1)!\,\pi_e(\lambda) + (l-2)![\lambda, \pi_{l-1}(\lambda)] + \delta\mu_{l-1}(\lambda).$$

Obviously, $(l-2)![\lambda, \pi_{l-1}] + \delta\mu_{l-1}$ is a bracket polynomial on $\pi_1, \ldots, \pi_{l-1}$.

The lemma stated below permits us to survey all admissible sequences. Let us again consider a free chronological algebra Λ with graduation defined by posing $\deg \lambda = 1$. The operation of commutation defined as $(x, y) \to xy - yx$, where $x, y \in \Lambda$, brings into Λ a structure of graded Lie algebras; the Jacobi identity ensures from the relation; $x(yz) - y(xz) = (xy - yx)z$. We denote this Lie algebra $[\Lambda]$; as usual $(\operatorname{ad} x)y = xy - yx$, $\forall x, y \in [\Lambda]$.

LEMMA 3.2. *Let $\mathscr{L}$ be a Lie subalgebra in $[\Lambda]$, generated by the elements $\delta_i(\lambda)$,*

$i = 1, 2, \ldots$, and $\mathscr{L}_k$ be a set of all homogeneous elements of degree k in $\mathscr{L}$. The

$$\mathscr{L}_2 = \mathrm{span}\{\lambda^2\}, \; \mathscr{L}_{k+1} = \delta\mathscr{L}_k + (\mathrm{ad}\,\lambda)\mathscr{L}_k, \quad k \geq 2.$$

Proof. Evidently, $\mathscr{L}$ is a minimal Lie subalgebra of $[\Lambda]$, containing λ and invariant relative to the derivation δ. This fact instantly implies the inclusion $\mathscr{L}_{k+1} \supseteq \delta\mathscr{L}_k + (\mathrm{ad}\,\lambda)\mathscr{L}_k$.

In order to get inverse inclusion, let us consider the vector space $[\Lambda] \oplus R\delta$ and provide a graduated Lie algebra structure on it assuming

$$\deg \delta = 1 \quad \text{and} \quad -[x, \delta] = [\delta, x] = \delta x, \quad \forall x \in [\Lambda].$$

The Jacobi identity arises from the Leibnitz rule.

$$[[\delta, x]y] + [[x, y]\delta] + [[y, \delta]x] = [\delta x, y] - \delta[x, y] + [x, \delta y] = 0.$$

Obviously, the Lie subalgebra of $[\Lambda] + R\delta$, one generated by λ and δ, coincides with $\mathscr{L} + R\delta$, and its homogeneous component of the kth degree coincides with $\mathscr{L}^k$. It follows from the standard properties of Lie algebras (see [7, §1.IV.8]) that any element of $(\mathscr{L} + R\delta)_{k+1}$ may be presented as

$$(\mathrm{ad}\,\lambda)x_k + (\mathrm{ad}\,\delta)y_k = \mathrm{ad}\,\lambda x_k + \delta y_k$$

for some $x_k, y_k \in \mathscr{L}_k$, so the inclusion $\mathscr{L}_{k+1} \supseteq \delta\mathscr{L}_k + (\mathrm{ad}\,\lambda)\mathscr{L}_k$ is proved.

Let us now turn to the control system (2.1).

PROPOSITION 3.3 *Suppose that for a given $t \in R$ and $u(\cdot) \in L_\infty[\mathrm{o}, t]$ the equalities $q_0 \circ \delta_i(\int_0^\cdot h_\tau u(\tau)\,\mathrm{d}\tau)_t = 0$, $i = 1, \ldots, k-1$ hold. Then*

$$G_t(u(\cdot)) = q_0 + q_0 \circ \frac{(k-1)!}{k}\, \pi_k\Big(\int_0^\cdot h_\tau u(\tau)\,\mathrm{d}\tau\Big)_t + O\Big(\int_0^t |u(\tau)|\,\mathrm{d}\tau\Big)^{k+1} \tag{3.16}$$

and

$$G_t(u(\cdot)) = q_0 + q_0 \circ \frac{1}{k}\, \delta_k\Big(\int_0^\cdot h_\tau u(\tau)\,\mathrm{d}\tau\Big)_t + O\Big(\int_0^t |u(\tau)|\,\mathrm{d}\tau\Big)^{k+1}, \tag{3.17}$$

(δ_k, π_k are the polynomials with respect to chronological multiplication '$*$').

Formulae (3.16) and (3.17) arise from (3.9), (3.12) and (3.15) (here $\int_0^t |u(\tau)|\,\mathrm{d}\tau$ plays the part of small parameter ϵ) with regard for $Q_\tau(0) = \mathrm{Id}$, because the right-hand side of (2.1) vanishes when $u = 0$.

Thus, we have ascertained that the kth derivative of the mapping G_t in the case when the previous derivatives are null, can be expressed by means of brackets of the vector fields h_τ, $\tau \in [0, t]$; moreover, an explicit formula was presented above.

4. Flags in Tangent Space and Nilpotent Lie Algebras

The results of this section are, to some extent, inspired by the papers of Hermes and Sussmann (see [8, 9], especially by Sussmann's proof of one hypothesis of Hermes.

Let $0 = E^0 \subset E^1 \subseteq E^2 \subseteq \cdots \subseteq E^l = T_{q_0}M$ be an arbitrary flag in $T_{q_0}M$; $0 < \dim E_1 \leqslant \cdots \leqslant \dim E^l = n$. In other respects, the dimensions of E^i are arbitrary, some of them may coincide, and the case $n < l$ is not excluded. Let us put $k_i = \dim(E^i/E^{i-1})$, $i = 1, \ldots,$ and let $R^n = R^{k_1} \oplus R^{k_2} \oplus \cdots \oplus + R^{k_l}$ be a standard representation of R^n as a direct sum (R^{k_1} is a span of the first k_1 vectors from standard basis, R^{k_2} – the span of the next k_2 vectors, etc.). Let $M \supset 0_{q_0}$ be a coordinate neighborhood of q_0, and smooth vector-functions $x_i = (x_{i1}, \ldots, x_{ik_i})^T : 0_{q_0} \to R^{k_i}$; $i - 1, \ldots,$ be such that the mapping $x = (x_1, \ldots, x_l)^T : 0_{q_0} \to R^n$ defines the local coordinates on M which meets the conditions

$$x(q_0) = 0, \qquad x_*|_{q_0} E^i = R^{k_1} \oplus \cdots \oplus R^{k_i} \subseteq R^n.$$

Any differential operator localized on 0_{q_0} takes in these coordinates a form

$$\sum_\alpha \varphi_\alpha(x) \frac{\partial^{|\alpha|}}{\partial x^\alpha}, \quad \text{where } \varphi_\alpha(x) \in C^\infty(R^n),$$

α is a multi-index:

$$\alpha = (\alpha_1, \ldots, \alpha_l),\ \alpha_i = (\alpha_{i_1}, \ldots, \alpha_{ik_i}),$$

$$|\alpha_i| = \sum_{j=1}^{k_i} \alpha_{ij}, \quad i = 1, \ldots, l;\ |\alpha| = \sum_{i=1}^{l} |\alpha_i|.$$

The space $D(R^n)$ of all differential operators on R^n may be considered as an associative algebra with the composition of operators as a product. Differential operators with polynomial coefficients (all of $\varphi_\alpha(x)$ are polynomials) form a subalgebra of $D(R^n)$ with 1, x_{ij}, $\partial/\partial x_{ij}$, $i = 1, \ldots, l$, $j = 1, \ldots, k_i$ as generators. Let us introduce a $\mathbb{Z}$-graduation in this subalgebra by setting the weight 'ν' of the generators

$$\nu(1) = 0, \qquad \nu(x_{ij}) = i, \quad \nu\left(\frac{\partial}{\partial x_{ij}}\right) = -i.$$

Respectively,

$$\nu\left(x^\alpha \frac{\partial^{|\beta|}}{\partial x^\beta}\right) = \sum_{i=1}^{l} (|\alpha_i| - |\beta_i|)i,$$

where α, β are multi-indices. A differential operator with polynomial coefficients is called homogeneous of the weight r, if all the monomials contained in it have weight r. It can be easily seen that $\nu\ (D_1D_2) = \nu(D_1) + \nu(D_2)$ for any homogeneous differential operators D_1, D_2. As long as the vector fields are differential operators, then all the facts mentioned above are valid for them. Certainly, $\nu([X_1, X_2]) = \nu(X_1) + \nu(X_2)$ for any two homogeneous fields X_1, X_2 on R^n. Let us also note that the weight of the differential operator of Nth order is not less than $(-Nl)$; in particular, the weight of nonzero vector fields is not less than $(-l)$.

Let us introduce a special notation for the linear span of the set of homogeneous vector fields with negative weights by putting

$$V^{-}(k_1, \dots, k_l) = \operatorname{span}\left\{x^{\alpha} \frac{\partial}{\partial x_{ij}} \,\middle|\, \nu(x^{\alpha}) < i, \quad 1 \leq j \leq k_i\right\}.$$

It can be easily seen that $V^{-}(k_1, \dots, k_l)$ is a nilpotent Lie subalgebra (of length l) of the Lie algebra $\operatorname{Der}(R^n)$. Nilpotency is due to the fact that the (negative) weights of the fields are added when we make them commute.

Let

$$D = \sum_{\alpha} \varphi_{\alpha}(x) \frac{\partial^{|\alpha|}}{\partial x^{\alpha}}$$

be an arbitrary differential operator optionally with polynomial coefficients. The Taylor expansions of $\varphi_{\alpha}(x)$ in powers of x_{ij} after grouping of the monomials with identical weights gives us the following representation

$$D \sim \sum_{r=-Nl}^{+\infty} D^{(r)}, \tag{4.1}$$

where $D^{(r)}$ is a homogeneous differential operator of the weight r, which we shall call a 'weight r'-derivative of D. Representation (4.1) makes it possible to define a descending filtration in the algebra $D(R^n)$ by putting

$$\mathscr{D}^{r}(k_1, \dots, k_l) = \{D \in \mathscr{D}(R^n) \mid D^{(i)} = 0, \quad \text{when } i < r\}.$$

We have evidently

$$\mathscr{D}^{r_1}(k_1, \dots, k_l) \subset \mathscr{D}^{r_2}(k_1, \dots, k_l), \quad \text{when } r_2 < r_1,$$

and

$$\mathscr{D}^{r_1}(k_1, \dots, k_l) \circ \mathscr{D}^{r_2}(k_1, \dots, k_l) \subseteq \mathscr{D}^{r_1 + r_2}(k_1, \dots, k_l), \quad \forall r_1, r_2.$$

We shall say that a nonzero operator $D \in D(R^n)$ has a weight equal to r, if $D \in D^{r}(k_1, \dots, k_l) \backslash D^{r+1}(k_1, \dots, k_l)$.

Let us set also

$$\operatorname{Der}^{r}(k_1, \dots, k_l) = \operatorname{Der} R^n \cap D^{r}(k_1, \dots, k_l);$$

clearly $\operatorname{Der}^{r}(k_1, \dots, k_l) = \operatorname{Der} R^n$ when $r < -l$, and

$$[\operatorname{Der}^{r_1}(k_1, \dots, k_l), \operatorname{Der}^{r_2}(k_1, \dots, k_l)] \subseteq \operatorname{Der}^{r_1 + r_2}(k_1, \dots, k_l) \quad \forall r_1, r_2.$$

There exists a trivial isomorphism between the graded algebra

$$\bigoplus_{r} (D^{r}(k_1, \dots, k_l) / D^{r+1}(k_1, \dots, k_l))$$

and the algebra of all differential operators with polynomial coefficients, being graded by means of the weight 'ν'. Restricting ourself to the vector fields, we get an isomorphism between the graded Lie algebra

$$\bigoplus_{r=-l}^{+\infty} (\mathrm{Der}^r(k_1, \ldots, k_l)/\mathrm{Der}^{r+1}(k_1, \ldots, k_l)$$

and the Lie algebra of polynomial vector fields on R^n with appropriate graduation. Considering only the terms with $r<0$, we shall get the isomorphism

$$\bigoplus_{r=-l}^{-1} (\mathrm{Der}^r(k_1, \ldots, k_l)/\mathrm{Der}^{r+1}(k_1, \ldots, k_l)) \approx V^-(k_1, \ldots, k_l),$$

which is especially important for our purposes. Let us formulate an obvious proposition.

