# SMOOTH NONAUTONOMOUS NORMALIZATIONS FOR CONTRACTIVE MAPPINGS

Sergeĭ A. Dovbysh

Institute for Mechanics, Moscow State University
Michurinskij prospekt, d.1, 117192, Moscow, Russia

Abstract. The construction of smooth nonautonomous normalizations for a sequence of contractive mappings is discussed which generalizes the recent construction by M. Gyusinsky and A. Katok [20]. The associated Normal Form Theorem and Semi-Centralizer Theorems for extensions of homeomorphisms on compact spaces by smooth contraction mappings are considered. The approach utilized is essentially geometrical and is based on the usual tools of hyperbolic theory for mappings in Banach spaces. Some remarks for the general case of formal normal forms without the contractness assumption are also made.

Contents



Recently, M. Gyusinsky and A. Katok [20] (see also [25]) studied the non-autonomous normal forms for smooth contractions on $\mathbb{R}^n$. The relevant centralizer theorem has been applied to establish the differential rigidity of Anosov actions [25]. In the most general framework, extensions of topological dynamical systems by smooth ($C^\infty$) contractions of $\mathbb{R}^n$ are considered. The usual considerations based on the contraction argument (see Remark 11 below) show the presence of a unique invariant continuous section. The conjugacies of the extensions by a continuous family of local $C^\infty$-diffeomorphisms of the fibers preserving the invariant section were dealt with. Some our considerations are parallel to those in the paper cited.

However, the present work generalizes essentially the paper [20] because of the following reasons. Firstly, the authors of [20] restrict themselves by studying normal forms that correspond to a rather special stable set $\mathcal{R}$ (in terms of the present paper). They denoted these forms as *"sub-resonance normal forms"*. All the consideration were essentially adopted to this particular case, although the general case can be easily treated by the geometric method developed below. (The considerations in [20] are more analytic.) We will show that the narrow band spectrum condition introduced by M. Gyusinsky and A. Katok is not only sufficient but also necessary for the possibility to construct nonautonomous normal form with good properties. Secondly, the

*Key words and phrases.* Normal forms, normalization, (semi)conjugacy, hyperbolic fixed point, hyperbolic theory, itinerary scheme.

argument used to transfer from the formal normalization to the actual one (the analytic part of the normalization procedure) will be not optimal in the case of contraction mappings of a finite smoothness. In [25] the authors announced some results by M. Gyusinsky for the case of a finite differentiability but the smoothness assumptions are seen to be not optimal. Thirdly, the authors of [20] assumed that the invariant subspaces $M_{\bar{s}}$ are spectrally separated rather than exponentially separated. This means that conditions i) and ii) of the Proposition 3 below are imposed on the union of all the spectra of operators $J_i$ rather than on each spectrum. Notice in this connection that, due to Remark in Appendix A of [14], the case of the exponential separation is indeed more typical in applications than the case of the spectral separation[1]. Notice also that, in contrast to [20] and [13] (see below), we prefer to solve appearing functional equations by the direct contraction arguments rather than via analyzing the equations obtained by iteratively substituting along the orbit. We state three Theorems (in fact, the Normal Form Theorem and the two Semi-Centralizer Theorems) which generalize the corresponding two results of [20] related to the existence of "sub-resonance normal forms" and to the structure of the centralizers for these normal forms.

Under the absence of resonances (the case $\mathcal{R} = \emptyset$), the nonautonomous normalization constructed below is reduced to the nonautonomous linearization that has been discussed in detail in [14]. A paper by Y. Yomdin [38] was cited therein, in which the construction of the nonautonomous linearization for a bi-infinite sequence of contractive mappings has appeared. It happened, however, that the nonautonomous normalizations (in particular, linearizations) for systems of differential equations analytic in phase variables have been considered earlier by Y. Sibuya [35] and V.V. Kostin [27, 28] (see also a preprint-like textbook [29]). On the other hand, these authors proved the convergence of the formal normalizing expansions (the analytic part of the normalization procedure) by the use of the classical analytic tools, in contrast to Y. Yomdin who applied the hyperbolic theory geometric in character which can be utilized in the smooth (non-analytic) case (see Section 3). So, Y. Sibuya dealt with the general case of bounded coefficients and constant linear part and looked for normalizing transformations bounded over the whole infinite time interval. The possibility of including parameters was discussed and the convergence of the normalizing transformation in the case where the spectrum of the linear part lies in the left complex half-plane was proved. V.V. Kostin considered the case where the linear part depends, generally speaking, on time but it takes the triangular or even Jordan form and the boundedness conditions are not assumed (in particular, in view of sharp restrictions for terms left in the normal form). The main attention was drawn to the convergence problem for nonautonomous normalizing transformations. The convergence conditions obtained are of an integral form. Notice also that the papers of the authors cited here contain a transfer of the classical results [8] (described in Section 3 below) to the nonautonomous case. Inheriting the almost-periodicity of the system by its normal form and by the normalizing transformation was also shown. As for the paper by Y. Yomdin [38], we notice that the simplicity of the spectra and the one-dimensionality of the associated invariant subspaces that were assumed therein are not essential. On the other hand, he considered a slightly more general case where the sequences of mappings and linearizing transformations grow not faster than exponentially. This generalized situation can also be easily described in terms of itinerary schemes as it will be done in Remark 11 below. Y. Yomdin dealt also with some special situation where all the elements of the sequence of differentials coincide and possess a simple spectrum that lies in the Siegel domain,

[1] N. Fenichel [18, Part II] has proposed a simple but somewhat artificial example of bundles which are exponentially separated but not spectrally separated. In contrast to it, our example presented in [14] is very natural and typical in the smooth dynamical systems. I.U. Bronshtein [10] has noticed that one can construct an $C^1$-open set of diffeomorphisms on the two-dimensional torus (originating from an algebraic automorphism) which satisfy the exponential separation condition but are not spectrally separated.

and, moreover, all the mappings are analytic (we treat a situation where the spectra lie inside the unit circle). Then the presence of the nonautonomous linearization is proved via the usage of the KAM-theory.

The subject of the present paper is, on the one hand, an extension of the corresponding results of [14] to the general case of the presence of resonances, and, on the other hand, it is an immediate generalization and improvement of the results obtained by M. Gyusinsky and A. Katok. Thus, many definitions and results, including the basic Proposition 3, which are presented here are proper generalizations of the concepts of [14] to the general resonance case. The nonautonomous normalization constructed was already applied to the many-dimensional non-integrability problem, see announcement of some results in [15].

Among other nearby results, one should mention the paper by D. De Latte [13] who proved the non-autonomous variant of the Moser theorem on the normal forms for hyperbolic area-preserving two-dimensional maps. Both the analytic and $C^\infty$ cases are considered and the non-autonomous normal forms are shown to provide a countably many invariants that form obstructions to smooth conjugacy of symplectic Anosov diffeomorphisms on the torus. In fact, the general case of extensions of topological dynamical systems (homeomorphisms on compact spaces) by smooth hyperbolic area-preserving maps of $\mathbb{R}^2$ is also considered.

The structure of the paper is as follows. In Section 1 some preliminary definitions and results for number sets are presented. In the context of the paper, an important notion of a stable set $\widetilde{\mathcal{R}}$ is introduced. Section 2 is devoted to the concept of the narrow band spectrum condition which is reproduced from [20] (see also [25]). It is established that the validity of this condition for the spectrum is necessary and sufficient for the existence of an associated set $\widetilde{\mathcal{R}}$ which is finite and stable. In Section 3 some results on normal forms will be recalled and their nonautonomous analogs will be obtained. For this purpose, the basic Proposition 3 is stated. Section 4 is devoted to proving this Proposition and further discussions. The importance of the stability condition for $\widetilde{\mathcal{R}}$ (achieved provided that the nonautonomous version of the narrow band spectrum condition is satisfied) is indicated. Namely, it is shown that the (partial) normal forms associated to the stable set $\widetilde{\mathcal{R}}$ are the most adequate generalizations of the usual normal forms. In particular, the corresponding Semi-Centralizer Theorem generalizing the Centralizer Theorem in [8] (for the formal case) and in [5,20] (for the contractive smooth case) is established. In Section 5, the extensions of homeomorphisms on compact spaces by smooth contractive mappings are considered and the corresponding Normal Form Theorem and two Semi-Centralizer Theorems are stated and proved. These theorems generalize the results by M. Gyusinsky and A. Katok as was mentioned above. Some modifications of the conditions of these theorems are considered to avoid an explicit usage of a particular Lyapunov norm. Three Historical comments are included in the text of the paper, to review the related papers and to explain the novelty of the results presented.

The author is thankful to A. Katok and D. De Latte for sending offprints of their papers [20] and [13].

## 1. Preliminary definitions for number sets

Let complex numbers $\lambda_1, \dots, \lambda_n$ satisfy the condition $0 < |\lambda_j| < 1$. Consider partitions of the set $\{\lambda_j\}$ into classes $\Lambda_r$.

**Definition 1.** A partition is said to be *strongly ordered* if the classes $\Lambda_r$ are sets of numbers, $\lambda_j$, whose moduli lie, respectively, in some non-intersecting intervals.

The classes $\Lambda_r$ will be placed in the order of non-decreasing moduli of the numbers $\lambda \in \Lambda_r$.

Let $\{\lambda_{i,j} : 1 \leqslant j \leqslant n\}$ be some number sets consisting of $n$ elements such that $0 < |\lambda_{i,j}| < 1$. Suppose that there are strongly ordered partitions

$$\xi_i = \{\Lambda_{i,1}, \dots, \Lambda_{i,p}\}$$

of these sets. Moreover, assume that the quantities of classes, $p$, and the numbers of elements in each class $\operatorname{card} \Lambda_{i,s} = c_s$ $(1 \leqslant s \leqslant p)$ do not depend on $i$. Here card denotes the cardinality (the quantity of elements) of a set.

**Definition 2.** Strongly ordered partitions $\xi_i$ which satisfy these conditions will be said to be *concordant*.

Analogously, one can consider complex numbers $\mu_j$ such that $|\mu_j| > 1$ and introduce strongly ordered partitions and concordant partitions. Obviously, this case is conjugate with the previous one by the inversion operation $\mu_j \mapsto \lambda_j = \mu_j^{-1}$. Therefore, elements of partitions for this case will be placed in the order of non-increasing moduli.

In the sequel, we will write $z^m = \prod_{j=1}^n z_j^{m_j}$ and $|m| = \sum_j m_j$ for any collection $z = (z_1, \dots, z_n)$ and any multiindex $m = (m_1, \dots, m_n)$ with non-negative integer components $m_j$. Under discussing the nonautonomous normal forms of contractive mappings, the usual multiplicative Poincaré non-resonance conditions

$$\lambda_s \neq \lambda^m, \tag{1'}$$

where $\lambda = \{\lambda_j\}$ is the spectrum at the fixed point and the multiindex $m$ is such that $|m| = \sum_j m_j \geqslant 2$, will be replaced by that

$$|\lambda_s| \neq |\lambda^m| \tag{1}$$

(cf. [38,20,14]). These conditions are obviously satisfied if $|m| > D$ where

$$D = \ln \min_j |\lambda_j| / \ln \max_j |\lambda_j| \geqslant 1. \tag{2}$$

Let $\mathcal{R}$ be some finite (possibly, empty) set of couples $(s, m)$ with multiindex $m$ satisfying $|m| \geqslant 2$ and index $s$ such that all inequalities (1) with $(s, m) \notin \mathcal{R}$ are valid. In other words, all the couples $(s, m)$, for which the resonance conditions $|\lambda_s| = |\lambda^m|$ are met, belong to $\mathcal{R}$, i.e. $|\lambda_s| = |\lambda^m| \Rightarrow (s, m) \in \mathcal{R}$. Therefore, the set $\mathcal{R}$ will be conditionally said to *contain resonances*. Obviously, there is a smallest set, $\mathcal{R}_0$, containing resonances (the set *consisting of the resonances*). In Section 4, while discussing the usual normal forms, we will also use the *exact resonances* $\lambda_s = \lambda^m$ and the sets $\mathcal{R}$ which may be infinite (since the spectrum may lie from both the sides of the unit circle) and which contain these resonances, i.e., such that $\lambda_s = \lambda^m \Rightarrow (s, m) \in \mathcal{R}$. The corresponding smallest set, $\mathcal{R}_0^{(ex)}$, consisting of the exact resonances, will be also utilized therein.

For a given partition $\xi = \{\Lambda_1, \dots, \Lambda_p\}$, the sets $\mathcal{R}$ of couples $(s, m)$ satisfying the following condition have a great importance: if $(s, m) \in \mathcal{R}$ where $m = (m_1, \dots, m_n)$, the numbers $\lambda_s$, $\lambda_{s'}$ belong to the same class of $\xi$, and unordered set $\operatorname{St}(m') = \{\underbrace{1, \dots, 1}_{m'_1}, \dots, \underbrace{n, \dots, n}_{m'_n}\}$ is obtained from the set $\operatorname{St}(m) = \{\underbrace{1, \dots, 1}_{m_1}, \dots, \underbrace{n, \dots, n}_{m_n}\}$ via replacement of some numbers $j$ by numbers $j'$ such that $\lambda_j$ and $\lambda_{j'}$ belong, respectively, to the same classes of $\xi$, then $(s', m') \in \mathcal{R}$ where $m' = (m'_1, \dots, m'_n)$.

**Definition 3.** The set $\mathcal{R}$ satisfying this condition will be said to be *subordinate to the partition* $\xi$ or $\xi$*-subordinate*.

Notice that this definition allows some equal numbers $\lambda_j$ with different indices $j$ to belong to different classes of the partition (in fact, the partition of the set of indices $\{1,\dots,n\}$ is considered). Given an index $s$ and a multiindex $m=(m_1,\dots,m_n)$, introduce[2] $\tilde{s}\,(1\leqslant\tilde{s}\leqslant p)$, where $\lambda_s\in\Lambda_{\tilde{s}}$, and $\widetilde{m}=(\widetilde{m}_1,\dots,\widetilde{m}_p)$, where $\widetilde{m}_t=\sum\limits_{\lambda_j\in\Lambda_t} m_j\geqslant 0$, and denote $(\tilde{s},\widetilde{m})=\Psi(s,m)$. Obviously, $|\widetilde{m}|=|m|$. Then for any set $\widetilde{\Phi}$ of couples $(\tilde{s},\widetilde{m})$, the set $\Phi=\Psi^{-1}(\widetilde{\Phi})$ is subordinate to the partition $\xi$, and for any set $\Phi$ of couples $(s,m)$, the condition $\Phi=\Psi^{-1}\big(\Psi(\Phi)\big)$ means exactly that $\Phi$ is subordinate to the partition $\xi$. The operation $\Psi$ is determined by the partition $\xi$ and for concordant partitions the corresponding $\Psi$'s are obviously coincident. We will sometimes write $\Psi_\xi$ in order to emphasize the dependence of $\Psi$ on $\xi$.

Obviously, the set consisting of the resonances, $\mathcal{R}_0$, is $\xi_0$-subordinate and the set consisting of the exact resonances, $\mathcal{R}_0^{(ex)}$, is $\xi_0^{(ex)}$-subordinate, where $\xi_0$ is the partition into the classes of numbers having the equal moduli and $\xi_0^{(ex)}$ is the partition into the classes of equal numbers.

**Definition 4.** A partition $\xi$ of the set $\{\lambda_j\}$ is said to be *strongly* $\widetilde{\mathcal{R}}$*-right* if the set $\mathcal{R}=\Psi_\xi^{-1}(\widetilde{\mathcal{R}})$ contains the resonances of $\{\lambda_j\}$ and the following condition is satisfied: $(*)$ two numbers $\lambda_{j'},\lambda_{j''}$ belong to the same class if and only if in all inequalities (1) such that $(s,m)\notin\mathcal{R}$ the signs persist under replacement of all factors $\lambda_{j'}$ by $\lambda_{j''}$ or all factors $\lambda_{j''}$ by $\lambda_{j'}$ in the left-hand side and right-hand side simultaneously.

*Remark 1.* (cf. Remark 1 of [14].) For a given $\mathcal{R}$, the condition $(*)$ indeed determines an equivalence relation. However, the existence of a strongly $\widetilde{\mathcal{R}}$-right partition does not yet follow from this, since this partition was already used in the auxiliary definition of the subordinate set $\mathcal{R}$. Due to the same reason, the strongly $\widetilde{\mathcal{R}}$-right partition does exist for not any set $\mathcal{R}$ containing resonances. A strongly $\widetilde{\mathcal{R}}$-right partition, if exists, is unique and is also strongly ordered. The definition of the strongly $\widetilde{\mathcal{R}}$-right partition admits an equivalent formulation dealing with the replacement of an arbitrary part of the factors. This definition admits also a form which will be indicated in the proof of Proposition 1.

Another fact is that we prefer to say about $\widetilde{\mathcal{R}}$-right partitions rather than about $\mathcal{R}$-right ones. This complicates slightly the notations but will allow us to consider in the sequel $\widetilde{\mathcal{R}}$-right partitions for the sets of different cardinalities[3].

According to Definition 4, the attempt to construct the strongly $\widetilde{\mathcal{R}}$-right partition can lead to the necessity to modify the set $\mathcal{R}$ itself. Since the set $\mathcal{R}$ has to contain the resonances, one should go ahead to enlarge it (possibly, choosing originally the smallest set containing resonances). Generally speaking, the bigger set $\mathcal{R}$, the less non-resonance conditions appear in Definition 4 and, accordingly, the rougher is the strongly $\widetilde{\mathcal{R}}$-right partition. In the extreme case, the finally modified set $\mathcal{R}$ contains all the couples $(s,m)$ such that $|m|\leqslant N$ where $N>D$ and the corresponding strongly $\widetilde{\mathcal{R}}$-right partition consists of the single class. The following Proposition shows, however, that to any set containing resonances, one can associate in a natural manner the smallest ambient set $\mathcal{R}^*$ and, respectively, the finest strongly $\widetilde{\mathcal{R}}^*$-right partition.

**Proposition 1.** *Let $\mathcal{R}$ be a set containing resonances. Then there is a smallest set $\mathcal{R}^*\supseteq\mathcal{R}$, for which the strongly $\Psi_{\xi^*}(\mathcal{R}^*)$-right partition $\xi^*$ is defined. Precisely speaking, if $\mathcal{R}'\supseteq\mathcal{R}$ and*

[2]Introducing the index $\tilde{s}$, we correct a minor inaccuracy passed in [14].

[3]This is the difference from [14] which is made to remove an inaccuracy having been passed in condition ii) of Proposition 5 of [14] while stating Proposition 3 below.

*$\xi'$ is a strongly $\Psi_{\xi'}(\mathcal{R}')$-right partition then $\mathcal{R}^* \subseteq \mathcal{R}'$ and $\xi^* \succcurlyeq \xi'$ (the partition $\xi'$ is a roughing of $\xi^*$). If $\mathcal{R} = \emptyset$ then $\mathcal{R}^* = \emptyset$.*

*Proof.* If $\mathcal{R} = \emptyset$ then $\mathcal{R}$ is certainly subordinate to the partition, and the condition $(*)$ gives an equivalence relation that determines the partition sought for. Therefore, $\mathcal{R}^* = \emptyset$.

In the general case, the idea is to construct the set $\mathcal{R}^*$ sought for as a result of minimal enlargements of the set $\mathcal{R}$ whose necessity is stipulated by the subordination condition. We replace the condition $(*)$ by the following condition that determines (after a revision below) an equivalence relation for any set $\mathcal{R}$: $(**)$ two numbers $\lambda_{j'}$, $\lambda_{j''}$ belong to the same class if and only if there is a chain of numbers $\lambda_{j'} = \lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_{k-1}}, \lambda_{j_k} = \lambda_{j''}$ such that the numbers $\lambda_{j_s}$, $\lambda_{j_{s-1}}$ $(0 \leqslant s < k)$ constituting any segment satisfy condition $(*)$. However, the relation determined by $(**)$ can be non-symmetric because of the fact that, as a result of a transformation indicated in $(*)$, one can obtain from the inequality (1) with some couple $(s, m) \notin \mathcal{R}$ an analogous inequality with a couple $(s', m') \in \mathcal{R}$. Therefore, we modify condition $(*)$ by considering the transformation only, for which both the original couple $(s, m)$ and the resulted couple $(s', m')$ do not belong to $\mathcal{R}$. Now the relation determined by the condition $(**)$ will indeed be an equivalence relation. Notice that condition $(*)$ and both the modified and non-modified conditions $(**)$ are equivalent if the set $\mathcal{R}$ is subordinate to the partition. We will perform the following two-stage iterative process. Some partition $\xi$ of the set $\{\lambda_j\}$ is determined, accordingly to $(**)$, for the set $\mathcal{R}$. Then we perform the minimal enlargement of the set $\mathcal{R}$, after which $\mathcal{R}$ will become to be subordinate to the given partition $\xi$. Then repeat this procedure. At each step of this iterative process we will obtain a set $\mathcal{R}^{(i)}$ and a partition $\xi^{(i)}$ such that if $\mathcal{R}' \supseteq \mathcal{R}$ and $\xi'$ is a strongly $\Psi_{\xi'}(\mathcal{R}')$-right partition then $\mathcal{R}^{(i)} \subseteq \mathcal{R}'$ and $\xi^{(i)} \succcurlyeq \xi'$. It suffices to show only that the non-decreasing sequence of the sets $\mathcal{R}^{(i)}$ and the sequence of the non-refined partitions $\xi^{(i)}$ are stabilized at some step, thus providing the required $\mathcal{R}^*$ and $\xi^*$ (since the subordination condition will be now satisfied). But this is certainly the case because of the presence of natural bounds from above for the quantity of the elements in $\mathcal{R}^{(i)}$ and that from below for the quantity of the elements in $\xi^{(i)}$. $\square$

Some trouble is the fact that the set $\mathcal{R}^*$ can be much more wider than $\mathcal{R}$ even if $\mathcal{R}$ is the smallest set containing resonances, $\mathcal{R}_0$. A simplest example will be described below. A possible alternative is to allow a partition $\xi$ to be finer for a given set $\mathcal{R}$ since this will lead to not so intensive enlargement of $\mathcal{R}^{(i)}$ while constructing $\mathcal{R}^*$. This justifies the following

**Definition 4′.** A partition $\xi$ is said to be $\widetilde{\mathcal{R}}$*-right* if the set $\mathcal{R} = \Psi_{\xi}^{-1}(\widetilde{\mathcal{R}})$ contains the resonances of $\{\lambda_j\}$ and the following condition is satisfied: $(*')$ if two numbers $\lambda_{j'}, \lambda_{j''}$ belong to the same class then in all inequalities (2) such that $(s, m) \notin \mathcal{R}$ the signs persist under replacement of an arbitrary part of factors $\lambda_{j'}$ by $\lambda_{j''}$ or an arbitrary part of factors $\lambda_{j''}$ by $\lambda_{j'}$ in the left-hand side and right-hand side simultaneously.

Equivalent formulations (here Definition $4'_1$ and a slightly modified Definition $4'_2$ are convenient to apply when some classes $\Lambda_r$ are infinite sets; however in this case Definitions 4 and $4'_1$ are equivalent, but Definition $4'_2$ is not):

**Definition $4'_1$.** A partition $\xi = \{\Lambda_1, \dots, \Lambda_p\}$ is said to be $\widetilde{\mathcal{R}}$*-right* if all the inequalities

$$|\lambda_{\tilde{s}}| \neq |\underbrace{\lambda_{1,1} \cdots \lambda_{1,\widetilde{m}_1}}_{\widetilde{m}_1} \cdots \underbrace{\lambda_{p,1} \cdots \lambda_{p,\widetilde{m}_p}}_{\widetilde{m}_p}| \tag{3}$$

are satisfied for $(\tilde{s}, \widetilde{m}) \notin \widetilde{\mathcal{R}}$ and arbitrary $\lambda_{\tilde{s}} \in \Lambda_{\tilde{s}}$, $\lambda_{j,k} \in \Lambda_j$ and the signs of these inequalities do depend on $(\tilde{s}, \widetilde{m})$ only.

**Definition $4'_2$.** A partition $\xi = \{\Lambda_1, \dots, \Lambda_p\}$ is said to be $\widetilde{\mathcal{R}}$-*right* if

$$[\underline{\lambda}_{\tilde{s}}, \overline{\lambda}_{\tilde{s}}] \cap [\underline{\lambda}^{\widetilde{m}}, \overline{\lambda}^{\widetilde{m}}] = \emptyset$$

for all $(\tilde{s}, \widetilde{m}) \notin \widetilde{\mathcal{R}}$ where two collections $\underline{\lambda} = (\underline{\lambda}_1, \cdots, \underline{\lambda}_p)$ and $\overline{\lambda} = (\overline{\lambda}_1, \cdots, \overline{\lambda}_p)$ are introduced such that $\underline{\lambda}_{\tilde{s}} = \min_{\lambda_j \in \Lambda_{\tilde{s}}} |\lambda_j|$ and $\overline{\lambda}_{\tilde{s}} = \max_{\lambda_j \in \Lambda_{\tilde{s}}} |\lambda_j|$.

Now the partition $\xi_0$ into the classes of numbers having the equal moduli is always $\Psi_{\xi_0}(\mathcal{R}_0)$-right and the partition $\xi_0^{(ex)}$ into the classes of equal numbers is always $\Psi_{\xi_0^{(ex)}}(\mathcal{R}_0^{(ex)})$-right.

The key point of the motivation is as follows. Below one has to deal with the set $\mathcal{R}$ and the $\Psi_\xi(\mathcal{R})$-right partition $\xi$ such that the set $\mathcal{R}$ is not "too large" (i.e., it is stable in the sense of Definition 6 below). Therefore, one has to accept the refinement of the partition $\xi$.

**Definition 5.** Concordant and $\widetilde{\mathcal{R}}$-right partitions $\xi_i$ will be said to be $\widetilde{\mathcal{R}}$-*right concordant* if in addition the following condition holds: in all inequalities (3), where $\lambda_{\tilde{s}} \in \Lambda_{i,\tilde{s}}$, $\lambda_{j,k} \in \Lambda_{i,j}$, and $(\tilde{s}, \widetilde{m}) \notin \widetilde{\mathcal{R}}$ the signs do not depend on $i$.

Definitions 4 and 5 are the immediate generalizations of definitions in [14]. In particular, a strongly $\emptyset$-right partition is a *right* one in the sense of [14], and strongly $\emptyset$-right partitions that are $\emptyset$-right concordant are *right concordant* in that sense. Note that in the present case the entire parts of the numbers $D_i = D$ determined accordingly to (2) can be different for different $\widetilde{\mathcal{R}}$-right concordant sets. However, they possess a finite upper bound, due to the finiteness of $\mathcal{R}$.

These definitions are immediately extendible to sets of numbers $\mu_j$ such that $|\mu_j| > 1$.

It is convenient sometimes to consider also couples $(s, m)$ such that $|m| = 1$ (these couples will be said to be *linear*) and to supplement to the set $\mathcal{R}$, subordinated to the partition $\xi$, all the couples $(i, \mathbf{e}_j)$ such that $\lambda_i$ and $\lambda_j$ belong to the same element of the partition $\xi$ (these linear couples will be said to be $\xi$-*trivial*). Here $\mathbf{e}_i$ denotes the basis multiindex such that $(\mathbf{e}_i)_j = \delta_{i,j}$, the Kronecker symbol. Thus, for the given partition $\xi$, one can identify the $\xi$-subordinate sets $\mathcal{R}$ and the corresponding "extended" sets. Generally speaking, one can consider "extended" sets containing arbitrary linear couples $(s, m)$. Then the above notation of $\xi$-subordinate set remains valid. The only restriction in the sequel will be that for every index $s$ there is a couple $(s, m) \in \mathcal{R}$ (see also Remark 3 below).

Let $\mathcal{R}_1$, $\mathcal{R}_2$ be two such "extended" sets of couples $(s, m)$ subordinated to the same partition $\xi$. Put $(s, m) \ltimes \big((s^{(1)}, m^{(1)}), \dots, (s^{(q)}, m^{(q)})\big)$ to be equal to $(s, m')$ with $m' = \sum_{t=1}^{q} m^{(t)}$ if $q = |m|$ and every index $j$ is met exactly $m_j$ times among the numbers $s^{(1)}, \dots, s^{(q)}$. Otherwise the $\ltimes$-expression will be assumed to be undefined. In other words, the set $\mathrm{St}(m')$ is obtained from the set $\mathrm{St}(m)$ via replacement of each element $j$ by the corresponding subset $\mathrm{St}(m^{(k)})$ such that $s^{(k)} = j$ where different indices $k$ are associated to different elements $j$. Furthermore, let $\mathcal{R}_1 \ltimes \mathcal{R}_2$ be the set of all such the sensible $\ltimes$-expressions where $(s, m) \in \mathcal{R}_1$ and all $(s^{(t)}, m^{(t)}) \in \mathcal{R}_2$. Obviously, the set $\mathcal{R}_1 \ltimes \mathcal{R}_2$ is also subordinate to the partition $\xi$, $\mathcal{R}_1 \ltimes \mathcal{R}_2 \supseteq \mathcal{R}_1 \cup \mathcal{R}_2$ if $\mathcal{R}_1$, $\mathcal{R}_2$ contain linear $\xi$-trivial couples, and $\Psi(\mathcal{R}_1 \ltimes \mathcal{R}_2) = \Psi(\mathcal{R}_1) \ltimes \Psi(\mathcal{R}_2)$ (the sets in the right-hand side of the last equality are subordinate to the partitions into the one-element classes). Moreover, the set operation $\ltimes$ is associative, but not commutative which is evident from the following Remark 2 explaining the actual reason for introducing the set operation $\ltimes$.

