Normal Forms, Symmetry,

and Linearization of Dynamical Systems.

Dario Bambusi,
Dipartimento di Matematica, Università di Milano, via Saldini 50, 20133 Milano (Italy);
`bambusi@vmimat.mat.unimi.it`;

Giampaolo Cicogna,
Dipartimento di Fisica, Università di Pisa, P.zza Torricelli 2, 56126 Pisa (Italy);
`cicogna@ipifidpt.difi.unipi.it`;

Giuseppe Gaeta,
Department of Mathematics, Loughborough University, Loughborough LE11 3TU (G.B.);
`g.gaeta@lboro.ac.uk`;

Giuseppe Marmo,
Dipartimento di Scienze Fisiche, Università di Napoli, and *INFN, Sezione di Napoli, 80123 Napoli (Italy).*

Summary. 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.

23/5/96

Introduction.

It is well known that the same procedure – based on formal series of polynomial changes of coordinates – devised by Poincaré [1-4] to integrate linearizable dynamical systems[1] in the neighbourhood of a fixed point, can also be used to normalize, again in the neighbourhood of a fixed point, non-linearizable dynamical systems, i.e. system 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 analyzed, and asserted under certain condition, 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 – in the framework of the perturbative theory, i.e. in the theory of Poincaré-Dulac Normal Forms [1,2].

To this aim, we have naturally to consider the case of Normal Forms in the presence of symmetry. More specifically, it turns out that we have to consider Normal Forms 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 indeed linear, i.e. when the problem is already solved. Thus, rather than employing the classical theory of Normal Forms 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.

It should be mentioned that the connection between "suitable" symmetries and linearity of the Normal Form was actually already remarked – albeit *en passant* – in [9] (see remark 4 there); we want to discuss this in more detail for two reasons: on one side, for its practical relevance; and on the other because it shows how consideration of nonlinear symmetries in Normal Form theory really improves the results that can be obtained by considering only linear symmetries.

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

1. The main result.

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

$$X \ = \ \sum_{i=1}^n \ A_{ij} \ y_j \ \frac{\partial}{\partial y_i} \ , \tag{1}$$

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

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

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

$$S \ = \ \sum_{i=1}^n \ y_i \ \frac{\partial}{\partial y_i} \ , \tag{3}$$

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 linearizability of nonlinear systems.

Let us now consider a (formal) near-identity[2] nonlinear change of coordinates,

$$y_i = \phi_i(x, ..., x_n) \ . \tag{4}$$

Under this, the linear dynamical system

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

is changed in general into a nonlinear system

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

it is immediate to see that the functions $f^i(x)$ are given by

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

where J is the jacobian of the coordinate transformation,

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

[2] By this we mean that $(D\phi)(0) = I$.

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[3], so that (7) makes sense. Again by the near-identity of (4), we are guaranteed that $(Df)(0) = A$.

In the new coordinates, X is expressed as

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

Similarly, S is now expressed as

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

where the p_i are nonlinear functions given explicitely by

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

The geometrical relations between geometrical objects (in particular, vector fields) do not depend, however, on the choice of coordinates; thus, $[X,S] = 0$ continue to hold. We recall that, in terms of the components of the vector fields, this means that

$$\{f,p\}^i \equiv \left(f^j \cdot \partial_j\right) p^i - \left(p^j \cdot \partial_j\right) f^i = 0 \qquad \forall i = 1, ..., n \ . \tag{12}$$

`Remark 1.` 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, e.g., the 1-dimensional harmonic oscillator

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -q , \end{aligned}$$

and a (nonlinear) transformation

$$\begin{aligned} Q &= f(q^2+p^2)\ q \\ P &= f(q^2+p^2)\ p , \end{aligned}$$

[3] It should be noted that the theory can also be formulated in terms of Lie transformations [11,12]: in this case the coordinate transformation correspond 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 the sake of simplicity.

