ANNALES
POLONICI MATHEMATICI
LXXIV (2000)

Convergence of power series along vector fields and their commutators; a Cartan–Kähler type theorem

by B. Jakubczyk (Warszawa)

The first version of this paper was written in 1995 as a result of a Banach Center Symposium coorganized by Bogdan Ziemian. I devote the paper to his memory.

Abstract. We study convergence of formal power series along families of formal or analytic vector fields. One of our results says that if a formal power series converges along a family of vector fields, then it also converges along their commutators. Using this theorem and a result of T. Morimoto, we prove analyticity of formal solutions for a class of nonlinear singular PDEs. In the proofs we use results from control theory.

1. Convergence along families of vector fields. We denote by $\mathbb{R}[[x]]$ and $\mathbb{R}\{x\}$ the algebras of formal (respectively, converging) power series of n commuting variables $x = (x_1, \ldots, x_n)$, with real coefficients. A formal (respectively, local analytic) vector field X on $\mathbb{R}^n$ is, by definition, a derivation of this algebra. It can be written in the form

$$X = \sum_{i=1}^{n} a_i(x) \frac{\partial}{\partial x_i},$$

where $a_i \in \mathbb{R}[[x]]$ (respectively, $a_i \in \mathbb{R}\{x\}$). The formal (respectively, local analytic) vector fields on $\mathbb{R}^n$ form a Lie algebra with the commutator (Lie bracket) as product. For $h \in \mathbb{R}[[x]]$ we denote by $h(0)$ the constant term of h.

Consider a formal power series $h \in \mathbb{R}[[x]]$ and a family of formal vector fields $X_1, \ldots, X_r$ on $\mathbb{R}^n$. Iterative differentiation of h along the vector fields

2000 *Mathematics Subject Classification*: Primary 35A10, 35G20, 41A58; Secondary 93B29.

Key words and phrases: Cartan–Kähler theorem, power series, convergence, Cauchy estimates, vector fields, commutators, control systems.

Partially supported by Polish KBN grant 2P03A 035 16.

gives the numbers

$$(X_{i_1} \dots X_{i_k} h)(0)$$

which can be viewed as coefficients of a noncommutative Taylor series of h along $X_1, \dots, X_r$. Such series were studied e.g. in [F] and will be described in Section 3. The following result says that if the above coefficients satisfy a Cauchy type estimate, then the formal data h and $X_1, \dots, X_r$ can be replaced by local analytic $\widetilde{h}$ and $\widetilde{X}_1, \dots, \widetilde{X}_r$.

THEOREM 1. *There exist positive constants C and ϱ such that*

$$|(X_{i_1} \dots X_{i_k} h)(0)| \le C\varrho^k k! , \tag{A}$$

for any $1 \le i_1, \dots, i_k \le r$ and $k \ge 1$, if and only if there are analytic $\widetilde{h}$ and $\widetilde{X}_1, \dots, \widetilde{X}_r$ defined locally around zero such that

$$(\widetilde{X}_{i_1} \dots \widetilde{X}_{i_k} \widetilde{h})(0) = (X_{i_1} \dots X_{i_k} h)(0) \tag{B}$$

for any $k \ge 0$ and any $1 \le i_1, \dots, i_k \le r$.

This theorem in a stronger version (with infinite families of vector fields) will follow from the results of Section 3.

REMARK 1. The "if" part of Theorem 1 is classic. Namely, if h and $X_1, \dots, X_r$ are analytic, then the estimate (A) follows for a fixed multiindex $I = (i_1, \dots, i_k)$ from the Cauchy inequalities applied to the local analytic function

$$\begin{aligned}(t_1, \dots, t_k) \mapsto h \circ \exp(t_k X_{i_k}) \circ \dots \circ \exp(t_1 X_{i_1})(0) \\ = \sum_{p_1, \dots, p_k \ge 0} (X_{i_1}^{p_1} \dots X_{i_k}^{p_k} h)(0) \frac{t_1^{p_1} \dots t_k^{p_k}}{p_1! \dots p_k!}\end{aligned}$$

(by $\exp(tX)$ we denote the local flow of X). The independence of the constants C and ϱ of the multiindex I can be proved by using complex Cauchy integrals.

Before we state our next result, let $r = 2$ and take $Z_1 = X_1$, $Z_2 = [X_1, X_2]$, the commutator of X_1 and X_2. It is clear that if (A) holds for the family X_1, X_2, then a simple substitution of $Z_2 = X_1 X_2 - X_2 X_1$ to (A) implies the following Gevrey type estimate:

$$|(Z_{j_1} \dots Z_{j_m} h)(0)| \le C\varrho^{m_1 + 2m_2} (m_1 + 2m_2)! 2^{m_2},$$

where m_i is the number of occurrences of i in the sequence $j_1, \dots, j_m$.

In order to generalize this observation we consider the Lie algebra $L = \mathrm{Lie}\{X_1, \dots, X_r\}$ of formal vector fields on $\mathbb{R}^n$ generated by the vector fields $X_1, \dots, X_r$. Denote by L_q the linear subspace in L spanned by the Lie brackets of $X_1, \dots, X_r$ of order not greater than q (for example we have $X_i \in L_1$ and $[X_i, X_j] \in L_2$).

Fix a family $Z_1, \ldots, Z_N$ of formal vector fields in $\mathrm{Lie}\{X_1, \ldots, X_r\}$. Denote by $w(1), \ldots, w(N)$ the minimal natural numbers such that $Z_j \in L_{w(j)}$, $j = 1, \ldots, N$. For a given multiindex $J = (j_1, \ldots, j_r)$ we define

$$w(J) = w(j_1) + \ldots + w(j_r).$$

One can easily prove the following result.

PROPOSITION 1. *If formal vector fields $X_1, \ldots, X_r$ and a formal power series h satisfy the estimate* (A) *then the following Gevrey estimate holds*:

$$|(Z_{j_1} \ldots Z_{j_m} h)(0)| \le C \varrho_1^{w(J)} (w(J))!,$$

with some constant $\varrho_1 > 0$.