PROPOSITION 4.1. *For any $X \in Der^r(k_1, \ldots, k_l)$ the following inclusions hold:*

$$0 \circ X \in \bigoplus_{i=1}^{-r} R^{k_i}, \quad \text{if } -l \leqslant r < 0;$$

$$0 \circ X = 0, \quad \text{if } r \geqslant 0.$$

This will be more suitable to give the next definition not in coordinate form but directly on manifold M.

Let $\mathrm{Der}\, M \supset \mathscr{X}$ be an arbitrary set of vector fields. Let us define a flag in the tangent space $T_{q_0}M$, setting for $k \geqslant 0$

$$E^k = \mathscr{L}^k_{q_0}(\mathscr{X}) = \{q_0 \circ \mathrm{ad}\, X^1 \circ \cdots \circ \mathrm{ad}\, X^{i-1} X^i \mid X^j \in \mathscr{X}, \quad i \leqslant j \leqslant i,\ i \leqslant k\}.$$

E^k is the linear span of values at q_0 of all bracket polynomials (of degree $\leqslant k$) on elements of $\mathscr{X}$. We will suppose that $E_l = Tq_0M$ and say that $\mathscr{X}$ induces this flag.

EXAMPLE 4.1. Let

$$M = R^n = \bigoplus_{i=1}^{l} R^{k_i},\ \bar{E}^i = \bigoplus_{j=1}^{i} R^{k_j};$$

$\mathrm{Der}^{-1}(k_1, \ldots, k_l)$ is a subspace of the $\mathrm{Der}\, R^n$ we defined above. It follows from Proposition 4.1, that

$$\mathscr{L}^i_0(\mathrm{Der}^{-1}(k_1, \ldots, k_l)) = \bar{E}^i = \bigoplus_{j=1}^{i} R^{k_j}, \quad i = 1, 2, \ldots \tag{4.2}$$

In addition, if $I(x)$ is an identical mapping of R^n onto itself, then

$$0 \circ X^1 \circ \cdots \circ X^i I(x) \in \bar{E}^i, \quad i = 1, 2, \ldots.$$

The set $\mathrm{Der}^{-1}(k_1, \ldots, k_l)$ is a maximal subset of $\mathrm{Der}\, R^n$ possessing property (4.2), i.e., inducing the flag $\bar{E}^1 \subseteq \bar{E}^2 \subseteq \cdots \subseteq \bar{E}_l$. If $\mathscr{X} \supset \mathrm{Der}^{-1}(k_1, \ldots; k_l)$, then for some $i \leqslant l$, $\mathscr{L}^i_0(\mathscr{X}) \supset \bar{E}^i$. We will demonstrate below that there exists an infinite number of maximal subsets of R^n which induce the given flag; the correspondence between them will also be ascertained.

If a family $\mathscr{X}$ is contained in $\mathrm{Der}^{-1}(k_1, \dots, k_l)$ and induces the given flag, then there exists a natural nilpotent approximation of it; just this is valid

PROPOSITION 4.2. *Let*

$$\mathscr{X} \subset Der^{-1}(k_1, \dots, k_l)$$

and

$$\mathscr{L}_0^i(\mathscr{X}) = \bar{E}^i = \bigoplus_{j=1}^{i} R^{k_j}, \quad i = 1, \dots, l.$$

Let us select from Taylor expansion (at 0_{R^n}) of every vector field $X \in \mathscr{X}$ all the terms with weight $\nu = -1$; we will denote $\hat{X}$ the sum of these terms for the given $X \in \mathscr{X}$. If we put $\hat{\mathscr{X}} = \{\hat{X} \mid X \in \mathscr{X}\}$ then

(1) *the family $\hat{\mathscr{X}}$ of vector fields generates nilpotent Lie subalgebra of the Lie algebra of all vector fields;*
(2) $\mathscr{L}_0^i(\hat{\mathscr{X}}) = \bar{E}^i, \quad i = 1, 2, \dots, l;$
(3) *if $X^1, X^2, \dots, X^i \in \mathscr{X}$, then $0 \circ ad\, \hat{X}^1, \dots, ad\, \hat{X}^{i-1}\hat{X}^i \in R^{k_i}$ and*
(3) $0 \circ (ad\, X^1, \dots, ad\, X^{i-1}X^i) = 0 \circ (ad\, \hat{X}^1, \dots, ad\, \hat{X}^{i-1}\hat{X}^i)\ (mod\, \bar{E}^{i-1});$
(4) $0 \circ \hat{X}^1, \dots, \hat{X}^{i-1}\hat{X}^i I(x) \in R^{k_i}$ *and* $0 \circ \hat{X}^1, \dots, \hat{X}^i I(x) = 0 \circ X^1 \circ \cdots \circ X^i I(x)$ $(mod\, \bar{E}^{i-1})$
(5) $\hat{X}^1, \dots, \hat{X}^i I(x) = 0 \quad$ *for $i > l$.*

Proof. The property (1) is obvious as long as the implication (4)$\Rightarrow$(3)$\Rightarrow$(2). The first part of (4) follows easily from the fact that $\hat{X}^1 \circ \cdots \circ \hat{X}^i$ is a homogeneous differential operator of the weight $-i$. If we present every $X^i \in \mathscr{X}$ as a sum $X^i = \hat{X}^i + \tilde{X}^i$, where

$$\tilde{X}^i \in \mathrm{Der}^0(k_1, \dots, k_l) \subseteq D^0(k_1, \dots, k_l),$$

then evidently $X^1 \circ \cdots \circ X^i - \hat{X}^1 \circ \cdots \circ \hat{X}^i$ is a sum of compositions of i vector fields (each being a differential operator of the first order), where in every such composition, all the 'factors' lie in $D^{-1}(k_1, \dots, k_l)$ and at least one of them lies in $D^0(k_1, \dots, k_l)$. It implies that

$$X^1 \circ \cdots \circ X^i - \hat{X}^1 \circ \cdots \circ \hat{X}^i \in D^{-(i-1)}(k_1, \dots, k_l),$$

so

$$0 \circ (X^1 \circ \cdots \circ X^i - \hat{X}^1 \circ \cdots \circ \hat{X}^i) I(x) \in \bar{E}^{i-1}.$$

Property (5) arises from the fact that when $k > l$, the differential operator $\hat{X}^1 \circ \cdots \circ \hat{X}^k$ is homogeneous of weight $-k < -l$, so the result of its action on any x_{ij} (having weight $i \leqslant l < k$) is zero.

Thus, when the family $\mathscr{X}$ of vector fields inducing the given flag is contained in $\mathrm{Der}^{-1}(k_1, \dots, k_l)$, we have shown how to construct an approximation. It was also shown that when we calculate the value of a Lie bracket at zero or a composition (both containing i vector fields) of the fields from $\mathscr{X}$ and their approximations, the results will be identical (mod $\bar{E}^{i-1}$) (modulo the span of values at zero of brackets containing less than i fields).

Is it true that any family $\mathscr{X}$ of vector fields inducing the given flag must be contained in $\mathrm{Der}^{-1}(k_1, \dots, k_l)$? It is evidently true when $l = 2$. Indeed if $\mathscr{L}_0^1(\mathscr{X}) = \bar{E}^1 = R^{k_1}$, $\mathscr{L}_0^2(\mathscr{X}) = \bar{E}^2 = R^{k_1} \oplus R^{k_2}$, then $\mathrm{Der}\, R^n = \mathrm{Der}^{-2}(k_1, k_2)$ and the only vector fields which have the weight -2 are the constant one of a kind $\partial/\partial x_{2j}$, where x_{2j} $(j = 1, \dots, k_2)$ coordinatize R^{k_2}. Obviously vector fields from $\mathscr{X}$ cannot contain a term of this kind provided that $0 \circ X \in R^{k_1}$.

However, the situation is quite different for $l \geqslant 3$; here the principal distinction between the high-order approximations and the second-order ones becomes apparent.

EXAMPLE 4.2. Let

$$R^3 = \{x = (x_1, x_2, x_3) \mid x_1, x_2, x_3 \in R\},$$

$$\bar{E}^1 = \{x \mid x_2 = x_3 = 0\}, \qquad \bar{E}^2 = \{x \mid x_3 = 0\}, \quad \bar{E}^3 = R^3,$$

then $R^3 = R \oplus R \oplus R$ and

$$\mathrm{Der}^{-1}(1, 1, 1) = \left\{ a(x)\frac{\partial}{\partial x_1} + b(x)\frac{\partial}{\partial x_2} + c(x)\frac{\partial}{\partial x_3} \,\Big|\, b(0) = c(0) = \frac{\partial c}{\partial x_1}(0) = 0 \right\}.$$

At the same time, for any $\mu \in R$, the set

$$\mathscr{X}_\mu = \left\{ a(x)\frac{\partial}{\partial x_1} + b(x)\frac{\partial}{\partial x_2} + c(x)\frac{\partial}{\partial x_3} \,\Big|\, b(0) = c(0) = 0, \frac{\partial c}{\partial x_1}(0) = \mu a(0) \right\}$$

satisfies the condition $\mathscr{L}_0^i(\mathscr{X}_\mu) = \bar{E}^i (i = 1, 2, 3)$ and is not contained in $\mathrm{Der}^{-1}(1, 1, 1)$ when $\mu \neq 0$. (*Remark* that $\mathrm{Der}^{-1}(1, 1, 1) = \mathscr{X}_0$). Indeed, it is obvious that $\mathscr{L}_0^1(\mathscr{X}_\mu) = \bar{E}^1$, and if

$$X^i = a_i(x)\frac{\partial}{\partial x_1} + b_i(x)\frac{\partial}{\partial x_2} + c_i(x)\frac{\partial}{\partial x_3} \in \mathscr{X}_\mu \quad (i = 1, 2),$$

then the x_3-component of $0 \circ [X_1, X_2]$ is equal to

$$(X^1(0)c_2)(0) - (X^2(0)c_1)(0) = a_1(0)\frac{\partial c_2}{\partial x_1}(0) - a_2(0)\frac{\partial c_1}{\partial x_1}(0) = 0,$$

so $\mathscr{L}_0^2(\mathscr{X}_\mu) \subseteq \bar{E}^2$. On the other hand, the direct computation shows us that the brackets of three vector fields

$$X^1 = \frac{\partial}{\partial x_1} + x_1\frac{\partial}{\partial x_3}, \qquad X^2 = x_1\frac{\partial}{\partial x_2}, \qquad X^3 = x_2\frac{\partial}{\partial x_3}$$

being contained in $\mathscr{X}_\mu$ induce flag $\bar{E}^i$ $(i = 1, 2, 3)$.

So the family $\mathscr{X}$ inducing the given flag is not necessarily contained in a corresponding subspace $\mathrm{Der}^{-1}(k_1, \dots, k_l) \subset \mathrm{Der}\, R^n$. It is found that, in this case, the method of approximation we described in Proposition 4.2 (the truncation of all terms of nonnegative weight in Taylor expansions of all $X \in \mathscr{X}$) does not imply the desirable results. For example, properties (3) and (4) of Proposition 4.2 do not hold. Moreover, even property (2) may be broken, i.e., the approximating fields may induce a flag which is different from the original one. This effects are demonstrated by the following examples.