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

$$\begin{aligned} \dot{Q} &= P \\ \dot{P} &= -Q \ , \end{aligned}$$

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

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

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. ⊙

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 the Poincaré-Dulac theory fails to recognize immediately[4] the intrinsically linear nature of the system, as e.g. in the examples we consider below.

If we study the symmetry properties of (6), i.e. the vector fields[5] $Y = p^i(x)\partial_i$ for which (12) is satisfied, we know that *if (6) is linearizable, then (12) admits at least one solution with p given by (11).* Notice that, in particular, this means that – also in the x coordinates – the linearization of Y is just given by $(DY)(0) = I$; this is again a consequence of the fact (4) is near-identity. (Here and in the following we use the short notation $(DY)(0) = M$, with $Y = p^i\partial_i$ a vector field and M a matrix, to mean that $(Dp)(0) = M$).

It is interesting, and we want to remark, that the converse is also true; i.e., we have that:

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 $Y = p^i(x)\partial_i$, then X is linearizable by a formal near-identity change of coordinates. Moreover, there is a transformation 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 [9,10] (see also [13]), which we report here for completeness, in a form suitable for our present purpose[6]; in order to state this we have to introduce some terminology.

We will 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]).

---

[4] By this, we mean that the Normal Form unfolding allows for nonlinear terms. Obviously, if one was going to actually perform the normalization – rather than just determining the general Normal Form unfolding corresponding to the linear part – there would be a series of “miracoulous” cancellations, so that the coefficients of the nonlinear resonant terms would vanish.

[5] Here and in the following $\partial_i = (\partial/\partial x_i)$ and sum over repeated indices is understood.

[6] A stronger form of this theorem also exists, see [9,13].

We will also refer to (Semisimple) Joint Normal Forms: denote by $\mathcal{A}, \mathcal{B}, \mathcal{A}_s, \mathcal{B}_s$ the homological operators [1] associated to A, B, A_s, B_s[7]. We say that

$$X = \widetilde{f}^i(y)\,(\partial/\partial y_i) \quad , \quad Y = \widetilde{s}^i(y)\,(\partial/\partial y_i) \ , \tag{13}$$

where

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

are in Semisimple Joint Normal Form if both $\widetilde{F}$ and $\widetilde{G}$ are in $\mathrm{Ker}(\mathcal{A}_s) \cap \mathrm{Ker}(\mathcal{B}_s)$, and that they are in Joint Normal Form[8] if both $\widetilde{F}$ and $\widetilde{G}$ are in $\mathrm{Ker}(\mathcal{A}) \cap \mathrm{Ker}(\mathcal{B})$.

With this notation, we can state the

Proposition. *Let the polynomial vector fields X and Y in R^n, expressed in the x coordinates as*

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

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 $A = A_s + A_n$, $B = B_s + B_n$; let F, G be the nonlinear parts of f, g, so that $f = Ax + F$, $g = Bx + G$. Then, by a formal series of near-identity (Poincaré) transformations, it is possible to reduce f, g to Joint Semisimple Normal Form*; if A and B are normal matrices, then it is possible to reduce f, g to* Joint Normal Form.

`Proof of Theorem 1.` Under the hypotheses of the above Theorem, $B = I$, so that $B_s = B = I$; in this case in particular $\mathrm{Ker}(\mathcal{B}) = \mathrm{Ker}(\mathcal{B}^+) = \mathrm{Ker}(\mathcal{B}_s) = \mathcal{K}$, and, indipendently of A, we can, due to the proposition, transform f to $\widetilde{f}$ with $\widetilde{F} \in \mathcal{K}$. It is a general result that $\mathcal{K} = \mathrm{Ker}(\mathcal{B}^+)$ consists of vector polynomials which are *resonant with* B; these are characterized as follows [1]. Let $\lambda_1, ..., \lambda_n$ be the eigenvalues of B; then $\mathcal{K}$ is spanned by vectors $v = (v_1, ..., v_n)$ which have all components equal to zero except for v_r, which is given by $v_r(x) = x^{m_1} ... x_n^{m_n}$, where the m_i are nonegative integers[9] 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 $B = I$, $\widetilde{F} \in \mathcal{K}$ actually means that $\widetilde{F} = 0$, and therefore $\widetilde{f}(y) = Ay$. This proves the Theorem. $\odot$

---

[7] With the notation introduced in (12), $\mathcal{A} = \{Ax, .\}$, etc.