The above fact being trivial as a corollary of Theorem 1, it turns out that a much stronger Cauchy type estimate holds.

THEOREM 2. *If formal vector fields $X_1, \ldots, X_r$ and a formal power series h satisfy the estimate* (A) *and $Z_1, \ldots, Z_N$ are elements of* $\mathrm{Lie}\{X_1, \ldots, X_r\}$, *then there are constants $C_1 > 0$ and $\varrho_1 > 0$ such that*

$$|(Z_{j_1} \ldots Z_{j_m} h)(0)| \le C_1 \varrho_1^m m! \tag{C}$$

for any $1 \le j_1, \ldots, j_m \le N$ and $m \ge 1$.

Proof. We can replace our formal system $(h, X_1, \ldots, X_r)$ by a local analytic system $(\widetilde{h}, \widetilde{X}_1, \ldots, \widetilde{X}_r)$, using the "only if" part of Theorem 1. Consider the vector fields $\widetilde{Z}_1, \ldots, \widetilde{Z}_N$ obtained from $\widetilde{X}_1, \ldots, \widetilde{X}_r$ by the same bracketing pattern as $Z_1, \ldots, Z_N$ were obtained from $X_1, \ldots, X_r$. They are also analytic. Then the equalities (B) yield

$$(\widetilde{Z}_{j_1} \ldots \widetilde{Z}_{j_m} \widetilde{h})(0) = (Z_{j_1} \ldots Z_{j_m} h)(0)$$

and the estimate (C) follows from the analyticity of $\widetilde{h}$ and $\widetilde{Z}_1, \ldots, \widetilde{Z}_N$, and the "if" part of Theorem 1. ∎

As a consequence of this theorem we have a result which is analogous to the following celebrated theorem of Gabrielov [G]: *If the composition $h \circ \psi$ of a formal power series $h \in \mathbb{R}[[x]]$ with a germ of an analytic map ψ is convergent, and the generic rank of ψ is equal to n, then h is also convergent.*

COROLLARY 1. *Suppose that a formal system $(h, X_1, \ldots, X_r)$ satisfies the estimate* (A). *If additionally the vector fields $X_1, \ldots, X_r$ are convergent* (*analytic at zero*) *and the Lie algebra* $\mathrm{Lie}\{X_1, \ldots, X_r\}$ *is of rank n at zero, then the formal power series h is convergent.*

First proof. Our assumption on the rank of the Lie algebra generated by $X_1, \ldots, X_r$ says that there are elements $Z_1, \ldots, Z_n$ of this Lie algebra such

that the vectors $Z_1(0), \ldots, Z_n(0)$ are linearly independent. From Theorem 2 it follows that the function

$$\phi(t_1, \ldots, t_n) = h \circ \exp(t_1 Z_1) \circ \ldots \circ \exp(t_n Z_n)(0)$$

is analytic as a function of $(t_1, \ldots, t_n)$. Additionally, as the vector fields $X_1, \ldots, X_r$ and so $Z_1, \ldots, Z_n$ are analytic, it follows that the local mapping ψ defined by

$$(t_1, \ldots, t_n) \mapsto \exp(t_1 Z_1) \circ \ldots \circ \exp(t_n Z_n)(0)$$

is analytic. It is also locally invertible at zero, as $Z_1, \ldots, Z_n$ are linearly independent. Therefore, $h = \phi \circ \psi^{-1}$ is convergent. ∎

Second proof. We also present another proof which is independent of the above argument but it uses other results including the nontrivial Gabrielov theorem.

For any multiindex $I = (i_1, \ldots, i_k)$ we introduce the local map

$$\psi_I(t_1, \ldots, t_k) = \exp(t_k X_{i_k}) \circ \ldots \circ \exp(t_1 X_{i_1})(0).$$

Each such map is defined on a certain maximal open set. The set $\mathcal{R}$ of points reachable from $0 \in \mathbb{R}^n$ piecewise by trajectories of the vector fields $X_1, \ldots, X_r$ is equal to the union of the images of these maps,

$$\mathcal{R} = \bigcup_I \operatorname{Im} \psi_I.$$

From our rank assumption on the Lie algebra $\operatorname{Lie}\{X_1, \ldots, X_r\}$ and the Chow–Rashevskiĭ theorem (cf. [C], [R]) it follows that the set $\mathcal{R}$ has nonempty interior. As the sum defining $\mathcal{R}$ is countable, it follows that there exists a multiindex I_0 such that $\operatorname{Im} \psi_{I_0}$ has nonempty interior. It follows that the rank of the map ψ_{i_0} is maximal and equal to n, generically.

Consider now the formal power series

$$h \circ \psi_{I_0}(t_1, \ldots, t_k) = \sum_{p_1, \ldots, p_k \geq 0} (X_{i_1}^{p_1} \ldots X_{i_k}^{p_k} h)(0) \frac{t_1^{p_1} \ldots t_k^{p_k}}{p_1! \ldots p_k!}.$$

From the estimate (A) it follows that this series is convergent.

As $h \circ \psi_{I_0}$ is convergent and ψ_{I_0} has generic rank n, it follows from Gabrielov's theorem that h is convergent. ∎

Remark 2. The second proof can be refined so that the estimate (A) is needed only for sequences of the form

$$X_{i_1}^{p_1} \ldots X_{i_n}^{p_n},$$

where n is the dimension of our space and $i_1, \ldots, i_n$ are fixed. Namely, by an argument of A. Krener [K] there exists a sequence $i_1, \ldots, i_n$ of length n such that $\operatorname{Im} \psi_I$ has nonempty interior.

Our preceding corollary can be generalized to the case when the Lie algebra generated by our vector fields is not of maximal rank.

We will say that a formal power series h at $x_0 = 0 \in \mathbb{R}^n$ *is convergent on a submanifold* S at x_0 if the (formal) composition $h \circ \psi$ is convergent, where $\psi : \mathbb{R}^m \to \mathbb{R}^n$ is a local analytic immersion defining S, i.e. $\psi(0) = x_0$ and $S = \operatorname{Im} \psi$ locally.