EXAMPLE 4.3. Consider the family $\mathscr{X} \subset \operatorname{Der} R^3$

$$\mathscr{X} = \left\{X^1 = \frac{\partial}{\partial x_1} + \mu x_1 \frac{\partial}{\partial x_3},\ X^2 = x_1 \frac{\partial}{\partial x_2},\ X^3 = x_2 \frac{\partial}{\partial x_3} + x_3 \frac{\partial}{\partial x_3}\right\}.$$

Apparently, $\mathscr{X} \subset \mathscr{X}_\mu$ and $\mathscr{X}$ induces the following decomposition of R^3 into the direct sum: $R^3 = R \oplus R \oplus R$ (so the weight of a variable coincides with its number). If, in each field, we truncate the terms with nonnegative weight, we will get 'approximating' fields $\hat{X}^1 = X^1$, $\hat{X}^2 = X^2$, $\hat{X}^3 = x_2\ \partial/\partial x_3$. At the same time, the x_3-component of the value at zero of the Lie brackets $[X^1[X^1, X^3]] = \mu\ \partial/\partial x_3$ and $[\hat{X}^1[\hat{X}^1, \hat{X}^3]] = 0$ differs, i.e., are not equal (mod $\bar{E}^2$).

EXAMPLE 4.4. Let us consider the family $\mathscr{X} = \{X^1, \dots, X^5\} \subset \operatorname{Der} R^5$:

$$X^1 = \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_3} + x_3 \frac{\partial}{\partial x_5}, \qquad X^2 = x_1 \frac{\partial}{\partial x_2},$$

$$X^3 = \frac{x_1^2}{2} \frac{\partial}{\partial x_3} + \frac{x_1^3}{2} \frac{\partial}{\partial x_5}, \qquad X^4 = \frac{x_1^3}{6} \frac{\partial}{\partial x_4},$$

$$X^5 = \frac{x_1^4}{24} \frac{\partial}{\partial x_5} - \frac{x_1^3}{6} \frac{\partial}{\partial x_5} + x_5 \frac{\partial}{\partial x_5}.$$

The direct computation shows us that this family induces in R^5 the flag $\bar{E}^1 \subset \cdots \subset \bar{E}^5 = R^5$;

$$\bar{E}^i = \{x \mid x_{i+1} = \cdots = x_5 = 0\}, \quad i = 1, \dots, 4;$$

$\bar{E}^5 = R^5$; or equivalently, the decomposition $R^5 = R \oplus R \oplus R \oplus R \oplus R$ (the weight of a variable coincides with its number). When the terms with nonnegative weights are truncated, we get the following 'approximating' fields

$$\hat{X}^1 = X^1, \quad \hat{X}^2 = X^2, \quad \hat{X}^3 = X^3, \quad \hat{X}^4 = X^4, \quad \hat{X}^5 = \frac{x_1^4}{24} \frac{\partial}{\partial x_5} - \frac{x_1^3}{6} \frac{\partial}{\partial x_5}.$$

The computation shows that the bracket monomial $(\operatorname{ad} \hat{X}^1)^3 \hat{X}^5$ of the fourth degree is equal to $(x_1 - 1)\ \partial/\partial x_5$, so its value at zero does not lie in $\bar{E}^4$. Therefore, the 'approximating' family $\hat{\mathscr{X}}$ induces the decomposition $R^5 = R \oplus R \oplus R \oplus R^2$ which differs from the original.

These examples demonstrate that if $\mathscr{X}$ is not contained in $\operatorname{Der}^{-1}(k_1, \dots, k_l)$, then the method of approximation that we offered above, is not adequate. On the other hand, a family $\mathscr{X}$ inducing the given flag must not be contained in a corresponding subset $\mathrm{Der}^{-1}(k_1, \ldots, k_l)$.

Nevertheless, it proves that $\mathrm{Der}^{-1}(k_1, \ldots, k_l)$ is yet in some way universal among all the maximal families inducing the given flag: any other such family may be mapped onto $\mathrm{Der}^{-1}(k_1, \ldots, k_l)$ by means of some change of coordinates. The coordinate-free formulation of this result is given by the following theorem.

THEOREM 1. *For any subset $\mathscr{X}$ of $\mathrm{Der}\, M$ meeting the conditions $\mathscr{L}^i_{q_0}(\mathscr{X}) = E^i (i > 0)$, $E^l = T_{q_0}M$, there exists a coordinate mapping Φ such that*

$$q_0 \circ \Phi = 0, \ \Phi_*|_{q_0} E^i = \bar{E}^i = \bigoplus_{j=1}^{i} R^{k_j}, \tag{4.3}$$

$$\Phi_* \mathscr{X} = Ad\, \Phi^{-1} \mathscr{X} \subseteq Der^{-1}(k_1, \ldots, k_l). \tag{4.4}$$

This theorem is the most important result of the section; it is analogous, in a sense, to the theorems of linear algebra concerning the simultaneous triangulation of the family of matrices.

The following, very helpful fact seems to confirm the analogy with triangular matrices.

COROLLARY. *Suppose that the conditions of Theorem 1 are satisfied; then for each $i = 1, \ldots, l$, and for arbitrary $X^1, \ldots, X^i$, the following inclusion holds:*

$$q_0 \circ X^1 \circ \cdots \circ X^i \Phi \in \bigoplus_{j=1}^{i} R^{k_j} = \bar{E}^i.$$

Proof of Theorem 1. We will outline the main steps of the proof. Let us first consider an arbitrary coordinate mapping Φ_0 satisfying (4.3) and the corresponding family $\Phi_{0*}\mathscr{X} \subset \mathrm{Der}\, R^n$. Evidently, $\mathscr{L}^i_0(\Phi_{0*}\mathscr{X}) = \bar{E}^i$, $i = 1, \ldots, l$. This fact implies that it is sufficient to prove the theorem for the case, when $M = R^n$, $q_0 = 0$, $E^i = \bar{E}^i$.

It will be convenient to prove the strengthened statement of Theorem 1 with one of its conditions being weakened. We will suppose that the equality $\mathscr{L}^i_0(\mathscr{X}) = \bar{E}^i$ is satisfied only for $1 \leq i \leq l-1$. Our further proof will be carried out by induction with respect to l – the length of decomposition of R^n. When $l = 2$ and $R^n = R^{k_1} + R^{k_2}$, then the vector fields from $\mathscr{X}$ ought to satisfy the condition $\mathscr{L}^1_0(\mathscr{X}) = R^{k_1}$, may contain only the following kinds of monomials with negative weight: $\partial/\partial x_{1s}$, $x_{1s}\, \partial/\partial x_{2j}$ $(s = 1, \ldots, k_1, j = 1, \ldots, k_2)$. So for $l = 2$, the strengthened variant of Theorem 1 is true.

Let the strengthened variant of Theorem 1 be true for all $l \leq m$ and R^n be decomposed in a direct sum

$$R^n = \bigoplus_{i=1}^{m-1} R^{k_i} + R^{k_m} + R^{k_{m+1}}.$$

Denote V as the direct sum of the last two subspaces and consider the 'reduced' decomposition $R^n = \bigoplus_{i=1}^{m-1} R^{k_i} + V$. Evidently, the set $\mathscr{X}$ also meets the conditions of strengthened Theorem 1 for the reduced decomposition, because $\mathscr{L}_0^i(\mathscr{X}) = \bar{E}^i$ $(i = 1, \ldots, m-1)$.

The reduced decomposition provides a new 'reduced weight' $r\nu$ of the variables, differential operators and vector fields on R^n:

$$r\nu(1) = 0, \qquad r\nu(x_{ij}) = -r\nu\left(\frac{\partial}{\partial x_{ij}}\right) = i, \quad 1 \leqslant i \leqslant m, j = 1, \ldots, k_i,$$

$$r\nu(x_{m+1,j}) = -r\nu\left(\frac{\partial}{\partial x_{m+1,j}}\right) = m.$$

By the induction assumption, there exists some polynomial diffeomorphism Φ such that

$$\mathrm{Ad}\,\Phi^{-1}\mathscr{X} \subset \mathrm{Der}^{-1}(\mathrm{k}_1, \ldots, \mathrm{k}_{m-1}, \mathrm{k}_m + \mathrm{k}_{m+1}),$$

i.e., all the monomials in the Taylor expansions of the fields $Y \in \mathrm{Ad}\,\Phi^{-1}\mathscr{X}$ have the reduced weight $r\nu \geqslant -1$. *Remark* that for any monomial x^α

$$\nu\left(x^\alpha \frac{\partial}{\partial x_{ij}}\right) \geqslant r\nu\left(x^\alpha \frac{\partial}{\partial x_{ij}}\right) \quad \text{if } 1 \leqslant i \leqslant m,$$

and

$$\nu\left(x^\alpha \frac{\partial}{\partial x_{m+1,j}}\right) \geqslant r\nu\left(x^\alpha \frac{\partial}{\partial x_{m+1,j}}\right) - 1; \quad 1 \leqslant j \leqslant k_{m+1}.$$

Thus, by virtue of the induction assumption, we may suppose that the Taylor expansions of the fields $x \in \mathscr{X}$ contain only the terms $x^\alpha\, \partial/\partial x_{ij}$, $i = 1, \ldots, m+1$; $j = 1, \ldots, k_i$, of the weight $\nu \geqslant -1$ and (possibly) the terms $x^\alpha\, \partial/\partial x_{m+1,j}$ with the weight $\nu = -2$.

Thus, according to the induction assumption, any field $X \in \mathscr{X}$ may be presented in the following form;

$$X = \hat{X} + \sum_{j=1}^{k_{m+1}} Q_j(x) \frac{\partial}{\partial x_{j,m+1}} + \tilde{X},$$

where $\hat{X}$ is a sum of all the monomials $x^\alpha\, \partial/\partial x_{ij}$ $(i = 1, \ldots, m;\ j = 1, \ldots, k_i)$ with the weight $\nu = -1$. The next summand joins all the terms with the weight $\nu = -2$ (here $Q_j(x)$ are homogeneous (with respect to weight) polynomials of the weight $\nu = m-1$); $\tilde{X}$ joins all the monomials of the weight $\geqslant 0$ and the monomials $x^\alpha \partial/\partial x_{m+1,j}$ $(j = 1, \ldots, k_{m+1})$ of the weight $\nu = -1$. We will name $\hat{X}$ the regular part, $\sum_j Q_j(x)\, \partial/\partial x_{m+1,j}$ the irregular part, and their sum $\hat{X} + \sum_j Q_j(x)\, \partial/\partial x_{m+1,j}$ the principal part of a vector field $X \in \mathscr{X}$. For the sake of simplicity (in fact, without loss of generality), we will assume in what follows that $k_{m+1} = 1$ and denote y as the single variable of the weight $\nu = m+1$. Let $x = (x_1, \ldots, x_m)$. In this case, a field $X \in \mathscr{X}$ takes the form

$$X = X(x, y) = \hat{X}(x) + Q(x)\frac{\partial}{\partial y} + \tilde{X}(x, y), \tag{4.5}$$

where $Q(x)$ is a homogeneous polynomial of the weight $\nu = m-1$. Denote $\mathscr{X}^-$ as the set of principal parts of all fields X from $\mathscr{X}$. It can be easily seen that $\mathscr{X}^-$ generates some nilpotent graduated Lie algebra $\bar{\mathscr{X}} = \bigoplus_{i=1}^m (\mathscr{X}^-)^m$ being a Lie subalgebra of the algebra of polynomial vector fields.

The R-linear span of all the fields of a kind $Q(x)\,\partial/\partial y$, where $\nu(Q(x)) \leqslant m-1$ is an Abelian ideal in the Lie algebra $V^-(k_1, \ldots, k_l)$ of the polynomial vector fields with negative weights. Direct computation shows that the Lie bracket of i vector fields X of (4.5), has the same kind:

$$X^{(i)} = \hat{X}^{(i)} + Q^{(i)}(x)\frac{\partial}{\partial y} + \tilde{X}^{(i)}(x, y),$$

where $\hat{X}^{(i)} + Q^{(i)}(x)\,\partial/\partial y \in (\mathscr{X}^-)^i \subset \bar{\mathscr{X}}$, $\hat{X}^{(i)}$ is a homogeneous field of the weight $-i$ and $Q^{(i)}$ is the homogeneous polynomial of the weight $m-i$; $\tilde{X}^{(i)}$ consists of monomials $x^\alpha\,\partial/\partial x_{ij}$ having weights $> -i$ and the monomials $x^\alpha\,\partial/\partial y$ of the weight $\geqslant -i$. Evidently, the values at zero of the vector field $X^{(i)}$ and its principal part $\hat{X}^{(i)} + Q^{(i)}(x)\,\partial/\partial y$ are equal (modulo $\bar{E}^{i-1}$). Further, we will ignore the term $\tilde{X}$ in (4.5), dealing with the family $\mathscr{X}^-$ and the Lie algebra $\bar{\mathscr{X}} = \bigoplus_{i=1}^m (\mathscr{X}^-)^i$. Obviously, $\mathscr{L}_0^i(\mathscr{X}^-) = \bar{E}^i$, $i = 1, \ldots, m$.

