

## Normal forms, symmetry and linearization of dynamical systems

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1998 J. Phys. A: Math. Gen. 31 5065

(<http://iopscience.iop.org/0305-4470/31/22/009>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.122

The article was downloaded on 02/06/2010 at 06:54

Please note that [terms and conditions apply](#).

# Normal forms, symmetry and linearization of dynamical systems

Dario Bambusi<sup>†¶</sup>, Giampaolo Cicogna<sup>†+</sup>, Giuseppe Gaeta<sup>§\*</sup> and Giuseppe Marmo<sup>||‡</sup>

<sup>†</sup> Dipartimento di Matematica, Università di Milano, via Saldini 50, 20133 Milano, Italy

<sup>‡</sup> Dipartimento di Fisica, Università di Pisa, Piazza Torricelli 2, 56126 Pisa, Italy

<sup>§</sup> Department of Mathematics, Loughborough University, Loughborough LE11 3TU, UK

<sup>||</sup> Dipartimento di Scienze Fisiche, Università di Napoli, and INFN, Sezione di Napoli, Mostra d'Oltremare, 80123 Napoli, Italy

Received 3 July 1997

**Abstract.** We discuss how the presence of a suitable symmetry can guarantee the perturbative linearizability of a dynamical system—or a parameter-dependent family—via the Poincaré normal form approach. We discuss this formally at first, and later pay attention to the convergence of the linearizing procedure. We also discuss some generalizations of our main result.

## 1. Introduction

It is well known that the same procedure—based on the formal series of polynomial changes of coordinates—devised by Poincaré [1–4] to integrate linearizable dynamical systems<sup>††</sup> in the neighbourhood of a fixed point, can also be used to normalize, again in the neighbourhood of a fixed point, nonlinearizable dynamical systems, i.e. systems whose linearization at the fixed point present resonances. This fact suggests that the Poincaré procedure does not take full advantage of the peculiar nature of (locally) linearizable system, so that it is not impossible to obtain, in this specific case, some improvement over the general theory of Poincaré–Dulac normal forms.

It was recently proposed [5] that the linearizability of a dynamical system can be analysed, and asserted under certain conditions *non-perturbatively* by considering the *symmetries* which are associated with the linearity of the system in suitable coordinates.

Here we want to show how these symmetry properties come into play—to ensure linearizability of the system (by means of the formal, or possibly convergent, series of near-identity changes of coordinates)—in the framework of the *perturbative* theory, i.e. in the theory of Poincaré–Dulac normal forms [1, 2]. Thus, although we investigate the same kind of question as in [5], we operate in quite a different framework, and we can make little

¶ E-mail address: bambusi@vmimat.mat.unimi.it.

+ E-mail address: cicogna@ipifdpt.difi.unipi.it

\* Present address: Dipartimento di Fisica, Università di Roma, 00185 Roma, Italy. E-mail address: gaeta@roma1.infn.it

‡ E-mail address: gimarmo@napoli.infn.it

††By dynamical system, we mean a polynomial ODE in  $R^n$ , or equivalently a polynomial vector field on  $R^n$ . Similarly, by vector field we mean an analytic one.

use of the results obtained there, as will be clear from the following; we discuss the relation between [5] and the present work, i.e. between the global and the perturbative approach, in the appendix.

In order to discuss the linearizability of a system in the perturbative approach, we have naturally to consider the case of normal forms (NFs) in the presence of symmetry. More specifically, it turns out that we have to consider NFs in the presence of *nonlinear* symmetry: indeed, the symmetries associated with the linearizable nature of the system may be linear only in the coordinates in which the dynamical system is linear, i.e. when the problem is already solved; moreover, such coordinates could be defined only locally in a neighbourhood of the origin. Thus, rather than employing the classical theory of NFs in the presence of linear symmetries [2, 6–8], we will use recent results [9, 10], which deal with the general (i.e. nonlinear) case.

Needless to say, systems which can be linearized—albeit only locally—are highly non-generic, and correspond to ones that can be exactly solved locally. This fact shows at the same time the limited application of our result, as it cannot deal with generic systems, and its interest, as it deals with systems which are special but also specially interesting, both in themselves and as starting points for a perturbative analysis of more general ones.

It should be mentioned that the connection between ‘suitable’ symmetries and the linearity of the NF has already been remarked—albeit *en passant*—in remark 4 of [9]; we want to discuss this in more detail for two reasons: on the one hand, for its practical relevance; and on the other because it shows how the consideration of nonlinear symmetries in NF theory really improves the results that can be obtained by considering only linear symmetries.

## 2. The main result

Let us consider a vector field  $X$ , which in the coordinates  $(y_1, \dots, y_n)$  on  $R^n$  is simply given by

$$X = \sum_{i=1}^n A_{ij} y_j \frac{\partial}{\partial y_i} \quad (1)$$

where  $A$  is a real  $(n \times n)$  matrix. This vector field obviously commutes with all the vector fields

$$Y^{(k)} = \sum_{i=1}^n [A^k]_{ij} y_j \frac{\partial}{\partial y_i} \quad (2)$$

associated to the (non-negative) integer powers of the matrix  $A$ . In particular, whatever the matrix  $A$ ,  $X$  commutes with the vector field  $S \equiv Y^{(0)}$  which generates the dilations in  $R^n$

$$S = \sum_{i=1}^n y_i \frac{\partial}{\partial y_i} \quad (3)$$

and it is easy to see that, conversely, the only vector fields commuting with  $S$  are the linear ones. This observation is, of course, completely trivial; nevertheless, it is extremely useful in studying the linearizability of nonlinear systems.

Let us now consider a (formal) near-identity (this means that  $(D\phi)(0) = I$ ) nonlinear change of coordinates

$$y_i = \phi_i(x, \dots, x_n). \quad (4)$$

Under this, the linear dynamical system

$$\dot{y}_i = A_{ij}y_j \tag{5}$$

is changed, in general, into a nonlinear system

$$\dot{x}^i = f^i(x) \tag{6}$$

where it can be immediately seen that the functions  $f^i(x)$  are given by

$$f^i(x) = [J^{-1}(x)]_{ij}A_{jk}\phi_k(x) \tag{7}$$

where  $J$  is the Jacobian of the coordinate transformation

$$J_{ij} = \frac{\partial y_i}{\partial x_j} \equiv \frac{\partial \phi_i}{\partial x_j}. \tag{8}$$

Notice that, as we assumed (4) to be a near-identity transformation, the inverse of the Jacobian exists, at least in some neighbourhood of the origin, so that (7) makes sense. Again by the near-identity of (4), we are guaranteed that  $(Df)(0) = A$ .

*Remark 1.* It should be noted that the theory can also be formulated in terms of Lie transformations [11, 12]: in this case the coordinate transformation corresponds to the time-one flow of an analytic vector field, and a number of technical problems—in particular, concerning inverse transformations—are automatically taken care of. Here we stick to the usual setting for simplicity.

