Convergence of normal form transformations: The role of symmetries

Giampaolo Cicogna
Dipartimento di Fisica, Università di Pisa
Via Buonarroti 2, Ed.B, 56127 Pisa, Italy
cicogna@df.unipi.it

Sebastian Walcher
Zentrum Mathematik, TU München
80290 München, Germany
and
Institut für Biomathematik und Biometrie, GSF-Forschungszentrum
Postfach 1129, 85758 Neuherberg, Germany
walcher@mathematik.tu-muenchen.de

Abstract

We discuss the convergence problem for coordinate transformations which take a given vector field into Poincaré-Dulac normal form. We show that the presence of linear or nonlinear Lie point symmetries can guarantee convergence of these normalizing transformations, in a number of scenarios. As an application, we consider a class of bifurcation problems.

1 Normal forms and normalizing transformations.

Normal form theory [1, 2, 6, 7, 9, 24] (see also [14], where many other references can be found) was introduced by Poincaré in his thesis, as a tool to integrate nonlinear systems. While it is now known that integration is in general impossible, normal forms have proven to be among the most useful tools both in the qualitative and quantitative local analysis of dynamical systems.

We will consider an ordinary differential equation

$$\dot{x} \equiv \frac{\mathrm{d}x}{\mathrm{d}t} = f(x) \qquad\qquad x = x(t) \tag{1.1}$$

and the associated vector field X_f (we will sometimes call both X_f and f a vector field)

$$X_f \equiv \sum_{i=1}^{n} f_i(x) \frac{\partial}{\partial x_i} \qquad\qquad (x \in \mathbf{R}^n) \tag{1.2}$$

which we assume to be analytic in a neighbourhood of a stationary point x_0 (i.e., a point such that $f(x_0) = 0$; we can choose $x_0 = 0$). The basic idea is to introduce a near-identity change of coordinates in order to eliminate nonlinear terms in the given vector field; certain (appropriately chosen) terms which cannot be eliminated, the so-called "resonant terms", constitute the Poincaré-Dulac normal form. We will elaborate details in a moment. The coordinate transformations are usually obtained by means of iterative techniques: in general the normalizing transformations are actually formal power series transformations, and only special conditions can ensure their (local) convergence and the (local) analyticity of the normal form [1, 2, 6, 7].

The procedure is well known; we recall some essential points, to fix notations, and in view of the applications below (see also [15]).

Writing the system (1.1) in the form

$$\dot{x} = f(x) = Ax + F(x) \tag{1.1'}$$

we will always assume that the matrix $A = (Df)(0)$ is nonzero and semisimple, with eigenvalues $\lambda_1, \ldots, \lambda_n$. A normal form of f will be written

$$\widehat{f}(x) = Ax + \widehat{F}(x). \tag{1.3}$$

(The notation $\widehat{\cdot}$ will be always reserved for vector fields in normal form; there is no danger of confusion if x will also be used to denote the "new" coordinates.) The characterizing property of Poincaré-Dulac normal forms can be stated as follows: If $e_1, \ldots, e_n$ is an eigenbasis of A, and $x_1, \ldots, x_n$ denote the corresponding coordinates, then $\widehat{F}$ is a series in those monomials

$$x_1^{m_1} \cdots x_n^{m_n}\, e_j \qquad \text{with} \qquad m_1\lambda_1 + \ldots m_n\lambda_n - \lambda_j = 0. \tag{1.4}$$

We will say that the eigenvalues are in resonance if they satisfy a relation as above with nonnegative integers m_i, $\sum m_i \geq 2$.

There is always a formal power series transformation to normal form (but neither the normal form nor the normalizing transformation are unique in general, see e.g. [1, 5, 6, 7]). A sufficient condition which ensures convergence of a normalizing transformation is given by the following criterion (see [1, 2] for more details and extensions):

Theorem 1.1 (Poincaré) *If the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the matrix $A = (Df)(0)$ belong to the Poincaré domain, i.e. if the convex hull of the points $\lambda_1, \ldots, \lambda_n$ in the complex plane does not contain zero, then there is a convergent normalizing transformation (and an analytic normal form).*

Building on work by Siegel and Pliss, Bruno succeeded in providing deep insight into convergence and divergence issues. He formulated two fundamental conditions which ensure the convergence of a normalizing transformation (in an open neighborhood of 0). These are called Condition A and Condition ω (see [6, 7] for details). In order to avoid technicalities, we state Condition A in its simplest (and slightly too restrictive) form:

Condition A: A normal form $\widehat{f}$ is said to satisfy Condition A if $\widehat{f}$ has the form

$$\widehat{f} = Ax + \alpha(x)Ax$$

where $\alpha(x)$ is some scalar-valued power series (with $\alpha(0) = 0$).

It follows from $\widehat{f}$ being in normal form that $X_{\widehat{f}}(\alpha) = 0$; one has also $X_A(\alpha) = 0$ (in the case of linear vector fields $f = Ax$, we shall simply write X_A instead of X_{Ax}). Depending on the position of the eigenvalues of A in the complex plane, Condition A should be modified appropriately [6, 7]. However, in all the applications we are going to discuss, the above formulation is sufficient. Bruno's condition improves an earlier criterion given by Pliss [23], which requires some formal normal form to be linear.

The other condition is a (weak) arithmetic condition, generalizing a criterion given earlier by Siegel, and devised to control the appearance of small divisors.

Condition ω: Let $\omega_k = \min |(Q,\Lambda) - \lambda_j|$ for all $j = 1, \ldots, n$ and $n-$tuples of integers $q_i \geq 0$ such that $1 < \sum_{i=1}^n q_i < 2^k$ and $(Q,\Lambda) \equiv \sum_i q_i \lambda_i \neq \lambda_j$: then

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

With these two conditions, one has [6, 7]:

Theorem 1.2 (Bruno) *If $A = (Df)(0)$ satisfies Condition ω, and if f can be transformed, via a formal coordinate transformation, to some $\widehat{f}$ which satisfies Condition A, then there is a convergent normalizing transformation for f.*

Here and in the following, "convergent" stands for "convergent in some open neighbourhood of the stationary point $x_0 = 0$".

