Normal form in filtered Lie algebra representations

Jan A. Sanders
Vrije Universiteit
Faculty of Sciences
Department of Mathematics
De Boelelaan 1081a
1081 HV Amsterdam
The Netherlands

November 5, 2004

Keywords: Normal forms, spectral sequence, filtered Lie algebras

1. Introduction

In this paper I have tried to formulate a setup for normal form theory. It is by no means complete, but it indicates the direction.

On a more elementary level such a setup is given in (Murdock, 2003b), but even there the reader is supposed to have some idea what normal form theory is good for at the start. There one can find references and historical remarks, which this paper lacks completely.

Normal form theory is both a very practical theory with practical results and a theoretical science, which tries to abstract away the theory from the practice.

For a given problem it can deliver to you specific numbers, bifurcation diagrams and as such there is always a certain demand for it. Luckily, for most concrete problems, normal form theory is not very difficult and most people manage to invent it, or anyway use it with some success.

The theoretical part has seen a development toward concepts such as filtered Lie algebras and spectral sequences.

I have tried to connect the two parts. Not by illustrating everything by examples, but by giving serious attention to the computational problem. This leads to an analysis of the computational problem in terms of the spectral sequence. I hope in this way the reader can appreciate the use of the spectral sequence approach, which looks very abstract at first sight, but it is firmly rooted in the algorithmic practice.

Let me try to motivate some of the big words here for those who need it. The concept of a filtration is familiar to anyone who has ever done some perturbation calculation (and who hasn't?). Let us look at vectorfields around some equilibrium. If we write out the Taylor expansion of the vectorfield it looks in $\mathbb{R}$ like

$$\sum_{i=1}^{\infty} a_i x^i \frac{d}{dx}.$$

If one takes the Lie bracket of any two homogeneous terms, we obtain

$$[x^i \frac{d}{dx}, x^j \frac{d}{dx}] = (j-i)x^{i+j-1}\frac{d}{dx}.$$

This strongly suggests to define the degree by: $\deg(x^i \frac{d}{dx}) = i-1$. For indeed, $\deg([x^i \frac{d}{dx}, x^j \frac{d}{dx}] = (i-1)+(j-1) = \deg(x^i \frac{d}{dx}) + \deg(x^j \frac{d}{dx})$. We express this by saying that we have a **graded Lie algebra**. We can also consider this as a filtered Lie algebra $\mathfrak{F}^0 \supset \mathfrak{F}^1 \supset \cdots$. Its elements will be of the form

$$\sum_{i=k}^{\infty} a_i x^i \frac{d}{dx} \in \mathfrak{F}^{k-1},$$

and we find $[\mathfrak{F}^k, \mathfrak{F}^l] \subset \mathfrak{F}^{k+l}$. So in the graded case we look at the homogeneous parts, in the filtered case we drag along the tail.

In the beginning the shift by one (or two if you consider Hamiltonian functions) is very confusing. But if you adhere to it, it prevents you from giving talks about the following question: I'm interested in the term of degree 4 of the normal form. Which terms of the original vectorfield play a role in the formula for it? If you formulate this in terms of gradings, the question is trivial. I once had to sit through a talk which essentially treated this question.

Once the reader is convinced that the concepts of grading and filtering take care of the bookkeeping aspects, the following hurdle will be the introduction of the spectral sequence. When one computes the unique normal form one has to come up with transformations that start with low order terms in order to remove higher order terms. This is new with respect to first order normal form theory, where only the linear part of the vectorfield plays a role, and since it has degree zero, one can always work with homogeneous terms and do the linear algebra in the space of homogeneous terms. The spectral sequence was invented to do the bookkeeping of exactly this kind of situation. It collects in certain spaces (called Z_n^m) the vectorfields of filtering degree n and the transformations that take elements of filtering degree n and transform them to filtering degree $m + n$. Since normal form theory is about removing terms by transformations, one next defines spaces $\mathbf{E}_n^m$, dividing out whatever can be removed and some stuff that is not necessary at this filtering degree. We are now in a graded situation, where we can pose a very exact and relevant question. If we have determined $\mathbf{E}_n^m$ then we know the vectorfields in mth order normal form and the transformations that still can be used without ruining our earlier computations (this last condition is very important, since we work with transformations starting at a low degree).

REMARK 1. *If you know what a spectral sequence is, you may at this point wonder what the differential is. That is indeed the main question. But this question was answered by Arnol'd (see (*Arnol'd et al., 1985*) for the most accessible source). The fact that this answer was given in the context of singularity theory shows that abstraction is the way to go, since we have similar methods working in different problem areas (this is not to imply that singularity theory can simply be done using normal form theory, it is more complicated than that). The amazing thing is that the property that the differential is indeed a differential is completely trivial, and one would expect its consequences to be completely trivial too. Maybe they are, but they certainly do the bookkeeping right!*

*The first application of this idea to normal form theory of ordinary differential equations at equilibrium under orbital equivalence is given by Bogdanov in his treatment of the planar nilpotent unique normal form (*Bogdanov, 1979*). This paper was never translated into English and did not have the impact it should have had. It might make a fine research topic to see how far one can simplify the analysis by combining the spectral sequence techniques with the* $\mathfrak{sl}(2,\mathbb{R})$ *techniques as employed in (*Baider and Sanders, 1992*). Notice that in the last mentioned paper orbital equivalence is not considered.*

One should add here that the usual miracles of spectral sequences do not work in our case, since they are usually based on the finite dimensionality of the underlying topological problem (in this context the following theoretical question comes to mind: if $\mathbf{E}^m$ *is trivial for some* $m < \infty$*, how much does this tell us about the original problem, or, in other words, can one characterize those vectorfields which have a unique normal form as in Corollary 4?).*

*But some of the methods, like the Tic-Tac-Toe lemma (*Bott and Tu, 1982, *Proposition 12.1), are useful in our context. Another difference that should be remarked upon is the fact that the boundary operator is changing as we go along. This is very natural in normal form theory, but less so in topology. The whole setup is extremely sensitive to this, and for instance prohibits (as far as I can see) to set the whole thing up with quadratic convergence. The problem lies with Lemma 7.*

Once we have computed the spectral sequence to a certain order, we know exactly what kind of transformations we can still use at each stage of the actual computation. This is nice, since it allows us to estimate how fast we will converge in the filtration topology (where high degree means small neighborhood of zero). I have tried to give this analysis as detailed as I could, but here we find that the theory still needs to be expanded in order to describe the practical situation. Nevertheless, at the end we arrive at a situation where we can start computing and experiment with the different methods as suggested by the theory. I may add that the last sections grew out of my frustration with my attempts to implement in a FORM (Vermaseren, 1991; Vermaseren, 2000) program quadratic convergence for higher order normal form computations. Sometimes the program worked very nicely, and at other times, with different input, not at all. Having done the theoretical analysis I can now turn back to the calculations...

Since the theory seems rather abstract, I have included a few examples. The main example is published in (Sanders, 2003) and will not be repeated here.The reader is strongly advised to study this example along with this paper. The examples I included are the Hamiltonian

1 : 2-resonance, which was published originally in (Sanders and van der Meer, 1992), and the averaging over spatial variables, as treated for instance in (Sanders, 1977; Sanders, 1979). Both examples do not fit in the theory perfectly, but this illustrates that the assumptions made in the theoretical development are there for simplicity, they are not the hard boundaries of the framework. All these examples share the fact that they look completely trivial to start with, and are still more complicated than one might have thought. As the normal form of the Bogdanov-Takens singularity shows (Baider and Sanders, 1992), it requires a complicated paper to do one example and even that incompletely.

The present theory using spectral sequences does not claim to magically simplify all these examples. What it does claim is that it provides a natural language in which to state results and that it simplifies some of the more complicated filtration arguments that one finds in the literature. Nevertheless, translating the Bogdanov-Takens singularity analysis into the spectral sequence language is still a formidable undertaking, which would only make sense if one were to use it as a guide to work on three dimensional nilpotent singularities.

Another area of applications might be the theory of normal forms around manifolds, see (Bronstein and Kopanskiĭ, 1994).

ACKNOWLEDGMENT 1. *I would like to thank Jim Murdock for long e-mail discussions. Several of the results in his paper (*Murdock, 2003a*) (which he was writing at the same time the present paper was written) found their way into this paper (how would this result look in my notation?). I hope he can recognize them.*

2. Filtered Lie algebra representations

NOTATION 1. *We denote the elements of $\mathbb{Z}/2$ by $\{0,1\}$.*

We consider a filtered Lie algebra representation consisting of spaces $\mathfrak{F}_l^k$, with $l \in \mathbb{Z}/2$ and $k \in \mathbb{N}$, such that $\mathfrak{F}_l^0 \supset \mathfrak{F}_l^1 \supset \mathfrak{F}_l^2 \supset \cdots \supset \mathfrak{F}_l^k \supset \cdots$ with $\cap_{k=0}^{\infty} \mathfrak{F}_l^k = 0$. Here $\mathfrak{F}_0^1$ is the filtered Lie algebra: $[\mathfrak{F}_0^k, \mathfrak{F}_0^n] \subset \mathfrak{F}_0^{k+n}$, and $\mathfrak{F}_1^0$ is the filtered representation space: $\mathfrak{F}_1^0$ is a module or a vector space, for which there exists a representation ρ such that for any $\mathfrak{f}_0^k \in \mathfrak{F}_0^k$ one has a linear map $\rho(\mathfrak{f}_0^k) : \mathfrak{F}_1^m \to \mathfrak{F}_1^{m+k}$, with $\rho([\mathfrak{f}_0^k, \mathfrak{g}_0^m]) = [\rho(\mathfrak{f}_0^k), \rho(\mathfrak{g}_0^m)]$, where the last bracket is just the ordinary commutator $[A,B] = AB - BA$. One could write this as $\rho \circ \mathrm{ad} = \mathrm{ad} \circ \rho$, with $\mathrm{ad}(x)y = [x,y]$.

REMARK 2. *One could also set up the theory with $\mathfrak{F}_0^0$ instead of $\mathfrak{F}_0^1$, to include linear transformations. This would be a bit more directed* *toward the applications, but it introduces some technical complications that I would like to avoid here. See however section 10.*

This induces an adjoint representation of $K^0 = \mathfrak{F}_1^0 \oplus \mathfrak{F}_0^1$ into itself, which we denote by ∇: $\nabla_{(\mathfrak{f}_1^0,\mathfrak{f}_0^1)}(\mathfrak{g}_1^0,\mathfrak{g}_0^1) = (\rho(\mathfrak{f}_0^1)\mathfrak{g}_1^0 - \rho(\mathfrak{g}_0^1)\mathfrak{f}_1^0, \mathrm{ad}(\mathfrak{f}_0^1)\mathfrak{g}_0^1)$. Here the direct sum is the vector space or module direct sum. We write $x \in K^0$ as (x_1, x_0), but we write x_1 for $(x_1, 0)$ and x_0 for $(0, x_0)$. The construction is the standard one for a trivial extension (as a Lie algebra) of $\mathfrak{F}_0^1$ by $\mathfrak{F}_1^0$: one has to check the formula

$$[\nabla_{\mathfrak{f}}, \nabla_{\mathfrak{g}}] = \nabla_{\nabla_{\mathfrak{f}}\mathfrak{g}}.$$

We call (K^0, ∇) obeying this relation, but not necessarily the antisymmetry $\nabla_x y + \nabla_y x = 0$, a **Leibniz algebra** (Loday, 1991). There is no harm in replacing the word Lie algebra by Leibniz algebra in this paper, since antisymmetry is never used. The results formulated as corollaries for ∇ are valid for general filtered Leibniz algebras, if they make sense, that is, if everything converges. Of course, K^0 is in this specific case also a Lie algebra, with $\nabla_x = \mathrm{ad}(x)$ and convergence in the filtration topology is easy to verify.

