Non-perturbative linearization of dynamical systems

1996 J. Phys. A: Math. Gen. 29 5035

(http://iopscience.iop.org/0305-4470/29/16/024)

J. Phys. A: Math. Gen. **29** (1996) 5035–5048. Printed in the UK

Non-perturbative linearization of dynamical systems

G Gaeta†§ and G Marmo‡

† Department of Mathematics, Loughborough University, Loughborough LE11 3TU, UK
‡ Dipartimento di Scienze Fisiche, Università di Napoli, *and* INFN, Sezione di Napoli, 80123 Napoli, Italy

Received 2 February 1996, in final form 14 May 1996

Abstract. We consider the classical problem of linearizing a vector field X around a fixed point. We adopt a non-perturbative point of view, based on the symmetry properties of linear vector fields.

1. Introduction and statement of the problem

The problem of linearizing a nonlinear vector field in the neighbourhood of a fixed point by means of $\mathcal{C}^\infty$ transformations is a classical one, its perturbative treatment going back to Poincaré [1–4] in the general case and to Birkhoff for Hamiltonian systems [5, 6]; here we want to consider it from a non-perturbative point of view; moreover, we will not deal with general normal form transformations [1–6], but only consider linearizable systems.

Indeed, although it turns out that the Poincaré procedure for linearizing, a linearizable system is also successful in reducing a generic (nonlinearizable) system to its normal form, so that the problem of formal reduction to normal form is not more difficult in the general case than in the linearizable one, it is natural to expect that if we proceed non-perturbatively, the linearizable case will be much easier to treat. The considerations we will use in the following are specific to the linearizable case, and can not be extended to the general one.

Let us consider a *linear* dynamical system in R^n for which the origin is a fixed point,

$$\dot{x}^i = A^i{}_j x^j \tag{1}$$

(where $i, j = 1, \ldots, n$ and A is a $n \times n$ real matrix), and consider now an invertible (nonlinear) diffeomorphism‖ which identifies a change of coordinates

$$x^i = \Phi^i(y) \tag{2}$$

we will also denote by $y^i = \Psi^i(x)$ the inverse change of coordinates. Let us denote by Λ the Jacobian of this change of coordinates, and by Γ its inverse,

$$\Lambda^i{}_j = \frac{\partial \Phi^i}{\partial x^j} \equiv \frac{\partial x^i}{\partial y^j} \qquad \Gamma^i{}_j = \frac{\partial y^i}{\partial x^j} \equiv \frac{\partial \Psi^i}{\partial x^j} \qquad \Lambda^i{}_j \Gamma^j{}_k = \delta^i{}_k. \tag{3}$$

In the new coordinates (1) is written as

$$\dot{y}^i = f^i(y) \tag{4}$$

§ E-mail address: G.Gaeta@lboro.ac.uk
‖ It will be clear from the following discussion that we could as well consider a domain D—containing the origin—in R^n rather than the whole R^n; similarly we could as well consider Φ invertible only locally.

0305-4470/96/165035+14$19.50

where the f are now *nonlinear* functions given explicitly by

$$f^i(y) = \Gamma^i{}_k(y) A^k{}_j \Phi^j(y). \tag{5}$$

Suppose now that we have to study (4), with f given explicitly, so that we do not know about A and Φ. How can we find out that (4) corresponds actually to linear dynamics 'in the wrong coordinates'?

The purpose of the present note is indeed to answer this question; it will turn out that in the process of answering this we also answer the question of how to concretely linearize the system, i.e. to determine the linearizing change of coordinates Φ.

2. Symmetry approach

Clearly, a possible approach would consist in using the theory of (Poincaré–Dulac) *normal forms*; this would amount to a perturbative construction, order by order, of the inverse change of coordinates, thus mapping (4) back into (1). This approach is completely algorithmic and constructive, and moreover it is quite general, in that it works both for linearizable and nonlinearizable systems. However, in the linearizable case this approach has also several drawbacks, essentially amounting to its perturbative character:

(a) if A presents resonances, one would expect nonlinear resonant terms [1–4] to be present, so that one would realize the inherent linearity of the system only after checking a series (usually infinite) of 'miraculous' cancellations occurring in the normalized expansion;

(b) if Φ is not analytic—even if C^∞—we can not hope to linearize the system by the Poincaré procedure, which is inherently perturbative and polynomial (one could consider Ecalle's resurgent functions theory [7, 8], but again this means introducing very complicated tools for a simple problem);

(c) in any case, the procedure requires extensive computations, checks of the convergence of perturbative expansions, and so on; moreover, we should go to infinite order in perturbation theory to obtain exact linearization. Even in the most favourable case, in which one is sure *a priori* of the linearizability of the problem and of the convergence of the linearizing transformation (e.g. thanks to Siegel's theorem [2, 6] or to symmetry properties [9–14]), to compute the explicit linearizing change of coordinates one still has to go at infinite order in perturbation. In one word, it requires a huge amount of work to recognize the simple system (1).

Thus, we will look for a different, non-perturbative, approach for this problem. The natural idea would be to look for properties of the dynamical system, or equivalently of the vector field

$$X = A^i{}_j x^j \frac{\partial}{\partial x^i} = f^i(y) \frac{\partial}{\partial y^i} \tag{6}$$

which are invariant under changes of coordinates—i.e. they have a tensorial character—and which recognize the linear nature of the system. From this point of view, it is quite natural to look at the *symmetries*† of (1): indeed, if a vector field S commutes with X, the relation $[X, S] = 0$ will hold independently of any system of coordinates (here and in the following $[.,.]$ is the usual commutator of vector fields).

