Journal of Differential Equations **179**, 479–537 (2002)
doi:10.1006/jdeq.2001.4043, available online at http://www.idealibrary.com on IDEAL®

The Analytic and Formal Normal Form for the Nilpotent Singularity[1]

Ewa Stróżyna

Institute of Mathematics, Warsaw University of Technology, Pl. Politechniki 1, 00-661 Warsaw, Poland
E-mail: strozyna@mini.pw.edu.pl

and

Henryk Żołądek

Institute of Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warsaw, Poland
E-mail: zoladek@mimuw.edu.pl

We study orbital normal forms for analytic planar vector fields with nilpotent singularity. We show that the Takens normal form is analytic. In the case of generalized cusp we present the complete formal orbital normal form; it contains functional moduli. We interprete the coefficients of these moduli in terms of the hidden holonomy group.

Key Words: singular point of vector field; normal form; monodromy group.

1. INTRODUCTION

1.1. *Around the Takens prenormal form.* Our subjects of investigation are germs at $0 \in \mathbb{C}^2$ of analytic vector fields of the form

$$\dot{x} = 2y + \cdots, \ \dot{y} = \cdots, \tag{1.1}$$

i.e., with nilpotent linear part. (We will also write $\dot{x} = y + \cdots$.)

We will pay particular attention to the case when this system is close to the Hamiltonian system with the Hamilton function

$$H = y^2 - x^s, \qquad s \geqslant 3. \tag{1.2}$$

It means that

$$\dot{x} = 2y + \cdots, \qquad \dot{y} = sx^{s-1} + \cdots, \tag{1.3}$$

[1] Supported by Polish KBN Grant 2 P03A 022 08 and 2 P03A 041 15.

0022-0396/02 $35.00

where the dots mean the terms of higher order. Later we shall specify this notion.

A special case of the system (1.3) forms the Bogdanov–Takens singularity $\dot{x}=y$, $\dot{y}=x^2+axy+\cdots$, where $a\neq 0$ (see [B1]); here we skip the latter assumption. In the paper of Moussu [M2] such singularity (i.e., the system (1.3) with $s=3$) is called *dégénérée transverse*. In the work of Elizarov *et al.* [EISV] this class is denoted by J_*.

Some work was done on the formal orbital normal forms for the systems with nilpotent linear part. Takens [T] in 1974 proved that the system (1.1) can be formally reduced to

$$\dot{x}=y+a(x),\qquad \dot{y}=b(x) \tag{1.4}$$

where $a(x)=a_r x^r+a_{r+1}x^{r+1}+\cdots$, $b(x)=b_{s-1}x^{s-1}+b_s x^s+\cdots$ are some formal power series. Some authors use the following, equivalent to (1.4), prenormal form

$$\dot{x}=y,\qquad \dot{y}=b(x)+yc(x). \tag{1.5}$$

Since the forms (1.4) and (1.5) are not the complete normal forms, we shall call them the *Takens prenormal forms* (similarly, the systems (1.5) and (1.6) are called the *Bogdanov–Takens prenormal forms*).

In order to obtain the form (1.4) Takens uses only changes of variables x, y but not the time. Baider and Sanders [BS] have continued investigations in this direction and obtained some new results about nonorbital normal form. They still are not complete; the final (nonorbital) formal normal form is not known.

We have two possibilities: either $s=\infty$ (i.e., $b(x)\equiv 0$) or $s<\infty$. If $s=\infty$, then the system (1) has a nonisolated critical point and its phase portrait is the same as the flow-box foliation. Later we assume that $s<\infty$.

It is easy to reduce the form (1.4) to

$$\dot{x}=y+a(x),\qquad \dot{y}=sx^{s-1} \tag{1.6}$$

(see Lemma 3 in Section 3 below and [B2]). In the case $s<2r$ (i.e., the system (1.3)) an equivalent to (1.6) normal form is

$$\dot{x}=2y+2xc(x),\qquad \dot{y}=sx^{s-1}+syc(x) \tag{1.7}$$

(see [L3] and Lemma 3 in Section 3). The latter form has the property that the cusp curve $y^2=x^s$ is invariant.

In the present paper we shall show that the prenormal forms (1.4)–(1.7) can be chosen *analytic* (see Theorem 6 in the next section).

The form (1.6) is not the final orbital normal form. Some (infinitely many) coefficients in the Taylor expansion of $c(x)$ can be reduced to zero.

However, this reduction cannot be realized by means of analytic change of variables and time (see Remark 18 in Section 5.1). The analyticity property reaches its limit in the formula (1.6).

We shall reduce all the terms from $c(x)$ which can be reduced. Therefore we shall obtain the complete *formal* orbital normal form in the case $a_r b_{s-1} \neq 0$ and

$$s < 2r, \tag{1.8}$$

i.e., for the system (1.3) (see Theorem 7 in the next section). We will call this case the *generalized cusp case.*

(The case with $s > 2r$ will be called the *generalized saddle–node case* and the case with $s = 2r$ we call the *generalized saddle case.*)

The work of Bogdanov [B2] appeared with formal orbital normal forms for a large class of singularities including the nilpotent ones. His method uses some homological machinery (spectral sequences, etc.) and probably was not understood by other specialists. Moreover, our results contradict one of his theorems.

Results similar to [B2] with much simpler proofs, were obtained by Sadovski [S2]. However, in the case (1.8), he repeats the theorem of Bogdanov (with the same mistake and without proof).

Recently Loray [L3] obtained the same formal orbital normal form as the our form from Theorem 7. Our proof is different (more algebraic) than Loray's, so we present it here.

1.2. *Resolution of singularity and hidden holonomy.* During the past 20 years a lot of work has been done in the theory of analytic orbital normal forms (Ecalle–Voronin moduli, Martinet–Ramis moduli, Stokes operators, etc.). In particular, this theory was applied to the nilpotent cusp singularity, first by Moussu and Cerveau [CM, M2], next by the group of Il'yashenko *et al.* [EISV], and, then by Loray and Meziani [L1, L2, LM].

The geometrical picture is as follows. It is well known that a singular point of a vector field V from the class J_* is resolved in three elementary blowing-ups. In the general case the resolution of the vector field (1.1) satisfying (1.8) depends on the parity of the exponent s. This resolution is the same as the resolution of the Hamilton function $y^2 - x^s$ (see Fig. 1).

If $s = 2k+1$, then we need $k+2$ elementary blowing-ups. The last blowing-up gives the divisor E_{k+2} with the three singular points of the induced foliation near it (see Fig. 1):

p_0 with $1{:}-(4k+2)$ resonance and with a separatrix Γ representing an invariant analytic curve $y^2 - x^{2k+1} + \cdots = 0$,

$p_1 = E_{k+2} \cap E_k$ with $1{:}-2$ resonance and

$p_2 = E_{k+2} \cap E_{k+1}$ with $k{:}-(2k+1)$ resonance.

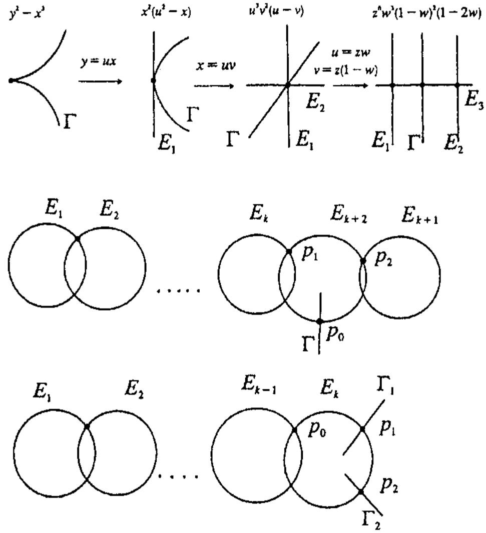

FIGURE 1

If $s = 2k$, then we need k blowing-ups. The last blowing-up gives the divisor E_k with the three singular points:

$p_0 = E_k \cap E_{k-1}$ with $(k-1){:}-k$ resonance and

$p_{1,2}$ with $1{:}-2k$ resonance and with separatrices $\Gamma_{1,2}$ representing two invariant analytic curves $y \pm x^k + \cdots = 0$.

There is a hidden holonomy group G associated with the germ of vector field V. It is the *monodromy group* (or *holonomy group*) G associated with the punctured divisor $E^* = E \setminus \{p_0, p_1, p_2\}$, where E is the last divisor appearing in the resolution. It is a subgroup of the group of germs of holomorphic diffeomorphisms of a holomorphic disk D transverse to E^* and is defined by lifts to the leaves of the corresponding foliation of loops in E^*.

This holonomy is called (by Moussu) hidden, in order to distinguish it from the holonomy maps associated with the separatrix (or separatrices). Two germs of vector fields may have the same holonomies associated with separatrices, but not be analytically equivalent. This holds in the case when G is solvable and s is odd.

The group G is also called the projective monodromy group.

In [CM] and [M2] series of results about the group G were proved. It is generated by two maps f_1 and f_2 corresponding to the simple loops around p_1 and p_2.

In the case $s=2k+1$ we have

$$f_1(z)=e^{i\pi}z+\cdots \tag{1.9}$$
$$f_2(z)=e^{2\pi ik/(2k+1)}z+\cdots$$
$$f_0(z)=f_1\circ f_2(z)=e^{-i\pi/(2k+1)}z+\cdots$$

and there are two relations

$$f_1^{[2]}=f_2^{[2k+1]}=id, \tag{1.10}$$

where $f^{[j]}=f\circ f\circ\cdots\circ f$ (j times).

In the case $s=2k$ we have

$$f_{1,2}(z)=e^{i\pi/k}z+\cdots \tag{1.11}$$
$$f_0(z)=f_1\circ f_2(z)=e^{2\pi i/k}z+\cdots$$

and there is one relation

$$f_0^{[k]}=id. \tag{1.12}$$

The relations (1.10) and (1.12) are consequences of the analytic linearizability of the corresponding singular points; that linearizability is obtained from contractibilities of some loops in certain divisors at Fig. 1: E_1 and E_{k+1}.

In what follows we shall mean by a monodromy group, a subgroup of the group of germs of analytic diffeomorphisms of $(\mathbb{C},0)$ with a distinguished system of its generators. Therefore we have $G=\langle f_1,f_2\rangle$, where $f_{1,2}$ satisfy (1.9)–(1.10) or (1.11)–(1.12).

The next theorem, which says that the pair (f_1,f_2) of germs of diffeomorphisms (modulo equivalence) constitutes a complete invariant of orbital classification of generalized cusp singularities, was proved by Moussu and Cerveau [M2, CM] (see also [EISV]). Before its formulation we give two definitions.

Definition 1. Two germs V,V' of analytic vector fields in $(\mathbb{C}^2,0)$ are *formally* (respectively *analytically*) *orbitally equivalent* iff there is a formal (respectively analytic) diffeomorphism $\mathcal{G}$ of $(\mathbb{C}^2,0)$ transforming the phase curves of V to the phase curves of V'. This means that there is a formal (respectively analytic) function ψ, $\psi(0)\neq 0$ such that $\psi\cdot V'=\mathcal{G}_*^{-1}V\circ\mathcal{G}$.

DEFINITION 2. Two groups $G = \langle f_1, \ldots, f_l \rangle$, $G' = \langle f'_1, \ldots, f'_l \rangle$ of germs of conformal diffeomorphisms of $(\mathbb{C}, 0)$ are *analytically equivalent* iff there is a germ h of analytic diffeomorphism of $(\mathbb{C}, 0)$ conjugating the corresponding maps from G and G'. Thus

$$f'_i = h \circ f_i \circ h^{-1}.$$

G and G' are *formally equivalent* if the above holds at the level of formal power series.

THEOREM 1. (a) *The monodromy group, modulo analytic equivalence, associated with the generalized cusp singularity constitutes an invariant of the orbital analytic classification for such germs of vector fields.*

(b) *Two germs with generalized cusp singularity having analytically equivalent monodromy groups are orbitally analytically equivalent.*

(c) *Each groups $G = \langle f_1, f_2 \rangle$ satisfying* (1.9)–(1.10) *or* (1.11)–(1.12) *is realized as a hidden monodromy group of a generalized cusp singularity.*

(d) *The same statements hold when we replace the orbital analytical equivalence of vector fields and the analytic equivalence of monodromy groups by the corresponding formal equivalences.*

1.3. *The theory of groups of germs of* 1*-dimensional diffeomorphisms and its application to generalized cusps.* Here we present results about classification of finitely generated groups of germs of analytic diffeomorphism of a complex line. They are standard for specialists. We present them in full generality (although we need them in a restricted form), in order to show their panorama for unexperienced readers. For details we refer the reader to the survey article by Elizarov *et al.* [EISV]).

THEOREM 2 (Formal classification of groups; [CM]). (a) *Any finitely generated* abelian *group G of germs of conformal diffeomorphisms of* $(\mathbb{C}, 0)$ *is* formally *equivalent either to a group consisting of linear maps*

$$z \to \lambda z, \qquad \lambda \in \mathbb{C}^*$$

or to a group consisting of maps of the form

$$z \to \lambda g^t_w, \qquad \lambda^p = 1,\ t \in \mathbb{C},$$

where g^t_w is the flow map generated by the vector field $w = w_{p,\mu} = [z^{p+1}/(1 + \mu z^p)]\, \partial_z$. Here the field $w_{p,\mu}$ is fixed for the whole group and at least one $t \neq 0$.

(b) *Any finitely generated* solvable nonabelian *group G of germs of conformal diffeomorphisms of* $(\mathbb{C}, 0)$ *is* formally *equivalent to a group consisting of maps of the form*

$$z \to \lambda g^t_{z^{p+1}}(z), \qquad \lambda \in \mathbb{C}^*,\ t \in \mathbb{C},$$

where $g^t_{z^{p+1}}(z) = z(1-ptz^p)^{-1/p}$ *is the flow map generated by the vector field* $w_{p,0} = z^{p+1}\partial_z$. *Here the integer p is fixed for the whole group and at least one* $t \neq 0$.

From this it follows that any abelian group defined by (1.9)–(1.10) is formally equivalent to $\langle \hat{f}_1, \hat{f}_2 \rangle = \langle -z, \lambda z \rangle$, $\lambda = e^{2\pi i k/(2k+1)}$ or to $\langle -z, \lambda g^1_w \rangle$, $w = w_{p,\mu}$, $p = (2k+1)\,m$. Any such solvable nonabelian group is formally equivalent to $\langle -z, \lambda g^1_{z^{p+1}} \rangle$.

Analogous groups defined by (1.11)–(1.12) are formally equivalent to $\langle e^{2\pi i/k}z, e^{\pi i/k}z \rangle$ (abelian); or $\langle e^{2\pi i/k}z, e^{\pi i/k}g^1_w \rangle$, $p = 2km$ (abelian); or $\langle e^{2\pi i/k}z, e^{\pi i/k}g^1_{z^{p+1}} \rangle$ (solvable nonabelian). We associate with a solvable group G two other groups: the additive group

$$T_G = \{t \colon g^t_w \in \hat{G}\}, \qquad w = w_{p,\mu} \quad \text{or} \quad w = w_{p,0}$$

and the multiplicative group (of multipliers if G)

$$\Lambda_G = \{f'(0) \colon f \in G\}.$$

Definition 3. (a) An abelian group G of conformal diffeomorphisms of $(\mathbb{C}, 0)$ is called *formally linear* iff it is formally equivalent to a group consisting of linear transformations.

If G is abelian but not formally linear, then we call it either *exceptional,* if the additive group T_G is cyclic, or *typical*, otherwise.

If G is formally linear and Λ_G is finite then we say that the group is *finite*; (it is finite in fact).

(b) A solvable nonabelian group of conformal diffeomorphisms of $(\mathbb{C}, 0)$ is called *exceptional* iff the additive group T_G is cyclic. A nonexceptional solvable group is called *typical.*

Remark 1. If G is formally linear, then it can be finite or infinite. If G is finite, then Λ_G consists of roots of unity and G is cyclic and analytically equivalent to Λ_G.

If G is abelian exceptional, then the additive group $T_G = \{t \colon g^t_w \in \hat{G}\}$ is cyclic and the multiplicative group Λ_G consists of roots of unity of order p, $\Lambda_G^p = \{1\}$.

If G is solvable nonabelian exceptional, then Λ_G is a subgroup of the group of roots of unity of degree $2p$, $\Lambda_G^p = \{\pm 1\}$.

In the solvable case with $p = 1$ the maps $z \to \lambda g_{z^2}^t = \lambda z/(1-tz)$ are conjugated with the affine automorphisms $\xi \to a\xi + b$ of the complex line. In the exceptional case the group T_G is identified with a subgroup $b_0\mathbb{Z} \subset \mathbb{C}$, which must be preserved by the group of multipliers $\Lambda_G =$ (group of a's). Thus, either $\Lambda_G = \{1\}$ and G is abelian or $\Lambda_G = \{\pm 1\}$ and G is nonabelian.

If $p > 1$, then one has only semiconjugation with a subgroup of the group of affine maps.

THEOREM 3 (Rigidity of groups; [R, CM]). *If two groups G and G' are either:*

(i) *finite, or*

(ii) *abelian typical, or*

(iii) *solvable typical, or*

(iv) *non-solvable*

and are formally equivalent, then they are also analytically equivalent.

Remark 2. In [CM] and [EISV] an analytic classification of exceptional abelian and solvable exceptional groups is given (Theorems 1.3 and 2.4 in [EISV]). In contrast to the typical solvable case they contain a functional modulus. It is the Ecalle–Voronin modulus of one of the generators of the group. In the cases of typical (abelian and solvable) groups the Ecalle–Voronin modulus turns out to be trivial; this is the reason for their rigidity.

We do not discuss here the problem of rigidity and analytic classification of groups which are formally linear, but infinite. This problem is related to the problem of small divisors (theorems of A. Briuno and J.-C. Yoccoz).

Application of the above results to the monodromy group of a generalized cusp singularity of vector fields has led the authors to the following result.

THEOREM 4 (Rigidity of generalized cusps; [CM, EISV, L1, LM]). (a) *Let $s = 2k+1$. Any group defined by* (1.9)–(1.10) *is either finite or solvable typical or non-solvable (i.e., it is rigid).*

(b) *Let $s = 2k$. Any group defined by* (1.11)–(1.12) *is either finite or abelian exceptional or solvable typical or nonsolvable.*

(c) *This implies the following (rigidity property): if two germs of analytic vector fields with generalized cusp singularity with non-exceptional monodromy groups are formally equivalent, then they are analytically equivalent (they are rigid).*

In particular, any generalized cusp with $s = 2k+1$ is rigid.

From Theorem 4 it follows that when the group is nonsolvable or solvable typical or finite and the formal orbital normal form of V is polynomial, then the conjugation transformation is always analytic.

Remark 3. Loray and Meziani [L1, LM] obtained analytical classification of germs of vector fields with solvable groups in the case of odd s. The corresponding orbital normal forms are polynomial and are similar to our formal normal form given in Theorem 7.

Remark 4. Some authors study also the topological rigidity of the groups defined by (1.9)–(1.10) (or (1.11)–(1.12)) and of generalized cusp singularities. We do not need them here and we omit formulation of the corresponding theorems.

We complete this point by presenting a criterion of solvability of the monodromy in the cases of generalized cusps of small multiplicity. We thank the referee for detecting a mistake in our previous formulation of this result (we claimed that an analogous criterion holds for any s).

THEOREM 5 (Solvability criterion). (a) *If* $s=3$, *then a group defined by* (1.9) *and* (1.10) *is solvable iff* $f_0^{[6]}=id$. *For vector field this means the following: its monodromy group is solvable iff the saddle* p_0 *of the resolved field is linearizable.*

(b) *If* $s=4$ *and* G *is defined by* (1.11) *and* (1.12), *then* $f_{1,2}^{[4]}=id$ *iff either* G *is solvable nonabelian or* G *is finite. In terms of the vector field, this property of its monodromy means that the saddles* $p_{1,2}$ *of the resolved vector field are linearizable.*

1.4. *About the results of this paper.* At first approach we hoped to obtain a polynomial formal normal form and then we planned to use the above analytic machinery to obtain results about analytical classification. However it turned out conversely; the normal form is not polynomial.