Let $\mathbb{P}[x]$ be the ring of polynomials (on $x = (x_1, \ldots, x_m)$. Let us construct a $\mathbb{P}[x]$-module $\mathfrak{M}$ generated by the vector fields from $\bar{\mathscr{X}}$. Evidently, $\mathfrak{M}$ is a Lie algebra; any vector field $X \in \mathfrak{M}$ may be presented in the form $\sum_j P_j^{(s_j)} Z_j^{(r_j)}$, where $P_j^{(s_j)}$ is a homogeneous (of the weight s_j) polynomial and $Z_j^{(r_j)}$ is a field from $(\mathscr{X}^-)^{r_j}$. We call $X \in \mathfrak{M}$ homogeneous with the regular weight i iff, in corresponding expression, $\sum_j P_j^{(s_j)} Z_j^{(r_j)}$ for all j: $s_j - r_j = i$. For any $X \in \mathfrak{M}$, we will call the sum of monomials of X, containing $\partial/\partial y$, the irregular part of X, while the regular part of X is the sum of the rest monomials.

LEMMA 4.3. *Let $Z^1, \ldots, Z^s$ be some vector fields from $\mathfrak{M}$, and the regular weights of monomials they contain are $\geqslant -i_1, \ldots, \geqslant -i_s$ correspondingly $(i_1 + \cdots + i_s = i)$. Then $0 \circ (\operatorname{ad} Z^1 \ldots \operatorname{ad} Z^{s-1} Z^s) \in \bar{E}^i$, if $i > 0$, and $0 \circ (\operatorname{ad} Z^1 \ldots \operatorname{ad} Z^{s-1} Z^s) = 0$ if $i \leqslant 0$.*

LEMMA 4.4. *For any i, j, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant k_i$; the module $\mathfrak{M}$ contains some field Z_{ij}, having $\partial/\partial x_{ij}$ as its regular part.*

Let us define for any $Z \in \mathfrak{M}$ the projections π^x and π^y of this vector field on its regular and irregular parts

$$\forall Z = \hat{Z} + Q\frac{\partial}{\partial y}: \quad \pi^x Z = \hat{Z}, \quad \pi^y Z = Q\frac{\partial}{\partial y}.$$

The following important lemma takes place.

LEMMA 4.5. *Ker* $\pi^x \subseteq$ *Ker* π^y, *i.e., there is no vector field* $Z \in \mathfrak{M}$ *with zero regular and nonzero irregular parts.*

Proof. Suppose by contradiction that such a field $Z \in \mathfrak{M}$ exists; then it may be presented as a sum of homogeneous (with respect to the regular weight) vector fields. Evidently, the regular part of every summand must be zero. Thus, we may assume without loss of generality the existence of a homogeneous (of regular weight i) vector field Z with zero regular and nonzero irregular parts, i.e., $Z = Q(x)\,\partial/\partial y$, $\nu(Q) = m - i$.

Let us choose in Q an arbitrary monomial $x_{i_1 j_1}, \dots, x_{i_s j_s} (i_1 + \cdots + i_s = m - i)$. By virtue of Lemma 4.4, for any $r = 1, \dots, s$ there exists some homogeneous vector field Z of regular weight $-i_r$ whose regular part is $\partial/\partial x_{i_r j_r}$. It can be easily shown that

$$0 \circ \left(\operatorname{ad} Z^1 \circ \cdots \circ \operatorname{ad} Z^r\left(Q(x)\frac{\partial}{\partial y}\right)\right) = \left(0 \circ \frac{\partial^s Q}{\partial x_{i_1 j_1}, \dots, \partial x_{i_s j_s}}\right)\frac{\partial}{\partial y}.$$

On the one hand, the summarized regular weight of the vector fields entering this Lie bracket is equal to $-i_1 - \cdots - i_s - i = -(m - i) - i = -m$, on the other hand, the value at zero of this bracket does not lie in E^m as long as

$$0 \circ \frac{\partial^s Q}{\partial x_{i_1 j_1}, \dots, \partial x_{i_s j_s}} \neq 0$$

and so we have obtained a contradiction to Lemma 4.3.

Thus Ker $\pi^x \subseteq$ Ker π^y and, therefore, there exists a $\mathbb{P}[x]$-linear mapping Λ: $\pi^x(\mathfrak{M}) \to \pi^y(\mathfrak{M})$ such that $\pi^y(Z) = \Lambda(\pi^x(Z))$.

By virtue of Lemma 4.4, $\pi^x(\mathfrak{M})$ coincides with the algebra of all polynomial vector fields (on $\bigoplus_{i=1}^m R^{k_i}$) and Λ maps the field from $\pi^x(\mathfrak{M})$ into the set of fields $\{Q(x)\,\partial/\partial y\}$. Let us consider a $\mathbb{P}[x]$-linear mapping Ω which maps $\pi^x(\mathfrak{M})$ (i.e., the set of all polynomial vector fields) into $\mathbb{P}[x]$. For any $X \in \mathfrak{M}$,

$$X = \hat{X} + Q(x)\frac{\partial}{\partial y}, \qquad \Omega(\hat{X}) = Q(x).$$

By virtue of the above statement, Ω is a 1-form with polynomial coefficients on $\bar{E}^m$.

LEMMA 4.6. *The* 1-*form* Ω *is closed*: $d\Omega = 0$.

Proof. If the vector fields

$$Z^j = \hat{Z}^j + Q^j(x)\frac{\partial}{\partial y} \in \mathfrak{M} \quad (j = 1, 2),$$

then by definition

$$d\Omega(\hat{Z}^1, \hat{Z}^2) = \hat{Z}^1(\Omega(Z^2)) - \hat{Z}^2(\Omega(Z^1)) - \Omega([\hat{Z}^1, \hat{Z}^2]).$$

Let us calculate $\Omega([Z^1, Z^2])$. As far as the Lie bracket is equal to

$$[Z^1, Z^2] = \left[\hat{Z} + Q^1 \frac{\partial}{\partial y}, \hat{Z}^2 + Q^2 \frac{\partial}{\partial y}\right] = [\hat{Z}^1, \hat{Z}^2] + ((\hat{Z}^1 Q^2) - (\hat{Z}^2 Q^1)) \frac{\partial}{\partial y},$$

so

$$\Omega([\hat{Z}^1, \hat{Z}^2]) = \hat{Z}^1 Q^2 - \hat{Z}^2 Q^1 = \hat{Z}^1(\Omega(\hat{Z}^2)) - \hat{Z}^2(\Omega(\hat{Z}^1)),$$

i.e.,

$$\mathrm{d}\Omega(\hat{Z}^1, \hat{Z}^2) = 0.$$

By virtue of the Poincaré lemma, the closed 1-form is strict, $\Omega = \mathrm{d}\tilde{\Phi}$. As far as Ω has polynomial coefficients, so $\tilde{\Phi} \in \mathbb{P}[x]$. If $X = \hat{X} + Q(x)\,\partial/\partial y$, and $\hat{X}$ is homogeneous, then as it was stated above $\nu\,(Q(x)) - \nu(\hat{X}) = m$, and since $\hat{X}\tilde{\Phi} = Q(x)$, so $\nu(\tilde{\Phi}) = \nu(Q) - \nu(\hat{X}) = m$. Thus, $\tilde{\Phi}$ is a homogeneous polynomial of weight m. So we have proved.

LEMMA 4.7. *There exists a homogeneous (of weight m) polynomial $\tilde{\Phi}$ such that all the vector fields $X \in \mathscr{X}$ (where $\mathscr{X}$ satisfies the conditions of Theorem* 1 *and induction assumption, i.e. so all X's are of the kind* (4.5)) *have a form*

$$X(x, y) = \hat{X}(x) + (\hat{X}\tilde{\Phi}) \frac{\partial}{\partial y} + \tilde{X}(x, y).$$

Concluding the proof of the theorem, let us define a polynomial diffeomorphism Φ: $R^n \to R^n$ by the formula

$$\Phi: (x, y) \to (x, y - \tilde{\Phi}(x)), \qquad \Phi^{-1}: (x, y) \to (x, y + \tilde{\Phi}(x))$$

and prove that this diffeomorphism is the one sought.

It can be easily seen that for $X = \hat{X} + (X\tilde{\Phi})\,\partial/\partial y + \tilde{X}(x, y)$:

$$\Phi_* \hat{X} = \mathrm{Ad}\,\Phi^{-1} \hat{X} = \hat{X} - (\hat{X}\tilde{\Phi}) \frac{\partial}{\partial y},$$

$$\Phi_*\left((\hat{X}\tilde{\Phi}) \frac{\partial}{\partial y}\right) = (\hat{X}\tilde{\Phi}) \frac{\partial}{\partial y}, \qquad \Phi_* X = \hat{X} + \Phi_* \tilde{X}.$$

On the other hand $\mathrm{Ad}\,\Phi^{-1} = \mathrm{e}^{\mathrm{ad}\,Z}$, where the vector field $Z = -\tilde{\Phi}(x)\,\partial/\partial y$ has weight $\nu = -1$. Let us consider an R-linear span of all monomials $x^\alpha\,\partial/\partial x_{ij}$ $i = 1, \dots, m$ of the weight $\nu \geqslant 0$ together with the nomials $x^\alpha\,\partial/\partial y$ of the weight $\nu \geqslant -1$. It is easy to ascertain that this span is a Lie subalgebra of $\mathrm{Der}^{-1}(k_1, \dots, k_{m+1})$. Since Z and $\tilde{X}$ lie in this subalgebra, then $\mathrm{e}^{\mathrm{ad}\,Z}\tilde{X}$ also lies in it, so $\mathrm{e}^{\mathrm{ad}\,Z}\tilde{X} \in \mathrm{Der}^{-1}(k_1, \dots, k_{m+1})$. Thus, $\forall X \in \mathscr{X}$

$$\Phi_* X = \hat{X} + \Phi_* \tilde{X} = \hat{X} + \mathrm{e}^{\mathrm{ad}\,Z}\tilde{X} \in \mathrm{Der}^{-1}(k_1, \dots, k_{m+1})$$

and the induction step, as well as the proof of theorem 1, is concluded.

Let us describe a recurrence algorithm for constructing, for a given set $\mathscr{X} \subset \operatorname{Der} M$, meeting the conditions of Theorem 1, a coordinate mapping Φ satisfying (4.3)–(4.4).

Let Φ_0 be an arbitrary coordinate mapping of $0_{q_0} \ni q_0$ into R^n satisfying (4.3). We will construct recursively the polynomial diffeomorphism

$$P: R^n \to R^n (0 \circ P = 0,\ P_* |_0 \bar{E}^i = \bar{E}^i, \quad i = 1, \ldots, l)$$

such that the composition $\Phi = \Phi_0 \circ P$ satisfies (4.4).