In the new coordinates,  $X$  is expressed as

$$X = \sum_{i=1}^n f^i(x) \frac{\partial}{\partial x_i} \equiv f^i \partial_i. \tag{9}$$

Similarly,  $S$  is now expressed as

$$S = \sum_{i=1}^n p^i(x) \frac{\partial}{\partial x_i} \equiv p^i \partial_i \tag{10}$$

where  $p_i$  are nonlinear functions given explicitly by

$$p^i(x) = [J^{-1}(x)]_{ij}\phi_j(x). \tag{11}$$

The geometrical relations between geometrical objects (in particular, vector fields) do not depend, however, on the choice of coordinates; thus,  $[X, S] = 0$  continues to hold. We recall that, if two vector fields  $X = f^i \partial_i$  and  $Y = g^i \partial_i$  satisfy  $[X, Y] = 0$ , this means that, in terms of the components of the vector fields

$$\{f, g\}^i \equiv (f^j \cdot \partial_j)g^i - (g^j \cdot \partial_j)f^i = 0 \quad \forall i = 1, \dots, n. \tag{12}$$

*Remark 2.* Let us remark, however, that a vector field  $X$  may be linear in two different coordinate systems even if they are connected by a *nonlinear* transformation: consider, for example, the one-dimensional harmonic oscillator

$$\dot{q} = p \quad \dot{p} = -q$$

and a (nonlinear) transformation

$$Q = f(q^2 + p^2)q \quad P = f(q^2 + p^2)p$$

where  $f$  is a smooth function such that  $f(0) = 1$ . In the new coordinate system  $Q$  and  $P$ , one gets

$$\dot{Q} = P \quad \dot{P} = -Q$$

and the dynamics is linear. Therefore, in this case, the dilation fields

$$S = q \frac{\partial}{\partial q} + p \frac{\partial}{\partial p} \quad \text{and} \quad S' = Q \frac{\partial}{\partial Q} + P \frac{\partial}{\partial P}$$

provide two *different* linear structures which linearize the vector field along with the respective symmetries, but are not taken into each other by the same transformation (indeed,  $S = [1 + 2(q^2 + p^2)f'/f]S'$ ).

Suppose now that we have to study the *nonlinear* system (6), without knowing about (4), and in particular we want to know if it is linearizable. It is quite easy to produce examples in which, as in those we consider later, the Poincaré–Dulac theory fails to recognize immediately the intrinsically linear nature of the system.

*Remark 3.* It should be specified that by this, we mean that the NF unfolding allows for nonlinear terms; obviously, if one was going to actually perform the normalization—rather than just determining the general NF unfolding corresponding to the linear part—there would be a series of ‘miraculous’ cancellations, so that the coefficients of the nonlinear resonant terms would vanish.

If we study the symmetry properties of (6), i.e. the vector fields  $Y = g^i(x)\partial_i$  for which (12) is satisfied, we know that *if (6) is linearizable, then (12) admits at least one solution with  $g = p$  given by (11)*. Notice that, in particular, this means that—also in the  $x$  coordinate—the linearization of  $Y$  is just given by  $(DY)(0) = I$ ; this is again a consequence of the fact that (4) is near-identity. (Here and in the following we use the short notation  $(DY)(0) = M$ , with  $Y = g^i\partial_i$  a vector field and  $M$  a matrix, to mean that  $(Dg)(0) = M$ .) It is interesting to remark, and this constitutes our first result, that the converse is also true; i.e. we have the following.

*Theorem 1.* Let  $X = f^i(x)\partial_i$  be a vector field in  $R^n$ . If the equation  $[X, Y] = 0$  admits a solution for which  $(DY)(0) = I$ , where  $Y = g^i(x)\partial_i$ , then  $X$  is linearizable by a formal series of near-identity changes of coordinates. Moreover, there is a formal series of near-identity changes of coordinates which linearizes  $X$  and which does also linearize  $Y$ , transforming it in the dilation field  $S$ , and this happens for any such solution.

The reason for this lies in a very simple consequence of a theorem given in [9] and [10], which we report here for completeness, in a form suitable for our present purpose (a stronger form of this theorem—not relevant to the present discussion—also exists, see [9, 13]); in order to state this we have to introduce some terminology.

We refer to the classical (S + N) decomposition of a matrix: this is the unique decomposition of a matrix  $M$  into a semisimple and a normal part, called respectively  $M_s$  and  $M_n$ , which moreover, satisfy  $[M_s, M_n] = 0$  (see, e.g., [14]). We also refer to (semisimple) joint normal forms: denote by  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{A}_s$  and  $\mathcal{B}_s$  the homological operators [1] associated to  $A$ ,  $B$ ,  $A_s$  and  $B_s$  (with the notation introduced in (12),  $\mathcal{A} = \{Ax, \cdot\}$ , and so on). We say that

$$X = \tilde{f}^i(y)(\partial/\partial y_i) \quad Y = \tilde{g}^i(y)(\partial/\partial y_i) \quad (13)$$

where

$$\tilde{f}(y) = Ay + \tilde{F}(y) \quad \tilde{g}(y) = B(y) + \tilde{G}(y) \quad (14)$$

are in semisimple joint normal form if both  $\tilde{F}$  and  $\tilde{G}$  are in  $\text{Ker}(\mathcal{A}_s) \cap \text{Ker}(\mathcal{B}_s)$ , and that they are in joint normal form if both  $\tilde{F}$  and  $\tilde{G}$  are in  $\text{Ker}(\mathcal{A}) \cap \text{Ker}(\mathcal{B})$ .

*Remark 4.* We recall [1, 2, 15] that for  $A$  normal,  $\text{Ker}(A) = \text{Ker}(A^+)$ ; we also recall that  $\text{Ker}(A)$  and  $\text{Ker}(A^+)$  are contained in  $\text{Ker}(A_s)$ .

With this notation, we can state the following.

*Proposition.* See [9, 10]. Let the polynomial vector fields  $X$  and  $Y$  in  $R^n$ , expressed in the  $x$  coordinates as

$$X = f^i(x)\partial_i \quad Y = g^i(x)\partial_i \tag{15}$$

commute, i.e.  $[X, Y] = 0$ . Let the linearization of  $X$  and  $Y$  at  $x = 0$  be given, respectively, by  $A = (DX)(0)$  and  $B = (DY)(0)$ , and let the matrices  $A$  and  $B$  have (S + N) decompositions  $A = A_s + A_n$  and  $B = B_s + B_n$ ; let  $F$  and  $G$  be the nonlinear parts of  $f$  and  $g$ , so that  $f = Ax + F$  and  $g = Bx + G$ . Then, by a formal series of near-identity (Poincaré) transformations, it is possible to reduce  $f$  and  $g$  to *joint semisimple normal form*; if  $A$  and  $B$  are normal matrices, then it is possible to reduce  $f$  and  $g$  to *joint normal forms*.

*Proof of theorem 1.* Under the hypotheses of theorem 1,  $B = I$ , so that  $B_s = B = I$ ; in this case in particular  $\text{Ker}(B) = \text{Ker}(B^+) = \text{Ker}(B_s) = \mathcal{K}$ , and, independently of  $A$ , we can, due to the proposition, transform  $f$  to  $\tilde{f}$  with  $\tilde{F} \in \mathcal{K}$ . It is a general result that  $\mathcal{K} = \text{Ker}(B^+)$  consists of vector polynomials which are *resonant with B*; these are characterized as follows [1]. Let  $\lambda_1, \dots, \lambda_n$  be the eigenvalues of  $B$ ; then  $\mathcal{K}$  is spanned by vectors  $v = (v_1, \dots, v_n)$  which have all components equal to zero except for  $v_r$ , which is given by  $v_r(x) = x^{m_1} \dots x_n^{m_n}$ , where  $m_i$  are non-negative integers† which satisfy the *resonance relation*

$$\sum_{i=1}^n m_i \lambda_i = \lambda_r. \tag{16}$$

Notice that this will be a polynomial of order  $m = \sum_i m_i$ .