According to Definition 8 below one can associate in a natural way to each $\xi$-subordinate ("extended") set $\mathcal{R}$ some class of the polynomial (or, more generally, formal power series) mappings $g : \mathbb{C}^n \to \mathbb{C}^n$, the so-called $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal forms. We explain briefly this notation in a self-contained and slightly simplified way. Let $\mathbf{e}_j$ $(1 \leqslant j \leqslant n)$ be some basis in $\mathbb{C}^n$ and $M_{\tilde{s}}$ be the subspace spanned by all the $\mathbf{e}_j$'s such that $\lambda_j \in \Lambda_{\tilde{s}}$. Let $\widetilde{\mathcal{R}} = \Psi(\mathcal{R})$ with the mapping $\Psi$

associated to $\xi$. A $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-*quasi-normal form* is a polynomial or formal power series mapping $g:\mathbb{C}^n\to\mathbb{C}^n$ which takes in the basis $\{\mathbf{e}_j\}$ the form

$$\big(g(x)\big)_s=\sum a_{s,m}x^m$$

such that $a_{s,m}\neq 0$ only if $(s,m)\in\mathcal{R}$.

*Remark 2.* Denote by $\mathcal{F}_{\widetilde{\mathcal{R}}}$ the set of the $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-quasi-normal forms where $\widetilde{\mathcal{R}}$ is any set of couples $(\tilde{s},\widetilde{m})$. Let $\widetilde{\mathcal{R}}_1$, $\widetilde{\mathcal{R}}_2$ be two sets of couples $(\tilde{s},\widetilde{m})$. Then $g_1\circ g_2\in\mathcal{F}_{\widetilde{\mathcal{R}}_1\ltimes\widetilde{\mathcal{R}}_2}$ if $g_i\in\mathcal{F}_{\widetilde{\mathcal{R}}_i}$ $(i=1,2)$, and this result is unimprovable in the sense that $\widetilde{\mathcal{R}}_1\ltimes\widetilde{\mathcal{R}}_2$ is the minimal set $\widetilde{\mathcal{R}}$ such that $g_1\circ g_2\in\mathcal{F}_{\widetilde{\mathcal{R}}}$ for all $g_i\in\mathcal{F}_{\widetilde{\mathcal{R}}_i}$. Moreover, the linear hull of all such $g_1\circ g_2$ coincides with the whole $F_{\widetilde{\mathcal{R}}_1\ltimes\widetilde{\mathcal{R}}_2}$.

**Definition 6.** The set $\mathcal{R}$ ("extended" or subordinated to a given partition) is said to be *stable* if $\mathcal{R}\ltimes\mathcal{R}=\mathcal{R}$.

Obviously, a set $\mathcal{R}$ subordinated to a given partition $\xi$ is stable if and only if the corresponding set $\Psi(\mathcal{R})$ is stable. The smallest set containing resonances $\mathcal{R}_0$ (containing exact resonances $\mathcal{R}_0^{(ex)}$, respectively), is always stable.

The set $\mathcal{R}$ can be not stable even if the partition $\xi$ is strongly $\Psi_\xi(\mathcal{R})$-right, which is seen in the following simple

**Example 1.** If $\{\lambda_j\}=\{2,4\}$ then the smallest set containing resonances, $\mathcal{R}_0$, has a single element, and for $\mathcal{R}=\mathcal{R}_0$, the partition $\xi^*$ consists of one class, and $\mathcal{R}^*$ consists of all the six couples $(s,m)$ such that $|m|=2$. The set $\underbrace{\mathcal{R}^*\ltimes\mathcal{R}^*\ltimes\cdots\ltimes\mathcal{R}^*}_{k}$ consists of all the couples $(s,m)$ such that $|m|\leqslant k+1$.

## 2. The narrow band spectrum condition

Recall an important concept introduced originally in [20] (see also [25]).

Let $\{\lambda_j\}$ be a finite collection of numbers such that $0<|\lambda_j|<1$ and let $\xi=\{\Lambda_1,\dots,\Lambda_p\}$ be a strongly ordered partition of the set $\{\lambda_j\}$, $\overline{\lambda}=\max\limits_j|\lambda_j|$, $\underline{\lambda}_s=\min\limits_{\lambda_j\in\Lambda_s}|\lambda_j|$, $\overline{\lambda}_s=\max\limits_{\lambda_j\in\Lambda_s}|\lambda_j|$ (then $\overline{\lambda}=\overline{\lambda}_p$).

**Definition 7.** The strongly ordered partition $\xi$ is said to satisfy the *narrow band* condition if $\overline{\lambda}_s\overline{\lambda}<\underline{\lambda}_s$ (or, equivalently, $\overline{\lambda}<\underline{\lambda}_s/\overline{\lambda}_s$) for all $1\leqslant s\leqslant p$.

This condition means that the partition $\xi$ is not too rough. In particular, $\xi$ may be the partition into the classes of numbers with equal moduli.

**Proposition 2.** *Given a partition $\xi$, there is a finite stable $\xi$-subordinate set $\mathcal{R}$ containing (exact) resonances and such that $\xi$ is $\Psi_\xi(\mathcal{R})$-right, if and only if $\xi$ satisfies the narrow band condition.*

*Remark 3.* In all the further applications, $\mathcal{R}$ will be a set containing (exact) resonances. Thus, all the $\xi$-trivial linear couples belong to $\mathcal{R}$. Therefore, without loss of generality, one can account that only $\xi$-trivial couples are the linear couples presented in $\mathcal{R}$, since removing the other linear couples from $\mathcal{R}$ does not violate the stability of $\mathcal{R}$.

We preface the proof of Proposition 2 by recalling some basic concepts and by the graphic interpretation as follows.

Firstly, we recall the definition of a *topological Markov chain* (TMC) and how it is determined by a directed graph. The term "topological Markov chain" was introduced by V. M. Alekseev [2, Part 1] although the corresponding mathematical object can be found in earlier papers. In the English mathematical literature it is usually referred to as *"subshift of finite type"* going back to S. Smale [36] (because there is a natural one-to-one topological equivalence between topological Markov chains and Smale's subshifts of finite type). Consider a finite set (*"alphabet"*) $\mathcal{L}$ containing $m < \infty$ elements and equipped with the discrete topology. One can put $\mathcal{L} = \{1, \dots, m\}$. Denote by $\Omega = \prod_{n=-\infty}^{+\infty} \mathcal{L}$ the Tychonoff product of a countable collection of copies of the space $\mathcal{L}$, i.e., the space of doubly-infinite sequences $\omega = \left[\omega_n \in \mathcal{L} : -\infty < n < +\infty\right]$ equipped with the topology such that

$$\omega^{(k)} = \left[\omega_n^{(k)} : n \in \mathbb{Z}\right] \to \omega = \left[\omega_n : n \in \mathbb{Z}\right] \quad \text{as} \quad k \to +\infty \iff \omega_n^{(k)} \to \omega_n \quad \text{as} \quad k \to +\infty \quad \text{for each} \quad n \in \mathbb{Z},$$

and by $T \colon \Omega \to \Omega$, the shift homeomorphism by one symbol to the left. Each zero-unit matrix $\Pi = \left(\pi_{i,j}\right)_{i,j=1}^{m}$ of size $m \times m$ defines a closed $T$-invariant compact subset $\Omega^{\Pi} \subset \Omega$ by the following condition:

$$\omega = [\omega_n] \in \Omega^{\Pi} \iff \pi_{\omega_n, \omega_{n+1}} = 1 \quad \text{for all} \quad n.$$

The restriction $T \mid \Omega^{\Pi}$ is called a TMC with $m$ states

$$\{1, \dots, m\} = \mathcal{L}$$

and *transition matrix* $\Pi$. Each sequence $\omega \in \Omega$ is called a *"word"*, and a sequence $\omega \in \Omega^{\Pi}$ (containing only admissible transitions determined by the matrix), an *"admissible word"*. If $\pi_{i,j} = 1$ for all the pairs $(i, j)$ (i.e., the transition from every state into every another one is possible) then $\Omega = \Omega^{\Pi}$ and the TMC $T \mid \Omega$ is called a *Bernoulli shift* on the set of $m$ elements (states).

It is convenient to represent a TMC by the directed graph with vertex set $\mathcal{L}$ and edges $\overrightarrow{(i,j)}$ whose initial vertices (origins), $i$, and terminal vertices (ends), $j$, form pairs $(i, j)$ such that $\pi_{i,j} = 1$. Without loss of generality one can suppose that a TMC is *"conservative"*, i.e., each vertex $i \in \mathcal{L}$ happens to be a *"transition"*. This means that there are vertices $j^+, j^- \in \mathcal{L}$ such that $\pi_{j^-,i} = \pi_{i,j^+} = 1$. Indeed, removing the remaining vertices, if any, does not lead to a change of the set $\Omega^{\Pi}$ of admissible words.

Now let $\mathcal{L}$ be the set of all the indices $\tilde{s}$. Introduce a non-negative matrix $\Pi = (\pi_{i,j})_{i,j \in \mathcal{L}}$ such that $\pi_{i,j} > 0$ if and only if $\widetilde{m}_j \geqslant 1$ for some $(i, \widetilde{m}) \in \widetilde{\mathcal{R}}$. Then, by some permutation of the indices, the matrix $\Pi$ can be represented in the *normal* block-triangle form (see, for instance, F.R. Gantmacher [19]). This means that there are some pairwise non-intersecting subsets $\mathcal{L}_l \subset \mathcal{L}$ such that $\pi_{i,j} > 0$ only if $i$ and $j$ belong to the same $\mathcal{L}_l$ or if $i < j$. One can associate to $\Pi$ a directed graph $\Gamma$ with the vertex set $\mathcal{L}$ which consists of all the edges $\overrightarrow{(i,j)}$ such that $\pi_{i,j} > 0$ and then the TMC as was described above. The edges whose origins and ends belong to the same $\mathcal{L}_l$ constitute a *connected* subgraph $\Gamma_l$, i.e., for any two vertices $a, b$ of the graph $\Gamma_l$, there is a path on $\Gamma_l$ with origin $a$ and end $b$. All the closed circuits (paths) form exactly the subgraphs $\Gamma_l$.

Introduce the notation $\ltimes^k \mathcal{R} = \underbrace{\mathcal{R} \ltimes \cdots \ltimes \mathcal{R}}_{\substack{k+1 \\ \text{copies of } \mathcal{R}}}$. The graph $\Gamma$ can be used to characterize particularly the set $\ltimes^k \widetilde{\mathcal{R}}$ as follows. Given $\tilde{s} \in \mathcal{L}$ and $k \geqslant 0$, one has $\widetilde{m}_j > 0$ for some $(\tilde{s}, \widetilde{m}) \in \ltimes^k \widetilde{\mathcal{R}}$ if and only if there is a path of length $(k+1)$ (i.e., consisting of $(k+1)$ edges) with the origin at $\tilde{s}$ and the end at $j$. In a more general framework, the correspondence that associates a TMC to a set $\widetilde{\mathcal{R}}$ possesses the following contravariance property: if the TMC $T_i$ corresponds to the set $\widetilde{\mathcal{R}}_i$, $i = 1, 2$, then $T_2 \circ T_1$ corresponds to $\widetilde{\mathcal{R}}_1 \ltimes \widetilde{\mathcal{R}}_2$.

One can control also $|\widetilde{m}|$ using a more detailed description based on trees as follows. Consider trees whose vertices belong to $\mathcal{L}$ (maybe, with repetitions) and such that from each tree's vertex $i$, placed not on the top of the tree, there grows a bunch of tree's edges which consists of exactly $\widetilde{m}_j$ edges $\overrightarrow{(i,j)}$ for some couple $(i,\widetilde{m}) \in \widetilde{\mathcal{R}}$ (determining uniquely this bunch). Thus, some edges in one bunch may be identical but their ends, $j$'s, will be considered as distinct vertices of the tree. Assume that the heights (the quantity of the successive edges leading from the root to the top) of all the branches of the tree are equal exactly to $(k+1)$. In other words, every maximal ordered path embedded into the tree has length $(k+1)$. Then all the elements of $\ltimes^k\widetilde{\mathcal{R}}$ are described by these trees so that $(\tilde{s},\widetilde{m}) \in \ltimes^k\widetilde{\mathcal{R}}$ if and only if the following holds: there is a tree $\mathcal{T}$ such that $\tilde{s}$ is the root vertex of $\mathcal{T}$ and, for any $j$, exactly $\widetilde{m}_j$ top vertices of $\mathcal{T}$ are placed at a given $j \in \mathcal{L}$.

*Proof of Proposition 2.* Obviously, $\mathcal{R}_{0,\xi} = \Psi^{-1}\big(\Psi(\mathcal{R}_0)\big)$ is the smallest $\xi$-subordinate set containing resonances and $\bigcup_{k=1}^{\infty} \ltimes^k\mathcal{R}_{0,\xi}$ is the smallest $\xi$-subordinate stable set containing resonances. Obviously, these considerations remain valid if only exact resonances are considered and the set $\mathcal{R}_{0,\xi}$ is replaced by $\mathcal{R}_{0,\xi}^{(ex)} = \Psi^{-1}\big(\Psi(\mathcal{R}_0^{(ex)})\big)$.

We will say that a $\xi$-subordinate set $\mathcal{R}$ (or the associated set $\widetilde{\mathcal{R}} = \Psi(\mathcal{R})$) *contains a circuit* $(\tilde{s}^{(1)},\widetilde{m}^{(1)}),\dots,(\tilde{s}^{(t)},\widetilde{m}^{(t)}) = (\tilde{s}^{(1)},\widetilde{m}^{(1)})$ if $(\tilde{s}^{(i)},\widetilde{m}^{(i)}) \in \widetilde{\mathcal{R}}$ and $\widetilde{m}^{(i)}_{\tilde{s}^{(i+1)}} \geqslant 1$ for all $1 \leqslant i < t$. In addition, this circuit will be said to be *linear* if $\widetilde{m}^{(i)} = \mathbf{e}_{\tilde{s}^{(i+1)}}$ for all $i$.

Then the result desired is implied immediately by the following two auxiliary facts:

**1)** given a $\xi$-subordinate set $\mathcal{R}$, the set $\ltimes^k\mathcal{R}$ is stable for some (sufficiently large) $k$ if and only if $\mathcal{R}$ contains only linear circuits;

**2)** the set $\mathcal{R}_{0,\xi}$ contains only linear circuits if and only if the partition $\xi$ satisfies the narrow band condition.

*Proof of fact* **2)**. If the narrow band condition is not satisfied then $\overline{\lambda}_{\tilde{s}}\overline{\lambda} \in [\underline{\lambda}_{\tilde{s}},\overline{\lambda}_{\tilde{s}}]$ for some $\tilde{s}$ and, consequently, $(\tilde{s},\widetilde{m}) \in \widetilde{\mathcal{R}}$ with $\widetilde{m}_j = \delta_{j,p} + \delta_{j,\tilde{s}}$, i.e., $(\tilde{s},\widetilde{m})$ constitutes a non-linear one-segment circuit. From the other hand, if the narrow band condition is satisfied then $\widetilde{\mathcal{R}}_{0,\xi} = \Psi(\mathcal{R}_{0,\xi})$ is easily seen to take the "triangular form with the linear diagonal part" in the sense that if $(\tilde{s},\widetilde{m}) \in \widetilde{\mathcal{R}}_{0,\xi}$ then $m_{\tilde{s}'} = 0$ for $\tilde{s}' < \tilde{s}$ and, moreover, $m_{\tilde{s}'} = \delta_{\tilde{s},\tilde{s}'}$ provided that $m_{\tilde{s}} \geqslant 1$. Obviously, this implies the absence of non-linear circuits. Moreover, for further purposes, we note that for any $\xi$-subordinate set $\mathcal{R}$ containing (exact) resonances and possessing only linear circuits, the set $\widetilde{\mathcal{R}} = \Psi(\mathcal{R})$ will take the "triangular form with the linear diagonal part" (since the other linear couples are removed from $\mathcal{R}$, according to Remark 3).

*Proof of fact* **1)**. Obviously, some circuit always exists (it may consist of a single couple) due to the finiteness of $\mathcal{L}$. Then, under the presence of the above $(t-1)$-segment circuit, the set $\ltimes^{t-1}\mathcal{R}$ will contain a couple $(s^{(1)},m)$ such that $m_{s^{(1)}} \geqslant 1$ and $|m| \geqslant \sum\limits_{i=1}^{t-1}\big(|m^{(i)}|-1\big)+1$. Therefore, if this circuit is not linear then the set $\ltimes^q\big(\ltimes^{t-1}\mathcal{R}\big) = \ltimes^{q(t-1)}\mathcal{R}$ contains a couple $(s^{(1)},m')$ such that $|m'| > q$ which excludes the stability of $\ltimes^k\mathcal{R}$ for any $k \geqslant 1$.

Now assume that all the circuits are linear and show that $\ltimes^k\mathcal{R}$ is stable for some $k$. Every (ordered) path of the graph $\Gamma$ cannot pass twice through a vertex in $\mathcal{L} \setminus \cup_l\mathcal{L}_l$ or reenter second time into some set $\mathcal{L}_l$. Therefore, there is a uniform upper bound for the quantity of the path edges not belonging to $\cup_l\Gamma_l$ (in fact, this quantity does not exceed $t = \operatorname{card}(\mathcal{L}\setminus\cup_l\mathcal{L}_l)+\operatorname{card}\{l\}-1$ with card denoting the cardinality of the set). Notice that any tree does not branch at the origin of its edge provided that this edge lies in some subgraph $\Gamma_l$. Consequently, we can conditionally cut out all such the edges from the tree, thus substituting a chain of edges belonging to the same $\Gamma_l$ by a single vertex. After this surgical operation, the tree's height will admit the above uniform upper estimate. This allows us to estimate $|\widetilde{m}|$ uniformly from above for all $(\tilde{s},\widetilde{m}) \in \ltimes^k\widetilde{\mathcal{R}}$ and arbitrary $k$ if $\widetilde{\mathcal{R}}$ is finite. (In fact, $|\widetilde{m}|$ does not exceed $M = g^t$ with $g$ being the maximal $|\widetilde{m}|$ for $(\tilde{s}, \widetilde{m}) \in \widetilde{\mathcal{R}}$.) Thus, the set $\ltimes^k\widetilde{\mathcal{R}}$ ranges a finite collection of sets, and, consequently, $\ltimes^k\widetilde{\mathcal{R}} = \ltimes^{k+q}\widetilde{\mathcal{R}}$ for some $k = k_0$ and $q > 0$. This implies that $\ltimes^k\widetilde{\mathcal{R}}$ depends $q$-periodically on $k$ (i.e., $\ltimes^k\widetilde{\mathcal{R}} = \ltimes^{k+q}\widetilde{\mathcal{R}}$) for all $k \geqslant k_0$. Therefore, $\ltimes^k\widetilde{\mathcal{R}}$ is stable provided that $k \geqslant k_0$ and $k$ is a multiple of $q$.

One could describe $k_0$ and $q$ in a more constructive manner applicable also to infinite sets $\widetilde{\mathcal{R}}$. Incidentally, this approach provides the smaller estimates for $k_0$ and $q$ which depend only on the finite graph $\Gamma$ (and are independent of the set $\widetilde{\mathcal{R}}$ itself). Let $q$ be a natural such that any vertex in $\cup_l \mathcal{L}_l$ belongs to some closed circuit of length $q$. Under replacement of the TMC $T$ by its $q$-th power, $T^q$ (which corresponds to the replacement of $\widetilde{\mathcal{R}}$ by $\ltimes^{q-1}\widetilde{\mathcal{R}}$), each set $\mathcal{L}_l$ decomposes, generally speaking, into a collection of the analogous subsets, say $\mathcal{L}_l^{(j)}$. (In fact, these subsets are closely related to the general formulation of the Spectral Decomposition Theorem for TMC, see [4].) For some $N$ every two vertices in each $\mathcal{L}_l^{(j)}$ can be joined by a path of length exactly equal to $Nq$ (the restriction of $T^q$ to the set of admissible words formed by the elements of $\mathcal{L}_l^{(j)}$ is a *primitive* TMC, see [4]). Therefore, if the length of a path in $\Gamma_l$ is $Q > q$ and $N' \geqslant N$ then one can replace an arbitrary subpath of length $[Q/q]q$ with $[\cdot]$ denoting the entire part, by a subpath of length $N'q$ with the same origin and end and with the edges lying also in $\Gamma_l$. Thus, if $Q \geqslant Q_0 = (N+1)q$ then one can shorten any path of length $Q$ in $\Gamma_l$ with the fixed endpoints by exactly $q$ edges.

Notice that among the vertices of each path of length $k+1$ constituting the original tree, at most $t'$ vertices do not belong to $\cup_l \mathcal{L}_l$ where $t' \leqslant \operatorname{card}(\mathcal{L} \setminus \cup_l \mathcal{L}_l)$. Therefore, if $k \geqslant t'$ then, certainly, each path passes through a vertex lying in $\cup_l \mathcal{L}_l$. Then one sees that $\ltimes^{k+q}\widetilde{\mathcal{R}} \supseteq \ltimes^k\widetilde{\mathcal{R}}$, due to the possibility to paste a closed circuit of length $q$ in each of these vertices. Furthermore, if $k \geqslant k_1 = t' + \operatorname{card}\{l\} \cdot (Q_0 - 1)$ then at each path, there is a chain of $Q_0$ successive edges belonging to $\cup_l \Gamma_l$. Replacing chains of these edges by chains of appropriate $(Q_0 - q)$ edges belonging again to $\cup_l \Gamma_l$, one sees that $\ltimes^{k-q}\widetilde{\mathcal{R}} \supseteq \ltimes^k\widetilde{\mathcal{R}}$. Thus, finally, $\ltimes^{k+q}\widetilde{\mathcal{R}} = \ltimes^k\widetilde{\mathcal{R}}$ as $k \geqslant k_0 = k_1 + q$.

The only trouble in these considerations is that every two paths in one tree have a common initial piece and all the tree paths should be modified simultaneously in a coherent way so that initial pieces are left identical if modified. Fortunately, this is easily achieved by the simple rule as follows: each path is modified at the nearest to the root place where it is possible. The sufficiency of this rule is seen from the following fact. Let the initial pieces of some paths be coincident up to the vertex $\tilde{s} \in \mathcal{L}_l$ (i.e., these paths belong to the same branch of the tree up to the vertex $\tilde{s}$) included. Then these paths have coincident maximal chains (maybe, empty) of successive edges belonging to $\Gamma_l$ and originated from $\tilde{s}$. In other words, the maximal common initial piece is ended at a vertex which is not the origin of a tree edge belonging to $\cup_l \Gamma_l$. $\square$

The following example generalizes that presented in [20] and shows that $\ltimes^k\mathcal{R}_{0,\xi}$ can "stabilize" for $k$ large enough.

**Example 2.** Let $\Delta_j = [a_j, b_j]$, $1 \leqslant j \leqslant p$, be a collection of $p$ closed intervals in $(0,1)$ such that $a_j \leqslant b_{j-1}$, $1 < j \leqslant p$, and $b_{j-1} < a_{j+1}$, $1 < j < p$, i.e., $\Delta_j \cap \Delta_{j'} \neq \emptyset \iff j = j'$ or $|j - j'| = 1$. Given natural numbers $r_j, 1 \leqslant j \leqslant p$, introduce the numbers $R_{\tilde{s}} = \prod_{j=\tilde{s}}^{p} r_j$ and the sets $[\underline{\lambda}_{\tilde{s}}, \overline{\lambda}_{\tilde{s}}] = \Delta_{\tilde{s}}^{1/R_{\tilde{s}}}$ (one could also set $R_{\tilde{s}} = \prod_{j=1}^{\tilde{s}} r_j$ with minor further modifications). Then $(\tilde{s}, \widetilde{m}) \in \widetilde{\mathcal{R}}_{0,\xi}$ if and only if

$$\Delta_{\tilde{s}} \cap \prod_j \Delta_j^{f_j} \neq \emptyset \tag{4}$$

where $f_j = \widetilde{m}_j/(r_j \cdots r_{\tilde{s}-1})$ for $j < \tilde{s}$, $f_{\tilde{s}} = \widetilde{m}_{\tilde{s}}$, and $f_j = (r_{\tilde{s}+1} \cdots r_j)\widetilde{m}_j$ for $j > \tilde{s}$. Here, the product $\prod_j A_j$ of the sets $A_j$ and the set $A$ raised to the power $t$ are understood as the collections of the corresponding numbers $\prod_j a_j$, $a_j \in A_j$, and $a^t$, $a \in A$. If all $\Delta_j$ are placed in a small enough vicinity of a given point of $(0,1)$ then (4) can be satisfied only when $\sum_j f_j = 1$ and $f_j = 0$ for $j > \tilde{s}$, due to the discreteness of the set of the possible $f_j$. Assume that this is the case and then choose the length, $b_{j-1} - a_j \geqslant 0$, of the intersection $\Delta_{j-1} \cap \Delta_j$ to be small enough for all $1 < j \leqslant p$. In fact, the appropriate upper bound for these lengths will be determined by the lower bounds for distances $a_{j+1} - b_{j-1}$ between the pairs of non-intersecting intervals $\Delta_{j+1}$, $\Delta_{j-1}$. Then (4) holds only if $f_j = 0$ for $j < \tilde{s} - 1$. Thus, the set $\widetilde{\mathcal{R}}_{0,\xi}$ is constituted exactly by all the couples $(\tilde{s}, \widetilde{m})$ such that $\widetilde{m} = \mathbf{e}_{\tilde{s}}$ or $\widetilde{m} = r_{s-1}\mathbf{e}_{\tilde{s}-1}$. Consequently, the set $\ltimes^k \widetilde{\mathcal{R}}_{0,\xi}$ becomes independent on $k$ only for some $k \geqslant p-1$.

*Historical comment 1.* In fact, M. Gyusinsky and A. Katok [20] (see also [25]) proceeded in their consideration from the set $\widetilde{\mathcal{R}} = \{(\tilde{s}, \widetilde{m}) : \underline{\lambda}_{\tilde{s}} \leqslant \overline{\lambda}^{\widetilde{m}}\}$ which is larger than $\widetilde{\mathcal{R}}_{0,\xi}$ because of the presence of additional elements $(\tilde{s}, \widetilde{m})$ such that $\overline{\lambda}_{\tilde{s}} \leqslant \underline{\lambda}^{\widetilde{m}}$. They have proved (in our terms) that $\ltimes^k \widetilde{\mathcal{R}}$ becomes stable for large $k > 0$ provided that the narrow band spectrum condition is met and have constructed the associated $(\{M_{\tilde{s}}\}, \ltimes^k \widetilde{\mathcal{R}})$-normal forms. Incidentally, the desired stability result is easily seen from the fact that $\widetilde{\mathcal{R}}$, analogously to $\widetilde{\mathcal{R}}_{0,\xi}$, takes the triangular form with the linear diagonal part. This proof is more simple and explicit than that presented in [20].

## 3. Nonautonomous normalization

In this Section we recall some results on the normal forms, in particular, the Sternberg theorem [37,5] (see also [23]) on polynomial resonant normal forms of contraction mappings and describe its nonautonomous analog. For this purpose, some quite natural class of particular normalizations will be introduced and discussed.

We preface a discussion by the following

*Remark 4.* It is well-known that “the problem of finding normal forms for smooth transformations splits into two parts: one of a purely formal nature and the other analytic in character” [37]. The formal part consists of the normalization of the diffeomorphism jet at the fixed point by means of a formal transformation, i.e., a transformation represented by a formal Taylor series. The formal normalizing transformation constructed should be the jet of the faithful normalizing transformation. The formal part provides also the normal form. The analytic part consists of proving the existence of a (smooth) normalizing transformation possessing the prescribed jet and carrying over the diffeomorphism into the prescribed normal form. The purpose of the formal normalization is to eliminate the non-resonant terms in the diffeomorphism jet. It is convenient to do this by the successive steps for increasing powers. Here, it is convenient to reduce beforehand the linear part to the Jordan form which requires, generally speaking, the use of complex variables. Since we consider real case, “if complex eigenvalues do occur, they occur in pairs of complex conjugates. In that case, it is convenient to modify” the mapping “by replacing the pair of real coordinates corresponding to the complex eigenvalues by a pair of conjugate complex variables and to allow complex formal power series satisfying certain symmetry (reality) conditions. This procedure is explained in detail in [7, p. 60–63]. In that follows, these symmetry conditions are automatically verified for the series that arise in virtue of the method of their construction” [37] (see also [9]). Thus, to the complex normal form constructed, there corresponds the so-called real normal form. In fact, the normalizing (formal or smooth) transformation is a result of the following two successive transformations: a real (formal or smooth) transformation reducing the original mapping to the so-called “real normal form” and the linear transformation which replaces some pairs of real coordinates by pairs of complex conjugate ones.

