On the computation of unique normal forms and quadratic convergence

Jan A. Sanders
Department of Mathematics, Faculty of Sciences
Vrije Universiteit, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
E-mail: jan.sanders.a@gmail.com

April 27, 2011

Abstract

An algorithm is given which computes the transformation that transforms the initial system to final normalized system to a given order, while reducing the order to which intermediate results need to be computed. The algorithm will convergence quadratically when one is only interested in obtaining the classical normal form and somewhere in between linearly and quadratically in the general case.

1 Introduction

In [San05] and [SVM07, section 13.8] it is noticed that there are possible obstructions to the quadratic convergence of normal form calculations, at least when one is interested in the unique normal form, as opposed to the classical normal form where one normalizes only with respect to the lowest order part of the object one is trying to normalize (typically the linear part of a differential equation at equilibrium). In the following discussion a scheme is formulated to make the best possible use of the ideas used when everything is working fine, but it is clear that the results are rather unpredictible.

In this paper an algorithm is given which computes the transformation that transforms the initial system to final normalized system to a given order, while reducing the order to which intermediate results need to be computed. The algorithm will convergence quadratically when one is only interested in obtaining the classical normal form and somewhere in between linearly and quadratically in the general case, with the standing assumption that **exponentiation is the computationally most expensive part of the algorithm**.

The algorithm is formulated in Theorem 3 in terms of the bookkeeping device of spectral sequences. These will be defined, and are done. but the reader unfamiliar with the concept should probably first read a selection of [Arn75, San03, San02, San05, Mur05, Mur04, BC06], since the introduction here will be rather terse. The notational conventions vary among authors. The convention among topologists is to keep the degree of the boundary operator equal to zero. This leads to rather skew diagrams and so there are now several other formats which claim to be closer to normal form thinking.

One usually considers in normal form theory a filtered Lie algebra and then one studies the action of a part of the is Lie algebra on a fixed element, both on the Lie algebraic as on the formal group level. In this paper we take the slightly more general view of studying the filtered action of a filtered Lie algebra on a filtered module. One motivation for pulling apart the space of formal transformation generators and the element one acts on is that this is natural when the element is equivariant under a finite group action on a Lie algebra (that is, $g[x, y] = [gx, gy]$ with x, y in the Lie algebra and g in the group). This happens for time-reversible systems: the system changes sign under the action of the reversor (which is then compensated for by reversing time). The transformation generators should be invariant under the reversor in order for the transformed element to be still time reversible: with $gx = x$ and $gy = -y$, $g[x, y] = [gx, gy] = -[x, y]$, so that the action of the symmetric element x maps the time-reversible element y to the time-reversible element $[x, y]$, see [vdMSV94].

Another motivational example might be the use of general, that is non-symplectic, coordinate changes on a Hamiltonian vectorfield, or the normalization of functions by coordinate transformations as in [Arn75].

In section 2 we show how a spectral sequence is constructed, starting with a filtered Lie algebra. In section 3 we formulate an algorithm to compute the unique normal form up to any given order. In section 4 we write out the details of this algorithm for the calculation up till order 8.

2 Unique normal form theory; spectral sequence

In this section a spectral sequence is constructed and all the details are given. There are no new results here. We start with a filtered Lie algebra $\mathfrak{f}^{0,0} \supset \mathfrak{f}^{1,0} \supset \cdots$, where the first superscript indicates the **order** and the second is a labeling that is consistent with the later cohomology computation. By definition one has: $[\mathfrak{f}^{i,0}, \mathfrak{f}^{j,0}] \subset \mathfrak{f}^{i+j,0}$. One can think of Taylor expansions of vector fields at equilibrium; then the order is the lowest degree terms in the expansion minus one. Next we consider a filtered $\mathfrak{f}^{0,0}$-module $\mathfrak{f}^{0,1}$, that is to say, there is a action ρ of $\mathfrak{f}^{0,0}$ on $\mathfrak{f}^{0,1}$ such that $\rho(\mathfrak{f}^{i,0})\mathfrak{f}^{j,1} \subset \mathfrak{f}^{i+j,1}$ and $\rho([\delta, \zeta]) = \rho(\delta)\rho(\zeta) - \rho(\zeta)\rho(\delta)$, $\delta, \zeta \in \mathfrak{f}^{0,0}$. Let $x \in \mathfrak{f}^{0,1}$. We act on x with formal transformations $\exp(\delta^1)$, where $\delta^1 \in \mathfrak{f}^{1,0}$ and $\exp(\delta^1)x = \sum_{k=0}^{\infty} \frac{1}{k!}\rho^k(\delta^1)x$. Notice that we can extend the definition of exp to $\mathfrak{f}^{0,0}$ by replacing ρ with the adjoint representation ad . With this exponential action we view $\mathfrak{f}^{1,0}$ as a formal group, with the (noncommutative) addition given by the Dynkin formula (see Lemma 9), acting on $\mathfrak{f}^{0,\cdot}$. In the sequel we define spaces $\mathfrak{f}_r^{p,0}, \mathfrak{f}^{p,1}, \mathfrak{g}_r^{p,i}, i = 0, 1$, where $\mathfrak{f}_0^{p,0} = \mathfrak{f}^{p,0}$.

Notation 1. *Elements of* $\mathfrak{f}_r^{p,0}$ *will we written as* δ_r^p*, elements of* $\mathfrak{f}^{p,1}$ *as* x^p*.*

Definition 1. *We say, for* $x, y \in \mathfrak{f}^{0,1}$ *that* $x \stackrel{r}{\equiv} y$*, if there exists a* $\delta^1 \in \mathfrak{f}^{1,0}$ *such that*

$$y - \exp(\delta^1)x \in \mathfrak{f}^{r+1,1}.$$

We call δ^1 *the* **generator** *of the formal transformation carrying* x *to* y*.*

We now define recursively the concepts of allowable transformations and normal forms at any given level. The abstract formulation creates some problems, for instance the existence of (compatible) sections is not obvious in general. So we restrict the validity of the definition and the later results to those cases where the existence problem can be solved. In the usual normal form practice this poses no restrictions.