If the Lie algebra $L = \operatorname{Lie}\{X_1, \ldots, X_r\}$ generated by analytic vector fields is not of full rank at x_0, then the distribution defined by $\Delta(x) = \{V(x) : V \in L\}$ is integrable (Nagano [N], Sussmann [S1]). This means that through any point, in particular through x_0, there passes an integral manifold S of Δ. Additionally, it follows from the Chow–Rashevskiĭ theorem that the submanifold S coincides with the orbit $\operatorname{Orb}(x_0)$ from x_0 of the family $\{X_1, \ldots, X_r\}$ (here $\operatorname{Orb}(x_0)$ is the set of points reachable from x_0 piecewise by trajectories of the vector fields $X_1, \ldots, X_r$).

COROLLARY 2. *Let h be a formal power series and let $X_1, \ldots, X_r$ be vector fields defined and analytic in a neighborhood of $x_0 = 0$ in $\mathbb{R}^n$ such that the estimate* (A) *is satisfied. Then h is convergent on the submanifold $S = \operatorname{Orb}(x_0)$.*

One could attempt to analyse "analyticity directions" of a formal power series, using the above corollary.

2. Analytic solutions of nonlinear singular PDEs. The results of the preceding section can be used to improve regularity statements for solutions of a class of nonlinear, singular partial differential equations on nilpotent Lie groups as considered in Morimoto [Mo1], [Mo2]. We begin by recalling the notion of a formal Gevrey function as introduced in [Mo1].

Suppose that

$$g = g_1 \oplus \ldots \oplus g_s$$

is a graded nilpotent Lie algebra of finite dimension n, with weighting $w(X) = i$ if $X \in g_i$. Let G be the corresponding nilpotent Lie group. For any element $X \in g$ we denote by $\widehat{X}$ the corresponding left invariant vector field on G.

Fix a basis $X_1, \ldots, X_n$ of g adapted to the direct sum, i.e. $X_i \in g_{j(i)}$, for any $i = 1, \ldots, n$. Consider the corresponding vector fields $\widehat{X}_1, \ldots, \widehat{X}_n$ on G.

Fix an analytic coordinate system in a neighborhood of the identity $e \in G$. Consider a formal power series h on G (at the identity). Following [Mo1] we say that h is in the *formal Gevrey class* if there exist positive constants C and ϱ such that the following Gevrey type estimate holds:

$$\text{(GE)} \qquad |(\widehat{X}_{i_1} \ldots \widehat{X}_{i_k} h)(0)| \le C \varrho^{w(I)} w(I)!$$

for any $1 \le i_1, \ldots, i_k \le n$ and $k \ge 1$, where $w(I) = w(X_{i_1}) + \ldots + w(X_{i_k})$.

Corollary 3. *If a formal power series h on a nilpotent Lie group G satisfies a Gevrey estimate* (GE), *and g_1 generates the whole Lie algebra g, then h is convergent.*

Proof. After a possible renumbering we may assume that $X_1, \ldots, X_r$ span g_1. Taking into account that $w(X_i) = 1$ if $X_i \in g_1$, we see that the Gevrey estimate (GE) implies the estimate (A). The conclusion follows from Corollary 1. ∎

This corollary implies a theorem on existence of local analytic solutions of nonlinear systems of partial differential equations, of the type of the Cartan–Kähler theorem, as stated by T. Morimoto in [Mo2], Theorem EA. This is an extension of earlier theorems by Malgrange [Ma], Appendix, and Morimoto [Mo1]. We restate this theorem below and give its proof.

Let $X_1, \ldots, X_n$ be a basis of the graded Lie algebra g as above. Note that any left invariant differential operator on G can be written as a polynomial in the first order differential operators $\widehat{X}_1, \ldots, \widehat{X}_n$. For (a germ) of m functions $y = h(x) = (h_1(x), \ldots, h_m(x))$ on G and a multiindex $I = (i_1, \ldots, i_k)$ we define

$$y_I(x) = h_I(x) = X_I h = (X_{i_1} \ldots X_{i_k} h)(x)$$

(we take the derivatives componentwise) and $w(I) = w(X_{i_1}) + \ldots + w(X_{i_k})$. We consider the following differential equation on G:

$$\Psi(x, \{y_I\}_{w(I) \le p}) = 0. \tag{E}$$

Following Morimoto [Mo1], we define weighted jets of h as follows. Two functions $h, h' : G \to \mathbb{R}^m$ (defined locally around the identity e) are called *q-equivalent* at e if $(X_I h)(e) = (X_I h')(e)$ for any I such that $w(I) \le q$. The equivalence class of h is called the (*weighted*) *jet of order* q of h at e and is denoted by $j^q h$. A q-jet $j^q h$, $q \ge p$, is called a *q-jet solution* of (E) if

$$(X_J \Psi(x, \{y_I\}_{w(I) \le p}))(e) = 0$$

for all multiindices J such that $w(J) \le q - p$.

Following Malgrange [Ma], Appendix, one calls a q-jet solution $y^q = j^q h$ of (E) *strongly prolongable* if for any s-jet extension $j^s \widetilde{h}$ of $j^q h$, $s > q$, which is an s-jet solution of (E) there exists an $(s+1)$-solution $j^{s+1} \widehat{h}$ of (E) whose s-jet is $j^s \widetilde{h}$.

It was proved by Morimoto [Mo1], as an extension of a theorem by Malgrange [Ma], that if Ψ is analytic with respect to y_I and in the formal Gevrey class with respect to x then, for any $m \ge p$ and any finite jet solution $j^m \widetilde{h}$ of our equation which is strongly prolongable, there exists a solution h in the formal Gevrey class with $j^m h = j^m \widetilde{h}$.

From this result and Corollary 2 we obtain the following theorem which is Theorem EA of [Mo2] (in fact, a slight generalization). Let G_0 be the

subgroup of the Lie group G which corresponds to the Lie subalgebra $g^0 = \mathrm{Lie}\{g_1\}$ generated by g_1.