Symmetries which are related to the linear nature of (1) are those generated by powers of A, i.e. by vector fields of the form (in the x coordinates)

$$X_k = [A^k]^i{}_j x^j \frac{\partial}{\partial x^i} \tag{7}$$

† The symmetry approach to differential equations—both ODEs and PDEs—pioneered by Lie, has received recently diffused attention and is dealt with, and applied, in several books and many papers, see e.g. [15–21].

for k a non-negative integer. Clearly, as it follows from $[A^n, A^m] = 0$, these form an abelian algebra (generically of dimension n). Notice that for $k = 0$ (i.e. for $A^0 = I$) we have the generator of scalings, $X_0 = x^i \partial/\partial x^i$, which will be a symmetry for any linear system.

In the y coordinates, the X_k take the form

$$X_k = \Gamma^i{}_j(y)[A^k]^j{}_m \Phi^m(y) \frac{\partial}{\partial y^i}. \tag{8$'$}$$

These satisfy therefore, in particular,

$$X_{k+1} = (\Gamma A \Gamma^{-1}) X_k. \tag{8$''$}$$

Thus, even if we analyse the symmetry algebra of (4) and we detect in it an abelian algebra, it can be difficult to realize the vector fields in it are of this form, although of course the relation (8$''$) is easier to recognize than the form (8$'$).

The situation is slightly better if we consider X_0 alone: indeed, in this case we have to look for symmetries of the simpler form

$$X_0 = \Gamma^i{}_j(y)\Phi^j(y) \frac{\partial}{\partial y^i} \tag{9}$$

with Γ given by (3). Thus, a possible approach would consist of looking for solutions to the determining equation for symmetries of dynamical systems†

$$\varphi_t + (f \cdot \nabla)\varphi - (\varphi \cdot \nabla) f = 0 \tag{10}$$

in the form $\varphi(y, t) = (D\Phi^{-1})\Phi$. Recalling that $D\Phi^{-1} = -\Phi^{-1}(D\Phi)\Phi^{-1}$, this also reads $\varphi(y, t) = -\Phi^{-1}(D\Phi)$.

Notice that the fact that $x^i(\partial/\partial x^i)$ is a symmetry, without further assumptions on the X (e.g. analyticity), only ensures that in the x coordinates we have $X = \tilde{f}(x)\partial_x$ with $\tilde{f}$ homogeneous of order one; on the other hand, as we deal with non-singular f, and Φ invertible, we are guaranteed that in this setting $\tilde{f}$ is indeed linear. We have therefore:

Lemma 1. If the equation (4) admits a symmetry $\varphi^i(y)\partial/\partial y^i$ and φ can be written in the form $[D\Phi^{-1}(y)]^i{}_j \Phi^j(y)$, then by the change of coordinates $y = \Phi^{-1}(x)$, (4) is reduced to a system $\dot{x} = \tilde{f}(x)$ with $\tilde{f}$ linear.

We stress that lemma 1 does *not* require the determination of the full symmetry of (4), i.e. the most general solution to (10), but only a special solution with an appropriate form. Indeed, getting the full solution to (10) requires one to find the most general solution to the associated homogeneous PDE, namely to solve (4).

Another possibility stems from the obvious observation that (1) admits a linear superposition principle‡; this means, in particular, that

$$X_\xi = \xi^i(t) \frac{\partial}{\partial x^i} \tag{11}$$

generates a symmetry of (1), provided ξ obeys (1) itself, i.e. provided $\dot{\xi}^i(t) = A^i{}_j \xi^j(t)$. Indeed, one can easily check that this is the case by using equation (10) in the x coordinates, which in this case reduces to $\dot{\xi} = A\xi$.

In the y coordinates we have

$$X_\xi = \Gamma^i{}_j(y)\xi^j(t) \frac{\partial}{\partial y^i} \tag{12}$$

and therefore we have:

† We recall that if φ satisfies (10), then $X_\varphi = \varphi^i(y, t)(\partial/\partial y^i)$ is a symmetry of X [15–21].

‡ The idea of using this fact to characterize the linearizability of a nonlinear PDE belongs to Kumei and Bluman [16, 22–24]; here we are actually specializing their theory to the case of first order ODEs.

Lemma 2. If the equation (4) admits a symmetry $\varphi^i(y)\partial/\partial y^i$ for φ of the form $M^i_j(y)\xi^j(t)$ with ξ an arbitrary solution of the linear equation $\dot{\xi} = A\xi$, then by the change of coordinates $y = \Phi^{-1}(x)$, with $M = D\Phi^{-1}$, (4) is reduced to the linear system $\dot{x} = Ax$.

Here again, we stress that it is not required to know the most general solution of (10).

3. Intrinsic approach

It is possible to look for the linearization of a dynamical system in a slightly more general setting, making contact with the general theory of Nijenhuis operators [25–29].

We define a *separating set* of functions to be a finite collection $f_1, f_2, \ldots, f_p$ such that $f_a(x) = f_a(y)$ for any $a = 1, \ldots, p$ implies $x = y$; this means that we must have $p \geqslant n$.

Definition. A separating set of functions is said to be a *linearizing set* for X if it happens that

$$L_X f_a = A_a{}^b f_b. \tag{13}$$

(Here, L_X is the Lie derivative along X.)