In the present paper we give the complete formal orbital classification in the case (1.8) (Theorem 7). The proof of this formal orbital form relies on solution of the homological equations with respect to the monomial components of the changes of coordinates and time. The problem is reduced to the solution of series of linear algebraic equations.

The analyticity of the conjugating diffeomorphism is obtained only for the preliminary normal form, i.e., the Bogdanov–Takens form (1.6) (Theorem 6). This is the main result of the paper. It is also the most technical part of the work.

We show analyticity directly by estimating the coefficients of the Taylor series of the changes of variables. The key point of the proof is that the essential part of the linear operator acting on vector fields in the homological equation can be decomposed into a sum of a diagonal operator and a quasinilpotent operator. Next we use a variant of Newton's method and techniques from the KAM theory.

Knowing the formal normal form we are able to recognize the type of its projective holonomy (Theorem 8). In the cases of finite and solvable nonabelian monodromy group the formal normal forms are polynomial, which means that these forms are also analytic orbital normal forms.

In the case $s=3$, we interpret the first coefficient of the Taylor expansion of the formal functional modulus in terms of the saddle quantities of the saddle p_0 of the resolved field (Theorem 9(a)). In the even s case we interpret the first terms of the formal normal form in terms of the saddle quantities of the saddles $p_{1,2}$ of the blown-up field (Theorem 9(b)).

2. RESULTS

The first of our theorems is the main result of the work. It solves the old problem of analyticity of the prenormal form obtained by Takens.

THEOREM 6. *The Bogdanov–Takens prenormal form* $\dot{x}=y+a(x)$, $\dot{y}=sx^{s-1}$ *of the singularity* $\dot{x}=y+\cdots$, $\dot{y}=sx^{s-1}\cdots$ *(with the only restriction* $2\leqslant s<\infty$*) is* analytic.

Remark 5. Let $s=2$. Then the singular point is a saddle with the $1:-1$ resonance. Theorem 6 says that the Bogdanov–Takens prenormal form $(y\partial_x+x\partial_y)+c(x)(x\partial_x+y\partial_y)$ holds and is analytic. The first nonzero coefficient in the Taylor expansion of the function $c(x)$ is the *saddle quantity* (the analogue of the focus quantity).

The further results concern only the generalized cusp singularities, i.e., the vector fields in the form (1.7). Let

$$X_H=2y\partial_x+sx^{s-1}\partial_y$$

be the Hamiltonian vector field and

$$E_H=2x\partial_x+sy\partial_y$$

be the quasihomogenous Euler vector field. We shall deal with the vector fields $X_H+c(x)\,E_H$, where $c(x)$ are germs of analytic function vanishing at $x=0$. We can assume that $c(x)=x^{r-1}+\cdots$ (we fix the first coefficient by rescaling) or that $c(x)\equiv 0$ (i.e., $r=\infty$).

We introduce also the notation

$$n_0 = \frac{r}{s} - \frac{1}{2}. \tag{2.1}$$

THEOREM 7. *Any analytic germ* $X_H + c(x)\,E_H$ *is* formally *orbitally equivalent to one of the normal forms* $J^s_{r,\phi}$, *indexed by exponents*

$$r \neq 0 \pmod s \qquad \text{or} \qquad r = \infty$$

and by formal power series

$$\phi = \phi(x) = \sum{}' c_j x^j$$

(where the summation runs over some subset of integers), defined below:

(i) $J^s_{\infty,0}$: X_H;

(ii)

$$J^s_{r,\phi}\colon\ X_H + x^{r-1}(1+\phi(x))\,E_H \tag{2.2}$$

with

$$\sum{}' = \sum_{j \neq 0,\ -r \pmod s}$$

if $r < \infty$ *and* $n_0 \notin \mathbb{Z}$;

(iii) *the field* (2.2) *with*

$$\phi = c_{n_0 s} x^{n_0 s}$$

if $r < \infty$ *and* $n_0 \in \mathbb{Z}$;

(iv) *the field* (2.2) *with*

$$\sum{}' = \sum_{j \in \{n_0 s,\ j_0\}} + \sum_{\substack{j > j_0 \\ j \neq 0,\ -r \pmod s \\ j \neq j_0 + n_0 s}}$$

if $r < \infty$, $n_0 \in \mathbb{Z}$ *and there exists a nonzero coefficient* c_{j_0} *with* $j_0 \neq 0, -r \pmod s$ *(we mean that* $c_{j_0} \neq 0$ *is the first such coefficient).*

If two vector fields with the normal forms $J^s_{r,\phi}$ *and* $J^{s'}_{r',\phi'}$ *are formally orbitally equivalent, then* $r = r'$, $s = s'$ *and* $\phi'(x) \equiv \phi(\alpha x)$ *for some constant* α *which satisfies* $\alpha^{2r-s} = 1$ *(when* $r \neq \infty$*).*

Remark 6. The normal forms from Theorem 7 need explanations. We explain them on the examples with $s = 3$ and $s = 4$.

If $s=3$, then we have two series of normal forms (with one functional modulus)

$$X_H + x^{3m}(1+x\psi(x^3))\, E_H,$$
$$X_H + x^{3m+1}(1+x^2\psi(x^3))\, E_H$$

and one exceptional (of infinite codimension) singularity X_H.

If $s=4$, then r can take the values $1, 2, 3 \pmod 4$. If $r=1 \pmod 4$ or $r=3 \pmod 4$, then $n_0 \notin \mathbb{Z}$ and the normal form has two functional moduli. This gives two series of normal forms (ii)

$$X_H + x^{4m}(1+x\psi_1(x^4)+x^2\psi_2(x^4))\, E_H,$$
$$X_H + x^{4m+2}(1+x^2\psi_1(x^4)+x^3\psi_2(x^4))\, E_H.$$

If $r=2 \pmod 4$, then $n_0 \in \mathbb{Z}$, $r=4n_0+2$, and the term $c_{4n_0}x^{4n_0}$ cannot be eliminated from $\phi(x)$. We have two possibilities: either the whole $\phi(x)$ equals $c_{4n_0}x^{4n_0}$ (iii) or it contains some nonzero term $c_{j_0}x^{j_0}$, $j_0 \neq 0, -r \pmod 4$ (iv). We assume that j_0 is the minimal index with this property.

In case (iii) we get one series (without functional moduli)

$$X_H + x^{4n_0+1}(1+\mu x^{4n_0})\, E_H.$$

In case (iv) we have two series, (indexed by $r=4n_0+2$ and by $j_0=4m+1$, or by $j_0=4m+3$ and containing two functional moduli):

$$X_H + x^{4n_0+1}[1+c_{4n_0}x^{4n_0}+x^{4m+1}\psi_1(x^4)+x^{4m+3}\psi_2(x^4)]\, E_H,$$
$$X_H + x^{4n_0+1}[1+c_{4n_0}x^{4n_0}+x^{4m+3}\psi_1(x^4)+x^{4m+5}\psi_2(x^4)]\, E_H,$$

where $\psi_1(0) \neq 0 = \psi_1^{(n_0)}(0)$. This means that the monomial $x^{4n_0+j_0}$ is absent in $\phi(x)$.

There remains the exceptional form (of infinite codimension) X_H.

Note also that the change $(x, y) \to (\alpha^2 x, \alpha^s y)$ induces the changes $X_H \to \alpha^{s-2} X_H$ and $x^{r-1}E_H \to \alpha^{2r-2} \cdot x^{r-1}E_H$. Therefore when $\alpha^{2r-s}=1$, this change does not influence the orbital type of the rth jet of the field $J^s_{r,\phi}$. This means that the normal forms are unique modulo action of the finite group $\mathbb{Z}/(2r-s)\,\mathbb{Z}$. We thank the referee for insisting on this point.

The reason for the elimination of $x^{n_0 s+j_0}$ is the following. All eliminations (proved in Theorem 7) are made using linear homological equation with respect to change of variables. This means that, in each step of the reduction, we solve a system of linear algebraic equations. If n_0 is not an integer, then the corresponding system always has a unique solution.

If n_0 is an integer, then, in a certain step of the reduction, the determinant of the corresponding linear matrix vanishes. There is a term which cannot be eliminated (it is $x^{n_0 s}$ in $\phi(x)$) and there remains one term in the transformation which is not used in this step. We use the latter term to eliminate the additional term $x^{j_0+n_0 s}$ in $\phi(x)$.

Remark 7. The normal form from Theorem 7 can be written in the following equivalent way

$$J^s_{r,\phi}\colon\ [y+x^r(1+\phi(x))]\,\partial_x+x^{s-1}\partial_y.$$

In the papers [B2] and [S2] Bogdanov and Sadovski present their normal form as

$$\dot{x}=y+x^r\left(1+\sum g_j x^j\right),\qquad \dot{y}=x^{s-1},$$

where $g_j=0$ for $j=sm-r$, $j=sm$, $j>3$ (without restrictions onto the first term $x^r\partial_x$). In our form the case with $r=sm$ is absent and in the case of integer n_0 the normal form is more complicated.

Remark 8. In case (ii) of Theorem 7 an equivalent form is $X_H+x^{r-1}(1+\sum_{i=1,\,i\neq s-r}^{s-1} x^i\phi_i(H))\,E_H$.

In the case (iii) of Theorem 7 the more natural normal form, equivalent to $J^s_{r,\phi}$, $\phi=c_{n_0 s}x^{n_0 s}$, is

$$X_H+x^{s/2-1}H^{n_0}(1+\mu H^{n_0})^{-1}E_H. \tag{2.3}$$

Analogously one can rewrite the normal form in case (iv). This kind of normal form was obtained by Loray in [L3].

Remark 9. The result of Theorems 7 can be extended to the nonnilpotent case $s=2$, i.e., $1:-1$ resonant saddle. It is known that the formal orbital normal form of such a saddle can be written as $X_H+(H^m+\mu H^{2m})\,E_H$. It coincides case (iii) of Theorem 7 with $r=2m+1$, $n_0=m$.

Because the complete analytic normal form of a resonant vector field contains functional moduli and is different from the formal normal form, we get an explanation of why the form from Theorem 7 is not analytic.

In a previous version of this work we obtained a formal normal form with the set of exponents only like in case (ii). Consideration of the case $s=2$ allowed revealed the mistake. We also strived to show that the formal normal form is analytic, which cannot be true in the case $s=2$. A close examination of the proofs has shown that the proof of analyticity of the Bogdanov–Takens normal form (1.6) was correct but the reduction of additional terms from $c(x)\,E_H$ can be divergent.

In order to obtain a complete analytic normal form one must construct a sort of Martinet–Ramis moduli: divide a neighborhood of (0, 0) into conic-like domains and perform analytic reductions in the cones to some standard systems (e.g., with a solvable monodromy group). The differences between the reducing transformations in intersections of the cones could give functional moduli. Probably these moduli will not be flat (contrary to the Martinet–Ramis moduli) and the coefficients of their Taylor expansion should correspond to the coefficients c_j in the formal functional modulus ϕ.

Something like this is already done with the hidden holonomy group (see Theorem 1), But we do not think that the story ends here.

In the next theorem we classify the holonomy groups according to the formal classification from Theorem 7.

THEOREM 8. *The hidden holonomy group associated with the germ* $X_H + c(x)\, E_H$ *is:*

— *non-solvable in cases* (ii) *with* $\phi \not\equiv 0$ *and (iv) from Theorem* 7*;*

— *typical solvable (nonabelian) in case (ii) with* $\phi \equiv 0$*;*

— *abelian in the cases* (i) *(then it is finite) and* (iii) *(then it is exceptional).*

In the cases of solvable and finite monodromy groups the formal orbital normal form is the same as the analytic orbital normal form.

This theorem says that the coefficients in the formal function $\phi = \sum' c_j x^j$ from Theorem 7 are the obstacles to solvability of the monodromy group. We see that in a generic case the holonomy group is nonsolvable, the solvability is the phenomenon of infinite codimension (very rare). More precisely, the set of germs of type (1.7) with solvable monodromy forms a pro-algebraic subset (i.e., it is given by algebraic equations in finite jets) of infinite codimension.

We do not know the full interpretation of the coefficients c_j in terms of known objects. We have only partial results in this direction.

THEOREM 9. (a) *If* $s = 3$*, then the first coefficient of the Taylor expansion of the function* ϕ *plays the role of the saddle quantity of the saddle* p_0*; it is an obstacle to its linearization.*

(b) *If* $n_0 \in \mathbb{Z}$*, then the term* $x^{r-1} E_H$ *is the first obstacle to the linearization of the resolved vector field near any of the saddles* $p_1, p_2 \in E_k$*. In case* (iii) *the coefficient* $c_{n_0 s}$ *plays the role of the modulus of formal orbital classification of the resolved vector field near* p_1 *as well as near* p_2*.*

Remark 10. Stróżyna [St] has recently obtained analogous results in the generalized saddle–node, (i.e. when the inequality (1.8) is reversed).

After resolution of the singularity, $(x, y) \to (x, u = y/x^r)$, one obtains a $1:-r$ resonant saddle p_0: $u = x = 0$ and a saddle–node p_1: $x = 0$, $u = 1$ with nonzero eigenvalue in the direction of $x = 0$. Two germs of such vector fields are orbitally analytically (formally) equivalent iff their monodromy groups are analytically (formally) equivalent.

A natural analogue of the Bogdanov–Takens prenormal form is $(y - x^r)\,\partial_x + yc(x)\,\partial_y$ with analytic $c(x)$, $c(x) = x^{s-r-1} + \cdots$. Denote $E_H = x\partial_x + ry\partial_y$ and $n_0 = s/r - 2$.

The final orbital formal normal form is equal to either:

(i) $J_r^{\infty,0}$: $(y - x^r)\,\partial_x$, or

(ii)

$$J_r^{s,\phi}\colon\ (y - x^r)\,\partial_x + x^{s-r-1}(1 + \phi(x))\,E_H$$

where $\phi(x) = \sum' c_j x^j$ and $\sum' = \sum_{j \neq 0 \,(\mathrm{mod}\ r)}$, if $n_0 \notin \mathbb{Z}$, or

(iii) $J_r^{s,\phi}$ with $\phi = c_{n_0 r} x^{n_0 r}$, if $n_0 \in \mathbb{Z}$, or

(iv) $J_r^{s,\phi}$ with $\sum' = \sum_{j \in \{n_0 r,\, j_0\}} + \sum_{j \neq 0\,(\mathrm{mod}\ r),\, j \neq t_0 + n_0 r}$, if $n_0 \in \mathbb{Z}$ and there exists a (first) nonzero coefficient c_{j_0} with $j_0 \neq 0\,(\mathrm{mod}\ r)$.

The monodromy group associated with the singularity is nonsolvable in case (ii) with $\phi \not\equiv 0$ and (iv), solvable nonabelian in case (ii) with $\phi = 0$ (it can be typical or exceptional), and abelian in cases (i) (it is finite) and (iii) (it is exceptional).

The term $x^{s-r-1}E_H$ responds for the first term in the normal form near the saddle–node p_1. In case (iii) the coefficient $c_{n_0 r}$ is a formal invariant of p_1.

Remark 11. Theorem 6 has applications in the problem of nilpotent centers. Recall that the (real) system in the Bogdanov–Takens prenormal form

$$\dot{x} = y + a(x), \qquad \dot{y} = \epsilon x^{s-1},$$

$a(x) = a_r x^r + \cdots$, $\epsilon = \pm 1$, has center iff: (a) $\epsilon = -1$, (b) $s = 2k$ is even, (c) $r > k$ or $r = k$ and $ra_r < 4$, and (d) $a(x) \equiv a(-x)$. This result was obtained independently by Sadovski [S1] and by Moussu [M1].

Berthier and Moussu [BM] have proven the following results about such centers:

(1) *It is analytically reversible, which means that there is an analytic vector field V' in $(\mathbb{R}^2, 0)$ and a local holomorphic map Φ: $(\mathbb{R}^2, 0) \to (\mathbb{R}^2, 0)$ of the fold type, $(x, y) \to (x^2 + \cdots, y + \cdots)$, such that $V = F \cdot \Phi_*^{-1} V' \circ \Phi$ for some holomorphic function F.*

(2) *Two vector fields* V_1, V_2 *with nilpotent centers are analytically orbitally equivalent iff the corresponding vector fields* V'_1, V'_2 *are analytically orbitally equivalent.*

The first result follows from Theorem 6, because the function $a(x)$ is analytic and the existence of center means that $a(x) = A(x^2)$. Thus we can apply the map $\Phi(x, y) = (x^2, y)$. The proof of (1) in [BM] is different. It relies on a subtle analysis of the monodromy maps.

Sadovski in [S3] investigated the problem of the existence of an analytic first integral for nilpotent centers. He has shown that there is such an integral iff the system is formally equivalent to the case $J^s_{\infty, 0}$. Using Theorems 7 and 8 we can assert the following.

COROLLARY 1. *The system* $\dot{x} = y + \cdots$, $\dot{y} = \cdots$ *with center has local analytic first integral iff it is analytically orbitally equivalent to the Hamiltonian system*

$$\dot{x} = y, \qquad \dot{y} = -x^{2k-1}.$$

Sadovski informed the authors that the problem of the analytic first integral for nilpotent centers was stated by Lyapunov.

The plan of the remainder of the paper is following. In the next section we recall the proofs of the Takens and Bogdanov–Takens prenormal forms. In Section 4 we prove Theorem 6. In Section 5 we prove Theorem 7. Section 6 contains the proof of Theorem 8 and Theorem 9 is proved in Section 7.

3. THE TAKENS HOMOLOGICAL EQUATION

Here we prove the formulas (1.4)–(1.7) at the level of formal power series. In Section 4 we shall prove that they are analytic. We also present an expansion formula for dependence of a transformed vector field on the change of variables; this formula will be next used in Section 4.

LEMMA 1 [T]. *There is a formal change of the variables* (x, y) *which reduces the vector field* $y\partial_x + \cdots$ *to the Takens prenormal form*

$$[y + a(x)]\,\partial_x + b(x)\,\partial_y. \tag{3.1}$$

Proof. We use a composition of infinite series of transformations of the form $id + Z^{(k)}$, i.e.,

$$(x, y) \to (x_1, y_1) = (x, y) + Z^{(k)}(x, y),$$

$x_1 = x + z_1(x, y)$, $y_1 = y + z_2(x, y)$, where $z_1(x, y)$ and $z_2(x, y)$ are homogeneous polynomials of degree k. We shall treat also $Z^{(k)}$ as a vector field, $Z^{(k)} = z_1\partial_x + z_2\partial_y$.

Let $V = y\partial_x + \sum_{j>1} V^{(j)}(X)$, where $V^{(j)}$ are homogeneous summands of the vector field V. Comparing the terms of degree k in the transformed vector field

$$(id + Z^{(k)})^* V = (id + Z^{(k)})_*^{-1} V \circ (id + Z^{(k)})$$

we obtain the homological equation

$$[Z^{(k)}, y\partial_x] + V^{(k)} = (V^{(k)})^{Takens}, \tag{3.2}$$

where $(V^{(k)})^{Takens}$ is a term from the Takens prenormal form, $(V^{(k)})^{Takens} = a_k x^k \partial_x + b_k x^k \partial_y$.

The operator $[\,\cdot\,, y\partial_x]$, written in the components, acts on $z_1\partial_x + z_2\partial_y$ as follows

$$z_1\partial_x + z_2\partial_y \to \left(z_2 - \frac{\partial z_1}{\partial x}\, y \right)\partial_x - \frac{\partial z_2}{\partial x}\, y\partial_y. \tag{3.3}$$

Equation (3.2) is the *Takens homological equation* and the operator (3.3) is the *Takens homological operator*.

Applying the formula (3.3) to Eq. (3.2), we see that we can cancel all the terms in the y-component of $V^{(k)}$ divisible by y and, when z_2 is fixed, we can also cancel the terms divisible by y in the x-component. In this way we obtain the formal Takens prenormal form (3.1). ▮

From the proof of Lemma 1 it follows that the components $z_1(x, y)$ and $z_2(x, y)$ can be chosen as divisible by x, and with this choice they are determined uniquely. Let us formulate this property in a separate lemma.

