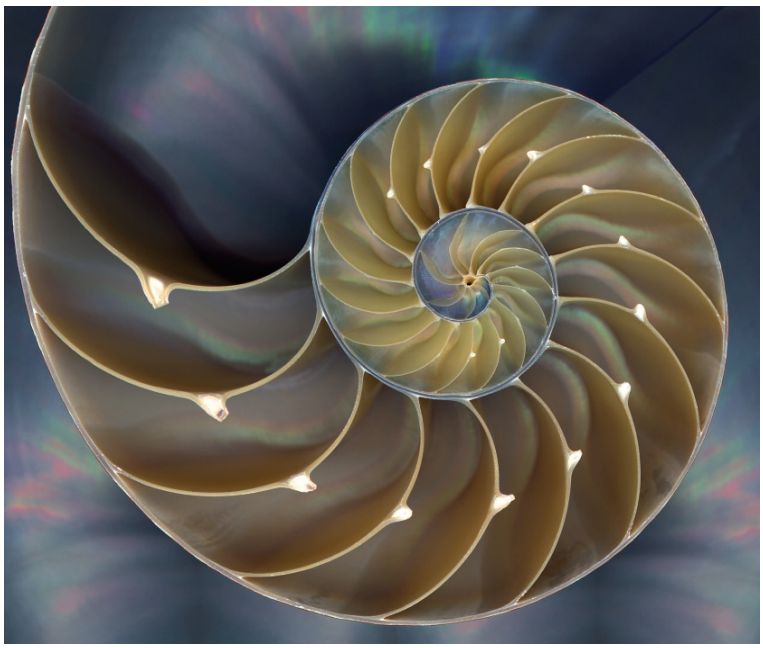
J. Calmet / W.M. Seiler / R.W. Tucker (Eds.)

Global Integrability of Field Theories



Proceedings of GIFT 2006
Cockcroft Institute, Daresbury (UK)
November 1-3, 2006

universitätsverlag karlsruhe

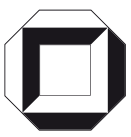
J. Calmet / W.M. Seiler / R.W. Tucker (Eds.)

Global Integrability of Field Theories

Global Integrability of Field Theories

Proceedings of GIFT 2006
Cockcroft Institute, Daresbury (UK),
November 1-3, 2006

Editors:
J. Calmet
W. M. Seiler
R. W. Tucker



universitätsverlag karlsruhe

Acknowledgement:

GIFT (Global Integrability of Field Theories) is a specific targeted research project in the NEST (New and Emerging Science and Technology) activity area within the 6th Framework Programme of the European Union.
Contract no.: 5006 (NEST-Adventure).

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de

Universitätsverlag Karlsruhe 2006
Print on Demand

ISBN 3-86644-035-9

Preface

The GIFT 2006 workshop covers topics related to the Global Integration of Field Theories. These topics span several domains of science including Mathematics, Physics and Computer Science. It is indeed an interdisciplinary event and this feature is well illustrated by the diversity of papers presented at the workshop.

Physics is our main target. A simple approach would be to state that we investigate systems of partial differential equations since it is widely believed that they provide a fair description of our world. The questions whether this world is Einsteinian or not, is described by String Theory or not are not however on our agenda. At this stage we have defined what we mean by field theories. To assess what global integrability means we surf on the two other domains of our interest.

Mathematics delivers the main methodologies and tools to achieve our goal. It is a trivial remark to say that there exist several approaches to investigate the concept of integrability. Only selected ones are to be found in these proceedings. We do not try to define precisely what global integrability means. Instead, we only suggest two tracks. The first one is by analogy with the design of algorithms, in Computer Algebra or Computer Science, to solve systems of differential equations. The case of ODEs is rather well understood since a constructive methodology exists. Although many experts claim that numerous results do exist to solve systems of PDEs, no constructive decision method exists. This is our first track. The second track follows directly since the real world is described by systems of PDEs, which are mainly non-linear ones. To be able to decide in such a case the existence of solutions would immediately increase the scope of new technologies applicable to industrial problems.

It is this latter remark that led to the European NEST project with the same name. The GIFT project aims at making progress in the investigation

of field theories through the use of very advanced mathematical tools. This workshop is part of the deliverables. It demonstrates that at least one goal has been reached: to foster the collaboration of physicists, mathematicians and computer scientists coming from different schools of thoughts.

We are very grateful to the European Commission for its support to GIFT and thus to the use of very advanced mathematical methods to investigate how field theories may generate innovative technologies. Our project officer in Brussels, Dr. Carlos Saraiva-Martins, deserves special thanks for his efficiency and his readiness to answer any of our queries. Special thanks are due to Magdalena Sałek and Marcus Hausdorf for their fully efficient management of the project, this workshop organization and proceedings preparation.

October 2006

Jacques Calmet

Werner M. Seiler

Robin W. Tucker

Organisation

Executive Committee

General Chair	Jacques Calmet
Program Chair	Robin W. Tucker
Organising Committee	Liz Kennedy, Jonathan Gratus
Project Support	Marcus Hausdorf, Magdalena Sałek

Program Committee

Jacques Calmet – Universität Karlsruhe (TH), Germany
Boris Kruglikov – University of Tromsø, Norway
Bernard Malgrange – Université Joseph Fourier, France
Juan J. Morales-Ruiz – Universitat Politécnica de Catalunya, Spain
Christiaan A.M. Peters – Université Joseph Fourier, France
Emmanuel Paul – Université Paul Sabatier, France
Jean-Pierre Ramis – Université Paul Sabatier, France
Julio Rubio – Universidad de La Rioja, Spain
Werner M. Seiler – Universität Kassel, Germany
Jan A. Sanders – Vrije Universiteit Amsterdam, The Netherlands
Francis Sergeraert – Université Joseph Fourier, France
Robin W. Tucker – Lancaster University, United Kingdom

Table of Contents

Symbolic Computation for Overdetermined Systems of Nonlinear Differential Equations

Evelyne Hubert

INRIA Sophia Antipolis
www.inria.fr/cafe/Evelyne.Hubert

The study of differential equations and differential systems has many facets and requires a large palette of tools, whether of computational or theoretical nature. Interaction among all of those is essential. Numerical integration based on analysis has held a prominent place in applications as approximating the solution of differential system is often the final aim. Symbolic computation brings some other tools for the qualitative investigation of differential systems. Symbolic algorithms handle algebraic structures. To study differential systems there is thus a need to *algebraicize* the problems, that is establish dictionaries between the analytic and the algebraic properties. A second stage in this line of research is then to develop algorithms for capturing the algebraic information, make them efficient and accessible through their implementations in symbolic systems.

On one hand symbolic computation can preprocess differential systems which are not given in a form that fits numerical schemes. But some academic and engineering problems of smaller size ask for more qualitative answers: What are the conditions on some parameters for the system to have a solution? What are the dimension of the solution space? What are the consistent set of initial condition that determine uniquely a solution? Can I deduce some unknowns from the knowledge of others? Those sample of questions on on over and under determined differential systems are addressed in what I would call *differential elimination and completion.*

There has been several algebraic theories developed for addressing those issues. They are amenable to symbolic computation. Cartan's introduction of exterior differential systems and involutivity has probably been the most celebrated [5] and is suited for geometric problems. The homological approach to involutivity for differential sytems, initiated by Spencer and Quillen, has shown its implication in mathematical physics, control theory and numerical analysis [4, 18, 20, 15]. The algebraic standpoint on differential equations of Ritt and Kochin [19, 12] started on different grounds. There the focus was on nonlinearity and the implications on the solution sets. A motivational question was: how do we define the general solution of an differential equation and what are its relations to the singular solutions. The subsequent theory of differential algebra has provided firm grounds for this question and has developed in an algebraic treatment of nonlinear differential systems.

My goal is not to expand on the inter-relationships between the algebraic theories for differential systems, though there is a real need for that [14]. Algorithms for systems of nonlinear differential systems based on differential algebra [8] have been implemented in widely available symbolic systems for 10 years now [2] and are still evolving [3, 9, 1, 7]. I want to promote here their use by introducing the underlying theory and giving a practical sense of of their applicability on examples. Indeed, symbolic computation software can be very user friendly but it is delusive to think that they can be used as black boxes. One needs to have basic understanding of the underlying theories to best use the software available.

Differential problems are sometimes better expressed with non commuting derivations. This is the case of some equivalence and classification problems [17]. This is also the case of differential systems with symmetry once they are rewritten in terms of the invariants of the symmetry group [16]. The computations are then more tractable in those terms. We shall base the presentation on the recent generalization of differential algebra to derivations with non trivial commuting rules [10]. This extension furthermore gives theoretical grounds for the investigation of the differ-

ential algebras formed by the differential invariants of a Lie group action [13, 6, 11].

References

[1] F. Boulier. *Réécriture algébrique dans les systèmes d'équations différentielles polynomiales en vue d'applications dans les sciences du vivant*. Habilitation Thesis. Université de Lille, 2006.

[2] F. Boulier and E. Hubert. DIFFALG*: description, help pages and examples of use*. Symbolic Computation Group, University of Waterloo, Ontario, Canada, 1998. `http://www.inria.fr/cafe/Evelyne.Hubert/diffalg`.

[3] François Boulier, François Lemaire, and Marc Moreno Maza. Pardi! In *Proceedings of the 2001 International Symposium on Symbolic and Algebraic Computation*, pages 38–47 (electronic), New York, 2001. ACM.

[4] R. L. Bryant, S. S. Chern, R. B. Gardner, H. L. Goldschmidt, and P. A. Griffiths. *Exterior differential systems*, volume 18 of *Mathematical Sciences Research Institute Publications*. Springer-Verlag, New York, 1991.

[5] E. Cartan. *Les systèmes différentiels extérieurs et leurs applications géométriquesérentiels extérieurs et leurs applications géométriques*. Actualités Sci. Ind., no. 994. Hermann et Cie., Paris, 1945.

[6] M. Fels and P. J. Olver. Moving coframes. II. Regularization and theoretical foundations. *Acta Appl. Math.*, 55(2):127–208, 1999.

[7] O. Golubitsky, M.V. Kondratieva, M Moreno Maza, and A. Ovchinnikov. Bounds for algorithms in differential algebra. 2006.

[8] E. Hubert. Notes on triangular sets and triangulation-decomposition algorithms II: Differential systems. In F. Winkler and U. Langer, editors, *Symbolic and Numerical Scientific Computing*, number 2630 in Lecture Notes in Computer Science, pages 40–87. Springer Verlag Heidelberg, 2003.

[9] E. Hubert. Improvements to a triangulation-decomposition algorithm for ordinary differential systems in higher degree cases. In *ISSAC 2004*. ACM press, 2004. 191-198.

[10] E. Hubert. Differential algebra for derivations with nontrivial commutation rules. *Journal of Pure and Applied Algebra*, 200(1-2):163–190, 2005.

[11] E. Hubert and I. A. Kogan. Rational invariants of a group action. construction and rewriting. *Journal of Symbolic Computation*, 2006. To appear.

[12] E. R. Kolchin. *Differential Algebra and Algebraic Groups*, volume 54 of *Pure and Applied Mathematics*. Academic Press, 1973.

[13] A. Kumpera. Invariants différentiels d'un pseudogroupe de Lie. In *Géométrie différentielle (Colloq., Univ. Santiago de Compostela, Santiago de Compostela,*

1972), pages 121–162. Lecture Notes in Math., Vol. 392. Springer, Berlin, 1974.

[14] B. Malgrange. Differential algebra and differential geometry. In *Lie groups, geometric structures and differential equations—one hundred years after Sophus Lie (Kyoto/Nara, 1999)*, volume 37 of *Adv. Stud. Pure Math.*, pages 197–204. Math. Soc. Japan, Tokyo, 2002.

[15] B. Malgrange. *Systèmes différentiels involutifs*, volume 19 of *Panoramas et Synthèses*. Société Mathématique de France, Paris, 2005.

[16] E. L. Mansfield. Algorithms for symmetric differential systems. *Foundations of Computational Mathematics*, 1(4):335–383, 2001.

[17] S. Neut. *Implantation et nouvelles applications de la méthode d'équivalence de Cartan*. PhD thesis, Université des Sciences et Technologie de Lille, `http://www.lifl.fr/~neut`, 2003.

[18] J.-F. Pommaret. *Partial differential control theory*, volume 530 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 2001. Vol. I. Mathematical tools; Vol. II. Control systems.

[19] J.F. Ritt. *Differential Algebra*, volume XXXIII of *Colloquium publications*. American Mathematical Society, 1950. `http://www.ams.org/online_bks`.

[20] W. Seiler. *Involution - The Formal Theory of Differential Equations and its Applications in Computer Algebra and Numerical Analysis*. Habilatation thesis. Universität Mannheim, 2001.

Dimension of the solutions space of PDEs

Boris Kruglikov Valentin Lychagin

Inst. of Mathematics and Statistics, University of Tromsø,
Tromsø 90-37, Norway,
kruglikov@math.uit.no, lychagin@math.uit.no

Abstract

We discuss the dimensional characterization of the solutions space of a formally integrable system of partial differential equations and provide certain formulas for calculations of these dimensional quantities.

Keywords: Solutions space, Cartan's test, Cohen-Macaulay module, involutive system, compatibility, formal integrability.

1 Introduction: what is the solutions space?

Let $\mathcal{E}$ be a system of partial differential equations (PDEs[1]). We would like to discuss the dimensional characterization of its solutions space.

However it is not agreed upon what should be called a solution. We can choose between global or local and even formal solutions or jet-solutions to a certain order. Hyperbolic systems hint us about shock waves as multiple-valued solutions and elliptic PDEs suggest generalized functions or sections.

A choice of category, i.e. finitely differentiable C^k, smooth C^∞ or analytic C^ω together with many others, plays a crucial role. For instance there are systems of PDEs that have solutions in one category, but lacks them in another (we can name the famous Lewy's example of a formally integrable PDE without smooth or analytic solutions, [L]).

[1] MSC numbers: 35N10, 58A20, 58H10; 35A30.

In this paper we restrict to local or even formal solutions. The reason is lack of reasonable existence and uniqueness theorems (in the case of global solutions even for ODEs). In addition this helps to overcome difficulties with blow-ups and multi-values.

If the category is analytic, then Cartan-Kähler theorem [Ka] guarantees local solutions of formally integrable equations [Go] and even predicts their quantity. We then measure it by certain dimension characteristics.

If the category is smooth, formal integrability yields existence of solutions only if coupled with certain additional conditions (see for instance [Ho]). Thus it is easier in this case to turn to formal solutions, which in regular situations give the same dimension characteristics. With this vague idea let us call the space of solutions $\mathrm{Sol}(\mathcal{E})$.

With this approach it is easy to impose a topology on the solutions space. However we shall encounter the situations, when the topological structure is non-uniform.

To illustrate the above discussion, let's consider some model ODEs (in which case we possess existence and uniqueness theorem). The space of local solutions for the ODE $y' = y^2$ is clearly one-dimensional, but the space of global solutions (continuous pieces until the blow-up) has two disconnected continuous pieces (solutions $y = (a - x)^{-1}$ for $a < 0$ or for $a > 0$) and a singular point (solution $y = 0$). Another example is the equation $y'^2 + y^2 = 1$, the solutions on$(-\varepsilon, \varepsilon)$ of which form S^1, but the space of global solutions is $\mathbb{R}^1$ (both united with two singular points in $\mathrm{Sol}(\mathcal{E})$).

We would like to observe the "biggest" piece of the space of $\mathrm{Sol}(\mathcal{E})$, so that in our dimensional count we ignore isolated and special solutions or their families and take those of connected components, that have more parameters in.

It will be precisely the number of parameters, on which a general solution depends, that we count as a dimensional characteristic. Let us discuss the general idea how to count it and then give more specified definitions.

Note that in this paper we consider only (over)determined systems of

PDEs. Most results will work for underdetermined systems, but we are not concerned with them.

2 Understanding dimension of the solutions space

Let us treat at first the case of linear PDEs systems (the method can be transferred to non-linear case). We consider formal solutions and thus assume the system of PDEs $\mathcal{E}$ is formally integrable. We also assume the system $\mathcal{E} = \mathcal{E}_k$ is of pure order k, which shall be generalized later.

Thus for some vector bundle $\pi : E(\pi) \to M$ we identify $\mathcal{E}$ as a sub-bundle $\mathcal{E}_k \subset J^k(\pi)$ (see [S, Go, KLV]) and let $\mathcal{E}_l \subset J^l(\pi)$ be its $(l-k)$-th prolongations, $l \geq k$. Then the fibres $\mathcal{E}^\infty_x \subset J^\infty_x(\pi)$ at points $x \in M$ can be viewed as spaces of formal solutions of $\mathcal{E}$ at $x \in M$. To estimate size of $\mathcal{E}^\infty_x$ we consider the spaces of linear functions on $\mathcal{E}_{l,x}$, i.e. the space $\mathcal{E}^*_{l,x}$. The projections $\pi_{l,l-1} : \mathcal{E}_{l,x} \to \mathcal{E}_{l-1,x}$ induce embeddings $\pi^*_{l,l-1} : \mathcal{E}^*_{l-1,x} \hookrightarrow \mathcal{E}^*_{l,x}$, and we have the projective limit

$$\mathcal{E}^*_x = \cup_l \mathcal{E}^*_{l,x}.$$

Remark that $\mathcal{E}^*$ is the module over all scalar valued differential operators on π, while the kernel of the natural projection $J^\infty_x(\pi)^* \to \mathcal{E}^*_x$ can be viewed as the space of scalar valued differential operators on π vanishing on the solutions of the PDEs system $\mathcal{E}$ at the point $x \in M$. Thus elements of $\mathcal{E}^*_x$ are linear functions on the formal solutions $\mathcal{E}^\infty_x$.

We would like to choose "coordinates" among them, which will estimate dimension of the formal solution space. To do this we consider the graded module associated with filtred module $\mathcal{E}^*_x$:

$$g^*(x) = \bigoplus_{l \geq 0} g^*_l(x),$$

where $g_l(x)$ are the symbols of the equation at $x \in M$:

$$g_l(x) = \mathcal{E}_{l,x}/\mathcal{E}_{l-1,x} \subset S^l T^*_x \otimes \pi_x$$

(we let $\mathcal{E}_l = J^l(\pi)$ for $l < k$), and reduce analysis of $\mathcal{E}^*_x$ to investigation of the symbolic module g^*_x.

This g^* is the module over the symmetric algebra $ST_xM = \oplus S^i(T_xM)$ and its support $\mathrm{Char}^{\mathbb{C}}_x(\mathcal{E}) \subset \mathbb{P}^{\mathbb{C}}T^*_xM$ is a complex projective variety consisting of complex characteristic vectors. The values $\mathcal{K}_p$ of the symbolic module g^*_x at characteristic covectors $p \in {}^{\mathbb{C}}T^*_x \setminus 0$ form a family of vector spaces over $\mathrm{Char}^{\mathbb{C}}_x(\mathcal{E})$, which we call *characteristic sheaf*.

By the Noether normalization lemma ([E]) there is a subspace $U \subset T_xM$ such that the homogeneous coordinate ring $ST_xM/\operatorname{Ann} g^*(x)$ of $\mathrm{Char}^{\mathbb{C}}_x(\mathcal{E})$ is a finitely generated module over SU. It follows that $g^*(x)$ is a finitely generated module over SU too.

If $g^*(x)$ is a Cohen-Macaulay module (see [E], but we recall the definition later in a more general situation, then $g^*(x)$ is a free SU-module (we called the respective PDEs systems $\mathcal{E}$ Cohen-Macaulay in [KL$_2$] and discussed their corresponding reduction).

Let σ be the rank of this module, and $p = \dim U$. By the above discussion these numbers can be naturally called *formal functional rank* and *formal functional dimension* of the solutions space $\mathcal{E}^\infty_x$ at the point $x \in M$, because they describe on how many functions of how many variables a general jet-solution formally depends (we shall omit the word "formally" later), or how many "coordinates" from $\mathcal{E}^*_x$ should be fixed to get a formal solution.

If the symbolic module is not Cohen-Macaulay, the module $g^*(x)$ over SU is not free, but finitely generated and supported on $\mathbb{P}^{\mathbb{C}}U^*$. Let $\mathbb{F}(U)$ be the field of homogeneous functions P/Q, where $P, Q \in SU$, $Q \neq 0$, considered as polynomials on U^*. Thus $\mathbb{F}(U)$ is the field of meromorphic (rational) functions on U^*.

Consider $\mathbb{F}(U) \otimes g^*(x)$ as a vector space over $\mathbb{F}(U)$. Keeping the same definition for σ, let us call the dimension of this vector space p *formal rank* of $\mathcal{E}$ at the point $x \in M$.

It is clear that for Cohen-Macaulay systems the two notions coincide. However since $g^*(x)$ over SU is not free, we would like to give more numbers to characterize the symbolic module.

Let us choose a base $e_1, \ldots, e_r$ of $\mathbb{F}(U) \otimes g^*(x)$ such that $e_1, \ldots, e_r$ are homogeneous elements of $g^*(x)$ and denote by $\Gamma_1 \subset g^*(x)$ the SU-submodule generated by this base. It is easy to check that Γ_1 is a free SU-module. For the quotient module $M_1 = g^*(x)/\Gamma_1$ we have the following property:

$$\operatorname{Ann} h \neq 0 \text{ in } SU, \text{ for any } h \in M_1.$$

Therefore $\operatorname{Ann} M_1 \neq 0$ and the support Ξ_1 of M_1 is a proper projective variety in $\mathbb{P}^{\mathbb{C}}U^*$.

We apply the Noether normalization lemma to Ξ_1, we get a subspace $U_1 \subset U$, such that M_1 is a finitely generated module over SU_1. Its rank will be the next number p_1 and we also get $\sigma_1 = \dim U_1$, which we can call the next formal rank and formal dimension.

Applying this procedure several more times we get a sequence of varieties Ξ_i and numbers (p_i, σ_i), which depends, in general, on the choice of the flag $U \supset U_1 \supset U_2 \supset \ldots$ and the submodules Γ_i of SU_{i-1}.

Thus we resolve our symbolic module via the exact 3-sequences

$$0 \to \Gamma_1 \to g^* \to M_1 \to 0 \text{ over } SU, \; 0 \to \Gamma_2 \to M_1 \to M_2 \to 0 \text{ over } SU_1, \ldots$$

(with $\operatorname{Supp} M_i = \operatorname{Supp} \Gamma_{i+1} \supsetneqq \operatorname{Supp} M_{i+1}$) etc.

3 Cartan numbers

In Cartan's study of PDEs systems $\mathcal{E}$ (basically viewed as exterior differential systems in this approach) he constructed a sequence of numbers s_i, which are basic for his involutivity test. These numbers depend on the flag of subspaces one chooses for investigation of the system and so have no invariant meaning.

The classical formulation is that a general solution depends on s_p functions of p variables, s_{p-1} functions of $(p-1)$ variables, $\ldots$, s_1 functions of 1 variable and s_0 constants (we adopt here the notations from [BCG^3]; in Cartan's notations [C] we should rather write s_p, $s_p + s_{p-1}$, $s_p + s_{p-1} + s_{p-2}$ etc). However as Cartan notices just after the formulation [C], this statement has only a calculational meaning.

Nevertheless two numbers are absolute invariants and play an important role. These are Cartan genre, i.e. the maximal number p such $s_p \neq 0$, but $s_{p+1} = 0$, and Cartan integer $\sigma = s_p$.

As a result of Cartan's test a general solution depends on σ functions of p variables (and some number of functions of lower number of variables, but this number can vary depending on a way we parametrize the solutions). Here general solution is a local analytic solution obtained as a result of application of Cartan-Kähler (or Cauchy-Kovalevskaya) theorem and thus being parametrized by the Cauchy data.

Hence we can think of p as of *functional dimension* and of σ as of *functional rank* of the solutions space $\mathrm{Sol}(\mathcal{E})$. In fact, we adopt this terminology further on in the paper, because as was shown in the previous section it correctly reflects the situation.

These numbers can be computed via the characteristic variety. If the characteristic sheaf over $\mathrm{Char}^{\mathbb{C}}(\mathcal{E})$ has fibers of dimension k, then

$$p = \dim \mathrm{Char}^{\mathbb{C}}(\mathcal{E}) + 1,\ \sigma = k \cdot \deg \mathrm{Char}^{\mathbb{C}}(\mathcal{E}).$$

The first formula is a part of Hilbert-Serre theorem ([H]), while the second is more complicated. Actually Cartan integer σ was calculated in [BCG^3] in general situation and the formula is as follows.

Let $\mathrm{Char}^{\mathbb{C}}(g) = \cup_\epsilon \Sigma_\epsilon$ be the decomposition of the characteristic variety into irreducible components and $d_\epsilon = \dim \mathcal{K}_x$ for a generic point $x \in \Sigma_\epsilon$. Then

$$\sigma = \sum d_\epsilon \cdot \deg \Sigma_\epsilon.$$

It is important that these numbers coincide with the functional dimension and rank of the previous section. Moreover the sequence of Cartan numbers s_i is related to the sequence (p_i, σ_i) of the previous section.

This can be seen from the general approach of the next and following sections, which treat the case of systems $\mathcal{E}$ of PDEs of different orders (we though make presentation for the symbolic systems, with interpretation for general systems being well-known [S, KLV, KL$_2$]).

4 Symbolic systems

Consider a vector space T of dimension n (tangent space to the set of independent variables, substitute to T_xM) and a vector space N of dimension m (tangent space to the set of dependent variables, substitute to $\pi_x = \pi^{-1}(x)$).

Spencer δ-complex is de Rham complex of polynomial N-valued differential forms on T:

$$0 \to S^kT^* \otimes N \xrightarrow{\delta} S^{k-1}T^* \otimes N \otimes T^* \xrightarrow{\delta} \cdots \xrightarrow{\delta} S^{k-n}T^* \otimes N \otimes \Lambda^nT^* \to 0,$$

where $S^iT^* = 0$ for $i < 0$. Denote by

$$\delta_v = i_v \circ \delta : S^{k+1}T^* \otimes N \to S^kT^* \otimes N$$

the differentiation along the vector $v \in T$.

The l-th prolongation of a subspace $h \subset S^kT^* \otimes N$ is

$$\begin{aligned} h^{(l)} =& \{p \in S^{k+l}T^* \otimes N \,:\, \delta_{v_1} \dots \delta_{v_l} p \in h \ \forall v_1, \dots, v_l\} \\ =& S^lT^* \otimes h \cap S^{k+l}T^* \otimes N. \end{aligned}$$

Definition. *A sequence of subspaces $g_k \subset S^kT^* \otimes N$, $k \geq 0$, with $g_0 = N$ and $g_k \subset g_{k-1}^{(1)}$, is called a symbolic system.*

If a system of PDEs $\mathcal{E}$ is given as $F_1 = 0, \dots, F_r = 0$, where F_i are scalar PDEs on M, then $T = TM, N \simeq \mathbb{R}^m$ and the system $g \subset ST^* \otimes N$ is given as $f_1 = 0, \dots, f_r = 0$, where $f_i = \sigma(F_i)$ are symbols of the differential operators at the considered point (or jet for non-linear PDEs).

With every such a system we associate its Spencer δ-complex of order k:

$$\begin{aligned} 0 \to g_k \xrightarrow{\delta} g_{k-1} \otimes T^* \xrightarrow{\delta} g_{k-2} \otimes \Lambda^2T^* \to \dots \\ \to g_i \otimes \Lambda^{k-i}T^* \xrightarrow{\delta} \cdots \xrightarrow{\delta} g_{k-n} \otimes \Lambda^nT^* \to 0. \end{aligned}$$

Definition. *The cohomology group at the term $g_i \otimes \Lambda^jT^*$ is denoted by $H^{i,j}(g)$ and is called the Spencer δ-cohomology of g.*

Note that $g_k = S^k T^* \otimes N$ for $0 \leq k < r$ and the first number $r = r_{\min}(g)$, where the equality is violated is called the minimal order of the system. Actually the system has several orders:

$$\mathrm{ord}(g) = \{k \in \mathbb{Z}_+ \,|\, g_k \neq g_{k-1}^{(1)}\}.$$

Multiplicity of an order r is:

$$m(r) = \dim g_{r-1}^{(1)}/g_r = \dim H^{r-1,1}(g).$$

Hilbert basis theorem implies finiteness of the set of orders (counted with multiplicities):

$$\mathrm{codim}(g) := \dim H^{*,1}(g) = \sum m(r) < \infty.$$

Starting from the maximal order of the system $k = r_{\max}$ we have:

$$g_{k+l} = g_k^{(l)}.$$

If we dualize the above construction over $\mathbb{R}$, then Spencer δ-differential transforms to a homomorphism over the algebra of polynomials ST and $g^* = \oplus_i g_i^*$ becomes an ST-module. This module is called a *symbolic module* and it plays an important role in understanding PDEs.

In particular, *characteristic variety* $\mathrm{Char}^{\mathbb{C}}(g) \subset \mathbb{P}^{\mathbb{C}}T^*$ is defined as the support of this module $\mathrm{Supp}(g^*) = \{[p] : (g^*)_p \neq 0\}$ and the *characteristic sheaf* $\mathcal{K}$ over it is the family of vector spaces, which at the point $p \in \mathrm{Char}^{\mathbb{C}}(g)$ equals the value of the module at this point $\mathcal{K}_p = g^*/p \cdot g^*$. For more geometric description see [S, KLV, KL_2].

5 Commutative algebra approach

We will study only local solutions of a system of PDEs $\mathcal{E}$, which we consider in such a neighborhood that type of the symbolic system does not change from point to point (on equation) in the sense that dimensions of g_k, of the characteristic variety $\mathrm{Char}^{\mathbb{C}}(g)$ and of the fibers of $\mathcal{K}$ are the same.

It should be noted that if a system $\mathcal{E}$ is not formally integrable and $\mathcal{E}'$ is obtained from it by the prolongation-projection method [K, M2, KL_2], then the numbers p, σ change in this process, i.e. either the functional dimension or the functional rank decrease. Thus from now on we suppose the system $\mathcal{E}$ is formally integrable.

The numbers p, σ can be described using the methods of commutative algebra. Recall ([AM]) that by Hilbert-Serre theorem the sum

$$f(k) = \sum_{i \le k} \dim g_i^*$$

behaves as a polynomial in k for sufficiently large k. This polynomial is called the *Hilbert polynomial* of the symbolic module g^* corresponding to $\mathcal{E}$ and we denote it by $P_{\mathcal{E}}(z)$. If $p = \deg P_{\mathcal{E}}(z)$ and $\sigma = P_{\mathcal{E}}^{(p)}(z)$, then the highest term of this polynomial is

$$P_{\mathcal{E}}(z) = \sigma z^p + \dots$$

(see [H] for the related statements in algebraic geometry, the interpretation for PDEs is straightforward).

A powerful method to calculate the Hilbert polynomial is resolution of a module. In our case a resolution of the symbolic module g^* exists and it can be expressed via the Spencer δ-cohomology. Indeed, the Spencer cohomology of the symbolic system g is $\mathbb{R}$-dual to the Koszul homology of the module g^* and for algebraic situation this resolution was found in [Gr]. It has the form:

$$\begin{aligned} 0 \to \oplus_q H^{q-n,n}(g) \otimes S^{[-q]} \xrightarrow{\varphi_n} \oplus_q H^{q-n+1,n-1}(g) \otimes S^{[-q]} \xrightarrow{\varphi_{n-1}} \dots \\ \to \oplus_q H^{q-1,1}(g) \otimes S^{[-q]} \xrightarrow{\varphi_1} \oplus_q H^{q,0}(g) \otimes S^{[-q]} \xrightarrow{\varphi_0} g^* \to 0, \end{aligned}$$

where $S^{[-q]}$ is the polynomial algebra on T_x^*M with grading shifted by q, i.e. $S_i^{[-q]} = S^{i-q}T_xM$, and the maps φ_j have degree 0.

Thus denoting $h^{i,j} = \dim H^{i,j}(g)$ and $\tau_\alpha = \dim S^\alpha TM = \binom{\alpha+n-1}{\alpha}$ we have:

$$\dim g_i = \sum_q \big(h^{q,0}\tau_{i-q} - h^{q,1}\tau_{i-q-1} + h^{q,2}\tau_{i-q-2} - \dots + (-1)^n h^{q,n}\tau_{i-q-n} \big).$$

Let also $j_\beta = \sum_{\alpha\le\beta}\tau_\alpha = \dim J^\beta_v M = \binom{\beta+n}{n}$ be the dimension of the fiber of the vertical jets $J^\beta_v M$, i.e. the fiber of the jet space $J^\beta M$ over M. Thus we calculate

$$\sum_{i\le k}\dim g_i = \sum_q \big(h^{q,0}j_{k-q} - h^{q,1}j_{k-q-1} + h^{q,2}j_{k-q-2} - \cdots \pm h^{q,n}j_{k-q-n}\big).$$

Finally we deduce the formula for Hilbert polynomial of the symbolic module g^*

$$P_{\mathcal{E}}(z) = \sum_q \Big(h^{q,0}\tbinom{z-q+n}{n} - h^{q,1}\tbinom{z-q+n-1}{n} + \\ + h^{q,2}\tbinom{z-q+n-2}{n} - \cdots + (-1)^n h^{q,n}\tbinom{z-q}{n}\Big).$$

Here

$$\tbinom{z+k}{k} = \frac{1}{k!}(z+1)\cdot(z+2)\cdots(z+k).$$

Denote $S_j(k_1,\dots,k_n) = \sum\limits_{i_1<\cdots<i_j} k_{i_1}\cdots k_{i_j}$ the j-th symmetric polynomial and let also

$$s^n_i = \frac{(n-i)!}{n!}\,S_i(1,\dots,n)$$

Thus

$$s^n_0 = 1, \quad s^n_1 = \frac{n+1}{2}, \quad s^n_2 = \frac{(n+1)(3n+2)}{4\cdot 3!}, \quad s^n_3 = \frac{n(n+1)^2}{2\cdot 4!},$$

$$s^n_4 = \frac{(n+1)(15n^3+15n^2-10n-8)}{48\cdot 5!} \quad \text{etc.}$$

If we decompose

$$\tbinom{z+n}{n} = \sum_{i=0}^n s^n_i \frac{z^{n-i}}{(n-i)!},$$

then we get the expression for the Hilbert polynomial

$$P_{\mathcal{E}}(z) = \sum_{i,j,q}(-1)^i h^{q,i}s^n_j\frac{(z-q-i)^{n-j}}{(n-j)!} = \sum_{k=0}^n b_k\frac{z^{n-k}}{(n-k)!},$$

where

$$b_k = \sum_{j=0}^{k} \sum_{q,i} (-1)^{i+j+k} h^{q,i} s_j^n \frac{(q+i)^{k-j}}{(k-j)!}.$$

6 Calculations for the Solutions space

We are going to compute the dimensional characteristics of two important classes of PDEs.

Involutive systems. These are such symbolic systems $g = \{g_k\}$ that all subspaces g_k are involutive in the sense of Cartan [C, BCG^3] (this definition for the symbolic systems of different orders was introduced in [KL$_5$]). Thanks to Serre's contribution [GS] we can reformulate this via Spencer cohomology as follows.

Denote by $g^{|k\rangle}$ the symbolic system generated by all differential corollaries of the system deduced from the order k:

$$g_i^{|k\rangle} = \begin{cases} S^i T^* \otimes N, & \text{for } i < k; \\ g_k^{(i-k)}, & \text{for } i \geq k. \end{cases}$$

Then the system g is involutive iff $H^{i,j}(g^{|k\rangle}) = 0$ for all $i \geq k$ (this condition has to be checked for $k \in \mathrm{ord}(g)$ only), see [KL$_5$].

In particular, $H^{i,j}(g) = 0$ for $i \notin \mathrm{ord}(g) - 1$, $(i,j) \neq (0,0)$, and the resolution for the symbolic module g^* as well as the formula for the Hilbert polynomial of $\mathcal{E}$ become easier.

Let us restrict for simplicity to the case of systems of PDEs $\mathcal{E}$ of pure first order. Then

$$\begin{aligned} P_{\mathcal{E}}(z) =& h^{0,0}\tbinom{z+n}{n} - h^{0,1}\tbinom{z+n+1}{n+1} + h^{0,2}\tbinom{z+n+2}{n+2} - \dots \\ =& b_1 \frac{z^{n-1}}{(n-1)!} + b_2 \frac{z^{n-2}}{(n-2)!} + \dots + b_0. \end{aligned}$$

Vanishing of the first coefficient $b_0 = 0$ is equivalent to vanishing of Euler characteristic for the Spencer δ-complex, $\chi = \sum_i (-1)^i h^{0,i} = 0$, and

this is equivalent to the claim that not all the covectors from ${}^{\mathbb{C}}T^* \setminus 0$ are characteristic for the system g.

The other numbers b_i are given by the general formulas from the previous section, but they simplify in our case. For instance

$$b_1 = \tfrac{n+1}{2}\, b_0 - \sum (-1)^i h^{0,i} i = \sum (-1)^{i+1} i \cdot h^{0,i}.$$

If $\operatorname{codim} \operatorname{Char}^{\mathbb{C}}(\mathcal{E}) = n - p > 1$, then $b_1 = 0$ and in fact then $b_i = 0$ for $i < n - p$, but $b_{n-p} = \sigma$.

Theorem. *If* $\operatorname{codim} \operatorname{Char}^{\mathbb{C}}(\mathcal{E}) = n - p$, *then the functional rank of the system equals*

$$\sigma = \sum_i (-1)^i h^{0,i} \frac{(-i)^{n-p}}{(n-p)!}.$$

Proof. Indeed one successively calculate the coefficients using the formula

$$b_k = \sum_i \sum_{\alpha=0}^{k} (-1)^{i+\alpha} h^{0,i} s^n_{k-\alpha} \frac{i^\alpha}{\alpha!}$$

and notes that b_k equals to the displayed expression plus a linear combination of $b_{k-1}, \ldots, b_0$. The claim follows. □

One can extend the above formula for general involutive system and thus compute the functional dimension and functional rank of the solutions space (some interesting calculations can be found in classical works [J, C]).

Cohen-Macaulay systems. A symbolic system g (and the respective PDEs system $\mathcal{E}$) is called Cohen-Macaulay ([KL_2]) if the corresponding symbolic module g^* is Cohen-Macaulay, i.e. (see [M1, E] for details)

$$\dim g^* = \operatorname{depth} g^*.$$

Consider an important partial case (we formulate the definition only for symbolic systems; PDEs are treated in [KL_4]):

Definition. *A symbolic system* $g \subset ST^* \otimes N$ *(*$n = \dim T$, $m = \dim N$*) of* $\operatorname{codim}(g) = r$ *is called a* generalized complete intersection *if*

- $m \leq r < n + m$;
- $\operatorname{codim}_{\mathbb{C}} \operatorname{Char}^{\mathbb{C}}(g) = r - m + 1$;
- $\dim \mathcal{K}_x = 1 \ \forall x \in \operatorname{Char}^{\mathbb{C}}(g) \subset P^{\mathbb{C}}T^*$.

Formal integrability of such systems are given by the compatibility conditions expressed via brackets (for scalar systems [KL_1, KL_3]) or multi-brackets (for vector systems [KL_4]). In this case we can calculate Cartan genre and integer directly.

Theorem. *Let $\mathcal{E}$ be a system of generalized complete intersection type and suppose it is formally integrable. Then the functional dimension of* $\operatorname{Sol}(\mathcal{E})$ *is*

$$p = m + n - r - 1$$

and the functional rank is

$$\sigma = S_{r-m+1}(k_1, \ldots, k_r) = \sum_{i_1 < \cdots < i_{r-m+1}} k_{i_1} \cdots k_{i_{r-m+1}},$$

the l-th symmetric polynomial of the orders $k_1, \ldots, k_r$ of the system.

Note that if the last requirement in the definition of generalized complete intersection is changed to $\dim \mathcal{K}_x = d$ everywhere on the characteristic variety, then the functional rank will be multiplied by d:

$$\sigma = d \cdot S_{r-m+1}(k_1, \ldots, k_r).$$

However the formal integrability criterion for generalized complete intersections is proved in [KL_4] under assumption that $d = 1$.

Proof. We shall consider the case of a system g of a pure order: $k_1 = \cdots = k_r = k$, $k_i \in \operatorname{ord}(g)$. The case of different orders is similar and will appear elsewhere.

The formula for functional dimension p follows directly from the definition of generalized complete intersection. Let's calculate σ.

We can use interpretation of the Cartan integer σ from §3. Recall that characteristic variety $\mathrm{Char}^{\mathbb{C}}(g)$ is the locus of the characteristic ideal $I(g) = \mathrm{Ann}(g)$, which the the annihilator of g^* in ST.

Since the module is represented by the matrix with polynomial entries (each differential operator Δ_i giving a PDEs system $\mathcal{E}$ is a column Δ_{ij}, $1 \le i \le r$, $1 \le j \le m$; so that their union is a $m \times r$ matrix $M(\Delta)$), its annihilator is given by the zero Fitting ideal (in fact, here we use the condition on grade of the ideal: $\operatorname{depth} \mathrm{Ann}(g) = r - m + 1$, which follows from the conditions of the above definition).

This ideal $\mathrm{Fitt}_0(g)$ is generated by all determinants of $m \times m$ minors of the corresponding to $M(\Delta)$ matrix of symbols $M(\sigma_\Delta)$. These minors are determined by a choice of m from r columns, so that there are $\binom{r}{m}$ determinants and each is a polynomial of degree k^{r-m+1}.

However not all the minors are required to determine $\mathrm{Char}^{\mathbb{C}}(g)$ and this is manifested by the fact, that we sum $\binom{r}{m-1}$ degrees k^{r-m+1} to get the functional rank σ. The easiest way to explain this is via the Hilbert polynomial of the symbolic module g^*.

This can be calculated since under the assumption of generalized complete intersection g^* possesses a resolution in the form of Buchsbaum-Rim complex (see [KL_4]):

$$0 \to S^{r-m-1}V^\star \otimes \Lambda^r U \to S^{r-m-2}V^\star \otimes \Lambda^{r-1}U \to \\ \cdots \to \Lambda^{m+1}U \to U \to V \to g^* \to 0,$$

where $V \simeq ST \otimes N^*$ (recall that $\dim N = m$ and $g \subset ST^* \otimes N$) and $U = \underbrace{ST \oplus \cdots \oplus ST}_{r \text{ terms}}$. Star $\star$ means dualization over ST and the tensor products are over ST as well.

Now the claim follows from the detailed investigation of degrees of the homomorphisms in the above exact sequence. To see this we suppose at first that $r = m + n - 1$ and use the following assertion.

Lemma. *The following combinatorial formula holds:*

$$m\tbinom{n+k(n+m-1)}{n} - (n+m-1)\tbinom{n+k(n+m-2)}{n} + \sum_{j=1}^{n-1}(-1)^{j-1}\tbinom{j+m-2}{m-1}\tbinom{n+m-1}{j+m}\tbinom{(k+1)n-k(1+j)}{n} = \tbinom{n+m-1}{n}k^n.$$

We would like to comment and interpret the sum on the left hand side of this formula. In our case the system is of finite type (g^* has finite dimension as a vector space) and $\sigma = \sum \dim g_i$ (the sum is finite).

Stabilization of the symbol occurs at the order $i = \sum k_i - 1 = k(n + m - 1) - 1$: $g_i = 0$. So we prolong $\mathcal{E}$ to the jets of order $k(n + m - 1)$ and the first term is just $\dim J_v^{k(n+m-1)}(M, N)$.

The next term is due to the fact that $\mathcal{E} \subset J^k(M, N)$ is proper. It is given by $r = n + m - 1$ equations of order k, we which we differentiate up to $k(n + m - 2)$ times along all coordinate directions (prolongation).

There are relations between these derivatives. These are compatibility conditions (1-syzygy of g^*), which appear in the form of multi-brackets [KL_4], in our case this bracket uses $(m + 1)$-tuples of Δ_i.

There are in turn relations among relations (2-syzygy of g^*), which are identities between multi-brackets (these we call generalized Plücker identities, to appear soon), in our case these latter use $(m + 2)$-tuples of the defining operators Δ_i etc.

Due to exact form of the relations (higher syzygies) we get factors $\binom{j+m-2}{m-1}$ in the summations formula of the lemma.

In the case $r < n + m - 1$ we should perform a reduction, which is possible by Theorem A [KL_2]. Then the functional dimension p grows, but the functional rank remains the same and the previous calculation works. □

7 Examples

Here we show some examples demonstrating the above results.

1. *Intermediate integral* of a system $\mathcal{E} \subset J^k\pi$ is such a system $\tilde{\mathcal{E}} \subset J^{\tilde{k}}\pi$ that $\tilde{k} < k$ and $\mathcal{E} \subset \tilde{\mathcal{E}}^{(k-\tilde{k})}$ (where $\mathcal{E}^{(i)}$ is the i-th prolongation of the system). Since every solution to the system $\mathcal{E}$ is a solution to $\tilde{\mathcal{E}}^{(k-\tilde{k})}$ we conclude: Whenever the functional dimension $p > 0$, we have $\tilde{p} = p$ and $\tilde{\sigma} = \sigma$.

Indeed the solutions of $\tilde{\mathcal{E}}^{(k-\tilde{k})}$ form a finite-dimensional parametric family, such that solutions of $\tilde{\mathcal{E}}$ appear for some fixed values of parameters (because we differentiate with respect to all variables to obtain the prolongation). Thus the number of functions of $p > 0$ variables, on which a general solution depends, will not be altered.

2. If the PDEs system $\mathcal{E}$ is underdetermined, then $p = n$ and $\sigma \geq 1$. Indeed, σ is precisely the under-determinacy degree, i.e. the minimal number of unknown functions that should be arbitrarily fixed to get a determined system. We assume we can do it to get a formally integrable system. If underdetermined system is not formally integrable, compatibility conditions can turn it into determined or over-determined and then decrease p and change σ.

A nice illustration is the Hilbert-Cartan system

$$z'(x) = (y''(x))^2.$$

It has $p = 1$, $\sigma = 1$. But even though a general solution depends on one function of one variable, it cannot be represented in terms of a function and its derivatives only (Hilbert's theorem).

3. As we noticed earlier the similar situation happens to overdetermined system: If $\mathcal{E}$ is not formally integrable, and $\tilde{\mathcal{E}}$ is obtained from $\mathcal{E}$ by prolongation-projection technique (sometimes it is said that $\tilde{\mathcal{E}}$ is the involutive form of $\mathcal{E}$, but this is not true, only a certain prolongation of $\tilde{\mathcal{E}}$ is), then $\tilde{p} < p$ or $[\tilde{p} = p$ and $\tilde{\sigma} < \sigma]$. Indeed, supplement of additional equations shrinks the solution space.

For instance if we consider two second-order scalar differential equations on the plane

$$F\big(x, y, u(x,y), Du(x,y), D^2u(x,y)\big) = 0,$$

$$G\big(x, y, u(x,y), Du(x,y), D^2u(x,y)\big) = 0,$$

such that F and G have no common complex characteristics, then the compatibility condition of this system $\mathcal{E}$ can be expressed via the Mayer bracket ([KL$_1$]): $H = [F, G]_{\mathcal{E}}$. If $H = 0$, then $p = 0$, $\sigma = 4$. If $H \neq 0$, then $p = 0$ and $\sigma \leq 3$, the equality being given by the Frobenius condition for the system $\tilde{\mathcal{E}} = \{F = 0, G = 0, H = 0\}$.

If the system has one common characteristic and is compatible, we have: $p = 1$, $\sigma = 1$. Pairs of such systems are basic examples of Darboux integrability.

4. Evolutionary equations $u_t = L[u]$ provide interesting examples, which usually "contradict" the theory. Consider for instance the heat equation

$$u_t = u_{xx}.$$

It is formally integrable and analytic. We can try to specify the initial condition $u|_{t=0} = \varphi(x)$ and then solve the Cauchy problem, so that we get $p = 1$, $\sigma = 1$. On the other hand we can let $u|_{x=0} = \psi_0(t), u_x|_{x=0} = \psi_1(t)$ and then get $p = 1$, $\sigma = 2$.

If we calculate the numbers using our definitions of functional dimension and functional rank (for instance, via Hilbert polynomial), it turns out that the second approach is correct. Indeed with the first idea we come into trouble with certain Cauchy data: Let, for instance, $\varphi(x) = (1-x)^{-1}$, which is an analytic function around the origin. Then the analytic solution should have the series

$$u(t,x) \doteqdot \frac{1}{1-x} + \frac{2}{1}\frac{t}{(1-x)^3} + \frac{4!}{2!}\frac{t^2}{(1-x)^5} + \cdots + \frac{(2n)!}{n!}\frac{t^n}{(1-x)^{2n+1}} + \cdots$$

which diverges everywhere outside $t = 0$. The reason why the second approach provides no problem is because the line $\{x = 0\}$ is non-characteristic

and we can solve our first order PDE by the classical method of Cauchy characteristics.