THEOREM 3. *Let Ψ be in the formal Gevrey class with respect to x and analytic with respect to y_I at (e, y^p). If $y^p = \{y_I\}_{w(I)\le p}$ is a strongly prolongable solution of* (E), *then there exists a solution h in the formal Gevrey class such that $j^p h = y^p$, which is convergent on G_0. In particular, if g_1 generates the whole Lie algebra, then h is convergent.*

Proof. It follows from [Mo1] that there exists a required infinite jet solution h which is in the formal Gevrey class. This implies, in particular, that this solution satisfies our estimate (A) with respect to the vector fields $\widehat{X}_1, \ldots, \widehat{X}_r$ generating g_1 (since $w(X) = 1$ if $X \in g_1$). To get the result it is enough to apply Corollary 2 and notice that the orbit of the family of vector fields $X \in g_1$ equals the subgroup G_0. ∎

Corollary 3 and Theorem 3 were conjectured (and proved in the case where g is the Heisenberg algebra) by T. Morimoto during the Banach Center Symposium on Differential Geometry and Mathematical Physics in April 1995. Later they were published in [Mo2]. Corollary 3 is stated in [Mo2] as Theorem G-A and its proof is reduced to two propositions (Propositions 1 and 2). To our knowledge, the proofs of these propositions have not been published (the proof of Proposition 2, given in the reference [St1] cited in [Mo2], is not correct, cf. [St2]).

3. Realizations of noncommutative power series. We will state an abstract version of Theorem 1, where the initial numbers $(X_{i_1} \ldots X_{i_k} h)(0)$ will be replaced by coefficients of a formal power series Q of noncommutative variables u (indexing vector fields).

Let U be a set with at least two elements, possibly infinite (we think of elements of U as letters or noncommutative formal variables). The *free semigroup* generated by U is the set of words

$$U^+ = \{u_1 \ldots u_k \mid k \ge 1,\ u_i \in U\}$$

with the natural product (concatenation). The *free monoid* U^* is obtained by adding the empty word λ to U^+. A formal power series of noncommuting variables $u \in U$, with real coefficients $\langle Q, w\rangle$, is written as a formal sum

$$\sum_{w \in U^*} \langle Q, w\rangle w,$$

with the constant term $\langle Q, \lambda\rangle \lambda$. The family of all such formal power series forms an associative algebra denoted by $\mathbb{R}\langle\langle U\rangle\rangle$, with the usual addition and product. The corresponding algebra of polynomials is denoted by $\mathbb{R}\langle U\rangle$. We

have the canonical inclusions

$$U \subset U^+ \subset U^* \subset \mathbb{R}\langle U\rangle \subset \mathbb{R}\langle\langle U\rangle\rangle.$$

The commutator $[Q, Q'] = QQ' - Q'Q$ defines Lie algebra structures in $\mathbb{R}\langle U\rangle$ and $\mathbb{R}\langle\langle U\rangle\rangle$. We also have a well defined scalar product

$$\langle Q, Q'\rangle = \sum_{w\in U^*} \langle Q, w\rangle\langle Q', w\rangle,$$

when the series converges (e.g. when one of the series is a polynomial).

We set $[u_1u_2] = u_1u_2 - u_2u_1$. More generally, for a word $w = u_1 \dots u_k \in U^+$ we define a polynomial, called the *Lie polynomial*, by

$$[w] = [u_1 \dots u_k] = [u_1, [u_2, \dots [u_{k-1}, u_k] \dots]].$$

We can now define the *Lie rank* (cf. [F], [J4]) of a series $Q \in \mathbb{R}\langle\langle U\rangle\rangle$ by

$$\operatorname{rank}_L Q = \sup \operatorname{rank}\{\langle Q, [w_p]z_q\rangle\}_{p,q=1}^k,$$

where $\{\ \}_{p,q=1}^k$ denotes a $k \times k$ matrix, and the supremum is taken over all $k \geq 1$ and all sequences $w_1, \dots, w_k \in U^+$ and $z_1, \dots, z_k \in U^*$. Similarly, for a family $\mathcal{Q} = \{Q_v\}_{v\in V}$ of formal power series we define its rank by

$$\operatorname{rank}_L \mathcal{Q} = \sup \operatorname{rank}\{\langle Q_{v_q}, [w_p]z_q\rangle\}_{p,q=1}^k,$$

where the supremum is taken over all $k \geq 1$ and all sequences $w_1, \dots, w_k \in U^+$, $z_1, \dots, z_k \in U^*$ and $v_1, \dots, v_k \in V$.

The following result is an abstract version of Theorem 1.

THEOREM 4. *The following conditions are equivalent for a given family of formal power series $\mathcal{Q}$.*

(i) *The family $\mathcal{Q}$ has finite rank and there are positive functions $C : V \to \mathbb{R}$ and $\varrho : U \to \mathbb{R}$ such that*

$$|\langle Q_v, u_1 \dots u_k\rangle| \leq C(v)\varrho(u_1) \dots \varrho(u_k)k!\,. \tag{A$'$}$$

(ii) *There exists n and a local analytic system $(\{X_u\}_{u\in U}, \{h_v\}_{v\in V})$ defined in a common neighborhood of $0 \in \mathbb{R}^n$ such that*

$$\langle Q_v, u_1 \dots u_k\rangle = (X_{u_1} \dots X_{u_k} h_v)(0)$$

for all $k \geq 0$, $v \in V$, and $u_1, \dots, u_k \in U$ (one can take $n = \operatorname{rank}_L \mathcal{Q}$).

In the case of $\mathcal{Q}$ consisting of a single series and U finite this theorem was stated in [F]. However, the proof given in [F] was not correct. Moreover, the main assumption $\operatorname{rank}_L \mathcal{Q} < \infty$ was not used in that proof (only the inequality $\operatorname{rank}_L \mathcal{Q} \geq n$ was used). We mentioned in [J3], Remark 2, that Theorem 2 in [J3] implied the analytic version of the main result in [F], and it implied Theorem 4 above. However, in order to make this implication explicit one has to show equality of two different ranks used in [F] and [J3].