Condition ω is satisfied by almost all (in the Lebesgue sense) $n-$tuples of eigenvalues λ_i. For the sake of simplicity, we will assume that the matrix A satisfies Condition ω, unless we explicitly say otherwise.

Bruno also stated and proved divergence theorems which show that weaker versions of Condition A and Condition ω are necessary to ensure convergence in the general setting. These theorems do not, however, address the convergence problem for a given analytic vector field, and there are many instances of convergence where A satisfies neither Condition A nor Condition ω. This is the central question to be discussed in this article, and symmetries play a fundamental role here.

2 Symmetries and invariants.

We now start to investigate the role of symmetry properties of the vector field in the convergence problem for the normalizing transformation [2, 16, 18] (see also [13, 14]). In this paper we shall consider only Lie point symmetries: given the vector field X_f, we will say that the vector field

$$X_g \equiv \sum_{i=1}^n g_i(x) \frac{\partial}{\partial x_i} \tag{2.1}$$

is an infinitesimal (Lie point) symmetry for X_f if

$$[X_f, X_g] \equiv X_f X_g - X_g X_f = 0 \tag{2.2}$$

or, equivalently,

$$[f,g](x) \equiv Dg(x)f(x) - Df(x)g(x) \;=\; 0, \qquad \text{all } x. \tag{2.2$'$}$$

We will then also say that X_g is in the centralizer of X_f (or that g is in the centralizer of f). The vector field X_g provides the Lie generator of a local one-parameter group of symmetries [21, 22] of the differential equation. We will have to deal with various classes of centralizers. By $\mathcal{C}_{\rm an}(f)$ we denote the Lie algebra of local analytic centralizer elements of f, while $\mathcal{C}_{\rm for}(f)$ is the algebra of all formal power series vector fields commuting with f.

Moreover, we call a scalar-valued function φ an invariant (or a first integral) of f if $X_f(\varphi) = 0$.

The presence of a symmetry can be of considerable help in the normalizing procedure, not only in the general problem of computing normal forms (symmetries may impose strong restrictions on the explicit expression of normal forms, see e.g. [13, 14]), but also in the study of the convergence of a normalizing transformation. Let us recall some well known and useful facts (see [13, 14, 15, 16, 18]).

First, the "resonant terms", which constitute the nonlinear part $\widehat{F}(x)$ that cannot be eliminated in the normal form $\widehat{f}(x) = Ax + \widehat{F}(x)$, are precisely the terms such that

$$\widehat{F}(x) \in \mathrm{Ker}(\mathrm{ad}\, A) \tag{2.3}$$

where $\mathrm{ad}\, A$ is the "homological operator" defined by

$$\mathrm{ad}\, A(h) = [Ax, h] \tag{2.4}$$

This characterization of the resonant terms leads to a characterization of normal forms in terms of symmetry properties:

Proposition 2.1 *The vector field $f = Ax + F$ is in normal form if and only if $[Ax, f] = 0$.*

Thus, vector fields in normal form $\widehat{f}$ always admit nontrivial commuting vector fields g (i.e. $[g, \widehat{f}] = 0$, $g \notin \mathbf{R} f$: this follows from the above proposition if $\widehat{f} \neq Ax$, in case $\widehat{f} = Ax$ there are nontrivial linear commuting vector fields, e.g. $g = A^k x$, or $g = Ix$ where I is the identity matrix.) A fundamental property of normal forms is the following [25]:

Proposition 2.2 *Every (infinitesimal) symmetry*

$$g(x) = Bx + G(x) \tag{2.5}$$

of a normal form $\widehat{f}$ is also a symmetry of the linear part Ax of $\widehat{f}$. Every invariant of $\widehat{f}$ is also an invariant of the linear part Ax.

Concerning the interplay between normal forms and symmetry properties, we have the following first results. Given a commuting field $g(x) = Bx + G(x)$, we will assume here, as we did for the matrix A, that B satisfies Condition ω and that B is semisimple.

Proposition 2.3 *If f admits a linear (infinitesimal) symmetry $g_B(x) = Bx$, then there is a normal form $\widehat{f}$ which also admits this symmetry. If f admits a (analytic or formal) symmetry $g = Bx + G(x)$, then Bx is a symmetry of some normal form $\widehat{f}$, thus $\widehat{F}$ also commutes with Bx: $\widehat{F} \in \mathrm{Ker}(\mathrm{ad}\, A) \cap \mathrm{Ker}(\mathrm{ad}\, B)$.*

Proof (sketch): Transform g to normal form $\widehat{g} = Bx + \dots$. Then this transformation sends f to some $\tilde{f} = Ax + \dots$ such that $[Bx, \tilde{f}] = 0$. Now there exists a normalizing transformation for $\tilde{f}$ which respects the commuting vector field Bx (see [14, 18]). □

If g can be transformed to normal form by a convergent transformation then, obviously, f will be transformed to some $\tilde{f}$ which admits the linear symmetry Bx. The next statement is more substantial.

Theorem 2.4 *Let the system $\dot{x} = f(x) = Ax + F$ admit an analytic symmetry $g(x) = Bx + G(x)$, as above. If $\mathrm{Ker}\,(\mathrm{ad}\, A) \cap \mathrm{Ker}\,(\mathrm{ad}\, B)$ contains only linear vector fields then there is a convergent transformation of f to normal form, and both transformed vector fields are linear.*

Proof (sketch): There is a formal transformation to normal form $\widehat{f}$ that also commutes with Bx. The hypothesis forces the normal form to be linear, hence $\widehat{f} = Ax$, and Condition A (even the Pliss condition [23]) is satisfied. □