Remark however that in the standard courses of mathematical physics the heat equation is solved with the first approach (by Fourier method). How is it possible?

Explanation is that we solve the heat equation then only for positive time $t \geq 0$. Doing the same method in negative direction blows up the solutions immediately (heat goes rapidly to equilibrium, but we cannot predict even closest past)! We here are interested in the solutions, which exist in an open neighborhood of the origin (like in Cauchy-Kovalevskaya theorem), and this contradicts the first approach.

5. Similar problems arise with Cauchy problems in other PDEs systems: one usually applies reduction or fixes gauge, but this can change dimensional characteristics.

For instance, consider the Cauchy problem for the Einstein vacuum equations, which is a system of 10 PDEs of 10 unknown functions. The system is over-underdetermined (i.e. it has compatibility conditions). In wave gauge [CB] its solution depends on several functions on a 3-dimensional space, which are subject to constraint equations, so that $p = 2$. On the other hand, the original Einstein system is invariants under diffeomorphisms and this yields $p = 4$.

One should also be careful with Cauchy data in higher order, since then the definition of characteristics becomes more subtle, see [KL_5].

6. Consider a system $\mathcal{E}$, which describes automorphisms of a given geometric structure. The corresponding symbolic system is $g \subset ST^* \otimes T$. The automorphism group has maximal dimension iff the system is formally integrable. Consider the examples, when the geometric structure is symplectic, complex or Riemannian (all these structures are of the first order).

Let at first g be generated by $g_1 = \mathrm{sp}(n) \subset T^* \otimes T$. Our tangent space $T = T_xM$ is equipped with a symplectic structure ω, and we can identify $T^* \overset{\omega}{\simeq} T$ and we get $g_1 = S^2T^* \subset T^* \otimes T^*$. The prolongations are

$g_i = S^{i+1}T^* \subset S^iT^* \otimes T$.

The system is easily checked to be involutive and the only non-vanishing Spencer δ-cohomology groups are

$$H^{0,i}(g) = \Lambda^{i+1}T^*.$$

Then one checks that the Euler characteristic is $\chi = 1 \neq 0$ and so $b_0 \neq 0$. Thus the functional dimension is $p = n$. Indeed the characteristic variety is $\mathbb{P}^{\mathbb{C}}T^*$ because each non-zero covector p is characteristic: $p^2 \in g_1 \simeq S^2T^*$. Next by a theorem from §6 one calculates the functional rank

$$\sigma = \sum_{i=0}^{n-1}(-1)^i\binom{n}{i+1}\frac{(-i)^0}{0!} = \chi = 1.$$

This result is easy to verify: an infinitesimal symplectic transformation has a generating function (Hamiltonian) and so it is determined by one function of n variables.

If we turn to (almost) complex structures J on M, then $g_1 = \mathrm{gl}(\frac{n}{2}, \mathbb{C}) = T^* \otimes_{\mathbb{C}} T$ (space of $\mathbb{C}$-linear endomorphisms of T) and the prolongations are $g_i = S^i_{\mathbb{C}}T^* \otimes_{\mathbb{C}} T$.

The characteristic variety is proper and one calculates that $p = \frac{n}{2}$, $\sigma = n$. The system is again involutive. The second Spencer cohomology is

$$H^{0,2}(g) = \Lambda^2_{\mathbb{C}}T^* \otimes_{\bar{\mathbb{C}}} T,$$

which is the space of $\mathbb{C}$-antilinear skew-symmetric $(2,1)$ tensors (Nijenhuis tensors).

The last example is the algebra of Riemannian isometries (i.e. T is equipped with a Riemannian structure) of a Riemannian metric q on M. The symbol is $g_1 = o(n)$ and the prolongations are zero $g_2 = g_3 = \cdots = 0$.

This system is not involutive. For instance,

$$H^{1,2}(g) = \mathrm{Ker}\big(S^2\Lambda^2T^* \to \Lambda^4T^*\big)$$

(the space of Riemannian curvatures) is non-zero (for $n = \dim T > 1$). Since the system is of finite type, the characteristic variety is empty and $p = 0$. The general solution (isometry) depends on $\sigma = \frac{(n+1)n}{2}$ constants.

We recall, that the above dimensional conclusions are correct if the system $\mathcal{E}$ is integrable, otherwise the space $\mathrm{Sol}(\mathcal{E})$ shrinks. In the above examples this means: the form ω is closed (with just non-degeneracy we have almost-symplectic manifold); the structure J is integrable (Nijenhuis tensor N_J vanishes); the manifold (M, q) has constant sectional curvature (so it is a spacial form).

References

[AM] M. F. Atiyah, I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley (1969).

[C] E. Cartan, *Les systèmes différentiels extérieurs et leurs applications géométriques* (French), Actualités Sci. Ind. **994**, Hermann, Paris (1945).

[CB] Y. Choquet-Bruhat (Foures-Bruhat), *Theoreme d'existence pour certains systemes d'equations aux derivees partielles non lineaires* (French), Acta Math. **88**, 141–225 (1952).

[BCG^3] R. L. Bryant, S. S. Chern, R. B. Gardner, H. L. Goldschmidt, P. A. Griffiths, *Exterior differential systems*, MSRI Publications **18**, Springer-Verlag (1991).

[E] D. Eisenbud, *Commutative algebra with a view toward algebraic geometry*, Springer-Verlag (1995).

[Go] H. Goldschmidt, *Integrability criteria for systems of nonlinear partial differential equations*, J. Diff. Geom., **1**(3) (1967), 269–307.

[Gr] M L. Green, *Koszul cohomology and geometry*, in: "Lectures on Riemann surfaces", World Sci. Publ., 177–200 (1989).

[GS] V. Guillemin, S. Sternberg, *An algebraic model of transitive differential geometry*, Bull. A.M.S., **70** (1964), 16–47.

[H] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer-Verlag (1977).

[Ho] L. Hörmander, *Linear partial differential operators*, Springer Verlag (1976).

[J] M. Janet, *Leçons sur les systèms d'équations*, Gauthier-Villers, Paris (1929).

[Ka] E. Kähler, *Einfürhung in die Theorie der Systeme von Differentialgleichungen*, B.G. Teubner, Leipzig (1934).

[KLV] I. S. Krasilschik, V. V. Lychagin, A. M. Vinogradov, *Geometry of jet spaces and differential equations*, Gordon and Breach (1986).

[KL$_1$] B. S. Kruglikov, V. V. Lychagin, *Mayer brackets and solvability of PDEs – I*, Diff. Geom. and its Appl. **17** (2002), 251–272.

[KL$_2$] B. S. Kruglikov, V. V. Lychagin, *Mayer brackets and solvability of PDEs – II*, Trans. Amer. Math. Soc. **358**, no.3 (2005), 1077–1103.

[KL_3] B. S. Kruglikov, V. V. Lychagin, *A compatibility criterion for systems of PDEs and generalized Lagrange-Charpit method*, A.I.P. Conference Proceedings, "Global Analysis and Applied Mathematics: International Workshop on Global Analysis", **729**, no. 1, 39-53 (2004).

[KL_4] B. S. Kruglikov, V. V. Lychagin, *Multi-brackets of differential operators and compatibility of PDE systems*, Comptes Rendus Math. **342**, no. 8 (2006), 557–561.

[KL_5] B. S. Kruglikov, V. V. Lychagin, *Spencer delta-cohomology, restrictions, characteristics and involutive symbolic PDEs*, ArXive: math.DG/0503124.

[K] M. Kuranishi, *On E. Cartan's prolongation theorem of exterior differential systems*, Amer. J. Math. **79** (1957), 1–47.

[L] H. Lewy, *An example of a smooth linear partial differential equation without solution*, Ann. of Math. (2) **66** (1957), 155–158.

[M1] H. Matsumura, *Commutative algebra*, New York: Benjamin (1970).

[M2] Y. Matsushima, *On a theorem concerning the prolongation of a differential system*, Nagoya Math. J. **6**, 1–16 (1953).

[S] D. C. Spencer, *Overdetermined systems of linear partial differential equations*, Bull. Amer. Math. Soc., **75** (1969), 179–239.

Effective Homology of Hopf Algebras[1]

Ainhoa Berciano-Alcaraz

Euskal Herriko Unibertsitatea
Dpto. Matemática Aplicada, Estadística e Investigación Operativa,
ainhoa.berciano@ehu.es

Abstract

In this paper we explain the importance of the notion of A_∞-Hopf algebra in the Algebraic Topology framework. In particular, we focus our efforts to show how, given a Hopf algebra, it is possible to weaken the structure in the computation of the homology. Furthermore, we will see that this process induces a new sophisticated structure, an A_∞-Hopf algebra.

Keywords: Homological perturbation theory, reduction, Basic Perturbation Lemma, Hopf algebras.

1 Introduction

Classification in Algebraic Topology frequently depends on *homology groups*, but these groups are most often difficult to reach. Various methods are available to facilitate and structure the calculation of these groups. In particular *reductions* between:

- Chain complexes which are richly structured (algebras, coalgebras, simplicial, cosimplicial. . .) but not of finite type;
- Simpler chain complexes poorly structured but of finite type;

play an important role, see [6].

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

In particular the A_∞-algebra and A_∞-coalgebra structures are defined in this way. A homology equivalence between a differential algebra A and a simple chaincomplex C defines an A_∞-algebra structure over C [10], the same for coalgebras. An A_∞-algebra is a sort of "weakened" algebra where the standard requirements for an algebra are satisfied only *up to homotopy*. The same for coalgebras.

The right definition for the notion of A_∞-Hopf algebra is a challenge for a long time, see in particular Ron Umble and Samson Saneblidze's papers [7, 8, 9]. We propose here an original point of view.

The notion of A_∞-algebra (resp. A_∞-coalgebra) greatly depends on the Bar (resp. Cobar) construction. A Hopf algebra is simultaneously an algebra and a coalgebra, so that it is tempting to define a "Bar-Cobar" construction for a Hopf algebra. It happens the Hopf relation explaining how algebra and coalgebra structures fit to each other is the key point allowing us to define the Bar-Cobar construction.

The Bar construction leads to the notion of A_∞-algebra, and the Bar-Cobar construction gives by an analogous process which seems a *natural* notion of A_∞-Hopf algebra.

2 Preliminaries and didactic examples

Here we will collect some basic definitions and results of homological algebra, as well as some simple examples.

Take a commutative ground ring with unit, Λ. A *differential graded module or dg-module*, (M, d) is a graded module, with a differential $d : M \to M$ (that is, a morphism of degree -1 such that $d\,d = 0$). A graded module M is *connected* whenever $M_0 = \Lambda$ and *simply connected* if it is connected and $M_1 = 0$. In such a case, the graded module $\overline{M}$ is defined as $\overline{M}_n = M_n$ for $n > 1$ and $\overline{M}_0 = 0$.

The *homology* of a dg-module M is the graded module $H_*(M)$, where

$$H_n(M) = \operatorname{Ker} d_n / \operatorname{Im} d_{n+1}.$$

A *dg–algebra*, (A, d_A, μ_A), is a dg–module endowed with an associative product with unit, compatible with the differential. Analogously, a *dg–coalgebra* (C, d_C, Δ_C) is a dg–module provided with a compatible coproduct and counit.

Furthermore, if $(H, d_H, \mu_H, \Delta_H)$ is a dg-algebra, a dg-coalgebra and product and coproduct verify the Hopf relation, i.e., $\Delta\mu = \mu^{\otimes 2}(1 \otimes T \otimes 1)\Delta^{\otimes 2}$, then it is a *dg-Hopf algebra*.

As simple examples of dg-Hopf algebras, let us mention

- The *polynomial algebra* $P(v, 2n)$, generated by v of degree $2n$, where n is a positive integer. The product is the usual one of monomials i.e., $v^i v^j = v^{i+j}$. The coproduct is $\Delta(v) = v \otimes 1 + 1 \otimes v$ and extended thanks to the Hopf relation.
- The *truncated polynomial algebra* $Q_p(v, 2n)$ is the quotient algebra $P(v, 2n)/(v^p)$, where p is a primer number.
- The *exterior algebra* $E(u, 2n+1)$, $n \geq 0$, with algebra generator u of degree $2n+1$ and trivial product $u^2 = 0$.
- The *divided polynomial algebra* $\Gamma(w, 2n)$, $n \geq 1$, generated by $\gamma_1(w) = w$ with the product given by $\gamma_k(w)\gamma_h(w) = \frac{(k+h)!}{k!\,h!}\,\gamma_{k+h}(w)$. The coproduct is defined by $\Delta(\gamma_k(w)) = \sum_{i+j=k} \gamma_i(w) \otimes \gamma_j(w)$.

A fundamental tool to relate two dg-modules is the next one:

A *reduction* $c : \{N, M, f, g, \phi\}$, from a dg–module (N, d_N) to another one (M, d_M) is a special type of homology equivalence given by the morphisms f, g and ϕ; where $f : N_* \to M_*$, $g : M_* \to N_*$ are two morphisms of degree zero and $\phi : N_* \to N_{*+1}$ is a homotopy operator. Apart from the conditions

$$(c1)\ fg = 1_M\,, \quad (c2)\ \phi d_N + d_N\phi + gf = 1_N\,,$$

the following ones must be satisfied

$$(c3)\ f\phi = 0\,, \quad (c4)\ \phi g = 0\,, \qquad (c5)\ \phi\phi = 0\,.$$

The most important consequence is that the homology groups of N are naturally isomorphic to the homology groups of M. But in general, if N is a dg-algebra, this structure is not transferred to M, which inherits an A_∞-algebra structure (the morphisms f and g *are not* isomorphisms and there is no coherent way in general to transfer the algebra structure to M).

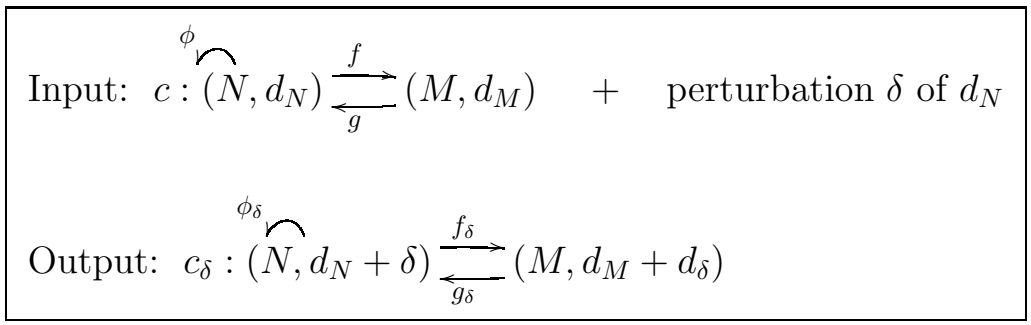
More explicitly, an A_∞-*algebra* is a dg-module with a family of operations $\mu_i : A^{\otimes i} \to A$, of degree $i-2$, such that for all $i \geq 1$ the following relations are satisfied

$$\sum_{n=1}^{i} \sum_{k=0}^{i-n} (-1)^{n+k+nk} \mu_{i-n+1}(1^k \otimes \mu_n \otimes 1^{\otimes i-n-k}) = 0.$$

In the same way, we can speak about an A_∞-coalgebra as a dg-module that has to verify the dual properties (in this case the operations are denoted by Δ_i).

Thanks to the perturbation theory, to speak about an A_∞-(co)algebra M is equivalent to give a reduction from a dg-(co)algebra to the dg-module M. Let us recall how it is possible.

The key in the homological perturbation theory is the **Basic Perturbation Lemma** (briefly, BPL [2]), which is an algorithm whose input is a reduction of dg–modules
$c : \{N, M, f, g, \phi\}$ and a perturbation datum δ of d_N whose output is a new reduction c_δ. The only requirement is the pointwise nilpotency of the composition $\phi\delta$, that guarantees that the sums involved on the series bellow are finite for each $x \in N$.



where $f_\delta, g_\delta, \phi_\delta, d_\delta$ are given by the formulas

$$d_\delta = f\,\delta\,\Sigma_c^\delta\,g; \qquad f_\delta = f\,(1-\delta\,\Sigma_c^\delta\,\phi); \qquad g_\delta = \Sigma_c^\delta\,g; \qquad \phi_\delta = \Sigma_c^\delta\,\phi;$$

and $\Sigma_c^\delta = \sum_{i\geq 0}(-1)^i\,(\phi\delta)^i$.

In particular, if we have a reduction from A to M, where A is a dg-algebra and M is a dg-module, through homological techniques, it is possible to compute the operations of the A_∞-algebra structure induced on M in terms of the product of A. This computation can be done in four steps (described *lightly*):

- To construct of a new reduction $T(s(c)) : \{T(sA), T(sM), Tf, Tg, T\phi\}$
- To use the BPL with perturbation datum a (*simplicial*) differential d_s depending on the product of A (see [5]).
- To extract the operations induced on M from the new differential in $T(sM)$.

Where, given a dg-module (M, d), the *tensor module* of M, $T(M)$, is

$$T(M) = \bigoplus_{n\geq 0} M^{\otimes n}.$$

The differential structure in $T(M)$ is provided by the *tensor differential*, d_t. $T(M)$ is endowed with both structures of dg–algebra and dg–coalgebra respectively, by a product $\mu((a_1\otimes\cdots\otimes a_n)\otimes(a_{n+1}\otimes\cdots\otimes a_{n+p})) = a_1\otimes\cdots\otimes a_{n+p}$; and a coproduct $\Delta(a_1\otimes\cdots\otimes a_n) = \sum_{i=0}^{n}(a_1\otimes\cdots\otimes a_i)\otimes(a_{i+1}\otimes\cdots\otimes a_n)$.

In particular, if A is a dga-algebra, one can construct the *reduced bar construction* of A, $\overline{B}(A)$, whose underlying module is the tensor module of the suspension of A, $T(sA)$.

The total differential $d_{\overline{B}}$ is given by $d_{\overline{B}} = d_t + d_s$, the component d_t being the tensor differential on the tensor module and d_s the *simplicial differential*, that depends on the product on A.

This dg–module is endowed with a structure of dg–coalgebra by the natural coproduct $\Delta_{\overline{B}} : \overline{B}(A) \to \overline{B}(A) \otimes \overline{B}(A)$ defined in the tensor module.

Analogously, if we have a reduction from C to M, where C is a dga-coalgebra and M is a dg-module, through homological techniques, it is possible to compute the operations of the A_∞-coalgebra structure induced in M.

A simple example of this fact is the algebraic structure of the homology groups of the bar construction of a truncated polynomial algebra with coefficients in $\mathbb{Z}$ (see [1]). Let us recall a little bit this example, in fact, let us see the scheme of the algebraic structure of the homology groups.

- There is an explicit reduction $c : \{B(Q_p(u, 2n)), E(v, 2n+1) \otimes \Gamma(\gamma, 2np+2), f, g, \phi\}$.
- Even though the bar construction has a coalgebra structure, this reduction does not preserve this structure, because the morphisms are not compatible with the coproducts; so we obtain an A_∞-coalgebra structure induced on $E \otimes \Gamma$. This structure is extremely simple, because it only has two operations non-null, Δ_2 and Δ_p.
- Thanks to the null differential in $E \otimes \Gamma$, the homology groups of $B(Q_p(u, 2n))$ are isomorphic to $E \otimes \Gamma$.
- So from the algebraic point of view, $H_*(B(Q_p(u, 2n)))$ has an A_∞-coalgebra structure defined by Δ_2 and Δ_p.

In particular, when we are speaking about the categories of algebras or coalgebras, we have to emphasize here that the category of Hopf algebras joins the two notions. But, until now, nothing was known about the analogous appropriate notion of A_∞-Hopf algebra. So, let us give a new way to understand them via perturbation.

3 The importance of the Hopf relation

Let (H, d, μ, Δ) be a *dga-Hopf algebra*. Because of the Hopf relation, it is possible to define a new algebraic object:

Definition 3.1. *Let us define a tensor module associated with H, $BC(H)$, Bar-Cobar of H, as*

$$\{BC(H)\}_{(p,q,n)} = (H^{\otimes p})_n^{\otimes q} = H_n^{p,q},$$

where an element is described as a matrix

$$(a_{ij}) = \begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix}$$

and the degree is $|(a_{ij})| = n + p - q$ where $n = \sum_{i=0,j=0}^{p,q} |a_{ij}|$.

$BC(H)$ is a differential graded module with three differential structures induced,

- the *tensor differential* $\mathbf{d_t}$: $H_n^{p,q} \to H_{n-1}^{p,q}$; $\quad d_t = -\sum_{i=1,j=1}^{p,q} (-1)^{P(i,j)} d_{i,j}$,

  where

$$d_{ij} = -(-1)^{P(i,j)} \begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & \vdots & \vdots \\ a_{i1} & \cdots d(a_{ij}) & a_{iq} \\ \vdots & \vdots & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix}$$

$$P(i,j) = \sum_{k<i \cup (k=i, l<j)} |a_{kl}| + (i-1) - (j-1)$$

- the *simplicial differential* $\mathbf{d_s}$: $H_n^{p,q} \to H_n^{p-1,q}$; $\quad d_s = \sum_{k=0}^{p} (-1)^k \delta_k$, depending on the product of H:

$$\delta_k = (-1)^{sg(k,k+1)+sg_B(k)} \begin{pmatrix} a_{11} & \cdots & a_{1q} \\ \vdots & \vdots & \vdots \\ \mu(a_{k1}, a_{k+1,1}) & \cdots & \mu(a_{kq}, a_{k+1,q}) \\ \vdots & \vdots & \vdots \\ a_{p1} & \cdots & a_{pq} \end{pmatrix}$$

where

$$\begin{aligned} sg(k,k+1) &= |a_{k+1,1}|(|a_{k,2}| + \cdots + |a_{k,q}|) \\ &+ |a_{k+1,2}|(|a_{k,3}| + \cdots + |a_{k,q}|) + \cdots + |a_{k+1,q-1}||a_{k,q}| \\ sg_B(k) &= \sum_{j=1,l=1}^{k,q} |a_{jl}| \end{aligned}$$

- the *cosimplicial* differential $\mathbf{d_c}$: $H_n^{p,q} \to H^{p,q+1}$; $\quad d_c = \sum_{i=0}^{q+1} (-1)^k \delta^k$, depending on the coproduct of H:

$$\delta^k = (-1)^{sg_C(k)} \begin{pmatrix} a_{11} & \cdots & a_{1,k-1} & \Delta(a_{1,k}) & a_{1,k+1} & \cdots & a_{1q} \\ \vdots & & & \vdots & & & \vdots \\ a_{p1} & \cdots & a_{p,k-1} & \Delta(a_{p,k}) & a_{p,k+1} & \cdots & a_{pq} \end{pmatrix}$$

where

$$sg_C(k) = \sum_{j=1,l=1}^{p,k-1} |a_{jl}|.$$

Theorem 3.2. *With the above definitions, the morphism*

$$\mathbf{d_{BC}} = \mathbf{d_t} + \mathbf{d_c} + (-\mathbf{1})^{\mathbf{q}} \mathbf{d_s} : H_n^{p,q} \to H_{n-1}^{p,q} \oplus H_n^{p-1,q} \oplus H_n^{p,q+1}$$

is a differential in the total complex of $BC(H)$.

Definition 3.3. *Given H a connected dga-Hopf algebra, it is possible to define a new algebraic object $\widehat{BC}H$ as a differential graded module*

$$\widehat{BC}(H) = \Lambda \oplus \overline{H}^{1,1} \oplus (\overline{H}^{2,1} \oplus \overline{H}^{1,2}) \oplus \cdots \oplus \sum_{i+j=k} \overline{H}^{i,j} \oplus \cdots$$

with the differentials d_t, d_s and d_c induced in a natural way from $BC(H)$ to $\widehat{BC}(H)$.

In the rest of the paper, given a dg-module M we will denote by $BC(M)$ the tensor module $\{M_n^{p,q}\}_{p,q,n}$ with the tensor differential induced on it. Analogously $\widehat{BC}(M)$ is the differential graded module

$$\widehat{BC}(M) = \Lambda \oplus \overline{M}^{1,1} \oplus (\overline{M}^{2,1} \oplus \overline{M}^{1,2}) \oplus \cdots \oplus \sum_{i+j=k} \overline{M}^{i,j} \oplus \cdots$$

with the tensor differential induced.

4 What about A_∞-Hopf algebras?

We can now formulate our main results related with A_∞-Hopf algebras. To start with, we make the following definition.

Definition 4.1. *An A_∞-Hopf algebra M is a dg-module with a family of operations*

$$h^{i,j} : M^{\otimes i} \to M^{\otimes j}$$

with $i, j \in \mathbb{N}$, of degree $i + j - 3$, such that

- *The family $\{h^{i,1}\} : M^{\otimes i} \to M$ defines an A_∞-algebra on M.*
- *The family $\{h^{1,j}\} : M \to M^{\otimes j}$ defines an A_∞-coalgebra on M.*
- *The extension of $\{h^{i,j}\}_{i,j\in\mathbb{N}}$ to $(\widehat{BC}(M), d_t)$ defines a differential on it.*

If we consider a reduction from H to M, where H is a dga-Hopf algebra and M is a dg-module, then the information about the A_∞-algebra structure induced on M and the information about the A_∞-coalgebra structure can be extracted, as well as the operations defining an A_∞-Hopf algebra structure.

Proposition 4.2. *Given H a Hopf algebra and $c : H \to M$ a reduction, this induces a new one of dg-module*

$$bc(c) : \{(BC(H), d_t), (BC(M), d_t), bc(f), bc(g), bc(\phi)\},$$

where the morphisms $bcf, bcg, bc\phi$ are defined by the formulas

$$\begin{aligned} bc(f)_{|BC(H)_n} &= \underbrace{f \otimes \cdots \otimes f}_{n \text{ times}}; \\ bc(g)_{|BC(M)_n} &= \underbrace{g \otimes \cdots \otimes g}_{n \text{ times}}; \\ bc(\phi)_{|BC(H)_n} &= \sum_{k=0}^{n-1} \underbrace{1 \otimes \cdots \otimes 1}_{k \text{ times}} \otimes \phi \otimes \underbrace{gf \otimes \cdots \otimes gf}_{n-k-1 \text{ times}}. \end{aligned}$$

Theorem 4.3. *Given $c : H \to M$ a reduction, where H is a simply connected dga-Hopf algebra and M is a dg-module, if we consider the reduction defined in proposition 4.2, $bc(c) : \{(\widehat{BC}(H), d_t), (\widehat{BC}(M), d_t), bc(f), bc(g), bc(\phi)\}$ together with the perturbation datum $d_c + (-1)^q d_s$, thanks to the basic perturbation lemma, it is possible to define a new reduction*

$$\begin{aligned} bc(c)_{d_c+(-1)^q d_s} :\{&(\widehat{BC}(H), d_t + d_c + (-1)^q d_s), (\widehat{BC}(M), d_t + d_\infty), \\ &bc(f)_\infty, bc(g)_\infty, bc(\phi)_\infty\} \end{aligned}$$

such that M inherits an A_∞-Hopf algebra structure, i.e.:

- *the projection over the elements of $\widehat{BC}(M)_*^{1,*}$, gives the A_∞-algebra of M;*
- *the projection over the elements of $\widehat{BC}(M)_*^{*,1}$, gives the A_∞-coalgebra of M;*

- *the homotopy operators of higher order of M are operations $h^{i,j}$: $M^{\otimes i} \to M^{\otimes j}$ of degree $i+j-3$, with $i>1, j>1$, defined by the formulas of the BPL.*

Indeed, the explicit formulas of $h^{i,j}$ are

$$h^{i,j} = f\delta(-1)^{i-1}\sigma(\underbrace{(\phi\delta)^{i-2}(\phi\delta')^{j-1}}_{\text{Permutations}})g + f\delta'(-1)^{i-1}\sigma(\underbrace{(\phi\delta)^{i-1}(\phi\delta')^{j-2}}_{\text{Permutations}})g$$

We obtain automatically the next result

Corollary 4.4. *Given H a simply connected dga-Hopf algebra with product μ, coproduct Δ, M a dga-module and $c : \{H, M, f, g, \phi\}$ a reduction between them. Then, M inherits an A_∞-Hopf algebra structure.*

5 Summary

As we have just seen, if we are interested in the algebraic structure of the homology of a Hopf algebra, in general, we know that this structure will not be a Hopf algebra, but yes an A_∞-Hopf algebra, that we can determine with simple methods explicitly, thanks to the perturbation theory.

References

[1] Berciano A. and Real P.: A_∞-coalgebra structure on the $\mathbb{Z}_p$-homology of Eilenberg-Mac Lane spaces, *Proceedings EACA*, 2004.

[2] Brown R.: The twisted Eilenberg-Zilber theorem, *Celebrazioni Archimedae del Secolo XX, Simposio di Topologia*, pp. 34–37, 1967.

[3] Gugenheim V. K. A. M., Lambe L. and Stasheff J.: Perturbation theory in Differential Homological Algebra, II. *Illinois J. Math.* **35** n. 3, pp. 357-373, 1991.

[4] Huebschmann J. and Kadeishvili T.: Small models for chain algebras. *Math. Zeit.* **v. 207**, pp. 245-280, 1991.

[5] Jiménez M.J.: A_∞–estructuras y perturbación homológica, *PhD Thesis of Universidad de Sevilla*, 2003.

[6] Rubio J. and Sergeraert F.: Constructive algebraic topology, *Bull. Sci. Math*, **126**, pp. 389-412, 2002.

[7] Saneblidze S. and Umble R.: A_∞-Hopf algebras, preprint.
[8] Saneblidze S. and Umble R.: The biderivative and A_∞-bialgebras, *Homology, Homotopy and Applications* **7** (2), pp. 161-177, 2005.
[9] Saneblidze S. and Umble R.: Diagonals on the permutahedra, multiplihedra and Associahedra, *Homology, Homotopy and Applications*, **6** (1), pp. 363-411, 2004.
[10] Stasheff J.D.: Homotopy associativity of H-spaces II, *Transactions of A.M.S.* **108**, pp. 293-312, 1963.

On integrability of the Euler–Poisson equations

A.D. Bruno [1] V.F. Edneral [2]

[1] Keldysh Institute of Applied Mathematics RAS
Moscow
[2] Skobeltsyn Institute of Nuclear Physics of
Lomonosov Moscow State University
edneral@theory.sinp.msu.ru

Abstract

We consider a special case of the Euler–Poisson system of equations, describing motion of a rigid body around a fixed point. We find 44 sets of stationary solutions near which the system is locally integrable. 10 of them are real. We study also number of these complex stationary solutions in 3-dimensional invariant manifolds of the system. We find that the number is 4, 2, 1 or 0.

Keywords: resonant normal form, Euler–Poisson equations, local integrability.

1 Local integrability

Now all cases of global integrability of classical ODE systems such as Euler–Poisson equations are known. So we are looking for cases of local integrability. The first such attempt was made by authors in [1], but it was not successful [2]. In [2] we also proposed new approach to the problem and found the first case of the local integrability of the Euler–Poisson equations. Here we develop the approach and find many new sets of such stationary solutions of the considered system near which the system is locally integrable.

Denote $\dot{} \stackrel{\text{def}}{=} d/dt$, $X = (x_1, \ldots, x_n)$. Autonomous system

$$\dot{x}_j = \varphi_j(X),\ j = 1, \ldots, n \tag{1.1}$$

with polynomials $\varphi_j(X)$ is *locally integrable* in domain $D \subset \mathbb{C}^n$, if in the domain it has enough number of independent first integrals of the form $a(X)/b(X)$, where functions $a(X)$ and $b(X)$ are analytic in the domain D. Certainly, in a neighborhood of a nonstationary point X^0, where $\Phi(X^0) \neq 0$, $\Phi \stackrel{\text{def}}{=} (\varphi_1, \ldots, \varphi_n)$, System (1.1) is integrable. So the question on integrability has a sense for domains, containing or adjoining a singularity: a stationary solution or a periodic solution and so on. Thus, we must study *local integrability* and *local nonintegrability* of a given system near its singularities. The best tool for that is normal form [3, 4, 5].

The paper is organized as follows. In Section 2 we give a short survey of the normal form theory. In Section 3 we describe the Euler–Poisson system of equations and select its special case which we consider here. In the case the system has two pairs of two-parameter families of stationary solutions (points). In Section 4 we consider the first pair of the families. In them we select all such points, near which we can guarantee the local integrability of the system. They form 12 sets, and 6 of them are real. In Section 5 we consider the second pair of the families and find 32 sets with the property of local integrability, 4 of them are real. In Section 6 we consider stationary points contained in a 3-dimensional invariant manifold and study a number of them with the property of local integrability. We used the package MATHEMATICA for most complicated computation.

2 Normal form

Below, vectors are lines, and asterisk * means transposition. Let $X^0 = 0$ be a stationary solution of System (1.1), i.e. $\Phi(0) = 0$:

$$\dot{X}^* = AX^* + \tilde{\Phi}^*(X). \tag{2.1}$$

Let the linear transformation

$$X^* = BY^* \tag{2.2}$$

bring the matrix A to the Jordan form $J = B^{-1}AB$ and (2.1) to

$$\dot{Y}^* = JY^* + \tilde{\tilde{\Phi}}^*(Y). \tag{2.3}$$

Let the formal change of coordinates

$$Y = Z + \Xi(Z), \tag{2.4}$$

where $\Xi = (\xi_1, \ldots, \xi_n)$ and $\xi_j(Z)$ are formal power series, transform (2.3) in the system

$$\dot{Z}^* = JZ^* + \Psi^*(Z). \tag{2.5}$$

We write it in the form

$$\dot{z}_j = z_j g_i(Z) = z_j \sum g_{jQ} Z^Q \ \text{ over } \ Q \in \mathbb{N}_j,\ j = 1, \ldots, n, \tag{2.6}$$

where $Q = (q_1, \ldots, q_n)$, $Z^Q = z_1^{q_1} \ldots z_n^{q_n}$,

$$\mathbb{N}_j = \{Q : Q \in \mathbb{Z}^n,\ \ Q + E_j \geq 0\},\ j = 1, \ldots, n,$$

E_j means the unit vector. Denote

$$\mathbb{N} = \mathbb{N}_1 \cup \ldots \cup \mathbb{N}_n. \tag{2.7}$$

The diagonal $\Lambda = (\lambda_1, \ldots, \lambda_n)$ of J consists of eigenvalues of the matrix A.

System (2.5), (2.6) is called the *resonant normal form* if:

a) J is the Jordan matrix,

b) in writing (2.6), there are only the *resonant terms*, for which the scalar product

$$\langle Q, \Lambda \rangle \stackrel{\text{def}}{=} q_1\lambda_1 + \ldots + q_n\lambda_n = 0. \tag{2.8}$$

Theorem 1. *There exists a formal change* (2.4) *reducing* (2.3) *to its normal form* (2.5), (2.6).

Let k be the number of linearly independent solutions $Q \in \mathbb{N}$ to Eq. (2.8), it is called the *multiplicity of resonance*. Integration of the normal form (2.6) is reduced to solving a system of order k respectively k *resonant variables*.

Property 1. If System (2.3) has the linear automorphism $\tilde{t} = \delta t$, $\tilde{Y}^* = SY^*$, then its normal form (2.6) has the same automorphism $\tilde{t} = \delta t$, $\tilde{Z}^* = SZ^*$.

In [3] there are conditions on the normal form (2.6), which guarantee the convergence of the normalizing transformation (2.4).

Condition A. *In the normal form* (2.6)

$$g_j = \lambda_j \alpha(Z) + \bar{\lambda}_j \beta(Z), \quad j = 1, \ldots, n,$$

where $\alpha(Z)$ *and* $\beta(Z)$ *are some power series.*

Let

$$\omega_k = \min |\langle Q, \Lambda \rangle| \text{ over } Q \in \mathbb{N}, \ \langle Q, \Lambda \rangle \neq 0, \ \sum_{j=1}^{n} q_j < 2^k, \ k = 1, 2, \ldots$$

Condition ω (on small divisors). *The series*

$$\sum_{k=1}^{\infty} 2^{-k} \log \omega_k > -\infty,$$

i.e. it converges.

It is fulfilled for almost all vectors Λ.

Theorem 2. *If vector* Λ *satisfies Condition* ω *and the normal form* (2.6) *satisfies Condition A then the normalizing transformation* (2.4) *converges.*

3 The Euler–Poisson equations

Motions of a rigid body around a fixed point is described by the system of six Euler–Poisson equations [6]

$$\begin{aligned} A\dot{p} + (C-B)qr &= Mg(z_0\gamma_2 - y_0\gamma_3),\\ B\dot{q} + (A-C)pr &= Mg(x_0\gamma_3 - z_0\gamma_1),\\ C\dot{r} + (B-A)pq &= Mg(y_0\gamma_1 - x_0\gamma_2), \end{aligned} \tag{3.1}$$

$$\dot{\gamma}_1 = r\gamma_2 - q\gamma_3, \quad \dot{\gamma}_2 = p\gamma_3 - r\gamma_1, \quad \dot{\gamma}_3 = q\gamma_1 - p\gamma_2, \tag{3.2}$$

where $A, B, C, M, g, x_0, y_0, z_0$ are real constants and A, B, C are positive and satisfy the triangle inequalities. System (3.1), (3.2) has 3 first integrals

$$\begin{aligned} &h \stackrel{\text{def}}{=} Ap^2 + Bq^2 + Cr^2 + 2Mg(x_0\gamma_1 + y_0\gamma_2 + z_0\gamma_3) = \text{const},\\ &g \stackrel{\text{def}}{=} Ap\gamma_1 + Bq\gamma_2 + Cr\gamma_3 = \text{const},\\ &\gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1. \end{aligned} \tag{3.3}$$

We consider the Euler–Poisson system (3.1), (3.2) in the case

$$A = B = 1, \quad C = c, \quad Mgx_0 = -1, \quad y_0 = z_0 = 0.$$

Then System (3.1) becomes

$$\dot{p} = (1-c)qr, \quad \dot{q} = (c-1)pr - \gamma_3, \quad \dot{r} = \gamma_2/c. \tag{3.4}$$

System (3.4), (3.2) has the linear automorphism

$$t, p, q, r, \gamma_1, \gamma_2, \gamma_3 \to -t, p, -q, r, \gamma_1, -\gamma_2, \gamma_3. \tag{3.5}$$

The unique parameter $c \in (0, 2]$. System (3.5), (3.2) has three first integrals

$$\begin{aligned} &h \stackrel{\text{def}}{=} p^2 + q^2 + cr^2 - 2\gamma_1 = \text{const},\\ &g \stackrel{\text{def}}{=} p\gamma_1 + q\gamma_2 + cr\gamma_3 = \text{const},\\ &I_3 \stackrel{\text{def}}{=} \gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1 \end{aligned} \tag{3.6}$$

and it is integrable in quadratures if it has one additional first integral, which is known in two cases: $c = 1$ (case of Lagrange) and $c = 1/2$ (case of S. Kovalevskaya). System (3.4), (3.2) is not integrable globally for other values c [7].

Theorem 3. *System* (3.4), (3.2) *has two pairs of two-parameter families of stationary solutions:*

$$\mathbf{S}_\sigma\colon\ p\ =\ p_0 \in \mathbb{C},\ q = 0,\ r = 0,\ \gamma_1 = \sigma,\ \gamma_2 = 0,\ \gamma_3 = 0,\ \sigma = \pm 1; \tag{3.7}$$

$$\begin{aligned} \mathbf{T}_\tau\colon\ p\ &=\ p_t \in \mathbb{C},\ q = 0,\ r_t = \tau\sqrt{1-(c-1)^2p_t^4}/((c-1)p_t), \\ \gamma_{1t}\ &=\ (c-1)p_t^2,\ \gamma_2 = 0,\ \gamma_{3t} = \tau\sqrt{1-(c-1)^2p_t^4},\ \tau = \pm 1; \end{aligned} \tag{3.8}$$

and all its stationary solutions belong to these families.

Families $\mathbf{S}_\sigma$ exist for any $p_0 \in \mathbb{C}$. Families $\mathbf{T}_\tau$ exist for $c \neq 1$ and $p_t \neq 0$. Intersections of the families are

$$\begin{gathered} \mathbf{S}_+ \cap \mathbf{S}_- = \varnothing, \\ \mathbf{T}_+ \cap \mathbf{T}_- = \mathbf{S}_\sigma \cap \mathbf{T}_\tau :\ p_0 = p_t,\ q = 0,\ r = 0, \\ \gamma_1 = (c-1)p_t^2 = \sigma = \pm 1,\ \gamma_2 = 0,\ \gamma_3 = 0. \end{gathered} \tag{3.9}$$

Families $\mathbf{S}_\sigma$ are real if $p_0 \in \mathbb{R}$. Families $\mathbf{T}_\tau$ are real if $p_t \in \mathbb{R}$ and $(c-1)^2p_t^4 \leq 1$.

4 Families $\mathbf{S}_\sigma$

First we will consider families $\mathbf{S}_\sigma$. In a neighborhood of each stationary point (3.7) we introduce local coordinates

$$P = p - p_0,\ \ q, r,\ \ \Gamma = \gamma_1 - \sigma, \gamma_2, \gamma_3. \tag{4.1}$$

In them System (3.4), (3.2) takes the form

$$\begin{aligned} &\dot P = (1-c)qr, \\ &\dot q = (c-1)p_0 r - \gamma_3 + (c-1)Pr, \\ &\dot r = \gamma_2/c, \\ &\dot\Gamma = r\gamma_2 - q\gamma_3, \\ &\dot\gamma_2 = -\sigma r + p_0\gamma_3 + P\gamma_3 - r\Gamma, \\ &\dot\gamma_3 = \sigma q - p_0\gamma_2 + q\Gamma - P\gamma_2. \end{aligned} \tag{4.2}$$

If coordinates (4.1) are denoted as $X = (x_1, \ldots, x_6)$:

$$x_1 = P, \quad x_2 = q, \quad x_3 = r, \quad x_4 = \Gamma, \quad x_5 = \gamma_2, \quad x_6 = \gamma_3, \tag{4.3}$$

then (4.2) is System (2.1): $X^* = AX^* + \tilde{\Phi}^*(X)$ with $n = 6$, where

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (c-1)p_0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1/c & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\sigma & 0 & 0 & p_0 \\ 0 & \sigma & 0 & 0 & -p_0 & 0 \end{pmatrix}. \tag{4.4}$$

Characteristic equation for matrix A is

$$\lambda^6 + a\lambda^4 + b\lambda^2 = 0, \tag{4.5}$$

where

$$a = \left(\sigma p_0^2 + 1 + 1/c\right)\sigma, \quad b = 1/c + \sigma p_0^2\,(1/c - 1)\,. \tag{4.6}$$

Eq. (4.5) has two zero roots and twin roots

$$\lambda_1 = \lambda_2 = 0, \quad \lambda_3 = -\lambda_4, \quad \lambda_5 = -\lambda_6. \tag{4.7}$$

Let $\tilde{\mathbf{S}}_\sigma$ be the part of the family $\mathbf{S}_\sigma$ with $p_0^2 \in \mathbb{R}$. Evidently $\tilde{\mathbf{S}}_\sigma \supset \operatorname{Re}\mathbf{S}_\sigma$. Below we will consider subfamilies $\tilde{\mathbf{S}}_\sigma$ only. The dependence of eigenvalues $\lambda_3, \lambda_4, \lambda_5, \lambda_6$ from two real parameters

$$c \in (0, 2], \quad y \stackrel{\text{def}}{=} \sigma p_0^2 \in \mathbb{R} \tag{4.7$'$}$$

is represented in Fig. 1. Three curves

$$C_1 = \{(cy + c - 1)^2 + 4c^2 y = 0\} = \{a^2 - 4b = 0\},$$

$$C_2 \cup C_3 = \{y(c-1) = 1\} = \{b = 0\}$$

divide the strip (4.7′) into five sets D_1, D_2, D_3, D_4, D_5. Curve C_1 touches the curve C_3 in the point

$$(c_0, y_0) = ((\sqrt{5} - 1)/2, -(\sqrt{5} + 3)/2) \approx (0.618, -2.618).$$

In the set D_1 eigenvalues $\lambda_3, \lambda_4, \lambda_5, \lambda_6$ are complex: $\lambda_4 = -\lambda_3$, $\lambda_5 = \bar{\lambda}_3$, $\lambda_6 = -\bar{\lambda}_3$; in sets D_2 and D_3 two of them are real and another two are pure imaginary, in D_4 they are pure imaginary, in D_5 they are real.

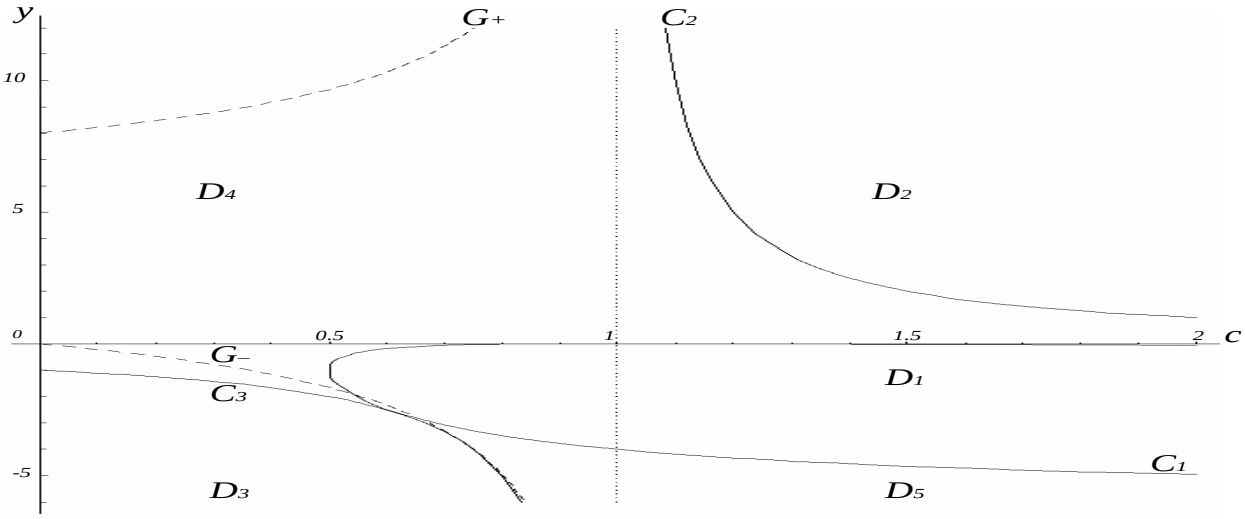


Figure 1: Sets of $\mathbf{S}_\sigma$ families.

Accordingly (3.5) and (4.1), System (4.2) has the automorphism

$$t, P, q, r, \Gamma, \gamma_2, \gamma_3 \to -t, P, -q, r, \Gamma, -\gamma_2, \gamma_3$$

or accordingly (4.3)

$$t, x_1, x_2, x_3, x_4, x_5, x_6 \to -t, x_1, -x_2, x_3, x_4, -x_5, x_6. \tag{4.8}$$

Let the linear transformation

$$X^* = BY^*$$

brings the matrix (4.4) to the diagonal form in sets D_1–D_5. Then it brings the automorphism (4.8) to the form

$$t, y_1, y_2, y_3, y_4, y_5, y_6 \to -t, y_1, y_2, y_4, y_3, y_6, y_5. \tag{4.9}$$

Theorem 4. *In sets D_1, D_2, D_3 the normalizing transformation of System* (4.2) *converges.*

Indeed, in these sets the ratio λ_3/λ_5 is not a real number. So the equation

$$\langle \Lambda, Q \rangle = 0$$

with vector $\Lambda = (0, 0, \lambda_3, \lambda_4, \lambda_5, \lambda_6)$ has only such real solutions $Q = (q_1, \ldots, q_6)$, where q_1, q_2 are arbitrary; $q_3 = q_4$, $q_5 = q_6$. Hence, in the normal form

$$\dot{z}_j = z_j g_j(Z), \quad j = 1, \ldots, 6 \tag{4.10}$$

series $g_j(Z)$ depend only on

$$z_1, \quad z_2, \quad \rho_1 \stackrel{\text{def}}{=} z_3 z_4, \quad \rho_2 \stackrel{\text{def}}{=} z_5 z_6, \tag{4.11}$$

which here are *resonant variables*. According to Property 1, the normal form (4.10) has the automorphism (4.9) where y_j are replaced by z_j. For the resonant variables (4.11) it gives the automorphism

$$t, \ z_1, \ z_2, \ \rho_1, \ \rho_2 \to -t, \ z_1, \ z_2, \ \rho_1, \ \rho_2. \tag{4.12}$$

Hence, in the normal form (4.10)

$$g_1 = g_2 \equiv 0, \quad g_4 = -g_3, \quad g_6 = -g_5. \tag{4.13}$$

It is not difficult to show that the normal form (4.10), (4.13) satisfies Condition A. Here Condition ω is also satisfied, because ratio λ_3/λ_5 is not a real number. According to Theorem 2, the normalizing transformation converges.