Let $\mathcal{G}_l^k = \mathfrak{F}_l^k / \mathfrak{F}_l^{k+1}$. Then we assume $\dim \mathcal{G}_l^k < \infty$. All our computations will be done in the $\mathcal{G}_l^k$, so the representations we work with will be finite-dimensional. This assumption is made so that we can guarantee a solution to the homological equation that will appear later on. In practice infinite dimensional spaces do play an important role, but as long as the homological equation can be solved, this presents no difficulty, see section 11. For more information about solving the homological equations, see the work of Belitskii, surveyed in (Belitskii, 2002).

We write elements in $\mathcal{G}_l^k$ as g_l^k. Notice that $\mathrm{ad}(\mathcal{G}_0^{k_1})\mathcal{G}_0^{k_2} \subset \mathcal{G}_0^{k_1+k_2}$ and $\rho(\mathcal{G}_0^{k_1})\mathcal{G}_1^{k_2} \subset \mathcal{G}_1^{k_1+k_2}$, that is, we now have a graded Lie algebra representation. The theory will apply to those cases where there is no $\mathbb{Z}/2$-grading, in which case the representation will be the adjoint representation. In that case we define ∇ by $\nabla_f g = [f, g]$ and call it ad. We call the elements in $\mathfrak{F}_0^1$ *transformations* (we think of them as infinitesimal generators of coordinate transformations) and the elements in $\mathfrak{F}_1^0$ *(vector)fields*, since local vectorfields around some equilibrium are the main example we are working from. We call $f^0 \in \mathcal{G}_1^0$ the *linear part* of the field, even though it may correspond to quadratic terms of a Hamiltonian as in section 10. This terminology will help to make things not too abstract, but on the other hand the reader should at some point appreciate the abstract approach and try to think of examples outside the main stream of normal form theory.

REMARK 3. *If we consider reversible vectorfields, then one chooses* $\mathfrak{F}_l$ *such that its elements change as* $R\mathfrak{f}_l = (-1)^l \mathfrak{f}_l$ *under the action of the* reversor R. *Thus* $\mathbb{Z}/2$ *is generated by the identity and the reversor. In this case one starts by defining* ∇ *as the adjoint representation on* K^0, *and deriving* ρ *by restriction.*

REMARK 4. *Another case where the representation has nothing to do with the adjoint, is that of symmetric forms under the action of the orthogonal group. More general, one could consider the problem of flattening the coordinate expression for a Riemannian metric by near-identity coordinate transformations. This is basically the way Riemann approaches the problem of normalizing the metric in his inaugural lecture in 1854, which leads to the definition of connection by looking at the transformation that does the best job, and the definition of curvature by looking at the obstruction (the normal form) to completely flatten the metric, cf. (*Spivak, 1979*). The reader may find it interesting to translate these classical results in terms of the spectral sequences to be defined later in this paper. For further nonstandard examples, see (*Meyer, 1994*).*

REMARK 5. *If one studies Hamiltonian vectorfields, one is often interested in properties (like eigenvalues) that remain invariant under formal diffeomorphisms. In such a situation one might choose for* $\mathfrak{F}_1^0$ *the space of formal Hamiltonians, and for* $\mathfrak{F}_0^1$ *the space of formal vectorfields, generating the formal diffeomorphisms.*

DEFINITION 1. *One can define for* $\mathfrak{h}_0^1 \in \mathfrak{F}_0^1$ *its* **exponential** *as the formal expression*

$$e^{\rho(\mathfrak{h}_0^1)} = \sum_{i=0}^{\infty} \frac{1}{i!} \rho^i(\mathfrak{h}_0^1).$$

This defines the filtered action of a formal group on $\mathfrak{F}_1^{\cdot}$. *If for given* $\mathfrak{f}_1^0, \mathfrak{g}_1^0 \in \mathfrak{F}_1^0$ *there exists an* $\mathfrak{h}_0^1 \in \mathfrak{F}_0^1$ *such that* $\mathfrak{g}_1^0 = \exp(\rho(\mathfrak{h}_0^1))\mathfrak{f}_1^0$, *then we say that* $\mathfrak{f}_1^0$ *and* $\mathfrak{g}_1^0$ *are formally equivalent and we write* $\mathfrak{f}_1^0 \equiv \mathfrak{g}_1^0$.

We consider the following problems here: to find all equivalence classes, to define a normal form for each equivalence class and to determine the transformation which brings a given $\mathfrak{f}_1^0$ in its normal form. This gives us a map from the representation space to its transformation group, which we call the *frame map*, since it maps the vectorfield onto the transformation that brings it into a certain normal form, that is, in a cross section of the bundle of all vectorfields with the group of formal transformations acting on it. This transformation can be identified with a moving frame in the sense of Cartan. We remark that computationally there is quite a difference between just finding a certain normal form with or without its transformation. This also reflects on the efficiency considerations.

NOTATION 2. *For fixed* $n \in \mathbb{N}$, *let* $\mathcal{H}_0^n = \mathfrak{F}_0^1/\mathfrak{F}_0^{n+1}$ *and* $\mathcal{H}_1^n = \mathfrak{F}_1^0/\mathfrak{F}_1^{n+1}$. *These are by assumption finite-dimensional spaces.*

3. First order normal form

To obtain some idea of what is involved here, let us write out the computation to the first order: Take $\mathfrak{f}_1^0 = \mathrm{f}_1^0 + \mathrm{f}_1^1 + \mathfrak{f}_1^2$, $\mathfrak{g}_1^0 = \mathrm{g}_1^0 + \mathrm{g}_1^1 + \mathfrak{g}_1^2$ and $\mathfrak{h}_0^1 = \mathrm{h}_0^1 + \mathfrak{h}_0^2$. Then

$$\mathrm{g}_1^0 + \mathrm{g}_1^1 + \mathfrak{g}_1^2 = \exp(\rho(\mathrm{h}_0^1 + \mathfrak{h}_0^2)(\mathrm{f}_1^0 + \mathrm{f}_1^1 + \mathfrak{f}_1^2) = \mathrm{f}_1^0 + \mathrm{f}_1^1 + \rho(\mathrm{h}_0^1)\mathrm{f}_1^0 \mathrm{mod} \mathfrak{F}_1^2.$$

It follows that $\mathrm{g}_1^0 = \mathrm{f}_1^0$ and $\mathrm{g}_1^1 = \mathrm{f}_1^1 + \rho(\mathrm{h}_0^1)\mathrm{f}_1^0$. We now define $\mathsf{d}^0\mathrm{f}_1^0 : \mathcal{H}^1 \to \mathcal{H}^1$ such that $\mathsf{d}^0\mathrm{f}_1^0(x) = -\nabla_x \mathrm{f}_1^0$ (if ρ is the adjoint representation, then $\mathsf{d}^0\mathrm{f}_1^0 = ad(\mathrm{f}_1^0)$). The choice of sign is unconventional in Lie algebra cohomology, but has to do with the role reversal during the computation, and doesn't matter anyway. Observe that $\mathrm{f}_1^0 \neq 0$ does not imply that $\mathsf{d}^0\mathrm{f}_1^0 \neq 0$; after all, the representation can be trivial.

We turn back to our computation, which now gives us the (homological) equation

$$\mathsf{d}^0\mathrm{f}_1^0(\mathrm{h}_0^1) = \mathrm{f}_1^1 - \mathrm{g}_1^1.$$

We see that, at least to first order, the equivalence class of f_1^1 lies in the the cokernel of $\mathsf{d}^0\mathrm{f}_1^0$. We now assume for the moment the existence of an *operator style* $\delta^0\mathrm{f}_1^0 : \mathcal{H}_1^1 \to \mathcal{H}_0^1$ such that $\mathcal{H}_1^1 = \ker \delta^0\mathrm{f}_1^0 \oplus \mathrm{im}\, \mathsf{d}^0\mathrm{f}_1^0$ and $\mathcal{H}_0^1 = \ker \mathsf{d}^0\mathrm{f}_1^0 \oplus \mathrm{im}\, \delta^0\mathrm{f}_1^0$. The choice of $\delta^0\mathrm{f}_1^0$ will in general not be unique and determines the *style* (cf. (Murdock, 2003b)) of the normal form. But it can always be made to any given finite order n, thanks to the fact that by assumption our spaces are finite dimensional and if an inner product is not already present (given by the original problem), we can construct it on $\mathcal{H}_0^1$ and $\mathcal{H}_1^1$, and define $\delta^0\mathrm{f}_1^0$ to be the adjoint of $\mathsf{d}^0\mathrm{f}_1^0$. We now say that f_1^1 is in **first order normal form** (with respect to $\delta^0\mathrm{f}_1^0$) if $\mathrm{f}_1^1 \in \ker \delta^0\mathrm{f}_1^0$ and that h_0^1 is in **first order conormal form** (with respect to $\mathsf{d}^0\mathrm{f}_1^0$) if $\mathrm{h}_0^1 \in \ker \mathsf{d}^0\mathrm{f}_1^0$. Instead of taking the adjoint one could use Jacobson-Morozov imbedding on $\mathcal{H}_1^1 \oplus \mathcal{H}_0^1$ (cf. (Sanders, 1994)), but this is computationally expensive, while the theoretical value is still unproven in the case that the normal form involves higher order terms, as it will in the following sections.