Example. With $x \equiv (x_1, x_2, x_3) \in \mathbf{R}^3$, consider the following system:

$$\begin{aligned} \dot{x}_1 &= \quad x_1 + a_1 x_1^3 x_2 + b_1 x_1 x_2^2 x_3 \\ \dot{x}_2 &= -3x_2 + a_2 x_1^2 x_2^2 + b_2 x_2^3 x_3 \\ \dot{x}_3 &= \quad 9x_3 + a_3 x_2^2 x_3^2 + b_3 x_2^2 x_3^2 \end{aligned}$$

where a_i, b_i are arbitrary constants. This system admits a linear commuting vector field

$$X_g = Bx \cdot \nabla = x_1 \partial_{x_1} - 2x_2 \partial_{x_2} + 4x_3 \partial_{x_3}$$

It can be verified that all the above assumptions are satisfied, in particular $\mathrm{Ker}(\mathrm{ad}\,A) \cap \mathrm{Ker}(\mathrm{ad}\,B)$ contains only linear vector fields, although both $\mathrm{Ker}(\mathrm{ad}\,A)$ and $\mathrm{Ker}(\mathrm{ad}\,B)$ have infinite dimension, and neither the eigenvalues of A nor those of B belong to a Poincaré domain. We conclude that this vector field can be linearized by a convergent transformation.

An especially important case occurs for $B = I$, the identity; see [4]:

Corollary 2.5 *A system $\dot{x} = f = Ax + F$ can be formally linearized if and only if it admits a formal symmetry $g = Bx + G$ such that $B = Dg(0) = I$. If g is analytic then there is a convergent transformation to normal form.*

As one more illustration of how normal forms may be influenced by symmetries, we quote without proof a recent result [17]:

Theorem 2.6 *Let $\mathcal{M}$ be the Lie algebra of a compact linear group, and suppose that the elements of $\mathcal{M}$ commute with the vector field $f = Ax + F$. If the elements of $\mathbf{R}\,Ax + \mathcal{M}$ admit no non-constant common polynomial invariant then every normal form $\widehat{f}$ of f is necessarily a polynomial.*

3 Dimension two.

In this section we will discuss analytic two-dimensional systems

$$\dot{x} = f(x) = Ax + F(x),$$

with A having eigenvalues λ_1 and λ_2, not both zero. A normal form of f will, as usual, be denoted by $\widehat{f}$. In dimension two, the picture is quite complete and satisfactory, and we will present the essential ideas and sketches of proofs (see also [9]).

To start, we present an example where Condition A is violated and prove directly (and in detail) that no convergent normalizing transformation exists. The example is quite old, going back to Horn (1899) in a somewhat different context, and is cited in Bruno's paper [6]. But it seems that a complete proof of the non-existence of a convergent normalizing transformation is not available in the literature.

Proposition 3.1 *The differential system*

$$\begin{aligned} \dot{x}_1 &= x_1^2 \\ \dot{x}_2 &= x_2 - x_1 \end{aligned} \tag{3.1}$$

does not admit a convergent transformation to normal form.

Proof. (i) An elementary computation shows that a normal form is given by

$$\widehat{f} = \begin{pmatrix} x_1^2 \\ x_2 \end{pmatrix},$$

after an intermediate step to transform f to

$$\begin{pmatrix} x_1^2 \\ x_2 - x_1^2 \end{pmatrix}$$

with diagonal linear part $A = \operatorname{diag}(0,\,1)$. Now assume there is a convergent normalizing transformation

$$\Phi(x) = \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}$$

with invertible $D\Phi(0)$. This implies

$$D\Phi(x)\begin{pmatrix} x_1^2 \\ x_2 \end{pmatrix} = \begin{pmatrix} \varphi_1^2 \\ \varphi_2 - \varphi_1 \end{pmatrix}.$$

(ii) We claim: If $\rho \neq 0$ is a series such that $\rho(0) = 0$ and $X_{\widehat{f}}(\rho) = \rho^2$ then ρ is a series in x_1 alone. (Note that ρ is a solution-preserving map to the one-dimensional equation $\dot{x} = x^2$.) As a consequence, we find $x_1^2\rho'(x_1) = \rho(x_1)^2$, and $\rho = x_1/(1 + cx_1)$ follows easily. Changing the constant c amounts to a coordinate transformation in $\mathbf{K}$, so we may take $c = 0$ and $\rho(x) = x_1$.

To prove the claim, we show $X_A(\rho) = 0$. Since $\widehat{f}$ is in normal form, we have

$$X_{\widehat{f}}X_A(\rho) = X_AX_{\widehat{f}}(\rho) = 2\rho\, X_A(\rho).$$

Now consider the Taylor expansion $\rho = \rho_r + \ldots$, with $\rho_r \neq 0$ homogeneous of degree r (and $r > 0$). Assume that there is a smallest integer $q \geq 0$ such that $X_A(\rho_{r+q}) \neq 0$. Then

$$2\rho\, X_A(\rho) = 2\rho_r\, X_A(\rho_{r+q}) + \ldots,$$

thus the term of smallest degree in this expansion has degree $2r + q$. On the other hand,

$$X_{\widehat{f}}X_A(\rho) = X_A^2(\rho_{r+q}) + \ldots$$

forces $X_A^2(\rho_{r+q}) = 0$ and then $X_A(\rho_{r+q}) = 0$, since A is semisimple. This yields a contradiction.

(iii) Applying (ii) to the series Φ, we may assume $\varphi_1(x) = x_1$. Then φ_2 must satisfy

$$x_1^2\frac{\partial\varphi_2}{\partial x_1} + x_2\frac{\partial\varphi_2}{\partial x_2} = \varphi_2 - x_1,$$

and this forces (see [6])

$$\varphi_2(x_1, 0) = \sum_{k\geq 1}(k-1)!\, x_1^k. \tag{3.2}$$

Therefore Φ is not convergent. $\square$

We now present the central theorem about convergence in the two-dimensional setting. It combines results of Markhashov [19], and Bruno and Walcher [8], and may be seen as a converse to what was stated in the observation following Proposition 2.1.