In the case of the identity matrix,  $\lambda_i = 1$ , and there is no resonance relation with  $m > 1$ . Thus, for  $B = I$ ,  $\tilde{F} \in \mathcal{K}$  actually means that  $\tilde{F} = 0$ , and therefore  $\tilde{f}(y) = Ay$ . This proves the theorem. □

*Remark 5.* It should be stressed that the above proof would also work if  $B$  was not the identity, but any matrix such that its semisimple part  $B_s$  does not admit resonances, as also discussed later.

### 3. Convergence of the linearizing transformation

It is well known that in general the Poincaré procedure is only formal, i.e. the series defining the coordinate transformations (called in the following the *normalizing transformation*, or NT) required to take the system (6) into NF is in general not convergent. Some special conditions which guarantee the convergence of the NT are known; these deal either with the structure of the spectrum of  $A = (Df)(0)$  (e.g. the condition that they belong to a Poincaré domain [1–3]), or with some symmetry property of  $f$  (see the Bruno–Markhashov–Walcher theory, [16–19]; see also [20, 21]).

Here we just recall that if  $A$  is real, its eigenvalues  $\sigma_1, \dots, \sigma_n$  belong to a Poincaré domain if and only if  $\epsilon \text{Re}(\sigma_i) > 0$  for all  $i$ , where the sign  $\epsilon = \pm 1$  is the same for all  $i$ . In this case, we are guaranteed of the convergence of the Poincaré normalizing transformation [1, 2].

† It is understood that we work in the space  $\mathcal{V}$  of polynomial vectors, so that  $\mathcal{A}, \mathcal{B}$ , etc are defined on these. The space  $\mathcal{V}$  is naturally graded by the degree of the polynomials.

*Theorem 2.* With the same notation and under the same hypotheses as in theorem 1, the series of near-identity changes of coordinates which takes  $X$  and  $Y$  into linear NF is convergent in a neighbourhood of the origin.

*Proof of theorem 2.* In the case of the matrix  $B = I$ , the eigenvalues are obviously in a Poincaré domain, and the NT is therefore guaranteed to be convergent. Notice that in general this NT would not be unique, being defined up to elements in the kernel of the homological operator  $\mathcal{B}$ ; however, for  $B = I$  this kernel is trivial, and the NT is unique. Theorem 1 guarantees that this transformation also takes  $X$  into NF, and thus that  $X$  can be linearized by means of a convergent change of coordinates.  $\square$

*Remark 6.* Similarly to that remarked earlier concerning theorem 1, the proof would work (and theorem 2 apply) for more general symmetries: for example, it would suffice to require that the eigenvalues of  $B$  belong to a Poincaré domain (see also section 5).

On the other hand, it is clear that the converse of theorem 2 (and of theorem 1 as well) is also true: indeed, the vector field  $X$ , once linearized, obviously admits the dilation symmetry  $S$ , and, if the linearizing transformation is convergent, there exists an analytic symmetry  $Y = p^i \partial_i$  with  $(DY)(0) = I$ , which is transformed into  $S$  by the transformation which linearizes  $X$ . So, focusing on the convergence of the normalizing transformation [16–20], the above arguments can be summarized and completed in the following form.

*Theorem 3.* A vector field  $X = f^i(x) \partial_i$  can be linearized if and only if it admits a (possibly formal) symmetry  $Y = p^i(x) \partial_i$  with  $(DY)(0) = B = I$ ; the NT which linearizes  $X$  is convergent in a neighbourhood of the origin if and only if this symmetry is analytic.

Notice that theorem 3 includes the (trivial) case that the vector field itself is such that  $(DX)(0) = I$ : the Poincaré criterion (mentioned earlier) is in this case sufficient to guarantee that the vector field can be linearized by a convergent transformation: we can see this case as one in which the symmetry requested by theorem 3 consists of the vector field itself.

#### 4. Families of vector fields

It should be stressed that the approach to the formal linearization, and the proof that this is actually not only formal, given previously, and based on symmetry properties of the vector field immediately extends to the case in which we have a family of vector fields;  $X_\mu$  depending smoothly on the real parameters  $\mu \in R^m$  such that there is a family  $Y_\mu$  of symmetry vector fields, i.e. of vector fields where  $[X_\mu, Y_\mu] = 0$ , provided  $Y_\mu$  have a linear part  $B(\mu) \equiv (DY_\mu)(0) = I$ .

We are also assuming, in view of remark 2, that the dynamic is linearizable with respect to the same linear structure (i.e. independently of the parameter  $\mu$ ). In general, the arguments given in theorems 2 and 3 remain valid considering changes of coordinates depending smoothly on  $\mu$ , and so we are guaranteed to have a  $\mu$ -dependent family of convergent NTs. In this framework, (1) and (4) would now be

$$X_\mu = \sum_{i=1}^n A_{ij}(\mu) y_j \frac{\partial}{\partial y_i} \quad (1')$$

$$y_i = \phi_i(x_1, \dots, x_n; \mu). \quad (4')$$

Under this, the linear dynamical system

$$\dot{y}_i = A_{ij}(\mu)y_j \tag{5'}$$

is changed into the nonlinear system

$$\dot{x}^i = f_\mu^i(x) \tag{6'}$$

and thus in new coordinates we have

$$X_\mu = \sum_{i=1}^n f_\mu^i(x) \frac{\partial}{\partial x_i}. \tag{9'}$$

We will not repeat our previous discussion for the parameter-dependent case here.

*Remark 7.* We would like to stress that—in the same way as in the previous sections—our discussion would also apply to the case in which  $B(\mu) = (DY_\mu)(0)$  is not the identity, provided, for example, the eigenvalues  $\lambda_i(\mu)$  of  $B(\mu)$  belong to a Poincaré domain for all values of  $\mu$ ; see also the next section.

*Remark 8.* In order to avoid possible confusion, we would like to briefly consider the case where we have a family of vector fields  $X_\mu$  admitting a common symmetry  $Y$ , i.e.  $[X_\mu, Y] = 0 \forall \mu$ . If  $B = (DY)(0)$  is the identity (or, however, does not admit resonances, see also the next section), then according to our discussion we would have a unique transformation which takes into NF the whole family of vector fields  $X_\mu$  at once, and this can appear surprising. Notice, however, that this is the case only if  $B$  does not admit resonances, and in this case  $X_\mu$  would become linear once transformed into NF. Thus, if  $B = I$ , we know that any linear vector field  $X_A = (Ax)^i \partial_i$  commutes with  $Y = (Bx)^i \partial_i$ ; if we change coordinates via a near-identity transformation  $x = \phi(y)$ , then  $Y = g^i(y) \partial_i$  is a symmetry of all the vector fields  $X_A = f^i(y) \partial_i$ , which depend on  $(n \times n)$  parameters, i.e. the entries of the matrix  $A$ .

### 5. Symmetries not having the identity as a linear part

In some cases, determining a symmetry whose linear part is *not*  $B = I$  can also suffice to guarantee the—formal or convergent—linearizability of the vector field, or at least the possibility to considerably simplify it. This fact was already pointed out in remarks 5–7, and we are now going to discuss it a little further, in particular with reference to the problem of convergence of the NT.

First, let us consider the case where  $B \neq I$  but  $\text{Ker}(B) = \{0\}$ , i.e.  $B$  does not admit resonances (for simplicity we assume  $B$  is semisimple): we know that  $f$  can then be linearized. If, in addition, the symmetry  $Y$  is analytic and the matrix  $B$  satisfies the ‘condition  $\omega$ ’ of Bruno [3], then we are guaranteed that the NT which linearizes  $f$  is convergent. This conclusion follows from the Bruno theorem [3]: indeed the other condition required by the Bruno theorem (‘condition A’) is in this case automatically satisfied, as the NF is in fact linear:  $Y = (Bx)^i \partial_i$ .