Obviously, the representation of R^n as $\bigoplus_{j=1}^l R^{kj}$ induces the representation of any field $X \in \operatorname{Der}(R^n)$ as a sum: $X = \sum_{j=1}^l X_j$; the values of X_j lie in R^{kj}. Let us denote $X_{(i)} = \sum_{j=1}^i X_j$; evidently $X_{(1)} = X_1$, $X_{(l)} = X$, and by condition for any $X \in \mathscr{X}$, $X_{(1)} = X_1 \in \operatorname{Der}^{-1}(k_1, \ldots, k_l)$. The construction of polynomial mapping P is carried out inductively, namely, assuming that $\forall X \in \mathscr{X}$ $X_{(i-1)} \in \operatorname{Der}^{-1}(k_1, \ldots, k_l)$ and $X_{(i)} \in \operatorname{Der}^{-r}(k_1, \ldots, k_l)$, $r > 1$, we construct a diffeomorphism P_{ir} such that $\forall X \in \mathscr{X}$,

$$(\operatorname{Ad} P_{ir}^{-1} X)_{(i-1)} \in \operatorname{Der}^{-1}(k_1, \ldots, k_l), \qquad (\operatorname{Ad} P_{ir}^{-1} X)_{(i)} \in \operatorname{Der}^{-r+1}(k_1, \ldots, k_l).$$

Indeed, by the condition of the theorem, there exists a collection of vector fields

$$Y^{\alpha\beta} (\alpha = 1, \ldots, i-1,\ \beta = 1, \ldots, k_\alpha)$$

possessing the following properties: (1) each $Y^{\alpha\beta}$ is a 'bracket polynomial' of degree α or an R-linear combination of brackets (of order $\leqslant \alpha$) involving fields from $\mathscr{X}$; (2) the values $0 \circ Y^{\alpha\beta} (\beta = 1, \ldots, k_\alpha)$ form a basis in $R^{k\alpha} (\alpha = 1, \ldots, i-1)$. By definition,

$$Y_{(i)}^{\alpha\beta} = Y_{(i-1)}^{\alpha\beta} + \sum_{j=1}^{k_i} Q_j^{\alpha\beta}(x) \frac{\partial}{\partial x_{ij}}.$$

Let us consider for every $Q_j^{\alpha\beta}(x)$, its McLaurin expansion and select the monomials with the weight $i - r - \alpha$; their sum is denoted $\hat{Q}_j^{\alpha\beta}(x)$. Let us denote $\tilde{Y}_{(i-1)}^{\alpha\beta}$ as an image of the field $Y_{(i-1)}^{\alpha\beta}$ under the standard projection

$$\operatorname{Der} R^n \to \operatorname{Der}^{-\alpha+1}(k_1, \ldots, k_l))$$

and put

$$\hat{Y}_{(i-1)}^{\alpha\beta} = Y_{(i-1)}^{\alpha\beta} - \tilde{Y}_{(i-1)}^{\alpha\beta}.$$

Evidently, vectors $0 \circ \hat{Y}_{(i-1)}^{\alpha\beta} (\beta = 1, \ldots, k_\alpha)$ form a basis in $R^{k_\alpha} (\alpha = 1, \ldots, i-1)$ as well as $0 \circ Y_{(i-1)}^{\alpha\beta}$.

Let us define a collection of differential 1-forms ω_j $(j = 1, \ldots, k_i)$ on the space $\bigoplus_{s=1}^{i-1} R^{k_s}$, putting

$$\langle \omega_j, \hat{Y}_{(i-1)}^{(\alpha\beta)} \rangle = \hat{Q}_j^{\alpha\beta}; \quad j = 1, \ldots, k_i,\ \alpha = 1, \ldots, i-1,\ \beta = 1, \ldots, k_\alpha.$$

The following facts prove to be true: (1) the differential 1-forms $\omega_j (j = 1, \ldots, k_i)$ are correctly defined, i.e., they do not depend on the choice of basis fields $Y^{\alpha\beta}$; (2) 1-forms ω_j are closed and, therefore, exact

$$\omega_j = \mathrm{d}F_j(x_{\alpha\beta})(\alpha = 1, \ldots, i-1; \beta = 1, \ldots, k_\alpha);$$

(3) the desired mapping P_{ir} is defined according to the formulae

$$x' = P_{ir}(x), \qquad x' = (x'_1, \ldots, x'_l)^T, \quad x = (x_1, \ldots, x_l)^T, \quad x'_k = x_k,$$

when $k \neq i$; $x'_{ij} = x_{ij} + F_j (j = 1, \ldots, k_i)$.

EXAMPLE 4.5. In the above-mentioned Example 4.2, we considered a family $\mathscr{X}_\mu \subset \operatorname{Der} R^3$:

$$\mathscr{X}_\mu = \Big\{X = a(x)\frac{\partial}{\partial x_1} + b(x)\frac{\partial}{\partial x_2} + c(x)\frac{\partial}{\partial x_3} \mid b(0) = c(0) = 0;$$
$$\frac{\partial c}{\partial x_1}(0) = \mu a(0), \ \mu \text{ is fixed}\Big\}.$$

Evidently $\forall \mu \neq 0$, $\forall X \in \mathscr{X}_\mu$,

$$X_{(2)} = a\frac{\partial}{\partial x_1} + b\frac{\partial}{\partial x_2} \in \operatorname{Der}^{-1}(1, 1, 1),$$

$$X_{(3)} = X \in \operatorname{Der}^{-2}(1, 1, 1) \quad \text{and} \quad X_{(3)} \notin \operatorname{Der}^{-1}(1, 1, 1),$$

if $\partial c/\partial x_1(0) \neq 0$. Let us consider the fields

$$Y = \frac{\partial}{\partial x_1} + \mu x_1 \frac{\partial}{\partial x_3}, \qquad Y' = x_1 \frac{\partial}{\partial x_2}$$

both belonging to $\mathscr{X}_\mu$; $[Y, Y'] = \partial/\partial x_2$. Evidently, the values of the fields

$$Y = \frac{\partial}{\partial x_1} + \mu x_1 \frac{\partial}{\partial x_3} \quad \text{and} \quad Z = [Y, Y'] = \frac{\partial}{\partial x_2}$$

at zero form basis in $E^2 = \{(x_1, x_2, 0) \mid x_1, x_2 \in R\}$. By definition,

$$Y_{(2)} = \hat{Y}_{(2)} = \frac{\partial}{\partial x_1}, \qquad Y_{(3)} = Y \qquad ,$$

$$Q = \hat{Q} = \mu x_1; \qquad Z_{(2)} = Z_{(3)} = Z = \frac{\partial}{\partial x_2}.$$

We define a 1-form ω putting

$$\left\langle \omega, \frac{\partial}{\partial x_1} \right\rangle = \mu x_1, \qquad \left\langle \omega, \frac{\partial}{\partial x_2} \right\rangle = 0$$

and get

$$\omega = \mu x_1 \, \mathrm{d}x_1 = \mathrm{d}\left(\frac{\mu x_1^{t2}}{2}\right).$$

The diffeomorphism

$$\Phi: (x_1, x_2, x_3) \to \left(x_1, x_2, x_3 - \mu \frac{x_1^2}{2}\right)$$

transforms the family $\mathscr{X}_\mu$ to the one

$$\mathscr{X}_\mu^1 = a \frac{\partial}{\partial x_1} + b \frac{\partial}{\partial x_2} + (c - \mu x_1 a) \frac{\partial}{\partial x_3}.$$

Obviously,

$$(c - \mu x_1 a)(0) = 0 \quad \text{and} \quad \frac{\partial}{\partial x_1}(c - \mu x_1 a)(0) = \frac{\partial c}{\partial x_1}(0) - a(0) = 0,$$

i.e., $\mathscr{X}_\mu^1 \subset \mathrm{Der}^{-1}(1, 1, 1)$.

EXAMPLE 4.6. Consider the family of vector fields $\mathscr{X} = \{X^1, \ldots, X^5\}$ we dealt with in Example 4.4:

$$X^1 = \frac{\partial}{\partial x_1} + x_1 \frac{\partial}{\partial x_3} + x_3 \frac{\partial}{\partial x_5}, \qquad X^2 = x_1 \frac{\partial}{\partial x_2},$$

$$X^3 = \frac{x_1^2}{2} \frac{\partial}{\partial x_3} + \frac{x_1^3}{2} \frac{\partial}{\partial x_5}, \qquad X^4 = \frac{x_1^3}{6} \frac{\partial}{\partial x_4},$$

$$X^5 = \frac{x_1^4}{24} \frac{\partial}{\partial x_5} - \frac{x_1^3}{6} \frac{\partial}{\partial x_5} + x_5 \frac{\partial}{\partial x_5}.$$

As already shown, $\mathscr{X}$ induces the decomposition of R^5 into

$$R \oplus R \oplus R \oplus R \oplus R \text{ and } \mathscr{X} \not\subset \mathrm{Der}^{-1}(1, 1, 1, 1, 1)$$

(because the fields X^1, X^3, X^5 contain the terms of weight -2). Omitting interstitial calculations, let us mention the sequence of the diffeomorphisms, mapping $\mathscr{X}$ into $\mathrm{Der}^{-1}(1, 1, 1, 1, 1)$ as a final result:

(1) an operator, Φ_{1*}, where

$$\Phi_1: (x_1, \ldots, x_5) \to \left(x_1, x_2, x_3 - \frac{x_1^2}{2}, x_4, x_5\right)$$

leaves the vector fields $X^2, \ldots, X^5$ fixed and transforms X^1 to

$$X^{1'} = \frac{\partial}{\partial x_1} + \left(x_3 + \frac{x_1^2}{2}\right) \frac{\partial}{\partial x_5}$$

(*Remark*, $X^{1'}$ contains the term of weight -3);

(2) an operator Φ_{2*}, where

$$\Phi_2: (x_1, \ldots, x_5) \to \left(x_1, \ldots, x_4, x_5 - \frac{x_1^3}{6}\right)$$

leaves the vector fields X^2, X^3, X^4 fixed and transforms X^1, X^5 to

$$X^{1''} = \frac{\partial}{\partial x_1} + x_3 \frac{\partial}{\partial x_5}, \qquad X^{5''} = \left(\frac{x_1^4}{24} + x_5\right)\frac{\partial}{\partial x_5};$$

(3) an operator Φ_{3*}, where

$$\Phi_3 : (x_1, \ldots, x_5) \to (x_1, \ldots, x_4, x_5 - x_1 x_3),$$

leaves the fields X^2, X^4 fixed and transforms $X^{1''}$, X^3, $X^{5''}$ to

$$X^{1'''} = \frac{\partial}{\partial x_1}, \qquad X^{3'''} = \frac{x_1^2}{2}\frac{\partial}{\partial x_3}, \qquad X^5 = \left(\frac{x_1^4}{24} + x_5 + x_1 x_3\right)\frac{\partial}{\partial x_5}.$$

Thus, the composition $\Phi_{3*} \circ \Phi_{2*} \circ \Phi_{1*}$ transforms the vector fields from the initial family to some canonical form

$$X^1 = \frac{\partial}{\partial x_1}, \qquad X^2 = x_1 \frac{\partial}{\partial x_2}, \qquad X^3 = \frac{x_1^2}{2}\frac{\partial}{\partial x_3},$$

$$X^4 = \frac{x_1^3}{6}\frac{\partial}{\partial x_4}, \qquad X^5 = \left(\frac{x_1^4}{24} + x_1 x_3 + x_5\right)\frac{\partial}{\partial x_5}.$$

The nilpotent approximations of these fields are

$$\hat{X}^1 = \frac{\partial}{\partial x_1}, \qquad \hat{X}^2 = x_1 \frac{\partial}{\partial x_2}, \qquad \hat{X}^3 = \frac{x_1^2}{2}\frac{\partial}{\partial x_3},$$

$$\hat{X}^4 = \frac{x_1^3}{6}\frac{\partial}{\partial x_4}, \qquad \hat{X}^5 = \left(\frac{x_1^4}{24} + x_1 x_3\right)\frac{\partial}{\partial x_5}.$$

correspondingly.