In the following definition the spaces $\mathfrak{g}_r^{p,0}$ measure the the availability of transformation generators of order p that can be used to remove terms of order $p+r$; the spaces $\mathfrak{g}_r^{p,1}$ measure the obstruction to remove the terms of order p by transformations. That everything is well-defined in this definition may not be immediately clear; this is shown in the following lemmas.

The main motivation for this eleborate scheme is the simple proof of Lemma 8, where the cohomological equation has to be solved as part of the normal form computation.

Definition 2. *Fix $x \in \mathfrak{f}^{0,1}$, the object to be put in normal form. Let, at level zero,*

- $\mathfrak{g}_0^{p,i} = \mathfrak{f}^{p,i}/\mathfrak{f}^{p+1,i}, \quad p \in \mathbb{N}, \quad i = 0, 1,$
- $\mathfrak{f}_0^{p,0} = \mathfrak{f}^{p,0}$,
- $\pi_0^{p,i} : \mathfrak{f}_0^{p,i} \to \mathfrak{g}_0^{p,i}, \quad i = 0, 1,$ *is the natural projection,*
- $\sigma_0^{p,i} : \mathfrak{g}_0^{p,i} \to \mathfrak{f}_0^{p,i}, \quad i = 0, 1,$ *is a section, that is* $\pi_0^{p,i}\sigma_0^{p,i} = id_{\mathfrak{g}_0^{p,i}}$,
- *Define* $\mathcal{N}_0^{0,i} = \sigma_0^{0,i}\pi_0^{0,i}$, *a projection operator acting on* $\mathfrak{f}^{0,i}$ *(In the vector field model,* $\mathcal{N}_0^{0,1}$ *can be defined as: take the linear part of the vector field.),*
- $x_0 = \mathcal{N}_0^{0,1}x$. *Notice that* $x \overset{0}{\equiv} x_0$, *since* $\pi_0^{0,1}(x - x_0) = \pi_0^{0,1}x - \pi_0^{0,1}\sigma_0^{0,1}\pi_0^{0,1}x = 0$ *and this implies* $x - x_0 \in \mathfrak{f}^{1,1}$.

For level $0 < r \in \mathbb{N}$ *we recursively define* $\mathfrak{f}_r^{p,0}$ *and* $\mathfrak{g}_r^{p,i}, i = 0, 1,$ x_r *and* $d^{r-1} : \mathfrak{f}_{r-1}^{p,0} \to \mathfrak{f}^{p+r-1,1}$*:*

- $d^{r-1}\delta_{r-1}^p = \rho(\delta_{r-1}^p)x_{r-1}$, $\delta_{r-1}^p \in \mathfrak{f}_{r-1}^{p,0}$,
- $\mathfrak{f}_r^{p,0} = \{\delta_{r-1}^p \in \mathfrak{f}_{r-1}^{p,0} | d^{r-1}\delta_{r-1}^p \in \mathfrak{f}^{p+r,1}\}, \quad \mathfrak{f}_r^{p,1} = \mathfrak{f}_{r-1}^{p,1}$,
- $\mathfrak{g}_r^{p,0} = \mathfrak{f}_r^{p,0}/\mathfrak{f}_{r-1}^{p+1,0}, \quad \mathfrak{g}_r^{p,1} = \mathfrak{f}^{p,1}/(d^{r-1}\mathfrak{f}_{r-1}^{p-r+1,0} + \mathfrak{f}^{p+1,1})$,
- $\pi_r^{p,i} : \mathfrak{f}_r^{p,i} \to \mathfrak{g}_r^{p,i}$ *is the natural projection,*
- $\sigma_r^{p,i} : \mathfrak{g}_r^{p,i} \to \mathfrak{f}_r^{p,i}$ *is a section, that is* $\pi_r^{p,i}\sigma_r^{p,i} = id_{\mathfrak{g}_r^{p,i}}$, *defining the normal form style, and subject to the* **compatibility condition** $\sigma_q^{p,i}\pi_q^{p,i}\sigma_r^{p,i} = \sigma_r^p$ *for* $q < r$,
- *Define* $\mathcal{N}_r^{p,i} = \sigma_r^{p,i}\pi_r^{p,i}$, *projection operators acting on* $\mathfrak{f}_r^{p,i}$, *with kernels* $\mathfrak{f}_{r-1}^{p+1,0}$ *and* $d^{r-1}\mathfrak{f}_{r-1}^{1,0} + \mathfrak{f}^{r+1,1}$ *for* $i = 0, 1$, *respectively.*
- *If* $x \overset{r-1}{\equiv} x_{r-1}$, *by* $\exp(\delta^1)x = x_{r-1} + x^r$, *define* $x_r = x_{r-1} + \mathcal{N}_r^{r,1}x^r$. *We say that* x_r *is the* r**th level normal form** *of* x *with style* $\sigma_0^{0,1}, \ldots, \sigma_r^{r,1}$.

Lemma 1. *Let the maps* $\mathcal{F}_{rq}^{p,i} : \mathfrak{g}_q^{p,i} \to \mathfrak{g}_r^{p,i}$ *be defined by* $\mathcal{F}_{rq}^{p,i} = \pi_r^{p,i}\sigma_q^{p,i}$. *Then* $\mathcal{F}_{rq}^{p,i}$ *is surjective if* $r \geq q$.