For resonant variables z_1, z_2, ρ_1, ρ_2, the normal form (4.10), (4.13) gives equations $\dot{z}_j = 0$, $\dot{\rho}_j = 0$, $j = 1, 2$. Hence, the normal form (4.10), (4.13) has four independent first integrals

$$z_1 = \text{const}, \quad z_2 = \text{const}, \quad \rho_1 = \text{const}, \quad \rho_2 = \text{const}. \tag{4.14}$$

As the normalizing transformation is analytic and invertible, then near sets D_1, D_2, D_3 System (4.2) has 4 local first integrals. Hence, it is locally integrable near these sets.

For fixed value $y = \sigma p_0^2$, parameters σ and p_0 have 4 variants of different sets of signs. So in two subfamilies $\tilde{\mathbf{S}}_\sigma$, sets D_1, D_2, D_3 give 12 different complex sets of the local integrability. If value $p_0 = \pm\sqrt{y\sigma}$ is real, then System (4.2) is real as well, and its normal form is real in appropriate coordinates, and 4 first integrals (4.14) are real. If $\sigma = +1$, value $p_0 = \pm\sqrt{y}$ is real for $y \geq 0$, i.e. the set D_2 is real. If $\sigma = -1$, value $p_0 = \pm\sqrt{-y}$ is real for $y \leq 0$, i.e. the sets D_1 and D_3 are real. Alltogether we have 6 real sets of the local integrability.

5 Families $\mathbf{T}_\tau$

Now we will consider families $\mathbf{T}_\tau$. In a neighborhood of each stationary point (3.8) we introduce local coordinates

$$P = p - p_t, \quad q, R = r - r_t, \quad \Gamma = \gamma_1 - \gamma_{1t}, \gamma_2, \Delta = \gamma_3 - \gamma_{3t}. \tag{5.1}$$

In them System (3.4), (3.2) takes the form

$$\begin{aligned}
&\dot{P} = (1-c)r_t q + (1-c)qP,\\
&\dot{q} = (c-1)(p_t R + r_t P) - \Delta + (c-1)PR,\\
&\dot{R} = \gamma_2/c,\\
&\dot{\Gamma} = r_t\gamma_2 - q\gamma_{3t} + R\gamma_2 - q\Delta,\\
&\dot{\gamma}_2 = p_t\Delta + P\gamma_{3t} - (r_t\Gamma + R\gamma_{1t}) + P\Delta - R\Gamma,\\
&\dot{\Delta} = q\gamma_{1t} - p_t\gamma_2 + q\Gamma - P\gamma_2.
\end{aligned} \tag{5.2}$$

If coordinates (5.1) are denoted as $X = (x_1, \ldots, x_6)$:

$$x_1 = P, \quad x_2 = q, \quad x_3 = R, \quad x_4 = \Gamma, \quad x_5 = \gamma_2, \quad x_6 = \Delta, \tag{5.3}$$

then (5.2) is System (2.1): $X^* = AX^* + \tilde{\Phi}^*(X)$ with $n = 6$, where

$$A = \begin{pmatrix}
0 & (1-c)r_t & 0 & 0 & 0 & 0\\
(c-1)r_t & 0 & (c-1)p_t & 0 & 0 & -1\\
0 & 0 & 0 & 0 & 1/c & 0\\
0 & -\gamma_{3t} & 0 & 0 & r_t & 0\\
\gamma_{3t} & 0 & -\gamma_{1t} & -r_t & 0 & p_t\\
0 & \gamma_{1t} & 0 & 0 & -p_t & 0
\end{pmatrix}. \tag{5.4}$$

Characteristic equation for matrix A is

$$\lambda^6 + a\lambda^4 + b\lambda^2 = 0, \tag{5.5}$$

where

$$\begin{gathered} c_1 \stackrel{\text{def}}{=} c - 1, \quad a = \left(1 + \frac{1}{c_1^2}\right)\frac{1}{p_t^2} - \frac{c_1(c_1^2 - 2)p_t^2}{1 + c_1}, \\ b = \frac{1}{c_1^2 p_t^4} + 2 - \frac{3}{1 + c_1} - \frac{3c_1^3 p_t^4}{1 + c_1}. \end{gathered} \tag{5.6}$$

Eq. (5.5) has two zero roots and twin roots

$$\lambda_1 = \lambda_2 = 0, \quad \lambda_3 = -\lambda_4, \quad \lambda_5 = -\lambda_6. \tag{5.7}$$

Let $\tilde{\mathbf{T}}_\tau$ be the part of the family $\mathbf{T}_\tau$ with $p_t^2 \in \mathbb{R}$. Evidently $\tilde{\mathbf{T}}_\tau \supset \operatorname{Re}\mathbf{T}_\tau$. Below we will consider subfamilies $\tilde{\mathbf{T}}_\tau$ only. The dependence of eigenvalues $\lambda_3, \lambda_4, \lambda_5, \lambda_6$ from two real parameters

$$c_1 = c - 1 \in (-1, 1], \quad V = c_1^2 p_t^4 > 0, \tag{5.8}$$

i.e. $p_t^2 \in \mathbb{R} \setminus \{0\}$, is represented in Fig. 2. Four curves

$$\begin{gathered} E_1 = \{V = 1\}, \\ E_2 = \{V = -\frac{c_1 + 1}{3c_1}\} \cap \{V > 0\}, \end{gathered}$$

$$\begin{gathered} E_3 \cup E_4 = \{(c_1 - 1)^2(c_1 + 1)^4 - 2c_1(c_1 - 1)(c_1 + 1)(2 + 4c_1 + \\ + c_1^2 + c_1^3)V + c_1^2(4 + 12c_1 + 8c_1^2 + c_1^4)V^2 = 0\} \cap \{V > 0\} \end{gathered} \tag{5.8′}$$

divide the halfstrip (5.8) into 8 sets F_1–F_8. Here $E_1 \cup E_2 = \{b = 0\} \cap \{V > 0\}$ and $E_3 \cup E_4 = \{a^2 - 4b = 0\} \cap \{V > 0\}$. Curves E_1–E_4 were drown by solid lines because they divide the halfstrip into sets F_1–F_8. Their vertical asymptotes are given by point lines; they have $c_1 = 0, -0.5188\ldots, -0.8225\ldots$ which are roots of the polynomial coefficient for V^2 in (5.8′). Curves E_1 and E_2 intersect in the point $c_1 = -1/4$, $V = 1$.

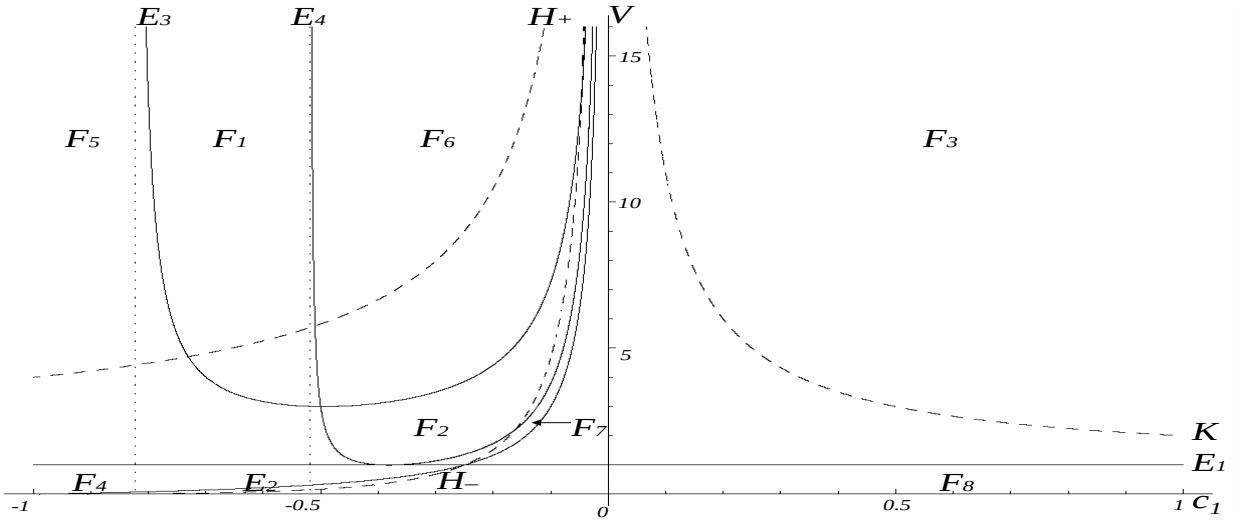


Figure 2: Sets of $\mathbf{T}_\tau$ families.

Curves E_3 and E_4 intersect in the point $c_1 = -0.5$, $V = 3$, because solution of Equation (5.8′) is

$$Vc_1(4 + 12c_1 + 8c_1^2 + c_1^4) = (c_1^2 - 1)(2 + 4c_1 + c_1^2 + c_1^3) \pm \\ \pm 2(c_1^2 - 1)(1 + 2c_1)\sqrt{-c_1}.$$

Curve F_4 touches the straight line $E_1 = \{V = 1\}$ in the point with $c_1 = (\sqrt{5} - 3)/2 = -0.3819\ldots$. Subfamilies $\tilde{\mathbf{T}}_\tau$ intersect subfamilies $\tilde{\mathbf{S}}_\sigma$ along the line E_1 in Fig. 2 and along curves C_2 and C_3 in Fig. 1; and $\sigma = \operatorname{sgn} c_1$ according to (3.9). The point $(c - 1, V) = ((\sqrt{5} - 3)/2, 1)$ corresponds to the point $(c, y) = (c_0, y_0) = ((\sqrt{5} - 1)/2, (\sqrt{5} + 3)/2)$.

In the sets F_1 and F_2 eigenvalues $\lambda_3, \lambda_4, \lambda_5, \lambda_6$ are complex: $\lambda_4 = -\lambda_3$, $\lambda_5 = \bar{\lambda}_3$, $\lambda_6 = -\bar{\lambda}_3$; in sets F_3 and F_4 two of them are real and another two are pure imaginary.

Accordingly (3.5) and (5.1), System (5.2) has the automorphism

$$t, P, q, R, \Gamma, \gamma_2, \Delta \to -t, P, -q, R, \Gamma, -\gamma_2, \Delta. \tag{5.9}$$

In coordinates (5.3) it is the automorphism (4.8). Let the linear transformation

$$X^* = BY^*$$

brings the matrix (5.4) to the diagonal form in sets F_1–F_8. Then it brings the automorphism (4.8) to the form

$$t, y_1, y_2, y_3, y_4, y_5, y_6 \to -t, y_1, y_2, y_4, y_3, y_6, y_5. \tag{5.10}$$

Theorem 5. *In sets F_1–F_4 the normalizing transformation of System* (5.2) *converges.*

Indeed, in these sets the ratio λ_3/λ_5 is not a real number. So the equation

$$\langle \Lambda, Q \rangle = 0$$

with vector $\Lambda = (0, 0, \lambda_3, \lambda_4, \lambda_5, \lambda_6)$ has only such real solutions $Q = (q_1, \ldots, q_6)$, where q_1, q_2 are arbitrary; $q_3 = q_4$, $q_5 = q_6$. Hence, in the normal form

$$\dot{z}_j = z_j g_j(Z), \quad j = 1, \ldots, 6 \tag{5.11}$$

series $g_j(Z)$ depend only on

$$z_1, \quad z_2, \quad \rho_1 \stackrel{\text{def}}{=} z_3 z_4, \quad \rho_2 \stackrel{\text{def}}{=} z_5 z_6, \tag{5.12}$$

which here are *resonant variables*. According to Property 1, the normal form (5.11) has the automorphism (5.10) where y_j are replaced by z_j. For the resonant variables (5.12) it gives the automorphism

$$t, \ z_1, \ z_2, \ \rho_1, \ \rho_2 \to -t, \ z_1, \ z_2, \ \rho_1, \ \rho_2. \tag{5.13}$$

Hence, in the normal form (5.11)

$$g_1 \equiv g_2 \equiv 0, \quad g_4 = -g_3, \quad g_6 = -g_5. \tag{5.14}$$

It is not difficult to show that the normal form (5.11), (5.14) satisfies Condition A. Here Condition ω is also satisfied, because ratio λ_3/λ_5 is not a real number. According to Theorem 2, the normalizing transformation converges.

For resonant variables z_1, z_2, ρ_1, ρ_2, the normal form (5.11), (5.14) gives equations $\dot{z}_j = 0$, $\dot{\rho}_j = 0$, $j = 1, 2$. Hence, the normal form (5.11), (5.14) has four independent first integrals

$$z_1 = \text{const}, \quad z_2 = \text{const}, \quad \rho_1 = \text{const}, \quad \rho_2 = \text{const}. \tag{5.15}$$

As the normalizing transformation is analytic and invertible, then near sets F_1, F_2, F_3, F_4 System (5.2) has 4 local first integrals. Hence, it is locally integrable near these sets.

For fixed values c_1, V from the halfstrip (5.8), parameter $\tau = \pm 1$ and parameter $p_t = \sqrt[4]{V/c_1^2}$ and they have alltogether 8 different pairs of values. So, in two subfamilies $\tilde{\mathbf{T}}_\tau$ sets F_1–F_4 give 32 different complex sets of the local integrability for $\tau = \pm 1$ and real $V > 0$. If the point (3.8) is real, then $V \leq 1$ and System (5.2) is real as well, and its normal form is real in appropriate coordinates, and 4 first integrals (5.15) are real. Real parameters $\tau = \pm 1$ and $p_t = \pm|p_t|$ have 4 variants of different sets of signs. Alltogether we have 4 real copies of the set F_4 of the local integrability.

6 Stationary points in invariant manifolds

Here we consider the question: How many locally integrable stationary points from subfamilies $\tilde{\mathbf{S}}_\sigma$ and $\tilde{\mathbf{T}}_\tau$ can be in one real invariant manifold $M = \{h = \text{const},\ g = \text{const}\}$ (see (3.6))? According to (3.7), on families $\mathbf{S}_\sigma$ integrals are

$$h = p_0^2 - 2\sigma, \quad g = \sigma p_0. \tag{6.1}$$

Hence,

$$\sigma = (g^2 - h)/2, \quad p_0 = \sigma g = g(g^2 - h)/2. \tag{6.2}$$

So, values σ and p_0 are uniquely determined by a pair (h, g), i.e. each pair of values (h, g), connected by $h = g^2 - 2\sigma$, corresponds to only one point from subfamilies $\tilde{\mathbf{S}}_\sigma$. Thus, one manifold M contains no more than one point from $\tilde{\mathbf{S}}_\sigma$.

On families $\mathbf{T}_\tau$ integrals (3.7) are

$$h = \frac{1 + c_1}{c_1^2 p_t^2} - 3c_1 p_t^2, \quad g = \frac{1 + c_1}{c_1 p_t} - c_1^2 p_t^3. \tag{6.3}$$

Resolving that system with respect to p_t, we obtain

$$p_t = \frac{h^2 g c_1 + 12g(1 + c_1)}{c_1[h^3 c_1 + 18g^2 - 4h(1 + c_1)]}. \tag{6.4}$$

Hence, a pair (h, g) with $g \neq 0$ determines the value p_t uniquely. And $\tau = \pm 1$ is arbitrary. So, the pair (h, g) corresponds to two points in subfamilies $\tilde{\mathbf{T}}_\tau$: one in $\tilde{\mathbf{T}}_+$ and another in $\tilde{\mathbf{T}}_-$, both with the same p_t.

If $g = 0$, then from (6.3)

$$V = \frac{1 + c_1}{c_1} = 1 + \frac{1}{c_1}. \tag{6.5}$$

As $V > 0$ then here $c_1 > 0$,

$$\sqrt{V} = c_1 p_t^2 = \sqrt{\frac{1 + c_1}{c_1}} > 0 \text{ and } h = -2\sqrt{\frac{1 + c_1}{c_1}} \tag{6.6}$$

and

$$p_t = \pm\sqrt{\frac{1}{c_1}\sqrt{\frac{1 + c_1}{c_1}}}. \tag{6.7}$$

So, the pair $(h, g = 0)$ with $h = -2\sqrt{(1 + c_1)/c_1}$ corresponds to 4 points in subfamilies $\tilde{\mathbf{T}}_\tau$:

$$p_t = \pm|p_t|, \;\; \tau = \pm 1.$$

We denote the curve (6.5) as K. It is shown in Fig. 2. It belongs to the set F_2. Corresponding manifolds M have four locally integrable stationary points. Accordingly (6.1), on families $\mathbf{S}_\sigma$ we have $h = -2\sigma$ for $g = 0$. So if subfamily of solutions (6.6) intersects families $\mathbf{S}_\sigma$ than $\sqrt{(1 + c_1)/c_1} = \sigma$, i.e. $1 + c_1 = c_1$. But it is impossible. Hence, the subfamily (6.6) (which is denoted as K) does not intersect families $\mathbf{S}_\sigma$ and all its stationary points are from $\tilde{\mathbf{T}}_\tau$ and are locally integrable.

Now let us consider possibility that a manifold M has one locally integrable point from subfamilies $\tilde{\mathbf{S}}_\sigma$ and two such points from subfamilies $\tilde{\mathbf{T}}_\tau$. For that we consider the system of four equations (6.1), (6.3). Excluding h and g from it, we obtain a system of two equations

$$p_0^2 - 2\sigma = \frac{1 + c_1}{c_1^2 p_t^2} - 3c_1 p_t^2, \;\; \sigma p_0 = \frac{1 + c_1}{c_1 p_t} - c_1^2 p_t^3. \tag{6.8}$$

Excluding p_0 from the system, we obtain equation $h = g^2 - 2\sigma$:

$$f \stackrel{\text{def}}{=} 1 + c_1 - 2\sigma u + (1 - 2c_1)V + c_1 V^2 = 0, \tag{6.9}$$

where $u = c_1 p_t^2 = \sqrt{V}$, i.e. u is real for $V > 0$. It has the four order with respect to u and

$$f = (u - \sigma)^2 (c_1 u^2 + 2\sigma c_1 u + c_1 + 1).$$

So Equation $f = 0$ has roots

$$u = \sigma \ \ (\text{i.e. } V = 1), \tag{6.10}$$

$$u = -\sigma \pm \frac{1}{\sqrt{-c_1}} \ \ \left(\text{i.e } V = 1 - \frac{1}{c_1} \mp \frac{2\sigma}{\sqrt{-c_1}}\right). \tag{6.11}$$

The line (6.10) does not intersect sets F_j (it belongs to their boundary).

Real solutions (6.11) exist only for $c_1 < 0$. As $u = c_1 p_t^2$, then $p_t^2 > 0$ for real p_t and $u < 0$. Hence, real solutions have the sign minus only, i.e. they are

$$u = -\left(\sigma + \frac{1}{\sqrt{-c_1}}\right) \text{ and } V = 1 + \frac{2\sigma}{\sqrt{-c_1}} - \frac{1}{c_1}. \tag{6.12}$$

In Fig. 2 curves (6.12) are given by dash lines with labels H_σ, i.e. H_+ for $\sigma = +1$ and H_- for $\sigma = -1$. The curve H_+ intersects the set F_1 in the interval

$$c_1 \in (-0.7304\ldots, -0.5097\ldots). \tag{6.13}$$

It intersects also sets F_5 and F_6. The curve H_- intersects the set F_2 in the interval

$$c_1 \in (-0.1624\ldots, -0.04267\ldots). \tag{6.14}$$

It intersects also sets F_8, F_7 and F_6. Substituting solutions (6.12) into the first equation (6.8), after some simplifications, we obtain

$$p_0^2 = 4\left(\sigma + \frac{1}{\sqrt{-c_1}}\right).$$

Hence,

$$y \stackrel{\text{def}}{=} \sigma p_0^2 = 4\left(1 + \frac{\sigma}{\sqrt{-c_1}}\right). \tag{6.15}$$

Curves (6.15) are denoted as G_σ and are given at Fig. 1 by dash lines. The curve G_- intersects the set D_1 in the interval

$$c \in (0.5327\ldots, 0.6637\ldots). \tag{6.16}$$

It also intersects sets D_4 and D_5. It touches the curve C_3 in the point $(c, y) = (3/4, -4)$. Interval (6.16) corresponds to the interval

$$c_1 \in (-0.4632\ldots, -0.3362\ldots). \tag{6.17}$$

Intervals (6.14) and (6.17) do not intersect. Hence, the invariant manifold M, containing locally integrable points from both subfamilies $\tilde{\mathbf{S}}_\sigma$ and $\tilde{\mathbf{T}}_\tau$ is absent. But manifolds M, containing only integrable points exist: they have either one point from sets D_1–D_3 that does not belong to the intersection $G_- \cap D_1$; either two points from sets F_1–F_4 which do not belong to the subfamily K, to intersections $H_+ \cap F_1$ and $H_- \cap F_2$; or 4 points from the subfamily K. Real manifolds M contain either 2 real points from F_4 either one real point from D_1–D_3 or no real point with property of local integrability.

7 Summary

The considered particular case of the Euler–Poisson system of equations has 44 sets of such complex stationary solutions near which the system is locally integrable. Among them 10 sets are real. Complex invariant manifold with fixed values of integrals of energy and momentum can have 4, 2, 1 or 0 stationary solutions near which the system is locally integrable, but the real manifold can have 2, 1 or 0 real solutions.

Acknowledgements

The work was supported by RFBR, grant 05-01-00050, and by grant of President RF Sc.Sch.-4476.06.02.

References

[1] Bruno, A.D. and Edneral, V.F.: Normal forms and integrability of ODE systems, *Programming and Computer Software*, **32**, No. 3, 1-6, 2006.
Bruno, A.D. and Edneral, V.F.: Normal forms and integrability of ODE systems, in: Ganzha, V.G., Mayr, E.W. and Vorozhtsov, E.V. (Eds.), *Proceedings of CASC 2005*, LNCS 3718, Springer-Verlag, Berlin, Heidelberg, 65-74, 2005.

[2] Bruno, A.D.: Theory of normal forms of the Euler-Poisson equations, *Preprint of the Keldysh Institute of Applied Mathematics of RAS*, Moscow, 2005, (In Russian).

[3] Bruno, A.D.: Analytical form of differential equations (I), *Trudy Moskov. Mat. Obsc.*, **25**, 119-262, 1971, (In Russian) = *Trans. Moscow Math. Soc.*, **25**, 131-288, 1971, (In English).

[4] Bruno, A.D.: *Local Methods in Nonlinear Differential Equations*, Springer-Verlag, Berlin–Heidelberg, 1989.

[5] Bruno, A.D.: *Power Geometry in Algebraic and Differential Equations*, Elsevier Science (North-Holland), Amsterdam, 2000.

[6] Golubev, V.V.: *Lectures on Integration of Equations of Motion of a Rigid Body Around a Fixed Point*, GITTL, Moscow, 1953, (In Russian).

[7] Ziglin, S.L.: Branching solutions and nonexistence of integrals in the Hamiltonian mechanics I, II, *Functional Analysis and Applications*, **16**, No. 3, 30-41, 1982; **17**, no. 1, 8-23, 1983.

Reducibility and Galois groupoid[1]

Extended abstract

Guy Casale

IRMAR, Université de Rennes 1
Campus de Beaulieu 35042 Rennes cedex, France
guy.casale@univ-rennes1.fr

In this talk two problems of reducibility/irreduciblitity of ordinary differential equations will be presented from a 'galoisian' point of view. The problem is to determine when an ordinary differential equation can be solved by means of classical functions as defined by H. Umemura in [7].

Definition 0.1 *(Painlevé, Umemura [7]) The field of classical functions over $\mathbb{C}(x)$ is a differential field which is the union of all the differential fields obtained by a tower of strongly normal extensions and algebraic extensions. Strongly normal extensions are :*

- *extensions by the entries of a fondamental solution of a linear ODE,*
- *extensions by an abelian function with classical functions as arguments.*

A common belief is that an answer to this kind of question should be given by a general nonlinear differential Galois theory. In [8, 4], general definitions of what should be a nonlinear Galois group (or groupoid) are given. Because of its geometric flavour we will focus on the Malgrange's Galois groupoid and use it to solve the two following problems.

[1] This work was done when the author was invited participant of the PEM program of the Newton Institute, Cambridge, UK. The author is partially supported by the ANR project no. JC05 41465 and by GIFT NEST-Adventure Project no. 5006.

Irreducibility of P_1: prove that no solution of the first Painlevé equation $y'' = 6y^2 + x$ is a classical function.

Reducibility of PP_6: explain why the Picard-Painlevé sixth equation

$$y'' = \frac{3y^2 - 2y(x-1) + x}{2y(y-1)(y-x)}(y')^2 + \left(\frac{1}{x-y} + \frac{1}{1-x} - \frac{1}{x}\right)y' + \frac{y(y-1)}{2x(x-1)(y-x)}$$

can be solved by a formula though most of its solutions are non classical. The formula to solve PP_6 looks pretty classical :

$$y = \wp(a\omega_1(x) + b\omega_2(x); \omega_1(x), \omega_2(x))$$

with a and b two constants and $\omega_{1,2}$ two periods of $z^2 = y(y-1)(y-x)$.

1 The Galois groupoid of a vector field in $\mathbb{C}^3$

Let X be a vector field in $\mathbb{C}^3$. In general it is not complete and its flows are only defined on open sets small enough. All the dynamic of this vector field is contained in the pseudogroup of transformations of $\mathbb{C}^3$ generated by these local flows. By keeping only the germs of diffeomorphisms from this pseudogroup one gets a groupoid, $TanX$, acting on $\mathbb{C}^3$.

The Galois groupoid of X is the Zariski closure of $TanX$ for a (nearly) obvious embedding of $TanX$ in a infinite dimensional algebraic variety. This variety is the space J^* of formal diffeomorphisms of $\mathbb{C}^3$ with its groupoid structure and its projections on the spaces J^*_q of order q jets of diffeomorphisms. The ring O_{J^*} of this variety is the commutative differential ring of nonlinear partial differential equations on germs of diffeomorphisms. The embedding is the Taylor expansion of elements of $TanX$.

Definition 1.1 *(Malgrange [4]) The Galois groupoid of X is defined by the ideal of O_{J^*} of all the PDEs satisfied by the flows of X.*

Using Lie-Cartan local classification of pseudogroups acting on $\mathbb{C}^2$ [1], one has the following proposition.

Proposition 1.2 *([2]) If X is divergence free and γ is the closed 2-form vanishing on X then one of the following situations occurs:*

- *$Gal(X)$* is imprimitive*: there exists an algebraic 1-form θ s.t. $\theta \wedge d\theta = 0$ and $\theta(X) = 0$,*
- *$Gal(X)$* is transversally affine*: there exists two algebraic 1-forms θ_1, θ_2 vanishing on X and a traceless matrix of 1-forms (θ_i^j), i, j =1 or 2, s.t. $d\theta_i = \theta_i^j \wedge \theta_j$ and $d\theta_i^j = \theta_i^k \wedge \theta_k^j$,*
- *the only transversal equations of $Gal(X)$ are those of the invariance of γ.*

2 Irreducibility of P_1

The discussion about the irreducibility to classical functions of the solutions of the first Painlevé equation depends on the transcendance degree of the differential field generated by these solutions over $\mathbb{C}(x)$.

This is a classical result of Painlevé that such a solution cannot be algebraic, and by the Kolchin-Kovacic lemma its transcendance degree must be two. Such a solution gives an inclusion of the field $\mathbb{C}(x, y, y')$ in a field $\mathbb{C}(x, h_i, \ldots, k_p, \ldots)$. Let's take $\mathbb{C}^3$ and $\mathbb{C}^N$ as model for these fields and let π be the dominate projection induced by the inclusion. The differential structure of the first field is given by the vector field

$$X_1 = \frac{\partial}{\partial x} + y' \frac{\partial}{\partial y} + (6y^2 + x) \frac{\partial}{\partial y'}$$

and because of its special construction the vector field on the second has the following shape

$$X_c = \frac{\partial}{\partial x} + \sum a_i^j(x) h_j \frac{\partial}{\partial h_i} + \sum b_p^q(x, h) k_q \frac{\partial}{\partial k_p} + \ldots$$

The projection of X_c by π gives X_1. The main tool to prove that this projection cannot exist is the following theorem.

Theorem 2.1 *The only transversal equations of $Gal(X_1)$ are those of the invariance of γ.*

Computations on the structural equation of the Galois groupoid of X_c show that

- a quotient of $Gal(X_c)$ is included in $Gal(X_1)$.
- such a quotient must be strictly smaller than $Gal(X_1)$.

On an other side, this quotient must contain $TanX_1$, this yields a contradiction.

3 Reducibility of PP_6

This equation is also divergence free but in this case one has the following theorem. Let X_{PP} be the vector field of this equation on $\mathbb{C}^3$.

Theorem 3.1 *$Gal(X_{PP})$ is transversally affine.*

To prove this we construct two first integrals in a Picard-Vessiot extension of the differential field $\left(C(x,y,y'); \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial y'}\right)$ following P. Painlevé [5]. If $y(x)$ is a solution of PP_6 the integral $\int_0^{y(x)} \frac{d\xi}{\sqrt{\xi(\xi-1)(\xi-x)}}$ is a period of $z^2 = y(y-1)(y-x)$. By pulling-back linear first integral of the linear order two equation of the periods (Picard-Fuchs) one gets:
for each solution of $v'' + \left(\frac{1}{x^2} + \frac{1}{(x-1)^2} - \frac{1}{x(x-1)}\right) v = 0$, the function

$$y'v\sqrt{\frac{x(x-1)}{y(y-1)(y-x)}} + \int \sqrt{\frac{x(x-1)}{y(y-1)(y-x)}} \left\{ \left[\frac{v}{2} \left(\frac{1}{x} + \frac{1}{x-1} + \frac{1}{y-x} \right) - v' \right] dy + v \frac{y(1-y)}{2x(y-x)(x-1)} dx \right\}$$

is a first integral of X_{PP}. The theorem follows easily.

The Galois groupoid shows that this equation is special even if its non algebraic solutions are non classical. In fact the first integrals are classical functions of three variables.

References

[1] Cartan, É. - *Les sous-groupes des groupes continus de transformations*, Ann. Sci. École Normale Sup. **25** (1908) 57–194

[2] Casale, G - *Le groupoïde de Galois de P_1 et son irréductibilité*, to appear in Commentarii Mathematici Helvetici

[3] Casale, G - *The Galois groupoid of the Picard-Painlevé VI equation*, in preparation

[4] Malgrange, B. - *Le groupoïde de Galois d'un feuilletage*, Monographie **38** vol **2** de L'enseignement mathématique (2001)

[5] Painlevé, P. - *Leçons de Stockholm* (1895), Oeuvres complètes Tome **1**, éditions du CNRS (1972)

[6] Picard, E. - *Mémoire sur la théorie des fonctions algébriques de deux variables*, Journal de Liouville **5** (1889) 135–319

[7] Umemura, H. - *Painlevé equations and classical functions* [translation of Sūgaku **47** (1995)] Sugaku Expositions. **11** (1998) 77–100

[8] Umemura, H. - *Galois theory of algebraic and differential equation*, Nagoya Math. J. vol **144** (1996) 1–58
- *Differential Galois theory of infinite dimension*, Nagoya Math. J. vol **144** (1996) 59–135

Some results *à l'Abel* obtained by use of techniques *à la Hopf*

Christian Costermans Hoang Ngoc Minh

University Lille II, 1, Place Déliot, 59024 Lille,
ccostermans@univ-lille2.fr
hoang@univ-lille2.fr

Abstract

In this work, we obtain some results *à l'Abel* dealing with noncommutative generating series of polylogarithms and multiple harmonic sums, by using techniques *à la Hopf*. In particular, this enables to explicit generalized Euler constants associated to divergent polyzêtas and to extract the constant part of (commutative and noncommutative) generating series of all polyzêtas.

Keywords: asymptotic expansion, generating series, multiple harmonic sums, polylogarithms, polyzêtas

1 Introduction

Let us consider the alphabet $Y = \{y_i\}_{i\in\mathbb{N}_+}$. To each word $w = y_{s_1}\dots y_{s_r}$ of the monoid Y^*, we associate the multiple harmonic sum $\mathrm{H}_w(N)$ and the polylogarithm $\mathrm{Li}_w(z)$

$$\mathrm{H}_w(N) = \sum_{N\geq n_1>\dots>n_r>0} \frac{1}{n_1^{s_1}\dots n_r^{s_r}}, \quad \mathrm{Li}_w(z) = \sum_{n_1>\dots>n_r>0} \frac{z^{n_1}}{n_1^{s_1}\dots n_r^{s_r}}. \qquad (1)$$

For $0 \leq N < r, \mathrm{H}_w(N) = 0$ and for the empty word ϵ, we put $\mathrm{H}_\epsilon(N) = 1$, for any $N \geq 0$. For $w \in Y^*\setminus y_1Y^*$, the limits $\lim_{z\to 1}\mathrm{Li}_w(z)$ and $\lim_{N\to\infty}\mathrm{H}_w(N)$

exist and, by an Abel theorem, are equal to the convergent polyzêta $\zeta(w)$

$$\zeta(w) = \sum_{n_1>\ldots>n_r>0} \frac{1}{n_1^{s_1}\ldots n_r^{s_r}}, \quad s_1 > 1. \tag{2}$$

In other cases, *i.e.* for $w = y_s w'$, the associated polylogarithm Li_w and polyzêta $\zeta(w)$ can be considered respectively as a polylogarithmic generating series and as Dirichlet series

$$\mathrm{Li}_{y_s w'}(z) = \sum_{N>0} \frac{p_N}{N^s} z^N \quad \text{and} \quad \zeta(y_s w') = \sum_{N>0} \frac{p_N}{N^s} \tag{3}$$

with $p_N = \mathrm{H}_{w'}(N-1)$. Both series can be obtained from the following generating series

$$\mathrm{P}_{w'}(z) = \sum_{N=0}^{\infty} \mathrm{H}_{w'}(N) z^N = \sum_{N\geq 0} p_{N+1} z^N, \tag{4}$$

respectively by the polylogarithmic transform and by the Mellin transform [1]

$$\mathrm{Li}_{y_s w'}(z) = \int_0^\infty \frac{\mathrm{P}_{w'}(ze^{-u})}{\Gamma(s)} \frac{du}{u^{1-s}} \quad \text{and} \quad \zeta(y_s w') = \int_0^\infty \frac{\mathrm{P}_{w'}(e^{-u})}{\Gamma(s)} \frac{du}{u^{1-s}}. \tag{5}$$

The generating series $\mathrm{P}_{w'}$ can also be expressed using the polylogarithm :

$$\mathrm{P}_{w'}(z) = (1-z)^{-1} \mathrm{Li}_{w'}(z). \tag{6}$$

The knowledge of the singular expansion of $\mathrm{P}_{w'}$ in the scale $\{(1-z)^a \log^b(1-z)\}_{a\in\mathbb{Z},b\in\mathbb{N}}$ enables then to get, on the first hand the asymptotic behaviour, as $N \to \infty$, of its Taylor coefficients $\mathrm{H}_{w'}(N)$ in the scale $\{N^\alpha \log^\beta N\}_{\alpha\in\mathbb{Z},\beta\in\mathbb{N}}$. Then, to deduce the behaviour of $\mathrm{H}_w(N)$, since $\mathrm{H}_w(N) = \sum_{i=1}^N \mathrm{H}_{w'}(i-1)/i^s$. This gives on the other hand, through a tauberian theorem, the singular expansion of Dirichlet series $\zeta(y_s w')$ considered then as a function of the complex variable s.

Both studies lead to apply another Abel theorem dealing with Dirichlet series [2]. Indeed, let us consider the partial sum S_N of coefficients of the ordinary generating series $\mathrm{P}_{w'}(z)$,

$$S_N = \sum_{i=1}^{N} p_{i+1} = \sum_{i=1}^{N} \mathrm{H}_{w'}(i). \tag{7}$$

If S_N admits a singular expansion of the following type

$$S_N = \sum_{j=1}^{k} B_j N^{\sigma_j} \log^{\alpha_j} N + O(N^{\beta}), \tag{8}$$

where, for all $j = 1, .., k, B_j$ is an arbitrary complex number, σ_j, α_j are arbitrary integers, and β is an integer such that $\beta > \sigma_k$, then the Dirichlet series $\zeta(y_s w')$ is convergent for $s > 1$ and even regular except in $\sigma_1, \ldots, \sigma_k$ which are its logarithmic singularities.

In order to adapt automatically these Abel techniques to polylogarithms $\{\mathrm{Li}_w\}_{w\in Y^*}$ and to multiple harmonic sums $\{\mathrm{H}_w\}_{w\in Y^*}$, we consider the *noncommutative generating series*

$$\Lambda(z) = \sum_{w\in Y^*} \mathrm{Li}_w(z)\, w \quad \text{and} \quad \mathrm{H}(N) = \sum_{w\in Y^*} \mathrm{H}_w(N)\, w. \tag{9}$$

Through algebraic combinatoric [3] and elements of topology of formal series in noncommutative variables [4], we show in Section 2.2 the existence of formal series over Y, Z_1 and Z_2 in non commutative variables with constant coefficients, such that

$$\begin{aligned} &\lim_{z\to 1} \exp\left[y_1 \log \frac{1}{1-z}\right] \Lambda(z) = Z_1 \text{ and} \\ &\lim_{N\to\infty} \exp\left[\sum_{k\geq 1} \mathrm{H}_{y_k}(N) \frac{(-y_1)^k}{k}\right] \mathrm{H}(N) = Z_2. \end{aligned} \tag{10}$$

Moreover, we have $Z_1 = Z_2$, both standing for the noncommutative generating series of all convergent polyzêtas $\{\zeta(w)\}_{w\in Y^*\setminus y_1 Y^*}$, (as shown by the

factorized form). This enables, in particular, to explicit generalized Euler constants associated to divergent polyzêtas $\{\zeta(w)\}_{w\in y_1Y^*}$ and to extract the constant part of generating series (commutative and noncommutative) of all polyzêtas.

Techniques presented in this paper can be applied to other fields, like polysystems occuring in physical problems, and enable to make the calculations easier. To illustrate this, we present in appendix some results to compute, thanks to such techniques, the solution of a linear differential system, with three singularities, that can be supposed to be $\{0, 1, \infty\}$, after an homographic transformation.

2 Polylogarithm and harmonic sum

2.1 Algebraic properties

2.1.1 Symmetric functions and harmonic sums

Let $\{t_i\}_{i\in\mathbb{N}_+}$ be an infinite set of variables. The elementary symmetric functions λ_k and the sums of powers ψ_k are defined by

$$\lambda_k(\underline{t}) = \sum_{n_1>\ldots>n_k>0} t_{n_1}\ldots t_{n_k} \quad \text{and} \quad \psi_k(\underline{t}) = \sum_{n>0} t_n^k. \tag{11}$$

They are respectively coefficients of the following generating functions

$$\lambda(\underline{t}|z) = \sum_{k>0} \lambda_k(\underline{t}) z^k = \prod_{i\geq 1}(1+t_i z) \text{ and} \tag{12}$$
$$\psi(\underline{t}|z) = \sum_{k>0} \psi_k(\underline{t}) z^{k-1} = \sum_{i\geq 1} \frac{t_i}{1-t_i z}.$$

These generating functions satisfy a Newton identity

$$d/dz \log \lambda(\underline{t}|z) = \psi(\underline{t}| - z). \tag{13}$$

The fundamental theorem from symmetric functions theory asserts that the $\{\lambda_k\}_{k\geq 0}$ are linearly independent, and remarkable identities give (putting $\lambda_0 = 1$) :

$$k!\lambda_k \;=\; (-1)^k \sum_{\substack{s_1,\ldots,s_k>0\\ s_1+\ldots+ks_k=k}} \binom{k}{s_1,\ldots,s_k}\left(-\frac{\psi_1}{1}\right)^{s_1}\ldots\left(-\frac{\psi_k}{k}\right)^{s_k} \qquad (14)$$

Let $w = y_{s_1}\ldots y_{s_r} \in Y^*$. The quasi-symmetric function F_w, of depth $r = |w|$ and of degree (or weight) $s_1 + \ldots + s_r$, is defined by

$$F_w(\underline{t}) = \sum_{n_1>\ldots>n_r>0} t_{n_1}^{s_1}\ldots t_{n_r}^{s_r}. \qquad (15)$$

In particular, $F_{y_1^k} = \lambda_k$ and $F_{y_k} = \psi_k$. As a consequence, the functions $\{F_{y_1^k}\}_{k\geq 0}$ are linearly independent and integrating differential equation (13) shows that functions $F_{y_1^k}$ and F_{y_k} are linked by the formula

$$\sum_{k\geq 0} F_{y_1^k} z^k = \exp\Big[-\sum_{k\geq 1} F_{y_k}\frac{(-z)^k}{k}\Big]. \qquad (16)$$

Remarkable identity (14) can be then seen as :

$$k!y_1^k = (-1)^k \sum_{\substack{s_1,\ldots,s_k>0\\ s_1+\ldots+ks_k=k}} \binom{k}{s_1,\ldots,s_k}\frac{(-y_1)^{\sqcup\!\sqcup s_1}}{1^{s_1}} \sqcup\!\sqcup \ldots \sqcup\!\sqcup \frac{(-y_k)^{\sqcup\!\sqcup s_k}}{k^{s_k}} \qquad (17)$$

Every $\mathrm{H}_w(N)$ can be obtained by specializing variables $\{t_i\}_{N\geq i\geq 1}$ at $t_i = 1/i$ and, for $i > N, t_i = 0$ in the quasi-symmetric function F_w [5]. In the same way, when $w \in Y^* \setminus y_1Y^*$, the convergent polyzêta $\zeta(w)$ can be obtained by specializing variables $\{t_i\}_{i\geq 1}$ at $t_i = 1/i$ in F_w [5]. The notation F_w is extended by linearity to all polynomials over Y.

If u (resp. v) is a word in Y^*, of length r and of weight p (resp. of length s and of weight q), $F_{u\sqcup\!\sqcup v}$ is a quasi-symmetric function of depth $r+s$ and of weight $p+q$, and we have $F_{u\sqcup\!\sqcup v} = F_u\, F_v$.

In consequence, $\mathrm{H}_{u \sqcup\!\sqcup v} = \mathrm{H}_u \, \mathrm{H}_v$ [5]. In the same way, when $u, v \in Y^* \backslash y_1 Y^*$, we also have $\zeta(u \sqcup\!\sqcup v) = \zeta(u) \, \zeta(v)$ [5].

Let us consider the noncommutative generating series $\mathrm{H}(N)$ of $\{\mathrm{H}_w(N)\}_{w \in Y^*}$ [14],

$$\mathrm{H}(N) = \sum_{w \in Y^*} \mathrm{H}_w(N) \, w = \prod_{1}^{l=N} \left(1 + \sum_{i>0} \frac{y_i}{l^i} \right). \tag{18}$$

since it verifies the difference equation

$$\mathrm{H}(N) = \left(1 + \sum_{i>0} \frac{y_i}{N^i} \right) \mathrm{H}(N-1), \quad \text{with} \quad \mathrm{H}(0) = 1. \tag{19}$$

2.1.2 Polylogarithms and polyzêtas

Let us denote by $\mathcal{C}$ the algebra $\mathbb{C}[z, 1/z, 1/(1-z)]$ of polynomial functions in $z, 1/z$ and $1/(1-z)$. We define two differential forms $\omega_0(z) = dz/z$ and $\omega_1(z) = dz/(1-z)$.

Let $w = x_0^{s_1-1} x_1 \ldots x_0^{s_r-1} x_1 \in X^* x_1$. One can check that the polylogarithm Li_w is also the value of the iterated integral over ω_0, ω_1 and along the integration path $0 \rightsquigarrow z$:

$$\mathrm{Li}_w = \int_{0 \rightsquigarrow z} \omega_0^{s_1-1} \omega_1 \ldots \omega_0^{s_r-1} \omega_1. \tag{20}$$

This provides an analytic continuation of the Li_w over the universal covering $\widetilde{\mathbb{C} - \{0,1\}}$ of $\mathbb{C}$ without points 0 and 1. We extend the definition of polylogarithms over X^* putting

$$\mathrm{Li}_{x_0^k}(z) = \log^k z / k!, \quad \text{for} \quad k \in \mathbb{N}. \tag{21}$$

Let $\mathrm{LI}_{\mathcal{C}} = (\mathcal{C}\{\mathrm{Li}_w\}_{w \in X^*}, .)$ be the smallest $\mathcal{C}$-algebra containing $\mathcal{C}$ and stable by differentiation and integration over ω_0, ω_1. It can be identified with the $\mathcal{C}$-module generated by polylogarithms. Thus, the polylogarithms are $\mathcal{C}$-linearly independent [6]. Hence, $(\mathcal{C}\{\mathrm{Li}_w\}_{w \in X^*}, .)$ is identified with the polynomial algebra $(\mathcal{C}\{\mathrm{P}_l\}_{l \in \mathcal{L}ynX}, .)$ [6].

The noncommutative generating series $\mathrm{L}(z) = \sum_{w\in X^*} \mathrm{Li}_w(z)\ w$ satisfies Drinfel'd differential equation [7, 8]

$$\begin{aligned} d\mathrm{L} &= (x_0\omega_0 + x_1\omega_1)\mathrm{L}, \text{ with the condition} \\ \mathrm{L}(\varepsilon) &= e^{x_0 \log \varepsilon} + O(\sqrt{\varepsilon}) \text{ for } \varepsilon \to 0^+. \end{aligned} \tag{22}$$

This enables to prove that L is the exponential of a Lie series [6]. So, applying a Ree theorem, it verifies Friedrichs criterion [6], *i..e* $\mathrm{Li}_{u \sqcup\!\sqcup v} = \mathrm{Li}_u\ \mathrm{Li}_v$ for $u, v \in X^*$. In particular, when $u, v \in x_0X^*x_1$, we also have $\zeta(u \sqcup\!\sqcup v) = \zeta(u)\ \zeta(v)$. From the factorization of monoid by Lyndon words, we obtain the factorization of the series L [6] :

$$\mathrm{L}(z) = e^{x_1 \log \frac{1}{1-z}} \left[\prod_{l\in \mathcal{L}ynX\setminus\{x_0,x_1\}}^{\searrow} e^{\mathrm{Li}_{S_l}(z)[l]} \right] e^{x_0 \log z}. \tag{23}$$

For all $l \in \mathcal{L}ynX \setminus \{x_0, x_1\}$, we have $S_l \in x_0X^*x_1$. So, let us put [6]

$$\mathrm{L}_{\mathrm{reg}} = \prod_{l\in \mathcal{L}ynX\setminus\{x_0,x_1\}}^{\searrow} e^{\mathrm{Li}_{S_l}[l]} \quad \text{and} \quad Z = \mathrm{L}_{\mathrm{reg}}(1). \tag{24}$$

Let σ be the monoid endomorphism verifying $\sigma(x_0) = -x_1, \sigma(x_1) = -x_0$, we also get [9]

$$\mathrm{L}(z) = \sigma[\mathrm{L}(1-z)]Z = e^{x_0 \log z}\sigma[\mathrm{L}_{\mathrm{reg}}(1-z)]e^{-x_1 \log(1-z)}Z. \tag{25}$$

In consequence, from (23) and (25), we get respectively

$$\mathrm{L}(z) \underset{z\to 0}{\sim} \exp(x_0 \log z) \quad \text{and} \quad \mathrm{L}(z) \underset{z\to 1}{\sim} \exp\left(x_1 \log \frac{1}{1-z}\right) Z. \tag{26}$$

Let $\pi_Y : \mathrm{LI}_{\mathcal{C}}\langle\langle X\rangle\rangle \to \mathrm{LI}_{\mathcal{C}}\langle\langle Y\rangle\rangle$ a projector s.t., for $f \in \mathrm{LI}_{\mathcal{C}}, w \in X^*, \pi_Y(f\ wx_0) = 0$. Then

$$\Lambda(z) = \pi_Y\mathrm{L}(z) \underset{z\to 1}{\sim} \exp\left(y_1 \log \frac{1}{1-z}\right) \pi_Y Z. \tag{27}$$

Definition 1 ([1]). *Let* $\zeta_{ш} : (C\langle\langle X\rangle\rangle, ш) \to (\mathbb{C}, .)$ *be the algebra morphism (*i.e. *for* $u, v \in X^*, \zeta_{ш}(u ш v) = \zeta_{ш}(u)\zeta_{ш}(v)$*) verifying for all convergent word* $w \in x_0X^*x_1, \zeta_{ш}(w) = \zeta(w)$*, and such that* $\zeta_{ш}(x_0) = \zeta_{ш}(x_1) = 0$.