Let us recall briefly, for completeness, what the requirements for  $B$  to satisfy ‘condition  $\omega$ ’ are: we ask that  $B$  is semisimple, and denote by  $\lambda_i$  its eigenvalues. Consider then the set of  $Q = (q_1, \dots, q_n)$  where  $q_i$  are integers such that  $q_i \geq -1$  and  $(Q, \Lambda) \neq 0$  (see [3] for full details), with

$$(Q, \Lambda) = q_1 \lambda_1 + \dots + q_n \lambda_n \tag{17}$$



let  $\omega_k = \min |(Q, \Lambda)|$  on  $Q$  such that  $(Q, \Lambda) \neq 0$  and  $1 < \sum q_i < 2^k$ . Then we say that condition  $\omega$  is satisfied if

$$\sum_{k=1}^{\infty} 2^{-k} \ln \omega_k^{-1} < \infty. \quad (18)$$

Let us then consider the case where again  $B \neq I$ , but its eigenvalues belong to a Poincaré domain; in such a case  $f$  can be transformed, by means of a convergent series of Poincaré transformations, into a very simple form (even if possibly nonlinear), i.e. can be brought to be in  $\text{Ker}(\mathcal{B})$ , which is in this case finite-dimensional. Indeed, in this case we can take  $Y$  into NF, and we are guaranteed that the required NT is convergent (due to the Poincaré domain condition). In doing this,  $X$  is transformed as well, and in the new coordinates  $X = f^i(x)\partial_i$ , with  $f \in \text{Ker}(\mathcal{B})$ .

Similarly, we can contemplate the case in which  $\text{Ker}(\mathcal{A}) \cap \text{Ker}(\mathcal{B}) = \{0\}$  (this situation is considered in example 3 later); in this case both  $X$  and  $Y$  are linearized when taken to the joint normal form—as mentioned in [9]—but the (joint) normalizing transformation is, without further assumptions, in general only formal.

If  $A$  satisfies condition  $\omega$ , then  $f$  can be taken into normal form—i.e. in  $\text{Ker}(\mathcal{A})$ —by a convergent transformation, but we are not guaranteed in general that the linearizing transformation is also convergent. Notice, however, that this linearizing transformation is in fact convergent in the particular case where  $Y$  is linear,  $Y = (B_{ij}x_j)\partial_i$ . Indeed, the symmetry condition  $[X, Y] = 0$  becomes in this case  $f \in \text{Ker}(\mathcal{B})$ , and it can be easily verified that this condition is preserved by any transformation taking  $f$  into NF; i.e.  $f$  in  $\text{Ker}(\mathcal{A})$  implies  $f \in \text{Ker}(\mathcal{A}) \cap \text{Ker}(\mathcal{B})$  and therefore the NF of  $f$  is necessarily linear. If it is instead  $B$  to satisfy condition  $\omega$ , then  $f$  can be taken to be in  $\text{Ker}(\mathcal{B})$  by a convergent transformation; unless  $B$  does not admit any resonance, this is obviously not sufficient to ensure the convergent linearization of  $f$ . Notice that in both these cases, ‘condition A’ is automatically satisfied.

These considerations can be summarized in the following form.

*Theorem 4.* Let the vector fields  $X = f^i\partial_i$  and  $Y = g^i\partial_i$  satisfy  $[X, Y] = 0$ , and assume the matrix  $B = (Dg)(0)$  is semisimple; then: (a) if  $\text{Ker}(\mathcal{B}) = \{0\}$  and  $B$  satisfies condition  $\omega$ , then  $X$  and  $Y$  can be linearized by a convergent transformation; (b) if  $\text{Ker}(\mathcal{A}) \cap \text{Ker}(\mathcal{B}) = \{0\}$ , then  $X$  and  $Y$  can be linearized (possibly by means of a non-convergent transformation), but this transformation is convergent in the case when  $Y$  is linear,  $Y = (Bx)^i\partial_i$ . Also, if  $A = (Df)(0)$  (respectively  $B = (Dg)(0)$ ) is semisimple and satisfies condition  $\omega$ , then there is a convergent transformation taking  $X$  (respectively  $Y$ ) into normal (not necessarily linear) form.

It is worth emphasizing the consequence of theorem 4 in the case of families of vector fields. Indeed, for vector fields of the form (6') one has that the eigenvalues of  $A$  vary continuously with  $\mu$ , so that generically [22], for almost all values of  $\mu$  the eigenvalues are strongly non-resonant, and therefore, by the Siegel theorem, the system can be linearized by an analytic change of coordinates. Moreover, there is a smooth (in the sense of Whitney) dependence of the linearizing transformation on  $\mu$ , when  $\mu$  varies in the Cantor set to which correspond strongly non-resonant frequencies. However, the Siegel theorem does not give any information on the behaviour of the system when the parameter belongs to the bad set (complementary to the Cantor set).

Theorem 4 ensures that, provided there exists at least one  $\mu$  dependent symmetry (with suitable properties), system (6') can be linearized for any value of the parameter, and the

linearizing transformation depends smoothly on  $\mu$ , when  $\mu$  belongs to an interval, i.e. a regular set, and not a Cantor set.

We would like, in concluding this section, to point out that it could happen that a symmetry with its linear part the identity cannot immediately be determined, but that its existence can be inferred from the presence of another symmetry (and the linear space structure of the Lie algebra of symmetries).

Thus, consider the case of a two-dimensional system  $X = f^i(x)\partial_i$  whose linear part  $A$  is semisimple and has distinct eigenvalues; assume that  $X$  admits a symmetry  $Y = g^i(x)\partial_i$  whose linear part  $B$  is semisimple and not proportional to  $A$ . (Obviously one does not have to seek explicitly all of these conditions: the existence of distinct eigenvalues ensures  $A$  is semisimple, and then  $[A, B] = 0$  implies  $B$  is semisimple as well.) In this case, we are always able to easily find a symmetry vector field  $Z$ —in particular, a linear combination  $Z = aX + bY$ —whose linear part is the identity (it is for this reason that in example 3 later we will have to consider a three-dimensional system).

The same argument is easily generalized to  $n$ -dimensional systems  $X$  having  $n$  symmetries—one of them being  $X$  itself—with independent linear part  $B_i$ , provided the matrices  $B_i$  are semisimple and span a nilpotent—for example Abelian—Lie algebra (for generalizations, not needed in the present discussion, see [13]). Notice that, in particular, if  $X$  is linearizable in the form (1) and  $A^k$  are independent, these could be the vector fields  $Y_{(k)}$  considered in (2).

### 6. Several symmetries having the identity as a linear part

When we consider a fixed vector field  $X$ , there can be—obviously—different vector fields  $Y$  which commute with it, i.e. which are symmetries for  $X$ ; the set of all vector fields which satisfy  $[X, Y] = 0$  is obviously a Lie algebra under the commutator; it is called the symmetry algebra of  $X$  and will be denoted by  $\mathcal{G}$ . In particular, it can happen that several  $Y_i \in \mathcal{G}$  satisfy  $(DY_i)(0) = I$ . Notice that in general these do not form an algebra; in particular, we know that  $(D[Y_i, Y_j])(0) = [(DY_i)(0), (DY_j)(0)]$  (as it is immediately seen by expanding  $Y$  in Taylor series around 0), and therefore  $[Y_i, Y_j]$  has a vanishing linear part; thus, in particular, this linear part is not the identity. Notice, however, that there is no reason for  $[Y_i, Y_j]$  to vanish in general.