Proof. Indeed, if $w \in \mathfrak{g}_r^{p,i}$, then let $v = \mathcal{F}_{rq}^{p,i}w$ and compute $\pi_r^{p,i}\sigma_q^{p,i}v = \pi_r^{p,i}\sigma_q^{p,i}\pi_q^{p,i}\sigma_r^{p,i}w = \pi_r^p\sigma_r^{p,i}w = w$, using the compatibility condition. □

Corollary 1. *If* $\mathfrak{g}_r^{p,i}$ *is a projective module, it can be mapped into* $\mathfrak{g}_q^{p,i}$ *with a section to* $\mathcal{F}_{rq}^{p,i}$, *that is, a higher level normal form can be seen as a lower level normal form.*

Theorem 1. *Define the* r*th level* **cohomological equation** *at order* $p \geq r + 1$ *by*

$$x^p - \mathcal{N}_{r+1}^{p,1}x^p = d^r\delta_{p+1:p}^{p-r} \mod \mathfrak{f}^{p+1,1}.$$

This equation can be solved with $\delta_{p+1:p}^{p-r} \in \bigoplus_{q=0}^r \mathcal{N}_{q+1}^{p-q,0}\mathfrak{f}_{q+1}^{p-q,0}$.

Proof. Since the left hand side is in the kernel of $\mathcal{N}_{r+1}^{p,1}$, which equals the kernel of $\pi_{r+1}^{p,1} = d^r \mathfrak{f}_r^{p-r,0} + \mathfrak{f}^{p+1,1}$, this equation can always be solved with $\delta_{p+1:p}^{p-r} \in \mathfrak{f}_r^{p-r,0}$. Write $\delta_{p+1:p}^{p-r} = \mathcal{N}_{r+1}^{p-r,0}\delta_{p+1:p}^{p-r} + (1-\mathcal{N}_{r+1}^{p-r,0})\delta_{p+1:p}^{p-r}$. Consider the term $\delta_r^{p-r+1} = (1-\mathcal{N}_{r+1}^{p-r,0})\delta_{p+1:p}^{p-r} \in \ker \mathcal{N}_{r+1}^{p-r,0} = \mathfrak{f}_r^{p-r+1,0}$. Write $\delta_r^{p-r+1} = \mathcal{N}_r^{p-r+1,0}\delta_r^{p-r+1} + (1-\mathcal{N}_r^{p-r+1,0})\delta_r^{p-r+1}$. Etcetera. This shows that $\delta_{r:r-1}^{p-r} \in \bigoplus_{q=0}^{r} \mathcal{N}_{q+1}^{p-q,0}\mathfrak{f}_{q+1}^{p-q,0}$. □

Definition 3. *Suppose $\mathcal{F}_{rq}^{p,i}$ and $\mathcal{G}_{rq}^{p,i}$ are the result of different choice of style. Then we say that they are equivalent if there exist $\mathcal{H}_s^{p,i} \in \mathrm{GL}(\mathfrak{g}_s^{p,i})$ such that $\mathcal{H}_r^{p,i}\mathcal{F}_{rq}^{p,i} = \mathcal{G}_{rq}^{p,i}\mathcal{H}_q^{p,i}$.*

Question 1. *Do there exist non-equivalent styles? (If the answer is yes, then this would be extremely interesting, since it would show the existence of another invariant in normal form theory besides the spectral sequence.)*

Conjecture 1. *Let $\mathfrak{h}_r^{p,i} = \mathfrak{g}_r^{p,i}/\mathcal{F}_{r,r+1}^{p,i}\mathfrak{g}_{r+1}^{p,i}$. Then the $\mathfrak{h}_r^{p,i}$ measure the change from increasing the level from r to $r+1$. If they are finite dimensional, their dimension is an invariant of the fixed x. If we let $\mathcal{G}_{rq}^{p,i}$ be an $\mathcal{H}$-equivalent style, then $\mathfrak{g}_r^{p,i}/\mathcal{F}_{r,r+1}^{p,i}\mathfrak{g}_{r+1}^{p,i} = \mathfrak{h}_r^{p,i} \equiv \mathcal{H}_r^i\mathfrak{h}_r^{p,i} = \mathcal{H}_r^i\mathfrak{g}_r^{p,i}/\mathcal{H}_r^i\mathcal{F}_{r,r+1}^{p,i}\mathfrak{g}_{r+1}^{p,i} = \mathcal{H}_r^i\mathfrak{g}_r^{p,i}/\mathcal{G}_{r,r+1}^{p,i}\mathcal{H}_{r+1}^i\mathfrak{g}_{r+1}^{p,i} \equiv \mathfrak{g}_r^{p,i}/\mathcal{G}_{r,r+1}^{p,i}\mathfrak{g}_{r+1}^{p,i}$.*

In the remainder of this section we show that this construction is well-defined and that we have a spectral sequence, that is $H^i(\mathfrak{g}_r^{p,\cdot}) = \mathfrak{g}_{r+1}^{p,i}$. Since $d^0\delta_0^p = \rho(\delta_0^p)x_0$, $\mathfrak{g}_1^{p,0}$ consists of those δ_0^p that do end up one order higher than expected, that is, $\rho(\delta_0^p)x_0 \in \mathfrak{f}^{p+1}$. These generators will not disturb the first level normal form terms at order p, so they are well suited for higher level computations. The first level obstructions live in $\mathfrak{g}_1^{p,1}$ and they are what is left of the order p terms once we remove those of the form $\rho(\delta_0^p)x_0$.

Lemma 2. *$\mathfrak{f}_r^{\cdot,0}$ is a filtered Lie algebra and $\mathfrak{f}^{\cdot,1}$ is a filtered $\mathfrak{f}_r^{\cdot,0}$-module.*