Then, any vector field X admitting a linearizing set f is f-related to a linear system on R^p, with a map $f : R^n \to R^p$ which is just given by

$$f : x \to (f_1(x), \ldots, f_p(x)). \tag{14}$$

If we denote by z_a the coordinates in R^p, the image of R^n under f will be given by $z_a = f_a(x)$. So, in the z coordinates our vector field X is f-related to

$$Y = A_a{}^b z_b \frac{\partial}{\partial z_a}. \tag{15}$$

When $p = n$, we get a linearization of our system in the usual sense†. Therefore, given a vector field X we can look for a linearizing set for X in the specific case $p = n$.

We consider now a vector field Z, and denote by ζ the semiflow under Z,

$$\frac{\mathrm{d}}{\mathrm{d}t}\zeta(t; y)\Big|_{t=0} = Z(y) \tag{16}$$

we will also denote by $B_\delta(y_0)$ the ball of radius δ centred in y_0.

Definition. The vector field Z is dilation-type if: (i) there exists a unique y_0 such that $Z(y_0) = 0$; (ii) there exist n functionally independent real functions $h_i : M \to R$ which are solutions of

$$Z(h) = h. \tag{17}$$

Notice that the h_is provide a linearizing set for Z, with matrix $A^b_a = \delta^b_a$. We say then that the h_i are a *diagonalizing* set of functions for Z.

Notice also that it would be natural to require that for a dilation field there is $\delta > 0$ such that

$$\lim_{t\to-\infty} \zeta(t; y) = y_0 \qquad \forall y \in B_\delta(y_0) \tag{18}$$

however, this is automatically satisfied when a linearizing set exists.

† The introduction of linearizing sets in the general case allows one to deal with more general situations [28, 29]; e.g. if we have a linear flow in R^n but we consider it on a nonlinear embedded submanifold M (the simplest case being that of $M = S^{n-1} \to R^n$), the flow on M cannot be globally linearized in the usual sense, but it is recognized as a linear flow by means of this approach.

Lemma 3. If Z is a dilation-type vector field, with $\{h_1, \ldots, h_n\}$ a diagonalizing set of functions for Z, then any f solution of $Z(f) = f$ can be uniquely written as

$$f(y) = \sum_{i=1}^{n} c^i h_i(y). \tag{19}$$

Indeed, since the h_is are functionally independent (which is a condition stronger than the separating condition) we have for any function f

$$\mathrm{d}f = g^i \, \mathrm{d}h_i \qquad g^i \in \mathcal{F}. \tag{20}$$

By requiring f to be a solution to (17) we get

$$f = \mathrm{d}f(Z) = g^i \mathrm{d}h_i(Z) = g^i h_i = L_Z f \tag{21}$$

and therefore $(L_Z g^i) h_i = 0$ implies $L_Z g^i = 0$, as the h_i are functionally independent. Due to the regularity requirement on the g^i in the neighbourhood of y_0, we have $g^i \in R$. Thus, we conclude that any solution to $Z(f) = f$ can be written in the form (19), with the c^i real constants; the lemma is proved.

Clearly, if Z is just X_0, then $y_0 = \Phi(0)$ and the h_i are nothing else than the x_i as functions of the y, i.e. $h_i = \Phi^i(y)$.

Using lemma 3, we have immediately:

Lemma 4. Let Z be a dilation-type vector field, with $\{h_1, \ldots, h_n\}$ a diagonalizing set of functions for Z. If $[X, Z] = 0$, then $\{h_1, \ldots, h_n\}$ is a linearizing set for X.

Indeed, if $[X, Z] = 0$, we have

$$L_X h_i = L_X(L_Z h_i) = L_Z L_X h_i \tag{22}$$

which also means

$$L_X h_i = A_i{}^j h_j \tag{23}$$

because of the properties of solutions to (17) and of lemma 3. As the h_is define a change of coordinates $x^i = h_i(y)$, the linearized vector field will be

$$Y = A_i{}^j h_j \frac{\partial}{\partial h_i}. \tag{24}$$

We finally notice that generically (i.e. under suitable non-degeneracy conditions, satisfied by generic vector fields) if h_1 is a solution to $L_Z h = h$, we may get new functionally independent solutions by applying repeatedly L_X to h_1. This simple fact can be of help in constructing the diagonalizing set for Z.

4. Symmetry and recursion operators

In this section, we would like to point out how the approach defined in the previous section is related to recursion operators and the Lax formalism for integrable systems. Notice, indeed, that a system which is linearizable by a change of coordinate (C-linearizable in the Calogero terminology) is by this definition also integrable.

When X is a linear vector field,

$$X = A^i{}_j x^j \frac{\partial}{\partial x^i} \tag{25}$$

we can associate to the matrix A a $(1,1)$ tensor field

$$R \equiv T_A = A_i{}^j \left[\mathrm{d}x^i \otimes \frac{\partial}{\partial x^j} \right]. \tag{26}$$

This tensor satisfies automatically the two equations

$$L_X(R) = 0 \tag{27}$$
$$N_R = 0 \tag{28}$$

where L_X is the Lie derivative along X, and N_R is the Nijenhuis tensor [25–29] associated with R, i.e.

$$N_R[X,Y] = [RX,RY] + R^2[X,Y] - R\,[[RX,Y] + [X,RY]]. \tag{29}$$

By applying R to X we get the vector fields $X_k = (R)^k X$, which have the property that

$$[X_k, X_m] = 0 \tag{30}$$

i.e. we can generate pairwise commuting symmetries.