We recall the basic result [8] related to the formal normalization (for simplicity, S. Sternberg [37], while discussing smooth contraction mappings, considered the case of the absence of the multiple elementary divisors of the linear part, which has no importance; a general case of the smooth contraction mappings was also treated by B. Anderson [5]). Consider formal power series $g : \mathbb{C}^n \to \mathbb{C}^n$ as

$$\big(g(x)\big)_s = \sum_{\substack{s,m\\ |m|\geqslant 1}} g_{s,m} x^m,$$

where $x = (x_1, \dots, x_n) \in \mathbb{C}^n$ and the linear part $dg|_0$ takes the (complex) Jordan form, $\big(dg|_0(x)\big)_j = \lambda_j x_j + \sigma_j x_{j+1}$ (where $\sigma_j \in \{0,1\}$ and the eigenvalues $\lambda_j$ are properly clustered). Recall that each nonlinear term $g_{s,m}x^m$ is said to be *resonant* (with respect to the linear part $dg|_0$) if $\lambda_s = \lambda^m$, and otherwise, *non-resonant*, where $\lambda = \{\lambda_1, \dots, \lambda_n\}$ denotes the spectrum. The *(complex) normal form* is a mapping $g$ vanishing at the origin whose linear part $dg\,|_0$ takes the (complex) Jordan form and which contains only nonlinear resonant terms. Now let the non-degenerate linear part, $J$, of the formal mapping,

$$\big(T(x)\big)_s = \big(J(x)\big)_s + \sum_{\substack{s,m\\ |m|\geqslant 2}} t_{s,m} x^m, \tag{5}$$

is already reduced to a required Jordan form (by a preliminary linear transformation) with the spectrum $\lambda$. Then some formal transformation $y \mapsto x$ with the identity linear part,

$$x_s = y_s + \sum_{\substack{s,m\\ |m|\geqslant 2}} \varphi_{s,m} y^m, \tag{6}$$

transfers the mapping (5) into the normal form. This transformation is called a *normalization* or *normalizing transformation*. Moreover, the coefficients $\varphi_{s,m}$ in the resonant terms of the normalizing transformation (6) (i.e., ones such that $\lambda_s = \lambda^m$) can be arbitrarily given. Then the normalizing transformation and the normal form are uniquely determined. We will present a simple proof of this result in Remark 15. For the contractive mappings of a sufficiently high but finite smoothness, the theorem by S. Sternberg ensures the presence of not only formal but also actual polynomial (real) normal forms. Then the result that the normalization is determined by its resonant terms is easily carried over to the actual normalizations as will be evident from Remark 6. The below-described construction will provide also the proper reformulations of this result to the situation of some weakened resonance conditions.

Now we recall in detail the Sternberg theorem. Let $T\colon U \to \mathbb{R}^n$ be a map of a vicinity, $U \subset \mathbb{R}^n$, of the fixed point $0 \in \mathbb{R}^n$ and let $J = dT(0)$ be its nondegenerate linear part, $\lambda = \{\lambda_1, \dots, \lambda_n\}$ be the spectrum of $J$ such that $0 < |\lambda_j| < 1$ for all $j$. Let $D$ be the number determined by the spectrum of $J$ in accordance with (2). If $T \in C^N$ where natural $N > D$ then, by some coordinate $C^N$-transformation called a $C^N$*-normalization* or $C^N$*-normalizing transformation*, the mapping $T$ can be reduced to polynomial (real) normal form (5). It is obvious that in this normal form $t_{s,m} = 0$ for $|m| > D$ because the corresponding non-resonance conditions $\lambda_s \neq \lambda^m$ are certainly met.

Now we describe the non-autonomous analog of the Sternberg theorem.

Under the non-resonance conditions (1′), the polynomial normal form, $\mathfrak{J}$, of a contraction mapping $T$ is reduced to the linear operator $J$. The corresponding coordinate transformation is called a *linearization*. The nonautonomous counterpart of the linearization (nonautonomous linearization) for a sequence of contraction mappings has been constructed by Y.Yomdin [38] and by the author [14]. The basic idea of [14] was as follows (Y.Yomdin used very similar considerations). Sternberg's proof [37] can be slightly rephrased in such a way that the linearization of a contraction mapping $T\colon U \to U$, $T(0)=0$, of a neighbourhood $U \subset \mathbb{R}^n$ of the origin (under conditions (1) that are sharper than conditions (1′) considered in [37] where $\{\lambda_j\}$ is the spectrum of the linear part, $J$) will be sought as a fixed point of a hyperbolic endomorphism $\mathcal{D}_T$ in the Banach space of $C^N$-mappings $f$. The endomorphism $\mathcal{D}_T$ is defined so that the following diagram is commutative:

$$\begin{array}{ccc} & \xrightarrow{\;T\;} & \\ {\scriptstyle \mathcal{D}_T(f+\mathrm{id})}\Big\downarrow & & \Big\downarrow{\scriptstyle f+\mathrm{id}} \\ & \xrightarrow[\;J\;]{} & \end{array} \tag{7}$$

where $T(0)=0$, $J=dT(0)$, $f\colon U\to\mathbb{R}^n$, and $f(0)=0$, $df(0)=0$. Then the operator $\mathcal{D}_T$ transforms functions defined in the image of the mapping $T$ into functions defined in the preimage, and $l=\mathrm{id}+f$ is the required linearization if $\mathcal{D}_T l = l$. In order to deal with the linear space of $C^N$-functions $f$ whose 1-jets vanish at the origin, it is convenient to represent the required $f$ as a fixed point of the mapping $(\cdot)\mapsto \overline{\mathcal{D}}_T(\cdot) = \mathcal{D}_T(\cdot)+d_T$, where $d_T = \mathcal{D}_T\,\mathrm{id}-\mathrm{id} = J^{-1}\circ T - \mathrm{id}$. Obviously, $\overline{\mathcal{D}}_{T_1\circ T_2} = \overline{\mathcal{D}}_{T_2}\circ\overline{\mathcal{D}}_{T_1}$, $\mathcal{D}_{T_1\circ T_2} = \mathcal{D}_{T_2}\circ\mathcal{D}_{T_1}$, i.e., the correspondences $T\mapsto\overline{\mathcal{D}}_T$, $\mathcal{D}_T$ are *contravariant functors*. Now, applying the standard tools of hyperbolic theory for mappings in Banach spaces, one can construct the desired nonautonomous normalization.

Now we admit the presence of some resonances. One should consider a modified diagram (7) where the linear map $J=dT(0)$ is replaced by the nonlinear normal form $\mathfrak{J}$. However, the proper generalization of the above consideration requires some accuracy.

Now we will present in detail the proper modification of the considerations of [14] related to the diagram (7). Let $\mathcal{R}$ be a finite set containing resonances of the spectrum, $\{\lambda_j\}$, of the linear part, $J$, of the mapping $T$, which is subordinate to some partition $\xi = \{\Lambda_1,\dots,\Lambda_p\}$ of $\{\lambda_j\}$. The only restriction is that all the multiple eigenvalues that correspond to one Jordan block ("geometrically identical" eigenvalues) belong to the same class of the partition $\xi$. In particular, $\mathcal{R}$ and $\xi$ can be, respectively, the smallest set containing resonances, $\mathcal{R}_0$, and the partition into the classes of numbers with equal moduli, $\xi_0$. We remove also from the set $\mathcal{R}$ all the couples $(s,m)$ such that $|m|\geqslant N$ where natural $N>D$ (recall that these couples are *a priori* non-resonant in the strengthened sense (1)). In the sequel, the multiindex $m$ will run over the finite set $2\leqslant |m| < N$. Denote $\Phi^0=\mathcal{R}$ and let the complementary set[4] $\Phi$ be the finite set of all the couples $(s,m)\notin\mathcal{R}$ such that $2\leqslant|m|<N$. Introduce notations $\widetilde{\Phi}=\Psi(\Phi)$ and $\widetilde{\Phi}^0=\Psi(\Phi^0)$ where the map $\Psi$ defined in Section 1 is associated with the partition $\xi$. Furthermore, denote by $M_{\tilde{s}}$ the invariant subspace corresponding to the Jordan blocks for the eigenvalues belonging to the element $\Lambda_{\tilde{s}}$ of the partition $\xi$ and let $\pi_{\tilde{s}}$ be the projection of $\mathbb{R}^n$ onto $M_{\tilde{s}}$ along $\bigoplus\limits_{l\neq\tilde{s}} M_l$, let $G_{\tilde{s},\widetilde{m}}$ be the linear space of the polynomials of the form $P\big(\pi_1(\cdot),\dots,\pi_p(\cdot)\big)$, where $P\colon M_1\times\dots\times M_p\to M_{\tilde{s}}$ is a homogeneous $M_{\tilde{s}}$-valued polynomial in $p$ variables that has degree $\widetilde{m}_t$ in the $t$-th argument. Let

$$F = \bigoplus_{(\tilde{s},\widetilde{m})\in\widetilde{\Phi}} G_{\tilde{s},\widetilde{m}}, \qquad F^0 = \bigoplus_{(\tilde{s},\widetilde{m})\in\widetilde{\Phi}^0} G_{\tilde{s},\widetilde{m}}\,.$$

Finally, let $E_{\mathcal{R}}$ be $F^0$ and $E_{\overline{\mathcal{R}}}$ be the space of $C^N$-functions that have at the origin $0\in\mathbb{R}^n$ $(N-1)$-jets lying in $F$, i.e., $E_{\overline{\mathcal{R}}} = F\oplus G_*$, where $G_*$ is the space of $C^N$-functions $f\colon U\to\mathbb{R}^n$

[4] Here and below the notations of [14] are slightly modified.

with zero $(N-1)$-jets at the origin. Then the space $E$ of $C^N$-functions $f\colon U\to\mathbb{R}^n$ whose 1-jets vanish at the origin decomposes into the direct sum $E=E_{\mathcal{R}}\oplus E_{\overline{\mathcal{R}}}$. So, every function $f\in E$ is decomposed into $f=f_{\mathcal{R}}\oplus f_{\overline{\mathcal{R}}}$ where $f_{\mathcal{R}}\in E_{\mathcal{R}}$ and $f_{\overline{\mathcal{R}}}\in E_{\overline{\mathcal{R}}}$. Here, summands $f_{\mathcal{R}}$ and $f_{\overline{\mathcal{R}}}$ represent the nonlinear terms which will be attribute to "resonant" and "non-resonant" ones. It would be more correct to denote these terms as *"(near-)neutral"* and *"hyperbolic"*, respectively, in view of the character of their behaviour under the action of the linear operator $\mathcal{D}_J$ (see below) which is an actual reason for introducing the subspaces $E_{\mathcal{R}}$ and $E_{\overline{\mathcal{R}}}$. For any function $f:U\to\mathbb{R}^n$ vanishing at the origin, $f_{\mathcal{R}}$ and $f_{\overline{\mathcal{R}}}$ will denote the corresponding summands of its nonlinear part, $f-df(0)$. The functional spaces $E_\nu=E_\nu(U)$, $\nu\in\{\mathcal{R},\overline{\mathcal{R}}\}$, do not depend on the choice of $U$ in the sense that for $U'\subset U''$, restricting the functions $f\in E_\nu(U'')$ onto $U'$ generates the natural projections of $E_\nu(U'')$ onto $E_\nu(U')$.

In fact, the decomposition $E=E_{\mathcal{R}}\oplus E_{\overline{\mathcal{R}}}$ is determined only by the generalized eigensubspaces $M_{\tilde{s}}$, whose direct sum is the whole space $\mathbb{R}^n$, and by the arbitrary finite subset $\widetilde{\mathcal{R}}=\Psi(\mathcal{R})=\widetilde{\Phi}^0$ of the couples $(\tilde{s},\widetilde{m})$ such that $2\leqslant|\widetilde{m}|<N$.

We notice also that one could consider, on the formal level, the general case of an arbitrary spectrum (not necessary lying from one side of the unit circle). Then the sets $\widetilde{\Phi}^0$ and $\widetilde{\Phi}$ can be infinite and $f$, $f_{\mathcal{R}}$, and $f_{\overline{\mathcal{R}}}$ are treated as the formal Maclaurin series.

**Definition 8.** We will say that a polynomial or formal power series mapping $g:\mathbb{R}^n\to\mathbb{R}^n$ is a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$*-quasi-normal form* if $g$ vanishes at the origin, possesses linear part $dg(0)$ for which all the subspaces $M_{\tilde{s}}$ are invariant, and is such that $g_{\overline{\mathcal{R}}}=0$. A $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-quasi-normal form $g$ will be called a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$*-normal form* if $\mathcal{R}=\Psi^{-1}(\widetilde{\mathcal{R}})$ is a set containing exact resonances for the spectrum, $\{\lambda_j\}$, of the linear part $dg(0)$, i.e., such that $\lambda_s=\lambda^m\Rightarrow(s,m)\in\mathcal{R}$. Any coordinate transformation reducing the mapping to a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal form will be called a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$*-normalization* or $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$*-normalizing transformation*. If $M_{\tilde{s}}$ are exactly the generalized eigensubspaces of some linear operator $l$ associated to its distinct eigenvalues (i.e., one considers the partition $\xi=\xi_0^{(ex)}$ of the spectrum of $l$ into the classes of equal numbers) and $\mathcal{R}=\mathcal{R}_0^{(ex)}$ is the set consisting of all the exact resonances for $l$ then we will denote a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-quasi-normal form as a *quasi-normal form with respect to $l$*. Analogously, a mapping which is a quasi-normal form with respect to its linear part will be called a *normal form*, and a coordinate transformation reducing the mapping to a normal form will be called a *normalization* or a *normalizing transformation*.

*Remark 5.* We notice that all the constructions are invariant, i.e., coordinate-free and that our Definition is consistent with the classic notation of normalization and normal form. The $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normalization is an example of a *particular normalization*, removing only some of the non-resonant terms.

In the above Definition, it is convenient to consider also linear couples $(s,m)$ and to replace the set $\mathcal{R}$ by the corresponding "extended" set via adding all the $\xi$-trivial linear couples. Then the condition $g_{\overline{\mathcal{R}}}=0$ will automatically imply the $dg(0)$-invariance of the subspaces $M_{\tilde{s}}$. (Now the decomposition $f=f_{\mathcal{R}}\oplus f_{\overline{\mathcal{R}}}$ is determined for all the functions $f$ vanishing at the origin.)

Recall that one should consider a modified diagram (7) where the linear map $J=dT(0)$ is replaced by the normal form. However, for further purposes, we have to allow a more general $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal form $\mathfrak{J}$ containing "$\mathcal{R}$-resonant" terms only, i.e., such that $\mathfrak{J}_{\overline{\mathcal{R}}}=0$. For convenience, under proving the required properties of the maps $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$, $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ described below, the lower arrow in the diagram will be inverted and the corresponding mapping $\mathfrak{J}$ will be replaced by its inverse, $\mathfrak{I}=\mathfrak{J}^{-1}$ satisfying the assumption that $\mathfrak{I}_{\overline{\mathcal{R}}}=0$. We mention in this connection the fact that if $\mathfrak{J}$ is the normal form in the usual sense then $\mathfrak{J}^{-1}$ is also the normal form with the same set of possible resonant term (i.e., with the same set $\mathcal{R}$), see

Remark 6. However, due to Remark 20 below, in the general case, both the conditions $\mathfrak{J}_{\overline{\mathcal{R}}} = 0$ and $\mathfrak{I}_{\overline{\mathcal{R}}} = 0$ are equivalent only if the set $\mathcal{R}$ is stable. Thus, we will pass to a completely analogous but not identical problem where the mappings $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ are replaced by their "counterparts", $\mathcal{E}^{\leftarrow}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{I}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$. The first point of the proof will be to establish the required properties of these "counterparts" and the second point, to transfer the results obtained to the original mappings $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$, $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$.

The required modification of the considerations of [14] is as follows. For given "resonant parts" $f_{\mathcal{R}}$ and $f'_{\mathcal{R}}$, there are nonlinear maps $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\overline{\mathcal{R}}}$ and $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\mathcal{R}}$ such that the diagram

$$\begin{array}{ccc} & \xrightarrow{\;T\;} & \\ l'\Big\downarrow & & \Big\downarrow l \\ & \xrightarrow[\;\mathfrak{J}\;]{} & \end{array} \tag{8}$$

with $l = f + \mathrm{id}$, $l' = f' + \mathrm{id}$ commutes when $f'_{\overline{\mathcal{R}}} = \mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$, $\mathfrak{J}_{\mathcal{R}} = \mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$, and $\mathfrak{J}_{\overline{\mathcal{R}}} \equiv 0$ (notice that $d\mathfrak{J}(0) = J$). Here, in contrast to [14], the nonlinearity of the maps introduced is caused by that of the $(\{M_{\bar{s}}\}, \widetilde{\mathcal{R}})$-normal form. However, the map $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ possesses good hyperbolicity properties over a given bounded set in $E_{\overline{\mathcal{R}}}$. The correctness of these definitions and the hyperbolicity property of the map $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ will be established in a more general framework while proving Proposition 3. Rigorously speaking, in Proposition 3 where a "uniform" variant of the above construction is considered which is required to build a "nonautonomous normalization", conditions imposed on $\xi$ will be somewhat stronger. However, the proof will be also valid under the above-mentioned assumptions concerning $\xi$.

In an auxiliary case of the inverted lower arrow of the latter diagram, we will show the existence of nonlinear maps $\mathcal{E}^{\leftarrow}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\overline{\mathcal{R}}}$ and $\mathcal{I}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\mathcal{R}}$ such that the diagram

$$\begin{array}{ccc} & \xrightarrow{\;T\;} & \\ l'\Big\downarrow & & \Big\downarrow l \\ & \xleftarrow[\;\mathfrak{I}\;]{} & \end{array}$$

with $l = f + \mathrm{id}$, $l' = f' + \mathrm{id}$ commutes when $f'_{\overline{\mathcal{R}}} = \mathcal{E}^{\leftarrow}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$, $\mathfrak{I}_{\mathcal{R}} = \mathcal{I}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$, and $\mathfrak{I}_{\overline{\mathcal{R}}} \equiv 0$ (notice that $d\mathfrak{I}(0) = I = J^{-1}$).

*Remark 6.* (See Remark 14 of [14].) Now we explain the proof of the Sternberg theorem and establish its locally uniform variant as follows. The normalizing $C^N$-transformation depends continuously on a $C^N$-diffeomorphism $T \colon U \to \mathbb{R}^n$ if all the mappings are treated in the $C^k$-topology for arbitrary $1 \leqslant k \leqslant N$ and $T$ remains uniformly $C^N$-bounded from above. This result remains valid for the case $k = 0$ provided that an additional assumption is met that the perturbation of the diffeomorphism $T$ is sufficiently small in the $C^1$-topology. Here $N \geqslant 2$ is determined by the inequality $N > D$. Actually, we adopt the original proof [37]. Moreover, we will extend these results to the most general case where the mapping $T$ is conjugated by a $C^N$-transformation $l$ possessing a prescribed $(N-1)$-jet $j_0^{N-1} l$ with a mapping $\mathfrak{J}$ possessing a prescribed $G_*$-component, $\mathfrak{J} - j_0^{N-1}\mathfrak{J}$. Here $l$ and $\mathfrak{J}$ will depend continuously on $T$, $j_0^{N-1} l$, and the $G_*$-component of $\mathfrak{J}$. A fixed point of the contractive map $T$ depends continuously on $T \in C^1$. Hence, without loss of generality, one can consider perturbations of $T$ that leave the fixed point to be located at the origin. There is a norm in $\mathbb{R}^n$ that induces an operator norm $\|\cdot\|$ such that $\|J\| < 1$ and $\mu = \left\|J^{-1}\right\| \|J\|^N < 1$ where $J = dT(0) \colon \mathbb{R}^n \to \mathbb{R}^n$. Let $U$ be a small enough ball in this norm centered at the origin and let $G_*$ be the space of $C^N$-functions $f\colon U \to \mathbb{R}^n$ with zero $(N-1)$-jets at the origin. Equip $G_*$ with the norm $\|\cdot\|_*$ being the exact upper bound for the norm of the $N$-th differential at all the points of $U$. Here the $N$-th differential is considered as a polylinear mapping $\underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{N} \to \mathbb{R}^n$ symmetrical in each cofactor. (The desired norm in the space of polylinear mapping is defined, for instance, in the textbook [12].) Then the affine map $\overline{\mathcal{D}}_T$ is well-defined and the key fact used in [14] is that $\|\mathcal{D}_T \mid G_*\|_* \leqslant \tilde{\mu}$, $\mathcal{D}_T$ being the linear part of $\overline{\mathcal{D}}_T$, the operator norm being induced by the norm $\|\cdot\|_*$ in $G_*$, and $\tilde{\mu}$ being close to $\mu$ if the neighbourhood $U$ is chosen to be sufficiently small. Indeed, elementary estimates show that $|\tilde{\mu}-\mu| \leqslant \varepsilon \cdot \|J^{-1}\| P_N\big(\|T\|_{C^N}, \varepsilon\big)$ where $P_N$ is some polynomial depending on $N$ only and $\varepsilon$ is the radius of the ball $U$.

The formal normalization of the $(N-1)$-jet of $T$ at the fixed point supplies the normal form $\mathfrak{J}$ sought for. The corresponding formal normalizing transformation is the $(N-1)$-jet for $C^N$-functions $R$ with identity linear part which satisfy the following condition: a $C^N$-transformation from the old coordinates $x \in U \subset \mathbb{R}^n$ to the new ones $y = R(x)$ transfers the mapping $T$ into the mapping $T' = R \circ T \circ R^{-1}$ such that $T' - \mathfrak{J} \in G_*$. Then in the new coordinates $y$, the desired normalizing transformation is sought for as $l' = \mathrm{id} + f$ where $f$ is a fixed point of the map $(\cdot) \mapsto \widetilde{\mathcal{E}}_{T',\mathfrak{J}}(\cdot) = \mathfrak{J}^{-1} \circ \big(\mathrm{id} + (\cdot)\big) \circ T' - \mathrm{id}$ in the space $G_*$. This map is well-defined and will be shown to be contractive over a bounded subset of $G_*$. In the original coordinates, the normalizing transformation is written as $l = l' \circ R$. (In [14], the author presented also another version of the above proof where the preliminary transformation of coordinates is not performed. However, this consideration is valid only if the normal form is linear.) One could notice that $\mathfrak{J} - J = \mathcal{J}_{T',0,0}(f)$ for every $f \in G_*$ and $\widetilde{\mathcal{E}}_{T',\mathfrak{J}} = \mathcal{E}_{T',0,0} \mid G_*$ under any choice of the set $\mathcal{R}$ containing resonances, since $T' - \mathfrak{J} \in G_*$. In the sequel, for brevity we will omit the subscript "prime" of $T$.

If the eigenvalues $\lambda_j$ are placed in the order of non-increasing moduli then the normal form takes the lower-triangular form, i.e., the nonlinear part of $\big(\mathfrak{J}(y)\big)_s$ is expressed via coordinates $y_{s'}$ only such that $s' < s$ [8]. It follows easily from this fact that the mapping $\mathfrak{I} = \mathfrak{J}^{-1}$ is polynomial and contains no resonant terms. Then $\widetilde{\mathcal{E}}_{T,\mathfrak{J}}(f) = \mathfrak{J}^{-1} \circ T - \mathrm{id} + \mathcal{D}_T f + h_{T,\mathfrak{J}}(f)$, where one puts $h_{T,\mathfrak{J}}(f) = \Theta_{T,\mathfrak{J}}(f \circ T)$ and $\Theta_{T,\mathfrak{J}}(g)$ is a finite sum of terms of the type

$$T_{\widetilde{w}_\alpha}(\cdot)\big(\widetilde{\widetilde{w}}_\beta \circ g(\cdot)\big)\mathbf{e}, \tag{9}$$

with the notation $T_w = w \circ T$ introduced, $\mathbf{e}$'s being some vectors in $\mathbb{R}^n$, and $\widetilde{w}_s$ $(s \geqslant 0)$, $\widetilde{\widetilde{w}}_s$ $(s \geqslant 1)$, homogeneous polynomials $\mathbb{R}^n \to \mathbb{R}$ of degree $s$. The polynomials $\widetilde{w}_\alpha$, $\widetilde{\widetilde{w}}_\beta$ included in each term are linear dependent on the homogeneous form $w_t$ of degree $t = \alpha + \beta \geqslant 2$ in the polynomial mapping $\mathfrak{I} = \mathfrak{J}^{-1}$. To each value of the index $t$, there corresponds, generally speaking, several (one-dimensional) terms of type (9).

The norm of the free term, $\|\mathfrak{J}^{-1} \circ T - \mathrm{id}\|_*$, admits a polynomial upper estimate via $J^{-1}$, $\|T\|_{C^N}$, and the radius of the ball $U$ denoted below as $\varepsilon$. The nonlinear term $h_{T,\mathfrak{J}}(f)$ that is caused by the non-linearity of the normal form vanishes at $f = 0$ and possesses a small enough Lipschitz constant over a given bounded subset of $G_*$[5]. Rigorously speaking, the following result holds. Let $R_* > 0$ be given and $B_*$ be the (closed) ball in $G_*$ of radius $R_*$ centered at the zero. Then $h_{T,\mathfrak{J}} \mid B_*$ possesses a Lipschitz constant not exceeding $\varepsilon Q_N\big(\|T\|_{C^N}, \|\mathfrak{I}\|_{C^N}, \varepsilon\big)$ where $Q_N$ is some polynomial depending on $N$ and $R_*$ only. Indeed, it suffices to notice that the map $h_{T,\mathfrak{J}}$ is differentiable due to below Lemma 1 and to obtain the required upper estimate for its

[5]In fact, S. Sternberg [37] stated that the Lipschitz constant is small enough over the whole space $G_*$ which is not true.

differential. The differential of the functional map $G_* \to G_*$ that determines by the formula $f \mapsto T_{\widetilde{w}_\alpha}(\cdot)\big(\widetilde{\widetilde{w}}_\beta \circ f \circ T(\cdot)\big)$ the coefficient in front of $\mathbf{e}$ in (9) takes the form

$$\delta f(\cdot) \mapsto g(\cdot) \cdot \delta f \circ T(\cdot) \tag{10}$$

where $g(\cdot) = T_{\widetilde{w}_\alpha}(\cdot) \cdot d\widetilde{\widetilde{w}}_\beta \mid_{f \circ T(\cdot)}$. The norm $\|g\|_{C^N}$ is estimated from above by a polynomial in $\|T\|_{C^N}$, $\|\mathfrak{I}\|_{C^N}$, and $\varepsilon$, and the function $g$ certainly vanishes at the origin due to that $t = \alpha + \beta \geqslant 2$. The operator (10) is the composition of two operators in the space $G_*$ which act according to the formulas

$$\delta f(\cdot) \mapsto \delta f \circ T(\cdot) \quad \text{and} \quad v(\cdot) \mapsto g(\cdot) \cdot v(\cdot). \tag{11}$$

It is easily seen that the norm of the first operator (11) does not exceed $S'_N(\|T\|_{C^N}, \varepsilon)$ and the norm of the second one, $\varepsilon S''_N(\varepsilon) \cdot \|g\|_{C^N}$ where $S'_N$ and $S''_N$ are analogous to $P_N$ and $Q_N$. The required estimate for the norm of the differential follows from these ones. The multiplier $\varepsilon$ in front of the second estimate is due to the fact that $g$ vanishes at the origin.

Thus, if $R_* > 0$ is chosen large enough ($R_* > R_*\big(\|T\|_{C^N}, \|\mathfrak{I}\|_{C^N}\big)$) and then the radius of the ball $U$ is set to be small enough ($\varepsilon < \varepsilon\big(\|T\|_{C^N}, \|\mathfrak{I}\|_{C^N}, R_*\big)$) then $\widetilde{\mathcal{E}}_{T,\mathfrak{J}} \mid B_*$ is a contractive map $B_* \to B_*$. Consequently, there is a unique fixed point $f \in B_*$.

Thus, *a priori* upper bound $R_*$ for the norm of the normalizing transformation determines the restriction from above to the radius $\varepsilon$ of the ball $U$ where this normalization is built. However, this does not lead to losing some possible solutions of the normalization problem and does not bound a domain where the normalization is defined, due to the following facts. Firstly, it is obvious that under increasing $R_*$ and the corresponding decreasing $\varepsilon$, one will obtain new normalizing transformations that are exactly the restrictions of the previous ones to narrower domains. Secondly, the iterations of the map $T$ and its normal form $\mathfrak{J}$ will just determine the continuation of the normalization onto the whole basin of attraction of the fixed point of $T$.

Arguments of [14] show that it suffices to prove the following result where all the mappings are considered in the $C^k$-topology. The unique fixed point $f \in B_*$ of the mapping $\widetilde{\mathcal{E}}_{T,\mathfrak{J}} \mid B_*$ depends continuously on $T$ provided that $T$ remains uniformly $C^N$-bounded. Here, in contrast to [14], the mapping $\widetilde{\mathcal{E}}_{T,\mathfrak{J}}$ is not affine. So, the considerations should be slightly modified as follows. The sequence $f_{(n)} = \widetilde{\mathcal{E}}^n_{T,\mathfrak{J}}(f)$, where $f \in B_*$, is uniformly convergent in $G_*$. This fact is due to the uniformity of the Lipschitz constant $< 1$ for $\widetilde{\mathcal{E}}_{T,\mathfrak{J}} \mid B_*$ and the uniform $C^N$-boundedness of $B_*$. Consequently, the sequence is also uniformly convergent in the $C^k$-norm and the desired result follows immediately from here because of the continuous dependence of each term on $T$ in the $C^k$-topology.

Now we consider the most general case mentioned in the beginning of this Remark. Let $w_N$ be the $G_*$-component of $\mathfrak{I} = \mathfrak{J}^{-1}$. The only distinction is that now a new summand $w_N \circ (\mathrm{id} + f) \circ T$ will appear in the expression for $\widetilde{\mathcal{E}}_{T,\mathfrak{J}}(T)$, in addition to the previous terms (9). By the Implicit Function Theory (in Banach spaces), one easily sees from $\mathfrak{I} \circ \mathfrak{J} = \mathrm{id}$ that the dependence of $\mathfrak{I}$ on $\mathfrak{J}$ is continuously differentiable. We have only to prove that this new term satisfies the above-described Lipschitz estimate. However, this is easily derived from the above arguments applied to the differential of this term, since $N \geqslant 2$.