Proof. Since $\mathfrak{f}_r^{p,0} \subset \mathfrak{f}^{p,0}$, it follows that $\rho(\mathfrak{f}_r^{p,0})\mathfrak{f}^{q,1} \subset \mathfrak{f}^{p+q,1}$. Then we have to show that $[\mathfrak{f}_r^{p,0}, \mathfrak{f}_r^{q,0}] \subset \mathfrak{f}_r^{p+q,0}$. Induction on r. This is certainly true for $r=0$. Then $[\mathfrak{f}_r^{p,0}, \mathfrak{f}_r^{q,0}] \subset [\mathfrak{f}_{r-1}^{p,0}, \mathfrak{f}_{r-1}^{q,0}] \subset \mathfrak{f}_{r-1}^{p+q,0}$ by the induction assumption. Furthermore, $d^{r-1}[\delta_r^p, \delta_r^q] = \rho(\delta_r^p)d^{r-1}\delta_r^q - \rho(\delta_r^q)d^{r-1}\delta_r^p$ since ρ is a representation, and this implies

$$\begin{aligned} d^{r-1}[\delta_r^p, \delta_r^q] &\in \rho(\mathfrak{f}_{r-1}^{p,0})d^{r-1}\mathfrak{f}_r^{q,0} + \rho(\mathfrak{f}_{r-1}^{q,0})d^{r-1}\mathfrak{f}_r^{p,0} \\ &\in \rho(\mathfrak{f}_{r-1}^{p,0})\mathfrak{f}^{q+r,1} + \rho(\mathfrak{f}_{r-1}^{q,0})\mathfrak{f}^{p+r,1} \\ &\subset \mathfrak{f}^{p+q+r,1}. \end{aligned}$$

This shows that $[\delta_r^p, \delta_r^q] \in \mathfrak{f}_r^{p+q,0}$. That ρ is a representation follows from the fact that it is a restriction (of $\mathfrak{f}^{p,0}$ to $\mathfrak{f}_r^{p,0}$) of a representation. □

Corollary 2. *$d^r[\delta_r^p, \delta_r^q] = \rho(\delta_r^p)d^r\delta_r^q - \rho(\delta_r^q)d^r\delta_r^p$.*

Lemma 3. *$\mathfrak{f}_{r-1}^{p+1,0} \subset \mathfrak{f}_r^{p,0}$.*

Proof. We use induction on r. For $r=1$ the statement reduces to $\mathfrak{f}_0^{p+1,0} \subset \mathfrak{f}_1^{p,0}$. First one has to check that $\mathfrak{f}_0^{p+1,0} \subset \mathfrak{f}_0^{p,0}$, but this follows form the filtration. Then one needs for $\delta^{p+1} \in \mathfrak{f}_0^{p+1,0}$ that $d^0\delta^{p+1} \in \mathfrak{f}_0^{p+1,0}$, but this follows immediately from the definition of d^0.

If $\delta_{r-1}^{p+1} \in \mathfrak{f}_{r-1}^{p+1,0}$ then $\delta_{r-1}^{p+1} \in \mathfrak{f}_{r-2}^{p+1,0}$ and $d^{r-2}\delta_{r-1}^{p+1} \in \mathfrak{f}^{p+r,1}$. To show that $\delta_{r-1}^{p+1} \in \mathfrak{f}_r^{p,0}$ one has to show that $\delta_{r-1}^{p+1} \in \mathfrak{f}_{r-1}^{p,0}$ and $d^{r-1}\delta_{r-1}^{p+1} \in \mathfrak{f}^{p+r,1}$. The first statement follows from the induction assumption. For the second we notice that $d^{r-1}\delta_{r-1}^{p+1} = d^{r-2}\delta_{r-1}^{p+1} + \rho(\delta_{r-1}^{p+1})\mathcal{N}_{r-1}^{r-1,1}x^{r-1} \in \mathfrak{f}^{p+r,1}$. □

Corollary 3. *$\mathfrak{g}_r^{p,0}$ is well defined as a linear space.*

Lemma 4. *The map $d^r : \mathfrak{f}_r^{p,0} \to \mathfrak{f}^{p+r,1}$ induces a unique $\mathbf{d}^r : \mathfrak{g}_r^{p,0} \to \mathfrak{g}_r^{p+r,1}$.*

Proof. Define $\mathbf{d}^r \pi_r^{p,0}\delta_r^p = \pi_r^{p+r,1} d^r \delta_r^p$. Furthermore, $d^r \delta_{r-1}^{p+1,0} \in \mathfrak{f}^{p+r+1,1}$. □

Lemma 5. *$\mathfrak{g}_r^{\cdot,0}$ is a graded Lie algebra.*

Proof. We have to show that $[\mathfrak{g}_r^{p,0}, \mathfrak{g}_r^{q,0}] \subset \mathfrak{g}_r^{p+q,0}$, but this follows from Lemma 2. Furthermore, the bracket should not depend on the choice of representants. Let $\delta_{r-1}^{q+1} \in \mathfrak{f}_{r-1}^{q+1,0}$. Then

$$[\delta_r^p + \delta_{r-1}^{p+1}, \delta_r^q + \delta_{r-1}^{q+1}] - [\delta_r^p, \delta_r^q] \in \mathfrak{f}_{r-1}^{p+q+1,0}$$

and therefore it is 0 in $\mathfrak{g}_r^{p+q,0}$. One can express this as $\pi_r^{p+q}[\delta_r^p, \delta_r^q] = [\pi_r^p \delta_r^p, \pi_r^q \delta_r^q]$. □

Lemma 6. *$d^{r-1}\mathfrak{f}_{r-1}^{p-r+1,0} \subset \mathfrak{f}^{p,1}$.*

Proof. This follows immediately from the definition of $\mathfrak{f}_{r-1}^{p-r+1,0}$. □

Corollary 4. *$\mathfrak{g}_r^{p,1}$ is well defined as a linear space.*

Lemma 7. *$\mathfrak{g}_r^{\cdot,1}$ is a graded $\mathfrak{g}_r^{\cdot,0}$-module.*