In the next two sections we restate a version of a realization theorem from [J3] and prove equality of both ranks. Another proof of Theorem 4 was given in [S2].

4. Realizations of causal functionals. Let U and V be fixed nonempty sets, where U has at least two elements and V is finite. For $t_1 \geq 0, \ldots, t_k \geq 0$ we denote by

$$a = (t_1, u_1) \ldots (t_k, u_k)$$

the piecewise constant function defined on the interval $[0, |a|)$ and taking values $a(t) = u_i \in U$ for $t \in [T_{i-1}, T_i)$, where $T_0 = 0$ and $T_i = t_1 + \ldots + t_i$, $i = 1, \ldots, k$, and $|a| = T_k$. The set of all such functions is denoted by S.

Note that two functions $a, b \in S$ can be concatenated, which defines the product ab, where

$$(ab)(t) = \begin{cases} a(t) & \text{for } t \in [0, |a|), \\ b(t - |a|) & \text{for } t \in [|a|, |a| + |b|). \end{cases}$$

This means that S forms a semigroup with concatenation as product.

Consider a positive function $\varrho : U \to \mathbb{R}$. We define the subset $S_\varrho \subset S$ by

$$S_\varrho = \{\, (t_1, u_1) \ldots (t_k, u_k) \in S : t_1 \varrho(u_1) + \ldots + t_k \varrho(u_k) < 1 \}$$

(here k is not fixed). Consider a family $P = \{P_v\}_{v \in V}$ of functions $P_v : S_\varrho \to \mathbb{R}$ called *causal functionals*.

In order to have existence of analytic realizations we impose the following conditions on our family P of causal functionals. Set $\mathbb{R}_+ = [0, \infty)$.

We say that P is *analytic* (respectively, *analytic and analytically extendable*) if, for any $u_1, \ldots, u_k \in U$ and $k \geq 1$, the maps

$$(t_1, \ldots, t_k) \mapsto P_v((t_1, u_1) \ldots (t_k, u_k))$$

are analytic on the subsets of $\mathbb{R}_+^k$ defined by

$$\{(t_1, \ldots, t_k) \in \mathbb{R}_+^k : t_1 \varrho(u_1) + \ldots + t_k \varrho(u_k) < 1\}$$

(respectively, they are analytic on these subsets and have analytic extensions to the subsets of $\mathbb{R}^k$ defined by

$$\{(t_1, \ldots, t_k) \in \mathbb{R}^k : |t_1| \widetilde{\varrho}(u_1) + \ldots + |t_k| \widetilde{\varrho}(u_k) < 1\},$$

where $\widetilde{\varrho} : U \to \mathbb{R}$ is a positive function which depends on P but is independent of the choice of k and $u_1, \ldots, u_k$).

Above, a function defined on a possibly non-open subset of $\mathbb{R}^n$ is called analytic if it has a Taylor series converging at any point of the set and is given by this series in a neighborhood of this point. In particular, analyticity of P means that the above maps have analytic extensions to open neighborhoods of the origin in $\mathbb{R}^k$. Thus, the expressions $P_v((t_1, u_1) \ldots (t_k, u_k))$ are also well defined for small negative $t_1, \ldots, t_k$. Analytic extendability is a slightly

stronger property requiring some uniformity of the domains of the extended maps.

We define the *rank* of the family P of causal functionals by

$$\operatorname{rank} P = \sup \operatorname{rank} \left\{ \frac{\partial}{\partial t_i} P_{v_j}(ab_j) \right\}_{i,j=1}^{k},$$

where the supremum is taken over all $k \geq 1$, all $a = (t_1, u_1) \dots (t_k, u_k) \in S$ and all $b_1, \dots, b_k \in S$ and $v_1, \dots, v_k \in V$ such that $ab_j \in S_\varrho$, $j = 1, \dots, k$. Consider a quadruple $\Gamma = (M, X, h, x_0)$, where $M = \mathbb{R}^n$, x_0 is a point in M, and $X = \{X_u\}_{u \in U}$, $h = \{h_v\}_{v \in V}$ are families of germs at x_0 of analytic vector fields on M and of analytic functions on M, respectively. We say that Γ is a *local analytic realization* of $P = \{P_v\}_{v \in V}$ if

$$P_v(a) = h_v \circ \exp(t_k X_{u_k}) \circ \dots \circ \exp(t_1 X_{u_1})(x_0)$$

for any $v \in V$, for any $k \geq 1$ and $u_1, \dots, u_k \in U$, with $t_i \geq 0$ small enough. Here $\exp(tX)$ denotes the local flow of X.

The realization Γ is called *locally minimal* if $\dim M = \operatorname{rank} P$.

THEOREM 5 ([J2], [J3]). *A family P of causal functionals has a local analytic realization Γ if and only if P is analytic, analytically extendable, and* $\operatorname{rank} P$ *is finite. Then there exists a locally minimal, local analytic realization of P.*

For global versions of this theorem, with M an arbitrary differentiable manifold, the reader may consult [J1], [CGM] and [J4].

REMARK 3. If the set U is finite then the requirement of analytic extendability is superfluous, which follows from the proof of Theorem 2 in [J3] and was independently proved in [CGM]. Theorem 5 remains true when the set V is infinite. This can be proved by choosing a finite subset $V_0 \subset V$ on which the rank of P is realized, constructing the realization $\Gamma = (\mathbb{R}^n, X, h, x_0)$ with $h = \{h_v\}_{v \in V_0}$, and showing that the remaining observation functions h_v, with $v \in V \setminus V_0$, can be constructed without changing the other elements of Γ.

REMARK 4. The family P of causal functionals defines a causal operator F by $F((t_1, u_1) \dots (t_k, u_k)v) = P_v((t_1, u_1) \dots (t_k, u_k))$, in the terminology used in [J3] and [J4]. For simplicity of notation the function ϱ used here in the definition of causal functionals was taken to be constant, equal to $1/T$, in [J3] (this implies that the operator F is defined on functions on the interval $[0, T)$). In the proofs in [J3] this assumption is not needed.