According to our theorem 1, we can choose each  $Y_i$ , and simultaneously linearize  $X$  and  $Y_i$ . It should be stressed that in this way we do not, in general, linearize other  $Y_j$ ; more precisely, not only do we not linearize them by the NT adapted to the pair  $(X, Y_i)$ , but in general we are not able to linearize simultaneously different  $Y$ . This fact has to do with the problem of simultaneously reducing to joint normal form a general algebra of vector fields [10, 23]. We mention the case of interest here, i.e. of algebras of vector fields having semisimple linear part (see [13] for a more general discussion): in this case, one can prove that a joint normal form is possible only if the algebra is nilpotent, which includes in particular the case of Abelian algebras [10].

It should be mentioned that the case of general—or even just solvable—algebras appears to be extremely hard; indeed, it is related to the problem of simultaneous reduction to the Jordan form of an algebra of matrices (the linear parts of the vector fields in question); or, this problem is not solved, neither it is known under which conditions it can or cannot be solved [24].

*Remark 9.* This is also an appropriate point to remark that in theorem 1 we could only guarantee the existence of a transformation linearizing both  $X$  and  $Y$ ; however, it is

quite clear—a *fortiori* in the light of the above considerations—that we could also have a transformation which linearizes  $X$  without linearizing  $Y$  (or *vice versa*). Similarly, in general the transformation which linearizes  $X$  and  $Y$  will not be unique.

## 7. Examples

We will now consider some very simple examples. Examples 1–3 show how we can apply our results to guarantee that a concrete nonlinear system can be linearized, without actually performing the Poincaré normalization; example 4 deals with the situation discussed in section 5, and the somewhat artificial example 5 (see later for its construction) attempts to show how the method can deal with ‘seriously wrong’ initial coordinates.

*Example 1.* Consider the dynamical system in  $R^2$  given by

$$\dot{x} = x \quad \dot{y} = 3y - x^2. \quad (19)$$

The linear part of this is given by the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \quad (20)$$

and the general NF corresponding to such a linear part is given by

$$\dot{x} = x \quad \dot{y} = 3y + \alpha x^3 \quad (21)$$

with  $\alpha$  an arbitrary real constant.

By explicit (standard) computations, one can check that the symmetry algebra of (19) is spanned by the vector fields

$$Z_1 = x\partial_x + 2x^2\partial_y \quad Z_2 = (y - x^2)\partial_y \quad (22)$$

(in particular,  $X = Z_1 + 3Z_2$  corresponds to (19) itself); if we choose  $Y = Z_1 + Z_2$ , i.e.

$$Y = x\partial_x + (y + x^2)\partial_y \quad (23)$$

we find a symmetry vector field  $Y$ , whose linearization is indeed  $(DY)(0) = I$ . This guarantees that our system can be linearized, and actually provides the linearizing transformation as well.

In the above example, we had a very simple situation, as no parameter appears in the vector field, and the eigenvalues belong to a Poincaré domain, so that the convergence of the NT is guaranteed. Thus, the only non-trivial result is that the term  $\alpha x^3$  in (21) actually disappears from the normalized form. In the following examples, we consider more complicated cases.

*Example 2.* Let us consider again a two-dimensional system, i.e.

$$\dot{x} = x + x^4y \quad \dot{y} = -2y - x^3y^2. \quad (24)$$

The linear part of this is given by the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \quad (25)$$

and the general NF corresponding to such a linear part is given by

$$\dot{x} = x + x\phi_1(x^2y) \quad \dot{y} = -2y + y\phi_2(x^2y) \quad (26)$$

where  $\phi_i$  are two arbitrary (smooth) functions.

A symmetry vector field of this is given by

$$Y = (x + 4x^4y)\partial_x + (y - 4x^3y^2)\partial_y \tag{27}$$

which has a linear part given by  $B = I$ , and thus we conclude that the system (24) can be reduced to its linear part by a convergent normalizing transformation.

Notice that in this case  $\text{Ker}(\mathcal{A})$  is infinite-dimensional, and the eigenvalues do not belong to a Poincaré domain.

*Example 3.* We consider now a (fourth-order) system in  $R^3$

$$\begin{aligned} \dot{x} &= x + a_1x^3y + b_1xy^2z \\ \dot{y} &= -3y + a_2x^2y^2 + b_2y^3z \\ \dot{z} &= 9z + a_3y^2z^2 + b_3y^2z^2 \end{aligned} \tag{28}$$

where  $a_i$  and  $b_i$  are arbitrary constants; thus,

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 9 \end{pmatrix}. \tag{29}$$

A (linear) symmetry for this DS is given by

$$Y = x\partial_x - 2y\partial_y + 4z\partial_z \tag{30}$$

and we have

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}. \tag{31}$$

In this case, both  $\text{Ker}(\mathcal{A})$  and  $\text{Ker}(\mathcal{B})$  are infinite-dimensional, but their intersection is just  $\{0\}$ . Notice also that their eigenvalues (both for  $A$  and  $B$ ) are not in a Poincaré domain.

According to the arguments in section 5, we can conclude that  $X$  can be linearized by a convergent transformation.

*Example 4.* In this example we merely want to illustrate the discussion of section 6, and consider a case strongly related to one already considered in example 6 of [5]. However, here we make different hypotheses concerning the linear part in which we have several non-commuting symmetries  $Y$  of a given vector field  $X$ , such that  $(DY)(0) = I$ . To avoid unnecessary complications, we consider  $X$  to be already in linear form; it is of course possible to rephrase the example by setting  $X$  to be nonlinear, by means of any suitable change of coordinates.

Consider the linear vector field

$$X = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}. \tag{32}$$

It can be immediately checked that any vector field of the form

$$Y = f(r^2)X + g(r^2)Z \tag{33}$$

is a symmetry of  $X$ , where

$$Z = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} \tag{34}$$

and  $r^2 = x^2 + y^2$ .

In particular, if we choose  $f$  and  $g$  such that

$$f(0) = 0 \quad g(0) = 1 \quad (35)$$

we have a symmetry vector field with linear part the identity.

We can rewrite such vector fields as

$$Y_{f,h} = f(r^2)X + (1 + h(r^2))Z \quad (36)$$

where it is understood that both  $f$  and  $h$  vanish in zero.

One can readily check that

$$[Y_{f,h}, Y_{\phi,\eta}] = 2r^2[(1 + h)\phi' - (1 + \eta)f']X + 2r^2[(1 + h)\eta' - (1 + \eta)h']Z \quad (37)$$

and thus that  $Y_{f,h}$  do not commute among themselves. More precisely,  $Y_{\phi,\eta}$  commutes with  $Y_{f,h}$  if and only if

$$\phi = c_1 f + c_2 \quad 1 + \eta = c_1(1 + h)$$

with  $c_1$  and  $c_2$  two arbitrary real numbers; it should be noticed that when we linearize  $Y_{f,h}$ , which reduces then to  $Z$ , the commuting fields  $Y_{\phi,\eta}$  also reduce to  $Z$ , as such fields also have the identity as its linear part.