Proof. We have to show that $\rho(\mathfrak{g}_r^{p,0})\mathfrak{g}_r^{q,1} \subset \mathfrak{g}_r^{p+q,1}$. First we see that $\rho(\mathfrak{f}_r^{p,0})\mathfrak{f}^{q,1} \subset \mathfrak{f}^{p+q,1}$. Let $\delta_{r-1}^{q+1} \in \mathfrak{f}_{r-1}^{q+1,0}$, $\delta_{r-1}^{p-r+1} \in \mathfrak{f}_{r-1}^{p-r+1,0}$ and $x^{q+1,1} \in \mathfrak{f}^{q+1}$. Then

$$\begin{aligned} & \rho(\delta_r^p + \delta_{r-1}^{p+1})(x^q + d^{r-1}\delta_{r-1}^{q-r+1} + x^{q+1}) - \rho(\delta_r^p)x^q = \\ = \; & \rho(\delta_r^p)d^{r-1}\delta_{r-1}^{q-r+1} + \rho(\delta_{r-1}^{p+1})d^{r-1}\delta_{r-1}^{q-r+1} + \rho(\delta_r^p)x^{q+1} + \rho(\delta_{r-1}^{p+1})x^q + \rho(\delta_{r-1}^{p+1})x^{q+1} \\ \in \; & \rho(\mathfrak{f}_r^{p,0})\mathfrak{f}^{q+1,1} + \rho(\mathfrak{f}_{r-1}^{p+1,0})\mathfrak{f}^{q,1} \subset \mathfrak{f}^{p+q+1,1} \end{aligned}$$

and this difference is zero in $\mathfrak{g}_r^{p+q,1}$. We can express this as $\pi_r^{p+q}\rho(\delta_r^p)x^q = \rho(\pi_r^p\delta_r^p)\pi_r^q x^q$. □

Corollary 5. *$\mathbf{d}^r[\pi_r^{p,0}\delta_r^p, \pi_r^{q,0}\delta_r^q] = \rho(\pi_r^{p,0}\delta_r^p)\mathbf{d}^r\pi_r^p\delta_r^q - \rho(\pi_r^{q,0}\delta_r^q)\mathbf{d}^r\pi_r^p\delta_r^p$.*

Proof.

$$\begin{aligned} \mathbf{d}^r[\pi_r^p\delta_r^p, \pi_r^q\delta_r^q] &= \mathbf{d}^r\pi_r^{p+q,0}[\delta_r^p, \delta_r^q] \\ &= \pi_r^{p+q+r,1}d^r[\delta_r^p, \delta_r^q] \\ &= \pi_r^{p+q+r,1}\rho(\delta_r^p)d^r\delta_r^q - \pi_r^{p+q+r,1}\rho(\delta_r^q)d^r\delta_r^p \\ &= \rho(\pi_r^{p,0}\delta_r^p)\pi_r^{q+r,1}d^r\delta_r^q - \rho(\pi_r^{q,0}\delta_r^q)\pi_r^{p+r,1}d^r\delta_r^p \\ &= \rho(\pi_r^{p,0}\delta_r^p)\mathbf{d}^r\pi_r^{q,0}\delta_r^q - \rho(\pi_r^{q,0}\delta_r^q)\mathbf{d}^r\pi_r^{p,0}\delta_r^p. \end{aligned}$$

□

Theorem 2. *$\mathfrak{g}_r^{\cdot,\cdot}$ is a spectral sequence, that is, $H^i(\mathfrak{g}_r^{\cdot,\cdot}) = \mathfrak{g}_{r+1}^{\cdot,i}$.*

Proof. We have to compute the cohomology at $i = 0, 1$.
We start with $i = 0$: $H^0(\mathfrak{g}_r^{p,\cdot}) = \ker \mathbf{d}^r = \{\delta_r^p \in \mathfrak{f}_r^{p,0} | d^r\delta_r^p \in d^{r-1}\mathfrak{f}_{r-1}^{p+1,0} + \mathfrak{f}^{p+r+1,1}\}$. That is, $d^r\delta_r^p = d^{r-1}\delta_{r-1}^{p+1} + x^{p+r+1}$. Put $\epsilon^p = \delta_r^p - \delta_{r-1}^{p+1}$. Then $d^r\epsilon^p = d^r\delta_r^p - d^r\delta_{r-1}^{p+1} = d^{r-1}\delta_{r-1}^{p+1} + x^{p+r+1} - d^r\delta_{r-1}^{p+1} = -\rho(\delta_{r-1}^{p+1})\mathcal{N}_r^{r,1}x^r + x^{p+r+1} \in \mathfrak{f}^{p+r+1,1}$. Thus $\epsilon^p \in \mathfrak{f}_{r+1}^{p,0}$. In other words, $\delta_r^p \in \mathfrak{f}_{r-1}^{p+1,0} + \mathfrak{f}_{r+1}^{p,0}$. Thus

$$H^0(\mathfrak{g}_r^{p,\cdot}) \quad \subset \quad (\mathfrak{f}_{r-1}^{p+1,0} + \mathfrak{f}_{r+1}^{p,0})/\mathfrak{f}_{r-1}^{p+1,0} = \mathfrak{f}_{r+1}^{p,0}/(\mathfrak{f}_{r-1}^{p+1,0} \cap \mathfrak{f}_{r+1}^{p,0}) \subset \mathfrak{f}_{r+1}^{p,0}/\mathfrak{f}_r^{p+1,0} = \mathfrak{g}_{r+1}^{p,0}.$$

On the other hand, if $\delta^p_{r+1} \in \mathfrak{g}^{p,0}_{r+1}$, then $d^r \delta^p_{r+1} \in \mathfrak{f}^{p+r+1,1}$ and this implies that $\delta^p_{r+1} \in \ker \mathbf{d}^r$. Next we take $i = 1$:

$$\begin{aligned} H^1(\mathfrak{g}^{p,\cdot}_r) &= \mathfrak{g}^{p,1}_r / \mathbf{d}^r \mathfrak{g}^{p-r,0}_r \\ &= (\mathfrak{f}^{p,1}/(d^{r-1}\mathfrak{f}^{p-r+1,0}_{r-1} + \mathfrak{f}^{p+1,1}))/((d^r \mathfrak{f}^{p-r,0}_r + \mathfrak{f}^{p+1,1})/(d^{r-1}\mathfrak{f}^{p-r+1,0}_{r-1} + \mathfrak{f}^{p+1,1})) \\ &= \mathfrak{f}^{p,1}/(d^r \mathfrak{f}^{p-r,0}_r + \mathfrak{f}^{p+1,1}) \\ &= \mathfrak{g}^{p,1}_{r+1}. \end{aligned}$$

This completes the proof. □

3 Computing the normal form

In this section a normal form algorithm is given that improves on the standard degree by degree computation and converges quadratically in the first level case.