Thus for a given X, the existence of a $(1,1)$ tensor field R such that (27) and (28) are satisfied is a necessary condition for X to be linearizable.

It would be possible [28, 29] to look for a separating set of functions by searching for invariant subspaces of exact 1-forms under the endomorphism associated to R on 1-forms, i.e.

$$R(\mathrm{d}f_a) = B_a{}^b \mathrm{d}f_b \tag{31}$$

with B a real matrix. Clearly, for the powers of R we have

$$(R)^k(\mathrm{d}f_a) = (B^k)_a{}^b \mathrm{d}f_b \tag{32}$$

where B^k is the kth power of the matrix B. Notice also that generalized eigenspaces of R are invariant subspaces for X.

It should be stressed, finally, that this R has all the properties of a *recursion operator* (in the sense encountered in the theory of integrable systems [15, 28]) for our finite-dimensional evolution equation; thus, it permits one to also obtain a Lax representation, as discussed, for example, in [28].

Rather than discussing this point here, we refer to [28, 29] for a general discussion, and more specifically to [28–35] for the geometry of Lax systems, to [25–29] for the geometry of Nijenhuis operators, to [25–29, 36–39] for how the Nijenhuis tensor describes the geometry of the tangent bundle; and to [36–40] for the geometry of the Nijenhuis tensor in relation with a distinguished vector field X on M describing dynamical evolution. Finally, for the Hamiltonian setting (shortly discussed in the appendix), see [31–35, 41, 42].

5. Examples: linearizable vector fields

We will now consider some examples of applications of our results. We will for each example consider, in order, the application of the methods based on lemma 1, lemma 2 and on lemma 4.

In the following we write all indices as lower ones, to avoid any confusion between indices and exponents.

Example 1. As a first although trivial test, we consider the case $n = 1$. Now we have for (1) $\dot{x} = ax$, and

$$f(y) = a\frac{\Phi}{\Phi'}. \tag{33}$$

Looking first for $\varphi = \varphi(y)$ as solution to (10), we get

$$\frac{\varphi_y}{\varphi} = g(y) \equiv \frac{f_y}{f} \tag{34}$$

and in this case we actually have

$$g(y) = \frac{\Phi'}{\Phi} - \frac{\Phi''}{\Phi'} \tag{35}$$

so that indeed

$$\varphi(y) = c\Phi(y)/\Phi'(y) = \tilde{c}f(y). \tag{36}$$

Thus, applying lemma 1 is just equivalent to determining directly if, given $f(y)$, there exists a Φ such that (33) is verified; obviously this just yields

$$\Phi(y) = c_1 \exp\left[\int_{y_0}^{y} [a/f(y)]\,\mathrm{d}y\right]. \tag{37}$$

To make a concrete example, in this way we immediately get that

$$\dot{y} = \frac{1+y^2}{1+3y^2}ay \tag{38}$$

is transformed into $\dot{x} = ax$ with

$$x = \Phi(y) = y + y^3. \tag{39}$$

Let us now look for

$$\varphi = \varphi(y,t) = \alpha(y)\xi(t) \tag{40}$$

so that (10) now reads

$$\alpha\dot{\xi} + f\xi\alpha_y = \alpha\xi f_y \tag{41}$$

which for α yields

$$\frac{\mathrm{d}\alpha}{\alpha} = \left[\frac{f_y}{f} - \frac{\dot{\xi}}{\xi}\frac{1}{f}\right]\mathrm{d}y. \tag{42}$$

If $\dot{\xi} = k\xi$, we get

$$\frac{\mathrm{d}\alpha}{\alpha} = \left[\left(1 - \frac{k}{a}\right)\frac{\Phi_y}{\Phi} - \frac{\Phi_{yy}}{\Phi}\right] \tag{43}$$

and choosing $k = a$ we get

$$\alpha = \frac{c_2}{\Phi_y}. \tag{44}$$

For a concrete example, one could use this approach to obtain Γ which, for f as in (38), yields the same Φ as in (39).

In the approach based on lemma 3 (and with the notation of section 3), Z is just given by $Z = g(y)\,\mathrm{d}/\mathrm{d}y$, see equation (34). With this, (18) just yields $f(y) = c\Phi(y)$, which corresponds to lemma 3.

Example 2. We consider now a two-dimensional system $\dot{y} = f(y)$, with only one nonlinear term:

$$\begin{aligned} \dot{y}_1 &= y_1 - \epsilon y_2 - y_2^2 \\ \dot{y}_2 &= y_2. \end{aligned} \tag{45}$$

Let us first try to apply lemma 1. From (10), we obtain that the vector field X_0 identified by

$$\varphi_1 = y_1 - y_2^2 \qquad \varphi_2 = y_2 \tag{46}$$

is a symmetry of our system; this is of the form $\varphi_i = \Gamma_{ij}\Phi_j$ with

$$\Phi = \begin{pmatrix} y_1 + y_2^2 \\ y_2 \end{pmatrix} \qquad \Gamma = \begin{pmatrix} 1 & 2y_2 \\ 0 & 1 \end{pmatrix} \tag{47}$$

which indeed satisfies $[\Gamma^{-1}]_{ij} = \partial\Phi_i/\partial y_j$; correspondingly we get the linearizing transformation Φ^{-1} as

$$y_1 = x_1 - x_2^2 \qquad y_2 = x_2 \tag{48}$$

and in this coordinate we get, as expected,

$$X_0 = x_i \frac{\partial}{\partial x_i}. \tag{49}$$

Let us now come to the procedure based on lemma 2, i.e. let us look for φ in the form $\varphi(y,t) = \Gamma_{ij}(y)\xi_j(t)$. The equations (10) are now†, assuming that $\dot{\xi}_i = A_{ij}\xi_j$ for some A and eliminating the common factor ξ_k,

$$f_j \frac{\partial \Gamma_{ik}}{\partial y_j} = \frac{\partial f_i}{\partial y_j}\Gamma_{jk} - \Gamma_{ij}A_{jk}. \tag{50}$$