Then, the noncommutative generating series $Z_{ш} = \sum_{w\in X^*} \zeta_{ш}(w)\ w$ verifies $Z_{ш} = Z$ [1]. In consequence, $Z_{ш}$ is the unique Lie exponential verifying $\langle Z_{ш}|x_0\rangle = \langle Z_{ш}|x_1\rangle = 0$ and $\langle Z_{ш}|w\rangle = \zeta(w)$, for any $w \in x_0X^*x_1$. Its logarithm is given by $\log Z_{ш} = \sum_{w\in X^*} \zeta_{ш}(w)\ \pi_1(w)$, where $\pi_1(w)$ is the Lie polynomial [3]

$$\pi_1(w) = \sum_{k\geq 1} \frac{(-1)^{k-1}}{k} \sum_{u_1,\cdots,u_k\in X^*\setminus\{\epsilon\}} \langle w|u_1 ш \cdots ш u_k\rangle\ u_1\cdots u_k. \tag{28}$$

The series Z shall be understood then as Drinfel'd associator Φ_{KZ} [7, 8] verifying duality, pentagonal and hexagonal relations. We also can obtain the expression of this associator given by Lê and Murakami [11] thanks to the following expansion [10]

$$Z = \sum_{k\geq 0} \sum_{l_1,\cdots,l_k\geq 0} \zeta_{ш}(x_1x_0^{l_1} \circ \cdots \circ x_1x_0^{l_k}) \prod_{i=1}^{k} \mathrm{ad}_{x_0}^{l_i}\, x_1, \tag{29}$$

where $\mathrm{ad}_{x_0}^{l}\, x_1$ stands for the iterated Lie bracket $[x_0, [\ldots, [x_0, x_1]\ldots]$, for $l > 0$ and $\mathrm{ad}_{x_0}^{0}\, x_1 = x_1$, the operation $\circ$ being defined as $x_1x_0^l \circ P = x_1(x_0^l ш P)$, for any $P \in \mathbb{C}\langle X\rangle$.

2.1.3 Ordinary generating series of harmonic sums

The functions $\{\mathrm{Li}_w\}_{w\in X^*}$ are $\mathcal{C}$-linearly independent. Thus, the functions $\{\mathrm{P}_w\}_{w\in Y^*}$ are also $\mathbb{C}$-linearly independent. In consequence, functions $\{\mathrm{H}_w\}_{w\in Y^*}$ are also $\mathbb{C}$-linearly independent [13, 12]. So,

Proposition 1 ([14]). *Extended by linearity, the application* $\mathrm{P} : u \mapsto \mathrm{P}_u$ *is an isomorphism from polynomial algebra* $(\mathbb{C}\langle Y\rangle, ⧢)$ *over Hadamard algebra* $(\mathbb{C}\{\mathrm{P}_w\}_{w\in Y^*}, \odot)$*. Moreover, the application* $\mathrm{H} : u \mapsto \mathrm{H}_u = \{\mathrm{H}_u(N)\}_{N\geq 0}$ *is an isomorphism from* $(\mathbb{C}\langle Y\rangle, ⧢)$ *over the algebra* $(\mathbb{C}\{\mathrm{H}_w\}_{w\in Y^*}, .)$.

Proof : Indeed, on the first hand, $\ker \mathrm{P} = \{0\}$ and $\ker \mathrm{H} = \{0\}$, and on the other hand, P is a morphism for Hadamard product (it inherits of H for the harmonic product) :

$$\begin{aligned} \mathrm{P}_u(z) \odot \mathrm{P}_v(z) &= \sum_{N\geq 0} \mathrm{H}_u(N)\mathrm{H}_v(N) z^N \\ &= \sum_{N\geq 0} \mathrm{H}_{u \sqcup\!\sqcup v}(N) z^N \\ &= \mathrm{P}_{u \sqcup\!\sqcup v}(z). \end{aligned}$$

Studying the equivalence between action of $\{(1-z)^l\}_{l\in\mathbb{Z}}$ over $\{\mathrm{P}_w(z)\}_{w\in Y^*}$ and this of $\{N^k\}_{k\in\mathbb{Z}}$ over $\{\mathrm{H}_w(N)\}_{w\in Y^*}$ [12], we have

Theorem 1. *The Hadamard $\mathcal{C}$-algebra of $\{\mathrm{P}_w\}_{w\in Y^*}$ can be identified with this of $\{\mathrm{P}_l\}_{l\in\mathcal{L}ynY}$. Identically, the algebra of harmonic sums $\{\mathrm{H}_w\}_{w\in Y^*}$ with polynomial coefficients can be identified with this of $\{\mathrm{H}_l\}_{l\in\mathcal{L}ynY}$.*

As for polylogarithms, we extend the definition of P_w putting $\mathrm{P}_w(z) = (1-z)^{-1}\mathrm{Li}_w(z)$, for any $w \in X^*$. The noncommutative generating series of $\{\mathrm{P}_w\}_{w\in X^*}$ is defined by

$$\mathrm{P}(z) = \sum_{w\in X^*} \mathrm{P}_w(z)\, w = \frac{\mathrm{L}(z)}{1-z}. \tag{30}$$

In consequence, by (23), we have

$$\mathrm{P}(z) = e^{-(x_1+1)\log(1-z)}\mathrm{L}_{\mathrm{reg}}(z)e^{x_0\log z}. \tag{31}$$

Lemma 1. *Let* $\mathrm{Mono}(z) = e^{-(x_1+1)\log(1-z)}$. *Then*

$$\mathrm{Mono} = \sum_{k\geq 0} \mathrm{P}_{y_1^k}\, y_1^k, \quad \textit{and} \quad \mathrm{Mono}^{-1} = \sum_{k\geq 0} \mathrm{P}_{y_1^k}\, (-y_1)^k.$$

Since the coefficient of z^N in the Taylor expansion of $\mathrm{P}_{y_1^k}$ is $\mathrm{H}_{y_1^k}(N)$ then

Lemma 2. *Let* $\mathrm{Const} = \sum_{k\geq 0} \mathrm{H}_{y_1^k}\, y_1^k$. *Then*

$$\mathrm{Const} = \exp\left[-\sum_{k\geq 1} \mathrm{H}_{y_k}\frac{(-y_1)^k}{k}\right] \quad \textit{and} \quad \mathrm{Const}^{-1} = \exp\left[\sum_{k\geq 1} \mathrm{H}_{y_k}\frac{(-y_1)^k}{k}\right].$$

Proof : This is a consequence of Formula (16).

Proposition 2. *For $k > 0, \mathrm{H}_{y_1^k}$ is polynomial in $\{\mathrm{H}_{y_r}\}_{1\leq r\leq k}$ (which are algebraically independent), and*

$$\mathrm{H}_{y_1^k} = \sum_{\substack{s_1,\ldots,s_k>0\\ s_1+\ldots+ks_k=k}} \frac{(-1)^k}{s_1!\ldots s_k!}\left(-\frac{\mathrm{H}_{y_1}}{1}\right)^{s_1}\ldots\left(-\frac{\mathrm{H}_{y_k}}{k}\right)^{s_k}.$$

Proof : From Identity (17), and applying the isomorphism H on the set of Lyndon words $\{y_r\}_{1\leq r\leq k}$, we get the expected result.

Example 1. $\mathrm{H}_{y_1^2} = (\mathrm{H}_{y_1}^2 - \mathrm{H}_{y_2})/2$, $\mathrm{H}_{y_1^3} = (\mathrm{H}_{y_1}^3 - 3\mathrm{H}_{y_2}\mathrm{H}_{y_1} + 2\mathrm{H}_{y_3})/6$.

Proposition 3 ([14])**.** *Let σ be the morphism verifying $\sigma(x_0) = -x_1, \sigma(x_1) = -x_0$.*

$$\mathrm{P}(z) = e^{x_0\log z}\left[\prod_{l\in\mathcal{L}ynX,\setminus\{x_0,x_1\}}^{\searrow} e^{\mathrm{Li}_{s_l}(1-z)\sigma([l])}\right]\mathrm{Mono}(z)Z,$$

Proof : On the first hand, from (31) and on the other hand, from (25), we get $\mathrm{P}(z) = e^{x_0\log z}\sigma[\mathrm{L}_{\mathrm{reg}}(1-z)]e^{-(x_1+1)\log(1-z)}Z$. Using the expressions of $\mathrm{L}_{\mathrm{reg}}(1-z)$ and of $\mathrm{Mono}(z)$, we get the expected results.

2.2 Asymptotic expansion

2.2.1 Results à l'Abel for generating series

Proposition 4. $\mathrm{P}(z) \underset{z\to 0}{\sim} e^{x_0\log z}$ *and* $\mathrm{P}(z) \underset{z\to 1}{\sim} \mathrm{Mono}(z)Z$.

Proof : From $\mathrm{P}(z) = e^{-(x_1+1)\log(1-z)}\mathrm{L}_{\mathrm{reg}}(z)e^{x_0\log z}$, we can deduce the behaviour of $\mathrm{P}(z)$ around 0. From Formula (25), we get the behaviour of $\mathrm{P}(z)$ around 1.

Corollary 1. *Let $\Pi(z) = \pi_Y\mathrm{P}(z) = \sum_{w\in Y^*}\mathrm{P}_w(z)\,w$. Then $\Pi(z) \underset{z\to 1}{\sim} \mathrm{Mono}(z)\pi_Y Z$.*

From this, we extract, taking care of Lemma 1, Taylor coefficients of P_w, and we get

Corollary 2. $\mathrm{H}(N) \underset{N\to\infty}{\widetilde{\ }} \mathrm{Const}(N)\pi_Y Z$.

Theorem 2.

$$\lim_{z\to 1}\exp\left(y_1\log\frac{1}{1-z}\right)\Lambda(z) = \lim_{N\to\infty}\exp\left(\sum_{k\geq 1}\mathrm{H}_{y_k}(N)\frac{(-y_1)^k}{k}\right)\mathrm{H}(N) = \pi_Y Z.$$

Proof : This is a consequence of Formula (27), of Lemma 2 and of Corollary 2.

From Proposition 3, we deduce

$$\mathrm{P}(z) = e^{x_0\log z}\left[\prod_{\substack{l\in\mathcal{L}ynX,\\ l\neq x_0,x_1}}^{\searrow} z\left(\sum_{k\geq 0}\mathrm{P}_{S_l^{\sqcup\!\sqcup k}}(1-z)\frac{(\sigma([l]))^k}{k!}\right)\right]\mathrm{Mono}(z)Z. \qquad (32)$$

Hence, the knowledge of Taylor expansion around 0 of $\{\mathrm{P}_w(1-z)\}_{w\in X^*}$ gives

Theorem 3 ([12])**.** *For all $g \in \mathcal{C}\{\mathrm{P}_w\}_{w\in Y^*}$, there exist algorithmically computable coefficients $c_j \in \mathbb{C}$, $\alpha_j \in \mathbb{Z}$ and $\beta_j \in \mathbb{N}$ such that*

$$g(z) \sim \sum_{j=0}^{+\infty} c_j(1-z)^{\alpha_j}\log^{\beta_j}(1-z) \quad \textit{for} \quad z\to 1.$$

In consequence, there exist algorithmically computable coefficients $b_i \in \mathbb{C}$, $\eta_i \in \mathbb{Z}$ and $\kappa_i \in \mathbb{N}$ such that

$$[z^n]g(z) \sim \sum_{i=0}^{+\infty} b_i n^{\eta_i}\log^{\kappa_i}(n) \quad \textit{for} \quad n\to\infty.$$

Corollary 3 ([12])**.** *Let $\mathcal{Z}$ the $\mathbb{Q}$-algebra generated by convergent polyzêtas and $\mathcal{Z}'$ the $\mathbb{Q}[\gamma]$-algebra generated by $\mathcal{Z}$. Then there exist algorithmically computable coefficients $c_j \in \mathcal{Z}$, $\alpha_j \in \mathbb{Z}$ and $\beta_j \in \mathbb{N}$ such that*

$$\forall w\in Y^*, \mathrm{P}_w(z) \sim \sum_{j=0}^{+\infty} c_j(1-z)^{\alpha_j}\log^{\beta_j}(1-z) \quad \textit{for} \quad z\to 1.$$

In consequence, there exist algorithmically computable coefficients $b_i \in \mathcal{Z}', \kappa_i \in \mathbb{N}$ *and* $\eta_i \in \mathbb{Z}$ *such that*

$$\forall w \in Y^*, \mathrm{H}_w(N) \sim \sum_{i=0}^{+\infty} b_i N^{\eta_i} \log^{\kappa_i}(N) \quad \textit{for} \quad N \to +\infty.$$

2.2.2 Generalized Euler constants associated to divergent polyzêtas

Definition 2. *Let* $\zeta_{ш} : (\mathbb{C}\langle Y\rangle, ш) \to (\mathbb{C}, .)$ *the algebra morphism (*i.e. *for all convergent word* $u, v \in Y^*, \zeta_{ш}(u ш v) = \zeta_{ш}(u)\zeta_{ш}(v)$*) verifying for* $w \in Y^* \setminus y_1 Y^*$, $\zeta_{ш}(w) = \zeta(w)$ *and such that* $\zeta_{ш}(y_1) = \gamma$.

Proposition 5.

$$\zeta_{ш}(y_1^k) = \sum_{\substack{s_1,\ldots,s_k>0 \\ s_1+\ldots+ks_k=k}} \frac{(-1)^k}{s_1!\ldots s_k!}(-\gamma)^{s_1}\left(-\frac{\zeta(2)}{2}\right)^{s_2}\cdots\left(-\frac{\zeta(k)}{k}\right)^{s_k}.$$

Proof : By (17) and applying the (surjective) morphism $\zeta_{ш}$, we get the expected result.

In consequence,

Theorem 4. *For* $k > 0$, *the constant* $\zeta_{ш}(y_1^k)$ *associated to divergent polyzêta* $\zeta(y_1^k)$ *is a polynomial of degree* k *in* γ *with coefficients in* $\mathbb{Q}[\zeta(2), \zeta(2i+1)]_{0<i\leq(k-1)/2}$. *Moreover, for* $l = 0, .., k$, *the coefficient of* γ^l *is of weight* $k - l$.

Example 2. $\zeta_{ш}(y_1^2) = [\gamma^2 - \zeta(2)]/2, \zeta_{ш}(y_1^3) = [\gamma^3 - 3\zeta(2)\gamma + 2\zeta(3)]/6$ *and* $\zeta_{ш}(y_1^4) = [80\zeta(3)\gamma - 60\zeta(2)\gamma^2 + 6\zeta(2)^2 + 10\gamma^4]/240$.

Let us consider (exponential) partial Bell polynomials partiels in the variables $\{t_l\}_{l\geq 1}$, $b_{n,k}(t_1, \ldots, t_{n-k+1})$, defined by the exponential generating series :

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n} b_{n,k}(t_1, \ldots, t_{n-k+1})\frac{v^n u^k}{n!} = \exp\left(u\sum_{l=1}^{\infty} t_l \frac{v^l}{l!}\right). \tag{33}$$

In particular, we have

Lemma 3. *Let* $t_m = (-1)^m(m-1)!\zeta_{\sqcup\!\sqcup}(m)$, *for* $m \geq 1$, *then*

$$\exp\left[\sum_{k\geq 1}\zeta_{\sqcup\!\sqcup}(k)\frac{(-y_1)^k}{k}\right] = 1+\sum_{n\geq 1}\left[\sum_{k=1}^{n} b_{n,k}(\gamma,\zeta(2),2\zeta(3),\ldots)\right]\frac{(-y_1)^n}{n!}.$$

Let us build the noncommutative generating series of $\zeta_{\sqcup\!\sqcup}(w)$ and let us take the constant part of the two members of $\mathrm{H}(N) \underset{N\to\infty}{\widetilde{\ }} \mathrm{Const}(N)\pi_Y Z$, we have

Proposition 6. *Let* $Z_{\sqcup\!\sqcup}$ *be the noncommutative generating series of the constants* $\zeta_{\sqcup\!\sqcup}(w)$, *i.e.* $Z_{\sqcup\!\sqcup} = \sum_{w\in Y^*}\zeta_{\sqcup\!\sqcup}(w)\, w$. *Then*

$$Z_{\sqcup\!\sqcup} = \left[1+\sum_{n\geq 1}\left(\sum_{k=1}^{n} b_{n,k}(\gamma,\zeta(2),2\zeta(3),\ldots)\right)\frac{(-y_1)^n}{n!}\right]\pi_Y Z.$$

Identifying coefficients of $y_1^k w$ in each member leads to

Corollary 4. *For all* $w \in Y^* \setminus y_1Y^*$ *and* $k \geq 0$, *we have*

$$\zeta_{\sqcup\!\sqcup}(y_1^k w) = \sum_{i=1}^{k}\frac{\zeta_{\sqcup\!\sqcup}(y_1^{k-i}w)}{i!}\left[(-1)^i\sum_{j=1}^{i} b_{i,j}(\gamma,\zeta(2),2\zeta(3),\ldots)\right].$$

Theorem 5. *In consequence, for* $w \in Y^* \setminus y_1Y^*, k \geq 0$, *the constant* $\zeta_{\sqcup\!\sqcup}(y_1^k w)$ *associated to* $\zeta(y_1^k w)$ *is a polynomial of de degree* k *in* γ *and with coefficients in* $\mathcal{Z}$. *Moreover, for* $l = 0,..,k$, *the coefficient of* γ^l *is of weight* $|w|+k-l$.

Corollary 5. *For* $s > 1$, *the constant* $\zeta_{\sqcup\!\sqcup}(1,s)$ *associated to* $\zeta(1,s)$ *is linear in* γ *and with coefficients in* $\mathbb{Q}[\zeta(2),\zeta(2i+1)]_{0<i\leq(s-1)/2}$.

Example 3. $\gamma = \dfrac{\zeta_{\sqcup\!\sqcup}(1,2)+2\zeta(3)}{\zeta(2)} = \dfrac{\zeta_{\sqcup\!\sqcup}(1,3)+\frac{1}{2}\zeta(2)^2}{\zeta(3)} = \dfrac{\zeta_{\sqcup\!\sqcup}(1,4)+3\zeta(5)-\zeta(2)\zeta(3)}{\frac{2}{5}\zeta(2)^2}$.

In other words, if we give to γ the weight[1] 1, then the constant $\zeta_{ш}(y_1^k w)$ associated to $\zeta(y_1^k w)$ would be an homogeneous polynomial of weight $|w|+k$.

Example 4.

$$\begin{aligned}
\zeta_{ш}(1,7) = &\quad \zeta(7)\gamma + \zeta(3)\zeta(5) - \tfrac{54}{175}\zeta(2)^4,\\
\zeta_{ш}(1,1,6) = &\quad \tfrac{4}{35}\zeta(2)^3\gamma^2 + \Big(\zeta(2)\zeta(5) + \tfrac{2}{5}\zeta(3)\zeta(2)^2 - 4\zeta(7)\Big)\gamma\\
&+\zeta(6,2) + \tfrac{19}{35}\zeta(2)^4 + \tfrac{1}{2}\zeta(2)\zeta(3)^2 - 4\zeta(3)\zeta(5),\\
\zeta_{ш}(1,1,1,5) = &\quad \tfrac{3}{4}\zeta(6,2) - \tfrac{14}{3}\zeta(3)\zeta(5) + \tfrac{3}{4}\zeta(2)\zeta(3)^2 + \tfrac{809}{1400}\zeta(2)^4\\
&-\Big(2\zeta(7) - \tfrac{3}{2}\zeta(2)\zeta(5) + \tfrac{1}{10}\zeta(3)\zeta(2)^2\Big)\gamma\\
&+\Big(\tfrac{1}{4}\zeta(3)^2 - \tfrac{1}{5}\zeta(2)^3\Big)\gamma^2 + \tfrac{1}{6}\zeta(5)\gamma^3.
\end{aligned}$$

2.2.3 Commutative generating series of polyzêtas

In Proposition 6, we explained how to extract the constant part of a non-commutative generating series of polyzêtas. Let us have a look now at following commutative generating series and corresponding to Ecalle's *Zig*-moulds [15] :

$$\begin{aligned}
\mathcal{Z}(t_1,\cdots,t_r) &= \sum_{s_1,\cdots,s_r>0} \zeta(s_1,\cdots,s_r)\, t_1^{s_1-1}\cdots t_r^{s_r-1} \qquad (34)\\
&= \sum_{n_1>\cdots>n_r>0} \frac{1}{(n_1-t_1)\cdots(n_r-t_r)}.
\end{aligned}$$

These commutative generating series can be encoded by series $\{S_j\}_{j=1,..,r}$ (or their projection over the alphabet Y) of the form [1]

[1] In the theory of periods, γ is conjectured to be non-period and so would be transcendent.

$$\begin{aligned} S_j =&(t_j x_0)^* x_1 \dots (t_r x_0)^* x_1 \qquad (35)\\ =& \sum_{s_j,\dots,s_r>0} x_0^{s_j-1} x_1 \dots x_0^{s_r-1} x_1 t_j^{s_j-1} \dots t_r^{s_r-1},\\ \pi_Y S_j =& \sum_{s_j,\dots,s_r>0} y_{s_j} \dots y_{s_r} t_j^{s_j-1} \dots t_r^{s_r-1} \qquad (36)\\ =&\left(\sum_{s_j\geq 1} y_{s_j} t_j^{s_j-1}\right) \dots \left(\sum_{s_r\geq 1} y_{s_r} t_r^{s_r-1}\right). \end{aligned}$$

Moreover, let $S_{r+1} = 1$. The series $\mathcal{Z}(t_1, \cdots, t_r)$ contain divergent terms of which we are looking for the constant part. We start from the following identity [1] due to convolution theorem [16]

$$S_1 = x_1^r + \sum_{j=1}^{r} t_j \sum_{i=0}^{j-1} x_1^i \sqcup x_0[(-x_1)^{j-1-i} \sqcup S_j], \qquad (37)$$

$$\Rightarrow \quad \pi_Y S_1 = y_1^r + \sum_{j=1}^{r} t_j \left(\sum_{s_j\geq 2} y_1^{j-1} y_{s_j} t_j^{s_j-1}\right) \pi_Y S_{j+1}. \qquad (38)$$

Proposition 7. *In consequence,*

$$\zeta_{\sqcup}(S_1) = \sum_{j=1}^{r} (-1)^{j-1} t_j \ \zeta[x_0(x_1^{j-1} \sqcup S_j)]. \qquad (39)$$

$$\begin{aligned} \zeta_{\text{ш}}(\pi_Y S_1) =&(-1)^r \sum_{\substack{s_1,\dots,s_r>0\\ s_1+\dots+rs_r=r}} \frac{(-\gamma)^{s_1}}{s_1!\dots s_r!} \prod_{j=2}^{r} \left(-\frac{\zeta(j)}{j}\right)^{s_j} \qquad (40)\\ &+ \sum_{j=1}^{r} \sum_{s_j\geq 2} t_j t_j^{s_j-1} \ \zeta_{\text{ш}}(y_1^{j-1} y_{s_j} \pi_Y S_{j+1}). \end{aligned}$$

In the following part, iterated integrals associated with words $w \in X^*$, along the path $0 \rightsquigarrow z$ and over differential forms ω_0, ω_1, will be denoted, as in [1], by $\alpha_0^z(w)$.

Example 5. *Since* $(tx_0)^*x_1 = x_1 + tx_0(tx_0)^*x_1$,

$$\alpha_0^z[(tx_0)^*x_1] = \alpha_0^z(x_1) + t\int_0^z \left(\frac{z}{s}\right)^t \mathrm{Li}_1(s)\frac{ds}{s} = \alpha_0^z(x_1) + t\sum_{n\geq 1}\frac{z^n}{n(n-t)}.$$

we get

$$\begin{aligned}
\sum_{s\geq 1}\zeta_{\sqcup\!\sqcup}(s)t^{s-1} &= \sum_{s\geq 2}\zeta(s)t^{s-1} = \sum_{n\geq 1}\left[\frac{1}{n-t}-\frac{1}{n}\right],\\
\sum_{s\geq 1}\zeta_{\sqcup\!\!\sqcup}(s)t^{s-1} &= \gamma+\sum_{s\geq 2}\zeta(s)t^{s-1} = \gamma+\sum_{n\geq 1}\left[\frac{1}{n-t}-\frac{1}{n}\right].
\end{aligned}$$

Example 6. *Identity (37) gives*

$$\begin{aligned}
S =& x_1^2 + t_1\, x_0(t_1x_0)^*x_1(t_2x_0)^*x_1 + t_2\, x_0\,(-x_1 \sqcup\!\sqcup (t_2x_0)^*x_1)\\
&+ t_2\, x_1 \sqcup\!\sqcup x_0(t_2x_0)^*x_1.
\end{aligned} \tag{41}$$

In the second member of the previous expression,

- *we have* $\zeta_{\sqcup\!\sqcup}(x_1^2) = 0$,
- *the first noncommutative rational series encodes, by convolution theorem [16], the following convergent integral*

$$\begin{aligned}
\alpha_0^1[(t_1x_0)^*x_0x_1(t_2x_0)^*x_1] &= \int_0^1\left(\frac{1}{s}\right)^{t_1}\frac{ds}{s}\int_0^s\frac{dr}{1-r}\sum_{n\geq 1}\frac{r^n}{n-t_2}\\
&= \sum_{n,m\geq 1}\frac{1}{(n+m)(n+m-t_1)(n-t_2)}\\
&= \sum_{n_1>n_2\geq 1}\frac{1}{n_1(n_1-t_1)(n_2-t_2)},
\end{aligned}$$

- *the following rational series encodes*

$$\begin{aligned}\alpha_0^1[x_0\,(-x_1 \sqcup\!\sqcup (t_2x_0)^* x_1)] &= -\int_0^1 \frac{ds}{s} \sum_{m\geq 1} \frac{s^m}{m} \sum_{n\geq 1} \frac{s^n}{n-t_2} \\ &= -\sum_{n,m\geq 1} \frac{1}{(n-t_2)(n+m)n} \\ &= -\sum_{n_1>n_2\geq 1} \frac{1}{n_1(n_1-n_2)(n_2-t_2)},\end{aligned}$$

- *the last noncommutative rational series corresponds to (but it has to be shuffled with x_1 for which $\zeta_{\sqcup\!\sqcup}(x_1)=0$)*

$$\alpha_0^1[x_0(t_2x_0)^* x_1] = \int_0^1 \frac{ds}{s} \sum_{n\geq 1} \frac{s^n}{n-t_2} = \sum_{n\geq 1} \frac{1}{n(n-t_2)}.$$

Thus,

$$\begin{aligned}\sum_{s,r\geq 1} \zeta_{\sqcup\!\sqcup}(s,r) t_1^{s-1} t_2^{r-1} = &\sum_{n_1>n_2\geq 1} \frac{t_1}{n_1(n_1-t_1)(n_2-t_2)} \\ &- \sum_{n_1>n_2\geq 1} \frac{t_2}{n_1(n_1-n_2)(n_2-t_2)}.\end{aligned} \tag{42}$$

Projecting S over alphabet Y, we get successively, since $y_1y_s = y_1 \sqcup\!\sqcup y_s - y_{s+1} - y_sy_1$

$$\begin{aligned}
\pi_Y S =& y_1^2 + \sum_{s,r\geq 2} y_s y_r t_1^{s-1} t_2^{r-1} + \sum_{s\geq 2} y_1 y_s t_2^{s-1} + \sum_{s\geq 2} y_s y_1 t_1^{s-1} \\
=& y_1^2 + \sum_{s,r\geq 2} y_s y_r t_1^{s-1} t_2^{r-1} + y_1 \sqcup\!\sqcup \sum_{s\geq 2} y_s t_2^{s-1} - \frac{1}{t_2}\left[y_1 + \sum_{s\geq 2} y_s t_2^{s-1}\right] \\
&+ \sum_{s\geq 2} y_s y_1 [t_1^{s-1} - t_2^{s-1}] \\
=& y_1^2 - y_1 t_2^{-1} + (y_1 - t_2^{-1}) \sum_{s\geq 2} y_s t_2^{s-1} + \sum_{s,r\geq 2} y_s y_r t_1^{s-1} t_2^{r-1} \\
&+ \sum_{s\geq 2} y_s y_1 [t_1^{s-1} - t_2^{s-1}].
\end{aligned}$$

In consequence,

$$\begin{aligned}
\sum_{s,r\geq 1} \zeta_{\sqcup\!\sqcup}(s,r) t_1^{s-1} t_2^{r-1} &= \frac{\gamma^2 - \zeta(2)}{2} - \frac{\gamma}{t_2} + (\gamma - t_2^{-1}) \sum_{n\geq 1} \left[\frac{1}{n - t_2} - \frac{1}{n}\right] \\
&+ \sum_{n_1 > n_2 \geq 1} \left[\frac{1}{n_1 - t_1} - \frac{1}{n_1}\right]\left[\frac{1}{n_2 - t_2} - \frac{1}{n_2}\right] \\
&+ \sum_{s\geq 2} \zeta(s,1)[t_1^{s-1} - t_2^{s-1}].
\end{aligned}$$

The last sum can be encoded by $x_0(t_1 x_0)^* x_1^2 - x_0(t_2 x_0)^* x_1^2$ *and can be obtained from*

$$\alpha_0^1[x_0(t_i x_0)^* x_1^2] = \int_0^1 \frac{ds}{s} \sum_{n_1 > n_2 \geq 1} \frac{s^{n_1}}{(n_1 - t_i) n_2} = \sum_{n_1 > n_2 \geq 1} \frac{1}{n_1 (n_1 - t_i) n_2}.$$

Appendix : application to polysystems [14]

Let $q_1, \ldots, q_n$ be commutative indeterminates over $\mathbb{C}$. We denote $Q = \{q_1, \ldots, q_n\}$. The algebra of formal power series (resp. polynomials) over Q with coefficients in $\mathbb{C}$ is denoted by $\mathbb{C}[[Q]]$ (resp. $\mathbb{C}[Q]$). An element of $\mathbb{C}[[Q]]$ is an infinite sum $f = \sum_{i_1,\ldots,i_n \geq 0} f_{i_1,\ldots,i_n} q_1^{i_1} \ldots q_n^{i_n}$.

Definition 3. *Let* $f \in \mathbb{C}[[Q]]$. *We set*

$$\begin{aligned} E(f) &= \{\rho \in \mathbb{R}_+^n : \exists C_{f\rho} \in \mathbb{R}_+ \text{ such that for all } i_1, \ldots, i_n \geq 0, \\ &\quad |f_{i_1,\ldots,i_n}| \rho_1^{i_1} \ldots \rho_n^{i_n} \leq C_{f\rho}\} \end{aligned}$$

$$\check{E}(f) : \text{interior of } E(f) \text{ in } \mathbb{R}^n.$$

$$\mathrm{CV}(f) = \text{convergence domain of } f = \{q \in \mathbb{C}^n : (|q_1|, \ldots, |q_n|) \in \check{E}(f)\}.$$

The power series f *is to be said* convergent *if* $\mathrm{CV}(f) \neq \emptyset$. *Let* $\mathcal{U}$ *be an open of* $\mathbb{C}^n$ *and let* $q \in \mathbb{C}^n$. *The power series* f *is to be said convergent on* q *(resp. over* $\mathcal{U}$*) if* $q \in \mathrm{CV}(f)$ *(resp.* $\mathcal{U} \subset \mathrm{CV}(f)$*). We set* $\mathbb{C}^{\mathrm{cv}}[[Q]] = \{f \in \mathbb{C}[[Q]] : \mathrm{CV}(f) \neq \emptyset\}$. *Let* $q \in \mathrm{CV}(f)$. *There exist some constants* $C_{f\rho}, \rho$ *and* $\check{\rho}$ *such that* $|q_1| < \check{\rho} < \rho, \ldots, |q_n| < \check{\rho} < \rho$ *and* $|f_{i_1,\ldots,i_n}| \rho^{i_1+\ldots+i_n} \leq C_{f\rho}$, *for* $i_1 \ldots, i_n \geq 0$. *The* convergence module *of* f *at* q *is* $(C_{f\rho}, \rho, \check{\rho})$.

Recall $D_1^{j_1} \ldots D_n^{j_n} f$ is the *partial derivation* of order $j_1, \ldots, j_n \geq 0$ of f and is given by

$$\frac{D_1^{j_1} \ldots D_n^{j_n} f}{j_1! \ldots j_n!} = \sum_{i_1 \geq j_1, \ldots, i_n \geq j_n} f_{i_1,\ldots,i_n} \prod_{l=1}^{n} \binom{i_l}{j_l} q_n^{i_l - j_l}.$$

Definition 4. *The* polysystem $\{A_i\}_{i=0,..,m}$ *is defined by the Lie derivations* $A_i = \sum_{j=1}^n A_i^j D_j$, *where* $A_i^j \in \mathbb{C}^{\mathrm{cv}}[[Q]]$. *It is* linear *if there exist* $\{M_i\}_{i=0,..,m} \in \mathcal{M}_{n,n}(\mathbb{C})$ *s.t.*

$$A_i = \begin{pmatrix} q_1 & \ldots & q_n \end{pmatrix} M_i \begin{pmatrix} D_1 \\ \vdots \\ D_n \end{pmatrix}.$$

Let $f \in \mathbb{C}^{\mathrm{cv}}[[Q]]$ and let $\{A_i\}_{i=0,1}$ be a polysystem. Let $(\rho, \check{\rho}, C_f)$ and let $(\rho, \check{\rho}, C_i)$, for $i = 0, 1$, be convergence modules of f and $\{A_i^j\}_{j=1,..,n}$ respectively at $q \in \mathrm{CV}(f) \Cap_{i=0,1,j=1,..,n} \mathrm{CV}(A_i^j)$. We denote by $(A_i f)_{|_q}$ the evaluation at q of $A_i f$. Let us consider the system

$$y(z) = f(q(z)), \text{ where } dq(z) = A_0(q)\omega_0(z) + A_1(q)\omega_1(z). \qquad (43)$$

Let us consider then the following generating series, $\sigma f_{|_{q(z_0)}}$ and the Chen series $S_{z_0 \rightsquigarrow z}$

$$\sigma f_{|_{q(z_0)}} = \sum_{w \in X^*} A_w f_{|_{q(z_0)}} w \quad \text{and} \quad S_{z_0 \rightsquigarrow z} = \sum_{w \in X^*} \alpha_{z_0}^z(w) w, \tag{44}$$

where $A_w = \mathrm{Id}, \alpha_{z_0}^z(w) = 1$ if $w = \epsilon$ and $A_w = A_v A_i, \alpha_{z_0}^z(w) = \int_{z_0 \rightsquigarrow z} \alpha_{z_0}^t(v)\omega_i(t)$ if $w = vx_i$.

Since $S_{z_0 \rightsquigarrow z}$ and $\mathrm{L}(z)\mathrm{L}(z_0)^{-1}$ satisfy (22) taking the same value at z_0 then $S_{z_0 \rightsquigarrow z} = \mathrm{L}(z)\mathrm{L}(z_0)^{-1}$. Hence, the asymptotic behaviour of L in (22) gives [6]

$$S_{\varepsilon \rightsquigarrow 1-\varepsilon} \sim e^{-x_1 \log \varepsilon} Z e^{-x_0 \log \varepsilon} \quad \text{for} \quad \varepsilon \to 0^+, \tag{45}$$

and the output y of (43) is given by $y(z) = \langle \sigma f_{|_{q(z_0)}} || S_{z_0 \rightsquigarrow z} \rangle = \sum_{w \in X^*} A_w f_{|_{q(z_0)}} \alpha_{z_0}^z(w)$ [14].

Let $\eta = q(z_0)$ and suppose that $f(q) = \lambda q$ with $\lambda \in \mathcal{M}_{1,n}(\mathbb{C})$. If $\{A_i\}_{i=0,1}$ is linear, then, by Definition 4, let $M_i = \mu(x_i)$, for $i = 0, 1$. Thus, $\sigma f_{|_{q(z_0)}} = \sum_{w \in X^*} [\lambda \mu(w) \eta] w$ is a rational power series of representation (λ, μ, η) and it is a generating series of the differential system of rank n, or equivalently of the linear differential equation of order n with singularities in $\{0, 1, \infty\}$.

Example 7 (hypergeometric equation)**.**

$$z(1-z)\ddot{y}(z) + [t_2 - (t_0 + t_1 + 1)z]\dot{y}(z) - t_0 t_1 y(z) = 0.$$

Let $q_1(z) = y(z)$ *and* $q_2(z) = z(1-z)\dot{y}(z)$*. One has*

$$\begin{pmatrix} dq_1 \\ dq_2 \end{pmatrix} = \left[\begin{pmatrix} 0 & 0 \\ -t_0 t_1 & -t_2 \end{pmatrix} \omega_0 - \begin{pmatrix} 0 & 1 \\ 0 & t_2 - t_0 - t_1 \end{pmatrix} \omega_1 \right] \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}.$$

Here $y = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$, $M_0 = -\begin{pmatrix} 0 & 0 \\ t_0 t_1 & t_2 \end{pmatrix}, M_1 = \begin{pmatrix} 0 & 1 \\ 0 & t_2 - t_0 - t_1 \end{pmatrix}$ *and* $\eta = \begin{pmatrix} q_1(z_0) \\ q_2(z_0) \end{pmatrix}$.

Thus, the solution of these equations can be obtained by examining the linear representation of generating series. The Drinfel'd equation allows to study the asymptotic behaviour, the functional equations and to compute the mondromy groups, the Galois differential groups.

References

[1] Hoang Ngoc Minh, Jacob, G., Oussous, N.E., and Petitot, M.: De l'algèbre des ζ de Riemann multivariées à l'algèbre des ζ de Hurwitz multivariées, *journal électronique du Séminaire Lotharingien de Combinatoire*, 44, 2001.

[2] Hardy, G.H.: *Divergent series*, Clarendon, 1949.

[3] Reutenauer, C.: *Free Lie Algebras*, London Math. Soc. Monographs, New Series-7, Oxford Science Publications, 1993.

[4] Berstel, J. and Reutenauer, C.: *Rational series and their languages*, Springer-Verlag, 1988.

[5] Hoffman, M.: The algebra of multiple harmonic series, *Journal of Algebra*, 1997.

[6] Hoang Ngoc Minh, Petitot, M. and Van der Hoeven, J.: Polylogarithms and Shuffle Algebra, *Proceedings of FPSAC'98*, 1998.

[7] Drinfel'd: Quasi-Hopf Algebras, *Leningrad Math. J.*, 1, 1419-1457, 1990.

[8] Drinfel'd: On quasitriangular quasi-hopf algebra and a group closely connected with $\mathrm{gal}(\bar{q}/q)$, *Leningrad Math. J.*, 4, 829-860, 1991.

[9] Hoang Ngoc Minh, Petitot, M., and Van der Hoeven, J.: L'algèbre des polylogarithmes par les séries génératrices, *Proceedings of FPSAC'99*, 1999.

[10] Hoang Ngoc Minh: Shuffle algebra and differential Galois group of colored polylogarithms, *Nuclear Physics B*, 2004.

[11] Lê, T.Q.T. and Murakami, J.: Kontsevich's integral for Kauffman polynomial, *Nagoya Math.*, 39-65, 1996.

[12] Costermans, C., Enjalbert, J.Y. and Hoang Ngoc Minh: Algorithmic and combinatoric aspects of multiple harmonic sums, *Discrete Mathematics & Theoretical Computer Science Proceedings*, 2005.

[13] Costermans, C., Enjalbert, J.Y., Hoang Ngoc Minh and Petitot, M.: Structure and asymptotic expansion of multiple harmonic sums, *Proceedings of International Symposium on Symbolic and Algebraic Computation*, 2005.

[14] Hoang Ngoc Minh: Finite polyzêtas, Poly-Bernoulli numbers, identities of polyzêtas and noncommutative rational power series, *Proceedings of* 4^{th} *International Conference on Words*, 232-250, 2003.

[15] Ecalle, J.: *La libre génération des multizêta et leur décomposition canonico-explicite en irréductibles*, séminaire IHP, 1999.

[16] Hoang Ngoc Minh: Summations of Polylogarithms via Evaluation Transform, *Mathematics and Computers in Simulations*, 42, 707-728, 1996.

Integrability of the Weyl-Lanczos and the Riemann-Lanczos Equations in General Relativity

P Dolan [1] A Gerber [2]

[1] Imperial College, Physics Department,
Quantum Optics and Laser Physics Group, London SW7 2BW, England
pdolan@inctech.com
[2] University of Rouen, LMRS Mathematics, UMR 6085 CNRS,
76801 Saint Etienne du Rouvray, France
agerber03@yahoo.co.uk

Abstract

Lanczos created two problems for systems of partial differential equations in general relativity, first in 1949 and then later in 1962. In his 1962 paper he presented the Weyl-Lanczos problem, where the Weyl conformal curvature tensor C_{abcd} is expressed by means of a third-order tensor potential $\tilde{L}_{abc}$. This system of PDEs is in involution and $\tilde{L}_{abc}$ satisfies a linear wave equation. In the next decade a second problem for systems of PDEs appeared in the relativity literature which is now called the Riemann-Lanczos problem. Here, the full Riemann curvature tensor R_{abcd} is determined from a three-index tensor potential L_{abc}. This problem is not in involution and requires prolongation to achieve involutivity. We give a review of both problems, the integrability of one and the attainment of integrability by the other.

Keywords: General Relativity, Curvature tensors, Lanczos Potentials, Integrability, Prolongation

1 Introduction

Lanczos [10, 11] made two attempts at deriving some or all of the Riemann curvature tensor R_{abcd} in general relativity from a three-index potential field

L_{abc}. His first attempt in 1949 [10] to derive all of the Riemann tensor from a potential L_{abc} failed. So, in 1962, he modified the problem to deriving only the conformally invariant part of the curvature tensor, C_{abcd}, or Weyl tensor from a potential tensor $\tilde{L}_{abc}$. This problem was a success and is now called the *Weyl-Lanczos problem*. In the period from the death of Lanczos in 1974 to 1983 there appeared in the Italian relativity literature [17] a Lanczos style curvature tensor problem which was simpler to present than either of Lanczos' attempts in 1949 or 1962. This problem is now called the *Riemann-Lanczos problem*. In this problem a tensor potential L_{abc} was proposed from which to derive all the Riemann curvature tensor components R_{abcd}. Lanczos' motivation in his 1962 paper [11] was a convincing set of analogies between electromagnetic field theory and gravitation in the form

$$A_a \longrightarrow \tilde{L}_{abc}\,, \tag{1}$$

$$F_{ab} = 2A_{[a,b]} \longrightarrow C_{abcd} = W_{abcd}(\tilde{L}_{fgh})\,, \tag{2}$$

$$\Box A_a + \cdots \longrightarrow \Box \tilde{L}_{abc} + \cdots\,, \tag{3}$$

where the last line indicates an analogy between the electromagnetic potential wave equation and a 'gravitational' type wave equation. In 1960, Penrose [16] discovered a wave equation for the Weyl tensor to give an analogy with electromagnetism which Lanczos missed due to lack of interest in the current literature

$$\Box F_{ab} + \cdots \longrightarrow \Box C_{abcd} + \cdots\,. \tag{4}$$

In 1982 Bampi and Caviglia [1], using a combination of Cartan's exterior differential systems (EDS) approach [3] and Vessiot's dual approach of vector fields, showed that the Weyl-Lanczos problem is a system of PDEs in involution. But they also showed that the Riemann-Lanczos problem *failed* to be in involution. However, in the following year, Bampi and Caviglia [2] returned to the Riemann-Lanczos problem and showed how to construct a suitable prolongation for this problem. The introduction of the methods of Cartan [3] and Vessiot into the relativity literature was a great

mathematical advance for problems that needed the greatest amount of generality for their resolution. Any re-examination of the work by Bampi and Caviglia is therefore best done through the *theory of jet bundles* over the spacetime manifold.

In the application of the theory of jet bundles the notion of *independent variables* is allowed to be much wider than it is usual in the general relativity literature and there are standard procedures to further widen this notion so as to incorporate necessary prolongations needed to make the system of PDEs reach involutivity.

Up to now the Lanczos curvature tensor problems have given rise to two dissimilar literatures:

(I) The Weyl-Lanczos literature can be easily related to many standard mathematical techniques of general relativity. There is a linear wave equation in every spacetime for the Lanczos potential $\tilde{L}_{abc}$ which becomes

$$\Box\tilde{L}_{abc} = 0 \tag{5}$$

in vacuo. This contrasts with Penrose's wave equation for the Weyl tensor which is *always* non-linear, even in empty space. In [7] Dolan and Kim derived Penrose's wave equation from Lanczos's wave equation inside matter in a manner which was completely general. However, the vacuum case given by (5) is of greater interest at present. References to the literature can be found in [15] of Dolan and Muratori.

(II) Since the pioneering papers of Bampi and Caviglia [1, 2] there was no jet bundle presentation of the Riemann-Lanczos problem until recently, when this subject was taken up again in articles such as [4, 5] or in [9]. This work, using the Lanczos potential tensor L_{abc} and its associated Monge variables, has shown that the Riemann-Lanczos system of PDEs is linear and remains linear after prolongation. The recent general theorem of Malgrange [12, 13] on prolongations tells us that at all generic points for the PDE system a finite number of prolongations will make the system attain

involutivity.

In papers by Massa and Pagani [14] and by Edgar [8], which is based on [14], there is a different approach which claims that non-linear 'constraints' arise in the Riemann-Lanczos problem. Their approaches, based on tensor calculus in local spacetime coordinate or frame methods, find the same 'non-linearities' which appear to make the Riemann-Lanczos problem insolvable in many general spacetimes. In [6] we show that the concept of 'partial derivative' that is used in such standard general relativity methods is inadequate for discussing existence problems of solutions for systems of PDEs, namely, whether the given system of PDEs is involutive or can be prolongated to involutivity.

2 The Weyl-Lanczos Problem

The Weyl conformal curvature tensor C_{abcd} has all the index symmetries of the Riemann tensor R_{abcd}, namely,

$$\begin{aligned} C_{abcd} &= -C_{bacd} = -C_{abdc} = C_{cdab} \,, \\ C_{a[bcd]} &= 0 \end{aligned} \tag{6}$$

as well as the vanishing of all its traces

$$C^{s}{}_{asb} = 0 \,. \tag{7}$$

A consequence of all these properties is the double-dual anti-symmetry

$${}^{*}C^{*}{}_{abcd} = -C_{abcd} \,. \tag{8}$$

In generating the Weyl tensor from a three-index tensor potential $\tilde{L}_{abc}$, we can get the index symmetries of equations (6) from the combination $\tilde{L}_{[ab][c;d]} + \tilde{L}_{[cd][a;b]}$ but to get the symmetries (7) and (8) we need the combination

$$C_{abcd} = \tilde{L}_{[ab][c;d]} + \tilde{L}_{[cd][a;b]} - {}^{*}\tilde{L}^{*}_{[ab][c;d]} - {}^{*}\tilde{L}^{*}_{[cd][a;b]} \,. \tag{9}$$

It has become traditional to add in the index symmetries

$$\tilde{L}_{abc} = \tilde{L}_{[ab]c} \,, \tag{10}$$

$$\tilde{L}_{[abc]} = 0 \,, \tag{11}$$

$$\tilde{L}_{as}{}^{s} = 0 \,. \tag{12}$$

If we evaluate all the double dual tensors in equation (9) we get

$$\begin{aligned} C_{abcd} = {} & \tilde{L}_{abc;d} - \tilde{L}_{abd;c} + \tilde{L}_{cda;b} - \tilde{L}_{cdb;a} + g_{bc}\tilde{L}_{(ad)} + g_{ad}\tilde{L}_{(bc)} \\ & - g_{bd}\tilde{L}_{(ac)} - g_{ac}\tilde{L}_{(bd)} + \frac{2}{3}\tilde{L}^{st}{}_{s;t}(g_{ac}g_{bd} - g_{ad}g_{bc}) \,, \end{aligned} \tag{13}$$

where $\tilde{L}_{ad} := \tilde{L}_{asd;}{}^{s} - \tilde{L}_{as}{}^{s}{}_{;d}$. There are 10 independent equations in (14) but equations (10), (11) and (12) tell us that we could have 16 independent components of $\tilde{L}_{abc}$.