The following important result which is an appropriately generalized reformulation of Proposition 5 of [14] will provide the nonautonomous normalization. Before stating it, we point out that if $T = J$ is linear and $f_{\mathcal{R}} \equiv 0$, $f'_{\mathcal{R}} \equiv 0$ then $\mathcal{J}_{J,0,0} = 0$, so that the "normal form" $\mathfrak{J}$ and the map $L = \mathcal{E}_{J,0,0} = \mathcal{D}_J \mid E_{\overline{\mathcal{R}}}$ are linear, $L(f) = J^{-1} \circ f \circ J$ for $f \in E_{\overline{\mathcal{R}}}$. One can define a map $L^0 : E_{\mathcal{R}} \to E_{\mathcal{R}}$ using also the latter formula. Introduce also a vector $d_T = J^{-1} \circ T - \mathrm{id}$ analogous to that used earlier.

**Proposition 3.** *Let the following "concordance" and "uniformity" conditions be satisfied for $C^N$-mappings $T_i\colon U \to \mathbb{R}^n$ of a domain $U \subset \mathbb{R}^n$ that have the common fixed point $0 \in U$:*

i) *the sets of eigenvalues for the operators $J_i = dT_i(0)\colon \mathbb{R}^n \to \mathbb{R}^n$ possess the $\widetilde{\mathcal{R}}$-right concordant partitions $\xi_i = \{\Lambda_{i,\tilde{s}} : 1 \leqslant \tilde{s} \leqslant p\}$, and, moreover, invariant subspaces $M_{\tilde{s}}$ corresponding to the elements $\Lambda_{i,\tilde{s}}$ of the partitions $\xi_i$ are coincided. The number $N$ satisfies the claim $N > D_i$ where $D_i$ are defined by the equalities of type* (2) *via the eigenvalues of the operators $J_i$;*

ii) *moreover, some norm in $\mathbb{R}^n$ and its restrictions to $M_{\tilde{s}}$ induce operator norms $\|\cdot\|$ such that $\|J_i\| < 1$ and the partition into classes $\left\{\|J_{i,\tilde{s}}\|, \|J_{i,\tilde{s}}^{-1}\|^{-1}\right\}$, $J_{i,\tilde{s}} = J_i \mid M_{\tilde{s}}$, that correspond to $\widetilde{\mathcal{R}}$-right concordant partitions $\xi_i$, are also $\widetilde{\mathcal{R}}$-right*[6]*. These conditions must be persistent under $\delta$-perturbations of the operators $J_{i,\tilde{s}}$, where $\delta > 0$ is some number;*

iii) *the $C^N$-norms of $T_i$ are uniformly bounded.*

*Then in a small enough ball (in the sense of the norm from condition* ii)*) $U$, centered at the origin $0 \in \mathbb{R}^n$, all the mappings $T_i$ are contractive. Besides, there exist a number $0 < \mu < 1$, a direct sum decomposition $E_{\overline{\mathcal{R}}} = E^+ \oplus E^-$ of the space $E_{\overline{\mathcal{R}}}$, and a norm in $E_{\overline{\mathcal{R}}}$ (that is equivalent to the original one) such that the following statements hold. Each linear operator $L_i = \mathcal{E}_{J_i,0,0} = \mathcal{D}_{J_i} \mid E_{\overline{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\overline{\mathcal{R}}}$ is hyperbolic with separatrices $E^+$ and $E^-$ and, moreover, $\left\|\left(L_i \mid E^{\pm}\right)^{\pm 1}\right\| \leqslant \mu$ for the corresponding operator norms. The functional spaces $E^{\pm} = E^{\pm}(U)$ do not depend on the choice of $U$ in the sense that for $U' \subset U''$, restricting the functions $f \in E^{\pm}(U'')$ onto $U'$ generates the natural projections of $E^{\pm}(U'')$ onto $E^{\pm}(U')$. The maps $\mathcal{E}_{T_i,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\overline{\mathcal{R}}}$ and $\mathcal{J}_{T_i,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\mathcal{R}}$ are well-defined and differentiable (in the sense of Frechét). Obviously, to the decomposition of $E_{\overline{\mathcal{R}}}$, there correspond the block forms of the linear operators $d\mathcal{E}_{T_i,f_{\mathcal{R}},f'_{\mathcal{R}}}$:*

$$d\mathcal{E}_{T_i,f_{\mathcal{R}},f'_{\mathcal{R}}} = \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix},$$

*where $A_i\colon E^+ \to E^+$, $B_i\colon E^- \to E^+$, $C_i\colon E^+ \to E^-$, and $D_i\colon E^- \to E^-$ are linear operators. Next, given bounded sets $B_{\mathcal{R}} \subset E_{\mathcal{R}}$, $B_{\overline{\mathcal{R}}} \subset E_{\overline{\mathcal{R}}}$ and a number $\varepsilon > 0$, the small enough neighbourhood $U$ of the origin and the norm $\|\cdot\|$ in $E_{\overline{\mathcal{R}}}$ can be chosen in such a manner that in addition the following conditions will hold: the set $B_{\overline{\mathcal{R}}}$ lies in the unit ball, $\|f_{\overline{\mathcal{R}}}\| < 1$, in $E_{\overline{\mathcal{R}}}$ and for all $i$ the norms of the block components of the differential $d\mathcal{E}_{T_i,f_{\mathcal{R}},f'_{\mathcal{R}}}$ at points $f_{\overline{\mathcal{R}}}$ of the unit ball satisfy estimates $\|A_i\| \leqslant \mu + \varepsilon$, $\|B_i\| \leqslant \varepsilon$, $\|C_i\| \leqslant \varepsilon$, and $\left\|D_i^{-1}\right\| \leqslant \mu + \varepsilon$, and the norms of the vectors $\mathcal{E}_{T_i,f_{\mathcal{R}},f'_{\mathcal{R}}}(0)$ do not exceed $\varepsilon$, provided that $f_{\mathcal{R}}, f'_{\mathcal{R}} \in B_{\mathcal{R}}$.*

**Corollary.** *If the subscript $i$ ranges over the set $\mathbb{N}$ and the polynomial mapping $f_i \in E_{\mathcal{R}}$ have uniformly bounded coefficients then, under the conditions of Proposition* 3*, there exists a unique sequence of uniformly $C^N$-bounded mappings $l_i$ possessing at the origin a tangency of (at least) first order with the identity map* id*, having the prescribed "resonant parts" $(l_i)_{\mathcal{R}} = f_i$ and such that the following diagram commutes:*

$$\begin{array}{ccccccccc} \cdots & \xrightarrow{T_{-1}} & & \xrightarrow{T_0} & & \xrightarrow{T_1} & & \xrightarrow{T_2} & \cdots \\ & & \Big\downarrow l_{-1} & & \Big\downarrow l_0 & & \Big\downarrow l_1 & & \\ \cdots & \xrightarrow[\mathfrak{J}_{-1}]{} & & \xrightarrow[\mathfrak{J}_0]{} & & \xrightarrow[\mathfrak{J}_1]{} & & \xrightarrow[\mathfrak{J}_2]{} & \cdots \end{array} \tag{12}$$

*where all $\mathfrak{J}_i$ are "resonant" in the sense that $(\mathfrak{J}_i)_{\overline{\mathcal{R}}} = 0$.*

[6]We correct here an inaccuracy passed in Proposition 5 of [14].

*Remark 7.* The construction of the mappings $\mathcal{E}_{T_i,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\overline{\mathcal{R}}}$, $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\mathcal{R}}$ and Proposition 3 can be generalized to the case where the differentials $dl(0)$, $dl'(0)$ are not necessary the identity mappings but are non-degenerate and possess $M_{\tilde{s}}$ as the invariant subspaces. The simplest way to show this fact is based on the following observation. The commutativity of diagram (8) is equivalent to the commutativity of the same diagram where the mappings $l$, $l'$, and $\mathfrak{J}$ are replaced, respectively, by the mappings $\big(dl(0)\big)^{-1} \circ l$, $\big(dl'(0)\big)^{-1} \circ l'$, and $\big(dl(0)\big)^{-1} \circ \mathfrak{J} \circ \big(dl'(0)\big)$. So, this general case is reduced to the particular one where $dl(0) = \mathrm{id}$ and $dl'(0) = \mathrm{id}$. To ensure the uniform choice of $U$ and $\|\cdot\|$ independently of the differentials $dl(0)$, $dl'(0)$, one should demand their uniform non-degeneracy in the sense that both $\|L\|$ and $\|L^{-1}\|^{-1}$ are uniformly bounded for $L = dl(0)$ and $dl'(0)$. Due to this generalization, the Corollary of Proposition 3 remains valid if all the prescribed differentials $dl_i(0)$ are uniformly non-degenerate and possess $M_{\tilde{s}}$ as the invariant subspaces.

*Remark 8.* Assume that there is a commutative diagram

$$\begin{tikzcd} \cdots \arrow[r, "J_{-1}"] & {} \arrow[r, "J_0"] \arrow[d, "t_{-1}"'] & {} \arrow[r, "J_1"] \arrow[d, "t_0"'] & {} \arrow[r, "J_2"] \arrow[d, "t_1"'] & \cdots \\ \cdots \arrow[r, "J'_{-1}"'] & {} \arrow[r, "J'_0"'] & {} \arrow[r, "J'_1"'] & {} \arrow[r, "J'_2"'] & \cdots \end{tikzcd} \tag{13}$$

of linear mappings such that $\{t_i\}$ are uniformly non-degenerate in the sense of the previous Remark, and $\{J_i\}$ and $\{J'_i\}$ satisfy, respectively, condition ii) of Proposition 3 with the same collection of invariant subspaces $\{M_{\tilde{s}}\}$. Then the subspaces $M_{\tilde{s}}$ are $t_i$-invariant for all $i$. Thus, in Remark 7, one can omit the assumption that the subspaces $\{M_{\tilde{s}}\}$ are invariant with respect to the linear parts of the mappings $l_i$ conjugating the original mappings $T_i$ with the normal forms $\mathfrak{J}_i$. Indeed, this assumption is automatically guaranteed by the uniform non-degeneracy of these linear parts and by the commutativity of the diagram (13) where $J_i$ and $J'_i$ are the linear parts of $T_i$ and $\mathfrak{J}_i$, respectively. For the proof of this result one uses only the fact that partition into the classes $\Big\{\|J_{i,\tilde{s}}\|, \|J_{i,\tilde{s}}^{-1}\|^{-1}\Big\}$ and the analogous partition into the classes $\Big\{\|J'_{i,\tilde{s}}\|, \|(J'_{i,\tilde{s}})^{-1}\|^{-1}\Big\}$ are strongly ordered. Indeed, this fact means that the subspaces $\{M_{\tilde{s}}\}$ are *exponentially separated* in the sense of Section 5. More precisely speaking, introduce the linear extension $\mathbf{J} = [J_i] : X \to X$, $X = \mathbb{Z} \times \mathbb{R}^n$, of the standard shift $\sigma : \mathbb{Z} \to \mathbb{Z}$, $\sigma(i) = i + 1$, by the mappings $J_i$ which acts as $(i, x) \mapsto \big(\sigma(i), J_i(x)\big)$. Then the trivial bundles $X_{\tilde{s}} = \mathbb{Z} \times M_{\tilde{s}} \to \mathbb{Z}$ are $\mathbf{J}$-invariant and exponentially separated. Moreover, they constitute a uniformly non-degenerate decomposition $X = \oplus_{\tilde{s}} X_{\tilde{s}}$ is the sense that the angles between vectors belonging to distinct fibers $M_{i,\tilde{s}} = M_{\tilde{s}}$ over the same point $i$ of the base $\mathbb{Z}$ are uniformly bounded away from zero for all $i \in \mathbb{Z}$. Note that, in contrast to Section 5, we deal with non-compact base $\mathbb{Z}$ but implicitly assume the required uniformity conditions. Obviously, since diagram (13) is commutative and the linear mappings $t_i$ are uniformly non-degenerate, they transform $\mathbf{J}$-invariant subbundles of $X$, which are exponentially separated and constitute a uniformly non-degenerate decomposition of $X$, into $\mathbf{J}'$-invariant subbundles of $X$ which possess the same properties, where $\mathbf{J}' = [J'_i] : X \to X$ is analogous to $\mathbf{J}$. But the invariant exponentially separated subbundles, which constitute a uniformly non-degenerate decomposition of the whole bundle, are uniquely determined by their dimensions, see Proposition 4 in Section 5. Thus, the subbundles $X_{\tilde{s}}$ are invariant under the linear automorphism $\mathbf{t} = [t_i]$, i.e., $t_i(M_{\tilde{s}}) = M_{\tilde{s}}$.

*Remark 9.* In the previous Remark, one can omit the condition that the linear maps $t_i$ are invertible by some sharpening the conditions imposed on $J$ and $J'$. Assume that for $v \in M_{\tilde{s}}$ one has

$$C^{-1}(g_{\bar{s}}^{-})^m \leqslant \frac{\left\| J_{i+m} \circ \cdots \circ J_{i+1}(v) \right\|}{\|v\|} \leqslant C(g_{\bar{s}}^{+})^m, \quad m \geqslant 0,$$

where $g_{\bar{s}}^{+} < g_{\bar{s}+1}^{-}$ and $C$ are positive, and also the analogous estimates with the same $g_{\bar{s}}^{\pm}$ where $J_i$ are replaced by $J_i'$. This means that the invariant subspaces $M_{\bar{s}}$ (more precisely, the invariant subbundles $X_{\bar{s}}$, $X_{\bar{s}}'$) are *spectrally separated* in the sense of Section 5. Moreover, if $t_i$ are uniformly bounded linear maps such that the diagram (13) commutes then the subspaces $M_{\bar{s}}$ are $t_i$-invariant for all $i$. This fact will be explained while proving Theorem 3′ in Section 5. Now we replace the condition ii) of Proposition 3 by the assumption that the spectral intervals $[g_{\bar{s}}^{-}, g_{\bar{s}}^{+}]$ are elements of a $\widetilde{\mathcal{R}}$-right partition. Then the result of the previous Remark is transferred to the present case. Thus, for any uniformly bounded maps $f_i \in E_{\mathcal{R}}$ with the prescribed differentials $t_i = df_i(0)$ such that diagram (13) commutes, there is a unique sequence of uniformly $C^N$-bounded mappings $l_i$ with the prescribed "resonant parts" $f_i = (l_i)_{\mathcal{R}}$ such that diagram (12) commutes. To prove this result, one can consider the commutative diagram (8) with $l = f+t$, $l' = f'+t'$, and the mappings $\mathcal{E}_{T,\mathfrak{J},t,t',f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{J}_{T,\mathfrak{J},t,t',f_{\mathcal{R}},f'_{\mathcal{R}}}$ completely analogous to those constructed in Proposition 3. The difference is that now the linear parts $J$, $J'$, $t$, and $t'$ of $T$, $\mathfrak{J}$, $l$, and $l'$ are such that $t$ and $t'$ may be even degenerate but all these linear maps should constitute a commutative diagram, i.e., $J' \circ t' = t \circ J$. We observe that for the mappings $L = J_{i+m} \circ \cdots \circ J_i$ and $L = J_{i+m}' \circ \cdots \circ J_i'$, the sets $\{\|L_{\bar{s}}\|^{1/m}, \|L_{\bar{s}}^{-1}\|^{-1/m}\}$, $L_{\bar{s}} = L \mid M_{\bar{s}}$, are contained in small vicinities of the intervals $[g_{\bar{s}}^{-}, g_{\bar{s}}^{+}]$ if $m > 0$ is large. Thus, replacing $T$ and $\mathfrak{J}$ by ${}_mT_i = T_{i+m} \circ \cdots \circ T_i$ and ${}_m\mathfrak{J}_i = \mathfrak{J}_{i+m} \circ \cdots \circ \mathfrak{J}_i$ for large $m > 0$, we obtain the mappings $\mathcal{E}_{{}_mT_i,{}_m\mathfrak{J}_i,t,t',f_{\mathcal{R}},f'_{\mathcal{R}}}$ possessing the desired hyperbolicity properties, which proves the result sought for.

*Remark 10.* In Remark 6 and in Proposition 3, the linear operators $\mathcal{D}_{T_i}$ and $L_i = \mathcal{D}_{J_i}$ in $E_*$ are not close, although their norms admit the close upper estimates. Indeed, let $T\colon U \to U$ be a $C^N$-mapping possessing a nondegenerate fixed point and let $J$ be the corresponding linear part of $T$. Assume that a mapping $T'\colon U \to U$ is $C^N$-close to $T$. Then, using the Chain Rule, one can represent the $N$-th differential of $(\mathcal{D}_{T'} - \mathcal{D}_T)(f)$ at a point $x \in U$ as a sum of monomials such that the norm of one of them,

$$J^{-1}\Big(d^N f \mid_{T'(x)} - d^N f \mid_{T(x)}\Big)\Big(\underbrace{dT \mid_x, \cdots, dT \mid_x}_{N}\Big),$$

can be estimated from above only in terms of the continuity modulus of $d^N f$, while the norms of the other monomials admit a suitable upper estimate $\|T' - T\|_{C^N} \cdot \left\|J^{-1}\right\|^2 P_N\big(\|T\|_{C^N}\big) \cdot \|f\|_{C^N}$ where $P_N$ is a polynomial depending on $N$ only. Therefore, Remark 6 is unimprovable in the above-mentioned sense if one imposes no restrictions on the continuity modulus of the $N$-th differential of $f$. Due to the same reason, the operator $\mathcal{D}_T$ and the mappings $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$, $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ in the topology of the uniform convergence over bounded sets do not depend continuously on $T \in C^N$ although $\mathcal{D}_T f$, $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$, $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$ do for each $f \in E$ or $f_{\overline{\mathcal{R}}} \in E_{\overline{\mathcal{R}}}$.

*Remark 11.* The previous Corollary is easily established by standard means of hyperbolic theory (the finite-dimensional case with invertible mappings $S_i$ is usually considered. We deal here with the *infinite-dimensional Banach case with non-invertible* $S_i$). We sketch the underlying geometrical ideas that are very natural and can be found in many works (see also a relevant discussion in [6] and [24]). These ideas are "nonautonomous" and many-dimensional variant of a geometrical approach used by J. Hadamard [21] to prove the presence of separatrices (invariant curves) of a hyperbolic fixed point of a two-dimensional diffeomorphism (see also an exposition in [33]; J. Hadamard considered only the two-dimensional case, but this is of no importance). This approach was denoted as *"graph transform"* in [24]. In essence, we will construct a mathematical object that was called an *itinerary scheme* by V. M. Alekseev [2, Part 1]. Different names were attached to the elements of the below-defined spaces $\Sigma^-$, $\Sigma^+$. For example, they can be called *unstable* and *stable slices*. Let $R > 0$ be large enough and $D = D^+ \times D^-$ be the "rectangle", where $D^\pm$ is an $R$-ball in $E^\pm$ (centered at zero). Consider spaces $\Sigma^\pm$ of continuous mappings $f^\pm\colon D^\pm \to D^\mp$ whose Lipschitz constants are $\leqslant \rho^\pm$, where the positive numbers $\rho^-$ and $\rho^+$ are such that $\rho^-\rho^+ < 1$. We will identify functions $f^\pm$ and their graphs. Let $S = S_i = \mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$. Then the correspondences $\alpha \mapsto S^{\mp 1}(\alpha) \cap D$, for $\alpha \in \Sigma^\pm$, will determine well-defined contracting (in the usual $C^0$-norm) mappings $\mathfrak{S}_i^\pm = \mathfrak{S}^\pm\colon \Sigma^\pm \to \Sigma^\pm$ (although the mappings $S_i$ are non-invertible) if small enough $\varepsilon > 0$ was chosen depending on $\mu$ and $\rho^\pm$. The Lipschitz constant $\lambda < 1$ of these contracting mappings and a lower bound for the radius $R$ depend only on $\mu$, $\rho^\pm$, and $\varepsilon$. In our case $R = 1/2$ is enough. Thus, to (12) there corresponds a double-infinite chain

$$\cdots \xleftarrow{\mathcal{E}_{T_{-1},f_{-1},f_{-2}}} \xleftarrow{\mathcal{E}_{T_0,f_0,f_{-1}}} \xleftarrow{\mathcal{E}_{T_1,f_1,f_0}} \xleftarrow{\mathcal{E}_{T_2,f_2,f_1}} \cdots . \tag{14}$$

Then the sets $\mathfrak{S}_{i+1}^- \circ \mathfrak{S}_{i+2}^- \circ \cdots \circ \mathfrak{S}_m^-(\Sigma^-)$ and $\mathfrak{S}_i^+ \circ \mathfrak{S}_{i-1}^+ \circ \cdots \circ \mathfrak{S}_{-m}^+(\Sigma^+)$ shrink, respectively, exponentially fast to some elements $\alpha_i^- \in \Sigma^-$ and $\alpha_i^+ \in \Sigma^+$ as $m \to +\infty$. The intersection $\alpha_i^- \cap \alpha_i^+$ consists of a single point due to the inequality $\rho^-\rho^+ < 1$ and coincides with the mapping $l_i - \mathrm{id}$.

We notice that the mapping $\mathfrak{S}^+$ of the stable slices is well-defined and contractive although $S$ may be non-invertible. A simple construction of $\mathfrak{S}^+$, that is based on the invertibility of the restrictions of the mapping $S$ onto unstable slices, can be found in [32] (somewhat other constructions can be also found in [24, Proof of Theorem 5.1; 22])[7] and will be shortly described here. A stable slice $\alpha \in \Sigma^+$ will be transformed into the set $\mathfrak{S}^+(\alpha) = S^{-1}(\alpha) \cap D \in \Sigma^+$ which is the graph of a function $g\colon D^+ \to D^-$ to be constructed as follows. The mapping $S$ transfers an unstable slice $\beta_\xi = \{\xi\} \times D^- \in \Sigma^-$, $\xi \in D^+$, into its image, whose piece which is located inside $D = D^+ \times D^-$ is an unstable slice $\widetilde{\beta_\xi} = \mathfrak{S}^-(\beta_\xi)$. Then $\bigl(\xi, g(\xi)\bigr) \in D^+ \times D^-$ is the pre-image of a unique point of intersection, $\alpha \cap \widetilde{\beta_\xi}$, under the invertible map (diffeomorphism) $S \mid \beta'_\xi\colon \beta'_\xi \to \widetilde{\beta_\xi}$, where $\beta'_\xi = S^{-1}(\widetilde{\beta_\xi}) \cap \beta_\xi$.

Incidentally, one could notice that Proposition 3 would be applicable even if the differentiability was treated in Gâteaux sense only. Indeed, the above estimates for the block components of the Gâteaux differential imply all the required properties of the induced mappings of stable and unstable slices.

In a slightly more general case[8], the construction of the itinerary scheme includes the sequence of identical copies $E_i$ of a Banach space $E = E^+ \oplus E^-$, a sequence of "rectangles" $D_i = D_i^+ \times D_i^- \subset E_i$ equipped with the spaces $\Sigma^-$ and $\Sigma^+$ of unstable and stable slices with some positive maximal "slopes" $\rho^-, \rho^+$ such that $\rho^-\rho^+ < 1$, and a sequence of mappings $S_i\colon D_i \to E_{i+1}$,

$$\cdots \xrightarrow{S_{-1}} \xrightarrow{S_0} \xrightarrow{S_1} \cdots$$

(in contrast to (14) the order is assumed to be direct) such that the deformed rectangle $S(D_i)$ and the rectangle $D_j$ have "good" intersection [30], i.e., they cross. Then the corresponding mappings $\mathfrak{S}_i^-\colon \Sigma_i^- \to \Sigma_{i+1}^-$ and $\mathfrak{S}_i^+\colon \Sigma_{i+1}^+ \to \Sigma_i^+$ are well-defined and contractive with some Lipschitz

[7]In essence, the itinerary scheme was constructed in [22, 32] to establish an analog of the well-known V.M. Alekseev's theorem [3] for a homoclinic structure of a smooth, possibly non-invertible map of a Banach manifold (the simplest case of a single homoclinic trajectory was discussed in [22, 32], but the result obtained is immediately transferable to the general case).

[8]We omit the most general case [2, Part 1] where the itinerary scheme is associated with an arbitrary TMC.

constant $\lambda = \lambda(\rho^+, \rho^-) < 1$ under a proper choice of $\rho^-$ and $\rho^+$. The precise formulations and strict estimates for these geometrically visual results can be found in [2, Part 1].

*Remark 12.* Let us consider the itinerary scheme just described above. Then the stable $\alpha_i^+ \in \Sigma_i^+$ and unstable $\alpha_i^- \in \Sigma_i^-$ slices, defined as above by the itinerary scheme, and their intersection, $r_i = \alpha_i^+ \cap \alpha_i^-$, will depend continuously on the sequence of mappings $S_j$. To attach the precise meaning to the last statement, one should consider $[S_i : i \in \mathbb{Z}]$ as an element of the corresponding Tychonoff product (over $i \in \mathbb{Z}$) of the spaces $V_i$, where $V_i$ is the space of mappings $S_i$ for which $S_i(D_i)$ and $D_{i+1}$ cross and define the contractive maps $\mathfrak{S}_i^- : \Sigma_i^- \to \Sigma_{i+1}^-$ and $\mathfrak{S}_i^+ : \Sigma_{i+1}^+ \to \Sigma_i^+$ with Lipschitz constant $\lambda$. Equip each $V_i$ with the usual $C^0$-norm and recall that $\Sigma_i^\pm$ are equipped with the analogous norms. Then, firstly, $\mathfrak{S}_i^- \beta^- \in \Sigma_{i+1}^-$ ($\mathfrak{S}_i^+ \beta^+ \in \Sigma_i^+$, respectively) depends continuously on the map $S_i \in V_i$ and the slice $\beta^- \in \Sigma_i^-$ ($\beta^+ \in \Sigma_{i+1}^+$). Next, recall that, secondly, the $C^0$-diameter of $\mathfrak{S}_{i-1}^- \circ \cdots \circ \mathfrak{S}_{-m}^-(\Sigma_{-m}^-)$ ($\mathfrak{S}_i^+ \circ \cdots \circ \mathfrak{S}_{m-1}^+(\Sigma_m^+)$, respectively) tends uniformly to zero as $m \to +\infty$ for all the mappings $S_j \in V_j$, and, thirdly, the intersection map $\varkappa_i \colon \Sigma_i^+ \times \Sigma_i^- \to D_i$ is continuous. It is easily seen from these results that $\alpha_i^\pm$ and $r_i = \alpha_i^+ \cap \alpha_i^-$ depend continuously on $[S_i]$. (Moreover, the mapping $\varkappa_i$ and the dependence of $\mathfrak{S}_i^\pm \beta^\pm$ on $S_i \in V_i$ are Lipschitz with some constants determined by $\rho^+$ and $\rho^-$ only. Therefore, the dependencies of $\alpha_i^\pm$ and $r_i$ on $S_j \in V_j$ are also Lipschitz with constant $C\lambda^{|i-j|}$ where $C = C(\rho^+, \rho^-) > 0$.) For further purposes, we notice that the analogous results can be easily proven if the spaces $\Sigma_i^\pm$ and $V_i$ are equipped with the pointwise convergence topologies. We emphasize that these topologies are finer than the $C^0$-topologies if $E^\pm$ or $E$ are infinite-dimensional, respectively. This is a direct consequence of the analogous result that in the space of linear operators $E^\pm \to \mathbb{R}$, whose norms do not exceed a given positive constant, the topology defined by the operator norm (the strong topology) is coarser than that of the pointwise convergence (the weak topology).

*Remark 13.* One can establish completely analogous to Remark 19 of [14] that the $C^N$-mappings $l_i$ depend continuously on the sequence of the $C^N$-mappings $T_i$ and on the sequence of the polynomial mappings $(l_i)_\mathcal{R}$ if all the mappings are treated in the $C^k$-topology, $1 \leqslant k \leqslant N$, and the sequences are treated in the Tychonoff topology. This result remains valid in the case $k = 0$ provided that the perturbations of $T_i$ are small enough in the $C^1$-sense. We omit the explanation of these facts since it consists of repeating the arguments described in Remark 19 of [14] and in Remark 6 above. Recall only that it is based on Remark 12 above.