REMARK 6. *From the equation* $\mathsf{d}^0\mathrm{f}_1^0(\mathrm{h}_0^1) = \mathrm{f}_1^1 - \mathrm{g}_1^1$ *it is clear that we can choose* h_0^1 *in a complement of the kernel of* $\mathsf{d}^0\mathrm{f}_1^0$ *and it seems natural* *to take* $\mathrm{h}_0^1 \in \mathrm{im}\ \delta^0\mathrm{f}_1^0$ *and this condition uniquely determines* h_0^1 *for given* $\mathrm{f}_1^1 - \mathrm{g}_1^1$*. Uniqueness in the transformation can be very important: if the system has a symmetry then* $\mathsf{d}^0\mathrm{f}_1^0$ *inherits this symmetry, and if* $\delta^0\mathrm{f}_1^0$ *has the same symmetry then the transformation and the normal form will inherit the symmetry from the original* $\mathfrak{f}_1$*, since there should at least be a symmetric solution to the equation, but this solution is unique (cf. remark 9). This explains the success of rather unsophisticated normalizing methods, treating Hamiltonian systems as ordinary vectorfields. In averaging theory the condition on* $\mathfrak{h}_0^1$ *translates to taking* $\mathfrak{h}_0^1$ *with zero average (so that it is in the image of* $\frac{\partial}{\partial t}$*), a condition that is well known to simplify many results, cf. (*Sanders and Verhulst, 1985*).*

Let $K^n = \mathfrak{F}_1^n \oplus \mathfrak{F}_0^n$ and $Z_n^0 = K^n$. We now define $d_n^0[\mathrm{f}_1^0] : K^n \to K^n$ by

$$d_n^0[\mathrm{f}_1^0](\mathfrak{g}_1^n, \mathfrak{h}_0^n) = \mathsf{d}^0\mathrm{f}_1^0((\mathfrak{g}_1^n, \mathfrak{h}_0^n)).$$

Observe the conceptual difference between $d_n^0[\mathrm{f}_1^0]$ and $\mathsf{d}^0\mathrm{f}_1^0$: the first is a boundary operator, which we usually denote by d_n^0, and the second is a 1-form on K^0, namely the coboundary of the zero form f_1^0. Let $\mathbf{E}_n^0 = Z_n^0/Z_{n+1}^0$. The map $d_n^0 : K^n \to K^n$ induces a $\mathbf{d}_n^0 : \mathbf{E}_n^0 \to \mathbf{E}_n^0$. We see that the kernel of this induced map contains the first order conormal forms, and the cokernel the first order normal forms. Clearly $\mathbf{d}_n^0 \cdot \mathbf{d}_n^0 = 0$.

Let us now formalize what we have done in a language that will be used to define a spectral sequence, the first terms of which we have already seen, namely $\mathbf{E}_n^0$. Let $Z_n^1 = \{x \in K^n | d_n^0 x \in K^{n+1}\}$. These are exactly the terms $(\mathfrak{g}_1^n, \mathrm{h}_0^n + \cdots)$ with $\rho(\mathrm{h}_0^n)\mathrm{f}_1^0 = 0 \in \mathcal{G}_1^n$. Now we divide out the image of d_n^0, that is we put the elements in Z_n^1 in first order normal form and we remove the elements in K^{n+1}, since they are in Z_n^1 no matter what f_1^0 is. This leads to

$$\mathbf{E}_n^1 = Z_n^1/(d_n^0 Z_n^0 + Z_{n+1}^0), \quad n \geq 0.$$

So this space contains terms of the form $(\mathrm{g}_1^n, \mathrm{h}_0^n)$, with g_1^n in first order normal form and h_0^n in first order conormal form. Our claim is now that $\mathbf{E}_n^1 = H^n(\mathbf{E}^0, \mathbf{d}^0) = \ker \mathbf{d}^0 | \mathbf{E}_n^0 / \mathrm{im}\ \mathbf{d}^0 | \mathbf{E}_\mathrm{n}^0$. This means that we can identify first order (co)normal form theory with cohomology of $\mathbf{E}^0$, and of course we would like to generalize this to higher order normal form theory. This definition of first order normal form is completely independent of the choice of style. Identifying $\mathbf{E}_n^1$ as a part of $\mathbf{E}_n^0$ is done by choosing $\delta^0\mathrm{f}_1^0$ (so it is style dependent) and saying that

$$\mathbf{E}_n^1 \cong \ker(\delta^0\mathrm{f}_1^0 | \mathbf{E}_n^0) \oplus \ker(\mathsf{d}^0\mathrm{f}_1^0 | \mathbf{E}_n^0).$$

4. Second order normal form

NOTATION 3. *We write* $\mathrm{f}_l^{q|p}$ *for* $\sum_{j=q}^{p} \mathrm{f}_l^j$*; we call* $\mathrm{f}_l^{k|m}$ *the* m*-jet of* $\mathfrak{f}_l^k$.

Usually one considers Lie algebras over a field, but we also want to include the case where the Lie algebra is over a local ring R, that is, a ring with a unique maximal ideal $\mathfrak{m}$. The field $R/\mathfrak{m}$ is called the *residue field*. We denote the projection of $\mathrm{f}_1^{0|m} \in \mathcal{H}_1^m$ on $\mathcal{H}_1^m/\mathfrak{m}\mathcal{H}_1^m$ by $\mathbf{f}_1^{0|m}$. We have in mind the situation where the original vectorfield contains parameters which may pass through zero, and so we cannot divide by them. So we let $\mathfrak{m}$ be generated by these parameters.

It is now clear how we can obtain the first order normal form of $\mathfrak{f}_1^0$ with respect to $\delta^0 \mathrm{f}_1^0$. Is this it? The answer to that question depends on the problem, but in principle is **no**. Suppose we have

$$e^{\rho(\mathfrak{h}_0^1)} \mathfrak{f}_1^0 = \mathfrak{g}_1^0.$$

If we now want to transform $\mathfrak{g}_1^0$, we better not undo our previous calculations. So whatever we do, we need the result to be again in first order normal form. This implies that the lowest order term of any $\mathfrak{k}_0^1$ that we use has to be in $\ker \mathsf{d}^0 \mathrm{f}_1^0$. So we take $\mathfrak{k}_0^1 = \mathrm{k}_0^1 + \mathrm{k}_0^2 + \mathfrak{k}_0^3, \mathrm{k}_0^1 \in \ker \mathsf{d}^0 \mathrm{f}_1^0$. Then we compute

$$\begin{aligned} &\exp(\rho(\mathfrak{k}_0^1))(\mathrm{g}_1^0 + \mathrm{g}_1^1 + \mathrm{g}_1^2 + \cdots) = \\ &= \ \mathrm{g}_1^0 + \mathrm{g}_1^1 + \rho(\mathrm{k}_0^1)\mathrm{g}_1^0 \\ &+ \ \mathrm{g}_1^2 + \rho(\mathrm{k}_0^2)\mathrm{g}_1^0 + \rho(\mathrm{k}_0^1)\mathrm{g}_1^1 + \tfrac{1}{2}\rho^2(\mathrm{k}_0^1)\mathrm{g}_1^0 \operatorname{mod} \mathfrak{F}_1^3 \\ &= \ \mathrm{g}_1^{0|1} + \mathrm{g}_1^2 + \rho(\mathrm{k}_0^1 + \mathrm{k}_0^2)\mathrm{g}_1^{0|1} \operatorname{mod} \mathfrak{F}_1^3. \end{aligned}$$

The second order normalization is now effectively done by the term $\rho(\mathrm{k}_0^1)\mathrm{g}_1^1$. The term k_0^2 is then used to put the remaining terms in first order normal form again.

Remark that we now have put the system in second order normal form by applying two transformations. Our goal is to find one transformation that does this and we will come back to this question in section 8 and 9.

We now define $d_n^1 : Z_n^1 \to Z_{n+1}^1$ by

$$d_n^1 x = \mathsf{d}^1 \mathbf{f}_1^{0|1}(x) = -\nabla_x \mathbf{f}_1^{0|1}, \quad x \in Z_n^1.$$

This induces a $\mathbf{d}_n^1 : \mathbf{E}_n^1 \to \mathbf{E}_{n+1}^1$. We see that ker $\mathbf{d}_n^1$ contains the second order conormal forms, and coker $\mathbf{d}_\mathrm{n}^1$ the second order normal forms. Clearly $d_n^1 \cdot d_n^1 = 0$.

Let $Z_n^2 = \{x \in K^n | d_n^1 x \in K^{n+2}\}$. These are exactly the terms $(\mathfrak{g}_1^n, \mathrm{h}_0^n + \mathrm{h}_0^{n+1} + \cdots)$ with $\rho(\mathrm{h}_0^n + \mathrm{h}_0^{n+1})\mathbf{f}_1^{0|1} = 0 \in \mathcal{G}_1^{n+1}$. Now we divide out the image of d^1_{n-1}, that is we put the elements in Z^2_n in first order normal form and we remove the elements in Z^1_{n+1}, since they are trivially in Z^2_n. This leads to

$$\mathbf{E}^2_n = Z^2_n/(d^1_{n-1}Z^1_{n-1} + Z^1_{n+1}), \quad n \geq 1.$$

Our claim is now that $\mathbf{E}^2_n = H^n(\mathbf{E}^1, \mathbf{d}^1) = \ker \mathbf{d}^1|\mathbf{E}^1_n/\mathrm{im}\ \mathbf{d}^1|\mathbf{E}^1_{\mathrm{n}}$. Identifying $\mathbf{E}^2_n$ as a part of $\mathbf{E}^1_n$ is done by choosing $\delta^1\mathbf{f}^{0|1}_1$ and saying that

$$\mathbf{E}^2_n \cong \ker(\delta^1\mathbf{f}^{0|1}_1|\mathbf{E}^1_n) \oplus \ker(\mathsf{d}^1\mathbf{f}^{0|1}_1|\mathbf{E}^1_n).$$

5. Definition of normal form

The definition of normal form will now be given as follows.

DEFINITION 2. *We say that* $\mathsf{d}^m\mathrm{f}^{0|m}_1 : \mathcal{H}^{m+1}_0 \to \mathcal{H}^{m+1}_1$ *and* $\delta^m\mathrm{f}^{0|m}_1 : \mathcal{H}^{m+1}_1 \to \mathcal{H}^{m+1}_0$ *define* operator style rules *if*

- $\mathcal{H}^{m+1}_1 = \ker \delta^m\mathrm{f}^{0|m}_1 \oplus \mathrm{im}\ \mathsf{d}^m\mathrm{f}^{0|m}_1$,
- $\mathcal{H}^{m+1}_0 = \ker \mathsf{d}^m\mathrm{f}^{0|m}_1 \oplus \mathrm{im}\ \delta^m\mathrm{f}^{0|m}_1$ *and, with* $n > m$,
- $\mathsf{d}^n\mathrm{f}^{0|n}_1|\mathcal{H}^{m+1}_0 = \mathsf{d}^m\mathrm{f}^{0|m}_1$,
- $\delta^n\mathrm{f}^{0|n}_1|\mathcal{H}^{m+1}_1 = \delta^m\mathrm{f}^{0|m}_1$.

DEFINITION 3. *We say that* $\mathrm{f}^{0|m+1}_1$ *is in* $m+1$*th order* **normal form** *if either it cannot be changed in* $\mathcal{H}^{m+1}_1$ *or* $\mathrm{f}^{0|m}_1$ *is in* m*th order normal form and* $\mathrm{f}^{0|m+1}_1 - \mathbf{f}^{0|m}_1$ *is in* $\ker \delta^m\mathbf{f}^{0|m}_1$*, where the* $\mathsf{d}^m\mathbf{f}^{0|m}_1$ *and* $\delta^m\mathbf{f}^{0|m}_1$ *respect the operator style rules in definition 2 and* $\mathsf{d}^m\mathrm{f}^{0|m}_1(x) = -\nabla_x\mathrm{f}^{0|m}_1$.

So a zeroth order normal form cannot be changed in $\mathcal{H}^0_1$, but this means its linear part cannot be changed, which is true for all vectorfields under the given action, so the concept of normal form is well defined.