[8] Recall [1,2,15] that for A normal, $\mathrm{Ker}(\mathcal{A}) = \mathrm{Ker}(\mathcal{A}^+)$; recall also that $\mathrm{Ker}(\mathcal{A})$ and $\mathrm{Ker}(\mathcal{A}^+)$ are contained in $\mathrm{Ker}(\mathcal{A}_s)$.

[9] 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.

`Remark 2.` It should be stressed that the above proof would also work if B was not the identity, but was a matrix such that its semisimple part B_s does not admit resonances; this is in particular the case when all the eigenvalues of B_s – and therefore of B – have positive (negative) real parts. $\odot$

2. 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 for short) required to take the system (6) into normal form 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]), either 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 $\sigma_1, ..., \sigma_n$ belong to a Poincaré domain if and only if $\epsilon \mathrm{Re}(\sigma_i) > 0$ for all the 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].

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 normal form is convergent.*

`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 be not 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. The theorem 1 guarantees that this transformation does also take X into normal form, and thus that X can be linearized by means of a convergent change of coordinates. $\odot$

`Remark 3.` Similarly to what remarked concerning theorem 1, again the above proof would work (and theorem 2 apply) for more general symmetries: e.g., it would suffice to require that the eigenvalues of B belong to a Poincaré domain (see also section 4). $\odot$

Changing point of view, and focusing on the convergence of the normalizing transformation [16-20], the above arguments can be summarized and completed in the following form (the proof of which is contained in the above discussion and therefore omitted):

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 normalizing transformation which linearizes X is convergent if and only if this symmetry is analytic.*

3. Families of vector fields.

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

We are also assuming that the dynamics is linearizable with respect to the same dilation field S and to the same linear structure (i.e., independently of the parameter μ), or also – more in general – that there is continuity with respect to μ. Therefore, the linearizing transformation will depend (smoothly) on μ; but the arguments given in theorem 2 remain valid, and we are guaranteed to have a μ-dependent family of convergent normalizing transformations.

We could also arrive at the same conclusion by repeating the discussion given at the beginning of section 1, considering now a μ-dependent linear system and a μ-dependent change of coordinates.

In this framework, (1) and (4) would now be

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

$$y_i = \phi_i(x_1, ..., x_n; \mu) \ . \tag{4'}$$

Under this, the linear dynamical system

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

is changed into the nonlinear system

$$\dot{x}^i \ = \ f^i_\mu(x) \ ; \tag{6'}$$

and thus in the new coordinates we have

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

We will not bore the reader by repeating any further our previous discussion in the parameter-dependent case.

`Remark 4.` 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 e.g. the eigenvalues $\lambda_i(\mu)$ of $B(\mu)$ belong to a Poincaré domain for all values of μ; see also the next section. $\odot$

`Remark 5.` In order to avoid possible confusion, we would like to briefly consider the case where we have a family of vector fields X_μ 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 normal form the whole family of vector fields X_μ at once; this appears surprising and quite paradoxical. Notice however that this is the case only if B does not admit resonances, and in this case the X_μ would be linear. 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. $\odot$

4. Symmetries not having the identity as 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 2-4 above, and we are now going to discuss it a little further, in particular with reference to the problem of convergence of the normalizing transformation.