*Remark 14.* The nonautonomous normalization $l_i$ constructed above admits an invariant coordinate-free description. The explanation of this fact repeats almost literally the arguments presented in item **c)** of Remark 20 in [14]. Indeed, consider a sequence $\{T_i\}$ that satisfies the conditions of Proposition 3. Let $U_i = U$ be the definition domain of the mapping $T_i \colon U_i \to U_{i+1}$. There exists a unique sequence $r_i = 0 \in U$ such that $T_i(r_i) = r_{i+1}$. The nonautonomous normalization is given by the mappings $l_i \colon U_i \to \mathbb{R}^n$ such that $l_i(r_i) = 0$, $dl_i(r_i) = \mathrm{id} \mid \mathbb{R}^n$. Using the "natural" coordinates in $U_i \subset \mathbb{R}^n$, one can identify the spaces $\mathbb{R}^n$ containing the images $\operatorname{im} l_i$ with the tangent spaces $T_{r_i}U_i$. Then the nonautonomous $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalization is just the sequence of $C^N$-mappings $l_i \colon U_i \to T_{r_i}U_i$ uniformly $C^N$-bounded and such that $l_i(r_i) = 0 \in T_{r_i}U_i$, $dl_i(r_i) = \mathrm{id} \mid T_{r_i}U_i \equiv \mathrm{id} \colon T_{r_i}U_i \circlearrowleft$ (the space $T_rU$ is equipped with the natural linear structure and, therefore, for each $\xi \in T_rU_i$ there is the canonical identification $T_rU_i \equiv T_\xi(T_rU_i)$ used here for $\xi = (r_i, 0)$, $0 \in T_{r_i}U_i$) and diagram (12) commutes, where $\mathfrak{J}_i \colon T_{r_i}U_i \to T_{r_{i+1}}U_{i+1}$ are the $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal forms. Nevertheless, the Corollary of Proposition 3 does not admit a completely coordinate-free description since the "resonant parts" $(l_i)_\mathcal{R}$ that should be prescribed depends on the coordinates used in the domain $U_i$. Indeed, recall from item **c)** of Remark 20 in [14] the following fact. If $l_i \colon \mathbb{R}^n \supset U \to \mathbb{R}^n = T_0U$ is an arbitrary mapping written in the natural coordinates then in the coordinates $y = g_i(u), u \in U_i$, where $g_i(0) = 0$, this mapping possesses the form $l_i' = \mathcal{D}_{g_i^{-1}} l_i$. Thus, generally speaking, the new "resonant part" $(l_i')_{\mathcal{R}}$ will depend on the old "non-resonant part" $(l_i)_{\overline{\mathcal{R}}}$.

## 4. Proof of Proposition 3 and further discussion

The proof of Proposition 5 in [14] requires some modifications only which will be described in all detail. At the end of the Section, we will indicate the importance of the stability condition for the set $\widetilde{\mathcal{R}}$ associated with the $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalizations. In particular, a result generalizing the Centralizer Theorem of [8,20] will be established.

First of all, we will state an important Lemma which is related to the smooth properties of the composition map Comp acting by the formula $(f, g) \mapsto f \circ g$ in the functional spaces. For any two domains $U$, $V$ in Banach spaces we denote by $C_b^r(U, V)$ the linear space of $C^r$-functions $f : U \to V$ with the finite $C^r$-norm $\|f\|_{C^r} = \sup_{x\in U} \sum_{t=0}^{r} \big\| d^t f \mid_x \big\|$. Here and below the differentiability is understood in the sense of Frechét.

**Lemma 1.** *Let $U$, $V$, $W$ be some domains in Banach spaces, $r \geqslant 0$, $s \geqslant 1$. The composition map* Comp : $C_b^{r+s}(V, W) \times C_b^r(U, V) \to C_b^r(U, W)$ *is linear in the first argument and $s$ times differentiable in the collection of the arguments at the points $(f, g)$ such that the $(r + s)$-th differential of $f$ is uniformly continuous in $V$. The associated differential is calculated by the formula*

$$d\,\mathrm{Comp} \mid_{(f,g)} (\delta f, \delta g)(\cdot) = \delta f \circ g(\cdot) + \nabla f \mid_{g(\cdot)} \cdot \delta g(\cdot) \tag{15}$$

*where $\nabla$ denotes the differential in the phase space (i.e., the gradient) and $\delta$, the differential in the functional spaces.*

The corresponding statements on the linearity in the first argument and the $s$-times differentiability in the second argument are local variants of the so-called Alpha- and Omega-Theorems. The global variant of these Theorems [1, 16] (see also [34]) are related to functional spaces of the mappings of compact manifolds (but not charts). (Notice also that we treat with infinite-dimensional charts, in contrast to compact manifolds that are covered by finite-dimensional charts; the compactness is needed only to provide the finiteness of the $C^r$-norms of smooth functions and the uniform continuity of continuous functions.) An independent proof which is less formal than that in [1] will be now presented.

*Proof of Lemma 1.* Let $r = 0$ and $s = 1$. Using the Chain Rule, the Mean Value Theorem, and the Lagrange Theorem, one can estimate from above the modulus of the difference of $\mathrm{Comp}(f + \delta f, g + \delta g) - \mathrm{Comp}(f, g)$ and the right-hand side of (15) as $\big\|\nabla(\delta f)\big\| \|\delta g\| + \omega_{\nabla f}\big(\|\delta g\|\big)\|\delta g\|$ where $\omega_h$ denotes the continuity modulus of $h$ and $\|\cdot\|$, the $C^0$-norm. This estimate is seen to be $o(\varepsilon)$ uniformly as $\varepsilon = \|\delta f\|_{C^1} + \|\delta g\|_{C^0} \to 0$. This verifies the differentiability and gives the equality (15) in the case $r = 0$, $s = 1$. Obviously, (15) remains valid in the general case provided that the differentiability of Comp is established. This is an elementary corollary of the fact that the convergence in the $\|\cdot\|_{C^l}$-norm implies that in the $\|\cdot\|_{C^t}$-norm where $t < l$.

Consider also a linear operator $\alpha_g : C_b^l(V, E) \to C_b^0(U, E)$ depending on $g \in C_b^0(U, V)$ and acting by the formula $\alpha_g(f) = \mathrm{Comp}(f, g)$. Here the domain $W$ is taken to coincide with the all ambient Banach space $E$. The above estimate shows that the operator $\alpha_g$ as being equipped with the natural operator norm is differentiable in $g$ when $l \geqslant 2$. This corresponds to the case where the first argument of Comp is assumed to be fixed and the differential looked for takes the form

$$d\alpha_g(\delta g)(f) = \alpha_g(\nabla f) \cdot \delta g.$$

We obtain by the induction that the operator $\alpha_g$ is $(l-1)$ times differentiable in $g$. The desired result for $r=0$ and arbitrary $s \geqslant 1$ will follow immediately from this. Indeed, since the right-hand side of (15) takes the form

$$\operatorname{Comp}(\delta f, g) + \operatorname{Comp}(\nabla f, g) \cdot \delta g,$$

the differential $d\operatorname{Comp}|_{(f,g)}$ being considered as the linear operator $(\delta f, \delta g) \mapsto C_b^r(U,E)$ will be $(s-1)$ times differentiable in $(f,g)$, by the induction in $s$.

Obviously, $T(f\circ g) = Tf \circ Tg$ where $Th : x \mapsto \big(x, dh|_x\big)$ is the tangent map that is associated to the mapping $h : x \mapsto h(x)$. Therefore, identifying $\|h\|_{C^{l+1}}$ with $\|Th\|_{C^l}$, one can transfer immediately the desired result to the general case of arbitrary $r \geqslant 0$. $\square$

*Proof of Proposition 3.* Recall briefly some basic ideas of the proof of Proposition 5 of [14]. The existence of the required ball $U$ follows immediately from conditions ii) and iii). The operators $L_i$ are hyperbolic and the separatrices $E^+$, $E^-$ for all of them coincide if condition i) is satisfied. Indeed, let $\Phi^-$ ($\Phi^+$, respectively) be the finite set of all the couples $(s,m) \notin \Phi^0 = \mathcal{R}$ such that $2 \leqslant |m| < N$ and $|\lambda_s| < |\lambda^m|$ ($|\lambda_s| > |\lambda^m|$, respectively) in inequality (1), where $\{\lambda_j\}$ is the spectrum of $J_i$. Then let $\Phi$ be constituted by the non-intersecting subsets $\Phi^-$ and $\Phi^+$. Introduce $\widetilde\Phi^0 = \Psi(\Phi^0)$, $\widetilde\Phi^\pm = \Psi(\Phi^\pm)$, and $\widetilde\Phi = \Psi(\Phi) = \widetilde\Phi^+ \cup \widetilde\Phi^-$ and notice that the correspondence $\Psi$ does not depend on the number $i$ due to condition i). Let the non-intersecting sets $F^-$, $F^+$ constituting $F$ be defined as

$$F^\pm = \bigoplus_{(\tilde s,\widetilde m)\in\widetilde\Phi^\pm} G_{\tilde s,\widetilde m}.$$

Then the outgoing separatrix $E^-$ is $F^-$ and the incoming separatrix $E^+$ is the space of functions that have at the origin $0\in\mathbb{R}^n$ $(N-1)$-jets lying in $F^+$, i.e., $E^+ = F^+ \oplus G_*$. Indeed, in the spaces $G_* \subset E$ and $G_{\tilde s,\widetilde m}$ invariant under $L_i$, the following norms $\|\cdot\|_*$ and $\|\cdot\|_{\tilde s,\widetilde m}$ will be induced. The norm $\|\cdot\|_*$ was defined above as the exact upper bound for the norm of the $N$-th differential, and $\|\cdot\|_{\tilde s,\widetilde m}$ is the norm in the space of polynomials $P$ that are considered as polylinear mappings $P\colon M_1^{\widetilde m_1}\times\cdots\times M_p^{\widetilde m_p} \to M_{\tilde s}$ symmetrical in each cofactor $M_t^{\widetilde m_t}$ (the case $p=1$ being discussed in [12] but the results being transferable immediately to the case of arbitrary $p\geqslant 1$). The corresponding operator norms (denoted by the same symbols) satisfy inequalities

$$\begin{aligned}
\Big\|\Big(L_i \mid G_{\tilde s,\widetilde m}\Big)^{-1}\Big\|_{\tilde s,\widetilde m} &\leqslant \mu \quad \text{for} \quad (\tilde s,\widetilde m)\in\widetilde\Phi^-,\\
\Big\|L_i \mid G_{\tilde s,\widetilde m}\Big\|_{\tilde s,\widetilde m} &\leqslant \mu \quad \text{for} \quad (\tilde s,\widetilde m)\in\widetilde\Phi^+, \qquad \text{and}\\
\Big\|L_i \mid G_*\Big\|_* &\leqslant \mu
\end{aligned}$$

for some $0<\mu<1$. In our case

$$\mu = \max\left\{ \max_{(\tilde s,\widetilde m)\in\widetilde\Phi^-} \big\|J_{i,\tilde s}\big\| \prod_{j=1}^p \big\|J_{i,j}^{-1}\big\|^{\widetilde m_j}; \quad \max_{(\tilde s,\widetilde m)\in\widetilde\Phi^+} \big\|J_{i,\tilde s}^{-1}\big\| \prod_{j=1}^p \big\|J_{i,j}\big\|^{\widetilde m_j} \right\}. \tag{16}$$

It follows from condition ii) that the number $\mu$ can be chosen independently of $i$. Define a norm in

$$E_{\overline{\mathcal{R}}} = E^+ \oplus E^- = \bigoplus_{(\tilde{s},\widetilde{m})\in\widetilde{\Phi}} G_{\tilde{s},\widetilde{m}} \oplus G_*$$

as the direct sum (or maximum) of the norms in $G_{\tilde{s},\widetilde{m}}$, $G_*$. Then $\left\|\left(L_i \mid E^{\pm}\right)^{\pm 1}\right\| \leqslant \mu$. The required norm in $E_{\overline{\mathcal{R}}}$ is built of the norms in $G_{\tilde{s},\widetilde{m}}$ and $G_*$ with corresponding weight coefficients $\varkappa_{\widetilde{m}}$ and $\varkappa_*$ that decrease quickly as $|\widetilde{m}|$ rises (one supposes $|\widetilde{m}| = N$ for $G_*$),

$$\Bigl\| \bigoplus_{(\tilde{s},\widetilde{m})\in\widetilde{\Phi}} u_{\tilde{s},\widetilde{m}} \oplus u_* \Bigr\| = \varkappa_* \|u_*\|_* + \sum_{(\tilde{s},\widetilde{m})\in\widetilde{\Phi}} \varkappa_{\widetilde{m}} \|u_{\tilde{s},\widetilde{m}}\|_{\tilde{s},\widetilde{m}}$$

(the sums in the right-hand side can be replaced by the maxima).

The main trouble now is to construct the nonlinear mappings $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and to prove the required properties of $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$. Let $G_t = \bigoplus_{\substack{(\tilde{s},\widetilde{m})\in\widetilde{\Phi} \\ |\widetilde{m}|=t}} G_{\tilde{s},\widetilde{m}}$ and $G_t^0 = \bigoplus_{\substack{(\tilde{s},\widetilde{m})\in\widetilde{\Phi}^0 \\ |\widetilde{m}|=t}} G_{\tilde{s},\widetilde{m}}$ for $2 \leqslant t < N$, $G_N = G_*$, and let $f_{\mathcal{R}} = \oplus_t u_t^0$, $f_{\overline{\mathcal{R}}} = \oplus_t u_t$, $f'_{\mathcal{R}} = \oplus_t v_t^0$, $f'_{\overline{\mathcal{R}}} = \oplus_t v_t$, $\mathfrak{J}_{\mathcal{R}} = \oplus_t z_t^0$, $\mathfrak{J}_{\overline{\mathcal{R}}} = 0$, $T_{\mathcal{R}} = \oplus_t \xi_t^0$, $T_{\overline{\mathcal{R}}} = \oplus_t \xi_t$, $(d_T)_{\mathcal{R}} = \oplus_t \eta_t^0$, $(d_T)_{\overline{\mathcal{R}}} = \oplus_t \eta_t$, where $u_t, v_t, \xi_t, \eta_t \in G_t$ and $u_t^0, v_t^0, z_t^0, \xi_t^0, \eta_t^0 \in G_t^0$ (the subscript $i$ is omitted for brevity). Then $\eta_t = J^{-1}\circ\xi_t$ and $\eta_t^0 = J^{-1}\circ\xi_t^0$. The crucial observation is that the mappings $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ have triangular form. Firstly, we will show this fact for their "counterparts", $\mathcal{E}^{\leftarrow}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{I}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$. Indeed, if $\mathfrak{J}_{\mathcal{R}} = \oplus_t w_t^0$, $w_t^0 \in G_t^0$, $\mathfrak{J}_{\overline{\mathcal{R}}} = 0$, then

$$\begin{aligned} v_t &= \eta_t + Lu_t + \Theta_t\Bigl(\bigoplus_{t'<t}(u_{t'} \oplus u_{t'}^0), \bigoplus_{t'<t} w_{t'}^0\Bigr), \quad t < N, \\ v_N &= \eta_N + \mathcal{D}_T u_N + \Theta_N\Bigl(\bigoplus_{t'<N}(u_{t'} \oplus u_{t'}^0), \bigoplus_{t'<N} w_{t'}^0, u_N \circ T\Bigr), \end{aligned} \tag{17}$$

where quantities $\Theta_t\,(2 \leqslant t \leqslant N)$ are affine (i.e., linear but non-homogeneous) in the second argument and polynomial in the first one. Here, free terms of $\Theta_t$'s with respect to the second argument are linear (homogeneous) in the first argument, $\Theta_N(\cdot, 0, g) = \pi_*\mathcal{D}_T(\cdot)$ with $\pi_* : E \to G_*$ being the natural projection. As for the dependence of $\Theta_t\,(t < N)$ on the mapping $T$, which is not explicitly indicated here, it is polynomial in $\oplus_{t'<t}(\xi_{t'} \oplus \xi_{t'}^0)$ and $J$, $J^{-1}$. Notice that $\Theta_t$'s $(t < N)$ take their values into the finite-dimensional spaces $G_t$ and are usual polynomial functions. In contrast to them, $\Theta_N(\overline{u}, \mathfrak{J}_{\mathcal{R}}, g)$ takes values in the space $G_*$ and, as a function of $y \in U$, is a finite sum of the above mentioned free term, $\pi_*\mathcal{D}_T(\overline{u})$, and terms of type (9) and

$$\bigl(\pi_* T_{w_{t'}^0}\bigr)(\cdot), \tag{18}$$

where the revised notation $T_w = w \circ (\mathrm{id} + \overline{u}) \circ T$ is used in both the formulas (9) and (18). Recall that the homogeneous polynomials $\widetilde{w}_\alpha$, $\widetilde{\widetilde{w}}_\beta$ $(\alpha \geqslant 0,\ \beta \geqslant 1)$ included in each term represented by (9) are linear dependent on the homogeneous form $w_{t'}^0$ of degree $t' = \alpha + \beta$, $2 \leqslant t' < N$ in the polynomial mapping $\mathfrak{J}$. To each value of the index $t'$, there correspond, generally speaking, several (one-dimensional) terms of type (9) and a single (multi-dimensional) term of type (18).

To arrange the map $\mathcal{I}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$, one should utilize the fact that

$$v_t^0 = w_t^0 \circ J + \eta_t^0 + L^0 u_t^0 + \Theta_t^0\Bigl(\bigoplus_{t'<t}(u_{t'} \oplus u_{t'}^0), \bigoplus_{t'<t} w_{t'}^0\Bigr), \quad t < N, \tag{19}$$

where $\Theta_t^0$'s are completely analogous to $\Theta_t$'s, and then determine quantities $w_t^0$ as polynomial functions of $u_t^0, v_t^0$ and $u_{t'}, u_{t'}^0, v_{t'}, v_{t'}^0, w_{t'}^0$ with $t' < t$.

Thus, applying formulas (17) and (19) one can successively express $v_t$ and $w_t^0$ $(t < N)$ via $u_t, u_t^0, v_t^0$ and $u_{t'}, u_{t'}^0, v_{t'}, v_{t'}^0, w_{t'}^0$ with less indices $t' < t$. As a result, one obtains the following formulas for the mapping $\mathcal{E}^{\leftarrow}_{T,f_\mathcal{R},f'_\mathcal{R}}$

$$\begin{aligned} v_t &= \eta_t + L u_t + \overline{\Theta}_t\Big(\bigoplus_{t'<t}(u_{t'} \oplus u_{t'}^0), \bigoplus_{t'<t} v_{t'}^0\Big), \quad t < N, \\ v_N &= \eta_N + \mathcal{D}_T u_N + \overline{\Theta}_N\Big(\bigoplus_{t'<N}(u_{t'} \oplus u_{t'}^0), \bigoplus_{t'<N} v_{t'}^0, u_N \circ T\Big), \end{aligned} \tag{20}$$

and for the mapping $\mathcal{I}_{T,f_\mathcal{R},f'_\mathcal{R}}$

$$w_t^0 = \big(v_t^0 - \eta_t^0 - L^0 u_t^0\big) \circ J^{-1} + \overline{\Theta}_t^0\Big(\bigoplus_{t'<t}(u_{t'} \oplus u_{t'}^0), \bigoplus_{t'<t} v_{t'}^0\Big), \quad t < N, \tag{21}$$

where $\overline{\Theta}_t$, $\overline{\Theta}_t^0$ $(2 \leqslant t < N)$ and $\overline{\Theta}_N$ are completely analogous to $\Theta_t$, $\Theta_t^0$ and $\Theta_N$. The mappings sought for are constructed for the case of the inverted lower arrow in the diagram.

In the sequel, for the brevity we will write $\mathcal{E}^{\leftarrow}$ and $\mathcal{I}$ instead of $\mathcal{E}^{\leftarrow}_{T,f_\mathcal{R},f'_\mathcal{R}}$ and $\mathcal{I}_{T,f_\mathcal{R},f'_\mathcal{R}}$. Now we establish the differentiability of the mappings $\mathcal{E}^{\leftarrow}$ and $\mathcal{I}$. Since the natural projections of $E$ onto $G_t$, $G_t^0$ are linear, the dependencies of $w_t^0$, $v_t$ $(2 \leqslant t < N)$ on $f_{\overline{\mathcal{R}}} \in E_{\overline{\mathcal{R}}}$ are differentiable (even $C^\infty$). One needs to prove only the differentiability of the simultaneous dependence of $v_N$ on $u_N$ and on the first two arguments of $\Theta_N$ which is expressed by the last formula (17). This desired result follows easily from Lemma 1. Thus, the mappings $\mathcal{E}^{\leftarrow}$ and $\mathcal{I}$ are differentiable and even continuously differentiable since their differentials depend continuously on $f_{\overline{\mathcal{R}}} \in E_{\overline{\mathcal{R}}}$.

Recall that the mapping $T$ is uniformly $C^N$-bounded and the operators $J$, $J^{-1}$ are uniformly bounded. At first, we explain briefly how to meet the required estimates for the differential of the map $\mathcal{E}^{\leftarrow}_{T,f_\mathcal{R},f'_\mathcal{R}}$ assuming that $f_\mathcal{R}, f'_\mathcal{R}$ and $f_{\overline{\mathcal{R}}}$ run over some given bounded subsets in $E_\mathcal{R}$ and $E_{\overline{\mathcal{R}}}$. Let the coefficients $\varkappa_{\widetilde{m}}, \varkappa_*$ be chosen in such a way that all the fractions $\varkappa_{\widetilde{m}'}/\varkappa_{\widetilde{m}''}$ are small enough (do not exceed $O(\varepsilon)$) when $|\widetilde{m}'| > |\widetilde{m}''|$. Then the differentials of the mappings $\mathcal{E}^{\leftarrow}_{T,f_\mathcal{R},f'_\mathcal{R}}$ and $\mathcal{I}_{T,f_\mathcal{R},f'_\mathcal{R}}$ can be evaluated with an error uniformly small in the induced operator norm via the use of formulas (20), (21) where the right-hand sides are simplified as follows. The quantities $\overline{\Theta}_t$, $\overline{\Theta}_t^0$ $(t < N)$ are omitted and the $(N-1)$-jet of $f_{\overline{\mathcal{R}}}$ at the origin, $\oplus_{t'<N} u_{t'}$, is considered as a constant while calculating the differential of $\overline{\Theta}_N$. (It is important that the subsequently narrowing the neighbourhood $U$ does not involve growing an upper bound for the error caused by the simplified calculation of the differential of $\overline{\Theta}_N$.) So, the differential of $\mathcal{E}^{\leftarrow}$ can be made to be arbitrarily close to the diagonal mapping and the differential of $\mathcal{I}$, to be arbitrarily small. The required estimate $\|\mathcal{D}_T \mid G_*\|_* \leqslant \mu + \varepsilon/3$ is provided by the arguments of Remark 6 if the neighbourhood $U$ is chosen to be of a size of order $\varepsilon$. One needs to prove only that the differential of $\overline{\Theta}_N(\overline{u}, \overline{v}^0, u_N \circ T)$ along $u_N \in G_*$, i.e., the partial derivative $(\delta/\delta u_N)\overline{\Theta}_N(\overline{u}, \overline{v}^0, u_N \circ T)$, is sufficiently small where the notations $\overline{u} = \oplus_{t'<N}(u_{t'} \oplus u_{t'}^0)$ and $\overline{v}^0 = \oplus_{t'<N} v_{t'}^0$ are introduced for the $(N-1)$-jets of the functions $f$ and $f'_\mathcal{R}$ at the origin. This result follows from the fact that $\overline{\Theta}_N(\overline{u}, \overline{v}^0, u_N \circ T) = \Theta_N(\overline{u}, \mathfrak{I}_\mathcal{R}, u_N \circ T)$ where $\mathfrak{I}_\mathcal{R} = \oplus_{t'<N} w_{t'}^0$ does not depend on $u_N$ and from the arguments of Remark 6 related to the function $\Theta_{T,\mathfrak{I}}$.

The construction of the required norm is achieved by more accurately choosing coefficients $\varkappa_{\widetilde{m}}$. It suffices to prove the desired estimates for the differential $\mathcal{E}^{\leftarrow}$ at all $f_{\overline{\mathcal{R}}} = \bigoplus_{(\tilde{s},\widetilde{m}) \in \widetilde{\Phi}} u_{\tilde{s},\widetilde{m}} \oplus u_*$ such that

$$\varkappa_{\widetilde{m}} \|u_{\tilde{s},\widetilde{m}}\|_{\tilde{s},\widetilde{m}} \leqslant 1, \quad \text{where} \quad (\tilde{s}, \widetilde{m}) \in \widetilde{\Phi}, \qquad \text{and} \quad \varkappa_* \|u_*\|_* \leqslant 1. \tag{22}$$

Indeed, the unit ball of the norm constructed lies entirely in the set (22). For a given $t < N$, let $\overline{\Theta}_t$ be a polynomial of degree $k$ in the first argument and let all the coefficients $\varkappa_{\widetilde{m}'}$ with $|\widetilde{m}'| = t' < t$ be already chosen. When, in order the norm of the contribution of the function $\overline{\Theta}_t$ into the differential $d\mathcal{E}^{\leftarrow}$ do not exceed $\varepsilon$, it suffices to demand the inequalities $\varkappa_{\widetilde{m}}/\varkappa^k_{\widetilde{m}'} < O(\varepsilon)$ for all $\widetilde{m}, \widetilde{m}'$ such that $|\widetilde{m}'| = t' < t = |\widetilde{m}|$. Here, the $O$-estimates are independent of the coefficients $\varkappa_{\widetilde{m}'}$. In an analogous way, the contributions of the terms in $\overline{\Theta}_N$ which do not depend on $u_*$ (i.e., the term $\pi_*\mathcal{D}_T\overline{u}$ free with respect to the second argument and the terms represented by formula (18)) can be made to be $\varepsilon$-small by choosing a small $\varkappa_*$. The analysis of the terms represented by formula (9) requires a more accuracy. Fortunately, partial derivatives of these terms in $u_t\,(t < N)$ are sums of terms which take the same form where $\widetilde{w}_\alpha$ or $\widetilde{\widetilde{w}}_\beta$ is replaced by its partial derivative, $d_{u_t}\widetilde{w}_\alpha$ or $d_{u_t}\widetilde{\widetilde{w}}_\beta$, or the function $T_{\widetilde{w}_\alpha}$ is replaced by its partial derivative $d_{u_t}\,|_{\widetilde{w}_\alpha=\mathrm{const}}\,T_{\widetilde{w}_\alpha}(\cdot) = T_{\nabla\widetilde{w}_\alpha}(\cdot)\,du_t \circ T(\cdot)$ calculated under constant $\widetilde{w}_\alpha$. We notice that $d_{u_t}\widetilde{w}_\alpha$ and $d_{u_t}\widetilde{\widetilde{w}}_\beta$ as well as $\widetilde{w}_\alpha$ and $\widetilde{\widetilde{w}}_\beta$ are homogeneous polynomials of degrees $\alpha$ and $\beta$ depending polynomially on $\overline{u}$ and $\overline{v}$. Later, the $C^N$-function $d_{u_t}\,|_{\widetilde{w}_\alpha=\mathrm{const}}\,T_{\widetilde{w}_\alpha}$ is analogous to $T_{\widetilde{w}_{\alpha'}}$ with $\alpha' = \alpha + t - 1 > \alpha$ in the sense that it has zero $(\alpha' - 1)$-jet at the origin and depends polynomially on $\overline{u}$ and $\overline{v}$. Therefore, these partial derivatives of terms (9) under discussion can be made to be arbitrarily small by choosing a small enough ball $U$ for a given $\varkappa_*$. Indeed, the terms constituting the partial derivatives under consideration are built with the use of the linear operators (11) in the space $G_*$ and the nonlinear map $(\cdot) \mapsto f \circ (\cdot)$ in $G_*$ where $f \in C^N$ in a vicinity of the origin, $f(0) = 0$ (in fact, $f = \widetilde{\widetilde{w}}_\beta$). It is easily seen that $\|f \circ v\|_*$ admits a uniform upper estimate $\|f\|_{C^N} S_N\big(\|v\|_*, \varepsilon\big)$ which can be multiplied by $\varepsilon$ if, in addition, $df(0) = 0$. Here, as above, $S_N$ is a polynomial depending on $N$ only, $\varepsilon$ is the radius of the ball $U$. The result sought for is implied immediately by these estimates and the above-written estimates for the norms of operators (11), taking into account the fact that certainly $\alpha' \geqslant \alpha \geqslant 1$ or $\beta \geqslant 2$. Later, the above-described arguments show that the norm of $\mathcal{D}_T \mid G_*$ will admit an upper estimate close to $\mu$ and the partial derivative $(\delta/\delta u_N)\overline{\Theta}_N(\overline{u}, \overline{v}^0, u_N \circ T)$ will be small provided that the size of the neighbourhood $U$ is chosen small enough. Finally, if all the coefficients $\varkappa_{\widetilde{m}}$, $\varkappa_*$ were chosen small enough then the given bounded set $B_{\overline{\mathcal{R}}}$ lies inside the unit circle of the norm constructed, $\|\cdot\|$, and the conditions $\big\|\mathcal{E}^{\leftarrow}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}(0)\big\| \leqslant \varepsilon$ are satisfied for all $f_{\mathcal{R}}, f'_{\mathcal{R}} \in B_{\mathcal{R}}$. The analog of Proposition 3 for the mappings $\mathcal{E}^{\leftarrow}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{I}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ is proven.

Now we will show that the structure of the original required mappings $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ and $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$ is analogous and this fact will enable us to extend the results just obtained to the latter mappings. Let $\mathfrak{I}_{\mathcal{R}} = \oplus_t w^0_t$, $\mathfrak{I}_{\overline{\mathcal{R}}} = \oplus_t w_t$, where $w_t \in G_t$ and $\mathfrak{I} = \mathfrak{J}^{-1}$. Completely analogously to the considerations above, we deduce from the relation $\mathfrak{I} \circ \mathfrak{J} = \mathrm{id}$ that

$$
\begin{aligned}
0 &= w_t \circ J + \Xi_t\Big(\bigoplus_{t'<t} z^0_{t'}, \bigoplus_{t'<t}(w_{t'} \oplus w^0_{t'})\Big),\\
0 &= I \circ z^0_t + w^0_t \circ J + \Xi^0_t\Big(\bigoplus_{t'<t} z^0_{t'}, \bigoplus_{t'<t}(w_{t'} \oplus w^0_{t'})\Big), \quad t < N,\\
0 &= w_N \circ \mathfrak{J} + \Xi_N\Big(\bigoplus_{t'<N} z^0_{t'}, \bigoplus_{t'<N}(w_{t'} \oplus w^0_{t'})\Big).
\end{aligned}
$$

Here $\Xi_t$, $\Xi^0_t$ are completely analogous to $\Theta_t$, $\Theta^0_t$, and $\Xi_N(\mathfrak{I}_{\mathcal{R}}, \overline{w})$ as a function of $y \in U$, is a polynomial.