With the explicit expression of f, and therefore $\partial f/\partial y$, given above, we have that indeed for Γ as in (47) above and

$$A = \begin{pmatrix} 1 & -\epsilon \\ 0 & 1 \end{pmatrix} \tag{51}$$

the equation is satisfied. This leads to the same linearizing transformation (48) as above, and the linear equation is indeed just $\dot{x} = Ax$.

Let us now consider, for the same problem, the approach of section 3; we take

$$Z = (y_1 - y_2^2)\frac{\partial}{\partial y_1} + y_2 \frac{\partial}{\partial y_2} \tag{52}$$

and the diagonalizing h_is are given by

$$h_1 = y_1 + y_2^2 \qquad h_2 = y_2. \tag{53}$$

One can check that in this case, solving (18)—e.g. by the method of characteristics—yields

$$f(y) = c_1 h_1(y) + c_2 h_2(y) \tag{54}$$

with arbitrary constants c_1, c_2; and moreover that

$$\begin{cases} X(h_1) = h_1 - \epsilon h_2 \\ X(h_2) = h_2 \end{cases} \tag{55}$$

so that $x_i = h_i(y)$ takes the system into the form $\dot{x} = Ax$ with the same A as in $X(h_i) = A_{ij}h_j$.

† Notice that A and Γ should be seen as the unknowns of the problem.

Example 3. We will consider again a system in R^2, given now by

$$f = \begin{pmatrix} 2(y_1 + y_2 + e^{y_1}) \\ (y_2 - y_1)(1 + 2\,e^{y_1}) + e^{y_1} - 2\,e^{2y_1} \end{pmatrix}. \tag{56}$$

If we look for solutions of (10) in the form $\varphi = \varphi(y)$, we are confronted with a system of two PDEs, i.e.

$$\begin{aligned} &2[y_1 + y_2 + e^{y_1}]\frac{\partial\varphi_1}{\partial y_1} + [(y_2 - y_1)(1 + 2\,e^{y_1}) + e^{y_1} - 2\,e^{2y_1}]\frac{\partial\varphi_1}{\partial y_2} = 2(1 + e^{y_1})\varphi_1 + 2\varphi_2 \\ &2[y_1 + y_2 + e^{y_1}]\frac{\partial\varphi_2}{\partial y_1} + [(y_2 - y_1)(1 + 2\,e^{y_1}) + e^{y_1} - 2\,e^{2y_1}]\frac{\partial\varphi_2}{\partial y_2} \\ &\qquad = -[(1 + e^{y_1}) + 2(y_1 + y_2)\,e^{y_1} + 4\,e^{2y_1}]\varphi_1 + (1 - 2\,e^{y_1})\varphi_2. \end{aligned} \tag{57}$$

It is quite clear that this is not an easy equation to solve; however, one can check that

$$\varphi = \begin{pmatrix} y_1 \\ y_2 + (1 - y_1)\,e^{y_1} \end{pmatrix} \tag{58}$$

provides a special solution to (57). This is indeed of the form (9), as required to apply lemma 1, with

$$\Phi = \begin{pmatrix} y_1 + y_2 + e^{y_1} \\ y_2 + e^{y_1} \end{pmatrix} \tag{59}$$

$$\Gamma = \begin{pmatrix} 1 & -1 \\ -e^{y_1} & 1 + e^{y_1} \end{pmatrix} = \left[\frac{\partial\Phi}{\partial y}\right]^{-1}. \tag{60}$$

In order to apply lemma 2 we would instead look for solutions to (10) in the form $\varphi = \Gamma_{ij}(y)\xi_i(t)$. The equations satisfied by the Γ_{ij} are now even more complicated, but we can take advantage of the freedom given by the A. Indeed, the Γ given above is also a solution to the set of equations one obtains in this way, provided one chooses

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 2 \end{pmatrix}. \tag{61}$$

Indeed, by (59) we have that, with the inverse change of coordinates

$$\begin{aligned} y_1 &= x_1 - x_2 \\ y_2 &= x_2 - e^{(x_1 - x_2)} \end{aligned} \tag{62}$$

we reduce $\dot{y} = f(y)$, with f given by (56), to $\dot{x} = Ax$ with A given by (61).

If we apply the approach of section 3, based on the existence of a dilation-type field Z commuting with X, the Z is given by $Z = \varphi_i\partial/\partial y_i$ with φ as in (58); the h_i are the Φ_i given in (59), and equation (18) gives f as in (54); again, it is easily verified that

$$X(h_i) = A_{ij}h_j \tag{63}$$

with A given explicitly by (61).

6. Examples: nonlinearizable vector fields

We will now give examples in which our results are used to show that a given vector field (dynamical system) can *not* be linearized; we consider systems in R^2 for simplicity.