5. Equivalent ranks. Our family $P = \{P_v\}_{v\in V}$ of analytic causal functionals defines a family $Q = \{Q_v\}_{v\in V}$ of formal power series via the formula

$$\langle Q_v, u_1 \dots u_k\rangle = \frac{\partial}{\partial t_1} \dots \frac{\partial}{\partial t_k} P_v((t_1, u_1) \dots (t_k, u_k))\Big|_{t_1=\dots=t_k=0}. \tag{1}$$

The following fact holds for this family of formal power series.

LEMMA 1. *If P is analytic, then both ranks defined above coincide, i.e.*

$$\operatorname{rank} P = \operatorname{rank}_L Q.$$

Before we prove the lemma we make the following observations.

OBSERVATION 1. The semigroup S can be extended to a group G defined as follows. Consider formal products

$$a = (t_1, u_1) \dots (t_k, u_k),$$

where $t_1, \dots, t_k \in \mathbb{R}$, $u_1, \dots, u_k \in U$, and $k \geq 0$. We introduce the relations

$$(t_1, u)(t_2, u) = (t_1 + t_2, u), \qquad (0, u) = e,$$

where e is the empty product (i.e. $k = 0$). The set of formal products factorized with respect to the equivalence relation generated by the above relations forms a group G, with multiplication defined by concatenation of products, and with inverse

$$((t_1, u_1) \dots (t_k, u_k))^{-1} = (-t_k, u_k) \dots (-t_1, u_1)$$

(for simplicity we omit equivalence classes in our notation). It is easy to see that the semigroup S defined in the preceding section can be identified with the set of equivalence classes of formal products $(t_1, u_1) \dots (t_k, u_k)$ with $t_1 \geq 0, \dots, t_k \geq 0$.

OBSERVATION 2. If P is a family of analytic causal functionals $P_v : S_\varrho \to \mathbb{R}$ then the expressions $P_v((t_1, u_1) \dots (t_k, u_k))$ are well defined for nonnegative $t_1, \dots, t_k$ and, by uniqueness of analytic extensions, they are also well defined for $(t_1, \dots, t_k)$ in a neighborhood of the origin in $\mathbb{R}^k$. The equalities

$$P_v((t_1, u_1) \dots (t', u)(t'', u) \dots (t_k, u_k)) = P_v((t_1, u_1) \dots (t' + t'', u) \dots (t_k, u_k))$$

and

$$\begin{aligned} P_v((t_1, u_1) \dots (t_i, u_i)(0, u)(t_{i+1}, u_{i+1}) \dots (t_k, u_k)) \\ = P_v((t_1, u_1) \dots (t_i, u_i)(t_{i+1}, u_{i+1}) \dots (t_k, u_k)) \end{aligned}$$

extend, by uniqueness of analytic extensions, from a set of nonnegative values of $t_1, \dots, t_k, t', t'' \in \mathbb{R}_+$, to arbitrary small values in $\mathbb{R}$. (This roughly means that the causal functionals $P : S_\varrho \to \mathbb{R}$ have well defined extensions

to a "neighborhood" of the identity in G.) In particular, our equations imply

$$P_v(abb^{-1}c) = P_v(ac)$$

for arbitrary elements $a = (t_1, u_1)\dots(t_k, u_k)$, $b = (\tau_1, v_1)\dots(\tau_m, v_m)$, and $c = (s_1, w_1)\dots(s_r, w_r)$, with t_i, τ_j, s_j small.

Observation 3. From the definition (1) of the series Q_v we easily see that

$$\begin{aligned}\langle Q_v, [u_1u_2]v_1\rangle &= \langle Q_v, u_1u_2v_1\rangle - \langle Q_v, u_2u_1v_1\rangle \\ &= \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}\frac{\partial}{\partial \tau_1} P_v((t_1,u_1)(t_2,u_2)(-t_1,u_1)(\tau_1,v_1))\Big|_{t_1=t_2=\tau_1=0}\end{aligned}$$

and, by easy induction,

$$\begin{aligned}(2)\quad &\langle Q_v, [u_1\dots u_{k+1}]v_1\dots v_m\rangle \\ &= \frac{\partial}{\partial t_1}\dots\frac{\partial}{\partial t_{k+1}}\frac{\partial}{\partial \tau_1}\dots\frac{\partial}{\partial \tau_m} P_v(a(t_{k+1},u_{k+1})a^{-1}b)\Big|_{t_1=\dots=t_{k+1}=\tau_1\dots=\tau_m=0},\end{aligned}$$

where

$$a = (t_1, u_1)\dots(t_k, u_k), \quad b = (\tau_1, v_1)\dots(\tau_m, v_m)$$

and $\langle\cdot,\cdot\rangle$ denotes the scalar product while $[u_1\dots u_{k+1}]$ denotes the Lie polynomial, both defined in Section 3. Using formula (2), the definition of the Lie polynomial $[(u_1)^{i_1}\dots(u_k)^{i_k}u_{k+1}]$, and the property

$$\frac{d^r}{ds^r}f(s)\Big|_{s=0} = \frac{\partial}{\partial s_1}\dots\frac{\partial}{\partial s_r}f(s_1+\dots+s_r)\Big|_{s_1=\dots=s_r=0}$$

we obtain the more general equalities

$$\begin{aligned}(3)\quad &\langle Q_v, [(u_1)^{i_1}\dots(u_k)^{i_k}u_{k+1}](v_1)^{j_1}\dots(v_m)^{j_m}\rangle \\ &= \left(\frac{\partial}{\partial t_1}\right)^{i_1}\dots\left(\frac{\partial}{\partial t_k}\right)^{i_k}\frac{\partial}{\partial t_{k+1}}\left(\frac{\partial}{\partial \tau_1}\right)^{j_1}\dots\left(\frac{\partial}{\partial \tau_m}\right)^{j_m} P_v(a(t_{k+1},u_{k+1})a^{-1}b),\end{aligned}$$

where $a = (t_1, u_1)\dots(t_k, u_k)$, $b = (\tau_1, v_1)\dots(\tau_m, v_m)$, and the derivatives on the right side are taken at $t_1 = \dots = t_{k+1} = 0$ and $\tau_1 = \dots = \tau_m = 0$.