Thus, one can successively express $w_t, w^0_t$ $(t < N)$ via $z^0_t$ and $z^0_{t'}, w_{t'}, w^0_{t'}$ with less indices

$t' < t$. As a result, one obtains

$$\begin{aligned} w_t &= \overline{\Xi}_t\Big(\bigoplus_{t'<t} z^0_{t'}\Big), \\ w^0_t &= -I \circ z^0_t \circ J^{-1} + \overline{\Xi}^0_t\Big(\bigoplus_{t'<t} z^0_{t'}\Big), \quad t < N. \end{aligned} \tag{23}$$

By the Implicit Function Theorem (in Banach spaces), one easily sees from $w_N \circ \mathfrak{J} = -\Xi_N$ that the dependence of $w_N$ on $\mathfrak{J}$ is continuously differentiable with the differential $\delta w_N(\cdot) = -\big(\nabla w_N(\cdot) \cdot \delta\mathfrak{J} \mid_{\mathfrak{J}(\cdot)} + \delta\Xi_N \mid_{\mathfrak{J}(\cdot)}\big)$. Obviously, this differential is uniformly bounded if $\mathfrak{J}_{\mathcal{R}}$ ranges over a bounded set.

Now, since the mapping $\mathfrak{J}$ may contain non-resonant terms, the following two changes take place in the right-hand sides of formulas (17), (19). Firstly, the terms $w^0_{t'}$ should be replaced by $w_{t'} \oplus w^0_{t'}$, and, secondly, in the expressions (17) for $v_t$ and $v_N$, there appear additional terms $w_t \circ J$ and $w_N \circ (\mathrm{id} + f) \circ T = w_N \circ (\mathrm{id} + \overline{u} + u_N) \circ T$, respectively, where $\overline{u} = \oplus_{t'<N}(u_{t'} \oplus u^0_{t'}) = j^{(N-1)}_0 f$. Substituting the expressions (23) for $w_t, w^0_t$ $(t < N)$ in the right-hand sides of these modified formulas and then expressing successively $v_t$ and $z^0_t$ $(t < N)$ via $u_t$, $u^0_t$, $v^0_t$ and $u_{t'}$, $u^0_{t'}$, $v_{t'}$, $v^0_{t'}$, $z^0_{t'}$ with less indices $t' < t$, one obtains the following formulas for the mapping $\mathcal{E}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$

$$\begin{aligned} v_t &= \eta_t + Lu_t + \overline{\Upsilon}_t\Big(\bigoplus_{t'<t}(u_{t'} \oplus u^0_{t'}), \bigoplus_{t'<t} v^0_{t'}\Big), \quad t < N, \\ v_N &= \eta_N + \mathcal{D}_T u_N + \overline{\Upsilon}_N\Big(\bigoplus_{t'<N}(u_{t'} \oplus u^0_{t'}), \bigoplus_{t'<N} v^0_{t'}, u_N \circ T\Big) + \\ &\quad + w_N \circ \Big(\mathrm{id} + \bigoplus_{t'<N}(u_{t'} \oplus u^0_{t'}) + u_N\Big) \circ T, \end{aligned} \tag{20$'$}$$

and for the mapping $\mathcal{J}_{T,f_{\mathcal{R}},f'_{\mathcal{R}}}$

$$z^0_t = -I^{-1} \circ \big(v^0_t - \eta^0_t - L^0 u^0_t\big) + \overline{\Upsilon}^0_t\Big(\bigoplus_{t'<t}(u_{t'} \oplus u^0_{t'}), \bigoplus_{t'<t} v^0_{t'}\Big), \quad t < N,$$

where $\overline{\Upsilon}_t$, $\overline{\Upsilon}^0_t$ $(2 \leqslant t < N)$ and $\overline{\Upsilon}_N$ are completely analogous to $\overline{\Theta}_t$, $\overline{\Theta}^0_t$ $(2 \leqslant t < N)$ and $\overline{\Theta}_N$. In view of the previous considerations for the maps $\mathcal{E}^{\leftarrow}$ and $\mathcal{I}$, now we have only to prove that the differential of $w_N \circ (\mathrm{id} + \overline{u} + u_N) \circ T$ along $u_N$, i.e., the partial derivative $(\delta/\delta u_N) w_N \circ (\mathrm{id} + \overline{u} + u_N) \circ T$, is sufficiently small. But this is easily seen from the arguments of Remark 6 related to the norm of the differential (10), since $N \geqslant 2$. Notice that a quite analogous consideration was applied to the "new" term at the end of Remark 6. Proposition 3 is proven. □

*Remark 15.* One can easily adopt the proof of Proposition 3 to show that in the general formal case, the $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalization $f$ with the identity linear part exists and is uniquely determined by its "resonant" part, $f_{\mathcal{R}}$, that can be chosen arbitrary. Indeed, to fulfill the successive steps of the formal $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalization procedure, one has to solve the first equation (20$'$) with respect to $u_t = v_t$, i.e., the linear equation $A_t u_t = \langle$terms found at the previous steps$\rangle$ where $A_t = L - \mathrm{id}$ is the operator in $G_t$. This finite-dimensional operator is non-degenerate since its spectrum, $\{\lambda^m/\lambda_s - 1 : (s, m) \notin \mathcal{R}, |m| = t\}$, does not contain the zero.

*Remark 16.* According to Remark 18 in [14], in the non-resonance case, $\mathcal{R} = \emptyset$, as a required norm in $E = E_{\overline{\mathcal{R}}}$, one can take the $C^N$-norm induced by the norm

$$\|x_1 \oplus \cdots \oplus x_p\|^* = C\varepsilon^{-1} \sum_{\tilde{s}=1}^{p} \|x_{\tilde{s}}\|_{\tilde{s}} \tag{24}$$

in $\mathbb{R}^n$, where $x_{\tilde{s}} \in M_{\tilde{s}}$, $\|\cdot\|_{\tilde{s}}$ are the restrictions of the original norm in $\mathbb{R}^n$ (which is used in condition ii)) to $M_{\tilde{s}}$, and $C > 0$ is some constant. (Recall that the $t$-th derivative of a mapping $\mathbb{R}^n \to \mathbb{R}^n$ is a polylinear symmetrical mapping $\left(\mathbb{R}^n\right)^t \to \mathbb{R}^n$ [12, 34] which allows one to define the $C^N$-norm in a coordinate-free way [34].) Moreover, one can take in the right-hand side of (24) a slightly more general expression $C\varepsilon^{-1} \sum\limits_{\tilde{s}=1}^{p} \alpha_{\tilde{s}} \|x_{\tilde{s}}\|_{\tilde{s}}$, where $\alpha_{\tilde{s}} \geqslant 1$.

*Remark 17.* The polylinear mappings $P : M_1^{\widetilde{m}_1} \times \cdots \times M_p^{\widetilde{m}_p} \to M_{\tilde{s}}$ symmetric in each cofactor $M_t^{\widetilde{m}_t}$ are identified with the linear mappings $P' : O^{\widetilde{m}_1} M_1 \otimes \cdots \otimes O^{\widetilde{m}_p} M_p \to M_{\tilde{s}}$. One can show that the induced norm, $\|\cdot\|_{\tilde{s},\widetilde{m}}$, in the space of the polylinear mappings coincides with the induced norm in the space of the above mentioned linear mappings if the linear space $O^{\widetilde{m}_1} M_1 \otimes \cdots \otimes O^{\widetilde{m}_p} M_p$ is equipped with a natural norm. The latter norm can be characterized as follows. Let $U$ and $V$ be two linear finite-dimensional spaces equipped with some norms both denoted as $\|\cdot\|$. Set $\|u \otimes v\|_1 = \|u\| \cdot \|v\|$ for arbitrary $u \in U$, $v \in V$, and

$$\|w\| = \inf \left\{ \sum_{i=1}^{g} \|u_i \otimes v_i\|_1 : \; w = \sum_{i=1}^{g} u_i \otimes v_i, \;\; u_i \in U, \, v_i \in V \right\}$$

for arbitrary $w \in U \otimes V$. Then $\|\cdot\|$ is a norm in $U \otimes V$. Incidentally, $\|u \otimes v\| = \|u \otimes v\|_1$ for all $u \in U$, $v \in V$ if the norms in both $U$ and $V$ are symmetric, i.e., the involutions $x \mapsto -x$ are metric isomorphisms. Indeed, let $B \subset U$ be the ball in $U$ centered at the origin with the radius $\|u\| > 0$ where $u \in U$. Due to the convexity of $B$, there is an affine hyperplane $u + \Pi_u$ such that $B$ lies entirely from one side of it, $\Pi_u$ being a linear subspace of codimension one. Choose an arbitrary basis $\overline{u}_1, \ldots, \overline{u}_p$ in $\Pi_u$ and denote $\overline{u}_0 = u$. Let $\Pi_v$ be the analogous subspace in $V$ associated to $v \in V$ with $\|v\| > 0$ and let $\overline{v}_1, \ldots, \overline{v}_q$ be a basis in $\Pi_v$, and $\overline{v}_0 = v$. Assume that

$$w = \sum_{i=1}^{g} u_i \otimes v_i, \;\; u_i \in U, \, v_i \in V \tag{25}$$

and denote by $u_{i,0}$ (by $v_{i,0}$, respectively) the $\overline{u}_0$-coordinate of $u_i$ in the basis $\{\overline{u}_j\}$ (the $\overline{v}_0$-coordinate of $v_i$ in the basis $\{\overline{v}_j\}$). Then

$$\|u_i\| \geqslant \|\overline{u}_0\| \cdot |u_{i,0}|, \quad \|v_i\| \geqslant \|\overline{v}_0\| \cdot |v_{i,0}|, \tag{26}$$

in view of the symmetry assumptions and, consequently, $\|u_i \otimes v_i\|_1 \geqslant \|u \otimes v\|_1 \cdot \left(|u_{i,0}| \cdot |v_{i,0}|\right)$. Assuming that $w = u \otimes v$ and expanding the both sides of (25) in the basis $\{\overline{u}_j \otimes \overline{v}_k\}$ of $U \otimes V$, one sees that $\sum\limits_{i=1}^{g} u_{i,0} \cdot u_{i,0} = 1$. Therefore, $\|u \otimes v\| \geqslant \|u \otimes v\|_1$ and, consequently, $\|u \otimes v\| = \|u \otimes v\|_1$. Now prove that in the general case $\|\cdot\|$ is a norm in $U \otimes V$. Let $\{\overline{u}_j\}$ and $\{\overline{v}_j\}$ be arbitrary two basis in $U$ and $V$, respectively. Then the inequalities (26) remain valid for all the coordinates with only some multiplier $c > 0$ in the right-hand sides: $\|u\| \geqslant c\|\overline{u}_j\| \cdot |u_j|$, $\|v\| \geqslant c\|\overline{v}_j\| \cdot |v_j|$, where $u_j$ and $v_j$ are the $\overline{u}_j$- and $\overline{v}_j$-coordinates of arbitrary $u \in U$ and $v \in V$, respectively. So, for any $w \in U \otimes V$ one obtains that $\|w\| \geqslant c^2 \|\overline{u}_j\| \cdot \|\overline{v}_k\| \cdot |w_{j,k}|$ where $w_{j,k}$ is the $\overline{u}_j \otimes \overline{v}_k$-coordinate of $w$ in the basis $\{\overline{u}_j \otimes \overline{v}_k\}$. Thus, $\|w\| > 0$ for $w \neq 0$ which guarantees that $\|\cdot\|$ is a norm in $U \otimes V$. Furthermore, for any three finite-dimensional linear normed spaces $U$, $V$, and $W$, the norm of a bilinear mapping $P : U \times V \to W$ coincides with the norm of the associated linear map $P' : U \otimes V \to W$. The inequality $\|P\| \leqslant \|P'\|$ is evident from that $P(u,v) = P'(u \otimes v)$ and $\|u \otimes v\| \leqslant \|u \otimes v\|_1$. For every $w \in U \otimes V$ and $\varepsilon > 0$ there is a representation (25) such that $\sum_i \|u_i \otimes v_i\|_1$ is $\varepsilon$-close to $\|w\|$. Then

$\|P'(w)\| = \|\sum_i P(u_i, v_i)\| \leqslant \sum_i \|P(u_i, v_i)\| \leqslant \|P\| \sum_i \|u_i \otimes v_i\|_1 \leqslant \|P\| \cdot (\|w\| + \varepsilon)$ which implies that $\|P'\| \leqslant \|P\|$. Thus, $\|P\| = \|P'\|$.

Now we are going to describe the norm looked for. The following results provide immediately this norm that is constructed inductively. For arbitrary linear finite-dimensional spaces $U$, $V$, and $W$ the natural isomorphisms of the linear spaces $U \otimes V \to V \otimes U$ and $(U \otimes V) \otimes W \to U \otimes (V \otimes W)$ are also the metric isomorphisms. This fact for the latter isomorphism is valid since

$$\|z\| = \inf \left\{ \sum_{i=1}^{g} \|u_i\| \cdot \|v_i\| \cdot \|w_i\| : \ z = \sum_{i=1}^{g} u_i \otimes v_i \otimes w_i, \ u_i \in U, v_i \in V, w_i \in W \right\}$$

for arbitrary $z \in U \otimes V \otimes W$. The norm in $O^m U$ is induced by the natural factorization $\otimes^m U \to \otimes^m U / L_m = O^m U$ [31] so that the distance between the subsets $x + L_m$, $y + L_m$ of $\otimes^m U$, which are just the elements of $O^m U$, is defined as the minimal distance between their points. Then the norm of a polylinear symmetric mapping $\otimes^m U \to V$ coincides with the norm of the associated linear map $O^m U \to V$.

*Remark 18.* Obviously, the mappings $(f_{\mathcal{R}}, f'_{\mathcal{R}}, f_{\overline{\mathcal{R}}}) \mapsto \mathcal{E}_{T, f_{\mathcal{R}}, f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$, $\mathcal{J}_{T, f_{\mathcal{R}}, f'_{\mathcal{R}}}(f_{\overline{\mathcal{R}}})$ are even of class $C^\infty$ in the collection of the variables. Moreover, they are polynomial in the sense that they can be represented as the finite Taylor series. Indeed, the elements $\mathfrak{I}_{\mathcal{R}} = \bigoplus_{t<N} w_t^0$ and $j_0^{(N-1)} f'_{\mathcal{R}} = \bigoplus_{t<N} v_t$ of the finite-dimensional spaces are polynomially dependent on the element $\overline{u} \oplus \overline{v}^0 = \big(\bigoplus_{t<N} u_t\big) \oplus f_{\mathcal{R}} \oplus f'_{\mathcal{R}}$ of the finite-dimensional space. As for the infinite-dimensional component of $f'_{\overline{\mathcal{R}}}$, $v_N$, the latter is seen to be a finite sum of terms polynomial in $u_N$ with coefficients polynomial in $\overline{u} \oplus \overline{v}^0$.

*Remark 19.* The equalities

$$\mathcal{E}_{T_1 \circ T_2, f_{\mathcal{R}}, f''_{\mathcal{R}}} = \mathcal{E}_{T_2, f'_{\mathcal{R}}, f''_{\mathcal{R}}} \circ \mathcal{E}_{T_1, f_{\mathcal{R}}, f'_{\mathcal{R}}}, \qquad \mathcal{J}_{T_1 \circ T_2, f_{\mathcal{R}}, f''_{\mathcal{R}}} = \mathcal{J}_{T_1, f'_{\mathcal{R}}, f''_{\mathcal{R}}} \circ \mathcal{J}_{T_2, f_{\mathcal{R}}, f'_{\mathcal{R}}} \tag{27}$$

are satisfied only if the set $\mathcal{R}$ is stable. Thus, in contrast to the non-resonant case $\mathcal{R} = \emptyset$ dealt with in [14], the correspondences $T \mapsto \mathcal{E}_{T, f_{\mathcal{R}}, f'_{\mathcal{R}}}$ and $T \mapsto \mathcal{J}_{T, f_{\mathcal{R}}, f'_{\mathcal{R}}}$ are not functors in any proper sense if $\mathcal{R}$ is not a stable set. Indeed, due to Remark 2, the second equality (27) and, consequently, the first one are always satisfied only if $\mathcal{R}$ is a stable set.

*Remark 20.* For an analogous reason, in addition to the previous Remark, the $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal forms with stable sets $\widetilde{\mathcal{R}}$ are the most adequate generalization of the usual (polynomial or formal) normal forms. Indeed, recall the following two facts of the classical theory (see [8] for the formal case and Historical comment 2 below). Firstly, if $f$ is a normal form with a non-degenerate linear part, then the inverse map, $f^{-1}$, is also a normal form, as it was mentioned in Remark 6. Secondly, assume that the linear parts of two normal forms possess the coincident spectra and the coincident generalized eigensubspaces associated to the equal eigenvalues. Then any transformation conjugating these normal forms will be a quasi-normal form with respect to their linear parts. Inversely, if $F$ is a quasi-normal form with respect to the linear part of a normal form and $dF(0)$ is non-degenerate then $F$ transforms this normal form into some normal form. These results remain valid if the set $\widetilde{\mathcal{R}}$ is stable and $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal forms and $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal forms are considered instead of the usual normal forms, $f$, and the quasi-normal forms with respect to $df(0)$.

Here one can consider any one of the following two cases. Firstly, the general formal case, in particular, the usual formal normal forms. As for the usual normal forms, we notice that the proof of the second fact (on the structure of the mapping conjugating two normal forms) which is presented below is more clear than the analytic approach utilized in [8]. Another simple geometric proof will be given in Historical comment 2. Secondly, the case of contractive $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal forms where one consider only $C^N$-smooth conjugacies with $N > D$, the number $D$ being determined by (2) via the spectrum of the linear part of these two $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal forms. B. Anderson [5] claimed this result for the particular case where self-conjugacies of the usual normal form are considered. However, his proof contains some inaccuracies and mistakes as will be discussed in Historical comment 2.

The key observation to prove these results is that $\mathcal{R}$ takes a "triangular form with the linear diagonal part" described in the proof of fact **2)** while proving Proposition 2. Indeed, the stability of $\mathcal{R}$ implies the absence of non-linear circuits which, in turn, implies that $\mathcal{R}$ takes this desired form (see a remark in the proof of fact **2)** above). Then, firstly, completely analogous to the usual normal forms, this form of $\mathcal{R}$ implies that $f^{-1} \in \mathcal{F}_{\widetilde{\mathcal{R}}}$ when $f \in \mathcal{F}_{\widetilde{\mathcal{R}}}$ and $df(0)$ is nondegenerate. Secondly, if $F \in \mathcal{F}_{\widetilde{\mathcal{R}}}$ and $dF(0)$ is non-degenerate then $F$ conjugates any $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal form with a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal form due to the stability of $\widetilde{\mathcal{R}}$. Inversely, the mapping $F$ that conjugates two $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal forms when being considered as a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normalization (reducing a given $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal form to an arbitrary $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal form) is uniquely determined by its $\mathcal{R}$-component, $F_{\mathcal{R}}$, that can be arbitrary. Therefore, $F_{\overline{\mathcal{R}}} = 0$ since this case is suitable.

Now we prove that the second fact can be generalized to the semi-conjugacies and, thus, the following result holds.

**Theorem 1.** *Let $\mathfrak{J}_1$ and $\mathfrak{J}_2$ be two $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-normal forms with coincident spectra of their linear parts and let the set $\widetilde{\mathcal{R}}$ (containing exact resonances) be stable. Then any mapping (maybe, non-invertible) $f$ such that $f \circ \mathfrak{J}_1 = \mathfrak{J}_2 \circ f$ is a $(\{M_{\tilde{s}}\},\widetilde{\mathcal{R}})$-quasi-normal form provided that $f$ is treated as a formal power series, in the general formal case, or as a $C^N$-mapping, in the smooth contractive case (where $\mathfrak{J}_1$ and $\mathfrak{J}_2$ are contractive, $N > D$ with $D$ being determined by (2) via the coincident spectra of $d\mathfrak{J}_i(0)$, $i = 1, 2$, and $|m| < N$ for all $(s,m) \in \mathcal{R}$).*

*Historical comment 2.* Now discuss in more detail the results by A.D. Bruno [8] and B. Anderson [5] related to conjugacies of the normal forms and mentioned in Remark 20.

Let a contraction $C^N$-mapping $f$ of a vicinity of the fixed point be written in the (real) normal form and let number $N > D$ where $D$ is determined by formula (2) via the spectrum of the linear part of $f$. According to B. Anderson [5], the $C^N$-centralizer $Z^N(f)$, i.e., the group of the $C^N$-diffeomorphisms commuting with $f$, contains only polynomials of degrees $< N$. This result is an immediate generalization of the result by N. Kopell [26] who treated the case of linear $f$. In fact, B. Anderson established that any element of $Z^N(f)$ is a quasi-normal form with respect to $df(0)$ in the sense of our Definition 8. His proof is essentially based on the consideration of auxiliary linear operators as follows. For any $C^k$-map $f : \mathbb{R}^n \to \mathbb{R}^n$ vanishing at the origin, the correspondence $g \mapsto g \circ f$ induces a linear operator in the space $J_0^k$ of the $k$-jets of mappings $g : \mathbb{R}^n \to \mathbb{R}$ at the origin. Denote by $\rho_k(f)$ the conjugate operator in $(J_0^k)^*$. If $e = (e_1, \dots, e_n)$ is a basis in $\mathbb{C}^n$ then $\{e^m\}$ is a basis in $(J_0^k)^*$ where the multiindex $m = (m_1, \dots, m_n)$ ranges over the values with non-negative integer components such that $|m| \leqslant k$ and $e^m$ denotes the corresponding symmetric product of the vectors. Now assume that $f$ is reduced to the complex normal form. B. Anderson has associated the normal form of $f$ in the basis $e$ with the upper-triangular forms of the operator $\rho_k(f)$ in the basis $\{e^m : |m| \leqslant k\}$ and has also pointed out that, in this case, the operator $\rho_k(f)$ possesses the generalized eigensubspaces $V_\Lambda^k$ associated with the eigenvalues $\Lambda$ and spanned by all the monomials $e^m$ such that $\lambda^m = \Lambda$ and $|m| \leqslant k$. The subsequent proof by B. Anderson consists of three points. Firstly, if $f$ and

$g$ commute then the same is valid for the corresponding linear operators $\rho_k(f)$ and $\rho_k(g)$ since the correspondence $h \mapsto \rho_k(h)$ is a contravariant functor. The crucial observation is that if two linear operators $F$ and $G$ commute then the generalized eigensubspaces of $F$ associated to distinct eigenvalues are also $G$-invariant. Therefore, the generalized eigensubspaces $V_\Lambda^{(N-1)}$ of the operator $\rho_{N-1}(f)$ are also $\rho_{N-1}(g)$-invariant, which means exactly that $g_{N-1}$ is a quasi-normal form with respect to $df(0)$ in the sense of our Definition 8. Secondly, if the normal form $f$ and a $C^{N-1}$-mapping $g$ commute then $f$ and $g_{N-1} = j_0^{N-1}g$ commute also where the $(N-1)$-jet is considered as a polynomial mapping. (B. Anderson omits an explanation of this fact.) Indeed, $\rho_{N-1}(h_1) = \rho_{N-1}(f \circ g) = \rho_{N-1}(g \circ f) = \rho_{N-1}(h_2)$ where $h_1 = f \circ g_{N-1}$ and $h_2 = g_{N-1} \circ f$. Note that the polynomial mappings $h_1$ and $h_2$ as being the compositions of the quasi-normal forms with respect to $df(0)$ are also the quasi-normal forms. Therefore, degree of $h_i$ does not exceed $(N-1)$ and, consequently, the operator $\rho_{N-1}(h_i)$ determines uniquely $h_i$. Thus, $h_1 = h_2$. Thirdly, if the normal form $f$ and a $C^N$-diffeomorphism $g$ commute then $g = g_{N-1}$. Indeed, it suffices to consider a $C^N$-diffeomorphism $h = g \circ g_{N-1}^{-1}$ that commutes with $f$ and has tangency of order $(N-1)$ with the identity mapping, id, at the origin and to show that $h = \mathrm{id}$. For this purpose, B. Anderson uses a Lemma of his Section 3 which is wrong. Indeed, for instance, it is easily seen that inequality (2) in this Lemma is not satisfied when $x \neq y = 0$ and natural $k$ is large enough, due to the following fact. If $F$ is a contraction map in the domain of $\mathbb{R}^n$ with a fixed point at the origin then the rate of the convergence of $\|F^m(x)\|$, $m \to +\infty$, to zero is slower than that of $\nu^m$, with $\nu > 0$ being any number less than the moduli of the eigenvalues of $F$ at the origin. For an analogous reason, inequality (1) in that Lemma seems to be far from optimal for large $k$. Fortunately, the desired result is immediately provided by the contraction property of the map $\widetilde{\mathcal{E}}_{f,f}$ discussed in Remark 6. Recall also in this connection that the proof by N. Kopell is almost identical to that for Sternberg's theorem in the non-resonant case which uses the contractive affine map $\overline{\mathcal{D}}_f = \widetilde{\mathcal{E}}_{f,f}$ (here, the first point of the above proof is superfluous, while the second point is evident).

Incidentally, the result under discussion is immediately carried out to the case of the $C^N$-mappings which commute with $f$ but whose linear parts at the origin may be degenerate. In order to show this, it suffices to apply a map $(\cdot) \mapsto f^{-1} \circ \big(g_{N-1} + (\cdot)\big) \circ f - g_{N-1}$ in $G_*$ whose contraction property is established in a completely analogous way (see the last formula (17) and the consequent analysis of the nonlinear term $\Theta_N$ in the proof of Proposition 3).

Thus, the proof by B. Anderson provides that any element of $Z^N(f)$ (i.e., a coordinate $C^N$-transformation conjugating the normal form $f$ with itself) is a quasi-normal form with respect to $df(0)$. The same remains valid for the transformations $g$ conjugating two normal forms $f_1$ and $f_2$, i.e., such that $g \circ f_1 = f_2 \circ g$, if $df_1(0) = df_2(0)$ (or, more generally, $df_1(0)$ and $df_2(0)$ possess the coincident spectra and the coincident generalized eigensubspaces associated to the equal eigenvalues). The formal part of this result (i.e., one related to the formal Maclaurin series) is provided by the first point in B. Anderson's proof without any contraction assumptions on $df(0)$. This is just the general classic result (see [8]) where, in addition, one sees that $df_1(0) = df_2(0)$ or $dg(0)$ may be degenerate (thus, $g$ may be a semi-conjugacy). We notice that the proof presented here is purely geometric in contrast to the analytic approach utilized in [8]. Returning to the above-discussed result on the (semi)conjugacy, notice that its "analytic" part is provided by the second point in B. Anderson's proof and by the third point via considering a map $(\cdot) \mapsto f_2^{-1} \circ \big(g_{N-1} + (\cdot)\big) \circ f_1 - g_{N-1}$ in $G_*$. The necessary minor modifications will be explained in the proof of Theorem 1. We emphasize that the linear part of $g$ is not assumed to be non-degenerate, i.e., $g$ may be a semi-conjugacy.

*Proof of Theorem 1.* Applying the basic Anderson's observation to the linear parts of $\mathfrak{J}_1$, $\mathfrak{J}_2$, and $f$, we see that all the subspaces $M_{\mathfrak{s}}$ are $df(0)$-invariant. Now, given $f_{\mathcal{R}}$, we will seek for $f_{\overline{\mathcal{R}}}$

as a unique fixed point of the functional map $\mathcal{E} : (\cdot) \mapsto \mathfrak{J}_2^{-1} \circ \big(f_{\mathcal{R}} + (\cdot)\big) \circ \mathfrak{J}_1 - f_{\mathcal{R}}$ in $E_{\overline{\mathcal{R}}}$. The arguments are completely analogous to those presented in Remark 6. Similarly to (17), the map $\mathcal{E}$ takes the triangular form

$$v_t = \eta_t + Lu_t + \Theta_t\big(\bigoplus_{t'<t}(u_{t'} \oplus u^0_{t'})\big),$$

and, in the $C^N$-smooth contractive case,

$$v_N = \eta_N + \mathcal{D}u_N + \Theta_N\big(\bigoplus_{t'<N}(u_{t'} \oplus u^0_{t'}), u_N \circ \mathfrak{J}_1\big).$$

Here $L(f) = J_2^{-1} \circ f \circ J_1$ and $\mathcal{D}(f) = J_2^{-1} \circ f \circ \mathfrak{J}_1$ where $J_i = d\mathfrak{J}_i(0)$. Some difference is that now the linear operators $J_1$, $J_2$ with coincident spectra and coincident generalized eigensubspaces can be distinct. Nevertheless, the arguments of Remarks 6 and 15 remain valid. We explain only the minor modifications required to establish the contraction property of $\mathcal{D}$. There are two norms $\|\cdot\|_i$, $i = 1, 2$, in $\mathbb{R}^n$ which induce operator norms denoted again as $\|\cdot\|_i$ such that $\|J_1\|_1 < 1$ and $\mu = \|J_2^{-1}\|_2 \|J_1\|_1^N < 1$. Then $U \subset \mathbb{R}^n$ will be chosen as a small enough ball in the norm $\|\cdot\|_1$ centered at the origin. To determine the norm $\|\cdot\|_*$ in $G_*$, we consider the $N$-th differential as a normed polylinear mapping $\underbrace{(\mathbb{R}^n, \|\cdot\|_1) \times \cdots \times (\mathbb{R}^n, \|\cdot\|_1)}_{N} \to (\mathbb{R}^n, \|\cdot\|_2)$ symmetrical in each cofactor. Then all the estimates of Remark 6 remain valid.