Theorem 1 can be essentially strengthened. Actually, the set $\mathscr{X}$, meeting the conditions of the theorem, uniquely defines a filtration $\mathscr{X}_{q_0}^0 \subset \mathscr{X}_{q_0}^{-1} \subset \cdots \subset \mathscr{X}_{q_0}^{-l} =$ Der M of the Lie algebra Der M, where $\mathscr{X} \subset \mathscr{X}_{q_0}^{-1}$ and mapping Φ_* transforms this filtration to the filtration

$$\mathrm{Der}^0(k_1, \ldots, k_l) \subset \mathrm{Der}^{-1}(k_1, \ldots, k_l) \subset \cdots \subset \mathrm{Der}^{-l}(k_1, \ldots, k_l) = \mathrm{Der}\, R^n.$$

DEFINITION. Let Der $M \supset \mathscr{X}$ be such that $q_0 \circ \mathscr{X} \neq 0$. Then we define for every $k \geq 0$, the set

$$\mathscr{X}_{q_0}^{-k} = \{Y \in \mathrm{Der}\, M \mid q_0 \circ (\mathrm{ad}\, X^1 \circ \cdots \circ \mathrm{ad}\, X^i) Y \in \mathscr{L}_{q_0}^{i+k}(\mathscr{X}), \forall X^j \in \mathscr{X},$$
$$1 \leq j \leq i,\ i \geq 0\}.$$

Evidently, $\mathscr{X}_{q_0}^{-k}$, $k \geq 0$ are $C^\infty(M)$-submodules of Der M, $\mathscr{X} \subset \mathscr{X}_{q_0}^{-1}$. The image of arbitrary $X \in \mathscr{X}$ under factorization $\mathscr{X}_{q_0}^{-1} \to \mathscr{X}_{q_0}^{-1}/\mathscr{X}_{q_0}^0$ is denoted X_{q_0}, and the whole set $\mathscr{X}_{q_0} = \{X_{q_0} \mid X \in \mathscr{X}\} \subset \mathscr{X}_{q_0}^{-1}/\mathscr{X}_{q_0}^0$ will be called a nilpotentization of the set $\mathscr{X}$ at q_0.

The meaning of the last definition is that the value

$$(q_0 \circ (\operatorname{ad} X^1, \dots, \operatorname{ad} X^{i-1} X^i) + \mathscr{L}^{i-1}_{q_0}(\mathscr{X})) \in \mathscr{L}^i_{q_0}(\mathscr{X}) / \mathscr{L}^{i-1}_{q_0}(\mathscr{X})$$

depend only on

$$X^j_{q_0}, j = 1, \dots i, \forall X^j \in \mathscr{X}, \forall i > 0.$$

THEOREM 1′. *For any set $\mathscr{X} \subset \operatorname{Der} M$ possessing the property*

$$\mathscr{L}^i_{q_0}(\mathscr{X}) = E^i, \forall i > 0,$$

there exists such a coordinate mapping

$$\Phi \colon 0_{q_0} \to R^n, \quad q_0 \circ \Phi = 0, \qquad \Phi_* |_{q_0} E^i = \bar{E}^i = \bigoplus_{j=1}^{i} R^{k_j}, \quad 1 \leqslant i \leqslant l,$$

that

$$\Phi_*(\mathscr{X}^{-k}_{q_0}) = \operatorname{Der}^{-k}(k_1, \dots, k_l), \quad k = 0, 1, \dots.$$

Theorem 1′ implies the concordance of the filtration

$$\mathscr{X}^0_{q_0} \subset \mathscr{X}^{-1}_{q_0} \subset \cdots \subset \mathscr{X}^{-l}_{q_0} = \operatorname{Der} M$$

with the Lie algebra structure of Der M; also

$$[\mathscr{X}^{-i}_{q_0}, \mathscr{X}^{-j}_{q_0}] \subseteq \mathscr{X}^{-i-j}_{q_0}$$

and mapping Φ_* induces the isomorphism of the graded nilpotent Lie algebra

$$\bigoplus_{i=1}^{l} (\mathscr{X}^{-i}_{q_0} / \mathscr{X}^{-i+1}_{q_0}) \approx V^-(k_1, \dots, k_l).$$

Let us denote

$$V^-_{q_0}(\mathscr{X}) = \bigoplus_{i=1}^{l} (\mathscr{X}^{-i}_{q_0} / \mathscr{X}^{-i+1}_{q_0}),$$

$V^-_{q_0}(\mathscr{X})$ will be called a local Lie algebra associated with the set $\mathscr{X}$. Then, $\mathscr{X}_{q_0} \subset \mathscr{X}^{-1}_{q_0} / \mathscr{X}^0_{q_0} \subset V^-_{q_0}(\mathscr{X})$ and

$$\begin{aligned}
&(q_0 \circ \operatorname{ad} X^1 \circ \cdots \circ \operatorname{ad} X^{i-1} X^i) + E^{i-1} \\
&\quad = (q_0 \circ \operatorname{ad} X^1_{q_0} \circ \cdots \circ \operatorname{ad} X^{i-1}_{q_0} X^i_{q_0}) + E^{i-1} \in E^i / E^{i-1}, \\
&\quad \forall X^j \in \mathscr{X}^{-1}_{q_0}, j = 1, \dots, i; \quad i = 1, \dots, l.
\end{aligned}$$

Let us return to the control system (2.1). We denote (see (2.1))

$$\mathscr{H} = \{h_\tau \mid \tau \in [0, t]\} \subset \operatorname{Der} M \quad \text{and} \quad E^i = \mathscr{L}^i_{q_0}(\mathscr{H}).$$

The image of the field h_τ under factorization $\mathscr{H}^{-1}_{q_0} \to \mathscr{H}^{-1}_{q_0} / \mathscr{H}^0_{q_0} \subset V^-_{q_0}(\mathscr{H})$ is denoted as $h_{\tau_{q_0}}$. In Section 3, special chronological monomials π_i, $i = 1, 2, \dots$, were introduced. Let us consider a homogeneous mapping of degree i, which maps $L_\infty[0, t]$ into E^i/E^{i-1} and looks like

$$\Delta^i_{q_0}(h.): u(\cdot) \to \frac{(i-1)!}{i}\left(q_0 \circ \pi_i\left(\left(\int_0^\cdot h_\tau u(\tau)\,d\tau\right)_t\right)\right) + E^{i-1}.$$

With regard to the equality (3.17), it is quite natural to call $\Delta^i_{q_0}(h\cdot)$ an ith *variation of the control system* (2.1) *at* $u(\cdot) = 0$. Evidently, $\Delta^i_{q_0}(h.) = 0$ for $i > l$. It can be easily seen, that $\Delta^i_{q_0}(h.) = \Delta^i_{q_0}(h._{q_0})$, i.e., all the variations of the control system (2.1) depend only on the image of the nonstationary field $h.$ in the local Lie algebra $V^-_{q_0}(\mathscr{H})$. The curve $h_{\tau_{q_0}}$, $\tau \in [0, t]$, situated in local Lie algebra $V^-_{q_0}(\mathscr{H})$, may be called a nilpotenization of system (2.1) at point q_0.

If Φ is a coordinate mapping which exists according to the formulations of Theorems 1 and 1′, then by virtue of isomorphism $\Phi_*: V^-_{q_0}(\mathscr{H}) \to V^-(k_1, \ldots, k_l)$, the curve $h_{\tau_{q_0}}$, $\tau \in [0, t]$ corresponds to the curve $\hat{h}_\tau = \Phi_* h_{\tau_{q_0}}$, lying in the space $\mathrm{Der}^{-1}(k_1, \ldots, k_l)$, the one, consisting of all homogeneous polynomial vector fields on R^n of the weight (-1). Mapping $\hat{G}_t: L_\infty[0, t] \to R^n$ defined as

$$\hat{G}_t(u(\cdot)) = 0 \circ \overrightarrow{\exp} \int_0^t \hat{h}_\tau u(\tau)\,d\tau,$$

is the polynomial of degree $\leq l$ with respect to $u(\cdot)$ (the C^∞ mapping of one Banach space to another is called the polynomial of degree $\leq l$ iff it coincides with its own Taylor expansion of lth order). In addition, all the variations of the system (2.1) and the following one

$$\dot{z} = \hat{h}_\tau(z)u, \quad z(0) = 0 \tag{4.6}$$

at $u(\cdot) = 0$ coincide, i.e.,

$$\Phi_*|_{q_0}\Delta^i_{q_0}(h.) = \Delta^i_0(\hat{h}.), \quad i = 1, \ldots, .$$

Thus, the control system (4.6) is an lth order approximation of the system (2.1) at point q_0 (and also is a kth-order approximation for any $k > l$, by virtue of the equality $\Delta^k_{q_0}(h.) = \Delta^k_0(\hat{h}.) = 0$ for $k > l$).

In order to get an approximation of ith order ($i < l$), it is necessary to project (the values of) the field $\Phi_* h_{\tau_{q_0}}$, defined on $R^n = \bigoplus_{j=1}^l R^{k_j}$, to the subspace $\bar{E}^i = \bigoplus_{j=1}^i R^{k_j}$. The resulting vector field $\hat{h}^i_\tau$, defined on R^n, is tangent to the subspace $\bar{E}_i \subset R^n$, therefore $\bar{E}^i$ is an invariant subspace of the control system $\dot{z} = \hat{h}^i_\tau(z)u$, $z(0) = 0$, $\tau \in [0, t]$, which is the ith order approximation for the system (2.1) at point q_0.

Remark 1. The curve $\hat{h}_\tau$ in contrast to $h_{\tau_{q_0}}$, depends upon a choice of coordinate mapping Φ. The difference between $h_{\tau_{q_0}}$ and $\hat{h}_\tau$ is of the same sort as the one between the k-jet of a smooth function at a point and its Taylor polynomial of kth degree at the same point.

Remark 2. The requirement of the equality $\mathscr{L}^l_{q_0}(\mathscr{H}) = Tq_0M$ is essential for the construction of higher-order approximation, but by virtue of the Nagano–Sussmann theorem, we can assume it to be fulfilled for all the analytic systems and, moreover, for all systems for which the local 'behaviour' is determined by their Taylor expansions.

5. Attainable Sets

As long as the variations of any order of system (4.6) coincide with the corresponding ones of system (2.1), then the attainable sets of nilpotent system (4.6) pretend to be a 'good' local approximation for the attainable sets of system (2.1) (or (1.1)). In this section, we shall give a strict sense to this thesis.

Everywhere below, Φ is the coordinate mapping which appeared in theorem 1.

Let us consider in $R^n = \bigoplus_{i=1}^{l} R^{k_i}$ the coordinates $x = (x_1, \ldots, x_l)$, where $x_i = (x_{i_1}, \ldots, x_{i_{k_i}})$ coordinatize $R^{k_i}(i = 1, \ldots, l)$. The properties of nilpotent approximations (see property (5) of Proposition 4.2) imply that for $k > l$

$$\hat{h}_{\tau_k} u(\tau_k) \circ \cdots \circ \hat{h}_{\tau_1} u(\tau_1) x_i = 0, \quad i = 1, \ldots, l,$$

and, hence, the mapping

$$\hat{G}_t(u(\circ)) = \overrightarrow{\exp} \int_0^t \hat{h}_\tau u(\tau)\, \mathrm{d}\tau I(x)$$

is a polynomial of the lth degree with respect to $u(\cdot)$; concretely

$$\hat{G}_t(u(\cdot)) = \bigoplus_{i=1}^{l} \left(\int_0^t \mathrm{d}\tau_1, \ldots, \int_0^{\tau_{i-1}} \mathrm{d}\tau_i \hat{h}_{\tau_i} u(\tau_i) \circ \cdots \circ \hat{h}_{\tau_1} u(\tau_1) \right) x_i . \tag{5.1}$$

In (5.1), the components of vector function x_i have weight i and the differential operator, acting on it, has weight $-i$; so the result is a vector function of weight 0, i.e., an element of R^{k_i}.

DEFINITION. For any $\epsilon > 0$ an image of $G_t(\hat{G}_t)$ restricted to the ϵ-neighbourhood $\mathcal{U}_\epsilon$ of zero in $L_\infty[0, t]$ will be called a time t ϵ-attainable set of system (2.1) (system (4.6)) from q_0 (from 0) and will be denoted $\mathfrak{A}_t(\epsilon)$ ($\hat{\mathfrak{A}}_t(\epsilon)$).

DEFINITION. A point $q \in \mathfrak{A}_t(\epsilon)$ ($x \in \hat{\mathfrak{A}}_t(\epsilon)$) is regularly attainable for the system (2.1) (for the system (4.6)) if $G_t^{-1}(q) \cap \mathcal{U}_\epsilon$ (correspondingly, $\hat{G}_t^{-1}(x) \cap \mathcal{U}_\epsilon$) contains the regular point of G_t (of $\hat{G}_t$).