Proof of Lemma 1. The following remark will be used in the proof. In the definitions of both ranks there appear ranks of finite-dimensional matrices. In order to show that $\operatorname{rank} P = \operatorname{rank}_L Q$ it is enough to show that for any $p \geq 1$ the following statement holds. If all the minors of order p of the matrices involved in the definition of the first rank vanish, then all the minors of order p of the matrices involved in the definition of the second rank vanish, and vice versa.

The matrices which appear in the definition of $\operatorname{rank} P$ depend analytically on some real parameters. We will use the following general property of determinants of matrices with coefficients depending on real parameters. Let

$\phi = \det\{a_{ij}\}_{i,j=1}^{p}$, where the coefficients $a_{ij} = a_{ij}(\overline{t}^i, \overline{\tau}^j)$ depend smoothly (analytically) on the parameters $\overline{t}^i = (t_1^i, \ldots, t_{k_i}^i)$ and $\overline{\tau}^j = (\tau_1^j, \ldots, \tau_{m_j}^j)$. Since determinant is a multilinear function with respect to rows, it follows that

$$\frac{\partial}{\partial t_s^k}\phi = \det\{\widetilde{a}_{ij}\},$$

where

$$a_{ij} = \widetilde{a}_{ij} \quad \text{for } i \neq k, \quad \text{and} \quad \widetilde{a}_{ij} = \frac{\partial}{\partial t_s^k} a_{ij} \quad \text{for } i = k.$$

A similar property holds for columns. Assume that

$$\det\{a_{ij}\} \equiv 0. \tag{4}$$

By consecutive differentiation of this identity and application of the above mentioned property of determinant we obtain

$$\det\{\widetilde{a}_{ij}\} \equiv 0, \tag{5}$$

where

$$\widetilde{a}_{ij} = \frac{\partial^{k_i+m_j}}{\partial t_1^i \ldots \partial t_{k_i}^i \partial \tau_1^j \ldots \partial \tau_{m_j}^j} a_{ij}. \tag{6}$$

We are ready for the main part of the proof.

We define

$$a_{ij} = \frac{\partial}{\partial t} P_{v_j}(a^i(t, w_{k_i+1}^i)(a^i)^{-1} b_j)\Big|_{t=0}, \tag{7}$$

where

$$a^i = (t_1^i, w_1^i) \ldots (t_{k_i}^i, w_{k_i}^i), \qquad b_j = (\tau_1^j, v_1^j) \ldots (\tau_{m_j}^j, v_{m_j}^j)$$

and $i, j = 1, \ldots, p$.

In order to prove $\operatorname{rank}_L Q \leq \operatorname{rank} P$ we assume that $\operatorname{rank} P = n < \infty$. Choosing the element a in the definition of $\operatorname{rank} P$ to be

$$a = a^1(s_1, w_{k_1+1}^1)(a^1)^{-1} \ldots a^p(s_p, w_{k_p+1}^p)(a^p)^{-1} \tag{8}$$

we have

$$\frac{\partial}{\partial s_i} P_{v_j}(ab_j)\Big|_{s_1=0,\ldots,s_p=0} = \frac{\partial}{\partial s_i} P_{v_j}(a^i(s_i, w_{k_i+1}^i)(a^i)^{-1} b_j)\Big|_{s_i=0} = a_{ij} \tag{9}$$

and so we get $\det\{a_{ij}\} = 0$ if $p > n$. This implies (5). Using the definition (7) and formula (2) we see that equality (5) takes the form

$$\det\{\langle Q, [w^i]v^j\rangle\}_{i,j=1}^{p} = 0, \tag{10}$$

where $w^i = w_1^i \ldots w_{k_i+1}^i$ and $v^j = v_1^j \ldots v_{m_j}^j$. As $w^i \in \Delta^+$ and $v^j \in \Delta^*$ are arbitrary, and p is any number $> n$, we conclude that $\operatorname{rank}_L Q \leq n$.

In order to prove the converse inequality $\operatorname{rank} P \le \operatorname{rank}_L Q$ assume that $\operatorname{rank}_L Q = n < \infty$. Then we have (10) for any $p > n$. Taking a_{ij} defined by (7), and then applying formula (2), we deduce from (10) that (5) holds at the origin (i.e., for all parameters equal to zero). The same can be proved with $\widetilde{a}_{ij}$ given by

$$\widetilde{a}_{ij} = \left(\frac{\partial}{\partial t_1^i}\right)^{p_1^i} \dots \left(\frac{\partial}{\partial t_{k_i}^i}\right)^{p_{k_i}^i} \left(\frac{\partial}{\partial \tau_1^j}\right)^{q_1^j} \dots \left(\frac{\partial}{\partial \tau_{m_j}^j}\right)^{q_{m_j}^j} a_{ij}$$

instead of (6). In this case one should start with

$$w^i = (w_1^i)^{p_1^i} \dots (w_{k_i}^i)^{p_{k_i}^i} w_{k_i+1}^i, \qquad v^j = (v_1^j)^{q_1^j} \dots (v_{m_j}^j)^{q_{m_j}^j}$$

in (10), and use formula (3).

In this way we deduce that any partial derivative of $\det\{a_{ij}\}$ with respect to the parameters $t_1^i, \dots, t_{k_i}^i$ and $\tau_1^j, \dots, \tau_{m_j}^j$ is equal to zero at the origin. This and analyticity give $\det\{a_{ij}\} \equiv 0$ for any $p \times p$ matrix with $p > n$ and a_{ij} given by (7).