LEMMA 2. *For any homogeneous vector field $V^{(k)}$, $k \geqslant 2$, the homological equation* (3.2) *has unique solution $Z^{(k)}$ in the class of (homogeneous) vector fields of the form*

$$Z^{(k)} = x\widetilde{Z}^{(k-1)}.$$

This means that the homological operator (3.3) *defines an isomorphism between the space $(x\mathbb{C}[[x, y]])^2$ of formal vector fields of the form $Z = x\widetilde{Z}$ and the space $(yC[[x, y]])^2$ of formal vector fields of the form $W = y\widetilde{W}$ (complementary to the space of Takens prenormal forms).*

Remark 12. The equivalent to (3.1) preliminary normal forms are

$$\dot{x} = y, \qquad \dot{y} = b(x) + yc(x)$$

and, if the condition $s < 2r$ holds,

$$\dot{x} = y + \mu x c(x),\ \dot{y} = b_1(x) + yc(x), \qquad \mu > 0. \tag{3.4}$$

Proof. In the first case we make the substitution $y_1 = y + a(x)$. (Note that here we do not need any restrictions).

In the second case we make the substitution $y_1 = y + d(x)$ with $d(x)$ satisfying the equation $\mu x d' = -d + a(x)$. We have

$$d(x) = \mu^{-1} x^{-1/\mu} \int_0^x s^{1/\mu - 1} a(s)\, ds$$

We see that, if $\mu > 0$, then $d(x)$ is well defined as a formal power series if $a(x)$ is formal, and as an analytic function if $a(x)$ is analytic.

We see also that $d(x) = \text{const} \cdot x^r + \cdots$ if $a(x) = \text{const} \cdot x^r + \cdots$. ▮

LEMMA 3 [B2]. (a) *Let $s < \infty$. There exists a change of the variables (x, y) and of time which reduces the vector field (3.1) to the Bogdanov–Takens prenormal form*

$$[y + a(x)]\, \partial_x + s x^{s-1} \partial_y.$$

(b) *Let $s < 2r < \infty$, i.e., the generalized cusp case. There exists a change of the variables (x, y) and of time which reduces the vector field* (3.1) *to the Bogdanov–Takens prenormal form*

$$(2y\partial_x + s x^{s-1} \partial_y) + c(x)(2x\partial_x + s y \partial_y), \tag{3.5}$$

which has the property that the cusp curve $y^2 = x^s$ is invariant.

(c) *The changes from the points* (a) *and* (b) *are analytic when the field* (3.1) *is analytic and are formal when* (3.1) *is formal.*

Proof. (a) Assume that $b(x) = s x^{s-1} b_1(x)$, $b_1(0) \neq 0$. Our aim is to reduce $b_1(x)$ to 1.

We apply the change $x_1 = x\lambda(x)$, $dt/dt_1 = \eta(x)$, which should satisfy the equations $dx_1/dt_1 = y + a_1(x_1)$, $dy/dt_1 = s x_1^{s-1}$. Comparing the terms $s x_1^{s-1} \partial_y$ and $y \partial_{x_1}$ we get the conditions

$$\eta(x) \cdot (x\lambda(x))' = 1, \qquad \eta(x)\, b_1(x) = s\lambda^{s-1}(x),$$

or the Bernoulli equation $x\lambda' + \lambda = \frac{1}{s} b_1(x)\, \lambda^{1-s}$. It has the solution

$$\lambda(x) = x^{-1} \left[\int_0^x \tau^{s-1} b_1(\tau)\, d\tau \right]^{1/s}.$$

This solution is an analytic function when $b_1(x)$ is analytic.

(b) We apply the change $x_1 = x\lambda(x)$, $y_1 = y + d(x_1)$, $dt/dt_1 = \eta(x) = (dx_1/dx)^{-1}$ (from the point (a) and Remark 12) to the vector field $[2y + a(x)]\, \partial_x + x^{s-1} b_1(x)\, \partial_y$. We write $a(x) = \tilde{a}(x_1)$, $\eta(x) = \tilde{\eta}(x_1)$.

We obtain the system

$$\frac{dx_1}{dt_1} = 2y_1 + [\tilde{a}(x_1) - 2d(x_1)]$$

$$\frac{dy_1}{dt_1} = \eta(x)[b(x) + d'(x_1)(\tilde{a}(x_1) - 2d(x_1))] + y_1 \cdot 2\tilde{\eta}(x_1)\, d'(x_1).$$

Following the proof of the formula (3.4), we choose $d(x_1)$ such that $\tilde{a}(x_1) - 2d(x_1) = 2x_1 c(x_1)$ and $y_1 \cdot 2\tilde{\eta}(x_1)\, d'(x_1) = s y_1 c(x_1)$.

By the assumption $s < 2r$ we have $d'(x_1)(\tilde{a}(x_1) - 2d(x_1)) = O(x^{2r-1}) = o(x^{s-1})$ and we can write $b(x) + d'(x_1)(\tilde{a}(x_1) - 2d(x_1))$ as $x^{s-1} \cdot [b_1(0) + x b_2(x, \lambda(x))]$. We obtain the following differential equation for $\lambda(x)$: $x\lambda' + \lambda = \frac{1}{s} \lambda^{1-s} [b_1(0) + x b_2(x, \lambda)]$.

Introduce the variable $\xi = \lambda^s$. Then the latter equation changes to

$$x\xi' + s\xi = b_1(0) + x\tilde{b}_2(x, \xi). \tag{3.6}$$

The initial condition (i.e., the limit as $x \to 0$) is $\xi(0) = b_1(0)/s$.

The graphic $\Gamma = \{(x, \xi(x))\}$ of the solution $\xi(x)$ to the latter problem forms a phase curve of the analytic vector field $\dot{x} = x$, $\dot{\xi} = b_1(u) - s\xi + x\tilde{b}_2$. We see that Γ forms a separatrix of the singular point $(0, \xi(0))$ of this vector field. This singular point is a saddle, which means that Γ is a unique analytic invariant curve through $(0, \xi(0))$. Therefore the function $\xi(x)$ (as well as $\lambda(x)$) is analytic if the function $\tilde{b}_2$ is analytic, i.e., if the system (3.1) is analytic.

If (3.1) is only formal, then Eq. (3.6) admits a solution in the form of a power series, which is unique. ∎

We finish this section by deriving the expansion of the changed vector field $(id + Z)^* V = ((I + DZ)^{-1} V) \circ (id + Z)$ with respect to Z.

LEMMA 4. *We have the following formula*

$$
\begin{aligned}
(id+Z)^* V = V & \\
&+[Z,V] \\
&+\tfrac{1}{2}\langle (D^2V\circ(id+\zeta Z))\cdot Z, Z\rangle \\
&-(DZ\cdot DV)\circ(id+\zeta' Z)\cdot Z \\
&-\langle D^2Z\circ(id+\zeta' Z)\cdot V\circ(id+\zeta' Z), Z\rangle \\
&+((DZ)^2\,(I+DZ)^{-1}\cdot V)\circ(id+Z).
\end{aligned}
\tag{3.7}
$$

Here DU and D^2U mean the first and second derivatives of a vector field $U=(u_1,u_2)$ with respect to x,y. The expressions like $U\circ(id+\zeta Z)$, $\zeta=(\zeta_1,\zeta_2)\in[0,1]\times[0,1]$, mean that both components of the field U take the form $u_1(x+\zeta_1 z_1, y+\zeta_1 z_2)$, $u_2(x+\zeta_2 z_1, y+\zeta_2 z_2)$ (the same index in ζ as in u).

Proof. Formula (3.7) is a consequence of the mean value theorem, which states that if $f(x)$, $x\in\mathbb{C}^k$ is an analytic scalar function and m is natural, then $f(x+h)=f(x)+Df(x)\cdot h+\frac{1}{2}\langle D^2f(x)\,h,h\rangle+\cdots+\frac{1}{m!}\langle D^mf(x+\zeta h),(h,\ldots,h)\rangle$ for some $\zeta\in[0,1]$.

The first row in the right–hand side of (3.7) contains the term constant with respect to Z. The second row contains the linear terms. The last row arises from the expansion $(I+DZ)^{-1}=\sum(-DZ)^n$ and takes into account only the summands with $n\geqslant 2$, $\sum_{n\geqslant 2}(-DZ)^n=(DZ)^2\,(I+DZ)^{-1}$.

The mean value theorem is applied to the expressions $V\circ(id+Z)=V+DV\cdot Z+\frac{1}{2}\langle DV\circ(id+\zeta Z)\cdot Z,Z\rangle$ and $-(DZ\cdot V)\circ(id+Z)=-DZ\cdot V-D(DZ\cdot V)\circ(id+\zeta' Z)\cdot Z$ (in fact, it is applied independently to each component).

The reader should note that in (3.7) the argument is shifted in V as well as in Z (and in their derivatives). ∎

4. THE ANALYTIC TAKENS PRENORMAL FORM

4.1. *Preliminaries.* In this section we prove Theorem 6. In order to show analyticity of the Takens prenormal form, we must prove convergence of the series of transformations described in Section 3.

There are two general methods of the proof of analyticity: estimation of the coefficients of the fields Z or reduction of the problem to a fixed point equation for some contractible operator in some space of analytic vector-valued functions. The first method is used in the linearization theorems in the Siegel domain, when there are no resonances (see [A]), while the second method was used by N. Brushlinskaya [Br] in the Poincaré domain. Our method is a combination of both.

The main ingredients of the proof are Newton's method of solution of nonlinear equations and a certain modification of the norm of nonlinear terms.

Let us make the problem more precise. Lemma 2 (from Section 3) says that the formal reduction of terms outside the Takens form, i.e., the terms $W^{(k)} = y\widetilde{W}^{(k-1)}$, is obtained by means of compositions of transformations $id + Z^{(k)}$, where $Z^{(k)} = x\widetilde{Z}^{(k)}$ are homogeneous of degree k. Moreover the Takens homological operator $Z \to [Z, y\partial_x]$ is an isomorphism between the spaces $(x\mathbb{C}[[x, y]])^2$ and $(y\mathbb{C}[[x, y]])^2$ of formal vector fields. Therefore we shall apply formal changes in the form $id + Z$, $Z = x\widetilde{Z}$ and try to show that the power series for $\widetilde{Z}$ is convergent.

Denote

$$\pi U(x, y) = U(x, y) - U(x, 0), \tag{4.1}$$

the projection onto the space of formal vector fields divisible by y (i.e., π: $(\mathbb{C}[[x, y]])^2 \to (y\mathbb{C}[[x, y]])^2$) and by

$$\frac{1}{y}, \tag{4.2}$$

the operator of division by y.

The condition that the nonlinear terms in the field $(id + Z)^* V = y\partial_x + \cdots$ depend only on x reads as $\pi[(id + Z)^* V - y\partial_x] = 0$. We rewrite it in the (homogeneous) form

$$\frac{1}{y} \circ \pi\{(I + Z)^* V - y\partial_x\} = 0, \qquad Z = x\widetilde{Z}. \tag{4.3}$$

It is an equation for $\widetilde{Z}$.

Equation (4.3) contains the linear part

$$\frac{1}{y} \circ \pi[x\widetilde{Z}, V],$$

a constant part (which has to be reduced), and a nonlinear part (see Lemma 4). The linear part strongly depends on the vector field V. This means that during iterations we shall deal with varying problems. In order to impose some stability, we distinguish the part of the latter expression associated with the Takens prenormal form part of V^{Takens}.

We have

$$\begin{aligned} V &= V^{Takens} + W, \\ V^{Takens} &= [y + a(x)]\,\partial_x + b(x)\,\partial_y, \\ W &= y\widetilde{W} = \pi(V - y\partial_x). \end{aligned} \tag{4.4}$$

Define the linear operator

$$L\colon \widetilde{Z} \to \frac{1}{y} \circ \pi[x\widetilde{Z}, V^{Takens}]. \tag{4.5}$$

The operator L is invertible in the space $(\mathbb{C}[[x, y]])^2$ of formal vector fields (by Lemma 2). It will vary during iterations, but it stabilizes with the number of iterations going to infinity. Moreover, the operator L is simpler than the commutator operator with the whole V and we will be able to estimate its inverse in a suitable Banach space.

By formula (3.7) from Lemma 4, Eq. (4.3) can be rewritten in the form

$$L\widetilde{Z} + \widetilde{W} + \frac{1}{y} \circ \pi[Z, W] + (nonlinear\ terms\ in\ \widetilde{Z}) = 0. \tag{4.6}$$

We shall solve Eq (4.6) successively, in each step getting better approximation to the normal form. In each step we solve it approximately. We put

$$\widetilde{Z} = -L^{-1}\widetilde{W}. \tag{4.7}$$

Therefore $\widetilde{W}$ is of the same order as $\widetilde{Z}$.

The transformed vector field $V' = (id + Z)^* V$ again has the form $V' = V'^{Takens} + W'$, $W' = \pi(V' - y\partial_x) = y\widetilde{W}'$. From the formula (3.7) we have

$$\begin{aligned} V'^{Takens} &= V^{Takens} + [Z, V]^{Takens} + (nonlin.\ terms)^{Takens}, \\ W' &= \pi[Z, W] + \pi(nonlin.\ terms), \end{aligned} \tag{4.8}$$

where π is the projection operator defined in (4.1) and $(nonlin.\ terms)$ denote the terms nonlinear in Z in (3.7). We see that W' is of second order with respect to W and $V'^{Takens} - V^{Takens}$ is of first order with respect to W.

Moreover, the stabilization of the Taylor expansions of V and V^{Takens} takes place. Namely, if W begins with terms of order m, then W' begins with terms of order $\geqslant m + 1$. It follows from Lemma 1.

In the further stages of the proof of Theorem 1 we shall assume that

$$s = 3, \tag{4.9}$$

i.e., $b(x) = x^2 + b_3x^3 + \cdots$. The general case ($s = 2$ or $s > 3$) is proved in the same way. Our restriction serves to simplify the exposition.

4.2. *The operator L.* The operator $\widetilde{Z} \to L\widetilde{Z}$ (see (4.5)) written in the basis $(x^iy^j\partial_x, x^iy^j\partial_y)$ is divided into two parts, *diagonal* and *quasinilpotent*. The latter means that for each element of the above basis some power of the operator nullifies this element. (In functional analysis by a quasinilpotent operator people mean an operator in a Banach space whose whole spectrum is concentrated at 0).

Let $\widetilde{Z} = \tilde{z}_1\partial_x + \tilde{z}_2\partial_y = (\tilde{z}_1, \tilde{z}_2)$. We have

$$L\widetilde{Z} = \frac{1}{y} \circ \pi[-y(x\widetilde{Z})'_x + x\tilde{z}_2\partial_x + x\tilde{z}_1a'\partial_x - a(x\widetilde{Z})'_x + x\tilde{z}_1b'\partial_y - xb\widetilde{Z}'_y].$$

The diagonal part is

$$L_0\widetilde{Z} = -\frac{1}{y}\pi[y(\widetilde{Z})'_x] = -(x\widetilde{Z})'_x$$

or

$$L_0x^iy^j\partial_{x,y} = -(i+1)\,x^iy^j\partial_{x,y}. \tag{4.10}$$

The operator $L - L_0$ is divided into three parts

$$L_1\widetilde{Z} = \frac{1}{y}\pi[x\tilde{z}_2\partial_x],$$

$$L_2\widetilde{Z} = \frac{1}{y}\pi[xa'\tilde{z}_1\partial_x - a(x\widetilde{Z})'_x + \tilde{z}_1b'\partial_y],$$

$$L_3\widetilde{Z} = \frac{1}{y}\pi[-xb\widetilde{Z}'_y].$$

We have

$$L = L_0(I + L_0^{-1}L_1 + L_0^{-1}L_2 + L_0^{-1}L_2).$$

LEMMA 5. *The expressions of the operators $L_0^{-1}L_j$ in the monomial basis are the following*

$$L_0^{-1}L_1x^iy^j\partial_y = \frac{-1}{i+2}x^{i+1}y^{j-1}\partial_x$$

$$L_0^{-1}L_1x^iy^j\partial_y = 0$$

$$
(4.11)\quad L_0^{-1}L_2x^iy^j\partial_x = \frac{i-1}{i+3}a_2x^{i+2}y^{j-1}\partial_x + \frac{i}{i+4}a_3x^{i+3}y^{j-1}\partial_x + \cdots
$$
$$
-\frac{2}{i+3}b_2x^{i+2}y^{j-1}\partial_y - \frac{3}{i+4}b_3x^{i+3}y^{j-1}\partial_y + \cdots
$$
$$
L_0^{-1}L_2x^iy^j\partial_y = \frac{i+1}{i+3}a_2x^{i+2}y^{j-1}\partial_y + \frac{i+2}{i+4}a_3x^{i+3}y^{j-1}\partial_y + \cdots
$$
$$
L_0^{-1}L_3x^iy^j\partial_{x,y} = \frac{j}{i+4}b_2x^{i+3}y^{j-2}\partial_{x,y} + \frac{j}{i+5}b_3x^{i+4}y^{j-2}\partial_{x,y} + \cdots
$$

Here we put zero whenever we get a negative power of y in the above formulas and $b_2 = 1$.

Figure 2 shows how the operators $L_0^{-1}L_j$ act.

4.3. *The weighted norm.* From Lemma 5 we see that the operators $L_{1,2}$ are bounded relatively to L_0. L_3 is not relatively bounded. However the fact that it is quasinilpotent allows us to treat it as bounded with respect to L_0. As in the finite–dimensional case, an operator of the form semisimple plus quasinilpotent is invertible. In the finite–dimensional case it is used to introduce a new norm such that the nilpotent part is small relatively to the semisimple part.

The natural norm in the space of vector–valued power series, giving the Banach space of series representing analytic vector fields, is $\|U\|_\rho = \sum_{ij}(|u_{1,ij}|+|u_{2,ij}|)\,\rho^{i+j}$ for some $\rho>0$ (representing the radius of convergence of the series $U=\sum_{i,j}(u_{1,ij}\partial_x+u_{2,ij}\partial_y)\,x^iy^j$); so $\|x^iy^j\partial_{x,y}\|_\rho=\rho^{i+j}$.

We modify this norm in the following way. We put $\|x^iy^j\partial_{x,y}\|_{w,\rho} = w_{i,j}\rho^{i+j}$, where $w_{i,j}$ are weights defined below.

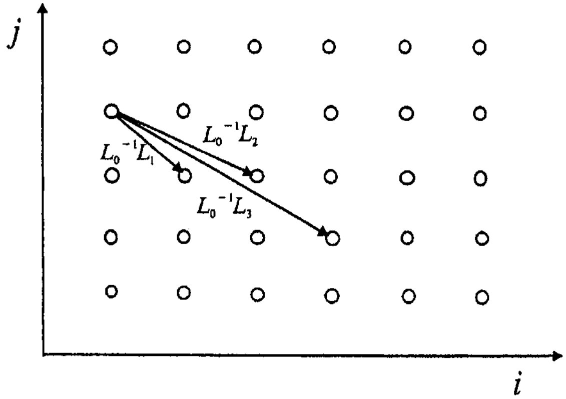

FIGURE 2

DEFINITION 4. Define rational numbers $\tilde{w}_{i,j}$, $(i, j) \in \mathbb{Z}_+^2$ as follows:

$$\begin{aligned} \tilde{w}_{i,0} &= 1, \\ \tilde{w}_{i,1} &= 1, \\ \tilde{w}_{i,j} &= \max\left\{1, \frac{j-2}{i+3}\right\} \tilde{w}_{i+3,j-2}, \qquad \text{for } j \geqslant 2. \end{aligned} \tag{4.12}$$

The *weight* $w_{i,j}$ is equal to

$$w_{i,j} = \Lambda^j \cdot \sqrt{j+1} \cdot \tilde{w}_{i,j}, \tag{4.13}$$

where $\Lambda > 1$ is some constant which will be fixed later.

We denote by $\mathcal{X}_{w,\rho}$ the Banach space of series $U = \sum_{i,j}(u_{1,ij}\partial_x + u_{2,ij}\partial_y)\, x^i y^j$ with the norm

$$\|U\|_{w,\rho} = \sum_{ij} (|u_{1,ij}| + |u_{2,ij}|)\, w_{ij}\rho^{i+j}.$$

Let us comment on this definition. The choice of the numbers $\tilde{w}_{i,j}$ is motivated by the requirement that the operator $L_0^{-1}L_3$ be bounded. The leading matrix elements of this operator (corresponding to $x^i y^j \partial_{x,y} \to x^{i+3}y^{j-2}\partial_{x,y}$) equal $\frac{j}{i+4}$. It is almost the same as $\frac{j-2}{i+3} = \tilde{w}_{i,j}/\tilde{w}_{i+3,j-2}$. However, the ratio $\frac{j-2}{i+3}$ has the geometrical interpretation as the slope coefficient of the radius vector of the endpoint of the segment S, joining the point (i, j) with $(i+3, j-2)$ and presented in Fig. 2.