Notation 2. *Let $\delta^1_{l:k}, k < l$, be a generator transforming $x_k + x^{k+1}$ to*

$$\exp(\delta^{1,0}_{l:k})(x_k + x^{k+1}) = x_l + x^{l+1}.$$

Lemma 8. *Given $x \in \mathfrak{f}^{0,1}_p$ one can determine $\delta^1_{p+1:p} \in \bigoplus_{q=0}^p \mathcal{N}^{p+1-q,0}_q \mathfrak{f}^{p+1-q,0}_q$ such that*

$$\exp(\delta^1_{p+1:p})(x_p + x^{p+1}) = x_{p+1} + x^{p+2}.$$

Proof. We see that $\exp(\delta^1_{p+1:p})(x_p + x^{p+1}) \in x_p + x^{p+1} + d^p \delta^1_{p+1:p} + \mathfrak{f}^{p+2,1}$. This can be solved as claimed according to Theorem 1. Doing this in a concrete problem can still be a lot of work, and books have been written about this, for instance [Mur03, SVM07]. □

Lemma 9. *Suppose $\delta^1_{l:k}$ transforms $x_k + x^{k+1}$ to $x_l + x^{l+1}$ and $\delta^1_{m:l}$ transforms $x_l + x^{l+1}$ to $x_m + x^{m+1}$. Let, with* ad *replacing ρ in the definition of* exp,

$$\delta^1_{m:l} \curlywedge \delta^1_{l:k} = \delta^1_{l:k} + \int_0^1 \psi[\exp(t\delta^1_{m:l})\exp(\delta^1_{l:k})]\delta^1_{m:l} dt \mod \mathfrak{f}^{m+1,0}_0,$$

where $\psi[z] = \log(z)/(z-1)$. Then $\delta^1_{m:k} = \delta^1_{m:l} \curlywedge \delta^1_{l:k}$ transforms $x_k + x^{k+1}$ to $x_m + x^{m+1}$.

Proof. Since both the original representation ρ and the adjoint representation play a role, we write in this proof $\exp(\pi(\delta)) = \sum_{n=0}^{\infty} \frac{1}{n!}\pi^n(\delta)$ instead of $\exp(\delta)$. Let $\delta^1, \zeta^1 \in \mathfrak{f}^1$. The goal is to find a formula for $\delta^1 \curlywedge \zeta^1$ such that

$$\exp(\rho(\delta^1 \curlywedge \zeta^1)) = \exp(\rho(\delta^1))\exp(\rho(\zeta^1)).$$

To this end we let $\exp(\rho(Z(t))) = \exp(t\rho(\delta^1))\exp(\rho(\zeta^1))$. Differentiating with respect to t gives (see [SVM07, p.330])

$$\rho(\frac{\exp(\text{ad } Z(t)) - 1}{\text{ad } Z(t)} \frac{dZ}{dt})\exp(\rho(Z(t))) = \rho(\delta^1)\exp(t(\rho(\delta^1))\exp(\rho(\zeta^1))$$

or

$$\frac{\exp(\text{ad } Z(t)) - 1}{\text{ad } Z(t)} \frac{dZ}{dt} = \delta^1.$$

Since $Z(0) = \zeta^1$, this leads to

$$\frac{dZ}{dt} = \frac{\text{ad } Z(t)}{\exp(\text{ad } Z(t)) - 1}\delta^1, \quad Z(0) = \zeta^1.$$

Then $\delta^1 \curlywedge \zeta^1 = Z(1) = \zeta^1 + \int_0^1 \psi(\exp(t\text{ad } \delta^1)\exp(\text{ad } \zeta^1)\delta^1 dt$. □

Remark 1. *This is the right-invariant formulation. The left-invariant formulation can be found in [Mur04]. The* ψ *there has an extra* x*, following from the* **big Ad-little ad** *lemma.*

Remark 2. *One has*

$$\psi(\exp(t\delta^1_{m:l})\exp(\delta^1_{l:k})) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m+1}(\exp(t\delta^1_{m:l})\exp(\delta^1_{l:k}) - 1)^m.$$

Notice that the infinite sum is not all that infinite for our purposes.

Notation 3. *Instead of* $\cdots$ *modulo* $\mathfrak{f}^{k,1}$ *we write* $\cdots + \mathcal{O}(k)$*. The* $\mathcal{O}(\cdot)$*-terms do not have to be computed.*

Theorem 3. *The following recursive procedure computes* $\delta^{m-r+1}_{m:k-1}, k \leq r \leq m$*, the generator putting* $x_{k-1} + x^k$ *into* r*th level normal form* $x_m + \mathcal{O}(m+1)$*, where the recursion in* $\delta_{p:q}$ *is with respect to* $q - p$*.*
Let $k = \sum_{i=0}^{\infty} \kappa_i 2^i$ *and* $m = \sum_{i=0}^{\infty} \mu_i 2^i$ *be the 2-adic expansions of* k *and* m*, that is,* $\kappa_i, \mu_i \in \{0, 1\}$*.*
Determine the minimal p *such that* $\sum_{i=k}^{\infty}(\kappa_i - \mu_i)2^i = 0$ *for* $k \geq p$*.*
Let $\lambda = \sum_{i=p}^{\infty} \mu_i 2^i$ *and* $l = 2^{p-1} + \lambda$*.*
Compute $\delta^{\max(1,k-r)}_{l-1:k-1}$ *using* $x_{k-1} + x^k + d^{\min(k-1,r)}\delta^1_{l-1:k-1} = x_{l-1} + x^l + \mathcal{O}(m+1)$*.*
Determine the order of the generator ν*, where* $\max(1, k-r) \leq \nu$*.*
If $k + \nu \leq m$*, recompute* x^l *with* $\exp(\delta^\nu_{l-1:k-1})(x_{k-1} + x^k) = x_{l-1} + x^l + \mathcal{O}(m+1)$*.*
Compute $\delta^{\max(1,l-r)}_{m:l-1}$ *using* $x_{l-1} + x^l + d^{\min(l-1,r)}\delta^1_{m:l-1} = x_m + \mathcal{O}(m+1)$*.*
Determine the order of the generator μ*, where* $\max(1, l-r) \leq \mu$*.*
Let $\delta^{\min(\mu,\nu)}_{m:k-1} = \delta^\mu_{m:l-1} \curlywedgeleft \delta^\nu_{l-1:k-1}$*, where* $\min(\mu, \nu) \geq \max(1, k-r)$*.*