Now we shall choose a^i and b_j in a special way. Take $a = (t_1, u_1) \dots \dots (t_k, u_k)$ and

$$a^i = (t_1, u_1) \dots (t_i, u_i), \qquad b_j = a\widetilde{b}_j \quad \text{and} \quad w_{k_i+1}^i = u_i.$$

Then we obtain

$$\begin{aligned} a_{ij} &= \frac{\partial}{\partial t} P_{v_j}(a^i(t, u_i)(a^i)^{-1} a\widetilde{b}_j)\Big|_{t=0} \\ &= \frac{\partial}{\partial t} P_{v_j}((t_1, u_1) \dots (t_i, u_i)(t, u_i)(t_{i+1}, u_{i+1}) \dots (t_k, u_k)\widetilde{b}_j)\Big|_{t=0} \\ &= \frac{\partial}{\partial t_i} P_{v_j}((t_1, u_1) \dots (t_k, u_k)\widetilde{b}_j), \end{aligned}$$

with $i, j = 1, \dots, p$. (With the above choice of a^i and b_j some of the earlier independent parameters are identified but our earlier conclusion $\det\{a_{ij}\} \equiv 0$ still holds.) Thus $\det\{a_{ij}\} = 0$ means that $\operatorname{rank} P \le n$. ∎

6. Proof of Theorem 4. (i)⇒(ii). We will show that the data of our theorem allow us to construct a family $P = \{P_v\}_{v \in V}$ of causal functionals which is analytic, analytically extendable, and has finite rank. We define

$$P_v((t_1, u_1) \dots (t_k, u_k)) = \sum_{i_1 \ge 0, \dots, i_k \ge 0} \langle Q_v, u_1^{i_1} \dots u_k^{i_k} \rangle \frac{t_1^{i_1} \dots t_k^{i_k}}{i_1! \dots i_k!}.$$

From the estimate (A′) it follows that the series is majorized by the series

$$\sum_{j \ge 0} C(v)(|t_1|\varrho(u_1) + \dots + |t_k|\varrho(u_k))^j$$

and so it is convergent for $\varrho(u_1)|t_1| + \ldots + \varrho(u_k)|t_k| < 1$. It follows that P_v are well defined analytic, analytically extendable (with the function $\widetilde{\varrho}$ equal to ϱ) causal functionals $S_\varrho \to \mathbb{R}$. From our assumption $\operatorname{rank}_L Q < \infty$ and Lemma 1 it follows that $\operatorname{rank} P = \operatorname{rank}_L Q < \infty$. Thus, Theorem 5 is applicable here and there exists a local analytic realization $\Gamma = (\mathbb{R}^n, \{X_u\}_{u\in U}, \{h_v\}_{v\in V}, x_0)$ of P. We may take $n = \operatorname{rank} P = \operatorname{rank}_L Q$ and $x_0 = 0$ and the equality in the statement (ii) of Theorem 4 follows by consecutive differentiation of the equality

$$P_v(a) = h_v \circ \exp(t_k X_{u_k}) \circ \ldots \circ \exp(t_1 X_{u_1})(x_0).$$

The proof of the implication (ii)⇒(i) uses Cauchy type estimates and is straightforward. ∎

References

[CGM] F. Celle, J.-P. Gauthier and E. Milani, *Existence of realizations of nonlinear input-output maps*, IEEE Trans. Automat. Control. AC-31 (1986), 378–381.

[C] W. L. Chow, *Ueber Systeme von linearen partiellen Differentialgleichungen erster Ordnung*, Math. Ann. 117 (1939), 98–105.

[F] M. Fliess, *Réalisation locale des systèmes non linéaires, algèbres de Lie filtrées transitives et séries génératrices non commutatives*, Invent. Math. 71 (1983), 521–537.

[G] A. M. Gabrielov, *Formal relations between analytic functions*, Math. USSR-Izv. 37 (1973), 1056–1088.

[J1] B. Jakubczyk, *Existence and uniqueness of realizations of nonlinear systems*, SIAM J. Control Optim. 18 (1980), 455–471.

[J2] —, *Réalisations locales des opérateurs causals non linéaires*, C. R. Acad. Sci. Paris 299 (1984), 787–789.

[J3] —, *Local realizations of nonlinear causal operators*, SIAM J. Control Optim. 24 (1986), 230–242.

[J4] —, *Realizations of nonlinear systems; three approaches*, in: Algebraic and Geometric Methods in Nonlinear Control Theory, M. Fliess and M. Hazewinkel (eds.), Reidel, 1986, 3–31.

[K] A. Krener, *A generalization of Chow's theorem and the bang-bang theorem to nonlinear control problems*, SIAM J. Control 12 (1974), 43–52.

[Ma] B. Malgrange, *Equation de Lie. II*, J. Differential Geometry 7 (1972), 117–141.

[Mo1] T. Morimoto, *Théorème de Cartan–Kähler dans une classe de fonctions formelles Gevrey*, C. R. Acad. Sci. Paris Sér. A 311 (1990), 433–436.

[Mo2] —, *Théorème d'existence de solutions analytiques pour des systèmes d'équations aux dérivées partielles non-linéaires avec singularités*, C. R. Acad. Sci. Paris Sér. I 321 (1995), 1491–1496.

[N] T. Nagano, *Linear differential systems with singularities and applications to transitive Lie algebras*, J. Math. Soc. Japan 18 (1966), 398–404.

[R] P. K. Rashevskiĭ, *On joining two points of a completely nonholonomic space by an admissible curve*, Uchen. Zapiski Pedagog. Inst. im. Liebknechta Ser. Fiz.-Mat. 1938, no. 2, 83–94 (in Russian).

[St1] R. S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geometry 24 (1986), 221–263.

[St2] —, *Corrections to "Sub-Riemannian geometry"*, ibid. 30 (1989), 595–596.

[S1] H. J. Sussmann, *Orbits of families of vector fields and integrability of distributions*, Trans. Amer. Math. Soc. 180 (1973), 171–188.

[S2] —, unpublished manuscript.

Institute of Mathematics
Polish Academy of Sciences
Śniadeckich 8
00-950 Warszawa, Poland
E-mail: jakubczy@impan.gov.pl

Reçu par la Rédaction le 30.7.1999
Révisé le 5.2.2000