Reduction of Normal Forms

Giuseppe Gaeta*
I.H.E.S., 91440 Bures sur Yvette (France)
(13/1/97)

We give an algorithm to simplify Poincaré normal forms, making use of higher order effects; this leads to a simplification in the local study of nonlinear systems. The procedure is easily extended to the hamiltonian case and Birkhoff normal forms.

PACS numbers: 02.30.Hq , 03.20.+i , 46.10.+z

I. INTRODUCTION AND MOTIVATION

Poincaré Normal Forms [1–4] are among the most powerful tools to study the local behaviour of Nonlinear Systems around a singular point (the behaviour around a nonsingular point is trivial, as follows from the rectification theorem [5]).

In general, given a nonlinear system of differential equations

$$\dot{\mathbf{x}} = f(\mathbf{x}) \ , \tag{1}$$

where we suppose $f : \mathbf{R}^n \to \mathbf{R}^n$ is a formal power series (e.g. the Taylor series corresponding to expansion around $\mathbf{x}_0 = 0$) such that $f(0) = 0$, we can transform this, by a sequence of formal changes of coordinates, collectively called the normalizing transformation, into a system which is in Normal Form (NF) up to any desired order [in the following, we will mean by NF the "infinite order" NF, and denote as "partial NF (of order m)" the NF of order m.] N, and formally for $N = \infty$. Thus, the study of the local behaviour of ODEs around a singular point can be reduced to the study of the local behaviour of ODEs which are in NF.

Two points should be stressed: *(i)* The equivalence between the original system and its NF is, in general, only formal (due to lack of convergence of the normalizing transformation; this is in turn due to small denominators): thus, in general a system is *not* conjugated to its NF; however, for N sufficiently small (depending on the system), the system *is* conjugated to its partial NF of order N; *(ii)* For a given system, the reduction to NF – although obtained by means of a well defined algorithm – becomes computationally very demanding with the increasing of N: the computations can be set in terms of linear algebra, but they require to consider – and invert – matrices of order rapidly growing with N.

Thus, in practice, when we analyze a *given* system by means of NF techniques, we consider the partial NF of some order m, study the truncation of this at order m – which is by construction in NF – and then resort to other kinds of considerations, typically m-th order averaging, to ensure that the trajectories of the truncated and the full system are near enough (say with a distance less than ε) for long times (typically for $t < 1/\varepsilon^m$) [1,6].

On the other side, if we want to consider the most general behaviour of systems with given linear part, we can operate a first reduction to (Jordan) normal form for its linear part, i.e. for the matrix $A = (Df)(0)$, and then study the most general NF compatible with this linear part (the determination of this is also referred to as the problem of NF expansion – or in mathematical literature, *unfolding* of the NF – for A). For short enough times, the study of the NF expansion gives the most general local behaviour (within suitable error) of systems with given linear part.

It is well known that the NF corresponding to a given system is in general *not unique*; correspondingly, the complete expansion (unfolding) of the NF for A is exhaustive, but in general not minimal: i.e., the Poincaré-Dulac theory provides a list of NFs such that any system can be (formally) reduced to one of the NFs of the list, but we are not guaranteed that the NFs are pairwise not conjugated.

It is thus of interest to provide a reduced classification of NFs, i.e. a list of NFs for A in which the redundancies in the Poincaré-Dulac classification, or at least some of them, have been eliminated: this automatically simplifies the local analysis of nonlinear systems.

This problem has been considered by several authors [7,8] in a mathematically sophisticated language (based on filtration of Lie algebras); several theoretical result exist, and it is actually also possible to define a *unique NF* [8]. However, all these results are of difficult concrete implementation.

The purpose of this note is indeed to propose a procedure of "further reduction" of NFs (which I call "renormalization") which does not lead to a unique reduced NF, but yields a significant simplification and – I would like to stress – is completely algorithmic, and not more difficult to implement in practice than the standard Poincaré-Dulac normalization.

Actually, the procedure I propose below is nothing else than a direct generalization of the one proposed by Poincaré; the main "new" ingredient will be to make use of the well known freedom in the choice of the generating functions for the coordinate transformations, and a control of higher-order effects in this transformation.

The theory esposed here does also apply to the hamiltonian case, i.e. to Birkhoff normal forms [4,6,9], as briefly shown below.