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

Let us recall briefly, for the sake of completeness, what are the requirements for B to satisfy "condition ω": we ask that B is semisimple, and denote by λ_i its eigenvalues. Consider then the set of $Q = (q_1, \ldots, q_n)$ such that $(Q,\ \Lambda) \neq 0$, where

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

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

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

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 non-linear) i.e. can be brought to be in $\mathrm{Ker}(\mathcal{B})$, which is in this case finite dimensional. Indeed, in this case we can take Y into normal form, and we are guaranteed the required normalizing transformation 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 $f \in \mathrm{Ker}(\mathcal{B})$ (see [13] for more detail).

In the same vein, we can contemplate the case in which $\mathrm{Ker}(\mathcal{A}) \cap \mathrm{Ker}(\mathcal{B}) = \{0\}$ (this situation is considered in example 3 below); in this case both X and Y are linearized when taken to the Joint Normal Form[10], but the (joint) normalizing transformation is, without further assumptions, in general only formal.

If A satisfies condition ω, then f can be taken into Normal Form – i.e. in $\mathrm{Ker}(\mathcal{A})$ – 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 $f \in \mathrm{Ker}(\mathcal{B})$, and it can be easily verified (see [13]) that this condition is preserved by any transformation taking f into Normal Form; i.e., f in $\mathrm{Ker}(\mathcal{A})$ implies $f \in \mathrm{Ker}(\mathcal{A}) \cap \mathrm{Ker}(\mathcal{B})$ and therefore the Normal Form of f is necessarily linear.

If it is instead B to satisfy condition ω, then f can be taken to be in $\mathrm{Ker}(\mathcal{B})$ 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.

The above 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 $\mathrm{Ker}(\mathcal{B}) = \{0\}$ and B satisfies condition ω, then X and Y can be linearized by a convergent transformation;* `b)` *If $\mathrm{Ker}(\mathcal{A}) \cap \mathrm{Ker}(\mathcal{B}) = \{0\}$, then X and Y can be linearized (possibly by means of a non-convergent transformation), but this transformation is convergent in the case Y is linear, $Y = (Bx)^i\partial_i$. Also, if $A = (Df)(0)$ (respectively $B = (Dg)(0)$) is semisimple and satisfies condition ω, then there is a convergent transformation taking X (respectively Y) into Normal (not necessarily linear) Form.*

It is worth to emphasize 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 continously with μ, so that generically [22], for almost all values of μ the eigenvalues are strongly nonresonant, and therefore, by Siegel therem the system can be linearized by an analytic change of coordinates. Moreover, there is smooth (in the sense of Whitney) dependence of the linearizing transformation on μ, when μ varies in the Cantor set to which correspond strongly nonresonant frequencies. However, Siegel theorem does not give any

[10] This remark was already presented in [9].

information on the behaviour of the system when the parameter belongs to the bad set (the complementary of the above Cantor set).

Theorem 4 ensures that, provided there exists at least one μ dependent simmetry (with suitable properties), system (6') can be linearized *for any value of the parameter*, and the linearizing transformation depends smoothly on μ, when μ belongs to an interval, i.e. a regular set, and not a Cantor set.

We would like, to conclude this section, to point out that it could happen that a symmetry with linear part the identity cannot immediately be determined, but 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 σ_1, σ_2; assume that X admits a symmetry $Y = s^i(x)\partial_i$ whose linear part B is semisimple and not proportional to A[11].

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[12].

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 (e.g. abelian) Lie algebra[13]. Notice that, in particular, if X is linearizable in the form (1) and the A^k are independent, these could be the vector fields $Y_{(k)}$ considered in (2).

5. Several symmetries having the identity as 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 the Y in Taylor series around

[11] Obviously one does not have to ask explicitely all of these conditions: the existence of distinct eigenvalues ensures A is semisimple, and then $[A, B] = 0$ implies B is semisimple as well.

[12] It is for this reason that in example 3 below we will have to consider a three dimensional system.

[13] For full details and generalizations, we refer again to [13].