The 6 differential gauge conditions

$$\tilde{L}_{abs;}{}^{s} = 0 \tag{14}$$

allow us to simplify equations (14) to

$$\begin{aligned} C_{abcd} = {} & \tilde{L}_{abc;d} - \tilde{L}_{abd;c} + \tilde{L}_{cda;b} - \tilde{L}_{cdb;a} + g_{bc}\tilde{L}_{sad;}{}^{s} + g_{ad}\tilde{L}_{sbc;}{}^{s} \\ & - g_{bd}\tilde{L}_{sac;}{}^{s} - g_{ac}\tilde{L}_{sbd;}{}^{s} \end{aligned} \tag{15}$$

or

$$C_{abcd} = W_{abcd}(x^{e}, \tilde{L}_{fgh}, \tilde{L}_{ijk,l}) \tag{16}$$

when the RHSs are now rewritten in terms of the partial derivatives $\tilde{L}_{ijk,l}$, the components of $\tilde{L}_{abc}$ and the Christoffel symbols. To this latter form (16) of the Weyl-Lanczos equations (16) we can apply the Janet-Riquier theory for systems of PDEs or else rewrite it as an EDS in the Cartan approach. The latter approach was taken by Bampi and Caviglia [1], where they showed that equations (16) represent a system in involution.

In [7] Dolan and Kim showed that the Weyl-Lanczos potential tensor $\tilde{L}_{abc}$ satisfies the linear tensor wave equation

$$\begin{aligned} J_{abc} = {} & \Box\tilde{L}_{abc} + 2R_{c}{}^{s}\tilde{L}_{abs} - R_{a}{}^{s}\tilde{L}_{bcs} - R_{b}{}^{s}\tilde{L}_{cas} - g_{ac}R^{ls}\tilde{L}_{lbs} \\ & + g_{bc}R^{ls}\tilde{L}_{las} - \frac{1}{2}R\tilde{L}_{abc} \,, \end{aligned} \tag{17}$$

where

$$J_{abc} = \frac{1}{2}R_{c[a;b]} - \frac{1}{6}g_{c[a}R_{;b]} \tag{18}$$

and $\Box\tilde{L}_{abc} = g^{sm}\tilde{L}_{abc;s;m}$. Penrose's wave equation for the Weyl tensor [16]

$$\Box C_{abcd} - C_{absm}C_{cd}{}^{sm} - 4C_{asm[c}C_{d]}{}^{sm}{}_{b} + \frac{1}{4}RC_{abcd}$$

$$= J_{[ab][c;d]} + J_{[cd][a;b]} - {}^{*}J^{*}_{[ab][c;d]} - {}^{*}J^{*}_{[cd][a;b]} \tag{19}$$

was derived by Dolan and Kim in [7] from the Lanczos wave equation (18) above thus completing the formal analogy with classical electromagnetism in equations (1), (2), (3) and (4).

3 The Riemann-Lanczos Problem

The first Lanczos paper of 1949 [10] attempted to find a three-index potential tensor field for the double dual of the Riemann curvature tensor ${}^{*}R^{*}_{abcd}$ in order to generate a wave equation for the potential. The complexity of this theory and the strangeness of the tensor ${}^{*}R^{*}_{abcd}$ to most relativists and also the lack of physical motivation for it was to lead to its abandonment. After Lanczos' death in 1974 Brinis Udeschini [17] introduced what was essentially an improved version of the 1949 theory by postulating the Riemann-Lanczos equations

$$R_{abcd} = L_{abc;d} - L_{abd;c} + L_{cda;b} - L_{cdb;a} \tag{20}$$

on the spacetime manifold M. We can rewrite this equation in the form

$$R_{abcd}(x^e) = K_{abcd}(x^e, L_{fgh}, L_{ijk,l}) \, , \tag{21}$$

where all the covariant derivatives have been written out in terms of partial derivatives $L_{ijk,l}$ and undifferentiated components of L_{abc}. We note that we can separate K_{abcd} into two distinct groups of terms as follows

$$K_{abcd}(x^e, L_{fgh}, L_{ijk,l}) = E_{abcd}(x^e, L_{fgh}) + F_{abcd}(L_{ijk,l}) \, , \tag{22}$$

where

$$\begin{aligned} E_{abcd} &= -\Gamma^n_{ad}(L_{nbc}+L_{ncb})+\Gamma^n_{bd}(L_{nac}+L_{nca})+\Gamma^n_{ac}(L_{nbd}+L_{ndb}) \\ &\quad -\Gamma^n_{bc}(L_{nad}+L_{nda}) \end{aligned} \tag{23}$$

and

$$F_{abcd} = L_{abc,d} - L_{abd,c} + L_{cda,b} - L_{cdb,a}\ . \tag{24}$$

The field components $L_{abc} = -L_{bac}$ give us 24 new independent field variables and the partial derivatives $L_{ijk,l}$ a further 96 new independent variables. Add to both of these new variables the 4 local coordinates (x^e) to give a total of $4 + 24 + 96 = 124$ independent variables in the jet bundle approach to the analysis of the Riemann-Lanczos problem on a 124-dimensional jet bundle. Bampi and Caviglia [1] showed that, in general, the Riemann-Lanczos PDE system cannot be in involution. However, a year later in 1983 they returned to the problem [2] and showed that, with a suitable prolongation, it could be made into a system in involution. Their work was based on Cartan's exterior differential systems (EDS) method with suitable Vessiot vector fields.

Using only essentially classical tensor calculus methods or their equivalent in terms of the covariant exterior differential operator Edgar [8] and Massa and Pagani [14] claimed to find non-linearities in the Riemann-Lanczos problem which made it inconsistent with general spacetime manifolds. An exhaustive analysis of the Riemann-Lanczos problem was recently made by Dolan and Gerber [6] using a jet bundle approach. Because of the splitting property of the quantities K_{abcd} given in equations (22), (23) and (24), which persist in comparable form into higher-order derivatives, no place in the calculations could be found in which non-linearities could appear in the original Riemann-Lanczos problem or in any derivative of its equations during prolongation. A recent theorem of Malgrange shows that all prolongations to involutivity [12, 13] will terminate in a finite number of steps in giving an involutory system at all generic points or else no such points exist.

4 Discussion

We gave a historic survey of the two problems called the Weyl-Lanczos problem and the Riemann-Lanczos problem. Both problems can be recast as an exterior differential system. The Weyl-Lanczos problem is in involution and many exact solutions of it exist for various spacetimes, see for example [15]. By contrast the Riemann-Lanczos problem is not in involution and needs prolongation. Prolongations are best carried out using the standard jet bundle approach and Monge variables. The Riemann-Lanczos system of PDEs is a linear system of PDEs and clearly stays that way during prolongation. Classical tensor calculus cannot exceed its 4 independent local coordinates (x^e). We pointed out that the jet bundle approach starts with 124 independent variables even before prolongation and the number of independent variables will increase considerably during prolongation. The alternative classical tensor-calculus-based approaches of the relativity literature [14, 8] miss (1) the linearity of the Riemann-Lanczos problem, which persists under all differentiations, and (2) the extra degrees of freedom which are necessary to locate integrability conditions.

References

[1] F. Bampi and G. Caviglia, Third-order tensor potentials for the Riemann and Weyl tensors, *General Relativity and Gravitation*, **15**, 375-386, (1983).

[2] F. Bampi and G. Caviglia, Third-order tensor potentials for the Riemann and Weyl tensors. II: singular solutions, *General Relativity and Gravitation*, **16**, 423-433, (1984).

[3] E. Cartan, *Les systémès différentiels extérieurs et leurs applications géométriques*, Hermann, Paris, (1945).

[4] P. Dolan and A. Gerber, Exterior differential systems, Janet-Riquier theory and the Riemann-Lanczos problems in two, three and four dimensions, *Journal of Mathematical Physics*, **44**, 3013-3034, (2003).

[5] P. Dolan and A. Gerber, The Weyl-Lanczos equations and the Lanczos wave equation in four dimensions as systems in involution, *Journal of Mathematical Physics*, **44**, 3035-3045, (2003).

[6] P. Dolan and A. Gerber, *The Riemann-Lanczos system of partial differential equations in general relativity*, preprint.

[7] P. Dolan and C.W. Kim, The wave equation for the Lanczos potential I, *Proceedings of the Royal Society London Series A* **447**, 557-575, (1994).
[8] S.B. Edgar, On effective constraints for the Riemann-Lanczos system of equations, *Journal of Mathematical Physics*, **44**, 5375-5385, (2004).
[9] A. Gerber, Prolongation methods and Cartan characters for the three-dimensional Riemann-Lanczos problem, *Journal of Mathematical Physics*, **44**, 3059-3070, (2003).
[10] C. Lanczos, Lagrange multiplier and Riemannian spaces, *Reviews of Modern Physics*, **21**, 497-502, (1949).
[11] C. Lanczos, The splitting of the Riemann tensor, *Reviews of Modern Physics*, **34**, 379-389, (1962).
[12] B. Malgrange, Generic involutiveness of analytic differential systems, *Comptes Rendus Acad. Sci. Paris Series I Math.*, **326**, (1998), pp. 863-866.
[13] B. Malgrange, Systèmes différentiels involutifs, *Panoramas et Synthèses*, **19**, Soc. Math. de France, (2005).
[14] E. Massa and E. Pagani, Is the Riemann tensor derivable from a tensor potential?, *General Relativity and Gravitation*, **16**, 805-333, (1984).
[15] P. Dolan and B.D. Muratori, The Lanczos potential for vacuum spacetimes with an Ernst potential, *Journal of Mathematical Physics*, **39**, 5406-5420, (1998).
[16] R. Penrose, A spinor approach to general relativity, *Annals of Physics*, **10**, 171-201, (1960).
[17] E. Brinis Udeschini, A third-order 'potential' tensor for the Riemann tensor. I, *Istituto Lombardo Accad. Sci. Lett. Rend. A*, **111**, 466-475, (1977).

Field Theories Found Geometrically From Embeddings in Flat Frame Bundles

Frank B. Estabrook

Jet Propulsion Laboratory,
California Institute of Technology,
4800 Oak Grove Drive, Pasadena, CA 91109
Frank.B.Estabrook@jpl.nasa.gov

Abstract

We present two families of exterior differential systems (EDS) for non-isometric embeddings of orthonormal frame bundles over Riemannian spaces of dimension $q = 2, 3, 4, 5....$ into orthonormal frame bundles over flat spaces of sufficiently higher dimension. We have calculated Cartan characters showing that these EDS satisfy Cartan's test and are well-posed dynamical field theories. The first family includes a constant-coefficient (cc) EDS for classical Einstein vacuum relativity ($q = 4$). The second family is generated only by cc 2-forms, so these are integrable (but nonlinear) systems of partial differential equations. These latter field theories apparently are new, although the simplest case $q = 2$ turns out to embed a ruled surface of signature (1,1) in flat space of signature (2,1). Cartan forms are found to give explicit variational principles for all these dynamical theories.

1 Introduction

We discuss two families of geometric field theories. By "geometric" we mean that these theories are given as exterior differential systems (EDS) for embedding of q-dimensional submanifolds R^q in flat homogeneous isotropic metric spaces E^N of higher dimension, say N. To formulate these EDS we in fact embed the orthonormal frame bundles over the submanifolds into the orthonormal frame bundles over the flat spaces, that is, into the groups ISO(N), which have dimension $N(N + 1)/2$. E. g., the $q = 4$

dimensional EDS are set using the 55 basis 1-forms of ISO(10). The fibers of the embedded bundles are subgroups of the O(N) fibers of ISO(N), thus inducing embedding maps of their q-dimensional bases R^q into the E^N bases of ISO(N).

By "field theories" we mean that each of these various EDS is shown, by an explicit numerical calculation of its Cartan characteristic integers [1] to have the property of being well-posed or, with the correct signature, of being"causal". This calculation uses a suite of *Mathematica* programs for EDS written by H. D. Wahlquist, and evaluates the Cartan characteristic integers with a Monte Carlo program to compute the ranks of the large matrices that arise [2]. The successive integers determine the dimension and well-posedness of the general solutions, and the Wahlquist programs also confirm the "involutory" property of certain fields in a solution, viz. those that can be adopted as independent variables.

A properly set EDS in a space with N variables is equivalent to a set of first order partial differential equations in N-q dependent variables, functions of q independent variables, and Cartan's technique of EDS is a deep approach to the Cauchy-Kowalewska analysis of such field theories.

Cartan's theory considers construction of a sequence of regular integral manifolds (of successively higher dimensions) of an EDS. His characteristic integers $s_i, i = 0, 1, .., q-1$, are calculated from the ranks of matrices that arise, and are diagnostic; they must pass Cartan's test [1] if the EDS (or an equivalent set of partial differential equations) is well posed. Then the final construction of the solution is determined solely from gauge and boundary data, and, at least in the analytic category, Cauchy existence and uniqueness are proved. We believe that, with proper attention to signature, the sets of partial differential equations following from such a well-posed EDS are those of a canonical field theory. Non-trivial embedding EDS that are well-posed, or causal, are not common. A key to their existence may be that for all the EDS we consider here we are also able to find Cartan q-forms from which the EDS may be derived by arbitrary variation.

There is a large literature, beginning with Lepage and Dedecker, on the

use of Cartan q-forms and their closure $q+1$-forms ("multisymplectic" forms). These are respectively the multidimensional field-theoretic extensions of classical Hamiltonian theory (the Cartan 1-form Ldt) and symplectic geometry. A short but essential bibliography can be found in Gotay [3]; cf. also Gotay et al [4], Bryant et al [5], Hermann [6] and Estabrook [7] [8]. The differential geometric setting for that work was for the most part (the structure equations of basis forms on) the first or second jet bundle over a base of q independent variables. Our use instead of basis forms and structure equations for embedding geometry, orthonormal frame bundles over flat metric geometries, is more in the spirit of string theory. It allows application of the variational techniques of field theory to the movable frames of general relativity, and can lead to interesting extensions.

In both these families of EDS we adapt the method usually used in the mathematical literature for isometric embedding, cf. e.g. [9], [10], in that we do not begin with a prior framing of the solution and prolong to higher bundles, but rather only use the bases of the embedding bundle. Such EDS have also been used in the theory of calibrated subspaces [1]. Perhaps this generalization of the customary isometric embedding can be called "dynamic embedding". The EDS that naturally arise are considerably more elegant, interpretable as field theories.

The Lie group ISO(N) (or one of its signature siblings ISO($N-1,1$) etc.) is the isometry group of N-dimensional flat space E^N (or a signature sibling). The group space is spanned by $P = N(N+1)/2$ canonical vector fields, and by a dual basis of left-invariant 1-forms that we first denote by $\theta^\mu, \mu = 1...N$, corresponding to translations, and $\omega^\mu{}_\nu$, that will correspond to rotations. Now the structure equations for general movable frames over an N-dimensional manifold are usually written covariantly (on the second frame bundle) as

$$d\theta^\mu + \omega^\mu{}_\nu \wedge \theta^\nu = 0 \quad (1)$$

$$d\omega^\mu{}_\nu + \omega^\mu{}_\sigma \wedge \omega^\sigma{}_\nu + R^\mu{}_\nu = 0. \quad (2)$$

These become the Cartan-Maurer equations of ISO(N) or one of its siblings when the curvature 2-forms $R^{\mu}{}_{\nu}$ are put equal to zero, and upper indices are systematically lowered using (for signature) a non-singular matrix of constants $\eta_{\mu\nu}$, after which imposing antisymmetry (orthonormality) $\omega_{\mu\nu} = -\omega_{\nu\mu}$. These structure equations then describe $N(N-1)/2$-dimensional rotation groups as fibers over N-dimensional homogeneous spaces E^N. (The $\eta_{\mu\nu}$, and other possible signatures in E^N, are often conveniently ignored in the following, and can be inserted later.)

We will write the two families of EDS using partitions $(n, m), n + m = N$, of the basis forms of ISO(N) into classes labeled respectively by the first n indices i, j, etc. $= 1, 2, ...n$ and the remaining indices A, B, etc. $= n + 1, n + 2, ...N$. So the basis forms are θ^i, θ^A, and, after lowering an index, $\omega_{ij} = -\omega_{ji}$, $\omega_{AB} = -\omega_{BA}$, $\omega_{iA} = -\omega_{Ai}$. Summation conventions on repeated indices will be used separately on each partition. The structure equations (1) (2) before lowering become

$$
\begin{aligned}
d\theta^i + \omega^i{}_j \wedge \theta^j &= -\omega^i{}_A \wedge \theta^A && (3)\\
d\theta^A + \omega^A{}_B \wedge \theta^B &= -\omega^A{}_i \wedge \theta^i && (4)\\
d\omega^i{}_j + \omega^i{}_k \wedge \omega^k{}_j &= -\omega^i{}_A \wedge \omega^A{}_j && (5)\\
d\omega^A{}_B + \omega^A{}_C \wedge \omega^C{}_B &= -\omega^A{}_i \wedge \omega^i{}_B && (6)\\
d\omega^i{}_A + \omega^i{}_j \wedge \omega^j{}_A + \omega^i{}_B \wedge \omega^B{}_A &= 0. && (7)
\end{aligned}
$$

The terms we have put on the right are interpreted as torsions and curvatures induced by an embedding; we will use them to set the EDS.

In Sections 2 and 3 we calculate the Cartan characteristic integers of the embedding EDS for the two families. We will report the results in a short tabular form: $P\{s_0, s_1, ...s_{q-1}\}q + CC$. This gives first the dimension P of the space in which we set the EDS, i.e. the total number of basis forms with whose structure equations we begin, then the series of Cartan integers found, $\{s_0, s_1, ...s_{q-1}\}$. q is the dimensionality of the base space of a general solution. Finally CC is the number of Cartan characteristic vectors (the number of auxiliary fields allowing us to write a cc system, fibers corresponding to variables that could in principle be eliminated from

the EDS). Cartan denotes $q + CC$ by g, the genus. The ultimately simple Cartan test showing the EDS to be well-posed and causal, calculated from these, is derived in [1] and the literature cited there. Here the test is simply that these integers satisfy $P - \sum_{i=0}^{q-1} s_i - q - CC := s_q \geq 0$. We will always have $s_q = 0$, which a physicist interprets as absence of a gauge group, and so according to Cartan solutions will depend on s_{q-1} functions of q-1 variables. We denote these theories as causal but that requires also adjusting the signatures, so that the final integration of solutions from this boundary data is hyperbolic.

It is a classic result [10] that smooth local embedding of Riemannian geometries of dimension $q = 3, 4, 5...$ is always possible into flat spaces of dimension respectively $N = q(q+1)/2 = 6, 10, 15...$, which motivates the partitions of our first family, viz. $(n, m) = (3, 3), (4, 6), (5, 10), ...$ The causal EDS we give determine submanifolds of ISO(N) which are themselves $O(n) \otimes O(m)$ bundles fibered over $q = n$-dimensional base spaces, say R^q and induce maps of these into E^N. The n θ_i remain independent ("in involution") when pulled back to a solution bundle, satisfying the structure equations of an orthonormal basis in any cross section, and Equations (5) and (7) express embedding relations that go back to Gauss and Codazzi. The solution bundle metric is the pullback of $\theta_i \theta_i$. We will present in Section 2 the family of Einstein-Hilbert Cartan forms from which the EDS of this first family are derived by variation. The EDS will require zero torsion for the θ_i but not insist on aligning the solutions with these orthonormal frames (the θ_A are *not* included in the EDS so it is not necessarily "isometric"), and from the induced map of bundle bases there is also a less interesting "ghost" metric which is the pullback of $\theta_i \theta_i + \theta_A \theta_A$. The induced curvature 2-forms are required by the EDS to satisfy "horizontality" 3-form conditions and also to have vanishing Ricci $n-1$-forms.

The field theories of our second family, of dimension $q = 2, 3, 4, 5...$ also arise from embeddings into flat spaces E^N of dimension $N = 3, 6, 10, 15, ...$ but the EDS use different partitions, viz.., $(n, m) = (1, 2), (2, 4), (3, 7),$ $(4, 11)$, etc. Solutions are ($O(n) \otimes O(m)$ bundles over) geometries of di-

mension $q = n + 1$ and can be called n-branes. They have rulings that are flat n-spaces. The EDS are generated only by cc sets of 2-forms (for vanishing torsion of both partitions) and are so-called "integrable systems". Again the embedding is dynamic, the partitioned frames are not required to be an orthonormal framing of the solution manifolds. In Section 3 we give the EDS and report the calculated Cartan characters showing them to be causal. The n θ_i when pulled back into a solution both determine a Riemannian submersion and geodesic slicing. This is either a theory of relativistic rigidity or perhaps of a Kaluza-Klein gravitational field, depending on N and the signature adopted. Cartan forms for those EDS are easily found.

As a sole illustration of introduction of explicit coordinates into such a frame bundle EDS, the simplest of these non-isometric geometric field theories, that based on partition $(1, 2)$, is integrated in Section 4. Its solutions turn out to be classically known, in the guise of geodesically ruled surfaces in E^3. We have only changed signature to show it as a stringy field causally evolving in time.

2 Einstein-Hilbert Action

The EDS of our first family arise from Cartan n-forms on ISO(N) expressing the Ricci scalars of $q = n$ dimensional submanifolds of E^N,

$$\Lambda = R_{ij} \wedge \theta_k \wedge ...\theta_p \epsilon_{ijk...p}, \tag{8}$$

where from the Gauss structure equation, Eq. (5), $2R_{ij} := -\omega_{iA} \wedge \omega_{jA}$ is the induced Riemann 2-form. The exterior derivative of the n-form field Λ on ISO(N), using Eq. (3) and (7), is quickly calculated to be the $n+1$-form (closed, multisymplectic)

$$d\Lambda = \theta_A \wedge \omega_{Ai} \wedge R_{jk} \wedge \theta_l \wedge ...\theta_p \epsilon_{ijkl...p}. \tag{9}$$

This $n + 1$-form is a sum of products of the m 1-forms θ_A and the m n-forms $\omega_{Ai} \wedge R_{jk} \wedge \theta_l \wedge ...\theta_p$ $\epsilon_{ijkl...p}$. A variational isometric embedding EDS

is generated by the m θ_A, their exterior derivatives for closure, and the m n-forms, since any vector field contracted on $d\Lambda$ yields a form in the EDS. That is, up to boundary terms, the arbitrary variation of Λ vanishes on solutions. We previously calculated Cartan's characteristic integers for these isometric embedding EDS showing them to be well set and causal [11] [12]. We denoted them as being "constraint free" geometries. Isometric embedding formulations of the Ricci-flat field equations then are obtained by adding in the closed $n-1$-forms for Ricci-flatness as constraints. The augmented EDS are again calculated to be causal. To be explicit, for partition (4, 6) the constraint-free Cartan character table was 55 $\{6, 6, 6, 12\}q = 4 + 21$ which, with the augumentation with four 3-forms became 55 $\{6, 6, 10, 8\}q = 4 + 21$. We now see that formulation as nevertheless somewhat unsatisfactory as field theory, since the Einstein-Hilbert action appears to have lead to equations which in fact mostly follow from the imposed constraints.

We have however noticed that there is another variational EDS belonging to a different quadratic factoring of the multisymplectic forms $d\Lambda$, Eq. (9). The θ_i will frame a Riemannian metric on an embedded space of dimension n so long as the induced torsion 2-forms of Eq.(2), $\omega_{iA} \wedge \theta_A$, vanish, and these factor Eq.(9) term-by-term, as products with the n $n-1$-forms for Ricci-flatness. The exterior derivatives of the torsion terms must be included; these are sometimes called conditions for horizontality. In sum, we have considered the following closed EDS (which now do *not* include the mathematically customary isometric embedding 1-forms θ_A)

$$(\omega_{iA} \wedge \theta_A, R_{ij} \wedge \theta_j, R_{ij} \wedge \theta_k \wedge ...\theta_l \epsilon_{ijk...lp}) \tag{10}$$

When n=4 this EDS is an exact parallel to the EDS for a moving frame formulation of vacuum relativity that used the 44 traditional intrinsic coordinates of tetrad frames and connections over 4-space, and had 10 gauge freedoms [13]. It had the same Cartan character table but no CC. The present formulation is set with more variables, viz. 55, but its solutions have 21 CC fibers (since ω_{ij} and ω_{AB} do not enter explicitly) and no gauge freedom; moreover it has the elegance of a cc EDS (no coordinate functions

appear in the generating forms) [14] [15] [16].

The calculation shows the EDS Eq.(10) to be well set and causal systems for embedding of $O(n) \otimes O(m)$ bundles over n space, for the partitions (3, 3), (4, 6), (5, 10) etc. as stated in the introduction. The embedding dimension, the computed Cartan characters, dimensionality and $O(n) + O(m)$ fiber dimension (CC) of the solutions for these cases are respectively $21\{0, 6, 3\}3 + 9, 55\{0, 4, 12, 14\}4 + 21, 120\{0, 5, 10, 20, 25\}5 + 55$, etc. The base spaces of the fibered solution manifolds are spanned by the 1-forms θ_i; evidently a solution is a bundle of orthonormal frames belonging to the Ricci-flat Riemannian connection ω_{ij}. The metric is $\theta_i\theta_i$. There is also present in the base space R^n another metric pulled back from the induced embedding of it in the base space E^N about which we know little: $\theta_i\theta_i + \theta_A\theta_A$. It is a ghost tensor field, perhaps with only indirect influence. The ideals we are writing are set on ISO(N), and their solutions are frame bundles embedded in ISO(N), and the induced embeddings of the base spaces seem to be of less interest.

The ideal Eq.(10) is contained in the augmented embedding ideal we have previously used, so solutions of the latter will be solutions of the former. This would seem to imply that our new dynamic embedding ideal will have additional solutions; indeed it implies fewer partial differential equations than does the isometric embedding ideal augmented with constraints for Ricci-flat geometry. Perhaps so-called singular solutions of the isometric embedding ideal–solutions which are not regular, that is, obtained by Cartan's sequential integrations–now appear as regular solutions, which could make this new formulation important for local numerical computation from boundaries.

3 Torsion-free n-brane Embedding

We have searched whether the torsion 2-forms induced in *both* the local partitions can together be taken as an EDS: ($\omega_{iA} \wedge \theta_A$, $\omega_{iA} \wedge \theta_i$). It can easily be checked that it is closed, and calculation of the characteristic integers indeed showed that *for just the values of (n, m) of the second*

family described in the introduction these EDS are causal, with $q = n + 1$ and fibers $O(n) \otimes O(m)$, dim $n(n-1)/2 + m(m-1)/2$. The results for the first five EDS are: $(n, m) = (1, 2)$, $6\{0, 3\}2 + 1$; $(2, 4)$, $21\{0, 6, 5\}3 + 7$; $(3, 7)$, $55\{0, 10, 9, 8\}4 + 24$; $(4, 11)$, $120\{0, 15, 14, 13, 12\}5 + 61$; $(5, 16)$, $231\{0, 21, 20, 19, 18, 17]6 + 130$; and the pattern seems evident.

Now well set EDS for geodesic flat dimension n submanifolds of flat N spaces are generated, using the partition (n, m), by the closed ideal of 1-forms (θ_A, ω_{Ai}). For example, if $N = 3$ and $n = 1$ and $m = 2$, geodesic lines in flat 3-space, the Cartan characteristic integers are $6\{4\}1 + 1$. If $N = 4$, for partition $(1, 3)$ we find $10\{6\}1 + 3$ (in all cases ω_{ij} and ω_{AB} give the Cauchy characteristic fibers). Similarly, the EDS for flat 2-dimensional submanifolds of flat N spaces are generated by the 1-forms with partitions $(2, N-2)$. For example if $N = 5$, $(n, m) = (2, 3)$, and the character table is $15\{9, 0\}2 + 4$. When $N = 6, (n, m) = (2, 4)$ and $21\{12, 0\}2 + 7$. The zeros can be ascribed to a gauge freedom. These constructions clearly continue. Our new torsion-free EDS $(\omega_{iA} \wedge \theta_A,\ \omega_{Ai} \wedge \theta_i)$ are contained in (θ_A, ω_{Ai}), so we see that the q-dimensional solutions of the torsion-free embedding theory must contain flat geodesic fibers of dimension $n = q - 1$. Thus the solutions are *ruled* spaces,

In a solution the θ_i remain independent (are "in involution") but fall short by one of being a complete basis. In addition to the slicing, they define there a vector field, say V, of arbitrary normalization (a congruence), by the relations $V \cdot \theta_i = V \cdot \omega_{ij} = V \cdot \omega_{AB} = 0$. Contracting V on the second torsion 2-form, since the θ_i remain linearly independent, gives also $V \cdot \omega_{iA} = 0$. It follows that the Lie derivatives with respect to V of θ_i, ω_{ij} and R_{ij} vanish on solutions. They live in (and are lifted from) an n dimensional quotient space of the solution, with metric $\theta_i\theta_i$ and Riemann tensor $\omega_{iA} \wedge \omega_{Aj}$. Cross sections of this quotient map are the rulings, geodesic n-dimensional subspaces calibrated by the volume form $\theta_i \wedge \theta_j ... \wedge \theta_k$.

In an earlier time we have discussed the problem of defining a rigid body in special and general relativity [17]. The kinematic quotient-space definition of rigidity due originally to Max Born (Riemannian submersion)

was shown by Herglotz and Noether to have only three degrees of freedom: the only Born-rigid congruences which were rotating (had vorticity) in Minkowski space were isometries of the space-time without time evolution. We showed this to be the case also for kinematic or "test" rigid bodies moving in vacuum Einstein spaces. It seemed to be impossible then to sensibly discuss the so-called "dynamic" rigid bodies envisioned by Pirani, which were to carry their own 3-dimensional geometry while distorting space-time. We are charmed by having now arrived at space-times, using dynamic embedding in the (3, 7) partition, having the greater dynamical freedom allowed by separation of the rôles of the induced 3-metrics in the cross sections and quotient space of a solution.

In the (4, 11) partition, the solutions are five dimensional, with a dynamically rigid congruence that projects to a metric quotient 4-space. This may be a well-posed causal variant of Kaluza-Klein theory, and merits further investigation.

Closed EDS generated only by cc 2-forms have a special structure, inasmuch as they can be equivalent to dual infinite Lie algebras of Kaĉ-Moody type and lead to hierarchies of so-called integrable systems. Lie groups have a duality between 2-form Cartan-Maurer structure equations for basis 1-forms and Lie commutator products of dual basis vector fields. This duality persists when the additional cc 2-forms of an EDS are imposed, but then the vector commutator table is incomplete. New vectors can be introduced in terms of the unknown commutators, and then more commutators calculated using the Jacobi identities. These allow adding 2-form structure equations for new dual 1-forms in higher dimensional spaces. If this expansion terminates, an embedding in a group has been found, the new 1-forms being potentials that integrate the original EDS. If the expansion continues, it leads to a Kaĉ-Moody algebra of finite growth. Such EDS belong to so-called integrable systems of partial differential equations. The prototype of this is the well-known Korteweg-de Vries equation, which both leads to [18], and belongs to, the hierarchy of the infinite Lie algebra ${A_1}^{(1)}$derived from SL(2, R). The Kaĉ-Moody algebras dual to our embed-

ding EDS remain to be worked out.

Finally, although we did not derive these EDS variationally, Cartan forms are easily found, at least for even dimensions. In particular, in the (3, 7) theory either the 2-forms $\tau_A = \omega_{Ai} \wedge \theta_i$ or $\sigma_i = \omega_{iA} \wedge \theta_A$ can be used to write a quadratic Cartan form as in some Yang-Mills theories:

$$\Lambda = \tau_A \wedge \tau_A, \text{ so } d\Lambda = 2\tau_A \wedge \omega_{Ai} \wedge \sigma_i \tag{11}$$

Every term of $d\Lambda$ contains both a τ_A and a σ_i so arbitrary variation yields the EDS. We also note that $\tau_A \wedge \tau_A + \sigma_i \wedge \sigma_i$ is exact.

4 The Partition (1, 2)

We will set this EDS on the frame bundle ISO(1, 2) over a flat 3-space with signature (-, +, -), so the structure equations of the bases are

$$\begin{aligned}
d\theta_1 + \omega_{12} \wedge \theta_2 + \omega_{31} \wedge \theta_3 &= 0 && (12)\\
d\theta_2 + \omega_{12} \wedge \theta_1 - \omega_{23} \wedge \theta_3 &= 0 && (13)\\
d\theta_3 - \omega_{31} \wedge \theta_1 - \omega_{23} \wedge \theta_2 &= 0 && (14)\\
d\omega_{12} - \omega_{31} \wedge \omega_{23} &= 0 && (15)\\
d\omega_{23} - \omega_{12} \wedge \omega_{31} &= 0 && (16)\\
d\omega_{31} + \omega_{23} \wedge \omega_{12} &= 0, && (17)
\end{aligned}$$

and the EDS to be integrated is generated by the three 2-forms $\omega_{iA} \wedge \theta^A, \omega_{Ai} \wedge \theta^i, i = 1, A = 2, 3$:

$$\omega_{12} \wedge \theta_2 + \omega_{31} \wedge \theta_3 \tag{18}$$

$$\omega_{12} \wedge \theta_1 \tag{19}$$

$$\omega_{31} \wedge \theta_1. \tag{20}$$

The characteristic integers are 6{0, 3} q =2 and CC = 1; O(2) fiber (since ω_{23} is not present). To introduce coordinates - scalar fields - we will successively prolong the EDS with potentials or pseudopotentials, checking at each step that it remains well-set and causal.

First, it is obvious that there is a conservation law, a closed 2-form that is zero mod the EDS, viz. $d\theta_1$. So we adjoin the 1-form

$$\theta_1 + dv, \tag{21}$$

introducing the scalar potential v. The characters are now $7\{1,3\}2+1$. Next we specialize to a particular, convenient, fiber cross-section making a choice of frame: we introduce two new fields ζ and η while prolonging with three 1-forms taken so that the original 2-forms in the EDS vanish (they have been "factored")

$$\omega_{12} - \zeta\theta_1 \tag{22}$$
$$\omega_{13} - \eta\theta_1\zeta \tag{23}$$
$$\zeta\theta_2 - \eta\theta_3 + (\eta+\zeta)\theta_1. \tag{24}$$

To maintain closure, however, three new 2-forms, exterior derivatives of these or algebraically equivalent, must also be adjoined:

$$(d\zeta - \eta\omega_{23}) \wedge dv \tag{25}$$
$$(d\eta - \zeta\omega_{23}) \wedge dv \tag{26}$$
$$(\eta d\zeta - \zeta d\eta) \wedge (\theta_2+\theta_3) - (\eta+\zeta)\omega_{23} \wedge (\eta\theta_2 - \zeta\theta_3). \tag{27}$$

Now we have $9\{4,3\}2$ with no CC. ω_{23} now appears in the EDS, but is conserved, $d\omega_{23} = 0$ mod EDS. Thus, we can introduce a pseudopotential variable x , and then further find another conserved 1-form and a final pseudopotential u. Which is to say we can adjoin

$$\omega_{23} - dx \tag{28}$$
$$\theta_2 + \theta_3 - e^x du, \tag{29}$$

without adding any 2-forms to the EDS. We have a total of 11 basis 1-forms: six in θ_i, θ_A, ω_{AB}, ω_{iA}, plus $d\zeta, d\eta, dx, du, dv$, and an EDS with $11\{6,3\}2$. The pulled-back original six bases are now all solvable in terms of coordinate fields on the solutions, and can be eliminated: $5\{0,3\}2$. We have

eliminated the CC. This is equivalent to a set of first order partial differential equations in 3 dependent variables and 2 independent variables. From the character table, we expect solutions to involve 3 arbitrary functions of 1 variable.

Taking x and v as independent in the solution, we can solve the first two 2-forms in Eq. (25) and (26) for η and ζ:

$$\eta = ae^x + be^{-x} \tag{30}$$
$$\zeta = ae^x - be^{-x}, \tag{31}$$

where a and b are arbitrary functions of v. The third 2-form then amounts to

$$e^x = 1/2(b/a)'\partial_x u \ , \tag{32}$$

which integrates to

$$e^x = 1/2B'(u - A(v)) \ . \tag{33}$$

We have put $b/a = B(v)$ and prime is derivation with respect to v.

The three arbitrary functions of v, a, b and A, give the general solution. On it the pulled-back bases of E^3 (no longer orthonormal or independent) are

$$\theta_1 = -dv \tag{34}$$
$$\theta_2 = dv + \frac{ae^x + be^{-x}}{2a}du \tag{35}$$
$$\theta_3 = -dv + \frac{ae^x - be^{-x}}{2a}du, \tag{36}$$

and the induced 2-metric from E^3 is

$$g = -\theta_1\theta_1 + \theta_2\theta_2 - \theta_3\theta_3 \tag{37}$$
$$= Bdu^2 + B'(u - A)dvdu - dv^2 \ . \tag{38}$$

This is, up to signature, the metric found classically from the construction of geodesically ruled surfaces in E^3, cf, e.g., Eisenhardt [10]. The surfaces are intrinsically characterized by a "line of striction", the locus

$u - A(v) = 0$, and a "parameter of distribution" $2B/B'$. The geodesic rulings, on which θ_2, θ_3, ω_{12}, and ω_{13} pull back to vanish, are the set of lines $u = \text{const}$. The rigid congruence is the set of lines on which V contracted with θ_1, ω_{12} and ω_{13} vanishes, hence $v = \text{const}$.

ACKNOWLEDGEMENTS

This work has at every point benefitted from conversations with my long time colleague Hugo D. Wahlquist. It follows on several lines of research we have jointly pursued. Wahlquist first suggested the new non-isometric. embedding EDS for vacuum relativity ($q = 4$). The Cartan characters reported were all obtained using his suite of *Mathematica* programs for manipulation of EDS. The author acknowledges a visiting appointment at the Jet Propulsion Laboratory, California Institute of Technology

References

[1] Ivey, T. A. and Landsberg, J. M., *Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems (Graduate Studies in Mathematics vol 61)* (American Mathematical Society: Providence, RI), 2003.

[2] Wahlquist, H. D., Monte Carlo calculation of Cartan characters in N. Breton, R. Capovilla and T.Matos, Eds. *Proceedings of the International Conference of Aspects of General Relativity and Mathematical Physics* (Ediciones CINVESTAV: Mexico DF), 168, 1994.

[3] Gotay, M. J., An exterior differential system approach to the Cartan form, symplectic geometry and mathematical physics, in P. Donato, J. Elhadad and G. M. Tuynmen, Eds., *Proc. Colloq. in Honor of J-M Souriau (Aix-en-Provence, 1990) (Progress in Mathematics)* (Birkhaüser: Boston MA), 99, 1991.

[4] Gotay, M. J., Isenberg, J., Marsden, J. and Montgomery, R., Momentum maps and classical relativistic fields, Part I: Covariant field theory, arXiv:physics/980901, 1998.

[5] Bryant, R., Griffiths, P. and Grossman, D., *Exterior Differential Systems and Euler-Lagrange Partial Differential Equations* (The University of Chicago Press: Chicago), 2003.

[6] Hermann, R., Differential form methods in the theory of variational systems and Lagrangian field theories, *Acta Appl. Math.**12***, 35-78, 1988.

[7] Estabrook, F. B., Differential geometry as a tool for applied mathematics, in R. Martini, Ed. *Lecture Notes in Math. **810*** (Springer: Berlin), 1980.

[8] Estabrook, F. B., Conservation laws for vacuum tetrad gravity, *Class. Quantum Grav.* ***114***, 2841-2848, 2006.

[9] Bryant, R., Griffiths, P. A.. and Yang, D., Characteristics and existence of isometric embeddings, *Duke Math. J.* ***50***, 893-994, 1983.

[10] Griffiths, P. A. and Jensen, G. R., Differential systems and isometric embeddings, *Annals of Mathematical Studies* ***114*** (Princeton University Press: Princeton NJ), 1987.

[11] Estabrook, F. B. and Wahlquist, H. D., Classical geometries defined by exterior differential systems on higher frame bundles, *Class. Quantum Grav.* ***6*** , 263-274, 1989.

[12] Estabrook, F. B., Robinson, R. S. and Wahlquist, H. D., Constraint-free theories of gravitation, *Class. Quantum Grav.* ***16*** , 911-918, 1999.

[13] Estabrook, F. B., Mathematical structure of tetrad equations for vacuum relativity, *Physical Review* ***D 71***, 044004, 2005.

[14] Estabrook, F. B., Moving frames and prolongation algebras, *Journal of Mathematical Physics* ***23***, 2071-76, 1982.

[15] Harrison, B. K., Ernst equation Backlund transformations revisited: New approaches and results, in Hu Ning, Ed. *Proceedings of the Third Marcel Grossman Meeting on General Relativity* (North-Holland, Amsterdam), 29-39, 1983.

[16] Harrison, B. K., Prolongation structures and differential forms, in C. Hoenselaers and W. Dietz, Eds.,*Solutions of Einstein's Equations: Techniques and Results, Lecture Notes in Physics* #205 (Springer, Berlin), 26-54, 1984.

[17] Wahlquist, H. D. and Estabrook, F. B., Rigid motions in Einstein space, *J. Math. Phys.* ***7,*** 894-905, 1966.

[18] Estabrook, F. B., Differential geometry techniques for sets of nonlinear partial differential equations, in R. Conte and N. Boccara, Eds., *Partially Integrable Evolution Equations in Physics* (Kluwer: Dordrecht), 413-434, 1990.

[19] Eisenhart, L. P., *A Treatise on the Differential Geometry of Curves and Surfaces* (Dover: New York), 241 *et seq*, 1909, 1960.

Vessiot Connections of Partial Differential Equations[1]

Dirk Fesser [1] Werner M. Seiler [2]

[1] Interdisziplinäres Zentrum für Wissenschaftliches Rechnen,
Universität Heidelberg, Im Neuenheimer Feld 368
69120 Heidelberg, Germany
dirk.fesser@iwr.uni-heidelberg.de
[2] Computational Mathematics Group, Universität Kassel
Heinrich-Plett-Str. 40, 34132 Kassel, Germany
seiler@mathematik.uni-kassel.de

Abstract

We provide a rigorous formulation of Vessiot's vector field approach to the analysis of general systems of partial differential equations and prove its equivalence to the formal theory.

Keywords: Vessiot distribution, integral element, involution

1 Introduction

Vessiot [13] proposed in the 1920s an approach to deal with general systems of partial differential equations which takes an intermediate position between the formal theory [8, 10] and the Cartan-Kähler theory of exterior differential systems [1, 4]: while still formulated in the language of differential equations (considered as submanifolds of a jet bundle), it represents essentially a dual, vector field based formulation of the Cartan-Kähler theory replacing exterior derivatives by Lie brackets.

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

The Vessiot theory has not attracted much attention. Presentations in a more modern language are contained in [2, 11]; applications have mainly appeared in the context of the Darboux method for solving hyperbolic equations, see e. g. [12]. While a number of textbooks provide a very rigorous analysis of the Cartan-Kähler theory, the above mentioned references (and also Vessiot's original work) are somewhat lacking in this respect. In particular, the question of what assumptions are needed has been ignored.

The purpose of the present article is to close this gap and simultaneously to relate the Vessiot theory with the key concepts of the formal theory like involution and formal integrability. We will show that the Vessiot construction succeeds, if and only if it is applied to an involutive system. This result is not surprising, given the well-known fact that the formal theory and the Cartan-Kähler theory are equivalent. However, to our knowledge an explicit proof has never been given. As a by-product, we will provide a new definition for integral elements based on the contact map making also the relations between the formal theory and the Cartan-Kähler theory more transparent.

2 Formal Theory

We cannot give here a detailed introduction into the formal theory. Our presentation and notations follow [10]; other general references are [5, 8]. For simplicity, we will mainly work in local coordinates, although the whole theory can be expressed in an intrinsic way.

Let $\pi : \mathcal{E} \to \mathcal{X}$ be a (smooth) fibred manifold. We call coordinates $\mathbf{x} = (x^1, \ldots, x^n)$ of $\mathcal{X}$ *independent variables* and fibre coordinates $\mathbf{u} = (u^1, \ldots, u^m)$ in $\mathcal{E}$ *dependent variables*. Sections[2] $\sigma : \mathcal{X} \to \mathcal{E}$ correspond locally to functions $\mathbf{u} = \mathbf{s}(\mathbf{x})$. Derivatives are written in the form $u^\alpha_\mu = \partial^{|\mu|} u^\alpha / \partial x_1^{\mu_1} \cdots \partial x_n^{\mu_n}$ where $\mu = [\mu_1, \ldots, \mu_n]$. Adding the derivatives u^α_μ up to order q (denoted by $\mathbf{u}^{(q)}$) defines a local coordinate system for the q-th order jet bundle $J_q\pi$ which may be considered as the space of truncated Taylor expansions.

[2] Although we will exclusively consider local sections, we will use throughout a "global" notation in order to avoid the introduction of many local neighbourhoods.

The jet bundle $J_q\pi$ admits a number of fibrations. For us particularly important are $\pi^q_{q-1} : J_q\pi \to J_{q-1}\pi$ and $\pi^q : J_q\pi \to \mathcal{X}$. To each section $\sigma : \mathcal{X} \to \mathcal{E}$ locally defined by $\sigma(\mathbf{x}) = \big(\mathbf{x}, \mathbf{s}(\mathbf{x})\big)$ we may associate its prolongation $j_q\sigma : \mathcal{X} \to J_q\pi$, a section of the fibration π^q given by $j_q\sigma(\mathbf{x}) = \big(\mathbf{x}, \mathbf{s}(\mathbf{x}), \partial_{\mathbf{x}}\mathbf{s}(\mathbf{x}), \partial_{\mathbf{xx}}\mathbf{s}(\mathbf{x}), \dots\big)$.

The geometry of $J_q\pi$ is to a large extent determined by its *contact structure*. It can be described in a number of ways. We will use three different approaches. The *contact codistribution* $\mathcal{C}^0_q \subseteq T^*(J_q\pi)$ consists of all one-forms such that their pull-back by a prolonged section vanishes. Locally, it is spanned by the contact forms[3]

$$\omega^\alpha_\mu = du^\alpha_\mu - u^\alpha_{\mu+1_i} dx^i \,, \qquad 0 \le |\mu| < q \,.$$

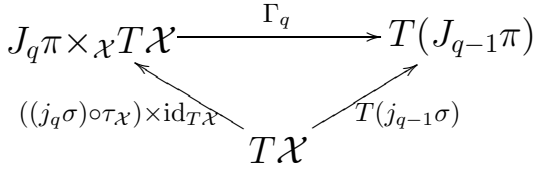
Dually, we may consider the *contact distribution* $\mathcal{C}_q \subseteq T(J_q\pi)$ consisting of all vector fields annihilated by $\mathcal{C}^0_q$. One easily verifies that it is generated by the fields

$$\begin{aligned} C^{(q)}_i &= \partial_i + u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu} \,, & 1 \le i \le n \,, \\ C^\mu_\alpha &= \partial_{u^\alpha_\mu} \,, & |\mu| = q \,. \end{aligned}$$

Note that the latter fields span the vertical bundle $V\pi^q_{q-1}$ of the fibration π^q_{q-1}. Thus the contact distribution can be split as $\mathcal{C}_q = V\pi^q_{q-1} \oplus \mathcal{H}$. Here the complement $\mathcal{H}$ is an n-dimensional transversal subbundle of $T(J_q\pi)$ and obviously not uniquely determined (though any local coordinate chart induces via the span of the vectors $C^{(q)}_i$ one possible choice). Note that any such complement $\mathcal{H}$ may be considered as the horizontal bundle of a connection on the fibred manifold $\pi^q : J_q\pi \to \mathcal{X}$ (*not* for the fibration π^q_{q-1}!). Following Fackerell [2], we call any connection on π^q whose horizontal bundle consists of contact fields a *Vessiot connection*[4].

As a third approach to the contact structure we consider, following [7], the *contact map*. It is the unique map $\Gamma_q : J_q\pi \times_{\mathcal{X}} T\mathcal{X} \to T(J_{q-1}\pi)$ such that the diagram

[3] Throughout the article we use the convention that a summation over repeated indices is understood.
[4] In the literature the name *Cartan connection* [6] is more common.



commutes for any section σ. Because of its linearity over π^q_{q-1}, we may also consider it as a map $\Gamma_q : J_q\pi \to T^*\mathcal{X}\otimes_{J_{q-1}\pi}T(J_{q-1}\pi)$ with the local coordinate form:

$$\Gamma_q : (\mathbf{x}, \mathbf{u}^{(q)}) \mapsto \big(\mathbf{x}, \mathbf{u}^{(q-1)}; dx^i \otimes (\partial_{x^i} + u^\alpha_{\mu+1_i}\partial_{u^\alpha_\mu})\big) \,. \tag{1}$$

Now one can see that $\operatorname{im}\Gamma_q = \mathcal{C}_{q-1}$ and hence $\mathcal{C}_q = (T\pi^q_{q-1})^{-1}(\operatorname{im}\Gamma_q)$.