II. CHANGE OF COORDINATES FORMULAS

We consider changes of coordinates $x \to \widetilde{x}$ defined by the time-one action of a smooth vector field

$$X_h = h^i(\mathbf{x})\ \partial/\partial x^i \equiv h^i \partial_i \ , \tag{2}$$

also known as *Lie transform* [10,11]. Thus, we have formally

$$\widetilde{\mathbf{x}} = \left[e^{\lambda X_h} \mathbf{x} \right]_{\lambda=1} \ ; \tag{3}$$

in this way, $X_0 = f^i(\mathbf{x})\partial_i$ is transformed into $\widetilde{X} = e^{X_h} X_0 e^{-X_h}$; this can be computed by means of the Baker-Campbell-Haussdorf formula [11,12], and its effect is that $f(\mathbf{x})$ changes into

$$\widetilde{f}(\mathbf{x}) = \sum_{n=0}^{\infty} \left[((-1)^n/n!)\, \phi_n(\mathbf{x}) \right] \ , \tag{4}$$

where $\phi_0 = f$ and $\phi_{n+1} = \{\phi_n, h\}$, having defined the bracket as $\{\phi, \psi\} = (\phi \cdot \nabla)\psi - (\psi \cdot \nabla)\phi$.

Let us denote by V_k $(k \ge 0)$ the vector functions whose components are homogeneous polynomials of order $(k+1)$ in the x^i; these will be seen as orthogonal subspaces of a space V (direct sum of the V_k's) and equipped with a scalar product [1,13] for the sake of later considerations. With $f_m \in V_m$, we rewrite (1) as

$$\dot{\mathbf{x}} = \sum_{m=0}^{\infty} f_m(\mathbf{x}) \ . \tag{5}$$

If we choose the h generating the change of coordinates to be $h_k \in V_k$, we have that the f_m with $m < k$ are unchanged, while for higher m we have (with square bracket denoting integer part)

$$\widetilde{f}_m = \sum_{s=0}^{[m/k]} (1/s!) \mathcal{H}^s \left(f_{m-sk} \right) \ , \tag{6}$$

where we have defined $\mathcal{H}(.) = \{h_k, .\}$. Notice that, as well known, $\widetilde{f}_k = f_k - \{f_0, h_k\}$.

III. NORMAL FORMS AND FURTHER REDUCTION

In the usual Poincaré normalization, we proceed by considering a sequence of changes of coordinates with generators h_k with $k = 1, 2, ...$; these h_k are chosen by solving the homological equations

$$\{f_0, h_k\} = \pi_0 \cdot \widehat{f}_k \ , \tag{7}$$

where π_0 is the projection on the range of the homological operator $\mathcal{L}_0 = \{f_0, .\}$ (once we have chosen a scalar product in the V_k, which we suppose to have done, this is a projection on the kernel of $\mathcal{L}_0^+$), and $\widehat{f}_k$ is the term in V_k after considering the transformation generated by $h_1, ..., h_{k-1}$. In this way we can eliminate any term in $\mathrm{Ran}(\mathcal{L}_0)$; we also say that we remain with *resonant* terms [2–5]. Notice that the solution h_k to (7) is only determined up to a term $(\delta h_k) \in \mathrm{Ker}(\mathcal{L}_0)$.

Notice however that, due to higher order effects [i.e. $s > 1$ in (6)], in this process we can – and in general will – generate terms in $[\mathrm{Ran}(\mathcal{L}_0)]^\perp$ which were not present in the original system; different choices of (δh_k) will give different higher order terms. By the same mechanism of "higher order effects", it would also be possible to eliminate (some of the) resonant terms initially present: this is unlikely to happen by random choice, but can be obtained by the procedure we are going to describe.

Suppose to have already taken into NF the terms f_1 and f_2, so that $\mathcal{L}_0^+(\widetilde{f}_1) = \mathcal{L}_0^+(\widetilde{f}_2) = 0$. We consider now a change of coordinates with generator $h_1^{(1)}$ such that $\widetilde{f}_1$ is not changed, which is the case if $h_1^{(1)} \in \mathrm{Ker}(\mathcal{L}_0)$; according to (7), $\widetilde{f}_2 \equiv f_2^{(1)}$ (we also write $\widetilde{f}_1 \equiv f_1^{(1)}$) is then changed into

$$f_2^{(2)} = f_2^{(1)} - \{\widetilde{f}_1, h_1^{(1)}\} \ ; \tag{8}$$

to conform with the general notation introduced below, we define the "higher homological operator" $\mathcal{L}_1 = \{f_1^{(1)}, .\}$, and $\mathcal{M}_1$ as the restriction of $\mathcal{L}_1$ to $\mathrm{Ker}(\mathcal{L}_0)$, so that (8) reads

$$f_2^{(2)} = f_2^{(1)} - \mathcal{M}_1 \left(h_1^{(1)} \right) \ . \tag{9}$$

In this way, we can eliminate from f_2 not only the terms in $\mathrm{Ran}(\mathcal{L}_0)$, but also those in $\mathrm{Ran}(\mathcal{M}_1)$. With the choice of a scalar product, we obtain a term $f_2^{(2)} \in \mathrm{Ker}(\mathcal{L}_0^+) \cap \mathrm{Ker}(\mathcal{M}_1^+)$. Notice that $h_1^{(1)}$ is determined by an "higher order homological equation"

$$\mathcal{M}_1 \left(h_1^{(1)} \right) = \pi_1 \cdot f_2^{(1)} \ , \tag{10}$$

where π_1 is the projection on the range of $\mathcal{M}_1$.