A first order normal form then is such that $\mathrm{f}^0_1 + \mathrm{f}^1_1 - \mathbf{f}^1_1 \in \ker \delta^0\mathbf{f}^0_1$, in accordance with the computations we did before (since there $\mathrm{f}^0_1 = \mathrm{f}^0_1 = \mathbf{f}^0_1$).

So far we have only been involved in how to define normal forms. But what about the actual computation? There are basically two approaches:

1. By linear algebra, that is direct elimination.

2. By spectral methods.

The advantage of the first method is that it involves very simple calculations and can always be used, there are no field extensions necessary, but the disadvantage is that this is linear algebra on high-dimensional spaces. The spectral method, which, by the way, has nothing to do with the spectral sequences announced in the abstract, requires the knowledge of the eigenvalues of $\mathsf{d}^0\mathrm{f}_1^0$ (assuming we can identify $\mathfrak{F}_0^1$ and $\mathfrak{F}_1^1$), which is a definite disadvantage, and may also involve computations in field extensions, which can complicate matters a lot. Its advantage is that it is rather efficient on small expressions. It is the spectral method that is most commonly associated with normal form theory and most of normal form theory is dedicated to it, but with the gradual changeover from hand calculations to computer calculations, the linear algebra method is becoming more and more interesting.

Both methods should lead to the following two maps $\partial_1^m : \mathcal{H}_1^m \to \mathrm{im}\, \mathsf{d}^{\mathrm{m}-1}\mathbf{f}_1^{0|\mathrm{m}-1}$ and $\pi_1^m : \mathcal{H}_1^m \to \ker \delta^{m-1}\mathbf{f}_1^{0|m-1}$: Given $\mathrm{f}_1^{n|m} \in \mathcal{H}_1^m$, find $\partial_1^m \mathrm{f}_1^{n|m}$ and $\pi_1^m \mathrm{f}_1^{n|m}$ such that

$$\mathrm{f}_1^{n|m} - \mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}(\delta^{m-1}\mathbf{f}_1^{0|m-1}(\partial_1^m \mathrm{f}_1^{n|m})) = \pi_1^m \mathrm{f}_1^{n|m}.$$

These maps exist thanks to the operator style properties given in definition 2. They will be used implicitly in the theoretical computations in section 8 and 9.

6. Some standard results

Here we formulate some well-known results, which we need in the next section.

PROPOSITION 1. *Let $x, y \in \mathfrak{F}_0^0$. Then for any representation ρ one has*

$$\rho^k(y)\rho(x) = \sum_{i=0}^{k} \binom{k}{i} \rho(\mathrm{ad}^i(y)x)\rho^{k-i}(y).$$

COROLLARY 1. *Let $x, y \in K^0$. Then for any filtered Leibniz algebra (K^0, ∇) one has*

$$\nabla_y^k \nabla_x = \sum_{i=0}^{k} \binom{k}{i} \nabla_{\nabla_y^i x} \nabla_y^{k-i}.$$

The following is known as the *big Ad–little ad* lemma:

PROPOSITION 2. *Let $x \in \mathfrak{F}_0^0$ and $y \in \mathfrak{F}_0^1$. Then for any filtered representation ρ one has*

$$\rho(e^{\mathrm{ad}(y)}x) = e^{\rho(y)}\rho(x)e^{-\rho(y)} = Ad(e^{\rho(y)})\rho(x).$$

COROLLARY 2. *Let $x, y \in L^0$, where L^0 is a filtered Leibniz algebra. Then*

$$\nabla_{e^{\nabla_y}x} = e^{\nabla_y}\nabla_x e^{-\nabla_y} = Ad(e^{\nabla_y})\nabla_x,$$

if both sides converge. In the case where $L^0 = K^0$ convergence is automatic.

COROLLARY 3. *Let $x, y, z \in L^0$, where L^0 is a filtered Leibniz algebra. Then*

$$\mathsf{d}\exp(\nabla_y)z(exp(\nabla_y)x) = \exp(\nabla_y)\mathsf{d}z(x).$$

PROOF 1. *Follows immediately from Corollary 2.*

7. Spectral sequences

The following discussion completely ignores the *style*. This has the advantage that it leaves no room for discussion, but in actual computations one does have to make a choice of *style*.

Let $K^n = \mathfrak{F}_1^n \oplus \mathfrak{F}_0^n$ and for $m \geq 1$ let $d_n^{m-1} : K^n \to K^n$ be defined by $d_n^{m-1}(\mathfrak{g}_1^n, \mathfrak{h}_0^n) = \mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}((\mathfrak{g}_1^n, \mathfrak{h}_0^n))$, where $\mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}(x) = -\nabla_x \mathbf{f}_1^{0|m-1}$. Let $Z_n^0 = K^n$ and

$$Z_n^m = \{x \in K^n | d_n^{m-1}x \in K^{n+m}\}, m \geq 1.$$

LEMMA 1. *The filtered Leibniz (or Lie) algebra $(K^\cdot, \nabla)$, leads to a filtered differential Leibniz (or Lie) algebra $(Z_\cdot^m, \nabla)$. That is, for $x \in Z_p^m$ and $y \in Z_q^m$, one has*

$$d_{p+q}^{m-1}\nabla_x y = \nabla_x d_q^{m-1}y - \nabla_y d_p^{m-1}x,$$

and, consequently, $\nabla_x y \in Z_{p+q}^m$.

REMARK 7. *We avoid here the usage of differential filtered or differential graded since that implies the occurrence of a term $(-1)^{|x|}\nabla_x d_q^{m-1}y$ in the relation (coming from $[x, y] = xy - (-1)^{|x|}yx$). So in our usage the words are taken together in a different order and with a different meaning. The reader may want to verify that, if we let $R(x, y) = \nabla_x\nabla_y - \nabla_y\nabla_x - \nabla_{\nabla_x y}$, then (just using the derivation rule $d\nabla_x y = \nabla_x dy - \nabla_y dx$) one finds $dR(x, y)z = R(x, y)dz - R(x, z)dy + R(y, z)dx$, so it seems to be the natural defining relation of a differential Leibniz algebra (where $R(x, y) = 0$ for all x, y).*

PROOF 2. *Indeed, it is easy to check that* $\nabla_x y \in K^{p+q}$. *Since* $\mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}(\cdot)$ *is a coboundary, it is automatically a cocycle, and*

$$\begin{aligned} d_{p+q}^{m-1}\nabla_x y &= \mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}(\nabla_x y) \\ &= \nabla_x \mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}(y) - \nabla_y \mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}(x) \\ &= \nabla_x d_q^{m-1} y - \nabla_y d_p^{m-1} x \in K^{p+q+m}, \end{aligned}$$

which proves the Lemma.

A similar lemma can be formulated for the spectral sequence $\mathbf{E}_{\cdot}^{m}$, to be defined next, that can be seen as a graded differential Leibniz (or Lie) algebra. Ultimately we want $\mathbf{E}_{\cdot}^{\infty}$ to be a trivial module with respect to the representation. We suppress in our notation the dependence of Z_n^m and d_n^{m-1} on $\mathbf{f}_1^{0|m-1}$. But to explicitly show the dependence on the formal group action we write now $Z_n^m[\mathbf{f}_1^{0|m-1}]$ and $d_n^{m-1}[\mathbf{f}_1^{0|m-1}]$for a while.

LEMMA 2.

$$Z_n^m[\exp(\nabla_{\mathsf{t}_0^1})\mathbf{f}_1^{0|m-1}] = \exp(\nabla_{\mathsf{t}_0^1}) Z_n^m[\mathbf{f}_1^{0|m-1}].$$

PROOF 3. *Indeed, Let* $x \in Z_n^m[\mathbf{f}_1^{0|m-1}]$. *Then, using Corollary 3,*

$$\begin{aligned} & d_n^{m-1}[\exp(\nabla_{\mathsf{t}_0^1})\mathbf{f}_1^{0|m-1}] \exp(\nabla_{\mathsf{t}_0^1}) x = \\ &= \mathsf{d}^{m-1} \exp(\nabla_{\mathsf{t}_0^1})\mathbf{f}_1^{0|m-1}(\exp(\nabla_{\mathsf{t}_0^1}) x) \\ &= \exp(\nabla_{\mathsf{t}_0^1})\mathsf{d}^{m-1}\mathbf{f}_1^{0|m-1}(x) \\ &= \exp(\nabla_{\mathsf{t}_0^1}) d_n^{m-1}[\mathbf{f}_1^{0|m-1}] x \in K^{n+m} \end{aligned}$$

and it follows that $\exp(\nabla_{\mathsf{t}_0^1}) x \in Z_n^m[\exp(\nabla_{\mathsf{t}_0^1})\mathbf{f}_1^{0|m-1}]$. *This gives us one inclusion, the other follows by symmetry.*

The following Lemma has been first formulated by Baider in (Baider and Sanders, 1991, Prop. 2.7).

LEMMA 3. *The removable space*

$$d_{n-m+1}^{m-1}[\mathbf{f}_1^{0|m-1}] Z_{n-m+1}^{m-1}[\mathbf{f}_1^{0|m-1}] / Z_{n+1}^{m-1}[\mathbf{f}_1^{0|m-1}],$$

is invariant under the formal group action.

PROOF 4. *One finds*

$$d^{m-1}_{n-m+1}[\exp(\nabla_{\mathfrak{t}^1_0})\mathbf{f}^{0|m-1}_1]Z^{m-1}_{n-m+1}[\exp(\nabla_{\mathfrak{t}^1_0})\mathbf{f}^{0|m-1}_1] \subset$$
$$d^{m-1}_{n-m+1}[\mathbf{f}^{0|m-1}_1]Z^{m-1}_{n-m+1}[\mathbf{f}^{0|m-1}_1] + Z^{m-1}_{n+1}[\mathbf{f}^{0|m-1}_1].$$

Indeed, let $x = d^{m-1}_{n-m+1}[\exp(\nabla_{\mathfrak{t}^1_0})\mathbf{f}^{0|m-1}_1]y$ *with* $y \in Z^{m-1}_{n-m+1}[\exp(\nabla_{\mathfrak{t}^1_0})\mathbf{f}^{0|m-1}_1]$. *Then, by Lemma 2,* $y = \exp(\nabla_{\mathfrak{t}^1_0})z$ *with* $z \in Z^{m-1}_{n-m+1}[\mathbf{f}^{0|m-1}_1]$. *We have*

$$\begin{aligned} x &= d^{m-1}_{n-m+1}[\exp(\nabla_{\mathfrak{t}^1_0})\mathbf{f}^{0|m-1}_1]\exp(\nabla_{\mathfrak{t}^1_0})z = \\ &= \exp(\nabla_{\mathfrak{t}^1_0})d^{m-1}_{n-m+1}[\mathbf{f}^{0|m-1}_1]z \\ &\in d^{m-1}_{n-m+1}[\mathbf{f}^{0|m-1}_1]Z^{m-1}_{n-m+1}[\mathbf{f}^{0|m-1}_1] + Z^{m-1}_{n+1}[\mathbf{f}^{0|m-1}_1]. \end{aligned}$$

Dividing out $Z^{m-1}_{n+1}[\exp(\nabla_{\mathfrak{t}^1_0})\mathbf{f}^{0|m-1}_1]$ *has the same effect (since we started with coboundaries) of dividing out* $Z^{m-1}_{n+1}[\mathbf{f}^{0|m-1}_1]$ *and we see that we have completed the proof.*

During the construction of the spectral sequence we will put some requirements on the form of the $\mathbf{f}^i_1$. The construction itself is independent on whether $\mathbf{f}^i_1$ is in normal form, but of course the whole thing only makes sense when it is, since we identify the process of normalization with the dividing out of subspaces in the following construction, since it is the normal form that appears in the computation during the exponentiation.