Proposition 2.1. *A section $\gamma : \mathcal{X} \to J_q\pi$ is of the form $\gamma = j_q\sigma$ for a section $\sigma : \mathcal{X} \to \mathcal{E}$, if and only if $\operatorname{im}\Gamma_q\big(\gamma(x)\big) = T_{\gamma(x)}\pi^q_{q-1}\big(T_{\gamma(x)}\operatorname{im}\gamma\big)$ for all points $x \in \mathcal{X}$ where γ is defined.*

Thus for any section $\sigma : \mathcal{X} \to \mathcal{E}$ the equality $\operatorname{im}\Gamma_{q+1}\big(j_{q+1}\sigma(x)\big) = \operatorname{im}T_x(j_q\sigma)$ holds and we may say that knowing the $(q+1)$-jet $j_{q+1}\sigma(x)$ of a section σ at some $x \in \mathcal{X}$ is equivalent to knowing its q-jet $\rho = j_q\sigma(x)$ at x plus the tangent space $T_\rho(\operatorname{im} j_q\sigma)$ at this point. This observation will later be the key for the Vessiot theory.

A *differential equation* of order q is a fibred submanifold $\mathcal{R}_q \subseteq J_q\pi$ locally described as the zero set of some smooth functions on $J_q\pi$:

$$\mathcal{R}_q \,:\, \Big\{ \; \Phi^\tau(\mathbf{x}, \mathbf{u}^{(q)}) = 0 \,, \quad \tau = 1, \dots, t \,. \tag{2}$$

(Note that we do not distinguish between scalar equations and systems.) We denote by $\iota : \mathcal{R}_q \hookrightarrow J_q\pi$ the canonical inclusion map. Differentiating every equation in (2) yields the *prolonged equation* $\mathcal{R}_{q+1} \subseteq J_{q+1}\pi$ defined by the equations $\Phi^\tau = 0$ and $D_i\Phi^\tau = 0$ with the formal derivative D_i, given by $D_i\Phi^\tau(\mathbf{x}, \mathbf{u}) = \frac{\partial\Phi^\tau}{\partial x^i}(\mathbf{x}, \mathbf{u}) + \frac{\partial\Phi^\tau}{\partial u^\alpha_\mu}(\mathbf{x}, \mathbf{u})u^\alpha_{\mu+1_i}$. Iteration of this process gives the higher prolongations $\mathcal{R}_{q+r} \subseteq J_{q+r}\pi$. A subsequent projection leads to $\mathcal{R}^{(1)}_q = \pi^{q+1}_q(\mathcal{R}_{q+1}) \subseteq \mathcal{R}_q$ which is a proper submanifold, whenever integrability conditions appear. $\mathcal{R}_q$ is *formally integrable*, if at any prolongation order $r > 0$ the equality $\mathcal{R}^{(1)}_{q+r} = \mathcal{R}_{q+r}$ holds.

A *solution* is a section $\sigma : \mathcal{X} \to \mathcal{E}$ such that its prolongation satisfies $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$. In local coordinates, this coincides with the usual notion of a solution. For formally integrable equations it is straightforward to construct order by order formal power series solutions; otherwise it is very hard to find solutions. A key insight of Cartan was to introduce *infinitesimal solutions* or *integral elements* at a point $\rho \in \mathcal{R}_q$ as subspaces $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ which are potentially part of the tangent space of a prolonged solution.

Definition 2.2. *Let $\mathcal{R}_q \subseteq J_q\pi$ be a differential equation. A linear subspace $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ is an* integral element *at the point $\rho \in \mathcal{R}_q$, if a point $\hat\rho \in \mathcal{R}_{q+1}$ exists such that $\pi_q^{q+1}(\hat\rho) = \rho$ and $T\iota(\mathcal{U}_\rho) \subseteq \operatorname{im}\Gamma_{q+1}(\hat\rho)$.*

The above definition of an integral element is not the standard one. Usually, one considers the pull-back $\iota^*\mathcal{C}_q^0$ of the contact codistribution or more precisely the differential ideal $\mathcal{I}[\mathcal{R}_q] = \langle\iota^*\mathcal{C}_q^0\rangle_{\text{diff}}$ generated by it (recall that algebraically $\mathcal{I}[\mathcal{R}_q]$ is thus spanned by a basis of $\iota^*\mathcal{C}_q^0$ and the exterior derivatives of the forms in this basis) and an integral element is a subspace on which this ideal vanishes.

Proposition 2.3. *Let $\mathcal{R}_q$ be a differential equation such that $\mathcal{R}_q^{(1)} = \mathcal{R}_q$. A linear subspace $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ is an* integral element *at $\rho \in \mathcal{R}_q$, if and only if $T\iota(\mathcal{U}_\rho)$ lies transversal to the fibration π_{q-1}^q and every differential form $\omega \in \mathcal{I}[\mathcal{R}_q]$ vanishes on $\mathcal{U}_\rho$.*

Proof. Assume first that $\mathcal{U}_\rho$ is an integral element. Thus there exists a point $\hat\rho \in \mathcal{R}_{q+1}$ such that $\pi_q^{q+1}(\hat\rho) = \rho$ and $T\iota(\mathcal{U}_\rho) \subseteq \operatorname{im}\Gamma_{q+1}(\hat\rho)$. This implies firstly that $T\iota(\mathcal{U}_\rho)$ is transversal to π_{q-1}^q and secondly that every one-form $\omega \in \iota^*\mathcal{C}_q^0$ vanishes on $\mathcal{U}_\rho$, as $\operatorname{im}\Gamma_{q+1}(\hat\rho) \subset (\mathcal{C}_q)_\rho$. Thus there only remains to show that the same is true for the two-forms $d\omega \in \iota^*(d\mathcal{C}_q^0)$.

Choose a section $\gamma : \mathcal{R}_q \to \mathcal{R}_{q+1}$ such that $\gamma(\rho) = \hat\rho$ and define a distribution $\mathcal{D}$ of rank n on $\mathcal{R}_q$ by setting $T\iota(\mathcal{D}_{\tilde\rho}) = \operatorname{im}\Gamma_{q+1}\big(\gamma(\tilde\rho)\big)$ for any point $\tilde\rho \in \mathcal{R}_q$. Obviously, by construction $\mathcal{U}_\rho \subseteq \mathcal{D}_\rho$. It follows from the coordinate form (1) of the contact map that locally the distribution $\mathcal{D}$ is spanned by n vector fields X_i such that $\iota_*X_i = C_i^{(q)} + \gamma^\alpha_{\mu+1_i}C^\mu_\alpha$ where the

coefficients γ^α_ν are the highest-order components of the section γ. Thus the commutator of two such vector fields satisfies

$$\iota_*\big([X_i,X_j]\big) = \big(C^{(q)}_i(\gamma^\alpha_{\mu+1_j}) - C^{(q)}_j(\gamma^\alpha_{\mu+1_i})\big)C^\mu_\alpha + \gamma^\alpha_{\mu+1_j}[C^{(q)}_i, C^\mu_\alpha] - \gamma^\alpha_{\mu+1_i}[C^{(q)}_j, C^\mu_\alpha] \ .$$

The commutators on the right side vanish whenever $\mu_i = 0$ or $\mu_j = 0$, respectively. Otherwise we obtain $-\partial_{u^\alpha_{\mu-1_i}}$ and $-\partial_{u^\alpha_{\mu-1_j}}$, respectively. But this implies that the two sums on the right side cancel each other and we find that $\iota_*\big([X_i,X_j]\big) \in \mathcal{C}_q$. Thus we find for any contact form $\omega \in \mathcal{C}^0_q$ that

$$\begin{aligned} \iota^*(d\omega)(X_i,X_j) &= d\omega(\iota_*X_i, \iota_*X_j) \\ &= \iota_*X_i\big(\omega(\iota_*X_j)\big) - \iota_*X_j\big(\omega(\iota_*X_i)\big) + \omega\big(\iota_*([X_i,X_j])\big) \ . \end{aligned}$$

Each summand vanishes, as all appearing fields are contact fields. Hence any form $\omega \in \iota^*(d\mathcal{C}^0_q)$ vanishes on $\mathcal{D}$ and in particular on $\mathcal{U}_\rho \subseteq \mathcal{D}_\rho$.

For the converse, note that any subspace $\mathcal{U}_\rho \subseteq T_\rho\mathcal{R}_q$ satisfying the imposed conditions is spanned by linear combinations of vectors v_i such that $T\iota(v_i) = C^{(q)}_i{}_{|\rho} + \gamma^\alpha_{\mu,i}C^\mu_\alpha{}_{|\rho}$ where $\gamma^\alpha_{\mu,i}$ are real coefficients. Now consider a contact form ω^α_ν with $|\nu| = q-1$. Then $d\omega^\alpha_\nu = dx^i \wedge du^\alpha_{\nu+1_i}$. Evaluating the condition $\iota^*(d\omega^\alpha_\nu)_{|\rho}(v_i,v_j) = d\omega\big(T\iota(v_i), T\iota(v_j)\big) = 0$ yields the equation $\gamma^\alpha_{\nu+1_i,j} = \gamma^\alpha_{\nu+1_j,i}$. Hence the coefficients are of the form $\gamma^\alpha_{\mu,i} = \gamma^\alpha_{\mu+1_i}$ and a section σ exists such that $\rho = j_q\sigma(x)$ and $T_\rho(\operatorname{im} j_q\sigma)$ is spanned by the vectors $T\iota(v_1),\ldots,T\iota(v_n)$. This implies that $\mathcal{U}_\rho$ is an integral element. □

For many purposes the purely geometric notion of formal integrability is not sufficient and one needs the stronger algebraic concept of involution. This concerns e. g. the derivation of uniqueness results but also the numerical integration of overdetermined systems [9]. An intrinsic definition of involution requires the Spencer cohomology. We give here only a simplified coordinate version requiring that one works in "good", so-called δ-regular, coordinates $\mathbf{x}$ (this is not a strong restriction, as generic coordinates are δ-regular and there are possibilities to construct systematically "good" coordinates – see e. g. [3]).

The *(geometric) symbol* of a differential equation $\mathcal{R}_q$ is $\mathcal{N}_q = V\pi^q_{q-1}|_{\mathcal{R}_q} \cap T\mathcal{R}_q$, i. e. the vertical part of the tangent space to $\mathcal{R}_q$. Locally, $\mathcal{N}_q$ is the

solution space of the following linear system of (algebraic) equations in the unknowns v^α_μ (coordinates on $S_q(T^*\mathcal{X}) \otimes V\mathcal{E}$):

$$\mathcal{N}_q \; : \; \left\{ \sum_{\alpha,|\mu|=q} \left(\frac{\partial \Phi^\tau}{\partial u^\alpha_\mu} \right) v^\alpha_\mu = 0 \, . \right. \tag{3}$$

The prolonged symbols $\mathcal{N}_{q+r}$ are simply the symbols of the prolonged equations $\mathcal{R}_{q+r}$.

The *class* of a multi-index $\mu = [\mu_1, \ldots, \mu_n]$ is the smallest k for which μ_k is different from zero. The columns of the symbol matrix (3) are labelled by the v^α_μ. After ordering them by class, i. e. a column with a multi-index of higher class is always left of one with lower class, we compute a row echelon form. We denote the number of rows where the pivot is of class k by $\beta^{(k)}_q$ and associate with each such row the *multiplicative variables* $x^1, \ldots, x^k$. Prolonging each equation only with respect to its multiplicative variables yields independent equations of order $q+1$, as each has a different leading term. If prolongation with respect to the non-multiplicative variables does not lead to additional independent equations of order $q+1$, i. e. if

$$\operatorname{rank} \mathcal{N}_{q+1} = \sum_{k=1}^{n} k \beta^{(k)}_q \; , \tag{4}$$

then the symbol $\mathcal{N}_q$ is *involutive* (Cartan test).

The differential equation $\mathcal{R}_q$ is called *involutive*, if it is formally integrable and its symbol is involutive. Involutive equations possess a number of pleasant properties; for our purposes the most important one is the *Cartan-Kähler theorem* asserting the existence and uniqueness of analytic solutions for a formally well-posed initial value problem with an analytic involutive differential equation and analytic initial data.

For notational simplicity, we will consider in our subsequent analysis mainly first-order equations $\mathcal{R}_1 \subseteq J_1\pi$. Furthermore, we will assume that any present algebraic (i. e. zeroth-order) equation has been explicitly solved, reducing thus the number of dependent variables. From a theoretical point of view this does not represent a restriction, as any differential

equation $\mathcal{R}_q$ can be transformed into an equivalent first-order one and under some mild regularity assumptions the algebraic equations can always be solved locally.

For later use, we define a local normal form, the *Cartan normal form*, for such a differential equation. It arises by solving each equation for a derivative u_j^α, the *principal derivative*, and eliminating this derivative from all other equations. Furthermore, the principal derivatives are chosen in such a manner that their classes are as great as possible. All the remaining derivatives are called *parametric*. Ordering the obtained equations by their class, we can decompose them into subsystems:

$$u_n^\alpha = \phi_n^\alpha(\mathbf{x}, \mathbf{u}, u_j^\gamma) \qquad \begin{cases} 1 \leq \alpha \leq \beta_1^{(n)} \\ 1 \leq j \leq n \\ \beta_1^{(j)} < \gamma \leq m \end{cases} \tag{5a}$$

$$u_{n-1}^\alpha = \phi_{n-1}^\alpha(\mathbf{x}, \mathbf{u}, u_j^\gamma) \qquad \begin{cases} 1 \leq \alpha \leq \beta_1^{(n-1)} \\ 1 \leq j \leq n-1 \\ \beta_1^{(j)} < \gamma \leq m \end{cases} \tag{5b}$$

$$\vdots$$

$$u_1^\alpha = \phi_1^\alpha(\mathbf{x}, \mathbf{u}, u_j^\gamma) \qquad \begin{cases} 1 \leq \alpha \leq \beta_1^{(1)} \\ 1 = j \\ \beta_1^{(j)} < \gamma \leq m \end{cases} \tag{5c}$$

Note that the values $\beta_1^{(k)}$ are exactly those appearing in the Cartan test (4), as the symbol matrix of a differential equation in Cartan normal form is automatically triangular with the principal derivatives as pivots. The *Cartan characters* of $\mathcal{R}_1$ are defined as $\alpha_1^{(k)} = m - \beta_1^{(k)}$ and thus equal the number of parametric derivatives of class k.

For a differential equation $\mathcal{R}_1$ in Cartan normal form, it is possible to perform an involution analysis in closed form. We remark that an effective test of involution proceeds as follows. Each equation in (5) is prolonged with respect to each of its non-multiplicative variables. The arising second-order equations are simplified modulo the original system and the

prolongations with respect to the multiplicative variables. The symbol $\mathcal{N}_1$ is involutive, if and only if after the simplification none of the equations is of second-order any more. The equation $\mathcal{R}_1$ is involutive, if and only if all new equations simplify to zero, as any remaining first-order equation would be an integrability condition.

In order to apply this test, we set $\mathcal{B} := \left\{(\alpha, i) \in \mathbb{N}^m \times \mathbb{N}^n \colon u_i^\alpha \text{ is a principal derivative } \right\}$ and for each $(\alpha, i) \in \mathcal{B}$ we define $\Phi_i^\alpha := u_i^\alpha - \phi_i^\alpha$. Now a straightforward calculation yields

$$D_j \Phi_i^\alpha = u_{ij}^\alpha - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i} \sum_{\gamma=\beta_1^{(h)}+1}^{m} u_{hj}^\gamma C_\gamma^h(\phi_i^\alpha) \; . \tag{6}$$

For $j > i$, the prolongation $D_j \Phi_i^\alpha$ is non-multiplicative, otherwise it is multiplicative.

Now let $j > i$, so that (6) is a non-multiplicative prolongation. According to our test, the symbol $\mathcal{N}_1$ is involutive, if and only if it is possible to eliminate on the right hand side of (6) all second-order derivatives by adding multiplicative prolongations. As a first step we note the following result arising after a tedious but fairly straightforward computation.

Lemma 2.4. *Expanding the formal derivatives and collecting all second-order derivatives in the right hand side of (6) yields that the difference* $D_j \Phi_i^\alpha - D_i \Phi_j^\alpha - \sum_{h=1}^{i} \sum_{\gamma=\beta_1^{(h)}+1}^{m} C_\gamma^h(\phi_i^\alpha) D_h \Phi_j^\gamma$ *equals*

$$\begin{aligned}
&C_i^{(1)}(\phi_j^\alpha) - C_j^{(1)}(\phi_i^\alpha) - \sum_{h=1}^{i} \sum_{\gamma=\beta_1^{(h)}+1}^{m} C_\gamma^h(\phi_i^\alpha) C_h^{(1)}(\phi_j^\gamma) \\
&- \sum_{h=1}^{i-1} \sum_{\delta=\beta_1^{(h)}+1}^{m} u_{hh}^\delta \left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C_\gamma^h(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \right] \\
&- \sum_{1 \le h < k < i} \left\{ \sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(k)}} u_{hk}^\delta \left[\sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C_\gamma^k(\phi_i^\alpha) C_\delta^h(\phi_j^\gamma) \right] \right.
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{\delta=\beta_1^{(k)}+1}^{m} u^{\delta}_{hk} \left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^k_\delta(\phi^\gamma_j) + \sum_{\gamma=\beta_1^{(k)}+1}^{\beta_1^{(j)}} C^k_\gamma(\phi^\alpha_i) C^h_\delta(\phi^\gamma_j) \right] \Bigg\} \\
&- \sum_{h=1}^{i-1} \Bigg\{ \sum_{\delta=\beta_1^{(h)}+1}^{\beta_1^{(i)}} u^{\delta}_{hi} \left[-C^h_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i) C^h_\delta(\phi^\gamma_j) \right] \\
&+ \sum_{\delta=\beta_1^{(i)}+1}^{m} u^{\delta}_{hi} \left[-C^h_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i) C^h_\delta(\phi^\gamma_j) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^i_\delta(\phi^\gamma_j) \right] \Bigg\} \\
&- \sum_{\substack{h=1 \\ i+1 \le k < j}}^{i-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u^{\delta}_{hk} \left[\sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^k_\delta(\phi^\gamma_j) \right] \\
&- \sum_{k=i}^{j-1} \sum_{\delta=\beta_1^{(k)}+1}^{m} u^{\delta}_{ik} \left[-C^k_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i) C^k_\delta(\phi^\gamma_j) \right] \\
&- \sum_{h=1}^{i-1} \sum_{\delta=\beta_1^{(j)}+1}^{m} u^{\delta}_{hj} \left[C^h_\delta(\phi^\alpha_i) + \sum_{\gamma=\beta_1^{(h)}+1}^{\beta_1^{(j)}} C^h_\gamma(\phi^\alpha_i) C^j_\delta(\phi^\gamma_j) \right] \\
&- \sum_{\delta=\beta_1^{(j)}+1}^{m} u^{\delta}_{ij} \left[C^i_\delta(\phi^\alpha_i) - C^j_\delta(\phi^\alpha_j) + \sum_{\gamma=\beta_1^{(i)}+1}^{\beta_1^{(j)}} C^i_\gamma(\phi^\alpha_i) C^j_\delta(\phi^\gamma_j) \right] .
\end{aligned}
$$

In the first line we collected all lower-order terms. Furthermore, none of the appearing second-order derivatives is of a form that it could be eliminated by adding some multiplicative prolongation. Hence the symbol $\mathcal{N}_1$ is involutive, if and only if all the expressions in square brackets vanish. The differential equation $\mathcal{R}_1$ is involutive, if and only if in addition the first line vanishes, as it represents an integrability condition. Thus Lemma 2.4 provides us with an explicit form of all obstructions to involution for $\mathcal{R}_1$.

3 The Vessiot distribution

By Proposition 2.1, the tangent spaces $T_\rho(\operatorname{im} j_q\sigma)$ of prolonged sections at points $\rho \in J_q\pi$ are always subspaces of the contact distribution $\mathcal{C}_q|_\rho$. If the section σ is a solution of $\mathcal{R}_q$, it furthermore satisfies by definition $\operatorname{im} j_q\sigma \subseteq \mathcal{R}_q$ and hence $T(\operatorname{im} j_q\sigma) \subseteq T\mathcal{R}_q$. These considerations motivate the following construction.

Definition 3.1. *The* Vessiot distribution *of a differential equation $\mathcal{R}_q \subseteq J_q\pi$ is the distribution $\mathcal{V}[\mathcal{R}_q] \subseteq T\mathcal{R}_q$ defined by*

$$T\iota\big(\mathcal{V}[\mathcal{R}_q]\big) = T\iota\big(T\mathcal{R}_q\big) \cap \mathcal{C}_q|_{\mathcal{R}_q} .$$

Again, this is not the usual definition found in the literature. But the equivalence to the standard approach is an elementary exercise in computing with pull-backs:

Proposition 3.2. *The Vessiot distribution satisfies $\mathcal{V}[\mathcal{R}_q] = (\iota^*\mathcal{C}_q^0)^0$.*

The Vessiot distribution is not necessarily of constant rank along $\mathcal{R}_q$; for simplicity, we will assume its rank does not vary over the differential equation. Note that the symbol $\mathcal{N}_q$ as the vertical part of $T\mathcal{R}_q$ is always contained in $\mathcal{V}[\mathcal{R}_q]$. In general, $\mathcal{V}[\mathcal{R}_q]$ is not involutive (an exception are formally integrable equations of finite type), but it may contain involutive subdistributions; among these, those of dimension n which are transversal (to the fibration $\mathcal{R}_q \to \mathcal{X}$) are of special interest for us.

Lemma 3.3. *If the section $\sigma : \mathcal{X} \to \mathcal{E}$ is a solution of the equation $\mathcal{R}_q$, then the tangent bundle $T(\operatorname{im} j_q\sigma)$ is an n-dimensional transversal involutive subdistribution of $\mathcal{V}[\mathcal{R}_q]|_{\operatorname{im} j_q\sigma}$. Conversely, if $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ is an n-dimensional transversal involutive subdistribution, then any integral manifold of $\mathcal{U}$ has locally the form $\operatorname{im} j_q\sigma$ for a solution σ of $\mathcal{R}_q$.*

This simple observation forms the basis of Vessiot's approach to the analysis of $\mathcal{R}_q$: he proposed to search for all n-dimensional, transversal involutive subdistributions of $\mathcal{V}[\mathcal{R}_q]$. Before we do this, we first show how integral elements appear in this program.

Proposition 3.4. *Let $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ be a transversal subdistribution of the Vessiot distribution of constant rank k. Then the spaces $\mathcal{U}_\rho$ are k-dimensional integral elements for all points $\rho \in \mathcal{R}_q$ if, and only if, $[\mathcal{U},\mathcal{U}] \subseteq \mathcal{V}[\mathcal{R}_q]$.*

Proof. Let $\{\omega_1, \dots, \omega_r\}$ be a basis of the codistribution $\iota^*\mathcal{C}_q^0$. Then an algebraic basis of the ideal $\mathcal{I}[\mathcal{R}_q]$ is $\{\omega_1, \dots, \omega_r, d\omega_1, \dots, d\omega_r\}$. Any vector field $X \in \mathcal{U}$ trivially satisfies $\omega_i(X) = 0$ by Proposition 3.2. For arbitrary fields $X_1, X_2 \in \mathcal{U}$, we have $d\omega_i(X_1, X_2) = X_1\big(\omega_i(X_2)\big) - X_2\big(\omega_i(X_1)\big) + \omega_i\big([X_1, X_2]\big)$. The first two summands on the right hand side vanish trivially and the remaining equation implies our claim. □

We call a subdistribution $\mathcal{U} \subseteq \mathcal{V}[\mathcal{R}_q]$ satisfying the conditions of Proposition 3.4 an *integral distribution*[5] for the differential equation $\mathcal{R}_q$. Note that generally an integral distribution is *not* integrable; the name only reflects the fact that it is composed of integral elements.

Since the symbol $\mathcal{N}_q$ of the equation $\mathcal{R}_q$ is contained in the Vessiot distribution, we can split the Vessiot distribution into $\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$ where $\mathcal{H}$ is some complement. By analogy to the above discussed decomposition of the full contact distribution, this leads naturally to connections: provided $\dim \mathcal{H} = n$, it may be considered as the horizontal bundle of a connection of the fibred manifold $\mathcal{R}_q \to \mathcal{X}$ and we call any such connection a *Vessiot connection* for $\mathcal{R}_q$. The existence of n-dimensional complements is connected to the absence of integrability conditions.

Proposition 3.5. *If the differential equation $\mathcal{R}_q$ satisfies $\mathcal{R}_q^{(1)} = \mathcal{R}_q$, then its Vessiot distribution possesses locally a decomposition $\mathcal{V}[\mathcal{R}_q] = \mathcal{N}_q \oplus \mathcal{H}$ with an n-dimensional complement $\mathcal{H}$.*

Proof. The assumption $\mathcal{R}_q = \mathcal{R}_q^{(1)}$ implies that to every point $\rho \in \mathcal{R}_q$ at least one point $\hat{\rho} \in \mathcal{R}_{q+1}$ with $\pi_q^{q+1}(\hat{\rho}) = \rho$ exists. We choose such a $\hat{\rho}$ and consider $\operatorname{im} \Gamma_{q+1}(\hat{\rho}) \subset T_\rho(J_q\pi)$. By definition of the contact map Γ_{q+1}, this is an n-dimensional transversal subset of $\mathcal{C}_{q\,|\rho}$. Thus there only remains to show that it is also tangential to $\mathcal{R}_q$, as then we can define a complement

[5] In the literature the terminology "involution" is common for such distributions which, however, is quite confusing in our opinion.

by $T\iota(\mathcal{H}_\rho) = \operatorname{im}\Gamma_{q+1}(\hat{\rho})$. But this tangency is a trivial consequence of $\hat{\rho} \in \mathcal{R}_{q+1}$; using for example the local coordinates expression (1) for Γ_q and a local representation $\Phi^\tau = 0$ of $\mathcal{R}_q$, one immediately sees that the vector $v_i = \Gamma_{q+1}(\hat{\rho}, \partial_{x^i}) \in T_\rho(J_q\pi)$ satisfies $d\Phi^\tau{}_{|\rho}(v_i) = D_i\Phi^\tau(\hat{\rho}) = 0$ and thus is tangential to $\mathcal{R}_q$.

Hence we have proven that it is possible to construct for each point $\rho \in \mathcal{R}_q$ a complement $\mathcal{H}_\rho$ such that $\mathcal{V}_\rho[\mathcal{R}_q] = (\mathcal{N}_q)_\rho \oplus \mathcal{H}_\rho$. Now we must show that these complements can be chosen in such a way that they form a distribution (which by definition is smooth). Our assumption $\mathcal{R}_q = \mathcal{R}_q^{(1)}$ implies that the restricted projection $\hat{\pi}_q^{q+1} : \mathcal{R}_{q+1} \to \mathcal{R}_q$ is a surjective submersion, i. e. it defines a fibred manifold. Thus if we choose a local section $\gamma : \mathcal{R}_q \to \mathcal{R}_{q+1}$ and then always take $\hat{\rho} = \gamma(\rho)$, it follows immediately that the corresponding complements $\mathcal{H}_\rho$ define a smooth distribution as required. □

Any n-dimensional complement $\mathcal{H}$ is obviously a transversal subdistribution of $\mathcal{V}[\mathcal{R}_q]$, but not necessarily involutive. Conversely, any n-dimensional subdistribution $\mathcal{H}$ of $\mathcal{V}[\mathcal{R}_q]$ is a possible choice as a complement. Hence we may reformulate Vessiot's goal as the construction of all *flat* Vessiot connections. Choosing a "reference" complement $\mathcal{H}_0$ with a basis $(X_1, \ldots, X_n)$, a basis for any other complement $\mathcal{H}$ arises by adding some vertical fields to the vectors X_i. We will follow this approach in the next section. For the remainder of this section we turn our attention to the choice of a convenient basis of $\mathcal{V}[\mathcal{R}_q]$ that will facilitate our computations.

Since the symbol $\mathcal{N}_q$ is an involutive distribution, there exists a basis $(Y_1, Y_2, \ldots Y_r)$ for it with $r = \dim \mathcal{N}_q$ whose Lie brackets vanish: $[Y_k, Y_\ell] = 0$ for all $1 \leq k, \ell \leq r$. Since the vertical bundle $V\pi_{q-1}^q$ is also involutive, we can decompose $V\pi_{q-1}^q = \mathcal{N}_q \oplus \mathcal{W}$ where $\mathcal{W}$ is again an involutive distribution. $\mathcal{W}$ can be spanned by vector fields $W_1, \ldots, W_s$ where $s = \sum_{k=1}^n \beta_q^{(k)}$ equals the number of principal derivatives which are chosen such that we have $[W_a, W_b] = 0$ for all $1 \leq a, b \leq s$. In local coordinates, a particularly convenient choice for the fields Y_k and W_a exists. We choose for any $1 \leq k \leq r$ a *parametric* derivative u_μ^α, that is $(\alpha, \mu) \notin \mathcal{B}$, and

$Y_k := Y^\alpha_\mu := \iota_*(\partial_{u^\alpha_\mu})$, and for any $1 \le a \le s$ a *principal* derivative u^α_μ such that $(\alpha, \mu) \in \mathcal{B}$ and $W_a := W^\alpha_\mu := \partial_{u^\alpha_\mu}$.

The reference complement $\mathcal{H}_0$ is chosen as follows. Any basis of it must consist of n transversal contact fields. Since the fields C^μ_α are vertical, we can always use a basis $(\tilde{X}_1, \dots, \tilde{X}_n)$ of the form $\tilde{X}_i = C^{(q)}_1 + \xi^\alpha_{i\mu} C^\mu_\alpha$ with some coefficient functions $\xi^\alpha_{i\mu}$ chosen such that $\tilde{X}_i$ is tangential to $\mathcal{R}_q$. The fields C^μ_α also span the vertical bundle $V\pi^q_{q-1}$ and hence we may exploit the above decomposition for a further simplification of the basis. By subtracting from each $\tilde{X}_i$ a suitable linear combination of the fields Y_k spanning the symbol $\mathcal{N}_q$, we arrive at a basis $(X_1, \dots, X_n)$ where $X_i = C^{(q)}_i + \xi^\alpha_i W_a$.

As already mentioned above, generally, the Vessiot distribution $\mathcal{V}[\mathcal{R}_q]$ is not involutive. Hence it is not surprising that its structure equations are going to be important later. Since the only non-vanishing Lie brackets of contact fields in $\mathcal{C}_q$ are

$$[C^{\nu+1_i}_\alpha, C^{(q)}_i] = \partial_{u^\alpha_\nu} \,, \qquad |\nu| = q - 1 \,, \tag{7}$$

we may extend the above chosen basis (X_i, Y_k) of $\mathcal{V}[\mathcal{R}_q]$ to a basis of the derived Vessiot distribution $\mathcal{V}'[\mathcal{R}_q]$ by adding vector fields Z_c, $1 \le c \le t = \dim \mathcal{V}'[\mathcal{R}_q] - \dim \mathcal{V}[\mathcal{R}_q]$, where, using (7), for each c we have some coefficients $\kappa^\alpha_{c\nu}$ such that $Z_c = \kappa^\alpha_{c\nu} \partial_{u^\alpha_\nu}$ with $|\nu| = q - 1$. By construction, the non-vanishing structure equations of $\mathcal{V}[\mathcal{R}_q]$ take now the form

$$[X_i, X_j] = \Theta^c_{ij} Z_c \qquad \text{and} \qquad [X_i, Y_k] = \Xi^c_{ik} Z_c \tag{8}$$

with smooth functions Θ^c_{ij} and Ξ^c_{ik}.

Remark 3.6. *In the next section we will have to analyse certain matrices which are built out of the coefficients Θ^c_{ij}, Ξ^c_{ik}. It will turn out that this analysis becomes simpler, if we pretend that all fields $\partial_{u^\alpha_\nu}$ with $|\nu| = q - 1$ are contained in $\mathcal{V}'[\mathcal{R}_q]$. Restricting to the first-order case $q = 1$, we write the right hand sides of (8) as $\Theta^\alpha_{ij} \partial_{u^\alpha}$ and $\Xi^\alpha_{ik} \partial_{u^\alpha}$, respectively. The new and the old coefficients are then related by*

$$\Theta^\alpha_{ij} = \Theta^c_{ij} \kappa^\alpha_c \,, \qquad \Xi^\alpha_{ik} = \Xi^c_{ik} \kappa^\alpha_c \,. \tag{9}$$

Knowing the (larger) sets of coefficients Θ^α_{ij}, Ξ^α_{ik}, we can reconstruct the true structure coefficients Θ^c_{ij}, Ξ^c_{ik} by solving the overdetermined linear system of equations (9). This is always possible since the fields Z_c are assumed to be linearly independent. Thus there exist some coefficents κ^c_α such that

$$\Theta^c_{ij} = \Theta^\alpha_{ij}\kappa^c_\alpha \,, \qquad \Xi^c_{ik} = \Xi^\alpha_{ik}\kappa^c_\alpha \,.$$

For a first-order equation $\mathcal{R}_1$ with Cartan normal form (5) satisfying the assumptions of Proposition 3.5 it is possible to perform this process explicitly. We choose as a reference complement $\mathcal{H}_0$ the linear span of the vector fields

$$X_i = C^{(q)}_i + \sum_{(\alpha,\mu)\in\mathcal{B}} C^{(q)}_i(\phi^\alpha_\mu) C^\mu_\alpha \,.$$

One easily verifies in a rather straightforward computation that this is a valid choice. Using this reference complement, we can explicitly evaluate the Lie brackets (8) on $\mathcal{R}_1$. As we are not able to determine a simple expression for the derived Vessiot distribution, we follow the approach taken in Remark 3.6 and obtain for the extended set of structure coefficients Θ^α_{ij}, Ξ^α_{ik} the following results: if $i < j$, then

$$\Theta^\alpha_{ij} = \begin{cases} 0 & : (\alpha,i) \notin \mathcal{B} \text{ and } (\alpha,j) \notin \mathcal{B} \,, \\ C^{(1)}_i(\phi^\alpha_j) & : (\alpha,i) \notin \mathcal{B} \text{ and } (\alpha,j) \in \mathcal{B} \,, \\ C^{(1)}_i(\phi^\alpha_j) - C^{(1)}_j(\phi^\alpha_i) & : (\alpha,i) \in \mathcal{B} \text{ and } (\alpha,j) \in \mathcal{B} \,, \end{cases} \tag{10}$$

and if we assume that $Y_k = Y^\beta_j$, then

$$\Xi^\alpha_{ik} = \begin{cases} 0 & : (\alpha,i) \notin \mathcal{B} \text{ and } (\alpha,i) \neq (\beta,j) \,, \\ -1 & : (\alpha,i) \notin \mathcal{B} \text{ and } (\alpha,i) = (\beta,j) \,, \\ -C^j_\beta(\phi^\alpha_i) & : (\alpha,i) \in \mathcal{B} \,. \end{cases} \tag{11}$$

We collect these coefficients into vectors Θ_{ij} and matrices Ξ_i where the m rows are ordered according to increasing α, and the $r = \dim \mathcal{N}_1$ columns are ordered, according to increasing j, into n blocks (empty for those j with $m = \beta^{(j)}_1$) and within each block according to increasing β (with $\beta^{(j)}_1 <$

$\beta \leq m$). Note that for a differential equation with constant coefficients all Θ_{ij} vanish and for a maximally overdetermined equation there are no Ξ_i.

The matrices Ξ_i have a special form. For any class i, the matrix Ξ_i has $\alpha_1^{(i)}$ rows where all entries are zero with only one exception: for each $1 \leq \ell_i < \alpha_1^{(i)}$ we have $\Xi_{i,k}^{\beta_1^{(i)}+\ell_i} = -\delta_{\ell\, k}$ where $\ell := \sum_{h=1}^{i-1} \alpha_1^{(h)} + \ell_i$. The entries in the remaining $\beta_1^{(i)}$ rows are $-C_\beta^j(\phi_i^\alpha)$. Some of these vanish, too: all of the parametric derivatives on the right side of an equation in the Cartan normal form (5) are of a class lower than that of the equation's left side as otherwise we would solve this equation for the derivative of highest class. This means $-C_\beta^j(\phi_i^\alpha) = 0$ whenever $j = \mathrm{class}(u_j^\beta) > \mathrm{class}(u_i^\alpha) = i$, and it follows that for each i, $1 \leq i \leq n$ the matrix Ξ_i looks like

$$\Xi_i = \begin{pmatrix} -C^1_{\beta^1}(\phi_i^\alpha) & \cdots & -C^{i-1}_{\beta^{i-1}}(\phi_i^\alpha) & -C^i_{\beta^i}(\phi_i^\alpha) & 0 \cdots 0 \\ 0 & \cdots & 0 & -\mathbb{1}_{\alpha_1^{(i)}} & 0 \cdots 0 \end{pmatrix} . \tag{12}$$

Here, for $1 \leq j \leq i$, we have $\beta_1^{(j)} + 1 \leq \beta^j \leq m$. The unit block of $\alpha_1^{(i)}$ rows leads immediately to the estimate

$$\alpha_1^{(i)} \leq \operatorname{rank} \Xi_i \leq \min\{m, \sum_{j=1}^{i} \alpha_1^{(j)}\} .$$

Since the matrices Ξ_i are made up of block matrices and we are going to calculate with these blocks, we introduce the following notation: let for any i, $1 \leq i \leq n$, ${}_a^b[\,\Xi_i\,]_c^d$ denote the block in Ξ_i consisting of the entries from the ath row to the bth row and from the cth column to the dth column.

4 Flat Vessiot connections

Recall that our goal is the construction of all n-dimensional transversal involutive subdistributions $\mathcal{U}$ within the Vessiot distribution $\mathcal{V}[\mathcal{R}_1]$. Taking (X_i, Y_k) as a basis for $\mathcal{V}[\mathcal{R}_1]$, we make for the basis $(U_1, \ldots, U_n)$ of such a distribution $\mathcal{U}$ the ansatz $U_i = X_i + \zeta_i^k Y_k$ with yet undetermined coefficient functions ζ_i^k. This ansatz follows naturally from our considerations above,

as the fields X_i span a reference complement to $\mathcal{N}_1$ and all fields Y_k are vertical. Since the fields U_i are in triangular form, the distribution $\mathcal{U}$ is involutive, if and only if their Lie brackets vanish, and using (8) this means:

$$\begin{aligned}[U_i, U_j] &= [X_i, X_j] + \zeta_i^k [Y_k, X_j] + \zeta_j^k [X_i, Y_k] + (U_i(\zeta_j^k) - U_j(\zeta_i^k)) Y_k \\ &= (\Theta_{ij}^c - \Xi_{jk}^c \zeta_i^k + \Xi_{ik}^c \zeta_j^k) Z_c + (U_i(\zeta_j^k) - U_j(\zeta_i^k)) Y_k = 0 \, .\end{aligned} \tag{13}$$

By definition of the Y_k and Z_a, these fields are linearly independent, so their coefficients must vanish for $\mathcal{U}$ to be involutive. Thus (13) yields two sets of conditions for the coefficient functions ζ_i^k: a system of algebraic equations

$$\Theta_{ij}^c - \Xi_{jk}^c \zeta_i^k + \Xi_{ik}^c \zeta_j^k = 0 \, , \qquad \left\{ \begin{array}{l} 1 \leq c \leq t \, , \\ 1 \leq i < j \leq n \end{array} \right. \tag{14}$$

and a system of differential equations

$$U_i(\zeta_j^k) - U_j(\zeta_i^k) = 0 \, , \qquad \left\{ \begin{array}{l} 1 \leq k \leq r \, , \\ 1 \leq i < j \leq n \, . \end{array} \right. \tag{15}$$

In (14) the true structure coefficients Θ_{ij}^c, Ξ_{jk}^c appear. For our subsequent analysis we follow Remark 3.6 and replace them by the extended set of coefficients Θ_{ij}^α, Ξ_{jk}^α. This corresponds to replacing (14) by an equivalent but larger linear system of equations which is simpler to analyse.

Remark 4.1. *The vector fields Y_k lie in the Vessiot distribution $\mathcal{V}[\mathcal{R}_1]$. Thus, according to Proposition 3.4, $\mathcal{U}$ is an integral distribution, if and only if the coefficients ζ_k^i satisfy the algebraic conditions (14). This observation permits us immediately to reduce the number of unknowns in our ansatz. Assume that we have values $1 \leq i, j \leq n$ and $1 \leq \alpha \leq m$ such that both (α, i) and (α, j) are not contained in $\mathcal{B}$, i. e. u_i^α and u_j^α are both parametric derivatives (and thus obviously the second-order derivative u_{ij}^α, too). Then there exist two symbol fields $Y_k = \iota_*(\partial_{u_i^\alpha})$ and $Y_l = \iota_*(\partial_{u_j^\alpha})$. Now it follows from the coordinate form (1) of the contact map that $\mathcal{U}$ can be an integral distribution, if and only if $\zeta_j^k = \zeta_i^l$.*

As the unknowns ζ_k^j may be understood as labels for the columns of the matrices Ξ_h, this identification leads to a contraction of these matrices. We introduce now contracted matrices $\hat{\Xi}_h$ which arise as follows: whenever $\zeta_j^k = \zeta_i^l$ then the corresponding columns of Ξ_h are added. Similarly, we introduce reduced vectors $\hat{\zeta}_h$ where the redundant components are left out. From now on we always understand that in the equations above this reduction has been performed.

Now the question arises, when the combined system (14,15) has solutions. We begin by analysing the algebraic part (14). As a system for the vectors $\hat{\zeta}_i$, we seek to build a solution step by step with i increasing. Thus we begin the construction of the integral distribution $\mathcal{U}$ by first choosing an arbitrary vector field U_1 and then aiming for another vector field U_2 such that $[U_1, U_2] \in \mathcal{V}[\mathcal{R}_q]$. During the construction of U_2 we regard the components of the vector $\hat{\zeta}_1$ as given parameters and the components of $\hat{\zeta}_2$ as the only unknowns of the system

$$\hat{\Xi}_1 \hat{\zeta}_2 = \hat{\Xi}_2 \hat{\zeta}_1 - \Theta_{12} \; . \tag{16}$$

Since the components of $\hat{\zeta}_1$ are not considered as unknowns, the system (16) must not lead to any restrictions for the coefficients $\hat{\zeta}_1^k$. Obviously, this is the case, if and only if

$$\operatorname{rank} \hat{\Xi}_1 = \operatorname{rank} (\hat{\Xi}_1 \;\; \hat{\Xi}_2) \; . \tag{17}$$

Assuming that (17) holds, the system (16) is solvable, if and only if it satisfies the augmented rank condition

$$\operatorname{rank} \hat{\Xi}_1 = \operatorname{rank} (\hat{\Xi}_1 \;\; \hat{\Xi}_2 \;\; -\Theta_{12}) \; . \tag{18}$$

Now we proceed by iteration. Given $i-1$ vector fields $U_1, U_2, \ldots U_{i-1}$ of the required form spanning an involutive subdistribution of $\mathcal{V}[\mathcal{R}_1]$, we construct the next vector field U_i by solving the system

$$\begin{aligned} \hat{\Xi}_1 \hat{\zeta}_i &= \hat{\Xi}_i \hat{\zeta}_1 - \Theta_{1i} \\ &\vdots \\ \hat{\Xi}_{i-1} \hat{\zeta}_i &= \hat{\Xi}_i \hat{\zeta}_{i-1} - \Theta_{i-1,i} \; . \end{aligned} \tag{19}$$

Again we consider only the components of the vector $\hat{\zeta}_i$ as unknowns and (19) must not imply any further restrictions on the components of the vectors $\hat{\zeta}_j$ for $1 \leq j < i$. The corresponding rank condition is

$$\operatorname{rank} \begin{pmatrix} \hat{\Xi}_1 \\ \hat{\Xi}_2 \\ \vdots \\ \hat{\Xi}_{i-1} \end{pmatrix} = \operatorname{rank} \begin{pmatrix} \hat{\Xi}_1 & \hat{\Xi}_i & & \\ \hat{\Xi}_2 & & \hat{\Xi}_i & 0 \\ \vdots & 0 & \ddots & \\ \hat{\Xi}_{i-1} & & & \hat{\Xi}_i \end{pmatrix} . \tag{20}$$

Assuming that it holds, (19) is solvable for the components of $\hat{\zeta}_i$, if and only if it satisfies the augmented rank condition

$$\operatorname{rank} \begin{pmatrix} \hat{\Xi}_1 \\ \hat{\Xi}_2 \\ \vdots \\ \hat{\Xi}_{i-1} \end{pmatrix} = \operatorname{rank} \begin{pmatrix} \hat{\Xi}_1 & \hat{\Xi}_i & & & -\Theta_{1i} \\ \hat{\Xi}_2 & & \hat{\Xi}_i & 0 & -\Theta_{2i} \\ \vdots & 0 & \ddots & & \vdots \\ \hat{\Xi}_{i-1} & & & \hat{\Xi}_i & -\Theta_{i-1,i} \end{pmatrix} . \tag{21}$$

The following theorem relates the satisfaction of these rank conditions and thus the solvability of the algebraic system (14) by the above described step by step process to intrinsic properties of the differential equation $\mathcal{R}_1$ and its symbol $\mathcal{N}_1$.

Theorem 4.2. *Assume that δ-regular coordinates have been chosen for the differential equation $\mathcal{R}_1$. The rank condition (20) is satisfied for all $1 \leq i \leq n$, if and only if the symbol $\mathcal{N}_1$ is involutive. The augmented rank condition (21) holds for all $1 \leq i \leq n$, if and only if the differential equation $\mathcal{R}_1$ is involutive.*

Proof. In order to prove (20), we transform the matrices into row echelon form. Since each matrix $\hat{\Xi}_i$ contains a unit block, there is an obvious way to do this. We describe the procedure using the above introduced notation for subblocks. As we shall see, the relevant entries in this row echelon form are the coefficients of the second-order derivatives u^δ_{hk} in Lemma 2.4 and therefore their vanishing is equivalent to involution of the symbol $\mathcal{N}_1$.

We start with $i = 1$, i. e. with (17). Since $\hat{\Xi}_1$ is a negative unity matrix of $\alpha_1^{(1)}$ rows with a $\beta_1^{(1)} \times \alpha_1^{(1)}$-matrix stacked upon it and only zeros for all other entries, we have $\operatorname{rank}(\hat{\Xi}_1) = \alpha_1$. Next, we transform the matrix $(\hat{\Xi}_1 \ \hat{\Xi}_2)$ into row echelon form using the special structure of the matrices $\hat{\Xi}_i$ as given in (12); the blocks are replaced in this way:

$$
\begin{aligned}
{}_1^{\beta_1^{(1)}}[\hat{\Xi}_1]_1^{\alpha_1^{(1)}} &\leftarrow {}_1^{\beta_1^{(1)}}[\hat{\Xi}_1]_1^{\alpha_1^{(1)}} + {}_1^{\beta_1^{(1)}}[\hat{\Xi}_1]_1^{\alpha_1^{(1)}} \cdot {}_{\beta_1^{(1)}+1}^{m}[\hat{\Xi}_1]_1^{\alpha_1^{(1)}} \ , \\
{}_1^{\beta_1^{(1)}}[\hat{\Xi}_2]_1^{\alpha_1^{(1)}} &\leftarrow {}_1^{\beta_1^{(1)}}[\hat{\Xi}_2]_1^{\alpha_1^{(1)}} + {}_1^{\beta_1^{(1)}}[\hat{\Xi}_1]_1^{\alpha_1^{(1)}} \cdot {}_{\beta_1^{(1)}+1}^{m}[\hat{\Xi}_2]_1^{\alpha_1^{(1)}} \ , \\
{}_1^{\beta_1^{(1)}}[\hat{\Xi}_2]_{\alpha_1^{(1)}+1}^{\alpha_1^{(1)}+\alpha_1^{(2)}} &\leftarrow {}_1^{\beta_1^{(1)}}[\hat{\Xi}_2]_{\alpha_1^{(1)}+1}^{\alpha_1^{(1)}+\alpha_1^{(2)}} + {}_1^{\beta_1^{(1)}}[\hat{\Xi}_1]_1^{\alpha_1^{(1)}} \cdot {}_{\beta_1^{(1)}+1}^{m}[\hat{\Xi}_2]_{\alpha_1^{(1)}+1}^{\alpha_1^{(1)}+\alpha_1^{(2)}} \ .
\end{aligned}
$$

If, for the sake of simplicity, we use the same names for the changed blocks, then we have

$$
{}_1^{\beta_1^{(1)}}[\hat{\Xi}_2]_1^{\alpha_1^{(1)}} = \left(-C_\delta^1(\phi_2^\alpha) + \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C_\gamma^1(\phi_1^\alpha) C_\delta^1(\phi_2^\gamma) \right)_{\substack{1 \le \alpha \le \beta_1^{(1)} \\ \beta_1^{(1)}+1 \le \delta \le m}} ,
$$

$$
{}_1^{\beta_1^{(1)}}[\hat{\Xi}_2]_{\alpha_1^{(1)}+1}^{\alpha_1^{(1)}+\alpha_1^{(2)}} = \left(C_\delta^1(\phi_1^\alpha) - C_\delta^2(\phi_2^\alpha) + \sum_{\gamma=\beta_1^{(1)}+1}^{\beta_1^{(2)}} C_\gamma^1(\phi_1^\alpha) C_\delta^2(\phi_2^\gamma) \right)_{\substack{1 \le \alpha \le \beta_1^{(1)} \\ \beta_1^{(2)}+1 \le \delta \le m}} .
$$

A comparison with the obstructions to involution obtained by applying Lemma 2.4 for $i = 1$ and $j = 2$ shows that all these entries vanish, if and only if the obstructions vanish. It follows that the first $\beta_1^{(1)}$ rows of the matrix $(\hat{\Xi}_1 \ \hat{\Xi}_2)$ are zero. The last $\alpha_1^{(1)}$ rows begin with the block $-\mathbb{1}_{\alpha_1^{(1)}}$ and hence $\operatorname{rank}(\hat{\Xi}_1 \ \hat{\Xi}_2) = \alpha_1^{(1)} = \operatorname{rank} \hat{\Xi}_1$. Thus we may conclude that the rank condition (17) holds, if and only if no non-multiplicative prolongation $D_2 \Phi_1^a$ leads to an obstruction of involution.