This construction can then be generalized to higher order terms f_p as follows.

For a given sequence $f_1, f_2, ...$ with $f_k \in V_k$, we define $L_p = \{f_p, .\}$; we define then the spaces $H^{(p)} \subseteq V$ by $H^{(0)} = V$ and, for $p \ge 1$,

$$H^{(p)} = H^{(p-1)} \cap \mathrm{Ker}(\mathcal{L}_{p-1}) \ ; \tag{11}$$

obviously $H^{(p+1)} \subseteq H^{(p)}$. Next, we define the $\mathcal{M}_k$ as the restriction of $\mathcal{L}_k$ to $H^{(p-1)}$, and π_k as the projection on $\mathrm{Ran}(\mathcal{M}_k)$. Finally, we define the spaces $F^{(p)}$ by $F^{(0)} = V$ and

$$F^{(p)} = F^{(p-1)} \cap \mathrm{Ker}\left(\mathcal{M}_p^+ \right) \ ; \tag{12}$$

hence these satisfy $F^{(p+1)} \subseteq F^{(p)}$.

With this notation, we can repeat the construction considered for $f_2^{(2)}$ at all orders: suppose to have already operated for $f_1, f_2, ...f_p$, so that for $k \le p$ we have $f_k^{(k)} \in F^{(k)} \cap V_k \equiv F_k^{(k)}$; we can then consider changes of coordinates with generators $h_{p-t}^{(t)} \in H^{(t)} \cap V_{p-t} \equiv H_{p-t}^{(t)}$ (with $t = 1, ..., p$): these will not affect the terms $f_1^{(1)}, ..., f_{p-1}^{(p-1)}$, but permit to eliminate terms in f_{p+1}, reducing this via

$$f_p^{t+1} = f_p^t - \mathcal{M}_t \left(h_{p-t}^{(t)} \right) \tag{13}$$

to a $f_p^{(p)} \in F_p^{(p)}$.

We say that the formal power series $f(\mathbf{x}) = \sum_{k=0}^\infty f_k$ (where $f_k \in V_k$) is in **Poincaré renormalized form** (PRF) up to order N if $f_k \in F_k^{(k)}$ for all $k \le N$; if this is satisfied for all k, we just say that f is in Poincaré renormalized form [15].

The above discussion shows (constructively) that: *any formal power series can be formally transformed into Poincaré renormalized form up to any given order by a formal series of changes of coordinates.* This transformation is concretely achieved by the algorithm described above. By analogy with the Poincaré normalizing transformation, the sequence of changes of coordinates considered here will be called the "Poincaré renormalizing transformation".

Notice that points *(i),(ii)* mentioned in the introduction still apply, and in practice one will consider partial PRFs up to a suitable order m, as for standard NFs.

Thus, we can consider renormalized forms instead than standard Poincaré normal forms. The advantage of doing so is that *(a)* the renormalized forms can present a substantial simplification, as the following examples will show; *(b)* the computation of PRF is not mor difficult than that of standard NF (going up to the same order), as it also amounts to solving linear algebraic equations (of the same order as those to be solved for NF at the same order). We also notice that, in considering convergence problem, the Bruno-Markhashov-Walcher theory [14] allows to affirm convergence based on symmetry considerations, making use of a system possessing higher symmetry than, but formally equivalent to, the original one: the PRFs are natural candidates for such a system.

The present theory can also be easily extended to Birkhoff normal forms, i.e. to the hamiltonian case [4,6,9], going essentially through the same construction: now we would consider a Hamiltonian H, V_k would be the space of functions from the phase space to $\mathbf{R}$ homogeneous of degree $(k+2)$, so that $H = \sum_k H_k$, $\{.,.\}$ would be the usual Poisson bracket, the role of A would be played by the quadratic part H_0 of the Hamiltonian, and $\mathcal{L}_k = \{H_k, .\}$; with this notation, the discussion would be formally identical to the one presented above (see [15] for details).

IV. EXAMPLES

We consider here some examples, comparing the expansion of PRF and of standard NFs for some systems with given linear part $f_0(\mathbf{x}) = A\mathbf{x}$; we only report results, referring for details on these – and in general for a complete discussion of PRFs – to [15].