Theorem 3.2 *Assume that f admits a nontrivial commuting analytic vector field g. Then there is a convergent normalizing transformation for f.*

Sketch of proof. Let λ_1, λ_2 be the eigenvalues of A; to be specific we assume that $\lambda_1 \neq 0$. Let $g = Bx + \ldots$; there is no a priori condition on B.
(i) If λ_1 and λ_2 are in the Poincaré domain (thus, λ_2/λ_1 is not a negative real number) then convergence is unproblematic by Poincaré's theorem.
(ii) Assume that $\lambda_2/\lambda_1 < 0$ and irrational. Then only linear vector fields commute with Ax; in particular the normal form is $\widehat{f} = Ax$. A formal normalizing transformation for f sends g to some formal series $\tilde{g} = Bx + \ldots$, and $[Ax, \tilde{g}] = 0$ forces $\tilde{g} = Bx$ and $B \neq 0$. Since A is diagonalizable with distinct eigenvalues and B commutes with A, the matrices A and B are simultaneously diagonalizable. If B is a multiple of A then f is already in normal form. Otherwise there are scalars σ_1, σ_2 such that $\sigma_1 A + \sigma_2 B = I$. Now $g^* := \sigma_1 f + \sigma_2 g$ is also a nontrivial commuting vector field for f, and we can apply Corollary 2.5.
(iii) Finally, assume that $\lambda_2/\lambda_1 \leq 0$ and rational. The last argument of (ii) shows that only the case $g = \sigma Ax + \ldots$, with a scalar σ, needs consideration. The critical point is to show that $g^* = f + \theta g$ satisfies Condition A for some scalar θ; see [8]. So we may assume that $g = \sigma Ax + \ldots$, with $\sigma \neq 0$ (see [8] for this), satisfies Condition A, and there is a convergent transformation of g to normal form $\widehat{g} = \sigma Ax + \ldots$. The same transformation takes f to some $\tilde{f}$, with $[\widehat{g}, \tilde{f}] = 0$, and Proposition 2.2 shows that $[\sigma Ax, \tilde{f}] = 0$. Hence $\tilde{f}$ is in normal form. $\square$

Example. Among the equations for which the theorem is applicable are the "holomorphic" 2-dimensional dynamical systems, i.e. equations of the form

$$\dot{x} = u(x,y) \qquad\qquad \dot{y} = v(x,y) \tag{3.3}$$

where, putting $z = x + iy$, the function $f(z) = u + iv$ is a holomorphic function of the complex variable z. It can be verified by a simple computation (using the Cauchy–Riemann equations) that such a system admits the non-trivial analytic infinitesimal symmetry

$$X_g = v(x,y)\frac{\partial}{\partial x} - u(x,y)\frac{\partial}{\partial y} \ . \tag{3.4}$$

The existence of a convergent normalizing transformation can also be seen from consideration of the (equivalent) one-dimensional complex equation $\dot{z} = f(z) = \alpha z + \ldots$ In the case $\mathrm{Re}\,\alpha = 0$, these systems are of some physical interest because they describe the "isochronous centers", i.e. planar systems possessing a family of periodic orbits, of the same period, around a stationary point.

There is a number of detailed (and deep) results about the complex analytic classification of germs of analytic vector fields, or rather the associated differential forms; we mention Martinet and Ramis [20]. Normal forms (and thus a formal classification) are but a first step towards analytic classification. Recall the correspondence: To a vector field f one assigns the differential form $f_1\,\mathrm{d}x_2 - f_2\,\mathrm{d}x_1$, but for any locally analytic σ with $\sigma(0) \neq 0$ the vector field $\sigma\, f$ will yield an equivalent differential form, with the same integral curves. Stated from a different perspective, the differential equations $\dot{x} = f(x)$ and $\dot{x} = \sigma(x)\, f(x)$ have the same local solution orbits (albeit with different parameterizations). Thus, the convergence problem on the differential form level amounts to convergence of a normalizing transformation for some $\sigma\, f$. We present a result that also follows from [20], but we supply a different (and elementary) proof. Call a function ρ an integrating factor of f if $\mathrm{div}\,(\rho\, f) = 0$ (whence $\rho\, f$ is locally a Hamiltonian vector field).

Proposition 3.3 *Let the analytic differential equation $\dot{x} = f(x)$ be given in a neighborhood of 0. Assume that $\lambda_2/\lambda_1 = -q/p$, with positive and relatively prime integers, or that $\lambda_1 \neq 0$ and $\lambda_2 = 0$. Then there is an analytic σ, with $\sigma(0) \neq 0$, such that σf admits a convergent transformation into normal form, if and only if there is an integrating factor φ^{-1}, with φ analytic in 0.*

Proof. We may assume that the formal normal form does not satisfy Condition A. (In the case $\hat{f}(x) = (1+\alpha(x))Ax$ one may take $\varphi(x) = (1+\alpha(x))x_1x_2$, so the assertion holds.) First assume that $\lambda_2/\lambda_1 = -q/p$. There are invariant analytic curves tangent to the eigenspaces of A (see, for instance, Bibikov [5], Theorem 3.2, which guarantees convergence to normal form on an invariant manifold), and we may therefore assume that

$$f(x) = \begin{pmatrix} x_1\beta_1(x_1,\, x_2) \\ x_2\beta_2(x_1,\, x_2) \end{pmatrix},$$

with $\beta_2(0)/\beta_1(0) = -q/p$. Let $f^* := \beta_2^{-1} f$; thus we have the orbit-equivalent vector field

$$f^*(x) = \begin{pmatrix} x_1\beta_1^*(x_1,\, x_2) \\ x_2 \end{pmatrix}.$$

According to [27], Theorem 2.3, since Condition A does not hold, there is a unique integrating factor $(x_1^{1+\ell q} x_2^{1+\ell p} \exp(\mu))^{-1}$ (with ℓ determined by the formal normal form), and [28], Prop. 1.1 shows that $g := \left(\begin{smallmatrix} x_1^{1+\ell q} x_2^{\ell p} \exp(\mu) \\ 0 \end{smallmatrix}\right)$ satisfies $[g,\, f^*] = \mu f^*$ for some analytic μ. On the other hand, a direct verification shows

$$[g,\, f^*] = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \begin{pmatrix} * \\ 0 \end{pmatrix} - \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} \begin{pmatrix} * \\ * \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix},$$

and this forces $\mu = 0$. It now follows from Theorem 3.2 that f^* admits a convergent transformation into normal form.