Proof. If $p = 0$ we are done, since then $l = m$ and we can apply Lemma 8. If $p > 0$ then $\kappa_{p-1} = 0$ and $\mu_{p-1} = 1$. By construction, $k = \sum_{i=0}^{\infty} \kappa_i 2^i = \lambda + \sum_{i=p-2}^{\infty} \kappa_i 2^i < \lambda + 2^{p-1} = l$. Also, $m = \sum_{i=0}^{\infty} \mu_i 2^i = \lambda + 2^{p-1} + \sum_{i=p-2}^{\infty} \mu_i 2^i \geq \lambda + 2^{p-1} = l$. This shows $k < l \leq m$. To get the picture, the reader may want to draw the corresponding binary tree with leaves $\delta^1_{p+1:p}, p = 0, \cdots, m-1$. The condition $k + \nu > m$ is enough to ensure that neither the nonlinear δ-terms nor the $\rho(\delta^1_{l-1:k-1})x^k$ terms can influence the result modulo $\mathfrak{f}^{m+1,1}$: observe that

$$\begin{aligned}\exp(\delta^\nu_{l-1:k-1})(x_{k-1} + x^k) &= x_{k-1} + x^k + \rho(\delta^\nu_{l-1:k-1})(x_{k-1} + x^k) + \frac{1}{2}\rho^2(\delta^\nu_{l-1:k-1})x_{k-1} + \cdots \\ &= x_{l-1} + \frac{1}{2}\rho^2(\delta^\nu_{l-1:k-1})x_{k-1} + \rho(\delta^\nu_{l-1:k-1})x^k + \cdots,\end{aligned}$$

where $\rho^2(\delta^\nu_{l-1:k-1})x_{k-1} \in \rho(\delta^\nu_{l-1:k-1})\mathfrak{f}^{l,1} \in \mathfrak{f}^{l+\nu,1}$ and $\rho(\delta^\nu_{l-1:k-1})x^k \in \mathfrak{f}^{k+\nu,1}$. □

Remark 3. *If one is only interested in the normal form with respect to* x_0*, the first level normal form, then one can replace* $\delta^1_{m:k-1}$ *by* $\delta^k_{m:k-1}$*. One can then compute* $\delta^1_{2^l-1:0}$ *by computing* $\delta^1_{2^{l-1}-1:0} \curlywedgeleft \delta^{2^{l-1}}_{2^l-1:2^{l-1}-1}$*. Now one has* $\nu = 2^{l-1}$*,* $k = 2^{l-1}$*,* $m = 2^l - 1$ *and*

$$k + \nu = 2^{l-1} + 2^{l-1} = 2^l > m.$$

So there is no need for exponentiation within the $\delta^{2^{l-1}}_{2^l-1:2^{l-1}-1}$*-computation. Since the exponentiation is only necessary after computing the* $\delta^1_{2^{l-1}-1:0}$*, the number of exponentiations equals the depth of the binary tree* l*. The Dynkin formula simplifies to*

$$\delta^1_{2^i-1:0} = \delta^1_{2^{i-1}-1:0} + \psi[\exp(\delta^1_{2^{i-1}-1:0})]\delta^{2^{i-1}}_{2^i-1:2^{i-1}-1} \mod \mathfrak{f}^{2^i,0}_0, \quad i = 1, \ldots, l.$$

This is called **quadratic convergence***, since at each step the accuracy is* $\mathcal{O}(2^i), i = 1, \ldots, l$*.*

Remark 4. *Apart from computing the complete transformation the present scheme has as an advantage that the order to which one needs to compute the* exp *is much lower than when one has to compute the* exp *to the final accuracy at every step along the way. The number of exponentations may also be lower, depending on the* $\min(k, \nu) + \nu \leq m$ *test. So the convergence in general will be between linear and quadratic convergence.*

4 The $\mathcal{O}(8)$ calculation in detail

We take (in Theorem 3) $k = 1$ and $m = 7$ and spell out the details. Take $\delta^1_{7:0}$. Then $k = 1, m = 7, p = 3$ and $\lambda = 0$. Thus $l = 4$ and we have to compute $\delta^1_{3:0}$ and $\delta^1_{7:3}$. Take $\delta^1_{3:0}$. Then $k = 1, m = 3, p = 2$ and $\lambda = 0$. Thus $l = 2$ and we have to compute δ^1_0 and $\delta^1_{3:1}$. Take $\delta^1_{7:3}$. Then $k = 4, m = 7, p = 2$ and $\lambda = 4$. Thus $l = 6$ and we have to compute $\delta^1_{5:3}$ and $\delta^1_{7:5}$. Etcetera.
Let us write out the $\curlywedge$ expressions:

$$\begin{aligned}
\delta^1_{7:0} &= \delta^1_{7:3} \curlywedge \delta^1_{3:0} \\
&= (\delta^1_{7:5} \curlywedge \delta^1_{5:3}) \curlywedge (\delta^1_{3:1} \curlywedge \delta^1_0) \\
&= ((\delta^1_6 \curlywedge \delta^1_5) \curlywedge (\delta^1_4 \curlywedge \delta^1_3) \curlywedge ((\delta^1_2 \curlywedge \delta^1_1) \curlywedge \delta^1_0)
\end{aligned}$$

We can read of the exponentials we have to compute by looking at the right hand sides of $\curlywedge$: $\exp(\delta^1_0), \exp(\delta^1_1), \exp(\delta^1_3), \exp(\delta^1_5), \exp(\delta^1_{5:3}), \exp(\delta^1_{3:0})$.
Let us now go through the scheme for an $\mathcal{O}(8)$-computation. We label the steps that only need to be taken in the higher level case by HL, and the ones that need to be taken instead in the first level case by 1L.