Example 4. We give first an example of a system which is not linearizable because it does not admit enough symmetries. We consider R^2 with coordinates (x, y) and the vector field

$$\Gamma = \varphi(x^2 + y^2) \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right). \tag{64}$$

If this vector field is linearizable, it has to admit at least two symmetries.

We notice that $X_0 = x\partial_y - y\partial_x$ and $\Delta = x\partial_x + y\partial_y$ are a basis for the module of vector fields which have the origin as a fixed point. Therefore our symmetries should have the form $Y = aX_0 + b\Delta$, with $a, b \in \mathcal{F}(R^2)$ smooth functions†.

We then look for a, b such that $[Y, \Gamma] = 0$; we have to require that (with L_X the Lie derivative along X)

$$[aX_0 + b\Delta, \varphi X_0] = (bL_\Delta\varphi - \varphi L_{X_0}a)X_0 - \varphi(L_{X_0}b)\Delta = 0. \tag{65}$$

Therefore—for the vanishing of the term along Δ—we need $L_{X_0}b = 0$, which implies $b = b(x^2 + y^2)$.

For the other term, i.e. from $bL_\Delta\varphi = \varphi L_{X_0}a$, we integrate both sides along a circle centred at the origin:

$$\int_{S^1} (bL_\Delta\varphi)\, \mathrm{d}\theta = \varphi \int_{S^1} (L_{X_0}a)\, \mathrm{d}\theta \tag{66}$$

where φ has been taken out of the integral because it depends only on $(x^2 + y^2)$, i.e. is a constant on S^1. By using this same argument we arrive at

$$2\pi(bL_\Delta\varphi) = \varphi(a(2\pi) - a(0)) \tag{67}$$

if the function a is regular, $a(2\pi) = a(0)$ and we get $bL_\Delta\varphi = 0$. Thus, either $b = 0$ or $L_\Delta\varphi = 0$ (or both). Now, Δ does not have any smooth constant of motion, and thus $L_\Delta\varphi = 0$ implies that φ is a constant, or otherwise it has to be $b = 0$.

It follows from this that $L_\Delta\varphi \neq 0$ requires $b = 0$, i.e. there is only one family of symmetries (depending on a constant) for our system, which therefore cannot be linearized.

Example 5. We will now consider examples in which we have the required number of symmetries, but none of them is a dilation-type vector field.

Let us write $r^2 = x^2 + y^2$, and consider the (Van der Pol-like) system

$$\begin{aligned} \dot{x} &= -(r^2 - 1)x + (r^2 - 2)y \\ \dot{y} &= -(r^2 - 2)x - (r^2 - 1)y \end{aligned} \tag{68}$$

we denote the corresponding vector field as X.

When we look for symmetries, i.e. for vector fields

$$Y = f(x, y)\partial_x + g(x, y)\partial_y \tag{69}$$

such that $[X, Y] = 0$, it turns out that the only solutions are of the form (with c_1, c_2 real constants)

$$Y_1 = c_1 X \qquad Y_2 = c_2(x\partial_y - y\partial_x). \tag{70}$$

It is clear that Y_2 is not a dilation-type vector field (it is just a homogeneous rotation), and Y_1 is just proportional to X (which is, by the way, not dilation-type as well). Thus, we can conclude—using any of the proposed approaches—that X is not linearizable.

Indeed, as for the first proposal, X does not admit symmetries depending on an arbitrary solution of a linear equation; for the second one, it does not admit a dilation-type symmetry.

† We recall that symmetries of a vector field with isolated fixed points should have the same points as fixed points.

Example 6. We consider now the following generalization (again in R^2) of the situation encountered in example 5:

$$\begin{aligned}\dot{x} &= \alpha(r)x - \beta(r)y \\ \dot{y} &= \beta(r)x + \alpha(r)y\end{aligned} \tag{71}$$

(where both $\alpha(r)$ and $\beta(r)$ are not identically zero) so that we deal with the vector field

$$X = (\alpha(r)x - \beta(r)y)\partial_x + (\beta(r)x + \alpha(r)y)\partial_y \tag{72}$$

or, as it is convenient to use polar coordinates (r, θ),

$$X = \alpha(r)\partial_r + \beta(r)\partial_\theta. \tag{73}$$

We write in full generality

$$Y = f(r,\theta)\partial_r + g(r,\theta)\partial_\theta \tag{74}$$

and now the condition $[X, Y] = 0$ gives two PDEs, i.e.

$$\alpha(r)\frac{\partial f}{\partial r} + \beta(r)\frac{\partial f}{\partial \theta} = f \cdot \alpha'(r) \tag{75}$$

$$\alpha(r)\frac{\partial g}{\partial r} + \beta(r)\frac{\partial g}{\partial \theta} = f \cdot \beta'(r). \tag{76}$$

These can be solved using the method of characteristics; for the first one we get

$$\frac{\mathrm{d}r}{\alpha(r)} = \frac{\mathrm{d}\theta}{\beta(r)} = \frac{\mathrm{d}f}{f\alpha'(r)} \tag{77}$$

equating the first and the third term we get

$$f(r,\theta) = \xi(\theta)\alpha(r) \tag{78}$$

and by using the other term—or going back to (75)—we get $\xi'(\theta) = 0$, i.e.