Next consider the case $\lambda_2 = 0$, $\lambda_1 \neq 0$. For the proof in this case (and any case where $\operatorname{div}(A) \neq 0$) one can argue like this: Given an integrating factor, choose g as in [28], Remark 1.2. Then $[g, \frac{1}{\operatorname{div}(f)} f] = 0$ can be directly verified. For the proof of the reverse direction, see [28], Prop. 1.1. $\square$

4 Finite dimensional centralizers.

Again, let $f = Ax + F$ be analytic, and $\widehat{f} = Ax + \widehat{F}$ a (formal) normal form of f. In this section we will discuss the role of the centralizers $\mathcal{C}_{\mathrm{an}}(f)$ (local analytic vector fields commuting with f) and $\mathcal{C}_{\mathrm{for}}(\widehat{f})$ (formal vector fields commuting with $\widehat{f}$). In the present section we always require that $\mathcal{C}_{\mathrm{for}}(\widehat{f})$ is a finite dimensional vector space.

Recall that $Ax \in \mathcal{C}_{\mathrm{for}}(\widehat{f})$, hence the formal centralizer of $\widehat{f}$ is not trivial. Let us first obtain a more precise description of the linear elements in $\mathcal{C}_{\mathrm{for}}(\widehat{f})$: Among the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A, we may assume that $\lambda_1, \ldots, \lambda_d$ are a maximal linearly independent system over the rational number field $\mathbf{Q}$. (This notation will be kept for the remainder of the section.)

Lemma 4.1 *Assume that A is diagonal.*
(a) There are linearly independent diagonal matrices $A_1, \ldots, A_d$ with rational entries such that

$$A = \lambda_1 A_1 + \ldots + \lambda_d A_d \qquad \text{and} \qquad [A_1 x,\, \widehat{f}] = \cdots = [A_d x,\, \widehat{f}] = 0.$$

(b) Let the complex numbers $\sigma_1, \ldots, \sigma_d$ be linearly independent over $\mathbf{Q}$. Then Ax and $(\sigma_1 A_1 + \ldots + \sigma_d A_d)x$ have the same nonlinear formal centralizer elements.

Sketch of proof. According to the hypothesis, there are rational numbers α_{ij} such that

$$\lambda_j = \sum_{i=1}^{d} \alpha_{ij} \lambda_i \qquad (j = d+1, \ldots, n).$$

Therefore one may take

$$A_1 = \operatorname{diag}(1, 0, \ldots, 0, \alpha_{1,d+1}, \ldots, \alpha_{1,n}), \qquad \ldots \qquad ,$$
$$A_d = \operatorname{diag}(0, \ldots, 0, 1, \alpha_{d,d+1}, \ldots, \alpha_{d,n}).$$

Moreover, the eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfy a resonance relation

$$m_1 \lambda_1 + \ldots m_n \lambda_n = \lambda_j$$

if and only if the eigenvalues of $A_1, \ldots, A_d$ satisfy this relation. $\square$

Clearly, the formal centralizer of $\widehat{f}$ contains $\widehat{f}$ itself as well as $A_1 x, \ldots, A_d x$. There may be more linear centralizer elements than the linear combinations of these; for instance if the eigenvalues of A are all rational, and $\widehat{f} = Ax$, then $d = 1$ but the space of linear centralizer elements has dimension $\geq n$. On the other hand, if the eigenvalues of A admit resonances (thus there are normal forms different from Ax), and if $\widehat{f}$ is sufficiently *generic* (to be precise: if sufficiently many of the resonant monomials occur in $\widehat{F}$ with nonzero coefficient) then the subspace of linear elements of $\mathcal{C}_{\text{for}}(\widehat{f})$ is spanned by the $A_i x$, as can be seen from the proof of the lemma.

Question. If the eigenvalues of A admit resonances, is it true for a generic f that $\mathcal{C}_{\text{for}}(\widehat{f})$ is spanned by $\widehat{f}$ and $A_1 x, \ldots, A_d x$?

In a number of cases this is true; see [26] for certain classes of equations, and also the examples later in this section. No counterexample is known, but no general proof seems to be known either. On the other hand, it can be seen that in the Hamiltonian case, the centralizer of the normal form is infinite dimensional. Indeed, let $H = H_0 + H_1$ be a given analytic Hamiltonian, where H_0 is the quadratic part in the canonical variables, and let $\dot{x} = J\nabla_x H = f(x) = Ax + F(x)$ be the associated dynamical system, with standard notations, and with A semisimple. Once in normal form, H_0 is a constant of motion for the Hamiltonian $\widehat{H}$ (see [3, 14, 24]), and therefore it is easy to see (cf. [11]) that any vector field of the form $\varphi(H_0)Ax$ belongs to $\mathcal{C}_{\text{for}}(\widehat{f})$, where φ is any scalar series. Actually, the Hamiltonian structure makes the Hamiltonian case "non-generic"; for a review of some results on the convergence of the normalizing transformations for this case see e.g. [14].

Now let us turn to convergence theorems involving centralizers. The following result combines work by Markhashov [19], Cicogna [10], and Walcher [26].

Theorem 4.2 *Let* $\dim \mathcal{C}_{\text{for}}(\widehat{f}) = k < \infty$*, and assume that there is a matrix* $A^{\#}$*, such that* $A^{\#}x$ *has the same formal centralizer as* Ax*, and satisfies Condition* ω*. If* $\dim \mathcal{C}_{\text{an}}(f) \geq k$ *then there is a convergent transformation of* f *to normal form.*

Sketch of proof. There is a formal transformation Ψ of f to normal form $\widehat{f}$; Ψ sets up a correspondence between the formal centralizers of f and $\widehat{f}$. The analytic centralizer of f is obviously contained in the formal centralizer, and by dimension assumptions this transformation induces a 1-1-correspondence between $\mathcal{C}_{\text{an}}(f)$ and $\mathcal{C}_{\text{for}}(\widehat{f})$ (see [26]). In particular there is an analytic $g = A^{\#}x + \dots$ that is transformed to $A^{\#}x$ via Ψ. Therefore g satisfies the Pliss condition, and there is a convergent transformation Φ sending g to $A^{\#}x$. The transformation Φ sends f to some analytic vector field commuting with $A^{\#}x$, hence commuting with Ax, hence in normal form. $\square$

The introduction of the matrix $A^{\#}$ is clearly useful in the case where A does not satisfy Condition ω.