0), and therefore $[Y_i, Y_j]$ has linear part the identity. However, there is no reason for $[Y_i, Y_j]$ to vanish in general.

According to our Theorem 1, we can choose each of the Y_i, and simultaneously linearize X and Y_i. It should be stressed that in this way we do not, in general, linearize the other Y_j's; more precisely, not only we do not linearize them by the normalizing transformation adapted to the pair (X, Y_i), but in general we are not able to linearize simultaneously different Y's.

This fact has to do with the problem of simultaneously reducing to Joint Normal Form a general algebra of vector fields [10,13,23]. Referring again to [13] for a full discussion, we mention here just the case of algebras of vector fields having semisimple linear part. 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.

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 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 6.` This is also maybe 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 viceversa). Similarly, in general the transformation which linearizes X and Y will not be unique. $\odot$

6. Examples.

We will now consider some very simple example, and show how we can apply our results to guarantee that a concrete nonlinear system can be linearized, without actually performing the Poincaré normalization.

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

$$\begin{aligned} \dot{x} &= x \\ \dot{y} &= 3y - x^2 \ . \end{aligned} \tag{19}$$

The linear part of this is given by the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \ , \tag{20}$$

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

$$\begin{aligned} \dot{x} &= x \\ \dot{y} &= 3y + \alpha x^3 \end{aligned} \tag{21}$$

with α an arbitrary real constant.

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

$$Z_1 = x\partial_x + 2x^2\partial_y \qquad ; \qquad Z_2 = (y - x^2)\partial_y \tag{22}$$

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

$$Y = x\partial_x + (y + x^2)\partial_y \ , \tag{23}$$

we have found 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 is appearing in the vector field, and the eigenvalues belong to a Poincaré domain, so that the convergence of the normalizing transformation is guaranteed. Thus, the only nontrivial result is that the term αx^3 in (21) does actually disappear from the normalized form. In the following examples, we consider more complicate cases.

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

$$\begin{aligned} \dot{x} =& x + x^4 y \\ \dot{y} =& -2y - x^2y^2 \ . \end{aligned} \tag{24}$$

The linear part of this is given by the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix} \ , \tag{25}$$

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

$$\begin{aligned} \dot{x} =& \ x + x\phi_1(x^2y) \\ \dot{y} =& -2y + y\phi_2(x^2y) \ . \end{aligned} \tag{26}$$

where ϕ_i are two arbitrary (smooth) functions.

A symmetry vector field of this is given by

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

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

Notice that in this case $\mathrm{Ker}(\mathcal{A})$ 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:

$$\begin{aligned} \dot{x} &= \ \ x + a_1 x^3 y + b_1 x y^2 z \\ \dot{y} &= - \ 3y + a_2 x^2 y^2 + b_2 y^3 z \\ \dot{z} &= 9z + a_3 y^2 z^2 + b_3 y^2 z^2 \ , \end{aligned} \tag{28}$$

where $a_i, \ b_i$ 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 $\mathrm{Ker}(\mathcal{A})$ and $\mathrm{Ker}(\mathcal{B})$ 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 Sect. 4, we can conclude that X can be linearized by a convergent transformation.

`Example 4.` In this example, we merely want to point out a case in which we have several noncommuting 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 is immediate to check 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, g such that

$$f(0) = 0 \qquad ; \qquad g(0) = 1 \tag{35}$$

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

We can rewrite such vector fields as

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

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

One can readily check that

$$[Y_{f,h}, Y_{\phi,\eta}] = 2r^2 \left(h(r^2)\phi'(r^2) - \eta(r^2)f'(r^2)\right) \tag{37}$$

and thus that the $Y_{f,h}$ do not commute among themselves.

Thus, in general – i.e. unless (37) vanishes – we are not able, as one would indeed expect, to simultaneously linearize different $Y_{f,h}$.