Note also that the numbers $\tilde{w}_{i,j}$ are chosen such that they are equal to 1 whenever $j \leqslant i+5$, i.e., below the shifted diagonal; (but above this shifted diagonal they are strictly increasing as we move along the lines with the slope $-2/3$ in the left direction). The reason for this is that the matrix elements of the operators $L_0^{-1}L_{1,2}$ are of order $O(1)$ and we cannot use the ratios $\frac{j-2}{i+3}$, which are smaller than 1.

The factor Λ^j allows us to control the norm of the operator $L_0^{-1}L_1$; the greater Λ is, the smaller the norm $\|L_0^{-1}L_1\|_{w,\rho}$ is. The constant Λ will be determined in Lemma 6 below. In fact, we will fix $\Lambda = 2$.

The factor $\sqrt{j+1}$ is responsible for the estimate from Lemma 8(e) (which is needed for estimation of norms of products of series); so it is introduced for technical reasons.

If the integer s from $b(x) = sx^{s-1} + \cdots$ is different than 3 (i.e., (4.9) does not hold), then the inductive definition of the numbers $\tilde{w}_{i,j}$ is modified as follows: $\tilde{w}_{i,j} = \max\{1, \frac{j-2}{i+s}\}\tilde{w}_{i+s,j-2}$.

Remark 13. The idea of changing the norm was used in [Br]. There the Lyapunov metric was introduced.

4.4. *The weighted norm at work.* In this point we demonstrate how useful our weighted norms are. Namely, we prove that the operator L is invertible in the Banach space $\mathscr{X}_{w,\rho}$ with bounded inverse and that the elements from $\mathscr{X}_{w,\rho}$ represent the Taylor series of analytic vector fields.

We will use the notation

$$U^{\geqslant m} = \sum_{i+j\geqslant m} (u_{1,ij}\partial_x + u_{2,ij}\partial_y)\, x^i y^j$$

for a vector field $U = \sum_{i,j}(u_{1,ij}\partial_x + u_{2,ij}\partial_y)\, x^i y^j$.

LEMMA 6 (Bound for L^{-1}). *Assume that* $V^{Takens} = [y + a(x)]\,\partial_x + b(x)\,\partial_y = y\partial_x + V^{\geqslant 2}$ *is an analytic vector field such that*

$$\|V^{\geqslant 2}\|_{w,\rho} = \sum_{n\geqslant 2} (|a_n| + |b_n|)\,\rho^n \leqslant \hat{C}\rho^2 \tag{4.14}$$

for small ρ's, where the constant $\hat{C}$ does not depend on ρ.

If the constant Λ (from Definition 4) is sufficiently large and ρ is sufficiently small, then the operator L^{-1} is bounded in $\mathscr{X}_{w,\rho}$ with the norm $\leqslant 4$.

Proof. We use the formula $L^{-1} = (I + L_0^{-1}L_1 + L_0^{-1}L_2 + L_0^{-1}L_3)^{-1}L_0^{-1}$.

Because L_0 is diagonal with the eigenvalues $-(i+1)$ (see (4.10)) then $\|L_0^{-1}\|_{w,\rho} = 1$.

We use the following properties of weights:

(i) $w_{i+1,j} \leqslant w_{i,j}$,

(ii) $w_{i+1,j-1} \leqslant \Lambda^{-1} w_{i,j}$.

Inequality (i) is equivalent to the inequality $\tilde{w}_{i+1,j} \leqslant \tilde{w}_{i,j}$ and follows from formula (4.12) (see also the point (a) of Lemma 8). Inequality (ii) constitutes the point (b) of Lemma 8 and will be proved later.

Note that (i) implies $w_{i+k,j} \leqslant w_{i,j}$ for any $k > 0$ and the both inequalities imply $w_{i+k,j-l} \leqslant \Lambda^{-l} w_{i,j}$ for $k > l \geqslant 0$.

Let $e = e_{i,j} = x^i y^j \partial_{x,y}$. By Lemma 5, inequality (ii), and Definition 4 of weights, we have

$$\begin{aligned}\frac{\|L_0^{-1}L_1 e\|_{w,\rho}}{\|e\|_{w,\rho}} &\leqslant \frac{1}{i+2}\cdot\frac{\|e_{i+1,j-1}\|}{\|e_{i,j}\|}\\ &= \frac{1}{i+2}\cdot\frac{w_{i+1,j-1}}{w_{i,j}}\\ &\leqslant \frac{1}{i+2}\Lambda^{-1}.\end{aligned}$$

We choose the constant Λ such that $\|L_0^{-1}L_1\|_{w,\rho} \leqslant 1/4$; i.e., we put $\Lambda = 2$.

Next, by Lemma 5, inequalities (i) and (ii), Definition 4, and formula (4.14), we have

$$\begin{aligned}\frac{\|L_0^{-1}L_2e\|}{\|e\|} &\leqslant |a_2|\,\rho\,\frac{w_{i+2,\,j-1}}{w_{i,\,j}}+|a_3|\,\rho^2\,\frac{w_{i+3,\,j-1}}{w_{i,\,j}}+\cdots\\ &\quad+|b_2|\,\rho\,\frac{w_{i+2,\,j-1}}{w_{i,\,j}}+|b_3|\,\rho^2\,\frac{w_{i+3,\,j-1}}{w_{i,\,j}}+\cdots\\ &\leqslant\frac{1}{\rho}\cdot\frac{w_{i+2,\,j-1}}{w_{i,\,j}}\cdot\sum(|a_n|+|b_n|)\,\rho^n\\ &\leqslant\frac{1}{\rho\Lambda}\,\|V^{\geqslant 2}\|_{w,\,\rho}\\ &\leqslant\hat{C}\rho/\Lambda.\end{aligned}$$

For small ρ we get $\|L_0^{-1}L_2\|\leqslant 1/4$.

Finally, we have

$$\begin{aligned}\frac{\|L_0^{-1}L_3e\|}{\|e\|} &\leqslant\frac{j}{i+4}\,|b_2|\,\rho\,\frac{w_{i+3,\,j-2}}{w_{i,\,j}}+\frac{j}{i+5}\,|b_3|\,\rho^2\,\frac{w_{i+4,\,j-2}}{w_{i,\,j}}+\cdots\\ &\leqslant\frac{j}{i+4}\cdot\frac{1}{\rho}\cdot\Lambda^{-2}\cdot\frac{\tilde{w}_{i+3,\,j-2}}{\tilde{w}_{i,\,j}}\cdot\sum|b_n|\,\rho^n\\ &\leqslant\left(\frac{j}{i+4}\min\left\{1,\frac{i+3}{j-2}\right\}\right)\cdot\frac{1}{\rho\Lambda^2}\cdot\|V^{\geqslant 2}\|_{w,\,\rho}\\ &\leqslant 3\,\frac{1}{\rho\Lambda^2}\,\|V^{\geqslant 2}\|\\ &\leqslant 3\hat{C}\rho/\Lambda^2.\end{aligned}$$

(we have used $\frac{j}{i+4}\cdot\min\{1,\frac{i+3}{j-2}\}\leqslant 3$). We get $\|L_0^{-1}L_3\|\leqslant 1/4$ for small ρ.

The reader can see that the same estimates hold in the case when $s=3$ is replaced by arbitrary integer $\geqslant 2$. The estimates for L_0, $L_0^{-1}L_1$, and $L_0^{-1}L_2$ are the same. There are only minor differences in the estimate for $L_0^{-1}L_3$. ∎

LEMMA 7 (Analyticity). *We have*

$$\rho^{i+j}\leqslant\|x^iy^j\partial_{x,\,y}\|_{w,\,\rho}\leqslant(C_1\rho)^{i+j}$$

for some constant C_1 which does not depend on i, j, ρ.

This implies that any element U of $\mathscr{X}_{w,\rho}$ represents an analytic vector field in the ball B_ρ with center at 0 *and radius ρ and that*

$$\sup_{(x,y)\in B_\rho} |U(x,y)| \leqslant \|U\|_{w,\rho}. \tag{4.15}$$

Moreover:

(a) *If a series $U(x,y)$ is such that $\|U^{\geqslant k}\|_{w,\rho} < const\cdot(\rho/\rho_0)^k$, $k=1,2,\ldots$, then U represents an analytic function in the ball B_{ρ_0}.*

(b) *Conversely, if a series $U(x,y)$ represents an analytic function convergent in an open set containing the ball B_{ρ_0}, $\rho_0 > C_1\rho$, then $U\in\mathscr{X}_{w,\rho}$ and $\|U^{\geqslant k}\|_{w,\rho} \leqslant K\cdot\lambda^k$, $\|DU^{\geqslant k}\|_{w,\rho} < K\cdot(2\lambda)^k$, where*

$$\lambda = C_1\rho/\rho_0 \tag{4.16}$$

and K is a constant depending on f.

Proof. This lemma follows directly from the inequality

$$1\leqslant w_{i,j}\leqslant C_1^{i+j}$$

which is proved in point (c) of Lemma 8 from the next section. ∎

4.5. *Estimates for the weights and the weighted norms.* This section is devoted to the formulation of additional technical lemmas necessary to obtain recursive estimates in the proof of Theorem 1. These lemmas are formulated in the same way for $s=3$ and for other s. In fact only in the proof of Lemma 8 are there minor differences for different s. The proofs are put off to the last part of this section.

LEMMA 8 (Properties of weights). *The weights $w_{i,j}$ satisfy the following properties:*

(a) $w_{i+k,j}\leqslant w_{i,j}$ *for* $k>0$,

(b) $w_{i+k,j-k}\leqslant \Lambda^{-k}w_{i,j}\leqslant w_{i,j}$ *for* $k>0$,

(c) $1\leqslant w_{i,j}\leqslant C_1^{i+j}$,

(d) $w_{i,j}\leqslant C_1(i+j+1)^{k+l}\,w_{i+k,j+l}$ *for* $k,l\geqslant 0$, $0\leqslant k+l\leqslant 2$,

(e) $w_{i+k,j+l}\leqslant C_1 w_{i,j}w_{k,l}$,

where C_1 is a constant not depending on the indices.

Inequalities (a), (b), and (c) from Lemma 8 have been used in the previous section. Estimates (b), (d), and (e) are successively used in the proofs of the Lemmas 9–11.

The further lemmas will deal with functions rather than with vector fields. Namely we introduce the Banach space $\mathscr{F}_{w,\rho}$ of scalar power series $f(x, y) = \sum_{i,j} f_{ij} x^i y^j$, $f_{ij} \in \mathbb{C}$ with the norm

$$\|f\|_{w,\rho} = \sum |f_{ij}|\, w_{i,j} \rho^{i+j}.$$

The estimates will be formulated in terms of elements from $\mathscr{F}_{w,\rho}$, but they are also true for vector–valued (and for matrix–valued) series. This approach was suggested by the referee and its aim is to simplify the exposition.

LEMMA 9 (Estimate of derivatives). *If $f \in \mathscr{F}_{w,\rho}$, $\rho'/\rho = 1-\delta \in (\frac{1}{2}, 1)$, and $k+l \leqslant 2$, then*

$$\|\partial_x^k \partial_y^l f\|_{w,\rho'} \leqslant \frac{C_3}{(\rho\delta^2)^{k+l}} \cdot \|f\|_{w,\rho},$$

where C_3 is a constant not depending on f, k, l, ρ, δ.

The proof, which uses the estimate (d) of Lemma 8, will be given later.

LEMMA 10 (Estimate of products). *Let f be a series in $\mathscr{F}_{w,\rho}$.*

(a) *If g is a function from $\mathscr{F}_{w,\rho}$, then fg is in $\mathscr{F}_{w,\rho}$ and*

$$\|fg\|_{w,\rho} \leqslant C_2 \cdot \|f\|_{w,\rho} \cdot \|g\|_{w,\rho}.$$

(b) *The series $\frac{1}{y} \circ \pi(f) \in \mathscr{F}_{w,\rho}$ (see* (4.1) *and* (4.2)*) and*

$$\left\| \frac{1}{y} \circ \pi(f) \right\|_{w,\rho} \leqslant \frac{C_2}{\rho} \cdot \|f\|_{w,\rho}.$$

Here the constant C_2 does not depend on f, g, ρ.

The proof uses estimate (e) of Lemma 8 and is put off to the last section.

LEMMA 11 (Estimate of compositions). *Let $Z \in \mathscr{X}_{w,\rho}$ be a vector field of the form $x\tilde{Z}$ and let $\rho'/\rho = 1-\delta \in (\frac{1}{2}, 1)$ be such that $\|\tilde{Z}\|_{w,\rho} \leqslant \delta/(2C_2)$, where C_2 is the constant from Lemma 8. If a function $f \in \mathscr{F}_{w,\rho}$, then $f \circ (id+Z) \in \mathscr{F}_{w,\rho'}$ and*

$$\|f \circ (id+Z)\|_{w,\rho'} \leqslant \left(1 + \frac{C_4\, \|\tilde{Z}\|_{w,\rho}}{\delta^3}\right) \|f\|_{w,\rho},$$

where C_4 is some constant not depending on f, Z, ρ, δ.

In particular, if we assume that $\|\widetilde{Z}\|_{w,\rho} \leqslant \delta^3/C_4$ *(what will take place in applications), then*

$$\|f \circ (id+Z)\|_{w,\rho'} \leqslant 2\,\|f\|_{w,\rho}. \tag{4.17}$$

The proof uses point (b) of Lemma 8 and will be given in the end of this section. We remark only here that the thesis of Lemma 11 ceases to be true when we remove the restriction onto the form of Z (i.e., for $Z \neq x\widetilde{Z}$). The reason is that when we expand $f \circ (id+Z)$ into series, then we obtain the terms $D^m f \cdot Z^m$ where the norms $\|D^m f\|$ behave badly for large m. When $Z = x\widetilde{Z}$, then $Z^m = (x^m)\cdot \widetilde{Z}^m$ and the large power of x gives small norm of Z^m (see Definition 4). This compensates the growth of $\|D^m f\|$.

4.6. *Description of the iterative procedure.* We begin with an application of the Lemmas 9–11 to the transformed vector field $(id+Z)^* V$. There we must replace scalar series by vector–valued or by matrix–valued series; this is not a problem.

COROLLARY 2. *There exists a constant* C_5 *not depending on* V, Z, ρ, δ *such that for* $0<\rho'<\rho< <1$ *satisfying* $\rho'/\rho = 1-\delta \in (\frac{1}{2}, 1)$*, any* $V \in \mathscr{X}_{w,\rho}$*, and any* $Z = x\widetilde{Z}$ *such that* $\|\widetilde{Z}\|_{w,\rho} \leqslant \delta^3/C_4$ *(see Lemma 11) the following inequalities hold*

$$\|[Z,V]\|_{w,\rho'} \leqslant \frac{C_5}{\delta^2}\,\|V\|_{w,\rho}\cdot\|\widetilde{Z}\|_{w,\rho},$$

$$\|(id+Z)^* V - V - [Z,V]\|_{w,\rho'} \leqslant \frac{C_5}{\delta^6}\,\|V\|_{w,\rho}\cdot(\|\widetilde{Z}\|_{w,\rho})^2.$$

Proof. The commutator $[Z,V]$ equals $DZ\cdot V - DV\cdot Z$. The expression $(id+Z)^* V - V - [Z,V]$ consists of the nonlinear terms from (3.7), i.e., $= (1/2)\langle D^2V \circ (id+\zeta Z)\cdot Z, Z\rangle - (DZ\cdot DV)\circ(id+\zeta' Z)\cdot Z - \langle D^2 Z \circ (id+\zeta' Z)\cdot V \circ (id+\zeta' Z), Z\rangle + ((DZ)^2\,(I+DZ)^{-1}\cdot V)\circ(id+Z)$. We shall use also the notion

$$(nonlin.terms) = (id+Z)^* V - V - [Z,V]. \tag{4.18}$$

We introduce the intermediate norm $\|\cdot\|_{w,\rho''}$, where $\rho'' = \frac{1}{2}(\rho+\rho')$. First we estimate the derivatives with nonshifted argument in this new norm using Lemma 9. Each derivative gives the factor $(1-\rho''/\rho)^{-2} < const\cdot\delta^{-2}$. Next, we apply Lemma 11 (the formula 4.17) to estimate the corresponding factors in summands with shifted argument (like $D^2V \circ (id+\zeta Z)$) using the norm $\|\cdot\|_{w,\rho'}$. Finally, we estimate the products in the norm $\|\cdot\|_{w,\rho'}$ using Lemma 10(a).

Recall that one summand takes the form $[(DZ)^2 (I+DZ)^{-1} \cdot V] \circ (id+Z)$. Due to the assumptions $\delta \ll 1$ and $\|\widetilde{Z}\|_{w,\rho} = O(\delta^3)$ the 2×2-matrix valued function DZ is small and the term $(I+DZ)^{-1}$ is bounded.

Since the nonlinear terms contain expressions where the derivative appears at most three times in each summand, we get the factor $1/\rho^6$. The linear terms acquire only the factor $1/\rho^2$. ∎

We pass to the description of iterations. Assume that we begin with the vector field $V = V_1$ which is analytic in a ball B_{ρ_0}. We have the splitting $V_1 = V_1^{Takens} + W_1$, $W_1 = y\widetilde{W}_1$. We can assume that W_1 has order 4 at the origin.

Choose a radius ρ_1 in such a way that the parameter

$$\lambda = C_1 \rho_1 / \rho_0$$

from Lemma 7 (the formula (4.16)) is small. We obtain the starting estimates (see Lemma 7)

$$\begin{aligned} \|V_1\|_{w,\rho_1} &\leqslant K\lambda, \\ \|V_1^{Takens} - y\partial_x\|_{w,\rho_1} &\leqslant K\lambda^2, \\ \|W_1\|_{w,\rho_1} &\leqslant K\lambda^4, \end{aligned} \tag{4.19}$$

where the constant K does not depend on ρ_1 (only on V).

Introduce the radii

$$\rho_k = \left(1 - \frac{1}{k^2}\right)\rho_{k-1}, \qquad k \geqslant 2. \tag{4.20}$$

We have $\lim \rho_k = \rho_\infty = \text{const} \cdot \rho_1 > 0$. ρ_∞ will be the radius of the ball B_{ρ_∞} of analyticity of the Takens prenormal form.

Define inductively the vector fields Z_k, V_k, V_k^{Takens}, $W_k = y\widetilde{W}_k$ and diffeomorphisms $\mathcal{G}_k$:

$$\begin{aligned} Z_k &= x\widetilde{Z}_k, \\ \widetilde{Z}_k &= -L^{-1}\widetilde{W}_k, \\ V_{k+1} &= (id+Z_k)^* V_k = V_{k+1}^{Takens} + W_{k+1}, \\ \mathcal{G}_k &= (id+Z_1) \circ \cdots \circ (id+Z_k). \end{aligned}$$

Recall that the operator L equals $\widetilde{Z} \to \frac{1}{y} \circ \pi[x\widetilde{Z}, V_k^{Takens}]$. Therefore L also depends on k, $L = L_k$.

The Taylor expansion of Z_k starts from terms of degree $\geqslant 3+k$. Hence the terms of degrees $\leqslant k+3$ in $\mathcal{G}_k$ are fixed. In this sense $\mathcal{G}_k$ tend to some formal map. The same holds for V_k. We have to show that $\mathcal{G}_k$ and V_k converge in the ball B_{ρ_∞}.

4.7. *Proof of Theorem* 6. This proof relies on the following proposition.

MAIN PROPOSITION. *For any $k \geqslant 1$ and λ small enough one has*

$$\begin{aligned} \|V_k^{Takens} - V_{k-1}^{Takens}\|_{w,\,\rho_k} &\leqslant K(2\lambda)^k, \\ \|W_k\|_{w,\,\rho_k} &\leqslant K\lambda^{k+3}, \\ \|Z_k\|_{w,\,\rho_k} &\leqslant K\lambda^{k+2}, \end{aligned} \tag{4.21}$$

so that

$$\begin{aligned} \|V_k\|_{w,\,\rho_k} &\leqslant \frac{4K\lambda}{1-2\lambda} \leqslant 8K\lambda, \\ \|V_k^{Takens} - y\partial_x\|_{w,\,\rho_k} &\leqslant 6K\lambda^2, \end{aligned} \tag{4.22}$$

where K is the constant from (4.19).