DEFINITION. System (2.1) (system (4.6)) is regularly locally controllable at point $q_0 \in M$ (at point $0 \in R^n$), if for any $\epsilon > 0$, point q_0 ($0 \in R^n$) is regularly attainable for system (2.1) (for system (4.6)) by means of some control $u(\cdot) \in \mathcal{U}_\epsilon$.

Let us define the dilatation of R^n as follows:

$$\Delta_\epsilon : (x_1, \ldots, x_l) \rightarrow (\epsilon x_1, \ldots, \epsilon^l x_l).$$

It follows from (5.1) that $\hat{G}_t(\epsilon u(\cdot)) = \Delta_\epsilon \hat{G}_t(u(\cdot))$, this, in turn, implies $\forall \epsilon, \epsilon_1, \epsilon_2$

$$\hat{\mathfrak{A}}_t(\epsilon) = \Delta_\epsilon \hat{\mathfrak{A}}_t(1), \qquad \Delta_{\epsilon_1} \hat{\mathfrak{A}}_t(\epsilon_2) = \hat{\mathfrak{A}}_t(\epsilon_1 \epsilon_2).$$

Let π_i be a projection of R^n on $R^{k_i} (i = 1, \ldots, l)$. If we take the Volterra expansion of chronological exponentials $\overrightarrow{\exp} \int_0^t h_\tau u(\tau)\, d\tau$ and $\overrightarrow{\exp} \int_0^t \hat{h}_\tau u(\tau)\, d\tau$ and compare the terms of the same degrees, then statement (4) of Proposition 4.2, when applied to the families $\mathscr{H}$ and $\hat{\mathscr{H}}$ of the vector fields, implies

$$\mathbb{R}^{k_i} \supset 0 \circ \int_0^t d\tau_1, \ldots, \int_0^{\tau_{i-1}} d\tau_i (\hat{h}_{\tau_i} u(\tau_i) \circ \cdots \circ \hat{h}_{\tau_1} u(\tau_1)) x_i$$

$$= q_0 \circ \int_0^t d\tau_1, \ldots, \int_0^{\tau_{i-1}} d\tau_i (h_{\tau_i} u(\tau_i) \circ \cdots \circ h\tau_1 u(\tau_1)) \Phi \pmod{\bar{E}^{i-1}},$$

and therefore

$$\|pr_i(\Phi(G_t(\epsilon u(\cdot)) - \hat{G}_t(\epsilon u(\cdot))\| = O(\epsilon^{i+1}) \quad (pr_i : \mathbb{R}^n \to \mathbb{R}^{k_i}, \quad i = 1, \ldots, l). \tag{5.2}$$

Let us formulate and prove the following proposition.

PROPOSITION 5.1. *Suppose that the point* $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_l) \in \hat{\mathfrak{A}}_t(1)$ *is normally attainable for system* (4.6). *Then, for all sufficiently small* $\epsilon > 0$, *the point* $q(\epsilon) = \Phi^{-1}(\Delta_\epsilon \bar{x})$ *belongs to* $\mathfrak{A}_t(\epsilon)$ *and is normally attainable for system* (2.1).

COROLLARY. *If* 0_{R^n} *is normally attainable for system* (4.6), *then* $q_0 = \Phi^{-1}(0)$ *is normally attainable for system* (2.1).

Remark. If the conditions of Proposition 5.1 are satisfied, then $\Delta_\epsilon \bar{x} \in \hat{\mathfrak{A}}_t(\epsilon)$, so (if $\bar{x}$ is normally attainable) Φ determines (for sufficiently small $\epsilon > 0$) a correspondence between the points $\Delta_\epsilon \bar{x} \in \hat{\mathfrak{A}}_t(\epsilon)$ and $\Phi^{-1}(\Delta_\epsilon \bar{x}) \in \mathfrak{A}_t(\epsilon)$.

Proof of Proposition 5.1. If $u(\cdot)$ is a regular inverse image of $\bar{x}$ for the mapping $\hat{G}_t$, then $\hat{G}_t(\epsilon u(\cdot)) = \Delta_\epsilon \bar{x}$. Let us prove that $\Delta_\epsilon \bar{x} \in \mathfrak{A}_t(\epsilon)$ if $\epsilon > 0$ is sufficiently small. To this end, we define the family of mappings

$$Q_t^\epsilon = \Delta_{\epsilon^{-1}}(\Phi(G_t(\epsilon u(\cdot))))$$

on a ball $\|u(\cdot)\|_{L_\infty} \leq 1$. With regard to identity $\Delta_{\epsilon^{-1}} \hat{G}_t(\epsilon u(\cdot)) = \hat{G}_t(u(\cdot))$, the equality (5.2) implies

$$\|Q_t^\epsilon - \hat{G}_t\|_{C^1} = O(\epsilon).$$

Consider an equation $\hat{G}_t(u(\cdot)) - \bar{x} = 0$; remember that $u(\cdot)$ is a regular point for $\hat{G}_t$ and $\hat{G}_t(\bar{u}(\cdot)) - \bar{x} = 0$. By virtue of the implicit function theorem, for any $\delta_1 > 0$ there exists $\delta > 0$ such that for any smooth mapping Q: $L_\infty \to R^n$ satisfying $\|Q - G_t\|_{C^1} < \delta$ there exists $u(\cdot)$: $\|u(\cdot) - \bar{u}(\cdot)\| < \delta_1$ satisfying the equality $Q(u(\cdot)) - \bar{x} = 0$. Let us take δ, δ_1 and ϵ_0 so small that, when

$$\|Q - \hat{G}_t\|_{C^1} \leq \delta, \qquad \|u(\cdot) - \bar{u}(\cdot)\| < \delta_1,$$

then the differential $Q'|_{u(\cdot)}$ is surjective, and for any $\epsilon \leqslant \epsilon_0$ $\|Q_t^\epsilon - \hat{G}_t\|_{C^1} < \delta$. So, taking $Q_t^\epsilon(\epsilon \leqslant \epsilon_0)$ as Q, we get the existence of $u_\epsilon(\cdot)$ $(\epsilon u_\epsilon(\cdot))$ being a regular point of Q_t^ϵ (of $\hat{G}_t$) such that $Q_t^\epsilon(u^\epsilon(\cdot)) = \bar{x}$ and, therefore, $\Delta_\epsilon \bar{x} = \Phi(G_t(\epsilon u_\epsilon(\cdot))$, i.e., $\Phi^{-1}(\Delta_\epsilon \bar{x}) \in \mathfrak{A}_t(\epsilon)$ for $\epsilon \leqslant \epsilon_0$.

For further development, we need to suppose that the vector fields f and g, determining system (1.1) (and correspondingly system (2.1)), are real analytic.

Remark that since $\tilde{u}(\cdot)$ must not be analytic, then the vector fields $h_\tau = \operatorname{Ad}\exp\int_0^\tau (f + gu(\theta))\,d\theta g$ are not, in general, analytic. In what follows, the analyticity of f and g is supposed without special stipulations.

When proving the following statement, we will deal with the families of mappings G_t and F_t (see Section 2). *Remark* that the 'point' $u(\cdot)$ is regular for the mapping G_t iff $\tilde{u}(\cdot) + u(\cdot)$ is regular for F_t.

PROPOSITION 5.2. *There exists $\epsilon > 0$ such that the set of regular points of the mapping G_t (and/or $\hat{G}_t$) is open and dense in the ϵ-neighborhood of zero in $L_1[0, t]$.*

Proof. An openness of the set of regular points is obvious. To prove its density, let us suppose for a time that $\tilde{u}(\cdot)$ is real analytic, then

$$h_\tau = \overrightarrow{\exp}\int_0^\tau \operatorname{ad}(f_\theta + gu(\theta))\,d\theta g$$

is an analytic (with respect to τ) family of real analytic vector fields.

By the condition $\mathscr{L}_{q_0}^l[\mathscr{H}] = T_{q_0}M$, obviously it is fulfilled for all q's from some sufficiently small neighborhood $Oq_0 \subset M$ of the point q_0.

Suppose by contradiction, that the set of regular points of G_t is not dense in any neighborhood of zero in $L_1[0, t]$, i.e., for any neighborhood $W \subset L_1[0, t]$ $(0 \in W)$, there exists an open ball $U \subset W$, which consists of critical points of G_t. Without loss of generality, we may assume that

(1) $G_t(U) \subset G_t(W) \subset Oq_0$;

(2) rkG_t' is constant and equal to $r < n = \dim M$ identically on U.

According to the rank theorem, the image $G_t(U)$ is an r-dimensional submanifold $N_t \subset M$: for any $u(\cdot) \in U$

$$q_0 \circ \exp\int_0^t h_\tau u(\tau)\,d\tau \in N_t \subset Oq_0.$$

The family G_τ possesses some property of 'monotonicity' with respect to τ. If $\tau_2 > \tau_1$, $W_1 \subset L_1[0, \tau_1]$, and $W_1^0 \subset L_1[0, \tau_2]$ is the set, consisting of functions $u(\cdot) \in W_1$, being prolonged by zero to the interval $(\tau_1, \tau_2]$, then, obviously, $G_{\tau_1}(W_1) = G_{\tau_2}(W_1^0)$. This implies that if $W \subset L_1[0, t]$, then $G_{\tau_1}(W) \subset G_{\tau_2}(W)$ and for any $u(\cdot) \in W$, $rkG_{\tau_1}'|u(\cdot) \leqslant rkG_{\tau_2}'|u(\cdot)$. By virtue of that stated above and because G_τ depends continuously (in C^1 metrics) on τ, then for $\epsilon > 0$ being sufficiently small and $\theta \in [t - \epsilon, t]$, mapping G_θ have constant rank (equal to r) on

U. So their images are r-dimensional submanifolds $N_\theta \subseteq N_t$. Evidently, $N_{\theta_1} \subseteq N_{\theta_2}$, when $\theta_1 \leq \theta_2$; in particular $N_\theta \supseteq N_{t-\epsilon}$ for $\theta \in [t-\epsilon, t]$.

Since for any $\theta \in [t-\epsilon, t]$ and $u(\cdot) \in U$, the inclusion $G_\theta(u(\cdot)) \in N_t$ is valid, then

$$\frac{\partial}{\partial \theta} G_\theta(u(\cdot)) = G_\theta(u(\cdot)) \circ h_\theta u(\theta)$$

is a tangent vector to N_t at point $G_\theta(u(\cdot))$ and, therefore, for any $\theta \in [t-\epsilon, t]$ the vector field h_θ is tangent to r-dimensional manifold $N_{t-\epsilon}$, being open subset of N_t.

Consider a family of vector fields $\mathscr{H}_\epsilon = \{h_\theta \mid \theta \in [t-\epsilon, t]\}$ and the Lie algebra $\mathrm{Lie}[\mathscr{H}_\epsilon]$ generated by this family. Evidently, all fields from $\mathrm{Lie}[\mathscr{H}_\epsilon]$ are tangent to N_t at every point of $N_{\tau-\epsilon} \subseteq N_t$. If $\bar{q} \in N_{t-\epsilon}$, then the linear space $\mathscr{L}^l_{\bar{q}}[\mathscr{H}_\epsilon]$ is r-dimensional ($r < n$), so there exists a covector $\psi \in T^*_{\bar{q}}M$ ($\psi \neq 0$) annihilating $\mathscr{L}^l_{\bar{q}}[\mathscr{H}_\epsilon]$. Hence, functions

$$\varphi_i(\theta_1, \ldots, \theta_i) = \langle \psi; \bar{q} \circ (\mathrm{ad}\, h_{\theta_1} \circ \cdots \circ \mathrm{ad}\, h_{\theta_{i-1}} h_{\theta_i}) \rangle, \quad i = 1, \ldots, l,$$

vanish identically on cubes $K^i_\epsilon = ([t-\epsilon, t])^i$ and, therefore, due to real analyticity, vanish identically on the whole cubes $K^i = ([0, t])^i$, $i = 1, \ldots, l$. This implies that ψ annihilates the linear space $\mathscr{L}^l_{\bar{q}}[\mathscr{H}]$, and we get a contradiction with the assumption $\forall q \in Oq_0$, $\mathscr{L}^l_q[\mathscr{H}] = T_qM$. This contradiction proves the proposition in the case when $\tilde{u}(\cdot)$ is analytic. We have proved that the regular points of G_t are dense in some sufficiently small neighborhood of zero (if $\tilde{u}(\cdot)$ is analytic) and also that the regular points of F_t are dense in some sufficiently small neighborhood of analytic control $\tilde{u}(\cdot)$.