Remark. If $\mathcal{C}_{\text{for}}(\widehat{f})$ is spanned by linear vector fields and $\widehat{f}$ then the condition of the theorem is also necessary for convergence [26].

Corollary 4.3 *If the eigenvalues of* A *are non-resonant (and pairwise different), then there is a convergent normalizing transformation if and only if* $\dim \mathcal{C}_{\text{an}}(f) = n$*.*

Proof. Choose $A^{\#}$ as a suitable diagonal matrix satisfying Condition A. $\square$

Remark. The results by Markhashov [19] and Cicogna [10] include more specific hypotheses on the centralizer elements. Actually, these conditions follow automatically from the assumptions on the dimension of $\mathcal{C}_{\text{an}}(f)$ and $\mathcal{C}_{\text{for}}(\widehat{f})$ in Theorem 4.2. However, the information on the special form of the centralizer elements is sometimes valuable.

The reader may ask a philosophical question here: What are these results good for? Principally, their value lies in the structural insight they provide: Convergence and existence of symmetries are closely related.

The analytic symmetries of f will generally not be accessible in an algorithmic manner. (Recall that first order ordinary differential equations are a big exception in that regard; see Olver [21].) One may find power series expansions for centralizer elements, but they pose the same convergence problems as in the normalization of f.) Therefore, one will generally have to resort to outside information when it comes to $\mathcal{C}_{\text{an}}(f)$. On the other hand, the formal centralizer of $\widehat{f}$ can be explicitly determined in a number of cases, even from a finite portion of the Taylor series; hence this part of the problem is computationally accessible. Basically, in these cases one can give an affirmative answer to the question we posed above.

Example. (See [26] for details.) Let A be diagonal and $\lambda_1, \ldots, \lambda_n$ be complex numbers with the following property: There are nonnegative integers $s_1, \ldots, s_n$, not all of them zero, such that $s_1\lambda_1 + \ldots s_n\lambda_n = 0$, and whenever $m_1, \ldots, m_n$ are nonnegative integers such that $\sum m_i\lambda_i - \lambda_j = 0$ for some j, $1 \leq j \leq n$, then $(m_1, \ldots, m_j - 1, \ldots, m_n) = k \cdot (s_1, \ldots, s_n)$ for some nonnegative integer k.

Then the elements of $\mathcal{C}_{\text{for}}(Ax)$ are exactly the vector fields which can be written as

$$\sum_{l\geq 0} \rho^l B_l x, \text{ with } \rho(x) := x_1^{s_1} \cdots x_n^{s_n}, \tag{4.1}$$

and all B_l diagonal matrices. Thus let $\widehat{f} = \sum_{l\geq 0} \rho^l C_l x$, and suppose the genericity condition $X_{C_1}(\rho) \neq 0$ holds. The linear vector fields in $\mathcal{C}_{\text{for}}(\widehat{f})$ are then expressed by means of diagonal matrices and admit the first integral ρ. (They obviously form a vector space of dimension $n-1$.)

Now let $g = Bx + \rho^r D_r x + \ldots$, with $D_r \neq 0$, be a nonlinear element of the centralizer of $\widehat{f}$ (and of Ax, according to Proposition 2.2). Then comparing terms of small degree shows

$$0 = [\rho C_1 x, \rho^r D_r x] = \rho \cdot r\rho^{r-1} X_{C_1}(\rho) D_r x - \rho^r \cdot X_{D_r}(\rho) C_1 x.$$

Since C_1 and D_r are diagonal, one has $X_{C_1}(\rho) = \alpha\rho$ for some $\alpha \neq 0$ and $X_{D_r}(\rho) = \beta\rho$ for some β. Substituting this in the above equation yields $r\alpha D_r - \beta C_1 = 0$, whence $\beta \neq 0$ and $X_{D_r} = (\beta/r\alpha) \cdot X_{C_1}$. Now the equality $X_{D_r}(\rho) = (\beta/\alpha)X_{C_1}(\rho) \neq 0$ shows that $r = 1$ and $D_r = (\beta/\alpha)C_1$. This is sufficient to conclude that the formal centralizer of f^* has dimension n.

More examples and a starting point for general investigations (the concept of "rigidity") can be found in [26].

The following result, originally due to Cicogna [11], may be seen as a counterpart to Theorem 4.2.

Theorem 4.4 *Assume that $\mathcal{C}_{\mathrm{for}}(\widehat{f})$ is spanned by $\widehat{f}$ and linear vector fields. If $\mathcal{C}_{\mathrm{an}}(f)$ contains a nontrivial vector field of the form $g = \beta Ax + G$ (with some scalar β; $g \neq \beta f$) then there exists a convergent normalizing transformation for f.*

Sketch of proof: We may assume $g = Ax + G$ (add f, if necessary). A formal normalizing transformation for f takes g to some $\tilde{g} = Ax + \ldots$ in the formal centralizer of $\widehat{f}$. The hypothesis now forces $\tilde{g} = Ax$, and we can proceed as in the proof of Theorem 4.2. $\square$

Example. Consider a 3-dimensional system

$$\dot{x} = f(x) = Ax + F(x) \qquad \text{with} \qquad A = \mathrm{diag}(1, 1, -2) \tag{4.2}$$

and let $f(x)$ possess the linear SO_2 symmetry generated by Lx, where

$$L = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.3}$$

i.e. f is equivariant under rotations in the plane (x_1, x_2). Putting $\rho = x_1^2 + x_2^2$, this implies that $F(x)$ must be of the form

$$F(x) = \varphi_0(\rho, x_3)Ax + \varphi_1(\rho, x_3)Ix + \varphi_2(\rho, x_3)Lx$$

where I is the identity matrix. If we now choose, for instance,

$$\varphi_0 = 0, \qquad \varphi_1 = a_1\rho x_3 + a_2 x_3^3, \qquad \varphi_2 = b\varphi_1$$

where a_1, a_2, b are constants $\neq 0$, then the differential equation (4.2) also admits the non-linear symmetry

$$g(x) = \rho x_3 (I + bL)x. \tag{4.4}$$

The assumption $a_2 \neq 0$ ensures that f is *not* in normal form, and that the above symmetry is not trivial. Now, the normal form of f is of the form

$$\widehat{f} = Ax + a_1 \rho x_3 (x + bLx) + \text{ higher order terms}, \tag{4.5}$$

and one can show in a manner similar to the previous example that the hypothesis of Theorem 4.4 is satisfied. We conclude that f admits a convergent normalizing transformation.