If the estimates (4.21) for Z_k hold, then by the bound (4.15) from Lemma 7 $|Z_k(x, y)| \leqslant K\lambda^{k+2}$ as $(x, y) \in B_{\rho_k}$. Moreover Z_k has small derivative in a slightly smaller ball $B_{(\rho_k+\rho_{k+1})/2}$, $|DZ_k(x, y)| \leqslant \|DZ_k\|_{w,\,(\rho_k+\rho_{k+1})/2} \leqslant const \cdot k^{-4} \cdot \lambda^{k+2}$ for $(x, y) \in B_{(\rho_k+\rho_{k+1})/2}$.

Therefore the map $id + Z_k$ represents a holomorphic diffeomorphism of the ball $B_{(\rho_k+\rho_{k+1})/2}$ to a domain containing the ball $B_{\rho_{k+1}}$. Indeed, $|G^{(k)}| \sim \lambda^k$ is much smaller than the difference of the radii $(\rho_k+\rho_{k+1})/2 - \rho_{k+1} = O(k^{-2})$ and then $(id+Z_k)(B_{(\rho_k+\rho_{k+1})/2}) \supset B_{\rho_{k+1}}$.

The infinite composition of the maps $id + Z_k$ is convergent in the ball B_{ρ_∞} and represents an analytic diffeomorphism $\mathcal{G}_\infty = \lim \mathcal{G}_k : B_{(\rho_k+\rho_{k+1})/2} \to \mathcal{G}_\infty(B_{(\rho_k+\rho_{k+1})/2})$.

The limit of V_k's defines a vector field V_∞ holomorphic in the ball B_{ρ_∞}. It is in the Takens prenormal form and the diffeomorphism $\mathcal{G}_\infty$ conjugates V with V_∞.

This completes the proof of Theorem 6. ∎

Remark 14. The convergence of the successive approximations is very rapid. If the perturbation is of order ε, then after n steps we would obtain error of order ε^{2^n}. Such convergence is characteristic for Newton's method of finding approximate zero of a differentiable functions. It was used in KAM theory (see also [A]). In our proof we generally have followed the

Newton's method but we did not use as good estimates as in [A] (Section 3, 12). We have used also some ideas from [P].

4.8. *Proof of main proposition.* The inequalities (4.22) are rather obvious.

The estimate for Z_k in (4.21) follows from the estimate for W_k. Indeed, we have $W_k = y\widetilde{W}_k$, $Z_k = x\widetilde{Z}_k$, and $\widetilde{Z}_k = -L^{-1}\widetilde{W}_k$. Using Lemma 10(b) we estimate $\|\widetilde{W}\|_{w,\rho_k} = \|\frac{1}{y} \circ \pi(W)\|$ by $C_2\rho_k^{-1}\,\|W_k\|_{w,\rho_k}$. The assumption of Lemma 7, i.e. that the nonlinear part of V_k^{Takens} is small enough, holds, $\|V_k^{Takens} - y\partial_x\|_{w,\rho_k} \leqslant 6K\lambda^2 \leqslant \frac{6K}{\rho_0^2}\rho_1^2 \leqslant \hat{C}\rho_k^2$, where $\hat{C} = \mathrm{const}\cdot K/\rho_0^2$. Therefore $\|L^{-1}\| \leqslant 4$ and $\|\widetilde{Z}_k\|_{w,\rho_k} \leqslant 4\,\|\widetilde{W}\|_{w,\rho_k}$. Next, we use Lemma 10(a) to estimate the product $x\cdot\widetilde{Z}_k$. We obtain

$$\|Z_k\|_{w,\rho_k} \leqslant 4C_2^2\,\|W_k\|_{w,\rho_k} \leqslant (4C_2^2K\lambda)\cdot\lambda^{k+2}.$$

For small λ we have $4C_2^2K\lambda \leqslant K$ and the third estimate from (4.21) holds.

The first two estimates in (4.21) for $k=1$ hold by (4.19): we have $\|V_1^{Takens} - V_0^{Takens}\| = \|V_1^{Takens}\|_{w,\rho_1} \leqslant K\lambda \leqslant K(2\lambda)^1$ and $\|W_1\|_{w,\rho_1} \leqslant K\lambda^{1+3}$.

The inductive estimates for$V_{k+1}^{Takens} - V_k^{Takens}$ and W_{k+1} follow from formulas (4.8) and Corollary 2. Indeed, we have $V_{k+1} = (id+Z_k)^*\,V_k = V_{k+1}^{Takens} + W_{k+1}$, where by (4.8)

$$\begin{aligned} V_{k+1}^{Takens} &= V_k^{Takens} + [Z_k, V_k]^{Takens} + (nonlin.\,terms)^{Takens},\\ W_{k+1} &= \pi[Z_k, W_k] + \pi(nonlin.\,terms). \end{aligned}$$

Here π is the projection operator (4.1) and the notion $(nonlin.\,terms)$ is defined in (4.18). We use Corollary 2 with $V = V_k$, $Z = Z_k$, $\rho = \rho_k$, $\rho' = \rho_{k+1}$, $\delta = (k+1)^{-2}$ and with the inductive assumptions $\|V_k\|_{w,\rho_k} \leqslant 8K\lambda$, $\|Z_k\|_{w,\rho_k} \leqslant K\lambda^{k+2}$. We get

$$\begin{aligned} \|V_{k+1}^{Takens} - V_k^{Takens}\|_{w,\rho_{k+1}} &\leqslant C_5(k+1)^4\cdot 8K\lambda\cdot K\lambda^{k+2} \qquad \leqslant K(2\lambda)^{k+1},\\ \|W_{k+1}\|_{w,\rho_{k+1}} &\leqslant C_5(k+1)^{12}\cdot 8K\lambda\cdot(K\lambda^{k+2})^2 \leqslant K\lambda^{(k+1)+3} \end{aligned}$$

for λ small.

The proof of the Main proposition is complete. ∎

4.9. *Proof of the technical lemmas.*

(A) *Proof of Lemma* 8. Thorough the proof we shall use the notion of asymptotic equivalence of functions if $f(t)$ and $g(t)$ are positive functions of $t \geqslant 0$, then $f(t) \sim g(t)$ iff there exists a constant $C > 0$ such that $\frac{1}{C}\,g(t) \leqslant f(t) \leqslant Cg(t)$ and $f(t) = O(g(t))$ when $f(t) \leqslant Cg(t)$ for some constant C.

For example, the Stirling formula says that $\Gamma(t+1) \sim t^te^{-t}\sqrt{t+1}$ and implies the formula $\Gamma(t+a+1) \sim (t+1)^a\,\Gamma(t+1)$ for a fixed or bounded.

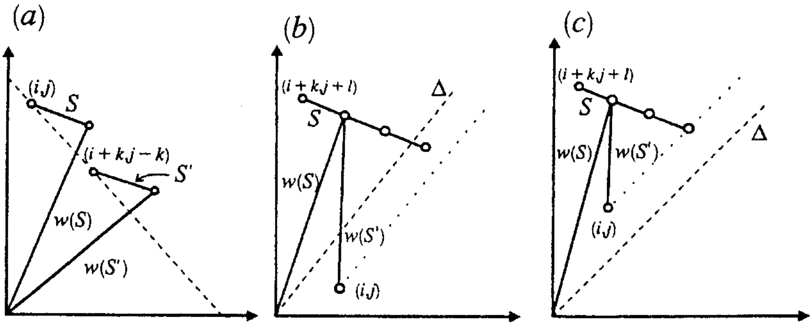

FIGURE 3

Recall the definition of weights,

$$w_{i,j} = \Lambda^j \sqrt{j+1}\, \tilde{w}_{i,j},$$

where $\Lambda > 1$ and

$$\tilde{w}_{i,0} = \tilde{w}_{i,1} = 1, \qquad \tilde{w}_{i,j} = \max\left\{1, \frac{j-2}{i+3}\right\} \tilde{w}_{i+3,j-2}.$$

We obtain that

$$\tilde{w}_{i,j} = \prod_{n=1}^{N} \frac{j-2n}{i+3n}, \tag{4.23}$$

where N is the integer defined by $N < \frac{j-i}{5} \leqslant N+1$ if $j-i > 5$ and the product is void, i.e., $\tilde{w}_{i,j} = 1$, if $j-i \leqslant 5$. Hence $N = \frac{j-i}{5} - \delta$, where $0 < \delta \leqslant 1$.

The factors $\frac{j-2n}{i+3n}$ have the interpretation of slopes $\tilde{w}(S_n)$ of the radius vectors from the origin to the right endpoints of the segments $S = S_n$ joining the points $(i+3n-3, j-2n+2)$ and $(i+3n, j-2n)$. The product in (4.23) is the product $\prod \tilde{w}(S_n)$ along the chain $\gamma = (S_1, ..., S_N)$ joining the point (i, j) with the "target" set $\Delta = \{j-i = 5\}$ (see Fig. 3).

Point (a) *of Lemma* 8 states that $w_{i+k,j} \leqslant w_{i,j}$ for $k > 0$. This inequality is equivalent to the inequality $\tilde{w}_{i+k,j} \leqslant \tilde{w}_{i,j}$, which easily follows from the above recurrent formula.

Point (b) *of Lemma* 8 states that $w_{i+k,j-k} \leqslant \Lambda^{-k} w_{i,j}$ for $k > 0$. Since $\sqrt{\frac{j+1}{j-k+1}} > 1$, this inequality follows from the inequality

$$\tilde{w}_{i+k,j-k} \leqslant \tilde{w}_{i,j}.$$

The latter inequality becomes evident when we use the slope interpretation of the quantities $\tilde{w}_{i,j}$. This situation is presented in Fig. 3a. We associate with each segment S from the chain γ starting at (i, j) the segments $S' = S + (k, -k)$ from the chain starting at $(i+k, j-k)$. We see that $w(S') < w(S)$.

Point (c) *of Lemma* 8 is the estimate $1 \leqslant w_{ij} \leqslant C_1^{i+j}$. Of course, it is enough to show that $1 \leqslant \tilde{w}_{i,j} \leqslant \tilde{C}^{i+j}$ for some constant $\tilde{C}$. This is a consequence of the following representation of the quantities $\tilde{w}_{i,j}$.

LEMMA 12. *Let* $\alpha = \frac{i}{i+j} \leqslant \frac{1}{2}$, $\beta = 1 - \alpha = \frac{j}{i+j}$. *We have*

$$\tilde{w}_{i,j} \sim \sqrt{\frac{i+1}{j+1}}\ \Phi(\alpha)^{i+j},$$

where $\Phi(\alpha) \geqslant 1$ *is uniformly bounded from the above continuous function.*

This means that $w_{i,j} \sim \sqrt{i+1}\ \Lambda^j \Phi(\alpha)^{i+j}$.

Proof. We use the formula (4.23) for $j > i+5$, which can be rewritten as

$$\begin{aligned} \tilde{w}_{i,j} &= \left(\frac{2}{3}\right)^N \cdot \frac{j/2-1}{i/3+1} \cdot \;\cdots\; \cdot \frac{j/2-N}{i/3+N} \\ &= \left(\frac{2}{3}\right)^N \cdot \frac{\Gamma(j/2)}{\Gamma(j/2-N)} \cdot \frac{\Gamma(i/3+1)}{\Gamma(i/3+N+1)} \qquad (4.24) \\ &\sim \left(\frac{2}{3}\right)^{(j-i)/5} \frac{\Gamma(i/3+1)\,\Gamma(j/2+1)}{\Gamma((2i+3j)/10+1)\,\Gamma((2i+3j)/15+1)} \cdot \frac{i+j+1}{j+1}. \end{aligned}$$

We use the Stirling formula. Putting $i = \alpha(i+j)$ and $j = \beta(i+j)$, after some calculation, we get $\tilde{w}_{ij} \sim \sqrt{\frac{i+1}{j+1}} \cdot \Phi(\alpha)^{i+j}$ where

$$\Phi(\alpha) = \alpha^{\alpha/3} \beta^{\beta/2} \left(\frac{5}{2\alpha+3\beta}\right)^{(2\alpha+3\beta)/6}, \qquad 0 \leqslant \alpha \leqslant \frac{1}{2}. \tag{4.25}$$

The function $\Phi(\alpha)$, $\alpha \in [0, \frac{1}{2}]$ is continuous and bounded: $1 = \Phi(\frac{1}{2}) \leqslant \Phi(\alpha) \leqslant \Phi(0) = \sqrt{\frac{5}{3}}$. It is Lipschitz continuous away from the point $\alpha = 0$; at $\alpha = 0$ the derivative is equal to $-\infty$.

In the case of general nilpotent singularity, not just $s = 3$, Lemma 12 is also true, but the function $\Phi(\alpha)$ takes the form $\alpha^{\alpha/s} \beta^{\beta/2}((s+2)/(2\alpha+s\beta))^{(2\alpha+s\beta)/2s}$. ∎

Point (e) *of Lemma* 8 says that the ratios

$$\frac{w_{i,j}}{w_{i+k,j+l}} \tag{4.26}$$

are bounded by $const\cdot(i+j+1)^{k+l}$ for $0\leqslant k+l\leqslant 2$. We have $w_{i,j}/w_{i+k,j+l}\sim \tilde{w}_{i,j}/\tilde{w}_{i+k,j+l}$. When $j-i\leqslant 5$ and/or $j+l-i-k\leqslant 5$ the numerator and the denominator of the latter ratio are uniformly bounded and the estimate is true. Otherwise we use the formulas (4.24) and $\Gamma(m+a+1)\sim(m+1)^a\,\Gamma(m)$. We obtain that (4.26) is $\leqslant const\cdot(\frac{i+j+1}{i+1})^{k/3}(\frac{i+j+1}{j+1})^{l/2}\leqslant const\cdot(\frac{i+j+1}{i+1})^{k/3}$ (since $i\leqslant O(j)$). This is $\leqslant const\cdot(i+j+1)^{k/3}\leqslant const\cdot(i+j+1)^{k+l}$.

Point (d) *of Lemma* 8 states that the ratio

$$R_{ijkl}=\frac{w_{i+k,j+l}}{w_{i,j}w_{k,l}}$$

is uniformly bounded. Due to the inequality $\sqrt{\frac{j+l+1}{(j+1)(l+1)}}\leqslant 1$, the problem reduces to the problem of estimating the ratio

$$\frac{\tilde{w}_{i+k,j+l}}{\tilde{w}_{i,j}\tilde{w}_{k,l}}$$

uniformly with respect to the indices.

We distinguish three cases:

(i) $j\leqslant i+5$ and $j+l\leqslant i+k+5$;

(ii) $j\leqslant i+5$ and $j+l>i+k+5$;

(iii) $j>i+5,\ l>k+5$; here we assume $\frac{j}{i}\leqslant\frac{l}{k}$.

In case (i) we have $\tilde{w}_{i,j}=\tilde{w}_{i+k,j+l}=1$, $\tilde{w}_{k,l}\geqslant 1$ and the inequality from Lemma 8(e) is obvious.

In case (ii) we have $\tilde{w}_{i,j}=1$ and the problem is to estimate the ratio $\tilde{R}=\tilde{w}_{i+k,j+l}/\tilde{w}_{k,l}$. Recall that $\tilde{w}_{i+k,j+l}$ is the product of slopes $w(S)=\frac{n-2}{m+3}$ over segments $S=[(m,n),(m+3,n-2)]$ from the chain γ joining the point $(i+k,j+l)$ with the diagonal.

The situation is presented in Fig. 3b. With each segment $S=[(i+m,j+n),(i+m+3,j+n-2)]$ from the chain γ we associate the unique segment $S'=[(m,n),(m+3,n-2)]=S-(i,j)$ from the chain γ' starting at (k,l). Because $w(S)<w(S')$ and $w(S'')\geqslant 1$ for other segments from the chain γ', we get $\tilde{w}_{i+k,j+l}\leqslant\tilde{w}_{k,l}$.

The situation from case (iii) is presented in Fig. 3c. With the segments S from the chain γ, which lie above the line $y=(j/i)\,x$, we associate the segment $S'=S-(i,j)$ from the chain γ'. We have $w(S)<w(S')$.

This shows that it is enough to prove the estimate from Lemma 8(e) in the case when $\alpha=\frac{i}{i+j}\approx\alpha'=\frac{k}{k+l}\approx\alpha''=\frac{i+k}{i+j+k+l}$. We get $w_{i,j}\sim\sqrt{i+1}\,\Lambda^j\Phi(\alpha)^{i+j}, w_{k,l}\sim\sqrt{k+1}\,\Lambda^l\Phi(\alpha)^{k+l}, w_{i+k,j+l}\sim\sqrt{i+k+1}\,\Lambda^{j+l}\Phi(\alpha)^{i+j+k+l}$ and hence $R_{ijkl}\sim\sqrt{\frac{i+k+1}{(i+1)(k+1)}}=O(1)$.

Unfortunately this is not the end of the proof of this point. It is correct when the equalities $\alpha=\alpha'=\alpha''$ are exact. However, they can be not exact and the function $\Phi(\alpha)$ from Lemma 12 is not Lipschitz continuous at 0. So we have to perform more estimates.

The differences between α's are of the form

$$\alpha'-\alpha=\frac{a}{k+l}, \qquad \alpha''-\alpha=\frac{a}{i+j+k+l},$$

where $|a|\leqslant 3$. This can be seen from Fig. 3c: a is the difference between the x-coordinates of the points $(i+k, j+l)$ and the point of intersection of the line $y=\frac{j}{i}x$ with the line $y+x=i+j+k+l$. We can assume that $i+j\leqslant k+l$. We have either $i>0$ (and then $\alpha\geqslant\frac{1}{i+j}$) or $i=0$ (and then $\alpha=0$, $0\neq k=a$ and $\frac{\alpha'}{\alpha''}\sim\frac{j+l}{l}$).

By (4.25) we have $\Phi(\alpha)=\alpha^{\alpha/3}\widetilde{\Phi}(\alpha)$, where the function $\widetilde{\Phi}$ is differentiable and hence Lipschitz continuous. Thus we have to estimate the two expressions:

$$\text{(4.27)}\qquad \left[\frac{(\alpha'')^{\alpha''(i+j+k+l)}}{\alpha^{\alpha(i+j)}(\alpha')^{\alpha'(k+l)}}\right]^{\frac{1}{3}} \quad\text{and}\quad \frac{\widetilde{\Phi}(\alpha'')^{i+j+k+l}}{\widetilde{\Phi}(\alpha)^{i+j}\,\widetilde{\Phi}(\alpha')^{k+l}}.$$

The logarithm of the first equals $\frac{1}{3}$ times

$$\text{(4.28)}\quad (i+j+k+l)[\alpha''\log\alpha''-\alpha\log\alpha]-(k+l)[\alpha'\log\alpha'-\alpha\log\alpha].$$

For $\alpha>0$ we represent $\alpha''\log\alpha''-\alpha\log\alpha$ as

$$\alpha''\log\frac{\alpha''}{\alpha}+(\alpha''-\alpha)\log\alpha=\frac{a}{i+j+k+l}(1+\log\alpha)+O\left(\frac{1}{\alpha(i+j+k+l)^2}\right).$$

Similarly we have $\alpha'\log\alpha'-\alpha\log\alpha=\frac{a}{k+l}(1+\log\alpha)+O(\frac{1}{\alpha(k+l)^2})$. Therefore (4.28) is bounded by $O(\frac{i+j}{k+l})=O(1)$, as $i+j\leqslant k+l$.

If $\alpha=0$, then (4.28) equals $a\log\frac{\alpha''}{\alpha'}\sim\log\frac{j+l}{l}\sim O(1)$, since $j\leqslant k+l=O(l)$.

Using the function $\Psi(\alpha)=\log\widetilde{\Phi}(\alpha)$ we represent the second ratio in (4.27) as the difference of two terms $(i+j+k+l)[\Psi(\alpha'')-\Psi(\alpha)]$ and $(k+l)[\Psi(\alpha')-\Psi(\alpha)]$. Both are bounded by La, where L is the Lipschitz constant of Ψ.