The claim for the augmented condition (18) follows from the explicit expression (10) for the entries Θ_{ij}^α. Performing the same computations as above described with the augmented system yields as additional relevant

entries exactly the integrability conditions arising from Lemma 2.4 applied for $i = 1$ and $j = 2$. Hence (18) holds, if and only if no non-multiplicative prolongation $D_2\Phi_1^a$ yields an integrability condition.

As one might expect from the above considerations for $i = 1$, the analysis of (20) for a general $1 \leq i \leq n$ will require the non-multiplicative prolongations $D_i\Phi_1^\alpha, D_i\Phi_2^\alpha, \ldots D_i\Phi_{i-1}^\alpha$. It follows trivially from the block form (12) of the matrices Ξ_i that the rank of the matrix on the left hand side of (20) is $\sum_{k=1}^{i-1} \alpha_1^{(k)}$.

For lack of space we skip the details for the general case. We follow the same steps as in the case $i = 1$. The transformation of the matrix on the right hand side of (20) can be described using block matrices, and the resulting matrix in row echelon form has as its entries in the rows where no unit block appears the coefficients of the second-order derivatives in Lemma 2.4. Thus we may conclude again that satisfaction of (20) is equivalent to the fact that in the non-multiplicative prolongations $D_i\Phi_1^\alpha, \ldots, D_i\Phi_{i-1}^\alpha$ no obstructions to involution arise. In the case of the augmented conditions (21), it follows again from the explicit expression (10) for the entries Θ_{ij}^α that the additional relevant entries are identical with the potential integrability conditions produced by the non-multiplicative prolongations $D_i\Phi_1^\alpha, \ldots, D_i\Phi_{i-1}^\alpha$.

At this point it becomes apparent why we had to introduce the contracted matrices $\hat{\Xi}_i$. As all functions are assumed to be smooth, partial derivatives commute: $u_{ij}^\alpha = u_{ji}^\alpha$. In Lemma 2.4 each obstruction to involution actually consists of two parts: one arises as coefficient of u_{ij}^α, the other one as coefficient of u_{ji}^α. While this decomposition does not show in Lemma 2.4 because both derivatives are collected into one term, the two parts appear in different columns of the matrices Ξ_i and in general the rank condition (20) will not hold, if we replace the contracted matrices $\hat{\Xi}_i$ by the original matrices Ξ_i (see the example below). The effect of the contraction is to combine the two parts in order to obtain the right rank. □

There remains to analyse the solvability, if we add the differential system (15). We first note that one can show in a straightforward computation

that (15) alone is again an involutive system. If the original equation $\mathcal{R}_1$ is analytic, then the quasi-linear system (15) is analytic, too. Thus we may apply the Cartan-Kähler theorem to it which guarantees the existence of solutions.

The problem is that the combined system (14,15) is in general not involutive, as the prolongation of the algebraic equations (14) leads to additional differential equations. Instead of analysing the effect of these integrability conditions, we proceed as follows. If we assume that $\mathcal{R}_1$ is involutive, then we know from Theorem 4.2 that the algebraic equations (14) are solvable. In the proof of the theorem we even produced an explicit row echelon form of the system matrix which we can now exploit to eliminate some of the unknowns $\hat{\zeta}_i^k$ as a linear combinations of the remaining ones.

Theorem 4.3. *Assume that δ-regular coordinates have been chosen for the differential equation $\mathcal{R}_1$ and that $\mathcal{R}_1$ is analytic. Then the combined system (14,15) is solvable.*

Proof. Following the strategy outlined above, we eliminate some of the unknowns $\hat{\zeta}_i^k$. Because of the simple structure of (15), it turns out that we must take a closer look only at those equations where the leading derivative is of one of the unknowns we eliminate. A somewhat lengthy but straightforward computation shows that these equations actually vanish. The remaining equations still form an involutive system. Thus we eventually arrive at an analytic involutive differential equation for the coefficient functions $\hat{\zeta}_i^k$ which is solvable according to the Cartan-Kähler theorem. □

Example 4.4. *Consider the first-order equation*

$$\mathcal{R}_1 : \left\{ \begin{array}{ll} u_t = v_t = w_t = u_s = 0 \,, & v_s = 2u_x + 4u_y \,, \\ w_s = -u_x - 3u_y \,, & u_z = v_x + 2w_x + 3v_y + 4w_y \,. \end{array} \right.$$

It is formally integrable, and its symbol is involutive with $\dim \mathcal{N}_1 = 8$. *Thus $\mathcal{R}_1$ is an involutive equation. For the matrices Ξ_i, all of which are*

3×8-*matrices, we find*

$$\Xi_1 = \begin{pmatrix} -1 & 0 & 0 & 0\,0\,0\,0\,0 \\ 0 & -1 & 0 & 0\,0\,0\,0\,0 \\ 0 & 0 & -1 & 0\,0\,0\,0\,0 \end{pmatrix}, \qquad \Xi_2 = \begin{pmatrix} 0\,0\,0 & -1 & 0 & 0 & 0\,0 \\ 0\,0\,0 & 0 & -1 & 0 & 0\,0 \\ 0\,0\,0 & 0 & 0 & -1 & 0\,0 \end{pmatrix},$$

$$\Xi_3 = \begin{pmatrix} 0 & -1 & -2 & 0 & -3 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \Xi_4 = \begin{pmatrix} 0 & 0\,0 & 0 & 0\,0\,0\,0 \\ -2 & 0\,0 & -4 & 0\,0\,0\,0 \\ 1 & 0\,0 & 3 & 0\,0\,0\,0 \end{pmatrix}, \quad \Xi_5 = 0_{3\times 8}\,.$$

For the first two steps in the construction of the fields U_i, the rank conditions are trivially satisfied even for the non-contracted matrices. But not so in the third step where we have in the row echelon form of the arising 9×32-matrix in the 7th row zero entries throughout except in the 12th column (where we have -2) and in the 17th column (where we have 2). As a consequence, we obtain the equality $\zeta_1^4 = \zeta_2^1$ and the rank condition for this step does not hold. However, since both u_x and u_y are parametric derivatives and in our ordering $Y_1 = \iota_*(\partial_{u_x})$ and $Y_4 = \iota_*(\partial_{u_y})$, this equality is already taken into account in our reduced ansatz and for the matrices $\hat{\Xi}_i$ the rank condition is satisfied.

Note that the rank condition is first violated when the rank reaches the symbol dimension (which is 8). From then on, the rank of the left matrix in (20) stagnates at $\dim \mathcal{R}_1$ while the rank of the augmented matrix may rise further. The entries breaking the rank condition differ by their sign, while their corresponding coefficients in Lemma 2.4 are collected into one sum and thus vanish.

References

[1] R.L. Bryant, S.S. Chern, R.B. Gardner, H.L. Goldschmidt, and P.A. Griffiths. *Exterior Differential Systems.* Mathematical Sciences Research Institute Publications 18. Springer-Verlag, New York, 1991.

[2] E.D. Fackerell. Isovectors and prolongation structures by Vessiot's vector field formulation of partial differential equations. In R. Martini, editor, *Geometric Aspects of the Einstein Equations and Integrable Systems*, Lecture Notes in Physics 239, pages 303–321. Springer-Verlag, Berlin, 1985.

[3] M. Hausdorf, M. Sahbi, and W.M. Seiler. δ- and quasi-regularity for polynomial ideals. These proceedings.

[4] T.A. Ivey and J.M. Landsberg. *Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems*. Graduate Studies in Mathematics 61. American Mathematical Society, Providence, 2003.

[5] I.S. Krasilshchik, V.V. Lychagin, and A.M. Vinogradov. *Geometry of Jet Spaces and Nonlinear Partial Differential Equations*. Gordon & Breach, New York, 1986.

[6] V.V. Lychagin. Homogeneous geometric structures and homogeneous differential equations. In V.V. Lychagin, editor, *The Interplay between Differential Geometry and Differential Equations*, Amer. Math. Soc. Transl. 167, pages 143–164. Amer. Math. Soc., Providence, 1995.

[7] M. Modugno. Covariant quantum mechanics. Unpublished manuscript, Dept. of Mathematics, University of Florence, 1999.

[8] J.F. Pommaret. *Systems of Partial Differential Equations and Lie Pseudogroups*. Gordon & Breach, London, 1978.

[9] W.M. Seiler. Completion to involution and semi-discretisations. *Appl. Num. Math.*, 42:437–451, 2002.

[10] W.M. Seiler. *Involution — The Formal Theory of Differential Equations and its Applications in Computer Algebra and Numerical Analysis*. Algorithms and Computation in Mathematics. Springer-Verlag, Berlin/Heidelberg, manuscript accepted for publication.

[11] O. Stormark. *Lie's Structural Approach to PDE Systems*. Encyclopedia of Mathematics and its Applications 80. Cambridge University Press, Cambridge, 2000.

[12] P.J. Vassiliou. Vessiot structure for manifolds of (p, q)-hyperbolic type: Darboux integrability and symmetry. *Trans. Amer. Math. Soc.*, 353:1705–1739, 2001.

[13] E. Vessiot. Sur une théorie nouvelle des problèmes généraux d'intégration. *Bull. Soc. Math. Fr.*, 52:336–395, 1924.

Towards an algorithmisation of the Dirac constraint formalism

Gerdt V [1] Khvedelidze A [1, 2, 3] Palii Yu [1]

[1] Laboratory of Information Technologies,
Joint Institute for Nuclear Research,
Dubna, 141980, Russia,[gerdt, palii]@jinr.ru
[2] Department of Theoretical Physics,
A.M.Razmadze Mathematical Institute,
Tbilisi, GE-0193, Georgia
[3] School of Mathematics and Statistics, University of Plymouth,
Plymouth, PL4 8AA, United Kingdom,
akhvedelidze@plymouth.ac.uk

Abstract

Cental issues of the Dirac constraint formalism are discussed in relation to the algorithmic methods of commutative algebra based on the Gröbner basis techniques. For a wide class of finite dimensional polynomial degenerate Lagrangian systems, we describe an algorithmic scheme of computation of the complete set of constraints, their separation into subsets of first and second class constraints as well as the construction of a generator of local symmetry transformations. The proposed scheme is exemplified by considering the so-called light-cone Yang-Mills mechanics with an $SU(2)$ gauge structure group.

Keywords: constrained Hamiltonian dynamics, commutative algebra, Gröbner basis.

1 Introduction

Lagrangians used for the description of fundamental particles, such as electrons and photons, as well as quarks and gluons have a *degenerate* Hessian functions. This rather unusual property, compared to standard Lagrangian mechanical models, profoundly modifies the whole mathematical description of classical evolution. It demands the physical interpretation of constrained variables (e.g. longitudinal components of the electromagnetic potential) and also requires the generalisation of the canonical quantisation scheme. From the mathematical point of view the new element of the Hamiltonian description of a degenerate Lagrangian system is the *involution analysis* of the differential equations of motion. Its pivotal ingredients in the *generalized Hamiltonian dynamics* [1]-[5] are realised in the form of the Dirac scheme to determine constraints. This is related to [6, 7, 8] the *formal theory* of differential equations [9]. The process of the determining all the *integrability conditions* that can not be derived using only the algebraic operations with the existing differential equations is just the "reproduction" of constraints in the Dirac formalism. Having a complete set of constraints we are able to identify the set of "truly" dynamical equations for this involutive system and therefore finally provide a deterministic classical evolution of the physical observables and perform the subsequent quantization.

Effective completion to involution of systems of differential equations needed in field theories represents a very complicated challenge requiring sophisticated computer-algebraic methods [10]. Similarly the generalized Hamiltonian formalism also needs an efficient algorithmisation and implementation in a proper computer algebra software.

In the present paper we apply the most universal algorithmic tool of commutative algebra, the *Gröbner bases* [11], as the main algorithmic ingredient of the generalized Hamiltonian dynamics for degenerate mechanical models with polynomial Lagrangians. In [12] it was already suggested to use the Gröbner bases for the computation and separation of constrains for such models. The underlying Dirac-Gröbner algorithm is based on the

facility of the Gröbner bases method to manipulate with a polynomial in the phase variables modulo constraint manifold, and, in particular, to check whether the polynomial vanishes on the manifold. In the present note we propose some further algorithmic improvements and extensions aiming at the computational realization of the Hamiltonian reduction of degenerate mechanical system possessing local symmetries.

It should be noticed that constructive ideas of the involution analysis of differential equations combined with those from the Gröbner bases technique have culminated in the concept of involutive bases [13] as a special type of Gröbner bases providing the efficient involutive algorithms [14] for construction of the involutive as well as the reduced Gröbner bases.

The plan of this paper is as follows. We start (Section 2) with a brief description of the main issues in the Dirac constraint formalism that should be put into an algorithmic form suitable for effective calculations. In Section 3 the ways to achieve this goal for finite-dimensional mechanical systems with polynomial Lagrangians are described. Then (Section 4) we consider the so-called light-cone $SU(n)$ Yang-Mills mechanics as an interesting example of constrained model for which the first algorithmic issue of the Dirac formalism, namely, construction of the primary constraints, can be performed for arbitrary n. The remaining algorithmic issues of the Dirac formalism are illustrated in Section 4 for this model specified [15] to the simplest nontrivial structure group $SU(2)$. Finally, in Section 5 some conclusions are presented.

2 The issues requiring algorithmisation

Here we sketch briefly the basic notions and definitions from the Dirac constraint formalism for a finite dimensional degenerate Lagrangian system and make a list of the main procedures requiring an algorithmic reformulation.

Consider an n-dimensional mechanical system whose configuration space is $\mathbf{R}^n$ and the Lagrangian $L(q, \dot{q})$ is defined on a tangent space as a function of the coordinates $q := q_1, q_2, \ldots, q_n$ and velocities $\dot{q} := \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$.

The Lagrangian system is called a *regular* one if the rank $r := rank\|H_{ij}\|$ of the corresponding Hessian function $H_{ij} := \partial^2 L/\partial\dot{q}_i\partial\dot{q}_j$ is maximal ($r = n$). In this case the Euler-Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial\dot{q}_i}\Big) - \frac{\partial L}{\partial q_i} = 0\,, \qquad 1 \le i \le n \tag{1}$$

rewritten explicitly as

$$H_{ij}\ddot{q}_j + \frac{\partial^2 L}{\partial q_j \partial\dot{q}_i}\,\dot{q}_j - \frac{\partial L}{\partial q_i} = 0$$

can be resolved with respect to the accelerations ($\ddot{q}$) and there is no *hidden constraints*. Otherwise, if $r < n$, the Euler-Lagrange equations are *degenerate* or *singular*. In this case not all differential equations (1) are of second order, namely there are $n - r$ independent equations, Lagrangian constraints, containing only coordinates and velocities. Passing to the Hamiltonian description via a Legendre transformation

$$p_i := \frac{\partial L}{\partial\dot{q}_i} \tag{2}$$

the degeneracy of the Hessian manifests itself in the existence of $n - r$ relations between coordinates and momenta, the *primary constraints*

$$\phi^{(1)}_\alpha(p, q) = 0\,, \qquad 1 \le \alpha \le n - r\,. \tag{3}$$

Equations (3) define the so-called *primary constraints* subset Σ_1. This definition is implicit and therefore the first algorithmisation topic is:

Issue I: *Find all primary constraints describing the subset* Σ_1 .

From (3) the dynamics is constrained by the set Σ_1 and by the Dirac prescription is governed by the *total* Hamiltonian

$$H_T := H_C + U_\alpha\phi^{(1)}_\alpha\,, \tag{4}$$

which differs from the *canonical* Hamiltonian $H_C(p, q) = p_i q_i - L$ by a linear combination of the primary constraints with the Lagrange *multipliers* U_α.

The next step is to analyze the dynamical requirement that classical trajectories remain in Σ_1 during evolution,

$$\dot{\phi}_\alpha^{(1)} = \{H_T, \phi_\alpha^{(1)}\} \stackrel{\Sigma_1}{=} 0\,. \tag{5}$$

In (5) the evolutional changes are generated by the canonical *Poisson brackets* with the total Hamiltonian (4) and the abbreviation $\stackrel{\Sigma_1}{=}$ stands for *a week equality*, i.e., the right-hand side of (5) vanishes modulo the constraints.

The consistency condition (5), unless it is satisfied identically, may lead either to a contradiction or to a determination of the Lagrange multipliers U_α or to new constraints. The former case indicates that the given Hamiltonian system is inconsistent.

In the latter case when (5) is not satisfied identically and is independent of the multipliers U_α the left-hand side of (5) defines the new constraints. Otherwise, if the left-hand side depends on some Lagrange multipliers U_α the consistency condition determines these multipliers, and, therefore, the constraints set is not enlarged by new constraints. The subsequent iteration of this consistency check ends up with the complete set of constraints and/or determination of some/or all Lagrange multipliers.

The number of Lagrange multipliers U_α which can be found is determined by the so-called *Poisson bracket matrix*

$$\mathbf{M}_{\alpha\beta} :\stackrel{\Sigma}{=} \{\phi_\alpha, \phi_\beta\}\,, \tag{6}$$

where Σ denotes the subset of a phase space defined by the all constraints including primary $\phi_\alpha^{(1)}$, *secondary* $\phi_\alpha^{(2)}$, *ternary* $\phi_\alpha^{(3)}$, etc., constraints, $\Phi := (\phi_\alpha^{(1)}\,, \phi_\alpha^{(2)}\,, \ldots\,, \phi_\alpha^{(k)}\,)$

$$\Sigma \;:\quad \phi_\alpha(p,q) = 0\,, \qquad 1 \le \alpha \le k\,. \tag{7}$$

The co-rank $s := k - rank(\mathbf{M})$ of matrix $\mathbf{M}$ represent the number of *first-class constraints* $\psi_1\,, \psi_2\,, \ldots\,, \psi_s$, linear combinations of constraints ϕ_α

$$\psi_\alpha(p,q) = \sum_\beta \mathrm{c}_{\alpha\beta}(p,q)\,\phi_\beta\,, \tag{8}$$

whose Poisson brackets are weakly zero

$$\{\psi_\alpha(p,q), \psi_\beta(p,q)\} \stackrel{\Sigma}{=} 0 \qquad 1 \leq \alpha\,, \beta \leq s\,. \tag{9}$$

The remaining functionally independent constraints form the subset of *second-class constraints*.

This method of constraints determination in the Dirac formalism represents the particular form of *completion* of the initial Hamiltonian equations to *involution*; the generated constraints are nothing else than the *integrability conditions* [6, 7, 8].

Therefore the second algorithmisation challenge can be formulated as

> **Issue II:** *Determine all integrability conditions and perform their separation into first and second class conditions.*

First-class constraints play a very special role in the Hamiltonian description: they provide the basis for *generators of local symmetry transformations*. The knowledge of a local symmetry transformation is important because according to physical requirement the physical observables are singlets under the gauge symmetry transformations.

So the next important algorithmisation problem is

> **Issue III:** *Construct the generator of local symmetry transformation and find the basis for singlet observables.*

The last problem has direct impact on the process of *Hamiltonian reduction*, that is a formulation of a new Hamiltonian system with a reduced number of degrees of freedom but equivalent to the initial degenerate one [2, 16, 17]. The presence of s first-class constraints and $r := k - s$ second-class constraints guarantees the possibility of local reformulation of the initial $2n$ dimensional Hamiltonian system as a $2n-2s-r$ dimensional reduced Hamiltonian system (cf. [7]).

Therefore, the final fourth algorithmisation challenge we formulate here as

> **Issue IV:** *Construct an equivalent unconstrained Hamiltonian system on the reduced phase space.*

3 How the algorithm works

Here we extend the main ideas of [12] and describe the algorithmic basics that can be used to solve the problems stated in the previous section. In doing so, we restrict our consideration to an arbitrary dynamical system with finitely many degrees of freedom whose Lagrangian is a polynomial in coordinates and velocities with rational (possibly parametric) coefficients[1] $L(q, \dot{q}) \in \mathbf{Q}[q, \dot{q}]$.

3.1 Primary constraints

For degenerate systems the primary constraints (3) are consequences of the polynomial relations (2). These relations generate the polynomial ideal in $\mathbf{Q}[q, \dot{q}, p]$

$$I_{p,q,\dot{q}} \equiv \mathrm{Id}(\cup_{i=1}^{n}\{p_i - \partial L/\partial \dot{q}_i\}) \subset \mathbf{Q}[p, q, \dot{q}]\,. \tag{10}$$

Thereby, primary constraints (3) belong to the radical $\sqrt{I_{p,q}}$ of the elimination ideal

$$I_{p,q} = I_{p,q,\dot{q}} \cap \mathbf{Q}[p, q]\,.$$

Correspondingly, for an appropriate term ordering which eliminates $\dot{q}$, a Gröbner basis of $I_{p,q}$ (denotation: $GB(I_{p,q})$) is given by [11, 18]

$$GB(I_{p,q}) = GB(I_{p,q,\dot{q}}) \cap \mathbf{Q}[p, q]\,.$$

This means that construction of the Gröbner basis for the ideal (10) with omitting elements in the basis depending on velocities and then constructing of $GB(\sqrt{I_{p,q}})$ allows us to compute the set of primary constraints. If $GB(\sqrt{I_{p,q}}) = \emptyset$ then the dynamical system is regular. Otherwise, the algebraically independent set Φ_1 of primary constraints is the subset $\Phi_1 \subset GB(\sqrt{I_{p,q}})$ such that

$$\forall \phi(p, q) \in \Phi_1 \;:\; \phi(p, q) \notin \mathrm{Id}(\Phi_1 \setminus \{\phi(p, q)\})\,. \tag{11}$$

[1] Throughout this section we use some standard notions and definitions of commutative algebra (see, for example, [18]).

Verification of (11) is algorithmically done by computing the following *normal form*: $NF(\phi, GB(\mathrm{Id}(\Phi_1 \setminus \{\phi\}))$.

Therefore, all the computational steps described above admit full algorithmisation by means of Gröbner bases. In addition, the canonical Hamiltonian $H_c(p,q)$ is computed as $NF(p_i q_i - L, GB(I_{p,q,\dot{q}}))$.

3.2 Complete set of constraints and their separation

The dynamical consequences (5) of a primary constraint can also be algorithmically analyzed by computing the normal form of the Poisson brackets of the primary constraint and the total Hamiltonian modulo $GB(\sqrt{I_{p,q}})$. Here the Lagrange multipliers U_α in (4) are treated as time-dependent functions. If the non-vanishing normal form does not contain U_α, then it is nothing else than the secondary constraint. In this case the set of primary constraints is enlarged by the secondary constraint obtained and the process is iterated. At the end either the complete set Φ of constraints (7) is constructed or inconsistency of the dynamical system is detected. The detection holds when the intermediate Gröbner basis, whose computation is a part of the iterative procedure, becomes $\{1\}$.

To separate the set $\Phi = \{\phi_1, \ldots, \phi_k\}$ into of first and second class constraints the Poisson bracket $k \times k$ matrix $\mathbf{M}$ (6) is built. Its entries are computed as normal forms of the Poisson brackets of the constraints modulo a Gröbner basis of the ideal generated by set Φ.

To construct $s := k - r$; where $r = rank(\mathbf{M})$ first-class constraints as linear combinations (8) of constraints (7) satisfying (9) it suffices to find the basis $P = \{p_1, \ldots, p_{k-r}\}$ of the null space (kernel) of the linear transformation defined by $\mathbf{M}$. Every vector $p \in P$ generates the first-class constraint of form $p_\alpha \phi_\alpha$.

Now consider the $s \times k$ matrix $(p_j)_\alpha$ composed of components of vectors in P and find a basis $T := \{t_1, \ldots, t_r\}$ of the null space of the corresponding linear transformation. For every vector $t \in T$ the second-class constraint is constructed as $t_\alpha \phi_\alpha$.

Thus the constraints separation can be done using linear algebra operations with the matrix $\mathbf{M}$ alone. Together with the Gröbner bases technique this implies full algorithmisation for computing the complete set of algebraically independent constraints and their separation (**Issues I** and **II** of Section 2).

3.3 Generator of local symmetry transformations

The local symmetries are generated by first-class constraints (cf. [5]) but the presence of the second-class constraints makes realization of the symmetry transformations very subtle. To overcome some of these difficulties one can effectively eliminate the second class constraints by changing the initial Poisson bracket to the new *Dirac bracket* defined as

$$\{f,g\}_D := \{f,g\} - \{f,\chi_\alpha\}\mathbf{C}^{-1}_{\alpha\beta}\{\chi_\beta,g\}\,,$$

where χ_α $(1 \le \alpha \le r)$ denotes the second-class constraints, and the invertible $r \times r$ matrix $\mathbf{C}_{\alpha\beta}$ is defined as

$$\mathbf{C}_{\alpha\beta} :\overset{\Sigma}{=} \{\chi_\alpha,\chi_\beta\}\,.$$

Since for an arbitrary function f it follows that $\{f,\chi_\alpha\}_D = 0$ the second-class constraints can be set to zero either before or after evaluating a Dirac bracket. This last evaluation, modulo the constraint functions, can be performed algorithmically exploiting the Gröbner bases. After elimination of all second-class constraints follow to the Dirac conjecture [1] the generator G of local transformations is expressed as a linear combination of all first-class constraints

$$G = \sum_{\beta=1}^{k_1} \varepsilon^{(1)}_\beta \phi^{(1)}_\beta \;+\; \sum_{\gamma=k_1+1}^{s} \varepsilon^{(2)}_\gamma \phi^{(2)}_\gamma\,, \tag{12}$$

and its action on phase space coordinates (q,p) is given now with the aid of the Dirac bracket

$$\delta q_i = \{G,q_i\}_D, \qquad \delta p_i = \{G,p_i\}_D\,.$$

In (12) the coefficients $\varepsilon_\beta^{(1)}$ and $\varepsilon_\gamma^{(2)}$ are functions of time t and the first sum includes k_1 primary first-class constraints while the second sum contains the all remaining first-class constraints. Not all of the functions ε in (12) are independent ones. Here we briefly state how following the method suggested in [19] one can extract the irreducible set of functions from the set of ε. The total time derivative of the gauge-symmetry generator (12) is given in terms of the Dirac bracket of G and the canonical Hamiltonian:

$$\frac{dG}{dt} = \frac{\partial G}{\partial t} + \{G, H_C\}_D \,. \tag{13}$$

Since the set of first-class constraints is complete, the Dirac bracket in the right-hand side of (13) is

$$\{\phi_\mu, H_C\}_D = \rho_{\mu\nu}\phi_\nu \,. \tag{14}$$

The unctions $\rho_{\mu\nu}$ can be algorithmically computed by using the Gröbner bases method. To perform this computation one can use, for example, the extended Gröbner basis algorithm [20]. Given a set of polynomials $F = \{f_1, \ldots, f_m\} \subset \mathbf{Q}[p, q]$ generating the polynomial ideal $\mathrm{Id}(F)$, this algorithm yields the explicit representation

$$g_i = h_{ij}\, f_j \tag{15}$$

of elements in the Gröbner basis $\{g_1 \ldots, g_n\}$ of this ideal in terms of the ideal generated by polynomials in F. Therefore, having computed a Gröbner basis for the ideal generated by the first-class constraints and the corresponding polynomial coefficients for the elements in the Gröbner basis as given in (15), the coefficients $\rho_{\mu\nu}$ are easily computed by reduction [11, 18, 20] of the Dirac bracket in (14) modulo the Gröbner basis expressed in terms of the first-class constraints ϕ_ν. Note that one can similarly compute the algebra of first-class constraints

$$\{\phi_\alpha, \phi_\beta\}_D = \varrho_{\alpha\beta\gamma}\phi_\gamma \,,$$

if the structure functions $\varrho_{\alpha\beta\gamma}$ are polynomials in p, q.

The generator of local transformation is conserved modulo the primary constraints

$$\frac{dG}{dt} \overset{\Sigma_1}{=} 0 \;\; \Rightarrow \;\; \dot{\varepsilon}^{(2)}_{\gamma}\phi^{(2)}_{\gamma} + \varepsilon^{(1)}_{\beta}\rho_{\beta\gamma}\phi^{(2)}_{\gamma} + \varepsilon^{(2)}_{\delta}\rho_{\delta\gamma}\phi^{(2)}_{\gamma} \overset{\Sigma_1}{=} 0\,. \tag{16}$$

Since, by their construction,the constraints $\phi^{(2)}_{\gamma}$ do not vanish on the primary-constraint manifold Σ_1, the relations (16) represent the following system of differential equations on the gauge functions $\varepsilon^{(1)}_{\beta}$ and $\varepsilon^{(2)}_{\gamma}$

$$\dot{\varepsilon}^{(2)}_{\gamma} + \varepsilon^{(1)}_{\beta}\rho_{\beta\gamma} + \varepsilon^{(2)}_{\delta}\rho_{\delta\gamma} = 0\,, \quad (k_1 + 1 \leq \gamma \leq s)\,, \tag{17}$$

where the index β runs from 1 to k_1, γ runs from $k_1 + 1$ to s and the functions $\rho_{\mu\nu}$ are projected on to the subset Σ_1.

Since the differential system (17) is underdetermined, one can express the functions $\varepsilon^{(1)}_{\beta}$ in terms of arbitrary functions $\varepsilon^{(2)}_{\gamma}(t)$ and their derivatives [19]. Since this last procedure is algorithmic, this completes the algorithmic construction of the generator of the local symmetry transformation.

The above described algorithmic procedures have been implemented as a Maple package (currently for Maple 10), and this package was used to perform the computations presented in the next section.

It is worth noting here that the remaining part of **Issue III** as well as **Issue IV** still require an algorithmisation.

4 Light-cone Yang-Mills mechanics

Now we discuss the application of the general scheme described above to a mechanical model originated from Yang-Mills gauge theory formulated on the light-cone under the assumption of spatial homogeneity of the gauge fields.

The standard action of Yang-Mills field theory with structure group $SU(n)$ in four-dimensional Minkowski space M_4, endowed with a metric η is

$$S := \frac{1}{g_0^2}\int_{M_4} \mathrm{tr}\, F \wedge *F\,, \tag{18}$$

where g_0 is a coupling constant and the $SU(n)$ algebra valued curvature two-form

$$F := dA + A \wedge A$$

is constructed from the connection one-form A. The connection and curvature, as Lie algebra valued quantities, are expressed in terms of the antihermitian algebra basis T^a,

$$A = A^a\, T^a\,, \qquad F = F^a\, T^a\ . \qquad a = 1, 2, \ldots, n^2 - 1\,.$$

The metric $\eta_{\gamma\delta}$ enters the action through the dual field strength tensor defined in accordance with the Hodge star operation
$*F_{\mu\nu} := \frac{1}{2}\sqrt{\det(\eta)}\,\epsilon_{\mu\nu\alpha\beta}\,F^{\alpha\beta}$, with totally antisymmetric tensor $\epsilon_{\mu\nu\alpha\beta}$.

The light-cone version of the theory is formulated using the frame where two vectors $e_\pm := \frac{1}{\sqrt{2}}\,(e_0 \pm e_3)$ tangent to the light-cone are combined with the orthogonal pair $e_k\,, k = 1, 2$. The corresponding coordinates are usually called (see, e.g. [21]) light-cone coordinates $x^\mu = \left(x^+, x^-, x^\perp\right)$

$$x^\pm := \frac{1}{\sqrt{2}}\,\left(x^0 \pm x^3\right)\,, \qquad x^\perp := x^k\,, \quad k = 1, 2\,.$$

The non-zero components of the metric are $\eta_{+-} = \eta_{-+} = -\eta_{11} = -\eta_{22} = 1$. The connection one-form in the light-cone basis is given as

$$A := A_+\, dx^+ + A_-\, dx^- + A_k\, dx^k\,. \tag{19}$$

By definition, the Lagrangian of light-cone Yang-Mills mechanics follows from the corresponding Lagrangian of Yang-Mills theory if one supposes that the components of the connection one-form A in (19) only depend on the light-cone "time variable" x^+

$$A_\pm = A_\pm(x^+)\,, \qquad A_k = A_k(x^+)\,.$$

Substitution of this *ansatz* into the classical action (18) defines the Lagrangian of light-cone Yang-Mills mechanics

$$L := \frac{1}{2g^2}\,\left(F^a_{+-}\,F^a_{+-} + 2\,F^a_{+k}\,F^a_{-k} - F^a_{12}\,F^a_{12}\right)\,, \tag{20}$$

where g is the "renormalized" coupling constant $g^2 = g_0^2/(\text{Volume})$ and the light-cone components of the field-strength tensor are given by

$$\begin{aligned} F^a_{+-} &:= \frac{\partial A^a_-}{\partial x^+} + \mathrm{f}^{abc}\, A^b_+\, A^c_- \,, \\ F^a_{+k} &:= \frac{\partial A^a_k}{\partial x^+} + \mathrm{f}^{abc}\, A^b_+\, A^c_k \,, \\ F^a_{-k} &:= \mathrm{f}^{abc}\, A^b_-\, A^c_k \,, \\ F^a_{ij} &:= \mathrm{f}^{abc}\, A^b_i\, A^c_j \,, \quad i,j,k = 1,2 \,. \end{aligned}$$

Therefore, (20) determines the $SU(n)$ Yang-Mills light-cone mechanics as $4(n^2-1)$- dimensional system with configuration coordinates $A_\pm$, A_k evolving with respect to the light-cone time $\tau := x^+$.

The Legendre transformation

$$\begin{aligned} \pi^+_a &:= \frac{\partial L}{\partial \dot{A}^a_+} = 0 \,, \\ \pi^-_a &:= \frac{\partial L}{\partial \dot{A}^a_-} = \frac{1}{g^2}\, \left(\dot{A}^a_- + \mathrm{f}^{abc}\, A^b_+\, A^c_- \right) \,, \\ \pi^k_a &:= \frac{\partial L}{\partial \dot{A}^a_k} = \frac{1}{g^2}\, \mathrm{f}^{abc}\, A^b_-\, A^c_k \end{aligned}$$

gives the canonical Hamiltonian

$$H_C = \frac{g^2}{2}\, \pi^-_a \pi^-_a - \mathrm{f}^{abc}\, A^b_+ \left(A^c_-\, \pi^-_a + A^c_k\, \pi^k_a \right) + \frac{1}{2g^2} F^a_{12} F^a_{12} \,. \tag{21}$$

The non-vanishing Poisson brackets between the fundamental canonical variables are

$$\{ A^a_\pm \,, \pi^\pm_b \} = \delta^a_b \,, \qquad \{ A^a_k \,, \pi^l_b \} = \delta^l_k \delta^a_b \,.$$

The Hessian of the Lagrangian system (20) is degenerate, $\det \| \frac{\partial^2 L}{\partial \dot{A} \partial \dot{A}} \| = 0$, and as a result there are primary constraints whose computation by the algorithm of Section 3.1 gives

$$\varphi^{(1)}_a := \pi^+_a = 0 \,, \tag{22}$$

$$\chi^a_k := g^2\, \pi^a_k + \mathrm{f}^{abc}\, A^b_- A^c_k = 0 \,. \tag{23}$$

The non-vanishing Poisson brackets between these constraints are

$$\{\chi_i^a\,,\chi_j^b\} = \,2\mathrm{f}^{abc}\eta_{ij}A_-^c\,.$$

According to the Dirac prescription, the presence of primary constraints affects the dynamics of the degenerate system. Now the generic evolution is governed by the total Hamiltonian

$$H_T := H_C + U_a(\tau)\varphi_a^{(1)} + V_k^a(\tau)\chi_k^a\,,$$

where the Lagrange multipliers $U_a(\tau)$ and $V_k^a(\tau)$ are unspecified functions of the light-cone time τ. Using this Hamiltonian the dynamical self-consistence of the primary constraints may be checked. From the requirement of conservation of the primary constraints $\varphi_a^{(1)}$ it follows that

$$0 = \dot{\varphi}_a^{(1)} = \{\pi_a^+\,,H_T\} = \mathrm{f}^{abc}\left(A_-^b\pi_c^- + A_k^b\pi_c^k\right)\,. \tag{24}$$

Therefore, there are three secondary constraints $\varphi_a^{(2)}$

$$\varphi_a^{(2)} := \mathrm{f}_{abc}\left(A_-^b\pi_c^- + A_k^b\pi_c^k\right) = 0 \tag{25}$$

which obey the $SU(n)$ algebra

$$\{\varphi_a^{(2)}\,,\varphi_b^{(2)}\} = \mathrm{f}_{abc}\,\varphi_c^{(2)}\,.$$

The same procedure for the primary constraints χ_k^a gives the following self-consistency conditions

$$0 = \dot{\chi}_k^a = \{\chi_k^a\,,H_C\} - 2\,g^2\,\mathrm{f}^{abc}\,V_k^b\,A_-^c\,.$$

A further issue, the identification of the first class constraints among the primary constraints χ_k^a, depends on the rank of the structure group. Below we specify to the simplest special unitary group of rank one.

4.1 The $SU(2)$ structure group

Here we present the results of our computations performed for the case of $SU(2)$ algebra where the structure constants are given by the totally antisymmetric three dimensional Levi-Civita symbol, $\mathrm{f}^{abc} = \epsilon^{abc}$.
Constraints and their separation. Computation of the complete set of constraints, as described in Section 3.2, gives nine primary constraints $\varphi_a^{(1)}, \chi_k^a$ and three secondary constraints $\varphi_a^{(2)}$, in accordance with (22) and (25). Performing the separation of the primary constraints (23) we find two additional first-class constraints

$$\psi_k := A_-^a \chi_k^a \,,$$

and four second class constraints

$$\chi_{k\perp}^a := \chi_k^a - \frac{\left(A_-^b \chi_k^b\right) A_-^a}{(A_-^1)^2 + (A_-^2)^2 + (A_-^3)^2} \,.$$

The new first class constraints ψ_i are abelian, $\{\psi_i \,, \psi_j\} = 0$, and also have zero Poisson brackets with all other constraints, while the second class constraints $\chi_{k\perp}^a$ have the following non-zero Poisson bracket relations

$$\begin{aligned} \{\chi_{i\perp}^a \,, \chi_{j\perp}^b\} &= 2\, g^2\, \epsilon^{abc}\, A_-^c\, \delta_{ij} \,, \\ \{\varphi_a^{(2)} \,, \chi_{k\perp}^b\} &= \epsilon^{abc}\, \chi_{k\perp}^c \,. \end{aligned}$$

Summarizing, there are 8 first-class constraints $\varphi_a^{(1)}, \psi_k, \varphi_a^{(2)}$ and 4 second-class constraints $\chi_{k\perp}^a$.
Generator of local symmetry transformations. The presence of two first class constraints ψ_i raises the question of the existence of new local symmetries as well as the expected $SU(2)$ gauge symmetry. To clarify this point we construct the corresponding generator of local symmetry transformation following Section 3.2. We start from the expression

$$G = \sum_{a=1}^{3} \varepsilon_a^{(1)} \varphi_a^{(1)} + \sum_{i=1}^{2} \eta_i \psi_i + \sum_{a=1}^{3} \varepsilon_a^{(2)} \varphi_a^{(2)} \,, \tag{26}$$

with the eight light-cone time-dependent functions $\varepsilon_a^{(1)}(\tau)$, $\varepsilon_a^{(2)}(\tau)$ and $\eta_i(\tau)$, then compute the functions ρ (see eq. (14)). Equation (16) reads now as

$$\left(\dot{\varepsilon}_a^{(2)} + \varepsilon_a^{(1)} - \epsilon_{abc}\varepsilon_b^{(2)}A_+^c - \eta_i A_i^a\right)\phi_a^{(2)} \stackrel{\Sigma_1}{=} 0\,.$$

Therefore expressing $\varepsilon_a^{(1)}$ in terms of the functions $\varepsilon_a^{(2)}$, the generator of local transformation takes the final form

$$G = \left(-\dot{\varepsilon}_a^{(2)} + \epsilon_{abc}\varepsilon_b^{(2)}A_+^c + \eta_i A_i^a\right)\phi_a^{(1)} + \eta_i\psi_i + \varepsilon_a^{(2)}\phi_a^{(2)}\,. \qquad (27)$$

Analyzing the changes of the canonical coordinates A^a and π^a generated by (27) we find that the abelian subgroup of the 5-parameter local symmetry is in some sense "inherited" from the rigid conformal symmetry of initial Yang-Mills theory. But now, instead of the conformal symmetry, the light-cone $SU(2)$ Yang-Mills mechanics has the $SL(2,R)$ dynamical group of symmetry. Moreover, the group action is accompanied by the abelian transformations generated by two constraints ψ_i. A detailed discussion of this symmetry realization will be given elsewhere.

Hamiltonian reduction to unconstrained system. Now that we have the generator of local transformation, we can address the question of finding a set of suitable coordinates part of which represent the invariants of these transformations. Solving this problem will let us project our system onto the constraint manifold and thus determine the unconstrained Hamiltonian system. We refer for details to [15], and here present the set of corresponding singlet variables (as an example of the solution of the second part of (**Issue III**). We also give a result of subsequent implementation of a Hamiltonian reduction (**Issue IV**) of the "redundant" degrees of freedom associated to the symmetries generated by constraints $\varphi_a^{(1)}$, $\varphi_a^{(2)}$ and ψ_a.

Let us pass to a matrix notation: the 3×3 matrix A_{ab} whose entries of the first two columns are A_i^a and the third column is composed by the elements A_-^a. Now one can verify that the elimination of local degrees of freedom associated with the three constraints $\varphi_a^{(2)}$ can be achieved by using the *polar representation* [22]

$$A = OS$$

where S is a positive definite 3×3 symmetric matrix and the orthogonal matrix O is parameterized by three Euler angles. It turns out that these three angles represent the pure gauge degrees of freedom corresponding to the constraints $\varphi_a^{(2)}$.

To find the gauge degrees connected with the remaining two abelian constraints ψ_1, ψ_2 one can pass to a principal axes representation for the symmetric matrix S

$$S = R^T \,\mathrm{diag}\,(q_1, q_2, q_3)\; R$$

with the orthogonal matrix $R(\chi_1, \chi_2, \chi_3)$ given in terms of the Euler angles (χ_1, χ_2, χ_3). Now again it turns out that the two angles χ_1 and χ_2 are pure gauge degrees of freedom.
Solving for the remaining second class constraints ${\chi_i^a}_\perp$ leads to an unconstrained system which represents a *free particle* or, considering the complex solutions to the second class constraints, to a more interesting model, the so-called *conformal mechanics*. In this case the diagonal variable q_1 and the angular variable χ_3 together with the corresponding conjugate momenta p_1 and p_{χ_3} are two unconstrained canonical pairs and their dynamics is governed by the reduced Hamiltonian

$$H = \frac{g^2}{2}\left(p_1^2 + \frac{p_{\chi_3}^2}{4}\,\frac{1}{q_1^2}\right), \tag{28}$$

which is a projection of the canonical Hamiltonian (21) to the constraints shell. Finally, noting that p_{χ_3} is a constant of motion, the Hamiltonian (28) coincides with the Hamiltonian of conformal mechanics with the coupling constant $p_{\chi_3}^2/4$.

5 Concluding comments

In this paper we have raised several issues for a constrained mechanical systems which require computational realization. We described how using the Gröbner basis technique the computation and separation of the complete

set of constraints as well as the construction of the local gauge transformations can be achieved in degenerate mechanical models whose Lagrangians are polynomials in coordinates and velocities. The remaining challenges, namely, the construction of a basis for singlet (gauge-invariant) variables as well as the subsequent Hamiltonian reduction still needs algorithmisation. However, a first step in this direction also has been performed. In systems with first-class constraints the configuration space should be factorized by the local symmetry group in order to find a gauge invariant basis. The infinitesimal structure of a local symmetry group is encoded in the generator of gauge transformations, and we have shown that its construction allows an effective algorithmisation.

As an example of the effectiveness of the proposed algorithms light-cone Yang-Mills mechanics with the $SU(2)$ structure group was analysed in details: we determined and separated constraints, constructed a local invariance transformation and found the equivalent unconstrained Hamiltonian system.

For the $SU(2)$ light-cone mechanics the computations with our implementation in Maple 10, which is an improved and extended version of that given in [12], takes about 1 minute on a machine with a 1.7 GHz processor. This uses the standard Gröbner package in the Maple library. Unfortunately, recent extensions of the Maple Gröbner bases facilities with the packages Gb and Fgb developed by J.C. Faugère [23] do not improve on the standard package. Gb is slower for our problems while Fgb cannot deal with the parametric coefficients. For the same reason we cannot use our software GINV [24] to implement the involutive algorithms [14] for involutive or/and Gröbner bases. Manipulation with parametric coefficients is essential for the Dirac formalism due to the presence of physical parameters (e.g. masses, coupling constants) in the initial Lagrangian, the Lagrange multipliers in the total Hamiltonian (4) and the time-dependent functions in the generator (12) of local symmetry transformations.

Consideration of light-cone mechanics with $n \geq 3$ is under current study. Here we note only that a recent paper [25] on geodesic motion on the

$SU(3)$ group provides us with a useful parametrization suitable for this investigation.

Acknowledgments

The authors are indebted to M. Lavelle, D. McMullan and D. Mladenov for helpful discussions concerning this work. The presented research was partially supported by grant No.04-01-00784 from the Russian Foundation for Basic Research. One of the authors (V.G.) was also supported by grant 5362.2006.2 from the Ministry of Education and Science of the Russian Federation.

References

[1] Dirac P.A.M.: Generalized Hamiltonian Dynamics, *Canad. J. Math.* **2**, 129–148, 1950; *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, Monographs Series, Yeshiva University, New York, 1964.

[2] Sundermeyer K.: *Constrained Dynamics*, Lecture Notes in Physics vol.169, Spinger-Verlag, Berlin Heidelberg New York, 1982.

[3] Hanson E.J., Regge T. and Teitlboim C.: *Constrained Hamiltonian Systems*, Accademia Nazionale dei Lincei, 1976.

[4] Gitman D.M. and Tyutin I.V.: *Quantization of Fields with Constraints*, Springer-Verlag, Berlin Heidelberg, 1990.

[5] Henneaux M. and Teitelboim C.: *Quantization of Gauge Systems*, Princeton University Press, Princeton, New Jersey, 1992.

[6] Hurtley D.H., Tucker R.W. and Tuckey P.: Constrained Hamiltonian Dynamics and Exterior Differential Systems. *J. Phys. A.* **24**, 5252–5265, 1991.

[7] Seiler W.M. and Tucker R.W.: Involution and Constrained Dynamics I: the Dirac Approach. *J. Phys. A.* **28**, 4431–4451, 1995.

[8] Seiler W.M.: Involution and Constrained Dynamics. II: the Faddeev-Jackiw Approach, *J. Phys. A.* **28**, 7315–7331, 1995.