$$f(r,\theta) = c_1\alpha(r). \tag{79}$$

Let us look at (76); this yields

$$\frac{\mathrm{d}r}{\alpha(r)} = \frac{\mathrm{d}\theta}{\beta(r)} = \frac{\mathrm{d}g}{c_1\alpha(r)\beta'(r)} \tag{80}$$

equating, here again, the first and the third term we get

$$g(r,\theta) = c_1\beta(r) + \xi(\theta) \tag{81}$$

and again, from the other term or going back to (76) we get $\xi'(\theta) = 0$, i.e.

$$g(r,\theta) = c_1\beta(r) + c_2. \tag{82}$$

Thus, we have only two symmetries,

$$Y_1 = c_1 X \tag{83}$$

$$Y_2 = c_2\partial_\theta \tag{84}$$

and we are in the same situation as in example 5, and we can derive the same conclusions, i.e. that X is not linearizable. Clearly, the present discussion does not apply if the condition $\alpha(r) \not\equiv 0 \not\equiv \beta'(r)$ is not satisfied.

Acknowledgments

This paper originated from discussions we had in the IASS (Vietri, Italy) in occasion of the workshop on *Completely integrable systems and separability* in September 1995; we would like to thank the IASS and the organizers for their hospitality on that occasion. GG would also like to thank the Mathematisches Forschungsinstitut Oberwolfach for hospitality while the final version of this paper was being prepared; this stay was supported by Volkswagen Stiftung under the RiP program.

Appendix. Linearization of Hamiltonian systems

In this appendix, we shortly consider the case of a Hamiltonian system, and how the discussion given in the present paper applies to it.

We recall that a dynamical system X is a Hamiltonian on a symplectic manifold (M, ω), where ω is a symplectic structure (i.e. a closed non-degenerate two-form) if

$$i_X\omega = -\mathrm{d}H \tag{85}$$

with H the Hamiltonian function.

It would be natural enough to ask what can be said about ω and H if we linearize X; as a matter of fact, the possibility that X admits alternative Hamiltonian descriptions [43] suggests that not much can be said, in the end, about them.

We recall that X admits alternative Hamiltonian descriptions if there exists symplectic structures ω_a and Hamiltonian functions H_a—where a belongs to some index set—such that

$$i_X\omega_a = -\mathrm{d}H_a. \tag{86}$$

To investigate our question, let us start with a coordinate system ξ^i which linearizes X, i.e.

$$X = (A^i{}_j\xi^j)\frac{\partial}{\partial\xi^i}. \tag{87}$$

In this the equations of motion are

$$\dot{\xi}^i = \omega^{ij}\frac{\partial H}{\partial\xi^j} = A^i{}_k\xi^k \tag{88}$$

where we have written ω as

$$\omega = \omega_{ij}\,\mathrm{d}\xi^i \wedge \mathrm{d}\xi^j \qquad \omega^{ij}\omega_{jk} = \delta^i{}_k. \tag{89}$$

We notice that for $\xi = 0$ (i.e. $\xi^k = 0$ for $k = 1, \ldots, 2n$) we have

$$\omega^{ij}(0)\frac{\partial H(0)}{\partial\xi^j} = 0 \tag{90}$$

with $\omega^{ij}(0)$ invertible.

If we differentiate (88) and evaluate it at the origin, we get

$$A^i{}_k = \omega^{ij}(0)\frac{\partial^2 H(0)}{\partial\xi^j\partial\xi^k}. \tag{91}$$

Thus, $\tilde{\omega}$ and $\tilde{H}$, given by

$$\tilde{\omega} = \omega_{jk}\,\mathrm{d}\xi^j \wedge \mathrm{d}\xi^k \qquad \tilde{H} = \xi^j\xi^k\frac{\partial^2 H(0)}{\partial\xi^j\partial\xi^k} \tag{92}$$

provide a Hamiltonian description for the dynamics.

As a byproduct of this discussion, we get that any decomposition of A into an invertible skew-symmetric matrix times a symmetric matrix provides an alternative Hamiltonian description of X. If we write

$$A = \omega_0^{-1} \cdot H_0 \tag{93}$$

by using any invertible matrix T which commutes with A we get

$$A = [T^{-1}\omega_0^{-1}(T^t)^{-1}] \cdot [(T^t)H_0T] \tag{94}$$

thus, if T is not a canonical transformation for ω_0, by this procedure we obtain alternative Hamiltonian descriptions.

It is not difficult to show that odd powers of A are matrices associated with Hamiltonian (with respect to ω_0) vector fields, while even powers are generators of symmetries for A, which are not canonical; therefore a linear Hamiltonian vector field always admits alternative Hamiltonian descriptions [44].

By using odd powers of A, for a generic matrix A which describes a linear Hamiltonian vector field, we get a maximal set of pairwise commuting quadratic first integrals; i.e. any linear Hamiltonian system (even if it is not generic) is always completely integrable in the Liouville sense [44].

We would like to summarize our discussion of the Hamiltonian case as follows. A linearizable Hamiltonian system does always admit (many) alternative Hamiltonian descriptions; out of a given Hamiltonian description for a given linearizable dynamical system described by a vector field X, we can always find one with a quadratic Hamiltonian (in a coordinate system in which the vector field X is linear); a linearizable Hamiltonian system is always integrable in the Liouville sense. As for the linearization process, we should first linearize the dynamical vector field X, and only afterwards take care of its Hamiltonian description.