Define, with d^{m-1}_{n-m+1} determined as before by the mth order normal form, let $\mathbf{E}^0_n = Z^0_n/Z^0_{n+1} = K^n/K^{n+1}$ and

$$\mathbf{E}^m_n = Z^m_n/(d^{m-1}_{n-m+1}Z^{m-1}_{n-m+1} + Z^{m-1}_{n+1}), \quad n \geq m-1 \geq 0.$$

Observe that our previous results show that if $x \in \mathbf{E}^m_n$ is equal to zero, then $\exp(\nabla_{\mathfrak{t}^1_0})x$ is also zero in $\mathbf{E}^m_n$, so $\mathbf{E}^m_n[\exp(\nabla_{\mathfrak{t}^1_0})\mathbf{f}^{0|m-1}_1] = \exp(\nabla_{\mathfrak{t}^1_0})\mathbf{E}^m_n[\mathbf{f}^{0|m-1}_1] = \mathbf{E}^m_n[\mathbf{f}^{0|m-1}_1]$ and the definition is invariant (thanks to the fact that $\mathbf{E}^m_n$ is graded) under the formal group action. We have the following result.

THEOREM 1. *Given any* $\mathfrak{f}^0_1 \in \mathfrak{F}^0_1$, *one can define its* m*th order normal form* $\mathbf{f}^{0|m}_1$ *and a coboundary operator* $d^m_n(\mathfrak{g}^n_1, \mathfrak{h}^n_0) = \mathsf{d}^m\mathbf{f}^{0|m}_1((\mathfrak{g}^n_1, \mathfrak{h}^n_0)) = -\nabla_{(\mathfrak{g}^n_1,\mathfrak{h}^n_0)}\mathbf{f}^{0|m}_1$. *There exists a coboundary map* $\mathbf{d}^m_n : \mathbf{E}^m_n \to \mathbf{E}^m_{n+m}$ *induced by* d^m_n, *but independent of the choice of style of* m*th order normal form. One has that* $H^n(\mathbf{E}^m) \simeq \mathbf{E}^{m+1}_n$ *canonically. We consider* $\mathbf{f}^m_1$ *as an element in* $\mathbf{E}^m_m, m \geq 0$.

We prove this theorem by stating and proving some simpler facts.

PROPOSITION 3. *The d_n^m are* **stable**, *that is, if $x \in K^n$ then $d_n^m x - d_n^{m-1} x \in K^{n+m}$.*

We could have made our life easier by defining d_n^m in terms of $\mathfrak{f}_1^0$ instead of its normal form jet, since they are equivalent anyway and then we can forget about the m dependence. The problem with this approach is that the definition of normal form (Definition 3) is inherently based on jets, and it says something about how the higher order terms behave with respect to the lower order jet (otherwise one might end up checking that indeed the vectorfield commutes with itself in some cases). So in any case one could only have a reasonable definition using $\mathfrak{f}_1^{0|m}$ which makes the proposition necessary.

PROOF 5. *The difference of the normal form of $\mathfrak{f}_1^0$ to order m and to order $m-1$ is an element of $\mathfrak{F}_1^m$. Then*

$$d_n^m x - d_n^{m-1} x \in K^{n+m}.$$

Remark that therefore it sits in Z_{n+m}^{m-1}, since applying d_n^{m-1} will kill it.

We now formulate two lemmas to show that the spectral sequence $\mathbf{E}_n^m$ is well defined.

LEMMA 4. $Z_{n+1}^{m-1} \subset Z_n^m$.

PROOF 6. *Let $x \in Z_{n+1}^{m-1}$, then $x \in K^{n+1} \subset K^n$ and $d_n^{m-2} x = d_{n+1}^{m-2} x \in K^{n+m}$. Therefore $d_n^{m-1} x = d_n^{m-2} x + K^{n+m-1}$ and this implies that $d_n^{m-1} x \in K^{n+m}$, that is, $x \in Z_n^m$.*

LEMMA 5. $d_{n-m+1}^{m-1} Z_{n-m+1}^{m-1} \subset Z_n^m$.

PROOF 7. *If $x \in d_{n-m+1}^{m-1} Z_{n-m+1}^{m-1}$ then there exists a $y \in Z_{n-m+1}^{m-1}$ such that $x = d_{n-m+1}^{m-1} y$. This implies $y \in K^{n-m+1}$ and $d_{n-m+1}^{m-2} y \in K^n$. It follows that $x \in K^n$, and, since $d_n^{m-1} x = 0 \in K^{n+m}$, $x \in Z_n^m$.*

LEMMA 6. *The coboundary operator d_n^m induces a unique (up to coordinate transformations) coboundary operator $\mathbf{d}_n^m : \mathbf{E}_n^m \to \mathbf{E}_{n+m}^m$.*

PROOF 8. *d_n^m maps Z_n^m into Z_{n+m}^m (by definition of Z_n^m and the fact that it is a coboundary operator). Furthermore, it maps $d_{n-m+1}^{m-1} Z_{n-m+1}^{m-1} + Z_{n+1}^{m-1}$ into $d_{n+1}^m Z_{n+1}^{m-1}$. Let $x \in d_{n+1}^m Z_{n+1}^{m-1}$. Then there is a $y \in Z_{n+1}^{m-1}$ such that $x = d_{n+1}^m y$. It follows from proposition 3 that $d_{n+1}^m y \in d_{n+1}^{m-1} Z_{n+1}^{m-1} + Z_{n+m+1}^{m-1}$. Since*

$$\mathbf{E}_{n+m}^m = Z_{n+m}^m / (d_{n+1}^{m-1} Z_{n+1}^{m-1} + Z_{n+m+1}^{m-1}),$$

it follows that d_n^m *induces a coboundary operator* $\mathbf{d}_n^m : \mathbf{E}_n^m \to \mathbf{E}_{n+m}^m$.

Now for the uniqueness: Consider $d_n^m[\tilde{\mathbf{f}}_1^{0|m}]$ *and* $d_n^m[\mathbf{f}_1^{0|m}]$. *Since* $\bar{\mathfrak{f}}_1 = \exp(\nabla_{\bar{\mathfrak{t}}_0^1})\mathfrak{f}_1$ *and* $\tilde{\mathfrak{f}}_1 = \exp(\nabla_{\tilde{\mathfrak{t}}_0^1})\mathfrak{f}_1$, *there exists, thanks to the Campbell-Baker-Hausdorff theory, cf. section 9, a* $\mathfrak{t}_0^1$ *such that* $\tilde{\mathfrak{f}}_1 = \exp(\nabla_{\mathfrak{t}_0^1})\bar{\mathfrak{f}}_1$, *and therefore one finds, using the proof of Lemma 2*

$$d_n^m[\tilde{\mathbf{f}}_1^{0|m}]\exp(\nabla_{\mathfrak{t}_0^1})(\mathfrak{g}_1, \mathfrak{g}_0) = exp(\nabla_{\mathfrak{t}_0^1})d_n^m[\mathbf{f}_1^{0|m}](\mathfrak{g}_1, \mathfrak{g}_0)$$

showing that $d_n^m[\exp(\nabla_{\mathfrak{t}_0^1})\mathbf{f}_1^{0|m}]\exp(\nabla_{\mathfrak{t}_0^1}) = \exp(\nabla_{\mathfrak{t}_0^1})d_n^m[\mathbf{f}_1^{0|m}]$, *that is,* $\exp(\nabla_{\mathfrak{t}_0^1})$ *is a morphism of the* $d_n^m[\mathbf{f}_1^{0|m}]$*-complex to the* $d_n^m[\exp(\nabla_{\mathfrak{t}_0^1})\mathbf{f}_1^{0|m}]$*-complex. Since* $\exp(\nabla_{\mathfrak{t}_0^1})$ *is invertible, the induced map* $\exp(\nabla_{\mathfrak{t}_0^1})_\star$ *defines an isomorphism of the cohomology groups. Using the result in Lemma 7 allows us to draw the final conclusion.*

LEMMA 7. *There exists on the bigraded module* $\mathbf{E}_n^m$ *a differential* $\mathbf{d}_n^m$ *such that* $H^n(\mathbf{E}^m)$ *is canonically isomorphic to* $\mathbf{E}_n^{m+1}$, $m \geq 0$.

PROOF 9. *We follow (*Godement, 1958*) with modifications to allow for the stable boundary operators. For* $x \in Z_n^m$ *to define a cocycle of degree* n *on* $\mathbf{E}_n^m$ *it is necessary and sufficient that* $d_n^m x \in d_{n+1}^{m-1}Z_{n+1}^{m-1} + Z_{n+m+1}^{m-1}$, *i.e.* $d_n^m x = d_{n+1}^{m-1}y + z$ *with* $y \in Z_{n+1}^{m-1}$ *and* $z \in Z_{n+m+1}^{m-1}$. *Putting* $u = x - y \in Z_n^m + Z_{n+1}^{m-1} \subset K^n$, *with* $d_n^m u = d_{n+1}^{m-1}y - d_{n+1}^m y + z \in K^{n+m+1}$, *one has* $u \in Z_n^{m+1}$. *In other words,* $x = y + u \in Z_{n+1}^{m-1} + Z_n^{m+1}$. *It follows that the* n*-cocycles are given by*

$$Z^n(\mathbf{E}_n^m) = (Z_n^{m+1} + Z_{n+1}^{m-1})/(d_{n-m+1}^{m-1}Z_{n-m+1}^{m-1} + Z_{n+1}^{m-1}). \tag{1}$$

The space of n*-coboundaries* $B^n(\mathbf{E}_n^m)$ *consists of elements of* $d_{n-m}^m Z_{n-m}^m$ *and one has*

$$B^n(\mathbf{E}_n^m) = (d_{n-m}^m Z_{n-m}^m + Z_{n+1}^{m-1})/(d_{n-m+1}^{m-1}Z_{n-m+1}^{m-1} + Z_{n+1}^{m-1}). \tag{2}$$

It follows, using Noether's isomorphy $U/(W + U \cap V) \simeq (U + V)/(W + V)$ *for submodules* $W \subset U$ *and* V *and* $(M/V)/(U/V) = M/U$, *that*

$$\begin{aligned} H^n(\mathbf{E}^m) &= (Z_n^{m+1} + Z_{n+1}^{m-1})/(d_{n-m}^m Z_{n-m}^m + Z_{n+1}^{m-1}) \\ &= Z_n^{m+1}/(d_{n-m}^m Z_{n-m}^m + Z_n^{m+1} \cap Z_{n+1}^{m-1}), \end{aligned} \tag{3}$$

since $d_{n-m}^m Z_{n-m}^m \subset Z_n^{m+1}$. *We now first prove that* $Z_n^{m+1} \cap Z_{n+1}^{m-1} = Z_{n+1}^m$. *Let* $x \in Z_n^{m+1} \cap Z_{n+1}^{m-1}$. *Then* $x \in K^{n+1}$ *and* $d_{n+1}^m x \in K^{n+m+1}$. *This implies* $x \in Z_{n+1}^m$.