`Example 5.` As a final example, we consider an apparently hopelessly complicate system, i.e.

$$
\begin{aligned}
\dot{x} = f_1(x,y,z) = & \left[\alpha x - y\right] - x^2 - \left[3xy^2 + 2\alpha y^3\right] \\
& - 6\left[x^3 y + \alpha x^2 y^2\right] - 3\left[x^5 + 2\alpha x^4 y + y^5\right] - \left[2\alpha x^6 + 15 x^2 y^4\right] \\
& - 30\left[x^4 y^3 + x^6 y^2\right] - 3\left[x^{10} + 5x^8 y\right] \\
\dot{y} = f_2(x,y,z) = & \left[x + \alpha y\right] - \left[\alpha x^2 - 2xy\right] + \left[2x^3 + y^3\right] \\
& + \left[9x^2y^2 + 4\alpha x y^3\right] + \left[15x^4 y + 12\alpha x^3 y^2\right] + \left[7x^6 + 12\alpha x^5 y + 6xy^5\right] \\
& + \left[4\alpha x^7 + 30 x^3 y^4\right] + 60x^5y^3 + 60x^7y^2 + 6\left[5x^9 y + x^{11}\right] \\
\dot{z} = f_3(x,y,z) = & + \beta z + \left[2xy + (2\alpha - \beta)\, y^2\right] \\
& + \left[2x^3 + 2\,(2\alpha-\beta)\, x^2 y + 3xy^2 + \alpha y^3 + (2\alpha - \beta)\, y^3\right] \\
& + \left[(2\alpha - \beta)\, x^4 - 3\alpha x^2 y^2 + 6xy^3 + 2y^4\right] + \left[6x^3y^2 + 8x^2y^3 + 3y^5\right] \\
& + \left[12x^4y^2 + 27x^2y^4 + 12\alpha x y^5\right] + 9\left[5x^4y^3 + 4\alpha x^3 y^4\right] \\
& + \left[2x^8 + 8x^6 y + 21 x^6 y^2 + 36\alpha x^5 y^3 + 18xy^7\right] + \left[12\alpha x^7 y^2 + 90 x^3 y^6\right] \\
& + 180x^5y^5 + 180x^7y^4 + 90x^9y^3 + 18x^{11}y^2 \ .
\end{aligned}
\tag{38}
$$

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 complicate form of (38), one can check explicitely 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 of this example, $(x_1,x_2,x_3) = (x,y,z)$), where

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

The linear part of this vector field corresponds just to the identity matrix, and thus our theorem guarantees immediately that X can be reduced to its linear part; i.e. in Normal Form 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

$$\begin{aligned} y_1 =& x_1 + \left(x_1^2 + x_2\right)^3 \\ y_2 =& x_2 + x_1^2 \\ y_3 =& x_3 - x_2^3 - \left(x_1^2 + x_2\right)^2 \ . \end{aligned} \tag{43}$$

and Y is, in the y coordinates, nothing else 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

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

which depends on α and β, and which obviously commutes with X (and with Y).

We stress that now we have a two parameters family of systems (42), and that as α, β 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.

7. Some further remarks.

`Remark 7.` 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 to obtain relevant informations 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. $\odot$

`Remark 8.` In general, one could try to determine perturbatively the functions $p^i(x)$, by expanding them in homogeneous terms as $p^i(x) = \sum_{m=1}^\infty p^i_m(x)$, where $p^i_m(ax) = a^m p^i_m(x)$, 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 always consider 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 the order k of the highest order resonance of the system (e.g., in example 1 above one would only had to go up to order three). Notice that we are guaranteed that this order is finite when the eigenvalues belong to a Poincaré domain. $\odot$

`Remark 9.` Finally, we would like to point out that the present approach is related to the study of integrability conducted by Marmo and collaborators, see e.g. [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]. $\odot$

Acknowledgement.

One of the authors (G.G.) would like to thank the Mathematische Forschunginstitut Oberwolfach for hospitality during part of this work. This stay was supported by the Volkswagen Stiftung under the RiP program.