Thus, the map $\mathcal{E}$ possesses a unique fixed point $f_{\overline{\mathcal{R}}} \in E_{\overline{\mathcal{R}}}$ and, obviously, $f_{\overline{\mathcal{R}}} = 0$, since this case is suitable, due to the stability of $\widetilde{\mathcal{R}}$. $\square$

## 5. Normalization and semi-centralizer theorems for smooth contractive extensions

The following three Theorems generalize results of [20]. In fact, we establish the non-autonomous versions of the two above-mentioned facts. These facts are the existence of a smooth $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal form for the smooth contractive map (the Normal Form Theorem), see Remark 15, and the result that any semi-conjugacy of two $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal forms with a stable set $\widetilde{\mathcal{R}}$ is a $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal form (two Semi-Centralizer Theorems), see Theorem 1. Then the conditions of the Normal Form Theorem will be modified to avoid the usage of a particular Lyapunov norm.

Let $\pi \colon X \to B$ be a finite-dimensional real or complex vector bundle with compact base $B$ and let $f \colon B \to B$, $F \colon X \to X$ be some homeomorphisms (or continuous maps, respectively). Recall some definitions. Homeomorphisms (continuous maps, respectively) $f$ and $F$ generate dynamical systems with discrete time $n \in \mathbb{Z}$ ($n \in \mathbb{Z}_+$), i.e., cascades (semi-cascades) $\{f^n\}$ and $\{F^n\}$. The system $\{F^n\}$ is called an *extension* of $\{f^n\}$ if $F^n$ "covers" $f^n$, i.e. $\pi \circ F = f \circ \pi$. We denote a fiber over $b \in B$ by $X(b) = \pi^{-1}(b)$. The extension $\{F^n\}$ is said to be *linear* if the restriction of $F$ to any fiber $X(b)$ is linear. The linear extension of a homeomorphism by non-degenerate linear fiber mappings is called a *linear automorphism* of the bundle. An arbitrary linear extension is called a *linear endomorphism*. One can identify the mappings $f$, $F$ and the systems $\{f^n\}$, $\{F^n\}$ generated by them. Therefore, $F$ can be called an extension of $f$. A *norm* in the vector bundle $X \to B$ is a continuous function $X \to \mathbb{R}$ that is a norm on each fiber $X(b)$. A continuous function $\|\cdot\|_* \colon X \to \mathbb{R}$ is called a *quasinorm* if the following conditions are satisfied: **1)** $\|x\|_* \geqslant 0$ for any $x \in X$ and $\|x\|_* = 0$ if and only if $x$ is the zero element of a fiber, **2)** $\|kx\|_* = |k| \cdot \|x\|_*$ for any element $x \in X$ and real number $k$. In contrast to a norm, one does not require here the validity of the triangle inequality $\|x + y\|_* \leqslant \|x\|_* + \|y\|_*$, $\pi(x) = \pi(y)$. Due to the compactness of the base $B$ and the finite-dimensionality of the bundle, any two quasinorms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent in the sense that inequalities $c^{-1}\|x\|_2 \leqslant \|x\|_1 \leqslant c\|x\|_2$, $x \in X$, hold for some $c > 0$.

Let $J : X \to X$ be a linear extension of homeomorphism $f : B \to B$ by non-degenerate linear mappings $J(b)$ (i.e., $J : X \to X$ is a linear automorphism) and let $X = X_1 \oplus X_2$ be the decomposition of the bundle $X$ into the Whitney sum of $J$-invariant subbundles. The subbundles $X_1$, $X_2$ (up to their permutation) are said to be *exponentially separated* if there are numbers $C > 0$ and $\nu \in (0, 1)$ such that

$$\frac{\|J^n(x_1)\|}{\|x_1\|} : \frac{\|J^n(x_2)\|}{\|x_2\|} \leqslant C\nu^n \tag{28}$$

for any $n \geqslant 0$ and $x_1 \in X_1$, $x_2 \in X_2$, $\pi(x_1) = \pi(x_2)$. This means that the vectors in the fibers of $X_2$ are "stretched" under iterations of $J$ faster, than those in the fibers of $X_1$. In this case we will write $X_1 \prec X_2$.

The subbundles $X_1$, $X_2$ are said to be *spectrally separated* if there are positive numbers $C$, $g_1^+$, $g_2^-$ such that $g_1^+ < g_2^-$ and

$$\frac{\|J^n(x_1)\|}{\|x_1\|} \leqslant C(g_1^+)^n, \quad \frac{\|J^n(x_2)\|}{\|x_2\|} \geqslant C^{-1}(g_2^-)^n \tag{28$'$}$$

for any $n \geqslant 0$ and $x_1 \in X_1$, $x_2 \in X_2$, $\pi(x_1) = \pi(x_2)$. This means that the growth rates for the vectors in the fibers of $X_1$ and $X_2$ under iterations of $J$ are uniformly separated. Simple examples show that the exponentially separated subbundles can not be spectrally separated. Both the definitions do not depend on the choice of the norm or quasinorm $\|\cdot\|$ because this choice affects $d$ only.

**Proposition 4.** *Let $J : X \to X$ be a linear automorphism and $X = X_1 \oplus \cdots \oplus X_p$ be the decomposition of the bundle $X$ into the direct sum of $J$-invariant subbundles.*

a) *The following two properties are equivalent:*

1) *any two vector subbundles $X_i$ and $X_j$ $(i \neq j)$ are exponentially (respectively, spectrally) separated;*
2) *the subbundles $X_s$, $1 \leqslant s \leqslant p$, can be renumbered in such a way that for any $s$, $1 \leqslant s < p$, subbundles*

$$Y_s^- = X_1 \oplus \cdots \oplus X_s, \quad Y_s^+ = X_{s+1} \oplus \cdots \oplus X_p$$

*are exponentially (respectively, spectrally) separated, so that $Y_s^- \prec Y_s^+$.*

*Moreover, estimates* (28) *(respectively, estimates* (28$'$)*) for pairs of subbundles $X_s$, $X_{s+1}$ and $Y_s^-$, $Y_s^+$ will be satisfied with the same $\nu$'s (respectively, $g_1^+$, $g_2^-$) but, maybe, with different $C$'s.*

b) *Let $X_s$, $1 \leqslant s \leqslant p$, be exponentially separated subbundles. Then the subbundles $X_s$ are uniquely determined by their dimensions, i.e., if $X_s'$, $1 \leqslant s \leqslant p$, are invariant subbundles such that $X_s' \prec X_{s+1}'$, $1 \leqslant s < p$, and $\dim X_s = \dim X_s'$ then $X_s = X_s'$.*

c) *Let $X_s$, $1 \leqslant s \leqslant p$, be spectrally separated subbundles such that $X_s \prec X_{s+1}$, $1 \leqslant s < p$, so that for $x_s \in X_s$ one has*

$$C^{-1}(g_s^-)^m \leqslant \frac{\|J^m(x_s)\|}{\|x_s\|} \leqslant C(g_s^+)^m, \quad m \geqslant 0,$$

*where $g_s^+ < g_{s+1}^-$, or, equivalently,*

$$\frac{\|J^{\pm m}(x_s)\|}{\|x_s\|} \leqslant C(g_s^\pm)^{\pm m}, \quad m \geqslant 0, \tag{29$_\pm$}$$

*for both the signs $\pm$. Then the subbundles $X_s$ are uniquely determined by estimates* $(29_\pm)$ *where $g_s^\pm$ are given and $C = C(x_s)$ may be an arbitrary constant dependent, generally speaking, on $x_s$.*

The interval $[g_s^-, g_s^+]$ will be called a *spectral interval* of the $J$-invariant subbundle $X_s$ if (29) holds for any $x_s \in X_s$.

*Proof of Proposition 4.* These results, especially item a), are rather simple and well-known. The relation $\prec$ is an *order relation* on the set $\{X_s\}$ and determines the required reordering of the elements of this set.

As for the items b) and c), it suffices to prove the desired results for the simplest case $p = 2$ and then apply it to $Y_s^-$, $Y_s^+$ for all $s$, $1 \leqslant s < p$. We present the following purely geometrical proof of item b) based on the "graph transform method" (see Remark 11 and cf., for instance, [24, points (2.11)–(2.13); 14]). Denote by $G_c\big(X(b)\big)$ the Grassmannian manifold of $c$-dimensional linear subspaces in $X(b)$, $b \in B$. Then for any $b \in B$ the non-degenerate linear operator $J(b)$ induces a diffeomorphism $G_c\big(X(b)\big) \to G_c\Big(X\big(f(b)\big)\Big)$ of the Grassmannian manifolds which will be denoted as $\mathcal{J}_c(b)$. The set $G_c(X) = \bigcup\limits_{b\in B} G_c\big(X(b)\big)$ is the total space of the bundle $G_c(X) \to B$ whose fiber is the Grassmannian manifold (the Grassmannian bundle). Then the linear automorphism $J$ of the vector bundle $X$ naturally induces the automorphism $\mathcal{J}_c$ of the Grassmannian bundle $G_c(X)$ that is the extension of $f$ by the fiber mappings $\mathcal{J}_c(b)$. Finally, for an arbitrary continuous bundle $\pi : Y \to B$ denote by $\Gamma(Y)$ the space of all the continuous sections of this bundle, i.e., the maps $\varphi : B \to Y$ such that $\pi \circ \varphi = \mathrm{id} \mid B$. The automorphism $\mathcal{J}_c$ of $G_c(X)$ induces an automorphism $\Gamma(\mathcal{J}_c)$ of $\Gamma\big(G_c(X)\big)$. Let $X = X^+ \oplus X^-$ be the decomposition into two exponentially separated subbundles of dimensions $c^+$ and $c^-$, respectively, such that $X^- \prec X^+$. Then $X^-$ (respectively, $X^+$) can be interpreted as a unique attracting (respectively, repelling) point of $\Gamma(\mathcal{J}_c)$ for $c = c^-$ $(c = c^+)$.

To prove item c), assume that $p = 2$ and $(28')$ is valid. Let $x = x_1 \oplus x_2$ where $x_i \in X_i$, $\pi(x_i) = b$. If $x_2 \neq 0$ then one easily sees that

$$\frac{\|J^m(x_s)\|}{\|x_s\|} : (g_1^+)^m \to \infty \quad \text{as} \quad m \to +\infty.$$

Thus, $(29_+)$ is not valid for $s = 1$. Analogously, if $x_1 \neq 0$ then one can replace $J$, $g_1^+$, and $g_2^-$ by $J^{-1}$, $(g_2^-)^{-1}$, and $(g_1^+)^{-1}$ with the simultaneous permutation of the subbundles $X_1$, $X_2$. As a direct consequence, we obtain that $(29_-)$ is not valid for $s = 2$. $\square$

The extension $F$ is called a *contraction* if there are numbers $d > 0$ and $\lambda \in (0,1)$ such that

$$\|F^n(x) - F^n(y)\| \leqslant d\lambda^n \|x - y\| \tag{30}$$

for any $n \geqslant 0$ and $x, y \in X$, $\pi(x) = \pi(y)$. Obviously, this definition is independent of the choice of the (quasi)norm. Analogously, one can consider extensions $F$ which are contractive over some $F$-invariant closed subset $Y \subset X$ covering the whole base $B$, i.e., such that $F(Y) \subset Y$, $\pi(Y) = B$, and (30) holds for $x, y \in Y$. The contractive extension possesses a single invariant section and this section is continuous. This result is an immediate corollary of the "graph transform" method, since the contractive extension induces a contractive map in the space of the continuous sections.

One can account in the sequel that the invariant section coincides with the zero section. Moreover, we will consider extensions which are smooth enough along the fibers. Precisely speaking, the mapping $F \mid X(b) : X(b) \to X\big(f(b)\big)$ of the fibers is of class $C^N$ and depends continuously in the $C^N$-topology on the point $b$ of the base where $N$ is some natural. Given a homeomorphism $f : B \to B$, a vector bundle $X \to B$, and a natural $N$, we denote by $E$ the space of all the extensions $F : X \to X$ of $f$ with the latter property and whose 1-jets in the fiber directions at the zero section vanishes. This is a natural generalization of the space $E$ of the $C^N$-function introduced above (for simplicity we do not change the notation). The linear structure in $E$ is naturally induced by the linear structure in the fibers.

For any contractive extension $F$, the linear extension $DF$ being the differential of $F$ in the fiber direction at the invariant zero section, is also contractive. From the other hand, given a contractive linear extension $J = DF$, there is a norm such that $d = 1$ in formula (30) for $DF$ (with an arbitrary bigger $\lambda$). This norm is called *Lyapunov* if $\lambda < 1$. Then (30) is valid for $F$ with $d = 1$ (and a bigger $\lambda$) with respect to this norm in some $F$-invariant neighbourhood of the zero section constituted by small balls in the fibers. In particular, the extension $F$ is contractive in this neighbourhood.

Let $X = X_1 \oplus \cdots \oplus X_p$ where all the (continuous) vector bundles $X_{\tilde{s}}$ are invariant with respect to the linear extension $DF$ being the differential of $F$ in the fiber direction at the zero section. Denote $J = DF$, $J(b) = J \mid X(b)$, $J_{\tilde{s}} = J \mid X_{\tilde{s}}$, $J_{\tilde{s}}(b) = J_{\tilde{s}} \mid X_{\tilde{s}}(b)$, $1 \leqslant \tilde{s} \leqslant p$, where $X_{\tilde{s}}(b) = X(b) \cap X_{\tilde{s}}$ are the fibers of $X_{\tilde{s}}$. Now we transfer the basic concepts of Definition 8 to the extensions, assuming in the sequel that the linear extension $J$ is an automorphism.

**Definition 9.** An extension $F$ will be called a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-*quasi-normal form* if all the fiber mapping $F \mid X(b)$ become $(\{M_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal forms when the linear spaces $X_{\tilde{s}}(b)$ are identified with $M_{\tilde{s}}$, respectively. This identification provides also the decomposition $E = E_{\mathcal{R}} \oplus E_{\overline{\mathcal{R}}}$ where $E_{\mathcal{R}}$ and $E_{\overline{\mathcal{R}}}$ are the spaces of all the extensions of $f$ whose fiber mappings are "resonant" and "non-resonant", respectively. A $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal form $F$ will be called a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-*normal form with respect to the norm* in the bundle $X$ if, in addition, the differential $J = DF$ at the zero section satisfies the following property: the partition $\xi(b)$ into classes $\left\{ \|J_{\tilde{s}}(b)\|, \left\|\left(J_{\tilde{s}}(b)\right)^{-1}\right\|^{-1} \right\}$ are $\widetilde{\mathcal{R}}$-right concordant for all $b \in B$ where $\|\cdot\|$ denotes the induced operator norms. Any continuous transformation of the bundle $X$ preserving the fibers and reducing the extension to a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal form with respect to some norm in $X$ will be called a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-*normalization* or $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-*normalizing transformation*.

Obviously, a linear extension $F$ is a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal form if and only if all the sub-bundles $X_{\tilde{s}}$ are $F$-invariant.

*Remark 21.* Obviously, these definitions do not depend on the choice of the identifying isomorphisms and can be easily reformulated to avoid the usage of $M_{\tilde{s}}$. The idea is to use $X_{\tilde{s}}(b)$ immediately instead of $M_{\tilde{s}}$. Thus, in the present context, $G_{\tilde{s},\widetilde{m}}$ will be the linear space of the continuous fiber mapping $X_1 \times \cdots \times X_p \to X_{\tilde{s}}$ whose restriction to each fiber, $X(b) = X_1(b) \times \cdots \times X_p(b)$, is homogeneous polynomial in $p$ variables that has degree $\widetilde{m}_t$ in the $t$-th argument ranging over $X_t(b)$. Analogously, $G_*$ will be the space of all the extensions of $f$ whose restrictions to the fibers are $C^N$-mappings depending continuously in the $C^N$-topology on the fiber and possessing the zero $(N-1)$-jets at the origin. Notice that the condition imposed on the partition $\xi(b)$ in the definition of a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal form with respect to the norm is an appropriate counterpart of condition ii) in Proposition 3. Here, the uniformity assumptions are automatically satisfied due to the compactness of the base $B$.

**Theorem 2 (Normal Form Theorem).** *Let the differential $J = DF$ of a contractive extension $F$ be a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal form with respect to some Lyapunov norm $\|\cdot\|$ and let $N > D(b)$ for all $b \in B$ where $D(b) = \ln\left(\|J^{-1}(b)\|^{-1}\right)/\ln\|J(b)\|$. Then, in a small enough neighbourhood of the zero section, there is a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalization $H$ of $F$ which is $C^N$ along the fibers. This normalization is uniquely determined by its "resonant part" $H_{\mathcal{R}}$ that can be arbitrarily chosen.*

*Proof.* The result can be immediately obtained via applying the Corollary of Proposition 3 to the fiber mappings over a particular orbit of $f : B \to B$. Here $T_i$, $l_i$, and $\mathfrak{J}_i$ are, respectively, the restrictions of $F$, $H$, and the normal form sought for to the fiber $X(b_i)$ where $\{b_i = f^i(b)\}$ is a trajectory of $f$. Then the mappings $l_i$ depend continuously on the sequences $[T_i]$ and $\big[(f_i)_{\mathcal{R}}\big]$, according to Remark 13, and, in turn, the latter sequences depend continuously on the initial point $b$ of the orbit.

Another possibility is to reformulate the proof of Proposition 3 in order to consider the whole cascade $f$ rather than its particular trajectories. Thus, given $H_{\mathcal{R}}$, we consider an extension $\mathcal{E}_{H_{\mathcal{R}}} : E_{\overline{\mathcal{R}}} \to E_{\overline{\mathcal{R}}}$ of the homeomorphism $f : B \to B$ by the hyperbolic mappings $\mathcal{E}_{H_{\mathcal{R}}}(b) = \mathcal{E}_{F|X(b), H_{\mathcal{R}}|X\left(f(b)\right), H_{\mathcal{R}}|X(b)} : E_{\overline{\mathcal{R}}}\big(f(b)\big) \to E_{\overline{\mathcal{R}}}(b)$. One could call this extension to be ***hyperbolic*** since its fiber mappings, $\mathcal{E}_{H_{\mathcal{R}}}(b)$, possess the uniform hyperbolic properties for all $b \in B$. Notice also that the fiber mapping depends continuously in the pointwise convergence topology on the point of the base. One can see that these properties guarantee the required result.

Recall that the uniformity conditions of Proposition 3 (in each of these two approaches) are automatically satisfied due to the compactness of the base $B$. $\square$

*Remark 22.* If the linear extension $DF$ is contractive then the invariant zero section of $F$ possesses some open basin of attraction. Then, if $DF$ satisfies the conditions of Theorem 2, the $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalization described therein exists and is extended over the whole basin of attraction.

*Remark 23.* In the present context, the given linear structure of the bundle $X \to B$ is not natural, since it is used only to determine the $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal form. Thus, in the spirit of Remark 14, one can describe the $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalization $H$ in a natural coordinate-free way as a fiber mapping $X \to T_{\mathbf{0}}X$ with the identity differential $DH = \mathrm{id} \mid T_{\mathbf{0}}X$ at the zero section $\mathbf{0} : B \to X$ where $T_{\mathbf{0}}X$ denotes the linear bundle constituted by the fibers $T_{\mathbf{0}(b)}X(b)$, $b \in B$. However, Theorem 2 cannot be reformulated in a completely coordinate-free way since the "resonant part" $H_{\mathcal{R}}$ does not admit such a description.

Now we state two Semi-Centralizer Theorems. The difference is that in the first Theorem, semi-conjugacies are considered which are extensions by diffeomorphisms of the fibers, while the second Theorem is related to semi-conjugacies which are arbitrary smooth extensions.

**Theorem 3.** *Let a contractive extension $F$ of a homeomorphism $f : B \to B$ be a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal form with respect to some Lyapunov norm $\|\cdot\|$ and let the number $N$ be as above. Furthermore, let $g : B \to B$ be a continuous map commuting with $f$ and $G$ be an extension of $g$ by $C^N$-diffeomorphisms of the fibers such that $G$ commutes with $F$. Then $G$ is a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal form if the set $\widetilde{\mathcal{R}}$ is stable.*

**Theorem 3′.** *Let the extension $F$, the number $N$, the homeomorphisms $f$, the continuous map $g$, and the set $\widetilde{\mathcal{R}}$ be as in Theorem 3, and let $G$ be an extension of $g$ by (maybe, non-invertible) $C^N$-mappings of the fibers such that $G$ commutes with $F$. Then the conclusion of Theorem 3 remains valid provided that (exponentially separated) subbundles $X_{\tilde{s}}$ are spectrally separated with the spectral intervals $[g_{\tilde{s}}^-, g_{\tilde{s}}^+]$ being the elements of a $\widetilde{\mathcal{R}}$-right partition.*

Notice that the assumption imposed on the subbundles $X_{\tilde{s}}$ in Theorem 3′ imply that the spectral intervals satisfy the narrow band condition

$$g_{\tilde{s}}^+ g_p^+ < g_{\tilde{s}}^- \quad \text{for all} \quad 1 \leqslant \tilde{s} \leqslant p. \tag{31}$$

The following very simple example shows that the immediate generalization of Theorem 3 to the semi-conjugacies $G$, which are extensions by non-invertible maps, will be not valid. Therefore, one has to impose an additional condition that the subbundles $X_{\tilde{s}}$ are spectrally separated.

**Example 3.** Let $B = \{b_1, b_2\}$, $f = \mathrm{id} \mid B$, and $g$ be the permutation of $b_1$ and $b_2$, the extension $F = J$ be linear, the number $p = 2$, a linear mapping $h : X_1(b_1) \to X_2(b_2)$ be a (semi)-conjugacy of $J_1(b_1) = J \mid X_1(b_1)$ and $J_2(b_2) = J \mid X_2(b_2)$. Construct the linear extension $G$ such that $G \mid X(b_1) = h \oplus \mathbf{0}_{X_2(b_1)\to X(b_2)}$ and $G \mid X(b_2) = \mathbf{0}_{X(b_2)\to X(b_1)}$ where $\mathbf{0}_{K\to L}$ denotes the zero mapping $K \to \{0\} \in L$ of the linear spaces $K$, $L$. Then $F$ and $G$ commute but $G$ is not a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal form for any set $\widetilde{\mathcal{R}}$ if $h$ is non-zero, since then the subbundle $X_1$ is not $G$-invariant.

*Proof of Theorem 3.* The condition that the partitions $\xi(b)$ are concordant means that the subbundles $X_{\tilde{s}}$ are exponentially separated. Under the hypothesis of Theorem 3, the linear automorphisms $dF$ and the linear extension $dG$ of $g$ by non-degenerate linear fiber mappings will commute. Thus, the subbundles $X_{\tilde{s}}$ are $dG$-invariant, since these subbundles are uniquely determined by their dimensions and the exponential separation condition, due to Proposition 4 (cf. Remark 8). For any particular orbit $\{b_i\}$ of $f$ we consider diagram (12) where $T_i = F \mid X(b_i)$, $\mathfrak{J}_i = F \mid X\big(g(b_i)\big)$, and $l_i = G \mid X(b_i)$. Thus, the upper and the lower arrows correspond, respectively, to the fiber mappings of $F$ over the orbit $\{b_i\}$ of $f$ and over the image of $\{b_i\}$ under $g$ which is also an orbit of $f$. Then $(l_i)_{\overline{\mathcal{R}}}$ and $\mathfrak{J}_i$ are uniquely determined by all the mappings $(l_i)_{\mathcal{R}}$ and $T_i$. The following arguments repeat those in Remark 20 and in the proof of Theorem 1. In our case one can choose $(l_i)_{\overline{\mathcal{R}}} = 0$ and $\mathfrak{J}_i = l_i \circ T_i \circ (l_{i-1})^{-1}$ for all $i$ since $(T_i)_{\overline{\mathcal{R}}} = 0$ and $\widetilde{\mathcal{R}}$ is stable. Therefore, $(l_i)_{\overline{\mathcal{R}}} = 0$ which means the desired result.

Notice also that one could utilize arguments based on the considerations of the whole cascade $f$ rather than its particular trajectories. $\square$

*Proof of Theorem 3′.* The key observation in the proof of Theorem 3 was the fact that the linear endomorphism $DG$ of the bundle $X \to B$, which commutes with $J = DF$ and whose linear fiber mappings are non-degenerate, will preserve the $J$-invariant exponentially separated subbundles $X_{\tilde{s}}$. The crucial point of the present proof is the observation that this result remains valid for any linear endomorphism $DG$ commuting with $J$ if the subbundles $X_{\tilde{s}}$ are spectrally separated. Indeed, if estimates ($29_\pm$) hold for a given vector $x_{\tilde{s}}$ then these estimates hold also for $y_{\tilde{s}} = DG(x_{\tilde{s}})$ with the same constants $g_{\tilde{s}}^{\pm}$ and with some number $C = C(y_{\tilde{s}}) > 0$. Accordingly to item c) of Proposition 4, we obtain the desired result that $x \in X_{\tilde{s}} \Rightarrow DG(x_{\tilde{s}}) \in X_{\tilde{s}}$ (cf. Remark 9). Now we complete the proof analogously to the proof of Theorem 3. $\square$

Now we are going to remove the choice of a particular Lyapunov norm from the conditions of Theorems 2, 3, and 3′. The properly modified conditions will be "asymptotic counterparts" of the original ones and indeed do not depend on the choice of a particular norm. Notice that a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-quasi-normal form $F$ is a $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normal form with respect to the norm in the bundle $X$ if and only if given any $(\tilde{s}, \widetilde{m}) \notin \widetilde{\mathcal{R}}$ one of the two conditions

$$\big\|J^n_{\tilde{s}}(b)\big\| \cdot \prod_{i=1}^{p} \Big\|\big(J^n_i(b)\big)^{-1}\Big\|^{\widetilde{m}_i} < 1 \qquad \text{or} \qquad \Big\|\big(J^n_{\tilde{s}}(b)\big)^{-1}\Big\| \cdot \prod_{i=1}^{p} \big\|J^n_i(b)\big\|^{\widetilde{m}_i} < 1$$

is satisfied for all the points $b \in B$. Here $J^n_i(b) = J_i\big(f^{n-1}(b)\big) \circ \cdots \circ J_i\big(f(b)\big) \circ J_i(b)$ denotes the fiber mapping $X(b) \to X\big(f^n(b)\big)$ of the linear extension $J^n$, $n \geqslant 0$. Consequently, $\big(J^n_i(b)\big)^{-1}$ is the fiber mapping $X\big(f^n(b)\big) \to X(b)$ of the linear extension $J^{-n}$. Notice that here and in the sequel, $J^n_i(b)$ for $n \geqslant 0$ can be also understood as the fiber mapping $X\big(f^{-n}(b)\big) \to X(b)$. This replacement requires only a minor modification of a long formula in the proof of Lemma 4.