There are several more results in this vein; we refer to [10, 11, 26]. To give an example (and to emphasize once again the role of symmetries in the convergence problem), one may assume the presence of additional nontrivial elements in $\mathcal{C}_{\mathrm{an}}(f)$: this may allow, in certain scenarios (see [11]), to conclude that $\mathcal{C}_{\mathrm{for}}(\widehat{f})$ is spanned by $\widehat{f}$ and linear vector fields, and then to directly apply Theorem 4.4.

5 An application: the "resonant bifurcation".

In this section we will apply normal form methods to show the existence of bifurcating solutions to dynamical systems, depending on real "control" parameters $\eta \in \mathbf{R}^p$. We consider a system of the form

$$\dot{x} = f(x, \eta) \equiv A(\eta)x + F(x, \eta) \tag{5.1}$$

with $f = f(x, \eta)$ analytic in a neighbourhood of $x_0 = 0$ and $\eta_0 = 0$ and with $f(0, \eta) = 0$, and assume that, for the "critical" value $\eta = \eta_0 = 0$ of the parameters, the matrix

$$A_0 = A(0) \tag{5.2}$$

is semisimple and its eigenvalues λ_i satisfy a resonance relation. We then show the existence – under suitable hypotheses – of a general class of bifurcating solutions in correspondence to this resonance. Details and complete proofs can be found in [12].

Theorem 5.1 *Consider the equation* (5.1) *and assume that for the value $\eta_0 = 0$ the eigenvalues λ_i of A_0 are distinct, real or purely imaginary, and satisfy a resonance relation. Assume also that $p = n - 1$, i.e. that there are $n - 1$ real parameters $\eta \equiv (\eta_1, \dots, \eta_{n-1})$, and finally that putting*

$$a_k^{(i)} = \frac{\partial A_{ii}(\eta)}{\partial \eta_k}\Big|_{\eta=0} \qquad (i = 1, \dots, n \; ; \; k = 1, \dots, n-1),$$

the $n \times n$ matrix

$$D := \begin{pmatrix} \lambda_1 & a_1^{(1)} & a_2^{(1)} & \dots & a_{n-1}^{(1)} \\ \lambda_2 & a_1^{(2)} & \dots & & \\ \dots & & & & \\ \lambda_n & a_1^{(n)} & \dots & & a_{n-1}^{(n)} \end{pmatrix} \tag{5.3}$$

is not singular. Then there is, in a neighbourhood of $x_0 = 0,\ \eta_0 = 0,\ t = 0$, a bifurcating solution of the form

$$x_i(t) = (\exp(\beta(\eta) A_0 t)) x_{0i}(\eta) + \text{h.o.t.} \qquad i = 1, \dots, n \tag{5.4}$$

where $\beta(\eta)$ is some function of the η's such that $\beta(\eta) \to 1$ for $\eta \to 0$, and h.o.t. *stands for higher order terms vanishing as $\eta \to 0$.*

Sketch of proof. The main idea is to transform the given system into normal form and to enforce that the normalizing transformation is convergent, using the convergence conditions given by Bruno [7] for normalizing transformations on certain subsets, which extend Theorem 1.2. Once in normal form, the equations can be easily integrated, and the solution (5.4) is obtained coming back to the initial coordinates by means of the inverse (convergent) transformation. Let us illustrate the idea in the case of dimension two: A formal normal form here will be of type

$$\widehat{f}(\eta, x) = A_0 x + \alpha(\eta, x)\, A_0 x + \beta(\eta, x)\, x, \tag{5.5}$$

with formal series α and β with zero constant term. Obviously Condition A is satisfied if and only if $\beta(\eta, x) = 0$. The decisive point is now that the equation $\beta = 0$ defines an analytic manifold (see Bruno [7], Theorem 2 on p. 204). The hypothesis on D ensures that this equation locally defines a function $\eta = \eta(x)$, and using this the assertion follows. □

The standard stationary bifurcation, Hopf bifurcation, and multiple periodic bifurcating solutions as well, are particular cases of the bifurcations obtained in this way. For instance, if $n = 2$ and with imaginary eigenvalues, it is easy to see that the condition on D in the theorem coincides with the familiar "transversality condition" $\mathrm{d}\ \mathrm{Re}\lambda(\eta)/\mathrm{d}\eta|_{\eta=0} \neq 0$ ensuring standard Hopf bifurcation. (In this context one should also mention Bibikov [5], §7, where similar arguments are used to ensure convergence of a certain transformation to "normal form on an invariant manifold".) A nontrivial example in dimension $n > 2$, and corresponding to the case of coupled oscillators with multiple frequencies, is given by the following corollary, which immediately follows from the theorem. For an explicit example, see [12].