(Another proof of the bound for R_{ijkl} uses the asymptotic formula $w_{ij}=O(1)\cdot\Gamma(\frac{i}{3}+1)\cdot j^{1/2-i/3}\cdot(\Lambda\Phi(0))^j$ for $i\ll j$. Under the assumption $\frac{i}{k}\approx\frac{j}{l}$ we get $R_{ijkl}\sim O(1)\cdot(i+1)^{-3/2}\cdot j^{-1/2}$.) ∎

Remark 15. The reader can notice here that the choice of definition of weights is not very rigid. For example, instead of the product $\prod w(S)$ in (4.23) one could use the formula (4.24) or its modification with $N=\frac{j-i}{5}$. Nevertheless, the requirements of finiteness of the norm $\|L_0^{-1}L_3\|$ says that they should take the form of product of terms $\sim\frac{j}{i}$. Another restriction is imposed by the just proven property (e) of Lemma 8.

(B) *Proof of Lemma* 9. Let $f=\sum f_{ij}x^iy^j$. We have to estimate the derivatives $\partial_x^k\partial_y^l f$ for $k+l\leqslant 2$.

We have

$$\begin{aligned}\|\partial_x^k\partial_y^l f\|_{w,\rho'} &\leqslant \sum_{i,j} i^kj^l\cdot|v_{ij}|\cdot w_{i-k,j-l}(\rho')^{i+j-k-l}\\ &\leqslant \|V\|_{w,\rho}\cdot(\rho')^{-k-l}\max_{ij}\left(i^kj^l\cdot\frac{w_{i-k,j-l}}{w_{ij}}\cdot(\rho'/\rho)^{i+j}\right).\end{aligned}$$

By point (e) of Lemma 8 $w_{i-k,j-l}/w_{i,j}<\text{const}\cdot(i+j+1)^{k+l}$. Thus

$$\begin{aligned}i^kj^l\cdot\frac{w_{i-k,j-l}}{w_{ij}}\cdot(1-\delta)^{i+j} &\leqslant \text{const}\cdot(i+j+1)^{2(k+l)}\,(1-\delta)^{i+j+1}\\ &\leqslant \text{const}\cdot\delta^{-2(k+l)}.\end{aligned}$$

(We have used the inequality $t^N(1-\delta)^t\leqslant(N/e\delta)^N$ for small δ.)

Now the estimate of the norm of the derivative by *const* $\cdot\|V\|/(\rho\delta^2)^{k+l}$ follows. ∎

(C) *Proof of Lemma* 10. (a) Let $f=\sum f_{ij}x^iy^j$, $g=\sum g_{ij}x^iy^j$, $f\cdot g=\sum h_{ij}x^iy^j$, $\|f\|_{w,\rho}=\sum|f_{ij}|\,w_{ij}\rho^{i+j}<\infty$, $\|g\|_{w,\rho}=\sum|g_{ij}|\,w_{ij}\rho^{i+j}<\infty$. Then we have

$$\|fg\|_{w,\rho}\leqslant\sum_{k,l}\sum_{i,j}\frac{w_{kl}}{w_{ij}\cdot w_{k-i,l-j}}(|f_{ij}|\,w_{ij}\rho^{i+j})\cdot(|g_{k-i,l-j}|\,w_{k-i,l-j}\rho^{k-i+l-j}).$$

By Lemma 8(d) $w_{kl}\leqslant C_1\cdot w_{ij}\cdot w_{k-i,l-j}$ for $0\leqslant i\leqslant k$, $0\leqslant j\leqslant l$. From this the first estimate from Lemma 10 follows.

(b) The second estimate is proved in the same way, using the inequality $w_{i,j-1}<w_{i,j}$. The latter is a consequence of the inequalities $w_{i,j-1}\leqslant w_{i+1,j-1}$ and $w_{i+1,j-1}\leqslant\Lambda^{-1}w_{i,j}<w_{i,j}$ given in Lemma 8 (the points (a) and (b)). ∎

(D) *Proof of Lemma* 11. Let $f=\sum f_{ij}x^iy^j$. We estimate $f\circ(id+Z)-f$ for $Z=x\tilde{Z}=(x\tilde{z}_1,x\tilde{z}_2)$.

Expanding this expression we get

$$\begin{aligned}\|f\circ(id+Z)-f\|_{w,\rho'} &\leqslant \sum_{i,j}|f_{ij}| \sum_{k+l\geqslant 1}\binom{i}{k}\binom{j}{l}\|(x\tilde{z}_1)^k(x\tilde{z}_2)^l x^{i-k}y^{j-l}\|_{w,\rho'}\\ &\leqslant \sum_{i,j}|f_{ij}| \sum_{k+l\geqslant 1}\binom{i}{k}\binom{j}{l}\|(x^{i+l}y^{j-l})\,\tilde{z}_1^k\,\tilde{z}_2^l\|_{\rho'}.\end{aligned}$$

Because $f\in\mathscr{F}_{w,\rho}$, we have $|f_{ij}|\leqslant\|f\|_{w,\rho}/(w_{ij}\rho^{i+j})$. By definition of the weighted norm, $\|x^{i+l}y^{j-l}\|_{w,\rho'}\leqslant w_{i+l,j-l}(\rho')^{i+j}$, where $w_{i+l,j-l}<w_{i,j}$ by Lemma 8(b). Using the estimate for products from Lemma 10(a), we get

$$\begin{aligned}\|f\circ(id+Z)-f\|_{w,\rho'} &\leqslant \|f\|\sum_{i,j}(\rho'/\rho)^{i+j}\sum_{k+l\geqslant 1}\binom{i}{k}\binom{j}{l}(C_2\,\|\tilde{Z}\|)^{k+l}\\ &\leqslant \|f\|\sum_{i,j}(1-\delta)^{i+j}\,[(1+C_2\,\|\tilde{Z}\|)^{i+j}-1]\\ &\leqslant \mathrm{const}\cdot\|f\|\cdot\|\tilde{Z}\|\sum_{i,j}(i+j)[(1-\delta)(1+C_2\,\|\tilde{Z}\|]^{i+j-1}\\ &\leqslant \mathrm{const}\cdot\|f\|\cdot\|\tilde{Z}\|\sum_{i,j}(i+j)[1-\delta/2]^{i+j-1}\\ &\leqslant \mathrm{const}\cdot\|f\|\cdot\|\tilde{Z}\|\sum_{n}(n+1)\,n[1-\delta/2]^{n-1}\\ &\leqslant \mathrm{const}\cdot\|f\|\cdot\|\tilde{Z}\|\cdot(2/\delta)^3\\ &\leqslant \frac{C_4}{\delta^3}\cdot\|f\|\cdot\|\tilde{Z}\|.\end{aligned}$$

Here we have used the inequalities $(1+t)^n-1\leqslant nt(1+t)^{n-1}$, $(1-\delta)(1+C_2\,\|\tilde{Z}\|)\leqslant 1-\delta/2$ (following from the assumption $\|\tilde{Z}\|\leqslant\delta/(2C_2)$) and $\sum_{n>0}(n+1)\,nt^{n-1}=2(1-t)^{-3}$. ∎

5. THE FORMAL NORMAL FORM

5.1. *Reduction of additional terms using the Hamiltonian part and solvable monodromy.* In Section 3 we obtained the Bogdanov–Takens prenormal form $[2y+a(x)]\,\partial_x+sx^{s-1}\partial_y$ (where, according to Theorem 6, the function $a(x)$ is analytic). In this section we assume the following restrictions.

We have $a(x)=a_rx^r+\cdots$, $a_r\neq 0$, $b(x)=sx^{s-1}$ with the condition

$$s<2r.$$

LEMMA 13. *The system*

$$\dot{x} = 2y + a_{sm}x^{sm} + O(x^{sm+1}), \qquad \dot{y} = sx^{s-1}$$

can be transformed to

$$\dot{x} = 2y + O(x^{sm+1}), \qquad \dot{y} = sx^{s-1}.$$

Proof. The smth order jet of this system can be transformed, via the change $(x, y) \to (h = y^2 - x^s, y)$, to

$$\dot{h} = -a_{sm}(y^2 - h)^m, \qquad \dot{y} = 1.$$

This is a nonsingular system with an analytic first integral $H = h + \cdots = y^2 - x^s + \cdots$. The function $H(x, y)$ has the $\mathbf{A}_{s-1}$ type singularity and is equivalent (by an analytic change of coordinates) to $Y^2 - X^s$ (see [AVG]). ▮

LEMMA 14. *If $r \neq 0 \pmod{s}$, then the terms of order r of the system*

$$\dot{x} = 2y + 2a_r x^r, \qquad \dot{y} = sx^{s-1} + sa_r x^{r-1} y$$

cannot be cancelled by means of any orbital transformation preserving the form $X_H + c(x)\, E_H$.

Proof. We can assume that $a_r = 1$ and we can consider only the rth jet of the vector field. We shall find a rational change of variables reducing it to some integrable vector field. This change is the same as the composition of the blowing-up maps described in the Introduction and the transformed system describes the resolved field of directions in some affine chart near the final divisor E of the resolution (see Introduction). The cases with odd and even s are treated separately.

(a) The case $s = 2k+1$. Here the change is the following

$$u = \frac{x^{2k+1}}{y^2}, \qquad v = \frac{y}{x^k} \tag{5.1}$$

(then $x = uv^2$, $y = u^k v^{2k+1}$). We obtain the system

$$\dot{u} = 2(2k+1)\, u(1-u), \qquad \dot{v} = [(2k+1)\, u - 2k]\, v + u^{r-k} v^{2(r-k)}. \tag{5.2}$$

The line $v = 0$ represents the affine line $E \setminus p_1$ (punctured divisor). The other singular points are $p_2 = (0, 0)$, $p_0 = (1, 0)$ and the separatrix Γ is given by $u = 1$; $(\infty, 0)$ represents the point p_1.

System (5.2) can be integrated. The equation for the phase curves is the Bernoulli equation with the first integral of the Darboux–Schwartz–Christoffel type

$$F = v^{2k-2r+1}u^{\alpha}(1-u)^{\beta} - \beta \int_0^u \tau^{\alpha+r-k-1}(1-\tau)^{\beta-1}\, d\tau, \tag{5.3}$$

where $\alpha = \frac{k(2k-2r+1)}{(2k+1)}$, $\beta = \frac{(2k-2r+1)}{(4k+2)}$.

Remark 16. Loray in his thesis calls the function (5.3) a hypergeometric function. We think that our notion is more suitable. The hypergeometric functions are given in certain integral form with the argument playing role of the parameter of the subintegral function. Their monodromies are usually nonsolvable. The argument in the integral in (5.3) stays in the limit and the function (5.3) can serve as a model example of a function with solvable monodromy.

By the assumption of Lemma 14, $r \neq 0 \pmod{s}$ and, as we shall see, the integral (5.3) is not of the Darboux type. This means that its monodromy group is solvable and nonabelian.

(By the *monodromy group* of a multivalued holomorphic function F on $\mathbb{C}^n$, with ramifications along some algebraic hypersurface S, we mean a certain group of permutations of a typical fiber $M_a = \pi^{-1}(a)$ of the covering $M \setminus S \to \mathbb{C}^n \setminus S$, where M is the Riemann surface of F with the natural projection π. These permutations are induced by analytic continuation of analytic elements (above a neighborhood of a) along loops in $\mathbb{C}^n \setminus S$ which start at a. In the case of solvable monodromy this group coincides with the group of deck automorphisms of the covering $M \setminus S \to \mathbb{C}^n \setminus S$).

The ramification curves of F are $u=0$, $u=1$, and $u=\infty$. The monodromy maps associated to loops around $u=0$ and $u=1$ are

$$h \to \mu_0 h, \qquad h \to \mu_1 h + c,$$

where $\mu_0 = e^{2\pi i\alpha}$, $\mu_1 = e^{2\pi i\beta}$, $c = \beta(1-\mu_1)\int_0^1 \tau^{\alpha+r-k-1}(1-\tau)^{\beta-1}\, d\tau$. The latter integral diverges at $\tau = 1$ (because $\beta < 0$). Therefore its value is given by a suitable regularization. Here we use the analytic continuation of the Euler betafunction $B(a, b) = \int_0^1 \tau^{a-1}(1-\tau)^{b-1}\, d\tau = \Gamma(a)\,\Gamma(b)/\Gamma(a+b)$ as a function of (a, b). If $r \neq 0 \pmod{s}$, then $c \neq 0$.

Of course, the monodromy group of the first integral is associated with the monodromy group of the divisor $E \setminus \{p_0, p_1, p_2\}$; one needs to parametrize the disk D transversal to E (at a point q) by F. Here the transition from the natural parametrization of the disk D by $v = v|_D$ to the parametrization $h = F|_D = Av^{2k-2r+1} + B$ is not one-to-one; it is semiconjugation of the monodromy group with a subgroup of the group of affine maps (see

Remark 1 in Section1). Nevertheless the fundamental properties of the groups (like solvability) are preserved under such semiconjugation.

The map f_1 (i.e., holonomy around p_1) corresponds to a loop around $u=\infty$, $f_1^{[2]}=\mathrm{id}$, the map f_2 corresponds to a loop around $u=0$, $f_2^{[2k+1]}=\mathrm{id}$, and $f_0=f_1\circ f_2$ corresponds to a loop around $u=1$. Applying the above semi-conjugation we get that $f_j(v)=\lambda_j v(1-C_j v^p)^{-1/p}$, $p=2r-2k-1$.

We have $f_0^{[4k+2]}=\mathrm{id}$, because the map f_0 in the chart h is a rotation around its fixed point. Thus by Theorem 5a (Section 1) the monodromy group is solvable. On the other hand it is nonabelian.

If there were a transformation reducing the rth order terms, then the monodromy group of this jet should be abelian. This shows the thesis of our lemma.

(b) The case $s=2k$. The change is of the form

$$u=\frac{y}{x^k}, \qquad v=x$$

and we obtain the system

$$\dot{u}=k(1-u^2), \qquad \dot{v}=uv+v^{r-k+1}. \tag{5.4}$$

The line $v=0$ represents the affine line $E\setminus\{\infty\}$, $p_0=(\infty,0)$, $p_{1,2}=(\pm 1,0)$ and the separatrices $\Gamma_{1,2}$ are $u=\pm 1$. The corresponding first integral equals

$$F=[v(1-u^2)^{1/2k}]^{k-r}+\frac{r-k}{k}\int^u (1-\tau^2)^{-(k+r)/2k}\,d\tau.$$

As in the previous case, we check that, for $r\neq 0 \pmod{k}$, the monodromy group of F is solvable nonabelian. (This case corresponds to the situation with $n_0\notin\mathbb{Z}$ in Theorem 7). Indeed, we get the situation exactly as in the case of odd s: F has ramifications at $u=\infty$, $u=\pm 1$ and the monodromy maps are of the form

$$h\to e^{\pi i(k-r)/k}h+c_{1,2},$$

where $c_{1,2}\neq 0$.

If $r=(2m+1)\,k$, then

$$F=v^{-2mk}(1-u^2)^{-m}+2m\int^u (1-\tau^2)^{-m-1}\,d\tau$$

contains logarithmic singularities $d_{1,2}\ln(1\pm u)$. The monodromy maps of F are $h\to h+c_{1,2}$ and generate an abelian group. The initial parametrization of the transversal D to the divisor E is given by v with the relation $h=F|_D=Av^{-2mk}+B$ and, after applying this semiconjugation, we obtain again the abelian group:

$$v\to e^{\pi i/k}v(1-C_{1,2}v^p)^{-1/p},\qquad p=2mk$$

We see that this group consists of maps $\lambda g^t_{v^{p+1}}$, $\lambda^p=1$, like the exceptional abelian group from Theorem 2(a) and Definition 3(a) from the Introduction (with the vector field $w=w_{p,0}$). Theorem 2(a) says that this group cannot be formally reduced to a finite abelian group. This is the case with $n_0\in\mathbb{Z}$ from Theorem 7; we have $m=n_0$.

Lemma 14 is complete. ∎

COROLLARY 3. *The monodromy group of system* (5.2) *or* (5.4) *(i.e., $J^s_{r,0}$) is solvable nonabelian for $n_0\notin\mathbb{Z}$ and exceptional abelian for $n_0\in\mathbb{Z}$. The generators take the forms*

$$f_0(v)=\lambda_0 g^{t_0}_{v^{p+1}},\qquad f_2(v)=\lambda_2 v,\qquad p=2(r-k)-1$$

for $s=2k+1$ and

$$f_{1,2}(v)=\lambda_{1,2}g^{t_{1,2}}_{v^{p+1}},\qquad p=r-k$$

for $s=2k$.

In particular, the monodromy group associated with any vector field $J^s_{r,\phi}$ cannot be formally linear.

Moreover, if $n_0\in\mathbb{Z}$, then the monodromy group of the initial germ $\dot{x}=2y+\cdots$, $\dot{y}=sx^{s-1}\cdots$ cannot be solvable nonabelian.

Remark 17. Consider the cases which are excluded from the assumption of Lemma 14.

Let $s=2k+1$. If r is of the form $(2k+1)\,m$, then the exponent α in the integral (5.3) is an integer, $\alpha=k(2k-2r+1)/(2k+1)=k(1-m)$, $\beta=1/2-m$. The integral $\int^u\tau^{(k+1)m-1}(1-\tau)^{\beta-1}\,d\tau$ in (5.3) can be expressed as $(1-u)^\beta\times$(polynomial). Because β is a half-integer, the square of the first integral F is a rational first integral for the system (5.2).

Let $s=2k$. If $r=2mk$, then the Schwartz–Christoffel integral in the expression for F takes the form

$$\int^u(1-\tau^2)^{-m-1/2}\,d\tau.$$

We claim that it is also of the form $(1-u^2)^{-m+1/2}\times$polynomial.

Indeed, calculation of the regularized integral $\mathcal{P}\int_{-1}^{1}(1-\tau^2)^{-m-1/2}$ (with the substitution $\imath=1-\tau^2$ and using the Beta–function) gives $-\int_0^1 1/(\imath^{m+1/2}(1-\imath)^{1/2})=-\Gamma(1/2)\,\Gamma(1/2-m)/\Gamma(-m)=0$. Thus the regularized "values" of the Schwarz–Christoffel integral at the branching points are equal. Putting this value equal to zero we extract from the Schwarz–Christoffel integral the factor $(1-u^2)^{1/2-m}$. The other factor must be a polynomial $P(u)$.

In fact we can find this polynomial explicitly as $P(u)=u(a_0+a_2u^2+\cdots+a_{m-1}u^{2(m-1)})$, where the coefficients satisfy the recurrent relations $a_0=1$, $a_j=-2\frac{m-j}{2j+1}a_{j-1}$. For example, $\int^u(1-\tau^2)^{-3/2}=u(1-u^2)^{-1/2}$, $\int^u(1-\tau^2)^{-5/2}=(u-2u^3/3)(1-u^2)^{-3/2}$. However $\int^u(1-\tau^2)^{-1/2}=\sin^{-1}(u)$.

This gives another explanation why the monodromy is abelian in the case $r=0\,(\mathrm{mod}\ s)$.

Notice that by repeating the proof of Lemma 13 we can reduce all the terms $a_{sm}x^{sm}\partial_x$. To be precise, we should take into account not the homogeneous filtration of the space of germs generated by the degree, but rather the quasi-homogeneous filtration given by the exponents

$$d(x)=2,\qquad d(y)=s,\qquad d(\partial_x)=-2,\qquad d(\partial_y)=-s.$$

Then the part of the vector field of the lowest quasihomogeneous degree (equal to $s-2$) is $X_H=2y\partial_x+sx^{s-1}\partial_y$. Because we can reduce $x^{sm}\partial_x$ using X_H, the same holds when we replace X_H by X_H+(higher order terms).

We see that the terms $x^{sm}\partial_x$ are the only terms which can be reduced when one uses the first part of the vector field, X_H.