- We take $x_0 + x^1 + \mathcal{O}(2)$ and determine $\mathcal{N}^{1,1}_1 x^1$ and $\delta^{1,0}_0$.
- We compute $\exp(\delta^1_0)(x_0 + x^1) = x_1 + x^2 + \mathcal{O}(4)$.
- We take $x_1 + x^2 + \mathcal{O}(3)$ and determine $\mathcal{N}^{2,1}_2 x^2$ and δ^1_1.

HL: We compute $\exp(\delta^1_1)(x_1 + x^2) = x_2 + x^3 + \mathcal{O}(4)$.

- We take $x_2 + x^3 + \mathcal{O}(4)$ and determine $\mathcal{N}^{3,1}_3 x^3$ and δ^1_2.
- We compute $\delta^1_{3:1} = \delta^1_2 \curlywedge \delta^1_1$.

HL We compute $\delta^1_{3:0} = \delta^1_{3:1} \curlywedge \delta^1_0$.

1L We compute $\delta^1_{3:0} = \sum_{i=1}^2 \delta^1_i \curlywedge \delta^1_0$.

- We compute $\exp(\delta^1_{3:0})(x_0 + x^1) = x_3 + x^4 + \mathcal{O}(8)$.
- We take $x_3 + x^4 + \mathcal{O}(5)$ and determine $\mathcal{N}^{4,1}_4 x^4$ and δ^1_3.

HL: We compute $\exp(\delta^1_3)(x_3 + x^4) = x_4 + x^5 + \mathcal{O}(6)$.

- We take $x_4 + x^5 + \mathcal{O}(6)$ and determine $\mathcal{N}^{5,1}_5 x^5$ and δ^1_4.

HL We compute $\delta^1_{5:3} = \delta^1_4 \curlywedge \delta^1_3$.

- We compute $\exp(\delta^1_{5:4})(x_3 + x^4) = x_5 + x^6 + \mathcal{O}(8)$.
- We take $x_5 + x^6 + \mathcal{O}(7)$ and determine $\mathcal{N}^{6,1}_6 x^6$ and δ^1_5.

HL: We compute $\exp(\delta^1_5)(x_5 + x^6) = x_6 + x^7 + \mathcal{O}(8)$.

- We take $x_6 + x^7 + \mathcal{O}(8)$ and determine $\mathcal{N}^{7,1}_7 x^7$ and δ^1_6.

HL We compute $\delta^1_{7:5} = \delta^1_6 \curlywedge \delta^1_5$.

HL We compute $\delta^1_{7:3} = \delta^1_{7:5} \curlywedge \delta^1_{5:3}$.

HL We compute $\delta^1_{7:0} = \delta^1_{7:3} \curlywedge \delta^1_{3:0}$.

1L We compute $\delta^1_{7:0} = \sum_{i=3}^6 \delta^1_i \curlywedge \delta^1_{3:0}$.

- We compute $\exp(\delta^1_{7:0})(x_0 + x^1) = x_7 + \mathcal{O}(8)$ and are done.

Notice that we assumed here the worst case scenario with linear convergence.

5 Concluding remarks

The failure of quadratic convergence was noticed in concrete normal form calculations and it seems best to first implement this method and then start experimenting. After all, the theory is that the practice should follow the theory. Hopefully, these computations will suggests theoretical improvements whereby the theory can follow the practice again.

References

[Arn75] V. I. Arnol'd. A spectral sequence for the reduction of functions to normal form. *Funkcional. Anal. i Priložen.*, 9(3):81–82, 1975.

[BC06] Martin Bendersky and Richard C. Churchill. A spectral sequence approach to normal forms. In *Recent developments in algebraic topology*, volume 407 of *Contemp. Math.*, pages 27–81. Amer. Math. Soc., Providence, RI, 2006.

[Mur03] James Murdock. *Normal forms and unfoldings for local dynamical systems.* Springer Monographs in Mathematics. Springer-Verlag, New York, 2003.

[Mur04] James Murdock. Hypernormal form theory: foundations and algorithms. *J. Differential Equations*, 205(2):424–465, 2004.

[Mur05] James Murdock. Refined computation of hypernormal forms. In *SPT 2004—Symmetry and perturbation theory*, pages 229–235. World Sci. Publ., Hackensack, NJ, 2005.

[San02] Jan A. Sanders. A spectral sequence approach to normal forms. In *SPT 2002: Symmetry and perturbation theory (Cala Gonone)*, pages 231–239. World Sci. Publishing, River Edge, NJ, 2002.

[San03] Jan A. Sanders. Normal form theory and spectral sequences. *J. Differential Equations*, 192(2):536–552, 2003.

[San05] Jan A. Sanders. Normal form in filtered Lie algebra representations. *Acta Appl. Math.*, 87(1-3):165–189, 2005.

[SVM07] J. A. Sanders, F. Verhulst, and J. Murdock. *Averaging methods in nonlinear dynamical systems*, volume 59 of *Applied Mathematical Sciences.* Springer, New York, second edition, 2007.

[vdMSV94] J.-C. van der Meer, J. A. Sanders, and A. Vanderbauwhede. Hamiltonian structure of the reversible nonsemisimple 1 : 1 resonance. In *Dynamics, bifurcation and symmetry (Cargèse, 1993)*, volume 437 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 221–240. Kluwer Acad. Publ., Dordrecht, 1994.