Finally, for what concerns the recursion operator, we mention that in this case we get a factorizable one [45, 46], by using ω_a to raise the indices of ω_b, with the additional requirement that the Poisson bracket

$$\{\xi^i, \xi^j\}_{a,b} = \omega_a^{ij} + \omega_b^{ij} \tag{95}$$

satisfies the Jacobi identity.

References

[1] Arnold V I 1982 *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 I—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)

[5] Arnold V I 1989 *Mathematical Methods of Classical Mechanics* (Berlin: Springer)

[6] Arnold V I, Kozlov V V and Neishtadt A I 1993 Mathematical aspects of classical and celestial mechanics *Dynamical Systems III—Encyclopaedia of Mathematical Sciences* 2nd edn, vol 3, ed V I Arnold (Berlin: Springer) pp 1–291

[7] Ecalle J 1992 *Ann. Inst. Fourier* **42** 73

[8] Il'yashenko Yu S and Yakovenko S Yu 1991 *Russ. Math. Surv.* **46** 3

[9] Belitsky G R 1978 *Russ. Math. Surveys* **33** 107; 1986 *Funct. Anal. Appl.* **20** 253; 1979 *Normal Forms, Invariants, and Local Mappings* (Kiev: Naukova Dumka) (in Russian)

[10] Elphick C, Tirapegui E, Brachet M E, Coullet P and Iooss G 1987 *Physica* **29D** 95

[11] Walcher S 1991 *Math. Ann.* **291** 293

[12] Bruno A D 1993 *Selecta Math.* (formerly *Sovietica*) **12** 13

[13] Bruno A D and Walcher S 1994 *J. Math. Anal. Appl.* **183** 571
[14] Cicogna G 1995 *J. Phys. A: Math. Gen.* **28** L179
[15] Olver P J 1986 *Applications of Lie Groups to Differential Equations* (New York: Springer)
[16] Bluman G W and Kumei S 1989 *Symmetries and Differential Tely Integrable Systems Equations* (New York: Springer)
[17] Stephani H 1989 *Differential Equations. Their Solution Using Symmetry* (Cambridge: Cambridge University Press)
[18] Ibragimov N H 1992 *Russ. Math. Surv.* **47** 89
[19] Winternitz P 1993 Lie groups and solutions of nonlinear PDEs *Integrable Systems, Quantum Groups, and Quantum Field Theory (NATO ASI C 409)* ed L A Ibort and M A Rodríguez (Dordrecht: Kluwer)
[20] Gaeta G 1994 *Nonlinear Symmetries and Nonlinear Equations* (Dordrecht: Kluwer)
[21] Olver P J 1995 *Equivalence, Invariants, and Symmetry* (Cambridge: Cambridge University Press)
[22] Kumei S and Bluman G W 1982 *SIAM J. Appl. Math.* **42** 1157
[23] Bluman G W 1990 *SIAM J. Appl. Math.* **50** 1689, 1706
[24] Bluman G W 1991 Linearization of PDEs *Group Theoretical Methods in Physics (Lecture Notes in Physics 382)* ed V V Dononov and V I Man'ko (Berlin: Springer)
[25] Frohlicher A and Nijenhuis A 1960 *Comm. Math. Helv.* **34** 227; 1956 *Ind. Math.* **18** 338
[26] Stone A P 1966 *Ann. Inst. Fourier* **15** 319
[27] Eiseman P R and Stone A P 1975 *Proc. Am. Math. Soc.* **53** 179
[28] Marmo G 1986 Nijenhuis operator in classical dynamics *Proc. of the Third Seminar on Group Theoretical Methods in Physics (Yurmala, May 22–24 1985)* ed M A Markov (Moscow: Nauka)
[29] De Filippo S, Marmo G, Salerno M and Vilasi G 1984 *Nuovo Cimento* B **83** 97
[30] Lax P J 1968 *Comm. Pure Appl. Math.* **21** 467
[31] Gel'fand I M and Dikii L A 1976 *Funct. Anal. Appl.* **10** 259
[32] Gel'fand I M and Dorfman I Ya 1979 *Funct. Anal. Appl.* **13** 248
[33] Magri F 1978 *J. Math. Phys.* **19** 1156
[34] Fokas A and Fuchsteiner B 1981 *Physica* **4D** 47
[35] Fuchsteiner B 1982 *Prog. Theor. Phys.* **68** 1082
[36] Klein J 1962 *Ann. Inst. Fourier* **12** 1
[37] Crampin M 1983 *J. Phys. A: Math. Gen.* **16** 3755
[38] Godbillon C 1969 *Geometrie Differentielle et Mecanique Analythique* (Paris: Hermann)
[39] De Filippo S, Marmo G and Vilasi G 1982 *Phys. Lett.* **117B** 418
[40] Carinena J F and Ibort L A 1984 *Lett. Math. Phys.* **8** 359
[41] Magri F, Morosi C and Ragnisco O 1985 *Comm. Math. Phys.* **99** 115
[42] Zakharov V E and Konopelchenko B G 1984 *Comm. Math. Phys.* **94** 483
[43] Marmo G, Saletan E J, Simoni A and Vitale B 1985 *Dynamical Systems. A Differential Geometric Approach to Symmetry and Reduction* (Chichester: Wiley)
[44] Giordano M, Marmo G and Rubano C 1993 *Inverse Problems* **9** 443
[45] Antonini P, Marmo G and Rubano C 1985 *Nuovo Cimento* B **86** 17
[46] Marmo G and Rubano C 1983 *Nuovo Cimento* B **78** 70