Example 1. Consider $\mathbf{R}^2$ and $A = \{\{0,-1\},\{1,0\}\}$ the standard symplectic matrix. The standard NF expansion corresponding to this is

$$f(\mathbf{x}) = A\mathbf{x} + \sum_{k=1}^\infty |\mathbf{x}|^{2k} \left[a_k I + b_k A \right] \mathbf{x} \tag{14}$$

with a_k, b_k two arbitrary sequences of real numbers. The expansion of PRF is, with μ, ν suitable positive integers (which can be explicitely determined [15]),

$$f(\mathbf{x}) = A\mathbf{x} + \left[|\mathbf{x}|^{2k\mu} c_1 I + |\mathbf{x}|^{4\mu} c_2 I + |\mathbf{x}|^{2\nu} c_3 A \right] \mathbf{x} \ , \tag{15}$$

and thus depends only on three arbitrary real constants. When (14) is hamiltonian (i.e. $a_k = 0$ for all k), this corresponds to a well known result [16].

Example 2. Consider $\mathbf{R}^2$ with coordinates (x,y), and A diagonal with entries $\{1,-n\}$ (with $n > 1$, as $n = 1$ is equivalent to example 1 and $n \le 0$ is trivial). The standard NF expansion is (sums are for $k = 1, ..., \infty$)

$$f(\mathbf{x}) = \begin{pmatrix} x + \sum \sigma_k (x^n y)^k x \\ -ny + \sum \rho_k (x^n y)^k y \end{pmatrix} \ ; \tag{16}$$

again this depends on two arbitrary sequences of real numbers. The expansion of PRF is, with p, q suitable positive integers (which can be explicitely determined [15]),

$$f(\mathbf{x}) = \begin{pmatrix} x + n \left[c_1 (x^n y)^q + c_2 (x^n y)^p + c_3 (x^n y)^{2p} \right] x \\ -ny + \left[c_1 (x^n y)^q + c_2 (x^n y)^p + c_3 (x^n y)^{2p} \right] y \end{pmatrix} \tag{17}$$

and thus it depends again on three real numbers.

Example 3. Consider $\mathbf{R}^3$ and A diagonal with entries $\{1,2,5\}$; the standard NF expansion depends on four real parameters, while the PRF expansion depends on three real parameters, of which at least one can be chosen to be zero.

Example 4. Consider $\mathbf{R}^3$ and A diagonal with entries $\{1,3,9\}$; the standard NF expansion depends on five real parameters, while for the PRF expansion there are seven possibilities, each one depending at most on two real parameters.

ACKNOWLEDGEMENTS

This note was written during my stay in IHES; I would like to express my hearty thanks to its Director, prof. J.P. Bourguignon, for his invitation, and to all the staff for the warm hospitality.

* On leave from Department of Mathematical Sciences, Loughborough University, Loughborough LE11 3TU (Great Britain); `G.Gaeta@lboro.ac.uk`

REFERENCES

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

[2] V.I. Arnold and Yu.S. Il'yashenko, "Ordinary differential equations"; in *Encyclopaedia of Mathematical Sciences - vol. 1, 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, 1996^2

[5] V.I. Arnold, "Ordinary differential equations", Springer, Berlin 1977, 1992^2

[6] V.I. Arnold, V.V. Kozlov and A.I. Neishtadt, "Mathematical aspects of classical and celestial mechanics"; in *Encyclopaedia of Mathematical Sciences - vol. 3, Dynamical Systems III*, (V.I. Arnold ed.), p. 1-291, Springer, Berlin 1988, 1993^2

[7] M. Kummer, *Nuovo Cimento B*, **1** (1971), 123; *Comm. Math. Phys.* **48** (1976), 53; J.C. van der Meer, "The Hamiltonian Hopf bifurcation", (*Lect. Notes Math.* **1160**), Springer, Berlin 1985

[8] A. Baider and R.C. Churchill, *Proc. R. Soc. Edinburgh A* **108** (1988), 27; A. Baider, *J. Diff. Eqs.* **78** (1989), 33

[9] V.I. Arnold, "Mathematical methods of classical mechanics", Springer, Berlin 1978, 1989^2

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

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

[12] L.D. Landau and I.M. Lifshitz, "Quantum Mechanics", Pergamon, London, 1959

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

[14] L.M. Markhashov, *J. Appl. Math. Mech.* **38** (1974), 788; A.D. Bruno, *Sel. Math.* **12** (1993), 13; A.D. Bruno and S. Walcher, *J. Math. Anal. Appl.* **183** (1994), 571; G. Cicogna, *J. Math. Anal. Appl.* **199** (1996), 243

[15] G. Gaeta, preprint IHES/P/96/74 (available through `http://www.ma.utexas.edu/mp_arc` as preprint `96-584`)

[16] C.L. Siegel and J.K. Moser, "Lectures on celestial mechanics", Springer, Berlin 1955 (reprinted in *Classics in Mathematics*, 1995). See also E. Forrest and D. Murray, *Physica D* **74** (1994), 181