Conclusion. Using the Hamiltonian part, we can reduce the system either to

$$\dot{x}=2y,\qquad \dot{y}=sx^{s-1}$$

(the case $J^s_{\infty,0}$) or to

$$\dot{x}=2y+x^r+\sum_{j\neq sm}a_jx^j,\ \dot{y}=sx^{s-1},\qquad r\neq 0\,(\mathrm{mod}\ s).$$

In the second case the rth order part of the system, i.e., $\dot{x}=2y+x^r$, $\dot{y}=sx^{s-1}$, does not determine the coordinate system uniquely. Namely, the changes $(x,y,t)\to(\alpha^2x,\alpha^sy,\alpha^{s-2}t)$ give the system $\dot{x}=2y+\alpha^{2r-s}x^r$, $\dot{y}=sx^{s-1}$. So the linear coordinates are fixed modulo action of the group $\mathbb{Z}/(2r-s)\,\mathbb{Z}$. This group acts on the remaining terms a_jx^j, and also after further reductions of some of them.

Further we assume the second possibility.

Remark 18. Here we explain why the method of the proof of analyticity of the Bogdanov–Takens prenormal form cannot be adapted in some way to get analyticity of the whole formal normal form.

Consider the reduction of the terms $x^{ms}\partial_x$ by means of L_{X_H}, the commutator with X_H. One can introduce the spaces of orbital changes $(id+Z, \chi)$ (where χ is a function) and of cancelled vector fields $y\widetilde{W}+\alpha(x^s)\,\partial_x$. It is possible to introduce a certain weighted norm in these spaces such that the operator $L_{X_H}^{-1}$ would be bounded and the series bounded in that norm would represent analytic functions.

Unfortunately, the operators $L_{x^j\partial_x}$ are neither bounded relatively to L_{X_H} nor quasinilpotent in this setting.

5.2. *The final formal reduction*

(A) *The general scheme of the reduction.* We use a formal orbital change

$$(X, dt) \to (\exp \hat{Z}(X), (1+\hat{\chi})\,dt),$$

where $X=(x, y)$, $\hat{Z}=\hat{z}_1\partial_x+\hat{z}_2\partial_y$ is a formal vector field, $\exp\hat{Z}=g^1_{\hat{Z}}$ is the phase flow diffeomorphism, and $\hat{\chi}(X)$ is a formal nonzero function. Applying it to a vector field V we obtain the field

$$\mathscr{P}_V(\hat{Z}, \hat{\chi}) = (Ad_{\exp \hat{Z}})_*\,(1+\hat{\chi})\,V.$$

The nonlinear map $(\hat{Z}, \hat{\chi}) \to \mathscr{P}_V(\hat{Z}, \hat{\chi})$ is clearly noninjective. For example, if $\hat{Z}=\hat{\kappa}\cdot V$ ($\hat{\kappa}$ – a formal function), then the map $\exp\hat{Z}$ preserves the phase portrait of V and the field $\mathscr{P}_V(\hat{\kappa}V, 0)$ is parallel to V and equals to $\hat{\eta}V$. In order to avoid this ambiguity, one uses the notion of a bivector field introduced by Bogdanov in [B2].

If $\hat{Z}=\hat{z}_1\partial_x+\hat{z}_2\partial_y$ and $V=P\partial_x+Q\partial_y$, then we define the *bivector field*

$$\hat{Z}\wedge V = \hat{\Omega}\,dx\wedge dy, \tag{5.5}$$

where $\hat{\Omega}=\Omega_V(\hat{Z})=Q\hat{z}_1-P\hat{z}_2$. One can say that $\hat{\Omega}$ measures the component of $\hat{Z}$ transversal to V. If $\Omega_V(\hat{Z})=0$ and V has isolated singularity, then $\hat{Z}=\hat{\kappa}V$ for some series $\hat{\kappa}$.

This shows that the problem of formal reduction of V is equivalent to the problem of formal reduction of terms transversal to V by means of $\hat{\Omega}$. One has the (non-linear) map from $\mathbb{C}[[x, y]]$ to $\mathbb{C}[[x, y]]$

$$\hat{\Omega} \to \mathscr{P}_V(\hat{Z}, 0)\wedge V/dx\wedge dy = \mathscr{L}_V\hat{\Omega}+\cdots, \tag{5.6}$$

where $\mathscr{L}_V$ denotes the linear part (the *bivector homological operator*) and $\hat{Z}$ satisfies the identity $\Omega_V(\hat{Z})=\hat{\Omega}$. $\mathscr{L}_V\hat{\Omega}$ is linear in V (for fixed $\hat{\Omega}$).

Next, one solves the corresponding (nonlinear) equation for $\hat{\Omega}$. Using the quasi-homogeneous filtration one reduces the nonlinear problem to the linear one, i.e., the description of the action of $\mathscr{L}_V$.

We begin with the operator $\mathscr{L}_{X_H}$. One describes a linear space $\mathscr{N}_{X_H}$ (of preliminary bivector normal forms) which is complementary to Im $\mathscr{L}_{X_H}$. This analysis repeats in a sense the calculations from the previous sections.

It turns out that $\mathscr{L}_{X_H}$ has a nontrivial kernel (infinite dimensional). This indicates that $\mathscr{N}_{X_H}$ may not be a good candidate for the space of bivector normal forms. Some $\hat{\Omega}$'s from Ker $\mathscr{L}_{X_H}$ may give contribution to the image of the non-linear operator (5.6) This contribution can be linear (from higher terms of V) or nonlinear.

For the next approximation of the homological operator we use $\mathscr{L}_{X_H+x^r\partial/\partial x}$. We define a linear space $\mathscr{N}_{X_H+x^r\partial/\partial x}$, complementary to Im $\mathscr{L}_{X_H+x^r\partial/\partial x}$. It turns out that in one case, when the number n_0 from Theorem 7 is not an integer, Ker $\mathscr{L}_{X_H+x^r\partial/\partial x}=0$. The standard arguments, involving the quasi-homogeneous filtration, show that $\mathscr{N}_{X_H+x^r\partial/\partial x}$ is a good space for bivector normal forms. It is transversal to the image of the operator (5.6).

If the number n_0 is an integer, then Ker $\mathscr{L}_{X_H+x^r\partial/\partial x}$ is 1–dimensional. In that case one uses the operator $\mathscr{L}_{X_H+x^r\partial/\partial x+a_t x^{r+t}\partial/\partial x}$ for some index t.

(B) *The formula for L_V and the bivector homological equation in the Bogdanov–Takens case.* The linear part of $\mathscr{P}_V(\hat{Z},0)$, $V=P\partial_x+Q\partial_y$, $\hat{Z}=\hat{z}_1\partial_x+\hat{z}_2\partial_y$ is

$$(\dot{\hat{z}}_1-P'_x\hat{z}_1-P'_y\hat{z}_2)\,\partial_x+(\dot{\hat{z}}_2-Q'_x\hat{z}_1-Q'_y\hat{z}_2)\,\partial_y,$$

where the dot denotes the derivative in the direction of V. From this, (5.5), and (5.6) one easily obtains that

$$\mathscr{L}_V\hat{\Omega}=\dot{\hat{\Omega}}-div\,V\cdot\hat{\Omega}. \tag{5.7}$$

If $W=w_1\partial_x+w_2\partial_y$ is the part of the vector field V devoted to killing, then we have the *bivector homological equation*

$$\mathscr{L}_V\hat{\Omega}+h=0,\qquad h=W\wedge V/dx\wedge dy. \tag{5.8}$$

In the Bogdanov–Takens case we have

$$\begin{aligned}V&=[2y+a(x)]\,\partial_x+sx^{s-1}\partial_y,\\ \hat{\Omega}&=sx^{s-1}\hat{z}_1-[2y+a(x)]\,\hat{z}_2,\\ h&=sx^{s-1}w_1-[2y+a(x)]\,w_2,\end{aligned}$$

and the bivector homological equation takes the form

$$\hat{\dot{\Omega}} - a'(x)\,\hat{\Omega} + h = 0. \tag{5.9}$$

What we are looking for is a *bivector normal form space* $\mathcal{N}_V \subset \mathbb{C}[[x, y]]$ complementary to the image of the linear map $\hat{\Omega} \to \hat{\dot{\Omega}} - a'(x)\,\hat{\Omega}$. Because we expect the normal form for V as $(2y + x^r(1 + \sum_{j \in J} a_j x^j))\,\partial_x + sx^{s-1}\partial_y$, the space $\mathcal{N}_V$ is identified with the space $\{x^{r+s-1}\sum_{j \in J} a_j\, x^j\}$ of series. Here the summation runs over the restricted set J of indices. Our problem is to describe the set J. We formulate this statement in a separate lemma.

LEMMA 15. *Assume that* $V = [2y + a(x)]\,\partial_x + sx^{s-1}\partial_y$, $a(x) = x^r + \cdots$ *has the property* $\ker \mathcal{L}_V = 0$.

Then the space $\mathcal{N}_V$ *can be chosen in the form* $\{x^{r+s-1}\sum_{j \in J} a_j x^j\}$. *In this case the formal orbital normal form for* V *is* $(2y + x^r(1 + \sum_{j \in J} a_j x^j))\,\partial_x + sx^{s-1}\partial_y$.

(C) *The image and the kernel of* $\mathcal{L}_{X_H}$. We have

$$\mathcal{L}_{X_H}\hat{\Omega} = \hat{\dot{\Omega}} = 2y\hat{\Omega}'_x + sx^{s-1}\hat{\Omega}'_y \tag{5.10}$$

and Eq. (5.8) is replaced by

$$2y\hat{\Omega}'_x + sx^{s-1}\hat{\Omega}'_y + h = 0. \tag{5.11}$$

Because L_{X_H} acts quasihomogeneously, we investigate its restrictions to the finite dimensional spaces of $\hat{\Omega} = \Omega$ with fixed quasihomogeneous degree.

Let this degree be $2(ms + i)$, $i = 0, 1, \ldots s - 1$; i.e.,

$$\Omega = \omega_0 x^{sm+i} + \omega_1 x^{s(m-1)+i} y^2 + \cdots + \omega_m x^i y^{2m}.$$

The function h in (5.11) is of degree $2(ms + i) + s - 2$; i.e.,

$$h = h_0 x^{ms+i-1} y + \cdots + h_m x^{i-1} y^{2m+1}$$

if $i \neq 0$ and

$$h = h_0 x^{s(m-1)+s-1} y + \cdots + h_{m-1} x^{s-1} y^{2m-1}$$

if $i = 0$.

The matrix of $\mathscr{L}_{X_H}$ expressed in these bases is equal to 2 times

$$\begin{bmatrix} ms+i & s & 0 & \dots & 0 & 0 & 0 \\ 0 & ms-s+i & 2s & \dots & 0 & 0 & 0 \\ 0 & 0 & ms-s+i & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2s+i & ms-s & 0 \\ 0 & 0 & 0 & \dots & 0 & s+i & ms \\ 0 & 0 & 0 & \dots & 0 & 0 & i \end{bmatrix}, \tag{5.12}$$

where there is no last row in the above matrix for $i=0$.

We see that if $i \neq 0$, then it is an invertible matrix. If $i=0$, then we get m rows and $m+1$ columns; so, it has a 1-dimensional kernel. This kernel equals $(x^s-y^2)\,\mathbb{C}$ (because then $\dot{\Omega}=0$).

If the quasihomogeneous degree of Ω is $2(ms+i)+s$ with

$$\begin{aligned} \Omega &= \omega_0 x^{ms+i} y + \cdots + \omega_m x^i y^{2m+1} \\ h &= h_0 x^{(m+1)\,s+i-1} + \cdots, && i \neq 0 \\ h &= h_0 x^{ms+s-1} + \cdots, && i=0 \end{aligned}$$

then the corresponding matrix is

$$\begin{bmatrix} s & 0 & 0 & \dots & 0 & 0 \\ 2(ms+i) & 3s & 0 & \dots & 0 & 0 \\ 0 & 2(ms-s+i) & 5s & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & (2m-1)\,s & 0 \\ 0 & 0 & 0 & \dots & 2(s+i) & (2m+1)\,s \\ 0 & 0 & 0 & \dots & 0 & 2i \end{bmatrix}, \tag{5.13}$$

where again the last row is absent for $i=0$.

We see that if $i \neq 0$, then we have an $(m+1)\times m$ matrix and the image of the corresponding operator is m-dimensional (and of codimension one). We project it to the space generated by the monomials $x^\alpha y^{\beta+1}$.

If $i=0$, then the matrix (5.13) (without the last row) is invertible.

Remark 19. Sadovski in [S2] considered the case $r=2s$ and wrote down an equation analogous to (5.11).

We have proven the following lemma, whose first part repeats the result of the previous section.

LEMMA 16. (i) *Equation* (5.11) *has solution* Ω *for any* h *of the form* $y\tilde{h}(x, y)+x^{s-1}\tilde{h}(x^s)$. *This means that the space* $\mathcal{N}_{X_H}=\{x^{s-1}\sum_{j\neq o(mod\ s)} a_j x^j\}$ *is complementary to Im* $\mathcal{L}_{X_H}$.

(ii) $\ker \mathcal{L}_{X_H}=\mathbb{C}[[y^2-x^s]]$. *This means that for any two solutions* Ω_1, Ω_2 *of* (5.11) *the difference* $\Omega_1-\Omega_2$ *is a function of* H.

(D) *Reduction by means of* $\mathcal{L}_{X_H+x^r\partial/\partial x}$. As in the previous point the problem reduces itself to a solution of the equation

$$(\tilde{\omega}+\tilde{\tilde{\Omega}}) \dot{-} div\, V\cdot(\tilde{\omega}+\tilde{\tilde{\Omega}})+h=0,$$

where $\tilde{\omega}$, $\tilde{\tilde{\Omega}}$, h are quasihomogeneous, $\partial\tilde{\omega}/\partial X_H\equiv 0$, and $V=X_H+x^r\partial_x$.

Thus $\Omega=\tilde{\omega}+\tilde{\tilde{\Omega}}=\omega_{-1}(x^s-y^2)^n+\omega_0 x^{sm+i}y+\cdots+\omega_m x^i y^{2m+1}$ and the quasihomogeneous part of $\mathcal{L}_{X_H+x^r\partial/\partial x}\,\Omega$ equals

$$[x^r\tilde{\omega}'_x-rx^{r-1}\tilde{\omega}]+[2y\tilde{\tilde{\Omega}}'_x+sx^{s-1}\tilde{\tilde{\Omega}}'_y]. \tag{5.14}$$

The quasihomogeneity implies

$$sn+r-1=s(m+1)+i-1$$

and $h=h_0x^{s(m+1)+i-1}+\cdots$. The corresponding matrix takes the form

$$\mathcal{A}=\begin{pmatrix} A & B\\ 0 & D\end{pmatrix}, \tag{5.15}$$

where A and D are equal to respectively

$$\begin{bmatrix}
(ns-r)\binom{n}{0} & s & 0 & \dots & 0 & 0\\
-(ns-s-r)\binom{n}{1} & 2(ms+i) & 3s & \dots & 0 & 0\\
(ns-2s-r)\binom{n}{2} & 0 & 2((m-1)\,s+i) & \dots & 0 & 0\\
\dots & \dots & \dots & \dots & \dots & \dots\\
\dots & \dots & \dots & \dots & 2(r+s) & (2n-1)\,s\\
\pm r\binom{n}{n} & 0 & 0 & \dots & 0 & 2r
\end{bmatrix}$$

$$\begin{bmatrix} 2((m-n)\,s+i) & (2n+3)\,s & 0 & \dots & 0 & 0 \\ 0 & 2((m-n-1)\,s+i) & (2n+5)\,s & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2(s+i) & (2m+1)\,s \\ 0 & 0 & 0 & \dots & 0 & 2i \end{bmatrix}.$$

LEMMA 17. *The determinant of the matrix $\mathscr{A}$ equals*

$$(5.16)\qquad \begin{aligned} &(-2s)^{m-n+1}\left(-\frac{i}{s}\right)_1^{m-n+1} \\ &\times s(-2s)^n(-r-1)\left(-\frac{r}{s}-\frac{3}{2}\right)_1^{n-1}\left[n-\frac{r}{s}-\frac{1}{2}\right], \end{aligned}$$

where $(a)_1^0 = 1$, $(a)_1^k = a(a-1)\cdots(a-k+1)$.

Proof. The first factor in the above formula is $\det D$. The second factor, i.e., $\det A$, is calculated using the expansion with respect to the first column (like in the calculation of the determinant of the characteristic polynomial for a higher order differential linear equation.) We have used the formulas

$$\sum_{k=0}^{n}\binom{n}{k}(a)_1^k(b)_1^{n-k} = (a+b)_1^n, \qquad \sum_{k=0}^{n} k\binom{n}{k}(a)_1^k\,(b)_1^{n-k} = na(a+b-1)_1^{n-1}$$

following from the identity

$$\begin{aligned} \sum_{k=0}^{n}\binom{n}{k}\mu^k(a)_1^k\,(b)_1^{n-k}\Big|_{\mu=1} &= \left(\mu\frac{\partial}{\partial\tau}+\frac{\partial}{\partial\nu}\right)^n \tau^a\nu^b\Big|_{\tau=\nu=\mu=1} \\ &= \left(\frac{\partial}{\partial\alpha}\right)^n(\alpha+\beta)^a(\alpha-\beta)^b\Big|_{\alpha=1,\,\beta=0} \\ &= (a+b)_1^n \end{aligned}$$

and its derivative with respect to μ, equal to $n(\mu\frac{\partial}{\partial\tau}+\frac{\partial}{\partial\nu})^{n-1}\frac{\partial}{\partial\tau}\tau^a\nu^b$. We have applied them with $a=-(r+1)/s$, $b=-1/2$. ∎

(E) *The final step.* Lemma 17 shows that $\det \mathscr{A} = 0$ iff $n = n_0 + 1$, where

$$n_0 = \frac{r}{s} - \frac{1}{2}$$

is the number defined in Theorem 7.

If n_0 is not an integer, then $\ker \mathscr{L}_{X_H + x^r \partial/\partial x} = 0$ and all the terms $x^{r+s-1} \cdot x^{(n-1)s}$ are in $\operatorname{Im} \mathscr{L}_{X_H + x^r \partial/\partial x}$; these terms correspond to $x^{r+(n-1)s} \partial_x$ in V. Thus $\mathscr{N}_{X_H + x^r \partial/\partial x} = \{x^{r+s-1} \sum_{j \neq 0, -r \pmod{s}} a_j x^j\}$. By Lemma 15 we have case (ii) of Theorem 7.

If n_0 is an integer, then the term $c_{n_0 s} x^{r+n_0 s} \partial_x$ remains in the normal form for V. Here we have two possibilities: either all the further terms vanish or there is a term $c_{j_0} x^{r+j_0} \partial_x \neq 0$, $j_0 \neq 0, -r \pmod{s}$.

In the first possibility we have

$$(2y + x^r(1 + \lambda x^{n_0 s}))\, \partial_x + s x^{s-1} \partial_y,$$

i.e., case (iii) of Theorem 7.

In the second case we use the bivector field $\tilde{\omega} = \omega_{-1}(x^s - y^2)^{n_0+1}$, applied to $x^{r+j_0} \partial_x$, to reduce the term

$$x^{r+j_0+n_0 s} \partial_x.$$

This gives case (iv) from Theorem 7.

Because there are no more possibilities of reductions, Theorem 7 is complete. ∎

Remark 20. From Loray's proof in [L3] it follows that the reductions performed in points (D) and (E) can be made analytic (provided that we start from an analytic form from Lemma 16). Therefore the only place where the analyticity fails is the reduction by means of $\mathscr{L}_{X_H}$.

6. PROOF OF THEOREM 8

(A) Theorem 8 says how to recognize the type of monodromy group from the formal orbital normal form. On the basis of this classification the cases when the formal normal form is the same as the analytic one are extracted. Below we successively associate to given classes of groups the corresponding formal normal forms and at the end of this section we discuss the problem of analyticity of these forms.

The projective monodromy group G associated with a germ of a nilpotent analytic vector field is either finite, exceptional abelian, solvable nonabelian, or non-solvable.

(B) From Corollary 3 (in Section 5.1) it follows that, if G is finite, then we have case (i) of Theorem 7. Here G is analytically isomorphic to a finite subgroup of the group of rotations.

(C) If G is exceptional abelian, then Corollary 3 implies that $n_0 \in \mathbb{Z}$ (and hence $s = 2k$). By Theorem 2(a) (Section 1) the generators of G are simultaneously formally equivalent to

$$f_{1,2} = \lambda g_w^{t_{1,2}},$$

where $\lambda = e^{\pi i/k}$ and $w = w_{p,\mu}$. One can see also from the proof of Lemma 14 that $p = 2n_0 k$ and $\lambda^p = 1$ (see Corollary 3).