From this short discussion, we can draw several conclusions. First,  $X$  admits a symmetry with its linear part the identity; however,  $X$  is already linear, so we should not enquire about it being linearizable. As for  $Y_{f,h}$ , these all admit the linear vector field  $X$  as symmetry, but  $(DX)(0) \neq I$ ; thus, the first part of theorem 1 cannot guarantee the linearizability of  $Y$ . On the other hand, the second part of the same theorem does guarantee that it is possible to take  $Y$  into linear form; this is not really a surprise, as the linear part  $Y_0$  of  $Y$  is associated to the identity matrix  $I$  (indeed  $Y_0(x) = Ix$ ), and the eigenvalues of this are not resonant in the Poincaré sense, but, moreover, belong to a Poincaré domain [1], so that the same result could be obtained by classical Poincaré theory.

Finally, for what concerns linearizing different  $Y$  at the same time, the above computation for the commutator shows that in general—i.e. unless (37) vanishes—we are not able to simultaneously linearize different  $Y_{f,h}$ .

*Remark 10.* The vector fields  $Y_{f,h}$  were also considered in example 6 of [5], and found to be nonlinearizable. In order to avoid possible misunderstandings, it should be stressed that the hypotheses made on the linear parts—and in particular on the role of the rotation component of the vector field in its linearization—are different in the present paper and in [5]. It should also be mentioned that in [5] we considered global linearization, while here we are in the (perturbative) framework of NF theory, and we consider only linearization in a neighbourhood  $U$  of the origin. This point is further discussed in the appendix.

*Example 5.* As a final example, we consider an apparently hopelessly complicated system (see later for how it was generated), i.e.

$$\begin{aligned} \dot{x} = f_1(x, y, z) &= [\alpha x - y] - x^2 - [3xy^2 + 2\alpha y^3] \\ &\quad - 6[x^3y + \alpha x^2y^2] - 3[x^5 + 2\alpha x^4y + y^5] - [2\alpha x^6 + 15x^2y^4] \\ &\quad - 30[x^4y^3 + x^6y^2] - 3[x^{10} + 5x^8y] \\ \dot{y} = f_2(x, y, z) &= [x + \alpha y] - [\alpha x^2 - 2xy] + [2x^3 + y^3] \\ &\quad + [9x^2y^2 + 4\alpha xy^3] + [15x^4y + 12\alpha x^3y^2] + [7x^6 + 12\alpha x^5y + 6xy^5] \\ &\quad + [4\alpha x^7 + 30x^3y^4] + 60x^5y^3 + 60x^7y^2 + 6[5x^9y + x^{11}] \end{aligned}$$

$$\begin{aligned} \dot{z} = f_3(x, y, z) = & \beta z + [2xy + (2\alpha - \beta)y^2] \\ & + [2x^3 + 2(2\alpha - \beta)x^2y + 3xy^2 + \alpha y^3 + (2\alpha - \beta)y^3] \\ & + [(2\alpha - \beta)x^4 - 3\alpha x^2y^2 + 6xy^3 + 2y^4] + [6x^3y^2 + 8x^2y^3 + 3y^5] \\ & + [12x^4y^2 + 27x^2y^4 + 12\alpha xy^5] + 9[5x^4y^3 + 4\alpha x^3y^4] \\ & + [2x^8 + 8x^6y + 21x^6y^2 + 36\alpha x^5y^3 + 18xy^7] + [12\alpha x^7y^2 + 90x^3y^6] \\ & + 180x^5y^5 + 180x^7y^4 + 90x^9y^3 + 18x^{11}y^2. \end{aligned} \tag{38}$$

In this case the linear part is given by

$$\dot{\mathbf{x}} = A\mathbf{x} \tag{39}$$

with  $\mathbf{x} = (x, y, z)$  and  $A$  the matrix

$$A = \begin{pmatrix} \alpha & -1 & 0 \\ 1 & \alpha & 0 \\ 0 & 0 & \beta \end{pmatrix}. \tag{40}$$

Despite the very complicated form of (38), one can check explicitly that  $X = f_i(\mathbf{x})(\partial/\partial x_i)$  commutes with  $Y = g_i(\mathbf{x})(\partial/\partial x_i)$  (here and in the following example,  $(x_1, x_2, x_3) = (x, y, z)$ ), where

$$\begin{aligned} g_1(\mathbf{x}) &= x - 2y^3 - 6x^2y^2 - 6x^4y - 2x^6 \\ g_2(\mathbf{x}) &= y - x^2 + 4xy^3 + 12x^3y^2 + 4x^7 + 12x^5y \\ g_3(\mathbf{x}) &= z + y^2 + 2x^2y + 2y^3 + x^4 - 3x^2y^2 + 12xy^5 + 36x^3y^4 + 36x^5y^3 + 12x^7y^2. \end{aligned} \tag{41}$$

The linear part of this vector field corresponds only to the identity matrix, and thus our theorem guarantees immediately that  $X$  can be reduced to its linear part; i.e. in NF coordinates we have

$$X = (A_{ij}y_j) \frac{\partial}{\partial y_i}. \tag{42}$$

Actually, the system (38) was obtained from (42) by the change of coordinates

$$y_1 = x_1 + (x_1^2 + x_2)^3 \quad y_2 = x_2 + x_1^2 \quad y_3 = x_3 - x_2^3 - (x_1^2 + x_2)^2 \tag{43}$$

and  $Y$  is, in  $y$  coordinates, nothing more than the dilation vector field, i.e.

$$Y = y_i \frac{\partial}{\partial y_i}. \tag{44}$$

Similarly, we could have considered the vector field

$$Z = ((A^2)_{ij}y_j) \frac{\partial}{\partial y_j} \tag{45}$$

which depends on  $\alpha$  and  $\beta$ , and which obviously commutes with  $X$  (and with  $Y$ ).

We stress now that we have a two-parameter family of systems (42), and that as  $\alpha$  and  $\beta$  are varied, this goes through resonances, and the eigenvalues  $\lambda_{\pm} = \alpha \pm i$  and  $\lambda_0 = \beta$  can be in a Poincaré domain or otherwise. However, our symmetry method is not sensitive to these facts, and works for all values of the parameters.

## 8. Some further remarks

*Remark 11.* In the previous example 1, we have been able to identify *all* the symmetries of our system. We would like to stress, however, that the only important fact from our point of view is that we are able to determine *one* symmetry with the required linear part: this is a much simpler task, and this is what has been done in the other examples. As it is generally the case with symmetry methods for differential equations, it is the possibility of obtaining relevant information from the knowledge of *one* symmetry (or a symmetry subalgebra), without the need to know the full—in general, infinite-dimensional—symmetry algebra of the dynamical system under study, which makes our method applicable.

*Remark 12.* In general, one could try to determine perturbatively the functions  $p^i(x)$ , by expanding them in homogeneous terms as  $p^i(x) = \sum_{m=1}^{\infty} p_m^i(x)$ , where  $p_m^i(ax) = a^m p_m^i(x)$ , and solving the determining equations order by order [25]; in particular, one should require  $p^i(x) = x_i$ . It should also be mentioned that if in this way we determine a solution (or a solution exists) only up to some order  $k$ , we can guarantee that the system can be linearized up to terms of order  $k$  [25].

This information, although more limited than a full linearization property, can equally be of great utility: first, because in actual computations one considers in most cases a truncation at some (high) order  $k$ ; and second, because if we combine such a result with the analysis of resonances, in order to guarantee the full linearizability of the system, it suffices to guarantee the existence of a symmetry  $Y$  with linear part the identity up to order  $k$  of the highest order resonance of the system (in example 1 one would only have to go up to order three). Notice that we are guaranteed that this order is finite when the eigenvalues belong to a Poincaré domain.