Corollary 5.2 *With the same notations as before, let* $n = 4$ *and* $\lambda_1 = -\lambda_2 = i\omega_0$, $\lambda_3 = -\lambda_4 = mi\omega_0$ *(with* $m = 2, 3, \ldots$*): with* $\eta \equiv (\eta_1, \eta_2, \eta_3) \in \mathbf{R}^3$*, let, after complexification of the space,* $A^C(\eta)$ *be conjugate to the matrix* $A(\eta)$ *such that* $A^C(0)$ *is diagonal. Putting* $a_k^{(i)} = \frac{\partial A_{ii}^C(\eta)}{\partial \eta_k}\Big|_{\eta=0}$ $(i = 1, \ldots, 4\ ;\ k = 1, 2, 3)$*, assume that*

$$\det D = \det \begin{pmatrix} a_1^{(1)} & a_2^{(1)} & a_3^{(1)} & 1 \\ a_1^{(2)} & \ldots & & -1 \\ \ldots & \ldots & & m \\ a_1^{(4)} & \ldots & & -m \end{pmatrix} \neq 0. \tag{5.6}$$

Then there is a multiple-periodic bifurcating solution preserving the frequency resonance $1 : m$.

In the case of multiple eigenvalues of the matrix A_0, the situation is a little bit more involved: for instance, the presence in this case of first integrals of $A_0 x$ of the form $\rho = x_i / x_j$, which may enter in the expression of normal forms, prevents the direct application of the arguments used before.

However, the presence of degenerate eigenvalues is typically connected to the existence of some symmetry property of the problem (indeed, in the absence of symmetries, the degeneration is "non-generic", being possibly removed by arbitrarily small perturbations); in this situation, the arguments used above are still applicable, with the same result [12]. An example, given by coupled oscillators with degenerate frequencies and in the presence of a rotation symmetry, is described in [12].

Acknowledgement. During the preparation of this paper the second named author was visiting the University of Cagliari. The hospitality of the Mathematics Department, and in particular of T. Gramchev, is gratefully acknowledged.

References

[1] V.I. Arnold, "Geometrical methods in the theory of ordinary differential equations", Springer-Verlag, Berlin (1982).

[2] V.I. Arnold and Yu.S. Il'yashenko, Ordinary differential equations, in "Dynamical systems I" (D.V. Anosov and V.I. Arnold, Eds.), Encyclopedia of Mathematical Sciences, Vol. 1, Springer-Verlag, Berlin (1988).

[3] V.I. Arnold, V.V. Kozlov and A.I. Neishtadt, Mathematical aspects of classical and celestial mechanics, in "Dynamical Systems III" (V.I. Arnold Ed.), Encyclopedia of Mathematical Sciences, Vol. 3, 2nd edition, Springer-Verlag, Berlin (1993)

[4] D. Bambusi, G. Cicogna, G. Gaeta, G. Marmo, Normal forms, symmetry and linearization of dynamical systems, *J. Phys. A* **31**, 5065 - 5082 (1998).

[5] Yu.N. Bibikov, "Local theory of nonlinear analytic ordinary differential equations", Lecture Notes in Mathematics **702**, Springer-Verlag, Berlin (1979).

[6] A.D. Bruno, Analytical form of differential equations, *Trans. Moscow Math. Soc.* **25**, 131 - 288 (1971).

[7] A.D. Bruno, "Local methods in nonlinear differential equations", Springer-Verlag, Berlin (1989).

[8] A.D. Bruno and S. Walcher, Symmetries and convergence of normalizing transformations, *J. Math. Anal. Appl.* **183**, 571 - 576 (1994).

[9] S.-N. Chow, C. Li and D. Wang, "Normal forms and bifurcations of planar vector fields", Cambridge University Press, Cambridge (1994).

[10] G. Cicogna, On the convergence of normalizing transformations in the presence of symmetries, *J. Math. Anal. Appl.* **199**, 243 - 255 (1996).

[11] G. Cicogna, Convergent normal forms of symmetric dynamical systems, *J. Phys. A* **30**, 6021 - 6028 (1997).

[12] G. Cicogna, Resonant bifurcations, *J. Math. Anal. Appl.* **241**, 157 - 180 (2000).

[13] G. Cicogna and G. Gaeta, Poincaré normal forms and Lie point symmetries, *J. Phys. A* **27**, 461 - 476 (1994).

[14] G. Cicogna and G. Gaeta, "Symmetry and perturbation theory in nonlinear dynamics", Lecture Notes in Physics, **m 57**, Springer-Verlag, Berlin (1999).

[15] G. Cicogna and G. Gaeta, Nonlinear symmetries and normal forms, in "SPT 98 - Symmetry and perturbation theory" (A. Degasperis and G. Gaeta, Eds.), World Scientific, Singapore (1999).

[16] C. Elphick, E. Tirapegui, M.E. Brachet, P. Coullet, and G. Iooss, A simple global characterization for normal forms of singular vector fields, *Physica D* **29**, 95 - 127 (1987).

[17] G. Gaeta and S. Walcher, Manuscript in preparation.

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

[19] L.M. Markhashov, On the reduction of an analytic system of differential equations to the normal form by an analytic transformation, *J. Appl. Math. Mech.* **38**, 788 - 790 (1974).

[20] J. Martinet and J.-P. Ramis, Classification analytique des équations differentielles non lineaires résonnantes du premier ordre, *Ann. Scient. Éc. Norm. Sup.* **16**, 571 - 621 (1983).

[21] P.J. Olver, "Applications of Lie groups to differential equations", Springer-Verlag, New York (1986).

[22] L.V. Ovsiannikov, "Group analysis of differential equations", Academic Press, London (1982).

[23] V.A. Pliss, On the reduction of an analytic system of differential equations to linear form, *Differential Equations* **1**, 153 - 161 (1965).

[24] J-C. van der Meer, "The Hamiltonian Hopf bifurcation", Lecture Notes in Mathematics **1160**, Springer-Verlag, Berlin (1985).

[25] S. Walcher, On differential equations in normal form, *Math. Ann.* **291**, 293 - 314 (1991).

[26] S. Walcher, On convergent normal form transformations in presence of symmetries, *J. Math. Anal. Appl.* **244**, 17 - 26 (2000).

[27] S. Walcher, On the Poincaré problem, *J. Differential Eqs.* **166**, 51 - 78 (2000).

[28] S. Walcher, Plane polynomial vector fields with prescribed invariant curves, *Proc. Royal Soc. Edinburgh* **130 A**, 633 - 649 (2000).