Let now $\tilde{u}(\cdot)$ be an arbitrary one from $L_\infty[0, t]$, satisfying all the conditions of Proposition 5.2. Then it follows from the results of [1], that for any $i = 1, \ldots, l$ the mappings $\mu_{\tau_1, \ldots, \tau_i}: L_\infty[0, t] \to T_{q_0}M$, where

$$\mu_{\tau_1, \ldots, \tau_i}(u(\cdot)) = q_0 \circ (\mathrm{ad}\, h_{\tau_1}(u(\cdot)) \circ \cdots \circ \mathrm{ad}\, h_{\tau_{i-1}}(u(\cdot)) h_{\tau_i}(u(\cdot))),$$

$$h_\tau = \overrightarrow{\exp} \int_0^\tau \mathrm{ad}(f + gu(\theta))\, d\theta\, g, \qquad \tau_s \in [0, t],\ s = 1, \ldots, i$$

are continuous on $u(\cdot)$ in metrics of $L_1[0, t]$. Hence, there exists sufficiently small neighborhood of zero in $L_1[0, t]$ (denoted W') such that any $u(\cdot) \in (\tilde{u}(\ ) + W') \cap L_\infty[0, t]$ satisfies the condition:

$$\mathrm{Span}\{\mu_{\tau_1, \ldots, \tau_i}(u(\cdot)) \mid \tau_s \in [0, t], s = 1, \ldots, i, i = 1, \ldots, l\} = T_{q_0}M.$$

As is well known, the real analytic functions are dense in $\tilde{u}(\cdot) + W'$ in metrics of $L_1[0, t]$. By virtue of that proved above, every real analytic $u(\cdot) \in \tilde{u}(\cdot) + W'$ possesses a small neighborhood in which the regular points of F_t are dense in metrics of $L_1[0, t]$. Hence, almost all points from $(\tilde{u}(\cdot) + W') \cap L_\infty[0, t]$ are regular ones of F_t, and almost all points from $W' \cap L_\infty[0, t]$ are regular ones of G_t. Proposition 5.2 is proved.

COROLLARY. *The set of regular attainable points is open and dense in* $\mathfrak{A}_t(\epsilon)$ (*in* $\hat{\mathfrak{A}}_t(\epsilon)$) *for all sufficiently small* $\epsilon > 0$.

Let us define a (nonsmooth) homeomorphism Γ: $R^n \to R^n$ by virtue of formula

$$\Gamma(x) = y;\ y_{ij} = |x_{ij}|^{(1-i)/i} x_{ij}, \quad \text{if } x_{ij} \neq 0;\ y_{ij} = 0, \quad \text{if } x_{ij} = 0.$$

Evidently, $\Gamma(\Delta_\epsilon x) = \epsilon\Gamma(x)$. *Remark* also that $\Gamma \circ \hat{G}_t(\epsilon u(\cdot)) = \Gamma \circ \Delta_\epsilon \circ G_t(u(\cdot))) = \epsilon\Gamma(\hat{G}_t(u(\cdot)))$, i.e. the composition $\Gamma \circ G_t$ is homogeneous of degree 1. It implies, that $\Gamma(\hat{\mathfrak{A}}_t(\epsilon)) = \epsilon\Gamma(\hat{\mathfrak{A}}_t(1))$ and image of mappings $\Gamma \circ \hat{G}_t$ is a cone (in general, nonconvex).

DEFINITION. Let $\mathfrak{M}(\epsilon)$ ($\epsilon \geqslant 0$) be a family of subsets of R^n, where $\mathfrak{M}(\epsilon_1) \subseteq \mathfrak{M}(\epsilon_2)$, when $\epsilon_1 < \epsilon_2$, $\mathfrak{M}(0) = 0$. We will say that vector $y \in R^n$ is interior to $\mathfrak{M}(\epsilon)$, if there exists such a neighborhood $O_y \ni y$, that $\{\alpha x \mid x \in O_y, 0 \leqslant \alpha \leqslant \epsilon\} \subset \mathfrak{M}(\epsilon)$ for all $\epsilon > 0$, being sufficiently small.

Our goal is a comparison of the sets of interior vectors of the families $\Phi(\mathfrak{A}_t(\epsilon))$ and $\hat{\mathfrak{A}}_t(\epsilon)$ at point $0 \in R^n$. It is worthwhile, however, to note that when the local controlability is lacking, then the case of $\Phi(\mathfrak{A}_t(\epsilon))$ and $\hat{\mathfrak{A}}_t(\epsilon)$ having no interior vectors, is rather typical (that is, the case when the tangent cones to these sets at zero of R^n are not solid). If this is the case, then the comparison of 'rectified' attainable sets $\Gamma(\Phi(\mathfrak{A}_t(\epsilon))$ and $\Gamma(\hat{\mathfrak{A}}_t(\epsilon))$ is more informative. *Remark*, meanwhile, that if x is an interior vector of $\hat{\mathfrak{A}}_t(\epsilon)$, then $\Gamma(x)$ is an interior vector of $\Gamma(\hat{\mathfrak{A}}_t(\epsilon))$. Besides, $\Gamma(\hat{\mathfrak{A}}_t(\epsilon))$ always has some interior vector, for example, the vectors $\Gamma(x)$, where x is the regularly attainable point of $\hat{\mathfrak{A}}_t(\epsilon)$ having no zero coordinates.

Propositions 5.1 and 5.2 imply the following result (below the words 'almost all' have the meaning 'all from an open dense subset of R^n').

THEOREM 2. *Almost all vectors being interior to the family of sets* $\Gamma(\hat{\mathfrak{A}}_t(\epsilon))$, $\epsilon \geqslant 0$, *are interior to the family* $\Gamma \circ \Phi(\mathfrak{A}_t(\epsilon))$.

In order to get the inverse inclusion, we need to impose a limitation on the 'rate of oscillation' of admissible controls. Let us put

$$\mathcal{U}(\epsilon, c) = \{u(\cdot) \in L_\infty[0, t] \mid \|u(\cdot)\|L_\infty \leqslant \epsilon, \mathrm{Var}_{[0,T]} u(\cdot) \leqslant C\|u(\cdot)\|_{L_1}\}.$$

It can be easily shown that $\mathcal{U}(\epsilon, c)$ is a precompact subset of $L_1[0, t]$. Let us denote $\mathfrak{A}_t(\epsilon, c) = G_t(\mathcal{U}(\epsilon, c))$; $\hat{\mathfrak{A}}_t(\epsilon, c) = \hat{G}_t(\mathcal{U}(\epsilon, c))$.

THEOREM 3. *For any* $c > 0$ *almost all vectors* y, *being interior to the family* $\Gamma \circ \Phi(\mathfrak{A}_t(\epsilon, c))$, $\epsilon \geqslant 0$, *are interior to* $\Gamma(\hat{\mathfrak{A}}_t(\epsilon, c))$, $\epsilon \geqslant 0$.

Remark. Nonrigorously speaking, we may explain the need for imposing the limitations on the rate of oscillations of controls $u(\cdot)$ as follows. When substituting into the Volterra series for $\overrightarrow{\exp} \int_0^t h_\tau u(\tau)\, d\tau$ the controls $u(\cdot)$ being fast-oscillating and small (of order ϵ in L_∞-norm), it may occur that due to fast-oscillating, the value of the ith term of the series is $O(\epsilon^i)$ and then may be majorized by the terms with greater numbers. Thus, the admission of fast-oscillating controls may mix the variations of different orders.

Thus, we have shown that the attainable sets of system (2.1) and its nilpotentization (4.6) 'have similar structures', where the strict sense of the last words is completely defined above.

Remark. After the first version of this paper (A.A. Agrachev, R.V. Gamkrelidze, A.V. Sarychev. Local invariants of smooth control systems. VINITI, 1986, No. 7020-V) was prepared, we got to know of the nice results of G. Stefani and R. M. Bianchini ([10–13]) containing ideas similar to the ones set forth in Sections 4 and 5 of this paper (see also [14]). However, we hope that the approach to nilpotent approximation presented here, is actual nowadays. It should also be noted that adjoining these subjects are some investigations of A. M. Vershik and V. Ja. Gershkovitch on nonholonomic variational problems (see survey [15], containing other references).

References

1. Agrachev, A. A. and Gamkrelidze, R. V.: The exponential representation of flows and chronological calculus, *Math. sbornik* **107** (149) (1978), 467–532; English transl. in *Math. USSR Sbornik*, **35** (1979), 727–785.
2. Agrachev, A. A. and Gamkrelidze, R. V.: The chronological algebras and the nonstationary vector fields, *Itogi Nauki*: *Problemy geometrii*, Vol. 11, VINITI, Moscow, 1980, pp. 135–176; English transl. in *J. Soviet Math*.
3. Agrachev, A. A., Vakhrameev, S. A., and Gamkrelidze, R. V.: The differential-geometric and group-theoretic methods in optimal control theory, *Itogi Nauki*: *Problemy geometrii*, Vol 14, VINITI, Moscow, 1983, pp. 3–56; English transl. in *J. Soviet Math*.
4. Sarychev, A. V.: Stability of mappings of Hilbert space and equivalence of control systems, *Matem sbornik*, **113** (155) (1980), 146–160; English transl. in *Math. USSR Sbornik*.
5. Arnol'd, V. I.: *Mathematical Methods in Classical Mechanics*, Nauka, Moscow, 1974; English transl., Springer-Verlag, Berlin and New York, 1978.
6. Sussmann, H. J.: A Lie-Volterra expansion for nonlinear systems, in *Lect. Notes Contr. and Inf. Sci*. No. 58, 1984, pp. 822–828.
7. Serre, J. P.: *Lie Algebra and Lie Groups*, Benjamin, New York, 1965.
8. Hermes, H.: Control systems which generate decomposable Lie Algebras. *J. Diff. Equations* **44** (1982), 166–187.
9. Sussmann, H. J.: Lie brackets and local controllability: a sufficient condition for scalar-input systems, *SIAM J. Control Optim*. **21**, (1983), 685–713.
10. Stefani, G.: Polynomial approximations to control systems and local controllability, *Proc. 24th IEEE Conf. Decision and Control, Ft. Lauderdale, Flo., Dec*. 11–13, 1985, Vol. 1. IEEE, New York, 1985, pp. 33–38.
11. Stefani, G.: On the local controllability of a scalar-input control systems, Theory and applied nonlinear control systems, Selected Paper *7th Int. Symp. Math. Theory Networks and Systems, Stockholm*, 10–14 *June* 1985'. Elsevier, North Holland, Amsterdam, 1986, pp. 167–179.
12. Stefani, G.: Local properties of nonlinear control systems, Sci. Pap. Inst. Techn. Cybern. Techn. Univ. Wroclow, 1985, No. 29, pp. 219–226.
13. Bianchini, R. M. and Stefani, G.: Sufficient conditions of local controllability, *Proc. 25th Conf. on Decision and Control, Athens, Greece*, 1986, Vol. 2, IEEE, New York, 1986, pp. 967–970.
14. Agrachev, A. A. and Sarychev, A. V.: Filtration of Lie algebra of vector fields and nilpotent approximation of control systems, *Dokl. Akad. Nauk USSR* **295** (1987), 777–781. English transl. in Soviet Math. Dokl., **36** (1988), 104–108.
15. Vershik, A. M. and Gershkovitch, V. Ja.: Nonholonomic dynamic systems. Geometry of distributions and variational problems, *Fundament. napravlenia* **16** (1987), 5–85; English transl. in *Encyclopaedia of Mathematical Sciences*, Vol. 16, Springer-Verlag (to appear).