*Remark 13.* It can be helpful to see theorem 1 from a slightly different perspective than the one used in section 2 (we use freely the notation employed there); this discussion actually repeats points already mentioned, and is thus also a way to summarize our argument. As  $(DY)(0) = I$ , we know [1] that  $Y$  is biholomorphically equivalent to its linear part in a neighbourhood  $U$  of the origin (this only depends on the fact that the eigenvalues of  $(DY)(0)$  belong to a Poincaré domain and there are no resonances, so that it would extend to more general linear parts  $(DY)(0) = B$ ). When we apply the NT we have, denoting by  $y$  the ‘new’ coordinates,  $Y = y^i \partial / \partial y^i$ , and  $X = \tilde{f}^i(y) \partial / \partial y^i$ ; however, the relation  $[X, Y] = 0$  is independent of the coordinate representation of  $X$  and  $Y$ , and thus in the neighbourhood  $U$  in which the NT is not only formal (actually, as mentioned previously, is biholomorphic) we necessarily have  $\tilde{f}^i(y) = A_j^i y^j$  with  $A$  a matrix, and actually  $A = (DX)(0)$ .

*Remark 14.* The previous remark also shows what results can be expected when we deal with  $C^k$  functions rather than with a formal power series: if  $Y$  is a  $C^k$  vector field with, for example,  $(DY)(0) = I$ , we can put it into NF—i.e. linearize it—by a  $C^{k-1}$  transformation; the rest of the argument follows as before, and thus we conclude that if  $X$  and  $Y$  are commuting  $C^k$  vector fields, with the same hypotheses as above concerning their linear parts, then they can be simultaneously linearized by a  $C^{k-1}$  NT. We thank N N Nekhoroshev for this observation.

*Remark 15.* Finally, we would like to point out that the present approach is related to the study of integrability conducted by Marmo and collaborators, see for example [26, 27]; the use of symmetries to study the linearizability of a dynamical system has been considered,

in a non-perturbative approach (related to the general theory of symmetry methods for differential equations [28–31]) in [5].

### Acknowledgments

We would like to thank the referees of a previous version of this paper, who suggested we clarify several points.

One of the authors (GG) would like to thank the Mathematisches Forschungsinstitut Oberwolfach (MFO) and the IHES (Bures sur Yvette) for hospitality during part of this work. The stay in the MFO was supported by the *Volkswagen Stiftung* under the RiP programme.

### Appendix. On perturbative and non-perturbative linearization

As already pointed out in the introduction and in the body of the paper, in the present work we discuss the same problem as in [5], i.e. the relation between symmetry and linearizability for dynamical systems; the main difference between the two approaches being that there we used a non-perturbative and global approach, while here we remain in the framework of perturbation theory. However, it is quite clear that (strong) relations have to be present between the two approaches, and this appendix is devoted to the discussion of these.

First, in the perturbative approach one should distinguish between formal linearization and an actual one, i.e.—in the present setting based on NFs—between the case in which the NT providing the linearization of the system is purely formal, and the case where it converges. It is quite clear that in the case of purely formal linearization, we should expect in general that the system is not linearizable when we consider global transformations, i.e. in the sense adopted in [5].

The second point is that, even when the NT is convergent, this transformation is in general by no means global, i.e. it is defined only on some neighbourhood  $U$  of the origin. Thus, even in this case, it is not surprising if we get results which do not agree with those obtained by the approach proposed in [5].

It should also be pointed out that it is possible that a system can be linearized, globally or only in a neighbourhood of the origin, but *not* by a near-identity transformation: in this case, it cannot be linearized by the NF approach. An example of this situation is provided by

$$\dot{x}_i = \sum_{j=1}^n A_{ij} \frac{\rho^2 x_j}{\rho^2 + x_i^2} \tag{A1}$$

(with  $\rho^2 = \sum_{i=1}^n x_i^2$ ), which is nothing else than

$$\dot{y}_i = A_{ij} y_j \tag{A2}$$

with the change of coordinates  $y_i = \rho^2 x_i$ .

Let us now come back to the differences between global and local linearization; a clarifying example in this respect is provided by a nonlinear oscillator described by (here  $r^2 = x^2 + y^2$ )

$$\dot{x} = -r^2 y - (r^2 - 1)x \quad \dot{y} = +r^2 x - (r^2 - 1)y \tag{A3}$$

which can be solved (e.g. passing to polar coordinates), and it is clear by its behaviour that it cannot be linearized in the global sense, according to [5], by the same arguments used in



[5]. On the other hand, its linear part is just the identity, and thus by the standard results in NF theory discussed here, it can be linearized by a convergent transformation (in fact, it is biholomorphically equivalent to its linear part in a neighbourhood  $U$  of the origin [1]), but it is obvious that such a neighbourhood  $U$  cannot include the limit cycle.

Another example, which we want to discuss at some length in the following, is provided by the vector fields  $Y_{f,h}$  considered in example 4 of section 7. Thus, we have that—not surprisingly—global and local linearization are quite different. However, the discussion conducted in [5] could be restricted to a neighbourhood  $U$  of the origin; in this case, the results obtained by the two approaches should be—when the NT is convergent—equivalent.

We are thus presenting here a short discussion of the restriction of the approach of [5] to such a neighbourhood; we will, for the sake of clarity, discuss systems as those encountered in example 4 (to avoid confusion with the notation used in this example and in (A3) as well, we are introducing here new and independent notation). Precisely, we want to discuss the problem of the existence of a diffeomorphism<sup>†</sup>  $\Phi : U \rightarrow U$  conjugating a vector field  $Y$ , such that  $Y(0) = 0$ , to the dilation vector field

$$\tilde{Y} = \eta_1 \frac{\partial}{\partial \eta_1} + \eta_2 \frac{\partial}{\partial \eta_2}. \quad (\text{A4})$$

As  $Y(0) = 0$  and  $\tilde{Y}(0) = 0$ , we also require that  $\Phi(0) = 0$ .

First, we notice that the eigenvalues of  $A = (DY)(0)$  are invariants under any such  $\Phi$  (of class at least  $\mathcal{C}^1$ ); indeed,  $Y$  induces its lift  $Y^T$  on  $TU$ , given (with  $v^i$  spanning a basis in  $T_x U$ ) by

$$Y^T = Y^i \frac{\partial}{\partial x^i} + \frac{\partial Y^i}{\partial x^j} v^j \frac{\partial}{\partial v^i} \quad (\text{A5})$$

and, as  $Y(0) = 0$ , we have

$$Y^T(0) = \frac{\partial Y^i}{\partial x^j}(0) v^j \frac{\partial}{\partial v^i} \equiv A_j^i v^j \frac{\partial}{\partial v^i}. \quad (\text{A6})$$

Also,  $\Phi : U \rightarrow U$  induces a diffeomorphism  $\Phi^T : TU \rightarrow TU$ ; thus for conjugated fields  $Y, \tilde{Y}, Y^T$  and  $\tilde{Y}^T$  are conjugated in the origin by  $\Phi^T$ . We conclude that all the algebraic invariants of  $(DY)(0)$  and of  $(D\tilde{Y})(0)$  coincide. Thus, we have that a *necessary* condition for  $Y$  and  $\tilde{Y}$  to be  $\mathcal{C}^1$ -conjugated is that the matrices  $A = (DY)(0)$  and  $\tilde{A} = (D\tilde{Y})(0)$  are conjugated.