On the other hand, if $x \in Z_{n+1}^m$ *we have* $x \in K^{n+1} \subset K^n$ *and* $d_{n+1}^{m-1}x \in K^{n+m+1} \subset K^{n+m}$. *Thus* $x \in K^n$ *and* $d_{n+1}^{m-1}x \in K^{n+m+1}$.

Again it follows that $d^m_{n+1}x \in K^{n+m+1}$*, implying that* $x \in Z^{m+1}_n$*. Furthermore* $x \in K^{n+1}$*,* $d^{m-1}_{n+1}x \in K^{n+m}$*, implying that* $d^{m-1}_{n+1}x \in K^{n+m+1}$ *from which we conclude that* $x \in Z^{m-1}_{n+1}$*.*

It follows that

$$H^n(\mathbf{E}^m) \;=\; Z^{m+1}_n/(d^m_{n-m}Z^m_{n-m} + Z^m_{n+1}) = \mathbf{E}^{m+1}_n. \tag{4}$$

In this way we translate normal form problems into cohomology.

This completes the proof of Theorem 1. Apart from the unicity issue, the results so far where first published in (Sanders, 2003). The unicity argument shows that the definition of the Hilbert-Poincaré series of the spectral sequence as it is given there is only dependent on f, the element to be put into normal form, itself and not on the choice of normal form (as it should be, of course).

8. Linear convergence, using the Newton method

We now show how to actually go about computing the normal form, once we can do the linear algebra correctly. We show how convergence in the filtration topology can be obtained by using Newton's method once the normal form stabilizes.

PROPOSITION 4. *Let* $y \in \mathfrak{F}^1_0$ *and* $x \in \mathfrak{F}^p_0, p \geq 1$*. Then we have the following equality modulo terms that contain more than one* $\rho(x)$*:*

$$e^{\rho(y+x)} - e^{\rho(y)} \simeq \rho(\frac{1-e^{-\mathrm{ad}(y)}}{\mathrm{ad}(y)}x)e^{\rho(y)}.$$

In the sequel we construct a sequence μ_m, starting with $\mu_0 = 0$. These μ_m indicate the accuracy to which we have a stable normal form $\mathbf{f}^{0|\mu_m}_1$. Obviously, this depends on the choice of f^0_1, but we do not express this in our notation.

In this section we want to consider the linear problem, that is, we want to consider an equation of the form

$$d^{\mu_{m-1}}\mathrm{h}^{1|\mu_m}_0 = \mathrm{f}^{\mu_{m-1}+1|\mu_m}_1 - \bar{\mathrm{f}}^{\mu_{m-1}+1|\mu_m}_1 \mathrm{mod}\mathfrak{F}^{\mu_m+1}_1,$$

where μ_m is to be determined. To this end we determine for $j = 1, 2, \cdots$ the lowest number r_j such that $\mathcal{T}^{\mu_{m-1}+j-r_j}_{r_j} \neq 0$ and we continue to do this as long as $j < r_j + 1$ and $j < \mu_{m-1}$. Then we put $\mu_m = \mu_{m-1} + j$ for the maximal j to satisfy the condition. Since $r_j \geq 1$, we have that $j \geq 1$, so the μ_m-sequence is strictly increasing. We have

$\mu_{m-1}+1<\mu_m+1\leq 2(\mu_{m-1}+1)$. This implies that $\mu_1=1$ and μ_2 equals 2 or 3. If $\mu_m=\mu_{m-1}+1$ we speak of linear convergence, when $\mu_m+1=2(\mu_{m-1}+1)$ of quadratic convergence at step m. The choice of this sequence is determined by our wish to make the computation of the normal form a completely linear problem. One can see the motivation by considering terms of the type $\rho(\mathrm{h}_0^{1|\mu_m})\mathfrak{f}_1^{\mu_{m-1}+1|\mu_m}$. These should not interfere with the linear computation, that is, they should be of order μ_m+1 at least.

Suppose now that we have for $m>1$, $\mathfrak{f}_1^0=\mathfrak{f}_1^{0|\mu_m}+\mathfrak{f}_1^{\mu_m+1}$. Let $\mathrm{h}_0^{1|\mu_m}$ be the transformation that brings $\mathfrak{f}_1^0$ in μ_{m-1}th order normal form with respect to $\delta^{\mu_{m-2}}\bar{\mathbf{f}}_1^{0|\mu_{m-2}}$ up to order μ_{m-1}, that is

$$\begin{aligned}\exp(\rho(\mathrm{h}_0^{1|\mu_m}))(\mathfrak{f}_1^{0|\mu_m}+\mathfrak{f}_1^{\mu_m+1})=\bar{\mathfrak{f}}_1^{0|\mu_m}+\bar{\mathfrak{f}}_1^{\mu_m+1},\\ \bar{\mathfrak{f}}_1^{0|\mu_{m-1}}-\bar{\mathbf{f}}_1^{0|\mu_{m-2}}\in\ker\delta^{\mu_{m-2}}\bar{\mathbf{f}}_1^{0|\mu_{m-2}}.\end{aligned}$$

We now construct $\mathrm{k}_0^{1|\mu_m}$ such that

$$\begin{aligned}\exp(\rho(\mathrm{h}_0^{1|\mu_m}+\mathrm{k}_0^{1|\mu_m}))(\mathfrak{f}_1^{0|\mu_m}+\mathfrak{f}_1^{\mu_m+1})=\hat{\mathfrak{f}}_1^{0|\mu_m}+\hat{\mathfrak{f}}_1^{\mu_m+1},\\ \hat{\mathfrak{f}}_1^{0|\mu_m}-\bar{\mathbf{f}}_1^{0|\mu_{m-1}}\in\ker\delta^{\mu_{m-1}}\bar{\mathbf{f}}_1^{0|\mu_{m-1}}.\end{aligned}$$

We compute (modulo $\mathfrak{F}_1^{\mu_m+1}$, which we indicate by $\equiv$):

$$\begin{aligned}&\exp(\rho(\mathrm{h}_0^{1|\mu_m}+\mathrm{k}_0^{1|\mu_m}))(\mathfrak{f}_1^{0|\mu_m}+\mathfrak{f}_1^{\mu_m+1})\equiv\\ &\equiv\ \exp(\rho(\mathrm{h}_0^{1|\mu_m}))\mathfrak{f}_1^{0|\mu_m}+\rho(\frac{e^{\mathrm{ad}(\mathrm{h}_0^{1|\mu_m})}-1}{\mathrm{ad}(\mathrm{h}_0^{1|\mu_m})}\mathrm{k}_0^{1|\mu_m})e^{\rho(\mathrm{h}_0^{1|\mu_m})}\mathfrak{f}_1^{0|\mu_m}\\ &\equiv\ \bar{\mathfrak{f}}_1^{0|\mu_m}+\rho(\frac{e^{\mathrm{ad}(\mathrm{h}_0^{1|\mu_m})}-1}{\mathrm{ad}(\mathrm{h}_0^{1|\mu_m})}\mathrm{k}_0^{1|\mu_m})\bar{\mathbf{f}}_1^{0|\mu_{m-1}}\end{aligned}$$

Let now $\mathrm{k}_0^{1|\mu_m}=\frac{\mathrm{ad}(\mathrm{h}_0^{1|\mu_m})}{e^{\mathrm{ad}(\mathrm{h}_0^{1|\mu_m})}-1}\mathrm{g}_0^{1|\mu_m}$ and solve the normal form equation for $\mathrm{g}_0^{1|\mu_m}\in\bigoplus_{j=1}^{\mu_m}\mathcal{T}_{r_j}^{\mu_{m-1}+j-r_j}$ as follows.

$$\begin{aligned}&\exp(\rho(\mathrm{h}_0^{1|\mu_m}+\mathrm{k}_0^{1|\mu_m}))\mathfrak{f}_1^0\equiv\\ &\equiv\ \bar{\mathfrak{f}}_1^{0|\mu_m}-d^{\mu_{m-1}}\mathrm{g}_0^{1|\mu_m}\\ &\equiv\ \hat{\mathfrak{f}}_1^{0|\mu_m}\end{aligned}$$

and we solve $\mathrm{g}_0^{1|\mu_m}$ from the relation $\bar{\mathfrak{f}}_1^{\mu_m}+\rho(\mathrm{g}_0^{1|\mu_m})\bar{\mathbf{f}}_1^{0|\mu_{m-1}}\in\ker\delta^{\mu_{m-1}}\bar{\mathbf{f}}_1^{0|\mu_{m-1}}$. Notice that the term $\rho(\mathrm{g}_0^{1|\mu_m})\bar{\mathbf{f}}_1^{0|\mu_{m-1}}$ may contribute to the normal form $\hat{\mathbf{f}}_1^{\mu_m}$, but $\hat{\mathbf{f}}_1^{\mu_{m-1}}=\bar{\mathbf{f}}_1^{0|\mu_{m-1}}$, since the allowable transformations cannot change the lower order terms of the normal form by definition.

From $g_0^{1|\mu_m}$ we compute $k_0^{1|\mu_m}$ and we are done. After exponentiation we can repeat the whole procedure with m increased by 1. It follows from the relation $\mu_{m-1} < \mu_m$ that we make progress this way, but it may be only one order of accuracy at each step, with $\mu_m = \mu_{m-1} + 1$.

REMARK 8. *So far we have not included the filtering of our local ring R in our considerations. There seem to be two ways of doing that.*

The first way to look at this is the following: we build a sieve, which filters out those terms that can be removed by normal form computations computing $\mathrm{mod}\,\mathfrak{m}^i\mathfrak{F}_1^0$ *starting with $i = 1$. We then increase i by one, and repeat the procedure on what is left. Remark that our transformations have their coefficients in $\mathfrak{m}R$, not in $\mathfrak{m}^iR$, in the same spirit of higher order normal form as we have seen in general. This way, taking the limit for $i \to \infty$, we compute the truly unique normal form description of a certain class of vectorfields. Of course, in the description of this process one has to make i an index for the spectral sequence that is being constructed. There seems to be no problem in writing this all out explicitly, but I have not done so in order to avoid unnecessary complications in the main text, but it might make a good exercise to do so for the reader.*

The second way is to localize with respect to certain divisors. For instance, if δ is some small parameter (maybe a detuning parameter), that is to say, $\delta \in \mathfrak{m}$, then one can encounter terms like $1 - \delta$ in the computation (we are not computing $\mathrm{mod}\,\mathfrak{m}\mathfrak{F}_1^0$ *here!). This may force one to divide through $1 - \delta$ and in doing so repeatedly, one may run into convergence problems, since the zeros of the divisors may approach zero when the order of the computation goes to infinity. Since this is very difficult to realize in practice, this small divisor problem is a theoretical problem for the time being, which may ruin however the asymptotic validity of the intermediate results if we want to think of them as approximations of reality.*

In general, at each step we can define the rate of progress as the number $\alpha_m \in \mathbb{Q}, m \geq 2$ satisfying $\mu_m = \alpha_m \mu_{m-1} + 1$. One has $1 \leq \alpha_m \leq 2$.