References

[1] V.I. Arnold, "Geometrical methods in the theory of differential equations", Springer, Berlin 1988

[2] V.I. Arnold and Yu.S. Il'yashenko, "Ordinary differential equations"; in *Encyclopaedia of Mathematical Sciences - vol. I, Dynamical Systems I*, (D.V. Anosov and V.I. Arnold eds.), p. 1-148, Springer, Berlin 1988

[3] A.D. Bruno, "Local methods in nonlinear differential equations", Springer, Berlin 1989

[4] F. Verhulst, "Nonlinear differential equations and dynamical systems", Springer, Berlin 1990

[5] G. Gaeta and G. Marmo, "Nonperturbative linearization of dynamical systems", Preprint 1996 (`mp-arc 96-27`)

[6] G. R. Belitsky, *Russ. Math. Surveys* **33** (1978) 107

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

[8] G. Iooss and M. Adelmeyer "Topics in bifurcation theory and applications, World Scientific, Singapore 1992

[9] G. Cicogna and G. Gaeta, *Journ. Phys. A: Math. Gen.* **27** (1994) 461

[10] G. Cicogna and G. Gaeta, *Journ. Phys. A: Math. Gen.* **27** (1994) 7115

[11] G. Benettin, L. Galgani and A. Giorgilli, *Nuovo Cimento B* **79** (1984), 201

[12] Yu. A. Mitropolsky and A.K. Lopatin, "Nonlinear mechanics, groups and symmetry", Kluwer, Dordrecht 1995

[13] G. Cicogna and G. Gaeta, "Nonlinear symmetries and Poincaré - Birkhoff normal forms", in preparation

[14] M.W.Hirsch and S. Smale "Differential equations, dynamical systems, and linear algebra", Academic Press, Orlando 1974

[15] A.N. Kolmogorov and S. Fomine, "Elements of the Theory of Functions", Pergamon

[16] L.M. Markhashov, *J. Appl. Math. Mech.* **38** (1974), 788

[17] S. Walcher, *Math. Ann.* **291** (1991), 293

[18] A.D. Bruno, *Selecta Mathematica* **12** (1993), 13

[19] A.D. Bruno and S. Walcher, *J. Math. Anal. Appl.* **183** (1994), 571

[20] G. Cicogna, *J. Phys. A* **28** (1995), L179; *J. Math. Anal. Appl.* **199** (1996), 243

[21] H. Ito, *Comm. Math. Helv.* **64** (1989), 412; *Math. Ann.* **292** (1992), 411

[22] V.G. Sprindzuk: "Metric Theory of Diophantine Approximation", Winston and Sons, Washington 1979

[23] D. Arnal, M. Ben Ammar and G. Pinczon, *Lett. Math. Phys.* **8** (1994), 467

[24] A.A. Kirillov, "Elements of the Theory of Representations", Springer, Berlin 1976

[25] G. Cicogna and G. Gaeta, *Nuovo Cimento B* **109** (1994), 989

[26] S. De Filippo, G. Vilasi, G. Marmo and M. Salerno, *Nuovo Cimento B* **83** (1984), 97

[27] G. Marmo, in "Proceedings of the third seminar on Group Theoretical Methods in Physics" (Yurmala, May 22-24 1985), Nauka, Moscow 1986

[28] P.J. Olver, "Applications of Lie groups to differential equations", Berlin, Springer 1986

[29] L.V. Ovsjannikov, "Group properties of differential equations" (Novosibirsk, 1962; English transl. by G. Bluman, 1967); "Group analysis of differential equations", Academic Press, New York 1982

[30] G.W. Bluman and S. Kumei, "Symmetries and differential equations", Springer, Berlin 1989

[31] G. Gaeta, "Nonlinear symmetry and nonlinear equations", Kluwer, Dordrecht 1994