LEMMA 18. *The exceptional abelian monodromy group corresponds to case* (iii) *from Theorem* 7.

Proof. We have to calculate the monodromy group associated with the normal form from case (iii) of Theorem 7. However, for this purpose that normal form is not the best choice. We shall choose the form from Remark 8 (Section 2); i.e., $X_H + x^{k-1} H^{n_0} (1 + \mu H^{n_0})^{-1} E_H$.

Here we can integrate the analogue of system (5.4), describing the phase portrait after resolution of the singularity. In the variables $u = y/x^k$, $H = y^2 - x^{2k} = (u^2 - 1)\, v^{2k}$ we have

$$\frac{dH}{du} = \frac{2H^{n_0+1}}{(1+\mu H^{n_0})(1-u^2)}. \tag{6.1}$$

Now it is clear that, when we parametrize the disk transversal to the exceptional divisor E by H, then the monodromy maps are the flow maps of the vector field $\dot{H} = H^{n_0+1}/(1+\mu H^{n_0})$. ∎

Remark 21. Equation (6.1) can be integrated. Its first integral takes the so-called generalized Darboux form

$$F = -H^{-n_0} + \mu n_0 \ln H - 2n_0 \ln[(1+u)/(1-u)].$$

(D) If G is solvable nonabelian, then it is formally conjugated to

$$f_1(z) = e^{i\pi}, \qquad f_2(z) = e^{2\pi i k/(2k+1)} z / \sqrt[p]{1 - pz^p}\,, \tag{6.2}$$

$s = 2k+1$, or to

$$f_0(z) = e^{2\pi i/k} z, \qquad f_1(z) = e^{\pi i/k} z / \sqrt[p]{1 - pz^p}\,, \tag{6.3}$$

$s=2k$ (see Theorem 2(b) in Section 1 and [CM]). It is also known that in the case $s=2k+1$ this group is typical (see Theorem 3 in Section 1).

Let $s=2k+1$. The typicality of G means that $p \neq 0 \pmod{s}$: $\Lambda_G=\{e^{\pi i l/2s}\colon l=0,1,\ldots\}$ and $\Lambda_G^p \neq \{\pm 1\}$ (see Remark 1 in Section 1).

This group is the holonomy group of the singularity $J_{r,0}^{2k+1}$; i.e., case (ii) of Theorem 7 with the formal functional invariant $\phi \equiv 0$. We have $p=2(r-k)-1$ (see Corollary 3 in Section 5.1) and the condition $r \neq 0 \pmod{s}$ from Theorem 7 is equivalent to the condition $p \neq 0 \pmod{s}$ (i.e., typicality of G).

Let $s=2k$. The case $p=0 \pmod{s}$ is excluded because then G is abelian: $\Lambda_G=\{e^{\pi i l/k}\colon l=0,1,\ldots\}$ and $\Lambda_G^p=\{1\}$ (see Remark 1). Moreover, because G is typical, we have $p \neq 0 \pmod{k}$.

This group is the holonomy group of the singularity $J_{r,0}^{2k}$. We have $p=r-k$ (see Corollary 3) and $n_0=r/s-1/2 \notin \frac{1}{2}\mathbb{Z}$ (for $p \neq 0 \pmod{k}$). Therefore, we have case (ii) of Theorem 7 with the formal functional modulus $\phi \equiv 0$. The conditions $r \neq 0 \pmod{s}$, $n_0 \notin \mathbb{Z}$ from Theorem 7 are equivalent to the typicality of solvable G.

(E) Because all solvable groups defined by (1.9)–(1.10) and by (1.11)–(1.12) are formally equivalent to the above ones, the other normal forms $J_{r,\phi}^s$, $\phi \not\equiv 0$ or $\phi \neq c_{n_0 s}x^{n_0 s}$ have nonsolvable holonomy group.

(F) In this point we prove analyticity of the normal form from Theorem 7 in the cases of solvable nonabelian and finite holonomy group.

Take a germ of analytic vector field V with solvable nonabelian monodromy group G. Because G is typical, it is analytically conjugated with its standard form (6.2) or (6.3) (see Theorem 3(iii) in the Introduction). On the other hand the polynomial field $J_{r,0}^s$ has the monodromy equal to (6.2) or (6.3) and is analytic. By Theorem 1(b) the analytic conjugation of the holonomies of V and $J_{r,0}^s$ implies analytic orbital conjugation of these vector fields.

The analyticity of the formal normal form $J_{\infty,0}^s$ (i.e., the Hamiltonian field) can be proven in three ways.

One way uses the rigidity of a finite group (Theorems 3(i) and 1(b)). This proof is the same as in the solvable case.

One can also use the following theorem of Mattei and Moussu [MM]: if a singular point of a planar holomorphic foliation has only finite number of separatrices (i.e., analytic invariant curves through the singularity) and any other leaf does not accumulate at the singular point, then the foliation has a local analytic first integral.

In the generalized cusp case we have only finite number of separatrices (one or two) and the finiteness of the holonomy group implies that the leaves are separated from the singular point. So, the assumptions of Mattei–Moussu's theorem hold and there is an analytic first integral. It has the form $y^2 = x^s + \cdots$. This function is equal to $Y^2 - X^s$ with analytic $X = x + \cdots$, $Y = y + \cdots$ (see [AVG]).

Below we propose another proof based on the following lemma.

LEMMA 19. *The case with a finite monodromy group occurs only when the Bogdanov–Takens normal form contains only powers of x^s; i.e., $X_H + (\sum b_m x^{sm-1})\, E_H$.*

Proof. Indeed, repeating the arguments from Section 5.1, we can show that any vector field of the form

$$X_H + \left(\sum_{m<M} c_{sm-1} x^{sm-1} + c_{r-1} x^{r-1} \right) E_H, \qquad Ms < r < (M+1)\, s$$

with $c_{r-1} \neq 0$ has either nonabelian or exceptional abelian monodromy group.

If $s = 2k+1$ is odd, then in the coordinates (5.1) we get

$$\frac{dv}{du} = A_1(u)\, v + \sum c_{sm-1} A_{sm-1}(u)\, v^{2(sm-k)} + c_{r-1} A_{r-1}(u)\, v^{2(r-k)}.$$

The fact that the terms $x^{sm-1} E_H$ can be removed means that, after applying certain change of the form $v \to v_1 = v + \sum_{m<M} B_m(u)\, v^{2(sm-k)}$, the $2(sM-k)$th part of the latter equation disappears. There remains the term $c_{r-1} A_{r-1}(u)\, v^{2(r-k)}$ and other terms which depend on the coefficients c_{sm-1}. However, the latter terms arise only from the action of the change $v \to v_1$ on the term with $v^{2(r-k)}$ and are of higher degree.

Therefore, $dv_1/du = A_1(u)\, v_1 + c_{r-1} A_{r-1}(u)\, v_1^{2(r-k)} + \cdots$ and the $2(r-k)$th jet of the monodromy group contains the map $v \to \mu_1 v + const \cdot c_r v^{2(r-k)} + \cdots$, where $const \neq 0$. This shows that the monodromy group is nonabelian.

The case with even s is analogous. ∎

If the (analytic) Bogdanov–Takens normal form is equal to $X_H + x f(x^s)\, E_H$, then repeating the proof of Lemma 13 (from Section 5.1) we see that it has analytic first integral $y^2 - x^s + \cdots$, which is analytically equivalent to $Y^2 - X^s$. Therefore the form $J^s_{\infty,0}$ is analytic.

This completes the proof of Theorem 8. ∎

7. PROOF OF THEOREM 9

In Theorem 8 we have given an interpretation of the first part of the formal orbital normal form in terms of the holonomy group. Theorem 9 is about the interpretation of certain higher order terms from the formal orbital normal form in terms of singularities of the resolved vector field and in terms of the holonomy. First we interpret the first coefficient of the formal functional modulus $\phi(x)$ for $s=3$ and then we interpret the coefficient $c_{n_0 s}$ in the exceptional abelian case.

(A) In the case $s=3$ we can prove the nonsolvability in a way other than in Section 6. Note that solvability of G means that the resolved vector field near the singular point p_0: $u=1, v=0$ is formally (and analytically) linearizable (see Theorem 5(a) in the introduction). Here we show that we can express the obstacles to the linearization in terms of the coefficients of the expansion of the functional modulus ϕ from Theorem 7.

Recall that by Theorem 7 the formal normal form is one of the two types: either ($r=3m+1$)

$$\dot{x}=2[y+x^{3m+1}(1+c_1x+c_4x^4+c_7x^7+\cdots)],$$
$$\dot{y}=3[x^2+x^{3m}y(1+c_1x+c_4x^4+c_7x^7+\cdots)],$$

or ($r=3m+2$)

$$\dot{x}=2[y+x^{3m+2}(1+c_2x^2+c_5x^5+c_8x^8+\cdots)],$$
$$\dot{y}=3[x^2+x^{3m+1}y(1+c_2x^2+c_5x^5+c_8x^8+\cdots)].$$

In the coordinates (5.1), when we additionally perform the change $u=1+z$, we get the equation

$$\frac{dv}{dz}=A_1(z)\,v+A_r(z)\,v^{2r-2}+d_tA_t(z)\,v^{2t-2}+\cdots,$$

where $A_1=-\frac{1+3z}{6z(1+z)}$, $A_r=\frac{(1+z)^{r-2}}{-6z}$, $A_t=\frac{(1+z)^{t-2}}{-6z}$, and d_t, $t\neq 0, r \pmod 6$ is the first non-zero coefficient of the Taylor expansion of ϕ (d_t is the first of c_j's).

Recall also that the formal orbital normal form of a $1:-6$ resonant saddle is the following

$$\frac{d\tilde{v}}{d\tilde{z}}=-\frac{\tilde{v}}{6\tilde{z}}(1+e_1(\tilde{v}^6\tilde{z})+e_2(\tilde{v}^6\tilde{z})^2+\cdots).$$

Here $\tilde{z}=z(1+\cdots)$, $\tilde{v}=v(1+\cdots)$.

Remark 22. The formal orbital normal form is simpler. There remain only two coefficients e_j, namely $e_k=1$ and e_{2k}, where the latter plays a role of the formal invariant. In our situation the range of changes is smaller. So, we get the richer normal form near p_0.

We know that if $d_t = d_{t+3} = \cdots = 0$, then all the coefficients e_j vanish. So, firstly we should apply a change which cancels the term $A_r v^{2r-2}$.

Denote $q = 2r-2$. The change is the following

$$v = w[1+B(z)\, w^{q-1}]^{-1/(q-1)},$$

where the function B satisfies the equation $B' + (q-1)\, A_1 B + (q-1)\, A_r = 0$ with the solution

$$B = \frac{q-1}{6}\, z^{(q-1)/6}(1+z)^{(q-1)/3} \int^z \tau^{-(q+5)/6}(1+\tau)^{(q-4)/6}\, d\tau = 1 + O(z).$$

The transformed equation becomes equal to

$$\frac{dw}{dz} = A_1 w + d_t A_t w^{2t-2}[1 + Bw^{2r-3}]^{2(r-t)/(2r-3)} + \cdots.$$

(Note that if $q = 6p+1$, then the expansion of $B(z)$ contains the logarithmic term

$$Mz^{\nu} \log z.$$

Namely the coefficients M before such logarithms will play a role of the saddle quantities. Above we have $q \neq 6p+1$, but we will use the coefficient M below.)

Expanding the last expression in the equation for w, we get the term $d_t A_t w^{2t-2}$ (which we reduce in the same way as $A_r v^{2r-2}$) and the next term of the form

$$A_t B w^{2r+2t-5}.$$

It turns out that, in our situation, $2r+2t-5 = 1 \pmod 6$ always either $r = 3m+1$, $t = 3n+2$ in the case $J^3_{3m+1,\phi}$, or $r = 3m+2$, $t = 3n+1$ in the case $J^3_{3m+2,\phi}$.

Repeating the proof of cancelling the A_r, with $q = 2r+2t-5$, we get a new $B(z)$, i.e.,

$$B_1 = \text{const} \cdot z^{(r+t-3)/3}(1+z)^{2(r+t-3)/3} \int^z \tau^{(3-r-t)/3}(1+\tau)^{2(3-r-t)/3} A_t(\tau)\, B(\tau)\, d\tau.$$

We are interested in the coefficient before τ^{-1} in the expansion of the subintegral function.

LEMMA 20. *This coefficient is nonzero.*

Proof. Let us write this subintegral function explicitly (up to a constant)

$$\tau^{-(2t+3)/6}(1+\tau)^{(t-3)/3}\int^{\tau}\sigma^{-(2r+3)/6}(1+\sigma)^{(r-3)/3}\,d\sigma.$$

In the case $J^3_{3m+1,\phi}$, $t=3n+2$, the expansion of $(1+\tau)^{\alpha}$ and the integration give

$$\tau^{-m-n-1}\left[\sum_{k=0}^{\infty}\binom{n-1/3}{k}\tau^k\right]\cdot\left[\sum_{l=0}^{\infty}\binom{m-2/3}{l}\frac{\tau^l}{l-m+1/6}\right]$$

and its residuum at $\tau=0$ is

$$\sum_{l=0}^{m+n}\binom{n-1/3}{m+n-l}\binom{m-2/3}{l}\frac{1}{l-m+1/6}.$$

It can be rewritten in the form

$$\begin{aligned}\binom{n-1/3}{m+n}&\left[\sum_{l=0}^{m+n}\binom{m+n}{l}\frac{(-1)^l}{l-m+1/6}\right]\\ &=\binom{n-1/3}{m+n}\mathscr{P}\int_0^1 t^{-m-5/6}(1-t)^{m+n}\,dt\\ &=\binom{n-1/3}{m+n}B(-m+1/6,m+n+1)\neq 0.\end{aligned}$$

The latter integral is divergent and $\mathscr{P}\int$ means its regularization. One uses the analytic continuation of the beta function $B(a,b)$.

In the case $J^3_{3m+2,\phi}$, the calculations are quite similar and one obtains the coefficient

$$\binom{n-2/3}{m+n+1}B(-m-1/6,m+n+1)\neq 0. \quad \blacksquare$$

Remark 23. In the proof of Proposition 1 we have applied two local (near p_0) changes and detected the resonant term. Unfortunately, this works only in the case $s=3$. If $s>3$, then we need many more local changes and the calculations become very complicated. We were not able to complete them.

Remark 24. Note that if the formal functional modulus $\phi \not\equiv 0$, then only the first nonvanishing coefficient a_l of the expansion of ϕ is important in nonsolvability

$$f_0^{[6]}(z) = z + a_l z^{6q+1} + \cdots.$$

Maybe the other coefficients of the expansion of ϕ are responsible for new relations in group G. Something like this: $f_1 \circ f_2 \circ f_1 \circ f_2 \cdots \circ f_2^{-1}(z) = z + a_n z^{r(n)} + \cdots$.

(B) Let $n_0 \in \mathbb{Z}$, $s = 2k$. Then we can take the vector field in the form $X_H + x^{k-1} H^{n_0} E_H + \cdots$. Its $(2n_0+1)\, s$–jet is integrable: $dH/du = 2H^{n_0+1}/(1-u^2)$ in the separating variables from the proof of Lemma 18. It is clear that the saddle $u = 1$, $v = 0$ in the divisor of resolution is not linearizable.

In case (iii) we get $dH/du = 2H^{n_0+1}/[(1+\mu H^{n_0})(1-u^2)]$. Here $\mu = \text{const} \cdot c_{n_0 s}$. It is clear that $c_{n_0 s}$ plays a role of the formal invariant of the saddle $u = 1$, $v = 0$.

Theorem 9 is complete. ∎

ACKNOWLEDGMENTS

The main problem for us during the work on this subject was our isolation. In this context we had very fruitful discussions with Frank Loray and Rafik Meziani during their visits in Warsaw. We thank Anton Sadovski for his remarks and Yuri Il'yashenko for critical comments about previous versions of the text. We thank the referee, whose clever suggestions helped to simplify the proof of Theorem 6 and avoid some mistakes.

REFERENCES

[A] V. I. Arnold, "Geometrical Methods in the Theory of Ordinary Differential Equations," Springer-Verlag, Berlin/Heidelberg/New York, 1983.

[AVG] V. I. Arnold, A. N. Varchenko, and S. M. Gusein-Zade, "Singularities of Differentiable Mappings," Monographs in Mathematics, Vol. 82, Birkhäuser, Boston, 1985. [In Russian; Vol. 1, Nauka, Moscow, 1982.]

[BS] A. Baider and J. Sanders, Further reduction of the Bogdanov–Takens normal form, *J. Differential Equations* **99** (1992), 205–244.

[BM] M. Berthier and R. Moussu, Réversibilité et classification des centres nilpotents, *Ann. Inst. Fourier* **44** (1994), 465–494.

[B1] R. I. Bogdanov, Versal deformation of singular point of vector field on a plane in the case of zero eigenvalues, *Trans. Petrovski Sem.* **2** (1976), 37–65. [In Russian]

[B2] R. I. Bogdanov, Local orbital normal forms of vector fields on a plane, *Trans. of the Petrovski Sem.* **5** (1979), 50–85. [In Russian]

[Br] N. N. Brushlinskaya, The finiteness theorem for families of vector fields in a neighborhood of singular point of the Poincaré type, *Funct. Anal. Appl.* **5** (1971), 10–15. [In Russian]

[CM] D. Cerveau and R. Moussu, Groupes d'automorphismes de $(\mathbf{C}, 0)$ et équations différentielles $ydy + \cdots = 0$, *Bull. Soc. Math. France* **116** (1988), 459–488.

[EISV] P. M. Elizarov, Yu. S. Il'yashenko, A. A. Shcherbakov, and S. M. Voronin, Finitely generated groups of germs of one-dimensional conformal mappings and invariants for complex singular points of analytic foliations of the complex plane, *in* "Nonlinear Stokes Phenomena" (Yu. S. Il'yashenko, Ed.), Advances in Sov. Math., Vol. 14, pp. 57–105, Amer. Math. Soc., Providence, RI, 1993.

[L1] F. Loray, "Feuilletages holomorphes à holonomie résoluble," Thèse, Université de Rennes, 1994.

[L2] F. Loray, Rigidité topologique pour des singularités de feuilletages holomorphes, *Publ. Inst. Rech. Math. Rennes* **37** (1995), 1–33.

[L3] F. Loray, Réduction formelle des singularités cuspidales de champs de vecteurs analytiques, *J. Differential Equation* **158** (1999), 152–173.

[LM] F. Loray and R. Meziani, Classification de certains feuilletages associés à un cusp, *Bol. Soc. Bras. Mat.* **25** (1994), 93–106.

[MM] J. F. Mattei and R. Moussu, Holonomie et intégrales premiéres, *Ann. Sci. Ecole Norm. Sup.* **13** (1988), 469–523.

[M1] R. Moussu, Symmétrie et forme normale des centers et foyers dégénérées, *Ergodic Theory Dynam. Systems* **2** (1982), 241–251.

[M2] R. Moussu, Holonomie évanescente des équations différentielles dégénérées transverses, *in* "Singularities and Dynamical Systems" (S. N. Pnevmaticos, Ed.), North-Holland, Amsterdam, 1985, pp. 161–173.

[P] J. Pöschel, On elliptic lower dimensional tori in Hamiltonian systems, *Math. Z.* **202** (1989), 559–608.

[R] J.-P. Ramis, Confluence et résurgence, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **36** (1989), 706–716.

[S1] A. P. Sadovski, On the problem of distinguishing center and focus, *Differential Equations* **12** (1976), 1237–1246. [In Russian]

[S2] A. P. Sadovski, Formal orbital normal forms of two-dimensional systems with non-zero linear part, *Differential Equations* **21** (1985), 1509–1516. [In Russian]

[S3] A. P. Sadovski, Remarks on existence of holomorphic integral in the critical case of double zero root, *Differential Equations* **23** (1987), 355–357. [In Russian]

[St] E. Stróżyna, The analytic and formal normal forms for the nilpotent singularity. The generalized saddle-node case, preprint, University of Warsaw, 1999.

[T] F. Takens, Singularities of vector fields, *Publ. Math. IHES* **43** (1974), 47–100.