Ideally, one can double the accuracy at each step in the normalization process which consists of solving a linear problem and computing an exponential at each step. Thus we can (ideally) normalize the $2\mu_{m-1} + 1$-jet $f_1^{\mu_{m-1}+1|2\mu_{m-1}+1}$. We proved we could normalize the μ_m-jet $f_1^{\mu_{m-1}|\mu_m}$. We therefore call $\Delta_m = 2\mu_{m-1} - \mu_m + 1 \overset{m\geq 2}{=} (2-\alpha_m)\mu_{m-1}$ the m-**defect**. If $\Delta_m \leq 0$, the obvious strategy is normalize up to $2\mu_{m-1} + 1$. Sooner or later we will either have a positive defect, or we are done normalizing, because we reached our intended order of accuracy. In the next section we discuss what to do in case of positive defect, if one still wants quadratic convergence.

THEOREM 2. *The transformation connecting $\mathfrak{f}_1^0 \in \mathfrak{F}_1^0$ with its normal form with coefficients in the residue field can be computed at a linear rate at least and at a quadratic rate at theoretical optimum.*

REMARK 9. *If the $\mathfrak{f}_1^0$ has an* **infinitesimal symmetry**, *that is, a $\mathfrak{s}_0^0 \in \mathfrak{F}_0^0$ (extending the transformation space to allow for linear group actions) then one can restrict one's attention to* $\ker \nabla_{\mathfrak{s}_0^0}$ *to set the whole thing up, so that the normal form will preserve the symmetry, since $\nabla_{\mathfrak{s}_0^0}\nabla_{\mathfrak{t}_0^1}\mathfrak{f}_1^0 = \nabla_{\mathfrak{t}_0^1}\nabla_{\mathfrak{s}_0^0}\mathfrak{f}_1^0 + \nabla_{\nabla_{\mathfrak{s}_0^0}\mathfrak{t}_0^1}\mathfrak{f}_1^0$. If one has two of these symmetries $\mathfrak{s}_0^0, \mathfrak{q}_0^0$, then $\nabla_{\mathfrak{s}_0^0}\mathfrak{q}_0^0$ is again a symmetry, that is $\nabla_{\nabla_{\mathfrak{s}_0^0}\mathfrak{q}_0^0}\mathfrak{f}_1^0 = 0$ so the set of all symmetries forms again a Leibniz algebra. By the way, it is not a good idea to do this for every symmetry of the original vectorfield (why not?).*

REMARK 10. *While we allow for the existence of a nonzero linear part of the vectorfield $\mathfrak{f}_1^0$, we do not require it: the whole theory covers the computation of normal forms of vectorfields with zero linear part.*

COROLLARY 4. *If for some m the representation ρ as induced on the graded Leibniz algebra $\mathbf{E}_{\cdot}^m$ becomes trivial (either for lack of transformations or because the possible transformations cannot change the normal form anymore), then $\mathbf{f}_1^{0|m}$ is the* **unique normal form**, *unique in the sense that if it is the normal form of some $\mathfrak{g}_1^0$ then $\mathfrak{f}_1^0 \equiv \mathfrak{g}_1^0$.*

9. Quadratic convergence, using the Dynkin formula

As we have seen in the last section, one can in the worst scenario only get convergence at a linear rate using the Newton method. In order to obtain quadratic convergence, we now allow for extra exponential computations within the linear step, hoping that these are less expensive since they are done with *small* transformations. To this end we now introduce the Dynkin formula, which generalizes the results from the last section.

LEMMA 8. *Let $\bar{\mathfrak{f}}_1^0 = exp(\rho(\bar{\mathfrak{h}}_0^1))\mathfrak{f}_1^0$ and $\hat{\mathfrak{f}}_1^0 = exp(\rho(\hat{\mathfrak{h}}_0^k))\bar{\mathfrak{f}}_1^0$, with $\bar{\mathfrak{f}}_1^0, \mathfrak{f}_1^0, \hat{\mathfrak{f}}_1^0 \in \mathfrak{F}_1^0, \hat{\mathfrak{h}}_0^k \in \mathfrak{F}_0^k, k \geq 1$ and $\bar{\mathfrak{h}}_0^1 \in \mathfrak{F}_0^1$. Then $\hat{\mathfrak{f}}_1^0 = exp(\rho(\mathfrak{h}_0^1))\mathfrak{f}_1^0$, where $\mathfrak{h}_0^1$ is given by*

$$\mathfrak{h}_0^1 = \bar{\mathfrak{h}}_0^1 + \int_0^1 \psi[\exp(\mathrm{ad}(\epsilon\ \hat{\mathfrak{h}}_0^k))\exp(\mathrm{ad}(\bar{\mathfrak{h}}_0^1))]\hat{\mathfrak{h}}_0^k \mathrm{d}\epsilon,$$

where $\psi(z) = log(z)/(z-1)$.

PROOF 10. *This is the right-invariant formulation, which is more convenient in our context, where we think of* $\hat{\mathfrak{h}}_0^k$ *as a perturbation of* $\bar{\mathfrak{h}}_0^1$*. A proof of the left-invariant formulation can be found in (*Hall, 2003*). Observe that in the filtration topology all the convergence issues become trivial, so one is left with checking the formal part of the proof, which is fairly easy. The idea is to consider* $Z(\epsilon) = \exp(\mathrm{ad}(\epsilon\ \hat{\mathfrak{h}}_0^k))\exp(\mathrm{ad}(\bar{\mathfrak{h}}_0^1))$*. Then* $\frac{dZ}{d\epsilon}Z^{-1}(\epsilon) = \mathrm{ad}(\hat{\mathfrak{h}}_0^k)$*, and the left hand side is invariant under right-multiplication of* $Z(\epsilon)$ *with some* ϵ*-independent invertible operator. One then proceeds to solve this differential equation.*

Since the first powers of two are the consecutive numbers $2^0, 2^1$, we can always start our calculation with quadratic convergence. Suppose now for some m, with $\mu_{m-1} = 2^p - 1$, we find $\Delta_m > 0$. So we have $\mu_m = 2^{p+1} - 1 - \Delta_m$ and

$$\bar{\mathfrak{f}}_1^0 = \exp(\rho(\bar{\mathfrak{h}}_0^1))\mathfrak{f}_1^0$$

Consider now $\bar{\mathfrak{f}}_1^{0|2\mu_{m-1}+1}$ as the vectorfield to be normalized up to order $2\mu_{m-1}+1$. In the next step, until we apply the Dynkin formula, we compute $\mathrm{mod}\mathfrak{F}_l^{2(\mu_{m-1}+1)}$.

We use the method from the last section to put $\bar{\mathfrak{f}}_1^{0|\mu_m}$ into μ_m order normal form and compute the induced vectorfield. Then we compute Δ_{m+1} and repeat the procedure until we have $\hat{\mathbf{f}}_1^{2^{p+1}-1}$ and the transformation $\hat{\mathfrak{h}}_0^k$ connecting $\bar{\mathfrak{f}}_1^0$ with the vectorfield in $2^{p+1}-1$ order normal form $\hat{\mathfrak{f}}_1^0$ by $\hat{\mathfrak{f}}_1^0 = \exp(\rho(\hat{\mathfrak{h}}_0^k))\bar{\mathfrak{f}}_1^0$.

Then we apply the Dynkin formula and continue our procedure with increased m, until we are done.

With all the intermediate exponentiations, one can not really call this quadratic convergence. Maybe one should use the term pseudo-quadratic convergence for this procedure. It remains to be seen in practice which method functions best. One may guess that the advantages of the method sketched here will only show at high order calculations. This has to be weighted against the cost of implementing the Dynkin formula. The Newton method is easy to implement, since it just involves a variation of exponentiation, and certainly better than just doing things term by term and exponentiating until everything is right. One should also keep in mind that the Newton method keeps on trying to double its accuracy: one may be better off with a sequence $1, 2, 3, 6$ than with $1, 2, 3, 4, 6$. The optimum may depend on the desired accuracy. In principle one could try and develop measures to decide these issues, but that does not seem to be a very attractive course. Computer algebra computations depend on many factors and it will not be easy to get a realistic cost estimate. If one can just assign some costs to the several actions, this will at best lead to upper estimates, but how to show that the best estimated method indeed gives the best actual performance? A more realistic approach is just to experiment with the alternatives, until one gets a good feel for their properties.

10. The Hamiltonian 1 : 2 resonance

In this section we analyze the spectral sequence of the Hamiltonian 1 : 2 resonance. This problem was considered in (Sanders and van der Meer, 1992), but this paper contains numerous typographical errors, which we hope to repair here. We work in $T^*\mathbb{R}^2$, with coordinates x_1, x_2, y_1, y_2. A Poisson structure is given, with basic bracket relations:

$$\{x_i, y_i\} = 1, \{x_i, x_j\} = \{y_i, y_j\} = 0, \quad i, j = 1, 2.$$

Hamiltonians are linear combinations of terms $x_1^k x_2^l y_1^m y_2^n$ and we put a grading deg on these terms by

$$\deg(x_1^k x_2^l y_1^m y_2^n) = k + l + m + n - 2.$$

One verifies that $\deg(\{f, g\}) = \deg(f) + \deg(g)$. The grading induces a filtering, and the *linear fields* consist of quadratic Hamiltonians. In our case, the quadratic Hamiltonian to be considered is

$$H_0^{\pm} = \tfrac{1}{2}(x_1^2 + y_1^2) \pm (x_2^2 + y_2^2).$$

We'll restrict our attention here to the H_0^+ for the sake of simplicity. The computation of $\mathbf{E}^1$ is standard. We have to determine $\ker \mathrm{ad}(H_0^+)$, and we find that it is equal to the direct sum of two copies of

$$\mathbb{R}[[B_1, B_2, R_1]] \oplus R_2\mathbb{R}[[B_1, B_2, R_1]],$$

where $B_1 = H_0^+$, $B_2 = H_0^-$ and

$$\begin{aligned} R_1 &= x_2(x_1^2 - y_1^2) + 2x_1y_1y_2 \\ R_2 &= 2x_1x_2y_1 - y_2(x_1^2 - y_1^2), \end{aligned}$$

and we have the relation

$$R_1^2 + R_2^2 = \tfrac{1}{2}(B_1 + B_2)^2(B_1 - B_2).$$

The Poisson brackets are (ignoring B_1, since it commutes with everything)

$$\begin{aligned} \{B_2, R_1\} &= -4R_2 \\ \{B_2, R_2\} &= 4R_1 \\ \{R_1, R_2\} &= 3B_2^2 + 2B_1B_2 - B_1^2. \end{aligned}$$

We now suppose that our first order normal form Hamiltonian is $\bar{\mathbf{H}}_1^{0|1} = H_0^+ + \epsilon_1 R_1 + \epsilon_2 R_2$, with $\epsilon = \sqrt{\epsilon_1^2+\epsilon_2^2} \neq 0$. For a complete analysis of this problem, one should also consider the remaining cases, but this has never been attempted, it seems. We now do something that is formally outside the scope of our theory, namely we use a linear transformation in the R_1, R_2-plane, generated by B_2, to transform the Hamiltonian to $\bar{\mathbf{H}}_1^{0|1} = H_0^+ + \epsilon R_1$. One should realize here that the formalism is by no means as general as could be, but since it is already intimidating enough, I have tried to keep it simple. The reader may want to go through the whole theory again to expand it to include this simple linear transformation properly. One should remark that it involves a change of topology, since convergence in the filtration topology will no longer suffice.