In the case of

$$Y \equiv Y_{f,h} = [1 + h(r^2)]Z + f(r^2)X \quad (\text{A7})$$

we have thus that a necessary condition for  $Y$  to be conjugated to  $\tilde{Y}$  is that  $A = (DY)(0)$  is conjugated to the identity matrix. But in this case  $A$  is just the identity matrix, as  $f(0) = h(0) = 0$ .

If we consider the more general family of vector fields (see example 6 of [5])

$$Y_{\alpha,\beta} = \alpha(r^2)Z + \beta(r^2)X \quad (\text{A8})$$

we have that

$$A = \begin{pmatrix} \alpha(0) & -\beta(0) \\ \beta(0) & \alpha(0) \end{pmatrix} \quad (\text{A9})$$

and thus we have as a necessary condition that the determinant and trace of  $A$  are equal to that of the identity matrix; these two conditions together mean that  $a(0) = 1$  and  $b(0) = 0$ ,

<sup>†</sup> In the present discussion, a diffeomorphism will be meant to be of class  $\mathcal{C}^1$ , and not necessarily  $\mathcal{C}^\infty$ .

i.e.  $A = I$  (obviously the identity matrix can only be conjugated to itself). This condition is verified in (A7), but not in the case considered in [5].

Although a necessary condition can be useful to ensure linearizability, it would be preferable to have a necessary and sufficient condition; this is provided by the existence of two solutions  $\eta_1$  and  $\eta_2$ , such that  $\eta_1(0) = \eta_2(0) = 0$  and which are functionally independent on  $U$ , to

$$L_Y \eta = \eta \tag{A10}$$

(where, as in the following,  $L_Y$  is the Lie derivative along the vector field  $Y$ ). We could, of course, attempt a solution of this by series expansion: this would be essentially equivalent to the Poincaré method.

In the case of  $Y = Y_{f,h}$  this equation reads

$$L_Z \eta + f(r^2)L_X \eta + h(r^2)L_Z \eta = \eta. \tag{A11}$$

In the general case, it can also be appropriate to observe that if  $Y$  is conjugated in  $U \subseteq R^n$  (with  $\{O\} \in U$ ) to the linear field

$$Y_A = A_i^j \eta_j \frac{\partial}{\partial \eta_i} \tag{A12}$$

then there will exist  $n$  functionally independent solutions to

$$L_Y \eta_i = A_i^j \eta_j. \tag{A13}$$

It should be stressed that in this way we can show that if  $Y$  is locally conjugated to the dilation field  $\tilde{Y}$ , then necessarily  $(DY)(0) = \pm I$ ; this is the converse to the statement, well known in NF theory, that if  $(DY)(0) = \pm I$  then  $Y$  is locally conjugated to the dilation field.

It should be stressed also, for completeness of discussion, that a vector field can be somehow correlated to the dilation field, for example being proportional to it, without being conjugated to it. Indeed, consider the vector field

$$Y = \frac{1}{3}x \frac{d}{dx} \tag{A14}$$

and let us look for solutions to (A10). This yields  $(1/3)x \, d\eta/dx = \eta$  and thus  $\eta(x) = x^3$ ; thus  $\eta$  is analytic, but it does not define a diffeomorphism (its inverse  $x = \eta^{1/3}$  is singular in the origin).

Finally, we would like to mention that, if we take the view adopted in the appendix, the arguments presented in the main body of the paper can be reinterpreted as a discussion of perturbative techniques, i.e. of the use of Poincaré theory, to solve (A11).

## References

- [1] Arnold V I 1988 *Geometrical Methods in the Theory of Differential Equations* (Berlin: Springer)
- [2] Arnold V I and Il'yashenko Yu S 1988 Ordinary differential equations *Dynamical systems, Encyclopaedia of Mathematical Sciences* vol I, ed D V Anosov and V I Arnold (Berlin: Springer) pp 1–148
- [3] Bruno A D 1989 *Local Methods in Nonlinear Differential Equations* (Berlin: Springer)
- [4] Verhulst F 1990 *Nonlinear Differential Equations and Dynamical Systems* (Berlin: Springer)
- Verhulst F 1996 *Nonlinear Differential Equations and Dynamical Systems* 2nd edn (Berlin: Springer)
- [5] Gaeta G and Marmo G 1996 *J. Phys. A: Math. Gen.* **29** 5035
- [6] Belitsky G R 1978 *Russ. Math. Surveys* **33** 107
- [7] Elphick C, Tirapegui E, Brachet M E, Couillet P and Iooss G 1987 *Physica* **29D** 95
- [8] Iooss G and Adelmeyer M 1992 *Topics in Bifurcation Theory and Applications* (Singapore: World Scientific)
- [9] Cicogna G and Gaeta G 1994 *J. Phys. A: Math. Gen.* **27** 461

- [10] Cicogna G and Gaeta G 1994 *J. Phys. A: Math. Gen.* **27** 7115
- [11] Benettin G, Galgani L and Giorgilli A 1984 *Nuovo Cimento B* **79** 201
- [12] Mitropolsky Yu A and Lopatin A K 1995 *Nonlinear Mechanics, Groups and Symmetry* (Dordrecht: Kluwer)
- [13] Cicogna G and Gaeta G, Symmetry and perturbation theory in nonlinear dynamics *Preprint IHES P/97/12*
- [14] Hirsch M W and Smale S 1974 *Differential Equations, Dynamical Systems, and Linear Algebra* (London: Academic)
- [15] Kolmogorov A N and Fomine S 1953 *Elements of the Theory of Functions* (Oxford: Pergamon)
- [16] Markhashov L M 1974 *J. Appl. Math. Mech.* **38** 788
- [17] Walcher S 1991 *Math. Ann.* **291** 293
- [18] Bruno A D 1993 *Selecta Mathematica* **12** 13
- [19] Bruno A D and Walcher S 1994 *J. Math. Anal. Appl.* **183** 571
- [20] Cicogna G 1995 *J. Phys. A: Math. Gen.* **28** L179  
Cicogna G 1996 *J. Math. Anal. Appl.* **199** 243  
Cicogna G 1997 *Preprint mp\_arc* 97–159  
Cicogna G 1997 *J. Phys. A: Math. Gen.* **30** 6021–8
- [21] Ito H 1989 *Comm. Math. Helv.* **64** 412  
Ito H 1992 *Math. Ann.* **292** 411
- [22] Sprindzuk V G 1979 *Metric Theory of Diophantine Approximation* (Washington: Winston and Sons)
- [23] Arnal D, Ben Ammar M and Pinczon G 1994 *Lett. Math. Phys.* **8** 467
- [24] Kirillov A A 1976 *Elements of the Theory of Representations* (Berlin: Springer)
- [25] Cicogna G and Gaeta G 1994 *Nuovo Cimento B* **109** 989
- [26] De Filippo S, Vilasi G, Marmo G and Salerno M 1984 *Nuovo Cimento B* **83** 97
- [27] Marmo G 1986 *Proc. 3rd Seminar on Group Theoretical Methods in Physics (Yurmala, May 22–24 1985)* (Moscow: Nauka)
- [28] Olver P J 1986 *Applications of Lie Groups to Differential Equations* (Berlin: Springer)
- [29] Ovsjannikov L V 1962 *Group Properties of Differential Equations* (Novosibirsk) (Engl. Transl. Bluman G 1967 *Group Analysis of Differential Equations* (New York: Academic))
- [30] Bluman G W and Kumei S 1989 *Symmetries and Differential Equations* (Berlin: Springer)
- [31] Gaeta G 1994 *Nonlinear Symmetry and Nonlinear Equations* (Dordrecht: Kluwer)