**Proposition 5.** *Let, as above, $F : X \to X$ be an extension of a homeomorphism $f : B \to B$ constituted by the fiber $C^N$-diffeomorphisms depending continuously in the $C^N$-topology of the* *point of the base and let $X = X_1 \oplus \cdots \oplus X_p$ be the decomposition into $J$-invariant (continuous) bundles where $J = DF$ is the differential of $F$ at the invariant zero section. (Then $J$ is a linear automorphism.) Assume that the extension $F$ and the natural $N$ satisfy the following conditions for some (and, consequently, for any) norm (or quasinorm) in $X$ where $\|\cdot\|$ will denote the induced operator norms:*

1) *given any $(\tilde{s}, \widetilde{m}) \notin \widetilde{\mathcal{R}}$ one of the two conditions*

$$\lim_{n\to+\infty} \|J_{\tilde{s}}^n(b)\| \cdot \prod_{i=1}^{p} \left\|\left(J_i^n(b)\right)^{-1}\right\|^{\widetilde{m}_i} = 0 \qquad or \qquad \lim_{n\to+\infty} \left\|\left(J_{\tilde{s}}^n(b)\right)^{-1}\right\| \cdot \prod_{i=1}^{p} \|J_i^n(b)\|^{\widetilde{m}_i} = 0$$

*is valid for all the points $b \in B$,*

2) $N > \lim_{n\to+\infty} \sup \dfrac{\ln\left(\left\|\left(J^n(b)\right)^{-1}\right\|^{-1}\right)}{\ln\|J^n(b)\|}$ *for all the points $b \in B$.*

*Then $J$ and $N$ satisfy the conditions of Theorem 2 under some choice of the norm. The desired norm can be chosen to be Euclidean (in the real case) or Hermitian (in the complex case). In particular, the subbundles $X_{\tilde{s}}$ are exponentially separated so that $X_i \prec X_j$ if and only if*

$$\lim_{n\to+\infty} \left\|\left(J_j^n(b)\right)^{-1}\right\| \|(J_i^n(b)\| = 0 \quad for\ all \quad b \in B.$$

*Moreover, the condition 2) is equivalent to that*

$$\lim_{n\to+\infty} \|J^n(b)\|^N \cdot \left\|\left(J^n(b)\right)^{-1}\right\| = 0 \quad for\ all\ the\ points \quad b \in B.$$

*Finally, the subbundles $X_{\tilde{s}}$ are spectrally separated so that $X_{\tilde{s}} \prec X_{\tilde{s}+1}$ if and only if there are numbers $g_{\tilde{s}}^{\pm}$, $1 \leqslant \tilde{s} \leqslant p$, such that $g_{\tilde{s}}^{+} < g_{\tilde{s}+1}^{-}$ and*

$$\lim_{n\to+\infty} \left\|\left(J_{\tilde{s}}^n(b)\right)^{-1}\right\| (g_{\tilde{s}}^{-})^{-n} = 0 \quad and \quad \lim_{n\to+\infty} \|J_{\tilde{s}}^n(b)\| (g_{\tilde{s}}^{+})^n = 0 \quad for\ all \quad b \in B.$$

*Moreover, $[g_{\tilde{s}}^{-}, g_{\tilde{s}}^{+}]$ is the spectral interval of $X_{\tilde{s}}$.*

First of all, we generalize the results and considerations of Appendix A of [14]. Given a quasinorm $\|\cdot\|$ in the bundle $X \to B$, for any $x \in X$, $x \neq 0$, and an integer $n$ introduce the notations

$$\Phi^{(n)}(x) = \frac{\|J^n(x)\|}{\|x\|}, \quad \Phi(x) = \Phi^{(1)}(x),$$
$$g^{-}(b) = \inf\{\Phi(x) \,:\, \|x\| = 1,\, \pi(x) = b\},$$
$$g^{+}(b) = \sup\{\Phi(x) \,:\, \|x\| = 1,\, \pi(x) = b\},$$

which are almost identical to those in the Appendix. Now for any $x_i \in X_i$, $x_i \neq 0$, $1 \leqslant i \leqslant p$, $\pi(x_1) = \cdots = \pi(x_p)$, a collection of reals $\alpha = \{\alpha_i : 1 \leqslant i \leqslant p\}$, and an integer $n$ we introduce the additional notations

$$\Phi_\alpha^{(n)}(x_1 \oplus \cdots \oplus x_p) = \prod_{\tilde{s}=1}^{p} \left(\Phi_{\tilde{s}}^{(n)}(x_{\tilde{s}})\right)^{\alpha_{\tilde{s}}}, \quad \Phi_\alpha(x) = \Phi_\alpha^{(1)}(x),$$
$$g_\alpha^{\pm}(b) = \prod_{\tilde{s}=1}^{p} \left(g_{\tilde{s}}^{\pm}(b)\right)^{\alpha_{\tilde{s}}},$$

where $\Phi_{\tilde{s}}^{(n)}$, $g_{\tilde{s}}^{-}$, and $g_{\tilde{s}}^{+}$ with index $\tilde{s}$, $1 \leqslant \tilde{s} \leqslant p$, are the functions $\Phi^{(n)}$, $g^{-}$, and $g^{+}$ for the invariant bundles $X_{\tilde{s}}$. In the sequel, $\alpha = \{\alpha_i\}$ will be collections of non-negative reals. Then obviously,

$$g_{\alpha}^{-}(b) = \inf\{\Phi_{\alpha}(x) \,:\, \|x\| = 1,\, \pi(x) = b\},$$
$$g_{\alpha}^{+}(b) = \sup\{\Phi_{\alpha}(x) \,:\, \|x\| = 1,\, \pi(x) = b\},$$

which generalize the above definitions of $g_{\tilde{s}}^{-}$ and $g_{\tilde{s}}^{+}$. Here $\Phi_{\tilde{s}}^{(n)} \equiv \Phi_{\alpha}^{(n)}$ and $g_{\tilde{s}}^{\pm} \equiv g_{\alpha}^{\pm}$ for $\alpha = \mathbf{e}_{\tilde{s}}$ where $(\mathbf{e}_{\tilde{s}})_i = \delta_{i,\tilde{s}}$.

Recall that a continuous mapping $G\colon B \times \mathbb{Z} \to \mathbb{R}_{+}$ is called a *multiplicative cocycle* of the cascade $\{f^n\}$ if $G(b, n+m) = G(b,n)G\big(f^n(b), m\big)$ for any $b \in B$ and integers $n, m$. Recall also an evident fact that a cocycle is uniquely determined by setting an arbitrary continuous everywhere positive function $g(\cdot) = G(\cdot, 1)$. Thus, the functions $g_{\tilde{s}}^{\pm}$ determine multiplicative cocycles $G_{\tilde{s}}^{\pm}$ via the relation $g_{\tilde{s}}^{\pm}(\cdot) = G_{\tilde{s}}^{\pm}(\cdot, 1)$. Then

$$G_{\tilde{s}}^{-}(b,n) \leqslant \Phi_{\tilde{s}}^{(n)}(x) \leqslant G_{\tilde{s}}^{+}(b,n), \quad n \geqslant 0, \quad x \in X_{\tilde{s}}(b),$$

and, consequently,

$$G_{\alpha}^{-}(b,n) \leqslant \Phi_{\alpha}^{(n)}(x) \leqslant G_{\alpha}^{+}(b,n), \quad n \geqslant 0, \quad x \in X(b),$$

where the multiplicative cocycle $G_{\alpha}^{\pm}(b,n) = \prod_{\tilde{s}=1}^{p}\big(G_{\tilde{s}}^{\pm}(b,n)\big)^{\alpha_{\tilde{s}}}$ is just associated with $g_{\alpha}^{\pm}(\cdot) = G_{\alpha}^{\pm}(\cdot, 1)$.

Let

$$h^{(n)}(\alpha^{+}, \alpha^{-}, b) = \prod_{\tilde{s}=1}^{p} \big\|J_{\tilde{s}}^{n}(b)\big\|^{\alpha_{\tilde{s}}^{+}} \Big\|\big(J_{\tilde{s}}^{n}(b)\big)^{-1}\Big\|^{\alpha_{\tilde{s}}^{-}},$$

where $\alpha^{+} = \{\alpha_i^{+}\}$, $\alpha^{-} = \{\alpha_i^{-}\}$. Then, obviously,

$$\max_{\substack{x \in X(b) \\ x' \in X(b)}} \Phi_{\alpha^{+}}^{(n)}(x) : \Phi_{\alpha^{-}}^{(n)}(x') = h^{(n)}(\alpha^{+}, \alpha^{-}, b)$$

and the condition that

$$h^{(n)}(\alpha^{+}, \alpha^{-}, b) \leqslant C\lambda^{n} \tag{32}$$

for some numbers $C > 0$ and $\lambda \in (0,1)$ and any $n \geqslant 0$ is an immediate generalization of the exponential separation condition that corresponds to the case $\|\alpha^{+}\| = 1$ and $\|\alpha^{-}\| = 1$. We will write $\alpha^{+} \prec \alpha^{-}$ if (32) is valid. This definition does not depend on the choice of the norm or quasinorm $\|\cdot\|$ because this choice affects $C$ only. The relation $\prec$ is a *partial order relation* on the set of all the collections $\alpha = \{\alpha_i : 1 \leqslant i \leqslant p\}$.

The following Lemma 3 is identical to Lemma A2 in the Appendix A of [14]. The Proposition 6 and Lemma 2 are also immediate generalizations of Proposition A2 and Lemma A1 in that Appendix. Their proofs are almost identical to those presented therein.

**Proposition 6.** *Let a finite collection of conditions* (32) *be met with some constants $\lambda = \lambda_j \in (0,1)$. Furthermore, let numbers $\mu_j$ lie in the intervals $(\lambda_j, 1)$. Then there is a norm $\|\cdot\|$ in $X$ such that all of these conditions* (32) *are valid with constants $C = 1$ and $\lambda = \mu_j$, respectively. The desired norm $\|\cdot\|$ can be chosen to be Euclidean (in the real case) or Hermitian (in the complex case).*

So, by slightly increasing $\lambda_j$, one is able to put all the constants $C$'s to be equal to 1.

**Lemma 2.** *Let the conditions of Proposition* 6 *be satisfied, so that a finite collection of conditions* $h^{(n)}(\beta^{j,+},\beta^{j,-},b)\leqslant C\lambda_j^n$ *is valid where* $\beta^{j,\pm}$ *denote some* $p$*-dimensional vectors* $\alpha=\{\alpha_i : 1\leqslant i\leqslant p\}$ *with non-negative entries. Then there are cocycles* $G^{\pm}_{\tilde s}$, $1\leqslant \tilde s\leqslant p$, *such that the following two properties hold:*

a) $g^{+}_{\beta^{j},+}(b) : g^{-}_{\beta^{j},-}(b)\leqslant \mu_j$, $b\in B$, *where* $g^{\pm}_{\alpha}$ *are determined via* $g^{\pm}_{\tilde s}(\cdot)=G^{\pm}_{\tilde s}(\cdot,1)$.

b) $g^{-}_{\tilde s}(b)\leqslant g^{+}_{\tilde s}(b)$ *and* $C_1^{-1}G^{-}_{\alpha}(b,n)\leqslant \Phi^{(n)}_{\alpha}(x)\leqslant C_1G^{+}_{\alpha}(b,n)$, $n\geqslant 0$, $x\in X(b)$, $b\in B$, *where constant* $C_1>0$ *depends on the choice of a quasinorm* $\|\cdot\|$.

*Proof.* Set $\|x\|_* = \left(\prod\limits_{n=0}^{N-1}\big\|J^n(x)\big\|_1\right)^{1/N}$, $\|\cdot\|_1$ being some (quasi)norm and $N$, a sufficiently large positive integer. Obviously, $\|\cdot\|_*$ is a quasinorm. However, $\|\cdot\|_*$ can be not a norm even if $\|\cdot\|_1$ is a norm (an appropriate example is easily constructed). Let $\Phi_{*,\alpha}$, and the desired functions $g^-_{\tilde s}$ and $g^+_{\tilde s}$ be the functions $\Phi_\alpha$, $g^-_{\tilde s}$, and $g^+_{\tilde s}$ determined via the constructed quasinorm $\|\cdot\|_*$. The functions $g^{\pm}_{\tilde s}$ determine the required cocycles $G^{\pm}_{\tilde s}$. The item b) of the Lemma is obviously satisfied. In order to prove item a), we recall that

$$\Phi^{(n)}_{\beta^{j},+}(x) : \Phi^{(n)}_{\beta^{j},-}(x')\leqslant C\lambda_j^n,\quad n\geqslant 0,\ \pi(x)=\pi(x'),$$

where $C>0$. One can easily check that

$$\Phi_{*,\alpha}(x)=\left(\Phi^{(N)}_{\alpha}(x)\right)^{1/N},$$

where $\Phi^{(N)}_{\alpha}$ corresponds to the original (quasi)norm (for which the assumed conditions (32) are met). Therefore

$$\Phi_{*,\beta^{j},+}(x) : \Phi_{*,\beta^{j},-}(x')=\left(\Phi^{(N)}_{\beta^{j},+}(x) : \Phi^{(N)}_{\beta^{j},-}(x')\right)^{1/N}\leqslant C^{1/N}\lambda_j\leqslant \mu_j$$

if $N>0$ is chosen to be large enough. The item a) of the Lemma follows immediately from this inequality. □

**Lemma 3.** *Let* $J : X\to X$ *be a linear automorphism and let* $G^{\pm}$ *be cocycles such that* $g^-(b)\leqslant g^+(b)$, $b\in B$, *where* $g^{\pm}(\cdot)=G^{\pm}(\cdot,1)$, *and*

$$C^{-1}G^-(b,n)\leqslant \Phi^{(n)}(x)\equiv \frac{\big\|J^n(x)\big\|}{\|x\|}\leqslant CG^+(b,n),\quad n\geqslant 0,\ x\in X(b),$$

$C>0$ *being some constant and* $\|\cdot\|$, *a quasinorm in* $X$. *Let* $\nu>1$. *Then there exists a norm* $\|\cdot\|_0$ *in* $X$ *such that*

$$\nu^{-n}G^-(b,n)\leqslant \frac{\big\|J^n(x)\big\|_0}{\|x\|_0}\leqslant \nu^nG^+(b,n),\quad n\geqslant 0,\ x\in X(b).$$

*The desired norm* $\|\cdot\|_0$ *can be chosen to be Euclidean (in the real case) or Hermitian (in the complex case).*

*Proof.* Without loss of generality, one can assume that $\|\cdot\|$ is a norm. Set

$$\|x\|_0=\sum_{n\geqslant 0}\big\|J^n(x)\big\|\big(G^+(b,n)\big)^{-1}\nu^{-n}+\sum_{n<0}\big\|J^n(x)\big\|\big(G^-(b,n)\big)^{-1}\nu^{n}.$$

Obviously, these two series converge uniformly and $\|\cdot\|_0$ is a norm. Then it is easily seen that

$$\nu^{-1} g^{-}(b) \leqslant \frac{\|J(x)\|_0}{\|x\|_0} \leqslant \nu g^{+}(b), \quad x \in X(b),$$

which implies that the norm $\|\cdot\|_0$ is the desired one. If the norm $\|\cdot\|$ was chosen to be Euclidean or Hermitian then the norm $\|\cdot\|_0$ possesses the same property. $\square$

*Proof of Proposition 6.* It follows easily from Lemmas 2 and 3. Apply Lemma 2, using some numbers $\mu_j'$ instead of numbers $\mu_j$. Then we apply Lemma 3 to each invariant subbundle $X_{\tilde{s}}$, assuming that $G^{\pm} = G^{\pm}_{\tilde{s}}$, $g^{\pm} = g^{\pm}_{\tilde{s}}$, $\nu = \nu_{\tilde{s}}$. As a result, some norms $\|\cdot\|_{\tilde{s}}$ in subbundles $X_{\tilde{s}}$ will be obtained. Finally, we require inequalities $\mu_j' \nu^{\beta_j^+ + \beta_j^-} \leqslant \mu_j$ where the notation $\nu^{\{\alpha_{\tilde{s}}\}} = \prod_{\tilde{s}} \nu_{\tilde{s}}^{\alpha_{\tilde{s}}}$ is used and construct the desired norm $\|\cdot\|$ in $X = X_1 \oplus \cdots \oplus X_p$ from the norms $\|\cdot\|_{\tilde{s}}$ in $X_{\tilde{s}}$. For instance, one can set

$$\|x_1 \oplus \cdots \oplus x_p\| = \left( \sum_{\tilde{s}=1}^{p} \|x_{\tilde{s}}\|_{\tilde{s}}^2 \right)^{1/2}, \quad x_{\tilde{s}} \in X_{\tilde{s}},\ \pi(x_{\tilde{s}}) = b. \quad \square$$

Finally, the following Lemma show that in the asymptotic conditions of Proposition 5, the decaying quantity possess an exponential upper estimate.

**Lemma 4.** *If $h^{(n)}(\alpha^+, \alpha^-, b) \to 0$ as $n \to +\infty$ for all $b \in B$ then there are $C > 0$ and $\nu \in (0,1)$ such that $h^{(n)}(\alpha^+, \alpha^-, b) \leqslant C\nu^n$ for all $n \geqslant 0$ and $b \in B$.*

*Proof.* For any $b \in B$ there is a natural $T(b)$ such that $h^{(n)}(\alpha^+, \alpha^-, b) < 1$ for $n = T(b)$. Then there is a neighbourhood $U(b)$ of $b$ such that $h^{(n)}(\alpha^+, \alpha^-, b') < 1$ for $n = T(b)$ and all $b' \in \overline{U(b)}$. Since $B$ is compact, we may choose finitely many points $b_1, \ldots, b_k$ such that $\{U(b_j) : 1 \leqslant j \leqslant k\}$ is a covering of $B$. Then there is $\nu \in (0,1)$ such that $h^{(n)}(\alpha^+, \alpha^-, b') < \nu^n$ for $n = T(b_j)$ and all $b' \in U(b_j)$, $1 \leqslant j \leqslant k$. Now let $b$ be given. Define a sequence of integers $\{i(k) : k > 0\}$ as follows. Choose $i(1)$ such that $b \in U(b_{i(1)})$. If $i(1), \ldots, i(k)$ have been chosen, let $\tau(k) = T(b_{i(1)}) + \cdots + T(b_{i(k)})$. Choose $i(k+1)$ such that $f^{\tau(k)} \in U(b_{i(k+1)})$. Now let $n > 0$ be given. Then there is $k$ such that $\tau(k) \leqslant n < \tau(k+1)$ and thus $n = \tau(k) + r$ where $0 \leqslant r < T = \max_j T(b_j)$. Therefore,

$$h^{(n)}(\alpha^+, \alpha^-, b) = \prod_{\tilde{s}=1}^{p} \|J_{\tilde{s}}^n(b)\|^{\alpha_{\tilde{s}}^+} \left\| \left(J_{\tilde{s}}(b)\right)^{-n} \right\|^{\alpha_{\tilde{s}}^-} \leqslant$$

$$\leqslant \prod_{\tilde{s}=1}^{p} \Bigg\{ \left[ \left\| J_{\tilde{s}}^r\left(f^{\tau(k)}(b)\right) \right\| \cdot \left\| J_{\tilde{s}}^{i(k)}\left(f^{\tau(k-1)}(b)\right) \right\| \cdots \left\| J_{\tilde{s}}^{i(2)}\left(f^{\tau(1)}(b)\right) \right\| \left\| J_{\tilde{s}}^{i(1)}(b) \right\| \right]^{\alpha_{\tilde{s}}^+} \times$$

$$\times \left[ \left\| \left(J_{\tilde{s}}^{i(1)}(b)\right)^{-1} \right\| \cdot \left\| \left(J_{\tilde{s}}^{i(2)}\left(f^{\tau(1)}(b)\right)\right)^{-1} \right\| \cdots \left\| \left(J_{\tilde{s}}^{i(k)}\left(f^{\tau(k-1)}(b)\right)\right)^{-1} \right\| \cdot \left\| \left(J_{\tilde{s}}^{r}\left(f^{\tau(k)}(b)\right)\right)^{-1} \right\| \right]^{\alpha_{\tilde{s}}^-} \Bigg\} =$$

$$= h^{(r)}\left(\alpha^+, \alpha^-, f^{\tau(k)}(b)\right) h^{(i(k))}\left(\alpha^+, \alpha^-, f^{\tau(k-1)}(b)\right) \cdots$$

$$\cdots h^{(i(2))}\left(\alpha^+, \alpha^-, f^{\tau(1)}(b)\right) h^{(i(1))}(\alpha^+, \alpha^-, b) \leqslant$$

$$\leqslant C\nu^n$$

where $C = \max\limits_{\substack{b \in B \\ 0 \leqslant r < T}} h^{(r)}(\alpha^+, \alpha^-, b)$. $\square$

*Proof of Proposition 5.* It follows immediately from Lemma 4 and Proposition 6. □

*Historical comment 3.* In fact, M. Gyusinsky and A. Katok [20] (see also [25]) have considered the $(\{X_{\tilde{s}}\}, \widetilde{\mathcal{R}})$-normalization with a rather special set $\widetilde{\mathcal{R}}$ (see Historical comment 2 at the end of Section 4) for the case where the subbundles $X_{\tilde{s}}$ are spectrally separated and satisfy the narrow band spectrum condition (31). For this case, they have stated also the corresponding Centralizer Theorem. As was mentioned above, we consider more general sets $\widetilde{\mathcal{R}}$. Moreover, our Theorems 2, 3, and 3′ generalize the results of these authors to the case where the subbundles $X_{\tilde{s}}$ are exponentially separated and the semi-conjugacy $G$ may be non-invertible. However, the assumption that the differential $DG$ is a linear automorphism is not essential in the considerations of [20] and the proof presented therein still works for arbitrary ($C^\infty$) smooth semi-conjugacies $G$. The idea to use the multiplicative cocycles in order to describe the growth rate of vectors in bundles is due to I. U. Bronshtein [11, 10]. He proved [11; 10, point 6.30] Proposition 6 for the simplest case of two exponentially separated subbundles (where $p = 2$ and a single relation (32) with $\alpha^+ = (1, 0)$, $\alpha^- = (0, 1)$ is presented). The generalization of his proof to a more general case (even to three exponentially separated subbundles) requires some accuracy. For this purpose, we state Lemma 3 on the existence of a norm satisfying the two-sided estimates, which generalizes the well-known result on the existence of the *adapted* or *Lyapunov* norm. The Lyapunov norm is the norm satisfying the one-sided estimate with $g^- \equiv \text{const} > 1$ or $g^+ \equiv \text{const} < 1$. Note that any cocycle $G$ and linear extension $\{F^n\}$ determine a linear extension $\{F^n_G\}$ via the formula $F^n_G(x) = G\big(n, \pi(x)\big)F^n(x)$. The possibility of constructing a norm satisfying the one-sided estimate follows immediately from the existence of the Lyapunov norm for the corresponding linear extension $\{F^n_G\}$. This idea was utilized in [10, 11] for proving Proposition 6 in the simplest case mentioned above. The case of the two-sided estimates with $g^- = \text{const}$, $g^+ = \text{const}$ is considered in [24, points 2.8–2.9; 20]. The corresponding proof generalizes the well-known construction of the Lyapunov norm. In turn, our proof is an immediate generalization of the latter proof. A simplified version of Lemma 2 which is related to the exponentially separated subbundles has been considered in [14]. The proof described therein was reproduced from [10]. For the present general case the proof remains almost identical. The Lemma 4 is a slightly generalized version of the Uniformity Lemma by N. Fenichel [17; 18, Part I] (see also [10]). The proof is reproduced from the sources cited.

## References


1. Abraham R., *Lectures of Smale on Differential Topology*, Mimeographed notes, Columbia, 1962.
2. Alekseev V.M., *Quasi-random dynamical systems.* I. *Quasi-random diffeomorphisms*, Mat. Sb. **76** (1968), no. 1, 72–134; English transl. in: Math. USSR-Sb. **5** (1968), no. 1, 73–128; II. *One-dimensional non-linear oscillations in a field with periodic perturbations*, Mat. Sb. **77** (1968), no. 4, 545–601; English transl. in: Math. USSR-Sb. **6** (1968), no. 4, 505–560; III. *Quasi-random oscillations of one-dimensional oscillators*, Mat. Sb. **78** (1969), no. 1, 3–50; English transl. in: Math. USSR-Sb. **7** (1969), no. 1, 1–43.
3. Alekseev V.M., *Final motions in the three-body problem and symbolic dynamics*, Uspekhi Mat. Nauk **36** (1981), no. 4, 161–176; English transl. in: Russian Math. Surveys **36** (1981), no. 4, 181–200.
4. Alekseev V.M., Yakobson M.V., *Symbolic dynamics and hyperbolic dynamical systems*, Appendix to R. Bowen, Methods of Symbolic Dynamics, Collected papers, "Mir", Moscow, 1979, pp. 196–240; English transl. in: Phys. Reports **75** (1981), no. 5, 287–325.
5. Anderson B., *Centralisers of $C^\infty$-diffeomorphisms*, Trans. Amer. Math. Soc. **222** (1978), 97–106.
6. Anosov D.V., *Geodesic Flows on Closed Riemannian Manifolds of Negative Curvature*, Trudy Matem. Inst. Steklova AN SSSR, Vol. 90, 1967; English transl. in: Proc. Steklov Inst. Math., Vol. 90, 1967.
7. Birkhoff G.D., *Dynamical Systems*, rev. ed., Amer. Math. Soc. Colloq. Publ., Vol. 9, Amer. Math. Soc., Providence, 1966.
8. Brjuno A.D., *Analytic form of differential equations*, Trudy MMO **25** (1971), 119–262; English transl. in: Trans. Moscow Math. Soc. **25** (1971), 131–288; *The analytic form of differential equations.* II, Trudy MMO **26** (1972), 199–238; English transl. in: Trans. Moscow Math. Soc. **26** (1972), 199–239.

9. Brjuno A.D., *Normal form of real differential equations*, Matem. Zametki **18** (1975), no. 2, 227–241; English transl. in: Math.Notes **18** (1975), no. 1–2, 722–731.
10. Bronshtein I.U., *Nonautonomous Dynamical Systems*, "Shtiintsa", Kishinëv, 1984. (Russian)
11. Bronshtein I.U., *Transversality implies structural stability*, Dokl. AN SSSR **257** (1981), no. 2, 265–268; English transl. in: Sov. Math. Dokl. **23** (1981), no. 2, 251–254.
12. Cartan H., *Calcul Différentiel*, Hermann, Paris, 1967.
13. De Latte D., *Nonstationary normal forms and cocycle invariants*, Random Comput. Dynamics **1** (1992/93), no. 2, 229–259.
14. Dovbysh S.A., *Transversal intersection of separatrices and branching of solutions as obstructions to the existence of an analytic integral in many-dimensional systems.* I. *Basic result: Separatrices of hyperbolic periodic points*, Collectanea Mathematicae **50** (1999), no. 2, 119–197.
15. Dovbysh S.A., *Intersection of separatrices and the non-integrability of multidimensional systems*, Uspekhi Mat. Nauk **55** (2000), no. 3, 179–180.
16. Eells J., *A setting for global analysis*, Bull. Amer. Math. Soc. **72** (1966), no. 5, 751–807.
17. Fenichel N., *Persistence and smoothness of invariant manifolds for flows*, Indiana Univ. Math. J. **21** (1971), no. 3, 193–226.
18. Fenichel N., *Asymptotic stability with rate conditions*, Indiana Univ. Math. J. **23** (1974), no. 12, 1109–1137; II, Indiana Univ. Math. J. **26** (1977), no. 1, 81–93.
19. Gantmacher F.R., *The Theory of Matrices*, Gostekhizdat, Moscow, 1953, 4th ed. by "Nauka", Moscow, 1988; English transl. in: Chelsea, New York, 1959.
20. Guysinsky M., Katok A., *Normal forms and invariant geometric structures for dynamical systems with invariant contracting foliations*, Math. Res. Letters **5** (1998), no. 1–2, 149–163.
21. Hadamard J., *Sur l'itération et les solutions asymptotiques des équations différentielles*, Bull. Soc. Math. France **29** (1901), 224–228.
22. Hale J. K., Lin Xiao-Biao, *Symbolic dynamics and nonlinear semiflows*, Ann. Math. Pura Appl. **144** (1986), 229–259.
23. Hartman P., *Ordinary Differential Equations*, Wiley, New York–London–Sydney, 1964, reprint 2nd ed. by Birkhäuser, Boston–Basel–Stuttgart, 1982.
24. Hirsch M.W., Pugh C.C., Shub M., *Invariant Manifolds*, Lecture Notes in Math., Vol. 583, Springer, Berlin–Heidelberg–New York, 1977.
25. Katok A., Spatzier R.J., *Differential rigidity of Anosov actions of higher rank Abelian groups and Abelian lattice actions*, Trudy Matem. Inst. Steklova RAN, Vol. 216, 1997, pp. 292–319; English transl. in: Proc. Steklov Inst. Math., Vol. 216, 1997, pp. 287–314.
26. Kopell N., *Commuting diffeomorphisms*, Global Analysis (S.-S. Chern, S. Smale, eds.), Proc. Symp. in Pure Math., Vol. 14, Amer. Math. Soc., Providence, 1970, pp. 165–184.
27. Kostin V.V., *Normal form of non-autonomous systems*, Dokl. AN USSR. Ser. A. (1973), no. 8, 693–696. (Russian)
28. Kostin V.V., Le Dinh Thuy, *Some tests for convergence of normalizing transformation*, Dopovidi AN URSR (Dokl. AN USSR). Ser. A. (1975), no. 11, 982–985. (Ukrainian)
29. Kostin V.V., *Normalizing Transformations of Nonautonomous Systems*, Preprint, Odessa Univ., 1975. (Russian)
30. Kušnirenko A.G., Katok A.B., Alekseev V.M., *Smooth dynamical systems*, Ninth Summer Mathematical School, "Naukova Dumka", Kiev, 1976, pp. 50–341; English transl.: *Three Papers on Dynamical Systems*, AMS Transl., Ser. 2. Vol. 116, 1981.
31. Lang S., *Algebra*, Addison–Wesley Series in Math., Addison–Wesley, Reading (Mass.), 1965.
32. Lerman L.M., Shil′nikov L.P., *Homoclinic structures in infinite-dimensional systems*, Sibir. Matem. Zh. **29** (1988), no. 3, 92–103; English transl. in: Siberian Math. J. **29** (1989), no. 3, 408–417.
33. Nemyckiĭ V.V., Stepanov V.V., *Qualitative Theory of Differential Equations*, Gostekhizdat, Moscow–Leningrad, 1947, 2nd ed., 1949; English transl. in: Princeton Math. Series, no. 22, Princeton Univ. Press, Princeton, 1960, reprint Dover, New York, 1989.
34. Nitecki Z., *Differentiable Dynamics. An Introduction to the Orbit Structure of Diffeomorphisms*, MIT Press, Cambridge (Mass.)–London, 1971.
35. Sibuya Y., *Non-linear ordinary differential equations with periodic coefficients*, Funkcialaj ekvacioj **1** (1958), no. 2, 121–204.
36. Smale S., *Differentiable dynamical systems*, Bull. Amer. Math. Soc. **73** (1967), no. 6, 747–817.
37. Sternberg S., *Local contractions and a theorem of Poincaré*, Amer. J. Math. **79** (1957), no. 4, 809–824.
38. Yomdin Y., *Nonautonomous linearization*, Dynamical Systems (J. C. Alexander, ed.), Lecture Notes in Math., Vol. 1342, Springer, Berlin–Heidelberg–New York, 1988, pp. 718–726.