Having done this, we now have to determine the image of $\mathrm{ad}(R_1$. One finds

$$\begin{aligned}
\mathrm{ad}(\bar{\mathbf{H}}_1^{0|1}) B_2^n R_1^k &= 4n B_2^{n-1} R_1^k R_2 \\
\mathrm{ad}(\bar{\mathbf{H}}_1^{0|1}) B_2^n R_1^k R_2 &= B_2^{n-1} R_1^k (-4n R_1^2 + 2n B_1^3) \\
&+ B_2^n R_1^k ((3-2n) B_2^2 + 2(1-n) B_1 B_2 + (2n-1) B_1^2))
\end{aligned}$$

The first relation allows one to remove all terms in $R_2 \mathbb{R}[[B_1, B_2, R_1]]$, while the second allows one to remove all terms in $B_2^2 \mathbb{R}[[B_1, B_2, R_1]]$, since $2n-3$ is never zero. The spectral sequence $\mathbf{E}^2$ looks now like the direct sum of $\mathbb{R}[[B_1, R_1]] + B_2 \mathbb{R}[[B_1, R_1]]$ and $\mathbb{R}[[B_1, R_1]]$ (the last statement follows from the first relation with $n = 0$). A moment of consideration shows that this is also the final result. It says that the unique normal form is of the form

$$\bar{\mathbf{H}}_1^{0|\infty} = H_0^+ + F_1(B_1, R_1) + B_2 F_2(B_1, R_1),$$

with $\frac{\partial F_1}{\partial R_1}(0,0) = \epsilon \neq 0$ and $F_2(0,0) = 0$. The symmetries of the system are $\mathbb{R}[[B_1, \bar{\mathbf{H}}_1^{0|\infty}]]$.

We have now computed the normal form of the $1:2$-resonance Hamiltonian under the formal group of symplectic transformations. The reader may want to expand the transformation group to include all formal transformations to see what happens, and to compare the result with the normal form given in (Broer et al., 2003, page 55).

11. Spatial averaging

For a theoretical method to be the right method, it needs to work in situations that arise in practice. Let us have a look at equations of the form

$$\begin{aligned}\dot{\theta} &= \Omega_0(x) + \sum_{i=1}^{\infty} \epsilon^i \Omega_i(\theta, x), \quad \theta \in S^1 \\ \dot{x} &= \sum_{i=1}^{\infty} \epsilon^i X_i(\theta, x), \quad x \in \mathbb{R}^n.\end{aligned}$$

This equation has a given filtering in powers of ϵ, and the zeroth order normal form is

$$\begin{aligned}\dot{\theta} &= \Omega_0(x), \quad \theta \in S^1 \\ \dot{x} &= 0, \quad x \in \mathbb{R}^n.\end{aligned}$$

This means that in our calculations on the spectral sequence level we can consider x as an element of the ring, that is the ring will be $C^\infty(\mathbb{R}^n, \mathbb{R})$ and the Lie algebra of periodic vectorfields on $S^1 \times \mathbb{R}^n$ acts on it, but in such a way that the filtering degree increases if we act with the original vectorfield or one of its normal forms, so that we can effectively assume that the x is a constant in the first order normal form calculations. The only thing we need to worry about is that we may have to divide through $\Omega_0(x)$ in the course of our calculations, thereby introducing resonances (Sanders, 1977; Sanders, 1979). This leads to interesting problems, but the formal computation of the normal form is not affected, as long as we stay outside the resonance domain. The first order normal form homological equation is

$$\Omega_0(x)\frac{\partial}{\partial\theta}\begin{pmatrix}\Phi_1 \\ Y_1\end{pmatrix} - \begin{pmatrix}Y_1 \cdot \nabla\Omega_0 \\ 0\end{pmatrix} = \begin{pmatrix}\Omega_1 - \bar{\Omega}_1 \\ X_1 - \bar{X}_1\end{pmatrix},$$

and we can solve this equation by taking (with $\mathrm{d}\phi$ the normalized Haar measure, that is, $\int_{S^1} \mathrm{d}\phi = 1$)

$$\bar{X}_1(x) = \int_{S^1} X_1(\phi, x)\mathrm{d}\phi,$$

and

$$Y_1(\theta, x) = \frac{1}{\Omega_0(x)}\int^{\theta} X_1(\phi, x) - \bar{X}_1 \mathrm{d}\phi.$$

We let $\bar{Y}_1(x) = \int_{S^1} Y_1(\phi, x)\mathrm{d}\phi$ and observe that it is not fixed yet by the previous calculations. We now put

$$\bar{\Omega}_1(x) = \int_{S^1} \Omega_1(\phi, x) + Y_1(\phi, x)\cdot\nabla\Omega_0(x)\mathrm{d}\phi = \int_{S^1} \Omega_1(\phi, x)\mathrm{d}\phi + \bar{Y}_1(x)\cdot\nabla\Omega_0(x),$$

and we observe that if $\nabla\Omega_0(x) \neq 0$ we can take $\bar{Y}_1(x)$ in such a way as to make $\bar{\Omega}_1(x) = 0$. All this indicates that the second order normal form computation will be messy, since there is still a lot of freedom in the choice of $\bar{Y}_1(x)$, and this will have to be carefully used. There do not seem to be any results in the literature on this problem apart from (Sanders and Verhulst, 1985, §6.3). We have the following

THEOREM 3. *Assuming that* $\Omega_0(x) \neq 0$ *and* $\nabla\Omega_0(x) \neq 0$ *one has that* $\mathbf{E}_1^1$ *is the direct sum of the space generated by vectorfields of the form*

$$\begin{aligned} \dot{\theta} &= 0, \quad \theta \in S^1 \\ \dot{x} &= \epsilon \bar{X}_1(x), \quad x \in \mathbb{R}^n, \end{aligned}$$

and transformations $\bar{Y}_1(x)$ *such that*

$$\bar{Y}_1(x) \cdot \nabla\Omega_0(x) = 0.$$

This illustrates that the computation of the spectral sequence is not going to be easy, but also that it mimics the usual normal form analysis exactly.

References

Arnol'd, V. I., S. M. Guseĭn-Zade, and A. N. Varchenko: 1985, *Singularities of differentiable maps. Vol. I.* Boston, MA: Birkhäuser Boston Inc. The classification of critical points, caustics and wave fronts, Translated from the Russian by Ian Porteous and Mark Reynolds.

Baider, A. and J. A. Sanders: 1991, 'Unique normal forms: the nilpotent Hamiltonian case'. *J. Differential Equations* **92**(2), 282–304.

Baider, A. and J. A. Sanders: 1992, 'Further reduction of the Takens-Bogdanov normal form'. *J. Differential Equations* **99**(2), 205–244.

Belitskii, G.: 2002, 'C^∞-normal forms of local vector fields'. *Acta Appl. Math.* **70**(1-3), 23–41. Symmetry and perturbation theory.

Bogdanov, R. I.: 1979, 'Local orbital normal forms of vector fields on the plane'. *Trudy Sem. Petrovsk.* **5**, 51–84.

Bott, R. and L. W. Tu: 1982, *Differential forms in algebraic topology.* New York: Springer-Verlag.

Broer, H., I. Hoveijn, G. Lunter, and G. Vegter: 2003, *Bifurcations in Hamiltonian systems*, Vol. 1806 of *Lecture Notes in Mathematics.* Berlin: Springer-Verlag. Computing singularities by Gröbner bases.

Bronstein, I. U. and A. Y. Kopanskiĭ: 1994, *Smooth invariant manifolds and normal forms*, Vol. 7 of *World Scientific Series on Nonlinear Science. Series A: Monographs and Treatises.* River Edge, NJ: World Scientific Publishing Co. Inc.

Godement, R.: 1958, *Topologie algébrique et théorie des faisceaux.* Paris: Hermann.

Hall, B. C.: 2003, *Lie groups, Lie algebras, and representations*, Vol. 222 of *Graduate Texts in Mathematics.* New York: Springer-Verlag. An elementary introduction.

Loday, J.-L.: 1991, *Cyclic Homology*, Vol. 301 of *Grundlehren der mathematischen Wissenschaften*. Berlin: Springer–Verlag.

Meyer, K. R.: 1994, 'Perturbation analysis of nonlinear systems'. In (Tournier, 1994), pp. 103–139. Papers from the conference (CADE-92) held in June 1992.

Murdock, J.: 2003a, 'Hypernormal form theory: foundations and algorithms'. Technical report, Mathematics Department, Iowa State University.

Murdock, J.: 2003b, *Normal forms and unfoldings for local dynamical systems*, Springer Monographs in Mathematics. New York: Springer-Verlag.

Sanders, J.: 1977, 'Are higher order resonances really interesting?'. *Celestial Mech.* **16**(4), 421–440.

Sanders, J.: 1979, 'On the passage through resonance'. *SIAM J. Math. Anal.* **10**(6), 1220–1243.

Sanders, J. A.: 1994, 'Versal normal form computations and representation theory'. In (Tournier, 1994), pp. 185–210. Papers from the conference (CADE-92) held in June 1992.

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

Sanders, J. A. and J.-C. van der Meer: 1992, 'Unique normal form of the Hamiltonian 1 : 2-resonance'. In: *Geometry and analysis in nonlinear dynamics (Groningen, 1989)*, Vol. 222 of *Pitman Res. Notes Math. Ser.* Harlow: Longman Sci. Tech., pp. 56–69.

Sanders, J. A. and F. Verhulst: 1985, *Averaging methods in nonlinear dynamical systems*, Vol. 59 of *Applied Mathematical Sciences*. New York: Springer-Verlag.

Spivak, M.: 1979, *A comprehensive introduction to differential geometry. Vol. II*. Wilmington, Del.: Publish or Perish Inc., second edition.

Tournier, E. (ed.): 1994, *Computer algebra and differential equations*, Vol. 193 of *London Mathematical Society Lecture Note Series*. Cambridge: Cambridge University Press. Papers from the conference (CADE-92) held in June 1992.

Vermaseren, J.: 1991, *Symbolic Manipulation with FORM, Tutorial and reference manual*. Amsterdam: CAN.

Vermaseren, J.: 2000, 'New features of FORM'. Technical report, Nikhef, Amsterdam. math-ph/0010025.