[9] Pommaret J.F.: *Partial Differential Equations and Group Theory. New Perspectives for Applications*, Kluwer, Dordrecht, 1994.

[10] Calmet J., Hausdorf M. and Seiler W.M.: A Constructive Introduction to Involution, in: R.Akerkar (Ed.) *International Symposium on Applications of Computer Algebra-ISACA 2000*, Allied Publishers, New Delhi, 33–50,2001.

[11] Buchberger B. and Winkler F. (Eds.): *Gröbner Bases and Applications*, Cambridge University Press, 1998.
[12] Gerdt V.P. and Gogilidze S.A.: Constrained Hamiltonian Systems and Gröbner Bases, in: V.G.Ganzha, E.W.Mayr and E.V.Vorozhtsov (Eds.), *Computer Algebra in Scientific Computing*, Springer-Verlag, Berlin, 138-146, 1999.
[13] Gerdt V.P., Blinkov Yu.A.: Involutive Bases of Polynomial Ideals. *Math. Comp. Simul.* **45**, 519–542, 1998, arXiv:math.AC/9912027.
[14] Gerdt V.P.: Involutive Algorithms for Computing Gröbner Bases, in: S.Cojocaru, G.Pfister and V.Ufnarovski (Eds), *Computational Commutative and Non-Commutative algebraic geometry*, NATO Science Series, IOS Press, 199–225, 2005, arXiv:math.AC/0501111.
[15] Gerdt V.P., Khvedelidze A.M. and Mladenov D.M.: *Light-cone $SU(2)$ Yang-Mills Theory and Conformal Mechanics*, arXiv:hep-th/0210022.
[16] Gogilidze S.A, Pervushin V.N. and Khvedelidze A.M.: Reduction in Systems with Local Symmetry, *Phys. Part. Nucl.* **30**, 160–208, 1999.
[17] Khvedelidze A.M.: On the Hamiltonian Formulation of Gauge Theories in Terms of Physical Variables, *J. Math. Sci.* **119**, 513–555, 2004.
[18] Cox D., Little J. and O'Shea D.: *Ideals, Varieties and Algorithms*. 2nd Edition, Springer-Verlag, New York, 1996.
[19] Gogilidze S.A., Sanadze V.V., Surovtsev Y.S. and Tkebuchava F.G.: Local Symmetries in Systems with Constraints, *J. Phys. A* **27** 6509–6524, 1994.
[20] Becker T. and Weispfenning V.: *Gröbner Bases. A Computational Approach to Commutative Algebra*. Graduate Texts in Mathematics **141**, Springer-Verlag, New York, 1993.
[21] Brodsky S.J., Pauli H.-C. and Pinsky S.S.: Quantum Chromodynamics and Other Field Theories on the Light-cone, *Phys. Rep.* **301**, 299–486, 1998.
[22] Zelobenko D.R.: *Compact Lie Groups and Their Representations*. Translations of Mathematical Monographs, Vol. 40, AMS, 1978.
[23] `http://fgbrs.lip6.fr/salsa/Software/`
[24] `http://invo.jinr.ru`
[25] Gerdt V., Horan R., Khvedelidze A., Lavelle M., McMullan D. and Palii Y.: On the Hamiltonian Reduction of Geodesic Motion on $SU(3)$ to $SU(3)/SU(2)$, *J. Math. Phys.* **47**, 2006, arXiv:hep-th/0511245.

The Electrodynamics of Charged Continua[1]

J Gratus, D A Burton, R W Tucker

Lancaster University and the Cockcroft Institute, UK

Abstract

The dynamic behaviour of a distribution of charged particles is explored in terms of a self-permeable continuum model interacting self-consistently with the Maxwell field in vacuo. The model is developed using intrinsic tensor field theory and exploits to the full the relativistic structure of Minkowski spacetime. The model predicts the dynamic formation of domains that separate multi-component currents. To determine the location of such domains one is confronted with a new type of electrodynamic problem in which the number of charged current components is indefinite and the state of a finite bunch of charge may approach a highly mixed configuration reminiscent of turbulence. In this paper a formalism is established to describe such a multi-component system in terms of a flow map between 4−manifolds. This map inter-relates a complex Euler description of electrodynamics on spacetime with a computational Lagrangian scheme on a 4-dimensional body-time manifold, the domain of the flow map.

Keywords: Maxwell, Charged Fluid, Multi-component Currents, Relativistic Dynamics, Self-permeable, Charged Bunches, Accelerator Science, Fields over maps, Lagrangian picture.

1 Introduction

Modern theories regard matter as being composed of interacting particles. A fruitful way to formulate these interactions is in terms of fields whose sources are related to the particles themselves [1], [2]. In most classical and quantum descriptions the fields and sources are sections of bundles over spacetime that fulfil the requirements placed on them by physical

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

laws compatible with observation. Although the notion of a point particle is probably no more than a convenient idealisation in classical physics it forms the conceptual basis of many models of charged sources in electrodynamics [3], [5]. However the physical laws of electrodynamics sit uneasily in such a framework and require awkward manoeuvres to eliminate the self-interaction effects attributed to the fields produced by particles acting on themselves [4]. Models of matter that adopt charged continua as fundamental concepts can evade these issues [6], [7]. They have the advantage that notions of continuity and differentiability can then be controlled mathematically in the field equations that determine the dynamics of such continua. Furthermore by regarding the motion of charged continua as a subset of spacetime on which a smooth 4-velocity vector field is defined [8], the notion of a particle history can be recovered by identifying it with a particular parameterised integral curve of such a vector field. The distribution of integral curves can be specified by a measure (the proper charge density) obtained in principle by solving Maxwell's equations in conjunction with a force law based on the vanishing of the divergence of the total stress-energy-momentum tensor of the complete system.

Although this program leads to a well defined differential system for the electromagnetic fields, source density and velocity field it is rare that initial conditions exist leading to a smooth vector field on spacetime. The existence of crossing integral curves after a finite time means that the premise on which the model is based breaks down. This is a common occurrence in many fluid models for flow fields. In neutral gas dynamics such occurrences are identified with the formation of shocks and appeal is made to dissipative effects to ameliorate singularities that arise as a result. Although energy dissipation can arise in many dynamic configurations of charged continua appeal to a similar amelioration is not available for systems controlled solely by electrodynamic forces and a new physical scenario must be accommodated in the model.

The approach adopted here is to regard a charged continuum subject to purely electrodynamic forces as a self-permeable structure that permits

self-penetration. It may be described as the continuum limit of a collection of dynamically interacting but non-colliding particles. Alternatively it may be regarded as a multi-component continuum described by a collection of vector fields on subsets of spacetime that have supports determined by the global dynamics of the entire system. At the interface of such subsets the measure describing the smooth source proper charge density may degenerate from a volume charge density to a surface or even line charge density. To determine the location of and interaction between these lower dimensional sources means that one is confronted with a new type of electrodynamic problem in which the number of charged current components is indefinite and the state of a finite bunch of charge may approach a highly mixed configuration reminiscent of turbulence.
In this paper a formalism is established to describe such a multi-component system in terms of a flow map whose properties follow from a generalisation of the single component scheme outlined above. This map inter-relates a complex Euler description of electrodynamics on spacetime with a computational Lagrangian scheme on a 4-dimensional body-time manifold, the domain of the flow map.

2 Fields over maps

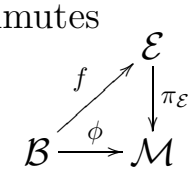
To establish notation the reader is reminded about the notion of a section over a map. Let $\phi : \mathcal{B} \to \mathcal{M}$ be a continuous map between manifolds $\mathcal{B}$ and $\mathcal{M}$ (assumed orientable), and let $\pi_{\mathcal{E}} : \mathcal{E} \to \mathcal{M}$ be a bundle over $\mathcal{M}$. The notation

$$f \in \Gamma(\phi, \mathcal{E})$$

means that

$$f : \mathcal{B} \to \mathcal{E} \qquad \text{and} \qquad \pi_{\mathcal{E}} \circ f = \phi \tag{1}$$

i.e. the following diagram commutes (2)



The map f may be referred to as a *section of* $\mathcal{E}$ *over* ϕ or an $\mathcal{E}$ *field over* ϕ. Scalar fields over $\mathcal{B}$ may also be regarded as scalar fields over ϕ i.e.

$$\Gamma\Lambda^0\mathcal{B} = \Gamma(\phi, \Lambda^0\mathcal{M}) \tag{3}$$

where $\Lambda^p\mathcal{M}$ is the bundle of exterior $p-$forms over $\mathcal{M}$ and in general $\Gamma\mathcal{E}$ denotes the space of sections of $\mathcal{E}$.
Let $\mathcal{M}$ be (Minkowski) $4-$dimensional spacetime with metric tensor g and associated Hodge map $\star$. The canonical measure on $\mathcal{M}$ is taken as $\star 1$. A general point in $\mathcal{M}$ will be written $p \in \mathcal{M}$. Denote $\mathcal{B} = \mathbb{R} \times \underline{\mathcal{B}}$ as the four dimensional *body-time* manifold, where $\underline{\mathcal{B}}$ is a three dimensional oriented *body manifold*. A general point in $\mathcal{B}$ will be written $P = (\tau, \underline{P}) \in \mathcal{B}$. Since $\mathcal{B} = \mathbb{R} \times \underline{\mathcal{B}}$, there exist projection maps

$$\tau : \mathcal{B} \to \mathbb{R}\,, \qquad \tau(\tau', \underline{P}) = \tau' \tag{4}$$

and

$$\underline{\pi} : \mathcal{B} \to \underline{\mathcal{B}}\,, \qquad \underline{\pi}(\tau', \underline{P}) = \underline{P} \tag{5}$$

These give rise to a preferred vector field $\mathcal{T} \in \Gamma T\mathcal{B}$ which may be written

$$\mathcal{T} = \partial_\tau \tag{6}$$

The model under consideration is constructed in terms of two fundamental fields: the flow field

$$C : \mathcal{B} \to \mathcal{M} \tag{7}$$

and the electromagnetic field

$$F \in \Gamma\Lambda^2\mathcal{M}, \tag{8}$$

each assumed to be continuous with degrees of differentiability as required.

3 Generic and Critical points

In general the flow map $C : \mathcal{B} \to \mathcal{M}$ is assumed neither surjective (onto) nor injective (one-to-one). As a result, for any point $p \in \mathcal{M}$, there may exists none, one, many or even an infinite number of roots $P \in \mathcal{B}$ which solve the equation $C(P) = p$. The set inverse of C is defined as

$$C^{-1}(U) = \left\{ P \in \mathcal{B} \,\middle|\, C(P) \in U \right\} \qquad \text{for} \qquad U \subset \mathcal{M} \tag{9}$$

Thus

$$C^{-1}(\{p\}) = \left\{ P \in \mathcal{B} \,\middle|\, C(P) = p \right\} \tag{10}$$

Let the function $N : \mathcal{M} \to \{0, 1, 2, \ldots\} \cup \{\infty\}$,

$$N(p) = \text{number of elements of } C^{-1}(\{p\}) \tag{11}$$

be the number of roots of C. If the number of roots is finite write

$$C^{-1}(\{p\}) = \left\{ P_{[1]}, \ldots, P_{[N(p)]} \right\} \tag{12}$$

Square bracketed subscripts are used to label the roots and any structure associated with the root $P_{[i]}$.
Let $p \in \mathcal{M}$ and $\left\{ P_{[1]}, \ldots, P_{[N(p)]} \right\} = C^{-1}(\{p\})$ with $0 < N(p) < \infty$. The point p is said to be *generic* if there exist open disjoint neighbourhoods $U^{\mathcal{B}}_{[1]}, \ldots, U^{\mathcal{B}}_{[N(p)]} \subset \text{interior}(\mathcal{B})$ and $U^{\mathcal{M}} \subset \mathcal{M}$ such that $P_{[i]} \in U^{\mathcal{B}}_{[i]}$, $p \in U^{\mathcal{M}}$,

$$\bigcup_{i=1}^{N(p)} U^{\mathcal{B}}_{[i]} = C^{-1}(U^{\mathcal{M}}) \tag{13}$$

and such that the maps

$$C_{[i]} = C|_{U^{\mathcal{B}}_{[i]}} : U^{\mathcal{B}}_{[i]} \to U^{\mathcal{M}} \tag{14}$$

are diffeomorphisms, i.e. $C_{[i]}$ is differentiable and invertible and $C^{-1}_{[i]}$ is differentiable.

For $p \in \mathcal{M}$ such that $N(p) = 0$, p is defined as generic if N is continuous at p. Recall that for integer valued functions, being continuous at a point is equivalent to being constant about that point.
It may be shown that a point $p \in \mathcal{M}$ is generic if and only if it obeys the following four conditions

- The set $\quad \left\{P_{[1]}, \ldots, P_{[N(p)]}\right\} = C^{-1}(\{p\}) \quad$ is finite (15)
- The Jacobian of C at $P_{[i]}) \neq 0$ (16)
- $P_{[i]} \notin \partial\mathcal{B}$ (17)
- $N : \mathcal{M} \to \{0, 1, 2, \ldots\} \cup \{\infty\} \quad$ is continuous at $\quad p$ (18)

All points which are not generic are called *critical*. Denote the set of all critical points $p \in \mathcal{M}$ by $\mathcal{M}_{\text{crit}}$ and the set of all generic points $p \in \mathcal{M}$ by $\mathcal{M}_{\text{gen}}$. From the definition of generic points the function N is continuous on $\mathcal{M}_{\text{gen}}$ but in general is not continuous on $\mathcal{M}$. For a given neighbourhood $U^{\mathcal{M}}$, all $p' \in U^{\mathcal{M}}$ have the same number of pre-images, $N(p')$, so $N(U^{\mathcal{M}}) = N(p')$ can be defined. The point $P \in \mathcal{B}$ will be called critical if it is the pre-image of a critical point $p \in \mathcal{M}$. Thus

$$P \quad \text{is critical in } \mathcal{B} \text{ if } \quad C(P) \quad \text{is critical in } \mathcal{M} \tag{19}$$

yielding the corresponding sets $\mathcal{B}_{\text{crit}}$ and $\mathcal{B}_{\text{gen}}$. The sets $\mathcal{M}_{\text{gen}} \subset \mathcal{M}$ and $\mathcal{B}_{\text{gen}} \subset \mathcal{B}$ are open sets and $\mathcal{M}_{\text{crit}} \subset \mathcal{M}$ and $\mathcal{B}_{\text{crit}} \subset \mathcal{B}$ are closed sets.

4 Differential equations for the flow field

The map (7) and the preferred vector field $\mathcal{T}$ give rise to the vector valued map

$$\dot{C} \in \Gamma(C, T\mathcal{M}) \, ; \qquad \dot{C}(P) = C_\star(\mathcal{T}|_P) \in T_{C(P)}\mathcal{M} \tag{20}$$

This is the push forward, under $C_\star$, of the vector $\mathcal{T}_P = \mathcal{T}|_P \in T_P\mathcal{B}$ to the tangent fibre $T_{C(P)}\mathcal{M}$. In terms of coordinate maps $(x^\mu, \dot{x}^\mu)$ for $T\mathcal{M}$ the field $\dot{C}$ over C is given by

$$x^\mu(\dot{C}(P)) = x^\mu(C(P)) \qquad \text{and} \qquad \dot{x}^\mu(\dot{C}(P)) = \mathcal{T}_P(x^\mu \circ C) \tag{21}$$

The field C also gives rise to a collection of curves given by

$$C_{\underline{P}} : \mathbb{R} \to \mathcal{M} \quad \text{for} \quad \underline{P} \in \underline{\mathcal{B}} \,, \qquad C_{\underline{P}}(\tau) = C(\tau, \underline{P}) \tag{22}$$

The total derivative of $C_{\underline{P}}$ is given by

$$\dot{C}_{\underline{P}} \in \Gamma(C_{\underline{P}}, T\mathcal{M}) \,, \qquad \dot{C}_{\underline{P}}(\tau') = \frac{dC_{\underline{P}}}{d\tau}(\tau') = C_{\underline{P}\star}(\partial_\tau|_{\tau'}) = \dot{C}(\tau', \underline{P}) \tag{23}$$

The coordinate τ is chosen so that $\dot{C}$ is a unit *timelike* field

$$g(\dot{C}, \dot{C}) = -1 \tag{24}$$

Furthermore let it be related to the Maxwell 2-form F on $\mathcal{M}$ by the equation of motion

$$\nabla_{\dot{C}} \dot{C} = \widetilde{i_{\dot{C}} F} \tag{25}$$

where for any vector field X on $\mathcal{M}$, the $1-$form $\tilde{X} \equiv g(X, -)$.
The above yields an ordinary differential system for the curves $C_{\underline{P}}$ given by

$$\nabla_{\dot{C}_{\underline{P}}} \dot{C}_{\underline{P}} = \widetilde{i_{\dot{C}_{\underline{P}}} F} \tag{26}$$

In the coordinates $(x^\mu, \dot{x}^\mu)$ on $T\mathcal{M}$ this becomes

$$\ddot{C}^\mu_{\underline{P}}(\tau) + \Gamma^\mu{}_{\alpha\beta}(p) \dot{C}^\alpha_{\underline{P}}(\tau) \dot{C}^\beta_{\underline{P}}(\tau) = F^{\mu\alpha}(p) \dot{C}^\beta_{\underline{P}}(\tau) g_{\alpha\beta}(p) \tag{27}$$

where $p = C(\tau, \underline{P})$, $C^\mu_{\underline{P}}(\tau) = x^\mu(C_{\underline{P}}(\tau))$

$$\dot{C}^\mu_{\underline{P}}(\tau) = \dot{x}^\mu(C_{\underline{P}}(\tau)) = \frac{dC^\mu_{\underline{P}}(\tau)}{d\tau}, \quad \ddot{C}^\mu_{\underline{P}}(\tau) = \frac{d\dot{C}^\mu_{\underline{P}}(\tau)}{d\tau} = \frac{d^2 C^\mu_{\underline{P}}(\tau)}{d\tau^2}$$

and where $\Gamma^\mu{}_{\alpha\beta}(p)$ denote the coordinate components of the Levi-Civita connection ∇ at p.

5 Proper Charge Density and the map Δ

A Lagrangian current 3-form $\mathcal{J}$ on $\mathcal{B}$ will be identified with a measure (non-vanishing 3−form) $\underline{\mathcal{J}} \in \Gamma\Lambda^3\underline{\mathcal{B}}$ on $\underline{\mathcal{B}}$

$$\mathcal{J} = \underline{\pi}^\star \underline{\mathcal{J}} \in \Gamma\Lambda^3\mathcal{B} \tag{28}$$

This induces a natural measure on $\mathcal{B}$ given by

$$d\tau \wedge \mathcal{J} \in \Gamma\Lambda^4\mathcal{B} \tag{29}$$

Note that $d\mathcal{J} = d\underline{\pi}^\star\underline{\mathcal{J}} = \underline{\pi}^\star d\underline{\mathcal{J}} = 0$ and $i_{\mathcal{T}}\mathcal{J} = i_{\mathcal{T}}\underline{\pi}^\star\underline{\mathcal{J}} = \underline{\pi}^\star i_{\underline{\pi}_\star\mathcal{T}}\underline{\mathcal{J}} = 0$. The measure (29) enables us to define the map Δ related to the Jacobian of the flow field C

$$\Delta \in \Gamma\Lambda^0\mathcal{B} = \Gamma(\phi, \Lambda^0\mathcal{M}) \,, \qquad \Delta\, d\tau \wedge \mathcal{J} = C^\star(\star 1) \tag{30}$$

Furthermore since τ is used to define the unit timelike field $C_\star\mathcal{T}$ this can be identified as the inverse of the partial proper charge density scalar, given by

$$\rho : \mathcal{B} \to \mathbb{R} \cup \{\infty\} \,, \qquad \rho = \frac{1}{|\Delta|} \tag{31}$$

It will be shown below that the pull back by C of $i_{\dot{C}} \star 1$ is the pull back by C of the total electric current 3−form $\star\tilde{J}$ on $\mathcal{M}$. Regions on $\mathcal{B}$ where $\Delta = 0$ and hence $\rho = \infty$ may be identified with loci having a finite surface or line charge density.

6 Example

Before considering a coupled problem in which the flow field depends on F through Maxwell's equations it is of interest to examine an artificial but non-trivial flow field that exhibits features that may be expected to arise in the coupled situation.

Let $\underline{\mathcal{B}} = I \times \mathbb{R}^2$ with coordinates (σ, Y, Z), where $\sigma \in I \subset \mathbb{R}$ is the closed-open interval $I = \{\sigma | 0 \leq \sigma < 1\}$. This interval will demonstrate the various types of critical points that can arise[2]. Let spacetime $\mathcal{M}$ have Cartesian coordinates (t, x, y, z) with metric $g = -dt \otimes dt + dx \otimes dx + dy \otimes dy + dz \otimes dz$ and choose the measure on $\underline{\mathcal{B}}$ to be

$$\underline{\mathcal{J}} = K(\sigma) d\sigma \wedge dY \wedge dZ \tag{32}$$

where $K(\sigma) > 0$. Define the flow map by

$$\begin{gathered}(t, x, y, z) = C(\tau, \sigma, Y, Z) = (\hat{t}(\tau, \sigma), \hat{x}(\tau, \sigma), Y, Z) \\ \text{where} \qquad \hat{t}(\tau, \sigma) = \sinh \tau + \sigma \,, \quad \hat{x}(\tau, \sigma) = \cosh \tau\end{gathered} \tag{33}$$

The map Δ then follows from

$$\begin{aligned}\Delta \, d\tau \wedge \pi^\star \underline{\mathcal{J}} &= \Delta \, d\tau \wedge K(\sigma) d\sigma \wedge dY \wedge dZ = C^\star(\star 1) \\ &= (\cosh \tau \, d\tau + d\sigma) \wedge (\sinh \tau \, d\tau) \wedge dY \wedge dZ\end{aligned}$$

Hence

$$\Delta = -\frac{\sinh \tau}{K(\sigma)} \qquad \text{and} \qquad \rho = \frac{K(\sigma)}{|\sinh \tau|} \tag{34}$$

The map C possesses various types of critical points which may be written

$$\mathcal{M}_{\text{crit}} = \mathcal{M}_{\text{crit}}^{\text{degen}} \cup \mathcal{M}_{\text{crit}}^{\text{closed}} \cup \mathcal{M}_{\text{crit}}^{\text{open}}$$

The set $\mathcal{M}_{\text{crit}}^{\text{degen}}$ correspond to the points where $\Delta = 0$ i.e. $\tau = 0$ and hence $x = 1$:

$$\mathcal{M}_{\text{crit}}^{\text{degen}} = \left\{ (t, x, y, z) \in \mathcal{M} \,\middle|\, x = 1 \text{ and } 0 \leq t \leq 1 \right\}$$

The set $\mathcal{M}_{\text{crit}}^{\text{closed}}$ is the image of $\partial\mathcal{B} = \{(\tau, 0) \in \mathcal{B}\}$

$$\mathcal{M}_{\text{crit}}^{\text{closed}} = \left\{ (t, x, y, z) \in \mathcal{M} \,\middle|\, x^2 - t^2 = 1 \text{ and } x \geq 1 \right\}$$

[2] Domains of this type are useful to accommodate fields which would otherwise have singularities in their domains of definition.

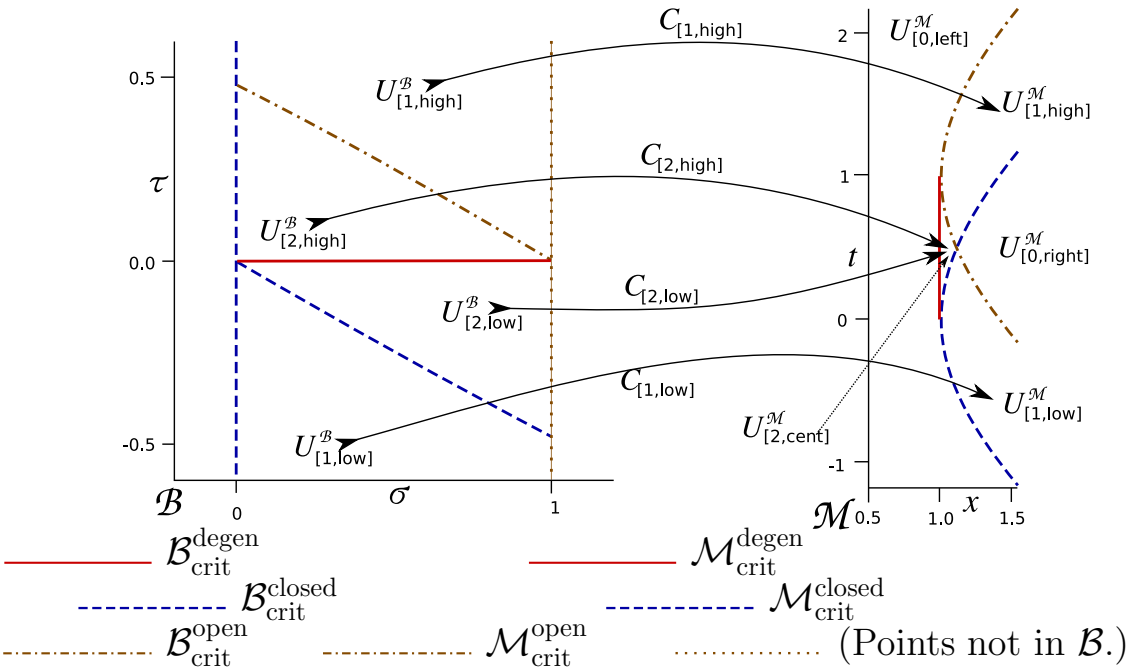


Figure 1: Anatomy of $\mathcal{B}$ and $\mathcal{M}$ with the map C between them. Coordinates Y, Z and y, z are suppressed.

The set $\mathcal{M}^{\text{open}}_{\text{crit}}$ must include the remaining critical points, i.e. those points where N changes but which are not in $\mathcal{M}^{\text{degen}}_{\text{crit}}$ or $\mathcal{M}^{\text{closed}}_{\text{crit}}$.

$$\begin{aligned}\mathcal{M}^{\text{open}}_{\text{crit}} &= \left\{\lim_{\sigma\to 1}(C(\tau,\sigma,Y,Z))\middle|\tau,Y,Z\in\mathbb{R}\right\}\\ &= \left\{(t,x,y,z)\in\mathcal{M}\,\middle|\,x^2-(t-1)^2=1 \text{ and } x\geq 1\right\}\end{aligned}$$

Some of the points in $\mathcal{M}^{\text{open}}_{\text{crit}}$ have pre-images (e.g. given by $\mathcal{B}^{\text{open}}_{\text{crit}}$ below) while others do not. All critical points in $\mathcal{M}_{\text{crit}}$ are indicated on the right of figure 1.
The generic points $\mathcal{M}_{\text{gen}}$ are then the remaining open sets. There are 5 disconnected components of $\mathcal{M}_{\text{gen}}$ shown in figure 1 labelled

$$\mathcal{M}_{\text{gen}} = U^{\mathcal{M}}_{[0,\text{left}]} \cup U^{\mathcal{M}}_{[0,\text{right}]} \cup U^{\mathcal{M}}_{[1,\text{low}]} \cup U^{\mathcal{M}}_{[1,\text{high}]} \cup U^{\mathcal{M}}_{[2,\text{cent}]}$$

The number in each case refers to the number of pre-images.

$$\begin{aligned}
U^{\mathcal{M}}_{[2,\text{cent}]} =& \Big\{ (t,x,y,z) \in \mathcal{M} \,\Big|\, x < \sqrt{1+t^2}\,,\ x < \sqrt{1+(t-1)^2}\,, \\
& \qquad\qquad x > 1 \text{ and } 0 < t < 1 \Big\}\,, \\
U^{\mathcal{M}}_{[0,\text{left}]} =& \Big\{ (t,x,y,z) \in \mathcal{M} \,\Big|\, x < \sqrt{1+t^2}\,, \\
& \qquad x < \sqrt{1+(t-1)^2} \text{ and } (t,x,y,z) \notin U^{\mathcal{M}}_{[2,\text{cent}]} \Big\}\,, \\
U^{\mathcal{M}}_{[0,\text{right}]} =& \Big\{ (t,x,y,z) \in \mathcal{M} \,\Big|\, x > \sqrt{1+t^2}\,,\ x > \sqrt{1+(t-1)^2} \Big\}\,, \\
U^{\mathcal{M}}_{[1,\text{high}]} =& \Big\{ (t,x,y,z) \in \mathcal{M} \,\Big|\, x < \sqrt{1+t^2}\,,\ x > \sqrt{1+(t-1)^2} \Big\}\,, \\
U^{\mathcal{M}}_{[1,\text{low}]} =& \Big\{ (t,x,y,z) \in \mathcal{M} \,\Big|\, x > \sqrt{1+t^2}\,,\ x < \sqrt{1+(t-1)^2} \Big\}
\end{aligned}$$

The critical points on $\mathcal{B}$ are given by

$$\mathcal{B}_{\text{crit}} = \mathcal{B}^{\text{degen}}_{\text{crit}} \cup \mathcal{B}^{\text{closed}}_{\text{crit}} \cup \mathcal{B}^{\text{open}}_{\text{crit}}$$

where

$$\begin{aligned}
\mathcal{B}^{\text{degen}}_{\text{crit}} &= \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \tau = 0 \big\}\,, \\
\mathcal{B}^{\text{closed}}_{\text{crit}} &= \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \sigma = 0 \big\} \cup \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \sigma = -2\sinh\tau \big\}\,, \\
\mathcal{B}^{\text{open}}_{\text{crit}} &= \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \sigma = 1 - 2\sinh\tau \big\}
\end{aligned}$$

These are shown on the left of figure 1. Thus the generic points of $\mathcal{B}$ are given by

$$\mathcal{B}_{\text{gen}} = U^{\mathcal{B}}_{[1,\text{high}]} \cup U^{\mathcal{B}}_{[1,\text{low}]} \cup U^{\mathcal{B}}_{[2,\text{high}]} \cup U^{\mathcal{B}}_{[2,\text{low}]}$$

where

$$\begin{aligned}
U^{\mathcal{B}}_{[1,\text{high}]} &= \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \sigma > 1 - 2\sinh\tau \big\}\,, \\
U^{\mathcal{B}}_{[1,\text{low}]} &= \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \sigma < -2\sinh\tau \big\}\,, \\
U^{\mathcal{B}}_{[2,\text{high}]} &= \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \sigma < 1 - 2\sinh\tau \text{ and } \tau > 0 \big\}\,, \\
U^{\mathcal{B}}_{[2,\text{low}]} &= \big\{ (\tau,\sigma,X,Y) \in \mathcal{B} \,\big|\, \sigma > -2\sinh\tau \text{ and } \tau < 0 \big\}
\end{aligned}$$

The diffeomorphisms from the components of $\mathcal{B}_{\rm gen}$ to the components of $\mathcal{M}_{\rm gen}$ are given by

$$\begin{aligned} &C_{[1,\rm high]} : U^{\mathcal{B}}_{[1,\rm high]} \to U^{\mathcal{M}}_{[1,\rm high]} \ , && C_{[1,\rm low]} : U^{\mathcal{B}}_{[1,\rm low]} \to U^{\mathcal{M}}_{[1,\rm low]} \ , \\ &C_{[2,\rm high]} : U^{\mathcal{B}}_{[2,\rm high]} \to U^{\mathcal{M}}_{[2,\rm cent]} \ , && C_{[2,\rm low]} : U^{\mathcal{B}}_{[2,\rm low]} \to U^{\mathcal{M}}_{[2,\rm cent]} \end{aligned} \tag{35}$$

If the Maxwell field generated by sources defined by C is ignored one can readily find a background electromagnetic field that generates this flow. The C given in (33) obeys the equations of motion with a prescribed constant electric field in the above frame. This follows since

$$\dot{C}(\tau,\sigma) = \big(\cosh\tau\,\partial_t + \sinh\tau\,\partial_x\big)\big|_{C(\tau,\sigma)} \tag{36}$$

so that C generates a normalised timelike velocity field (24) and furthermore setting the prescribed (external) electromagnetic field to

$$F_{\rm ext} = dt \wedge dx \tag{37}$$

confirms that

$$\ddot{C}|_{C(\tau,\sigma)} = (\sinh\tau\,\partial_t + \cosh\tau\,\partial_x)|_{C(\tau,\sigma)} = i_{\widetilde{\dot{C}(\tau,\sigma)}F_{\rm ext}}$$

Thus C is a flow field in the background external electromagnetic field (37).

7 Equations for the electromagnetic field.

In the coupled situation, where there is no external applied electromagnetic field on $\mathcal{M}$ then F in a domain containing generic points obeys Maxwell's equations

$$dF = 0 \tag{38}$$

and

$$d\star F = -\star\widetilde{J} \tag{39}$$

with a current vector field $J \in \Gamma T\mathcal{M}_{\text{gen}}$. At a generic point $p_0 \in \mathcal{M}_{\text{gen}}$ let $C_{[i]} : U^{\mathcal{B}}_{[i]} \to U^{\mathcal{M}}$ be the diffeomorphism given in (14). The partial current $J_{[i]}$ associated with C is defined by

$$J_{[i]} \in \Gamma T U^{\mathcal{M}} , \qquad J_{[i]} = \big(\rho \circ C^{-1}_{[i]}\big)\big(\dot{C} \circ C^{-1}_{[i]}\big) \tag{40}$$

that is for all $p \in U^{\mathcal{M}}$,

$$J_{[i]}|_p = \rho(P_{[i]})\dot{C}(P_{[i]}) \quad \text{where} \quad P_{[i]} = C^{-1}_{[i]}(p) \tag{41}$$

The total current J associated with C at all generic points is defined to be the sum of the partial currents

$$J|_{U^{\mathcal{M}}} = \sum_{i=1}^{N(U^{\mathcal{M}})} J_{[i]} \quad \in \Gamma T U^{\mathcal{M}} \tag{42}$$

The field F is then given at generic points by the Maxwell system above. Continuity conditions of F must be used to define F at critical points.
For an equivalent definition of the partial current, let $p \in \mathcal{M}_{\text{gen}}$, $P_{[i]} = C^{-1}_{[i]}(p)$ and observe that from (30)

$$\Delta(P_{[i]})\,(d\tau \wedge \mathcal{J})|_{P_{[i]}} = C^{\star}_{P_{[i]}}(\star 1)$$

where $C^{\star}_P : \Lambda^q_{C(P)}\mathcal{M} \to \Lambda^q_P\mathcal{B}$ is the pointwise pull back. Contracting with $i_{\mathcal{T}|_{P_{[i]}}}$ gives

$$\begin{aligned}
\Delta(P_{[i]})\,\mathcal{J}|_{P_{[i]}} &= i_{\mathcal{T}|_{P_{[i]}}} C^{\star}_{P_{[i]}}(\star 1) = C^{\star}_{P_{[i]}}(i_{C_{[i]\star}(\mathcal{T}|_{P_{[i]}})} \star 1) = C^{\star}_{P_{[i]}}(i_{\dot{C}(P_{[i]})} \star 1) \\
&= C^{\star}_{P_{[i]}}(\star\widetilde{\dot{C}(P_{[i]})}) = C^{\star}_{P_{[i]}}\Big(\star\,\widetilde{\dot{C}(P_{[i]})}\rho(P_{[i]})\,\big|\Delta(P_{[i]})\big| \Big) \\
&= C^{\star}_{P_{[i]}}(\star\widetilde{J_{[i]}|_p})\,\big|\Delta(P_{[i]})\big|
\end{aligned}$$

Since $C_{[i]}$ is a diffeomorphism, so that $(C^{\star}_{[i]})^{-1} = (C^{-1}_{[i]})^{\star}$, one has the equivalent form

$$\star\widetilde{J_{[i]}}|_p = \operatorname{sign}(\Delta(P_{[i]}))\, C^{-1\,\star}_{P_{[i]}}(\mathcal{J}|_{P_{[i]}}) \tag{43}$$

Since (43) is true for all $p \in U^{\mathcal{M}}$ write (43) in terms of the pull back $C_{[i]}^{-1\star} : \Gamma\Lambda^3 U_{[i]}^{\mathcal{B}} \to \Gamma\Lambda^3 U^{\mathcal{M}}$

$$\star\widetilde{J_{[i]}}|_{U^{\mathcal{M}}} = \mathrm{sign}(\Delta|_{U_{[i]}^{\mathcal{B}}})\, C_{[i]}^{-1\star}(\mathcal{J}) \tag{44}$$

From (44) it follows that the source of F is closed:

$$d \star \widetilde{J_{[i]}}|_{U^{\mathcal{M}}} = \mathrm{sign}(\Delta|_{U_{[i]}^{\mathcal{B}}})\, dC_{[i]}^{-1\star}(\mathcal{J}) = \mathrm{sign}(\Delta|_{U_{[i]}^{\mathcal{B}}})\, C_{[i]}^{-1\star}(d\mathcal{J}) = 0$$

hence at generic points

$$d \star \widetilde{J} = 0 \tag{45}$$

There is also an integral relation which inter-relates (39), (42) and (44). Integral formulae offer a practical method to implement numerical discretisations of the above dynamical equations.
If $\phi : \mathcal{B} \to \mathcal{M}$ is a diffeomorphism, $S \subset \mathcal{M}$ a hypersurface of dimension s and $\omega \in \Gamma\Lambda^s\mathcal{M}$, then the theory of integration gives

$$\int_S \omega = \kappa \int_{\phi^{-1}S} \phi^\star\omega \tag{46}$$

where $\kappa = 1$ if ϕ^{-1} preserves the orientation of S and $\kappa = -1$ otherwise.
Given a 3−dimensional spatial hypersurface $S \subset \mathcal{M}$ such that the set $S \cap \mathcal{M}_{\mathrm{crit}}$ has measure zero let $U_{[i]}^{\mathcal{B}}$ be one of the open sets in (13). Then from (46) and (44) the partial electric charge

$$\begin{aligned} Q_{[i]}[S \cap U^{\mathcal{M}}] &= \int_{S\cap U^{\mathcal{M}}} \star\widetilde{J_{[i]}} = \kappa \int_{C_{[i]}^{-1}(S\cap U^{\mathcal{M}})} C_{[i]}^{\star}(\star\widetilde{J_{[i]}}) \\ &= \mathrm{sign}(\Delta|_{U_{[i]}^{\mathcal{B}}})\kappa \int_{C_{[i]}^{-1}(S)\cap U_{[i]}^{\mathcal{B}}} \mathcal{J}|_{U_{[i]}^{\mathcal{B}}} \end{aligned}$$

Since $Q_{[i]}[S \cap U^{\mathcal{M}}]$ cannot change sign under evolution one must choose

$$\kappa = \mathrm{sign}(\Delta|_{U_{[i]}^{\mathcal{B}}})$$

and so

$$Q_{[i]}[S \cap U^{\mathcal{M}}] = \int_{C^{-1}_{[i]}(S)\cap U^{\mathcal{B}}_{[i]}} \mathcal{J}|_{U^{\mathcal{B}}_{[i]}} = \int_{C^{-1}_{[i]}(S)\cap U^{\mathcal{B}}_{[i]}} \mathcal{J}$$

Summing over all partial currents gives the total charge on $S \cap U^{\mathcal{M}}$

$$\int_{S\cap U^{\mathcal{M}}} \star\widetilde{J} = \sum_{i=1}^{N(U^{\mathcal{M}})} \int_{S\cap U^{\mathcal{M}}} \star\widetilde{J_{[i]}} = \sum_{i=1}^{N(U^{\mathcal{M}})} Q_{[i]}[S \cap U^{\mathcal{M}}] = \sum_{i=1}^{N(U^{\mathcal{M}})} \int_{C^{-1}_{[i]}(S)\cap U^{\mathcal{B}}_{[i]}} \mathcal{J}$$

Since the disjoint union

$$\bigcup_{i=1}^{N(U^{\mathcal{M}})} \left(C^{-1}_{[i]}(S) \cap U^{\mathcal{B}}_{[i]}\right) = C^{-1}(S \cap U^{\mathcal{M}})$$

one has $\int_{S\cap U^{\mathcal{M}}} \star\widetilde{J} = \int_{C^{-1}(S\cap U^{\mathcal{M}})} \mathcal{J}$. Taking the union of all the $U^{\mathcal{M}}$ yields $\int_{S\cap \mathcal{M}_{\text{gen}}} \star\widetilde{J} = \int_{C^{-1}(S\cap \mathcal{M}_{\text{gen}})} \mathcal{J}$ and since $S \cap \mathcal{M}_{\text{crit}}$ has measure zero $\int_S \star\widetilde{J} = \int_{C^{-1}(S)} \mathcal{J}$. But from (39) $\int_{\partial S} \star F = \int_S d \star F = -\int_S \star\widetilde{J}$. Hence

$$\int_{\partial S} \star F = -\int_{C^{-1}(S)} \mathcal{J} \tag{47}$$

This is a global identification of the total electric charge (associated with C) with the integral of $\star F$ over a regular 3−dimensional spacelike hypersurface $S \subset \mathcal{M}$.

8 Example Continued

It is of interest to compute from Maxwell's equations the field F for the flow field prescribed in the example in section 6 above. This of course ignores the back reaction of the electromagnetic field on the source which is taken into account in the fully coupled system.
The inverses of the maps (35) are given by

$$\begin{aligned}
C^{-1}_{[1,\text{high}]}(t,x,y,z) &= (\tau = \operatorname{arccosh}(x), \sigma = t - \sqrt{x^2-1}, Y = y, Z = z)\,,\\
C^{-1}_{[1,\text{low}]}(t,x,y,z) &= (\tau = -\operatorname{arccosh}(x), \sigma = t + \sqrt{x^2-1}, Y = y, Z = z)\,,\\
C^{-1}_{[2,\text{high}]}(t,x,y,z) &= (\tau = \operatorname{arccosh}(x), \sigma = t - \sqrt{x^2-1}, Y = y, Z = z)\,,\\
C^{-1}_{[2,\text{low}]}(t,x,y,z) &= (\tau = -\operatorname{arccosh}(x), \sigma = t + \sqrt{x^2-1}, Y = y, Z = z)
\end{aligned} \tag{48}$$

The partial current $J_{[2,\text{high}]} \in \Gamma TU^{\mathcal{M}}_{[2,\text{cent}]}$ is given by (41), (48), (34) and (36) as

$$\begin{aligned}
&J_{[2,\text{high}]}\big|_{(t,x,y,z)}\\
&\quad = \rho(\tau,\sigma,Y,Z)\dot{C}(\tau,\sigma,Y,Z)\\
&\quad = \rho(\operatorname{arccosh}(x), t-\sqrt{x^2-1}, Y, Z)\dot{C}(\operatorname{arccosh}(x), t-\sqrt{x^2-1}, Y, Z)\\
&\quad = \frac{K(t-\sqrt{x^2-1})}{\sinh(\operatorname{arccosh}(x))}(\cosh(\operatorname{arccosh}(x))\partial_t + \sinh(\operatorname{arccosh}(x))\partial_x)\\
&\quad = K(t-\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t + \partial_x\right)
\end{aligned}$$

and likewise $J_{[2,\text{low}]} \in \Gamma TU^{\mathcal{M}}_{[2,\text{cent}]}$ is given by
$J_{[2,\text{low}]}\big|_{(t,x,y,z)} = K(t+\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t - \partial_x\right)$.
Also $J_{[1,\text{high}]} \in \Gamma TU^{\mathcal{M}}_{[1,\text{high}]}$ is $J_{[1,\text{high}]}\big|_{(t,x,y,z)} = K(t-\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t + \partial_x\right)$
and $J_{[1,\text{low}]} \in \Gamma TU^{\mathcal{M}}_{[1,\text{low}]}$ is $J_{[1,\text{low}]}\big|_{(t,x,y,z)} = K(t+\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t - \partial_x\right)$.

Summing the partial currents in the five domains of $\mathcal{M}_{\text{gen}}$ gives

$$J|_{(t,x,y,z)} = \begin{cases} K(t-\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t + \partial_x\right) + & \\ \quad K(t+\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t - \partial_x\right) & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[2,\text{cent}]} \\ K(t-\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t + \partial_x\right) & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[1,\text{high}]} \\ K(t+\sqrt{x^2-1})\left(\frac{x}{\sqrt{x^2-1}}\partial_t - \partial_x\right) & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[1,\text{low}]} \\ 0 & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[0,\text{left}]} \\ 0 & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[0,\text{right}]} \end{cases}$$

Maxwell's equations (38) and (39) are solved with

$$F = E(t,x)dt \wedge dx \tag{49}$$

where

$$E(t,x) = \begin{cases} k(t+\sqrt{x^2-1}) - k(t-\sqrt{x^2-1}) + E_{-\infty} & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[2,\text{cent}]} \\ k(t+\sqrt{x^2-1}) + E_{-\infty} & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[1,\text{high}]} \\ k(1) - k(t-\sqrt{x^2-1}) + E_{-\infty} & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[1,\text{low}]} \\ E_{-\infty} & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[0,\text{left}]} \\ E = k(1) + E_{-\infty} & \text{if } (t,x,y,z) \in U^{\mathcal{M}}_{[0,\text{right}]} \end{cases}$$

with

$$k(\sigma) = \int_0^\sigma K(\sigma)$$

and $E_{-\infty}$ is a constant.

9 The Spherically Symmetric Coupled System

In this section the coupled system (24), (25), (38), (39) is explored where (40) and (42) define the dynamic sources. A spherically symmetric distribution of charge is considered to simplify the analysis. In spacetime $\mathcal{M}$

with standard spherical coordinates (t, r, θ, ϕ) and metric $g = -dt \otimes dt + dr \otimes dr + r^2 d\theta \otimes d\theta + r^2(\sin\theta)^2 d\phi \otimes d\phi$, write the electromagnetic field

$$F|_{(t,r,\theta,\phi)} = \frac{\mathcal{Q}(t,r)}{r^2} dt \wedge dr \tag{50}$$

for all $(t, r, \theta, \phi) \in \mathcal{M}$ with $r > 0$.
Let $\underline{\mathcal{B}} = I \times S^2$ where $I \subseteq \mathbb{R}_+$, with coordinates (σ, Θ, Φ). By spherical symmetry the solution can be described in terms of fields on and maps between 2-dimensional manifold. These will be shown in bold font.
Let $\boldsymbol{\mathcal{B}} = \mathbb{R} \times I$, coordinated by (τ, σ), be the projected body-time manifold and $\boldsymbol{\mathcal{M}} = \mathbb{R} \times \mathbb{R}_+$ coordinated by (t, r) be the projected spacetime manifold.
Let $\underline{\boldsymbol{\pi}} : \boldsymbol{\mathcal{B}} \to I$ be the projection and $\underline{\boldsymbol{\mathcal{J}}} \in \Gamma\Lambda^1 I$ be the choice of measure, so that $d\tau \wedge \boldsymbol{\mathcal{J}} \in \Gamma\Lambda^2\boldsymbol{\mathcal{B}}$ is a measure on $\boldsymbol{\mathcal{B}}$ where $\boldsymbol{\mathcal{J}} = \underline{\boldsymbol{\pi}}^\star \underline{\boldsymbol{\mathcal{J}}}$.
On $\boldsymbol{\mathcal{M}}$ the induced metric is $\boldsymbol{g} = -dt \otimes dt + dr \otimes dr$. This induces the flat Levi-Civita connection $\boldsymbol{\nabla}$ and Hodge map $\star$ with $\star 1 = dt \wedge dr$.
The projected flow map is

$$\boldsymbol{C} : \boldsymbol{\mathcal{B}} \to \boldsymbol{\mathcal{M}} ; \qquad \boldsymbol{C}(\tau, \sigma) = (\hat{t}(\tau, \sigma), \hat{r}(\tau, \sigma)) \tag{51}$$

Thus the 4-dimensional spherically symmetrical flow map C is given by

$$\begin{aligned} &t \circ C(\tau, \sigma, \Theta, \Phi) = \hat{t}(\tau, \sigma) , \quad r \circ C(\tau, \sigma, \Theta, \Phi) = \hat{r}(\tau, \sigma) , \\ &\theta \circ C(\tau, \sigma, \Theta, \Phi) = \Theta \quad \text{and} \quad \phi \circ C(\tau, \sigma, \Theta, \Phi) = \Phi \end{aligned} \tag{52}$$

Substituting (50) and (52) into the equations of motion (24) and (25) yields

$$\boldsymbol{g}(\dot{\boldsymbol{C}}, \dot{\boldsymbol{C}}) = -1 \tag{53}$$

$$\boldsymbol{\nabla}_{\dot{\boldsymbol{C}}} \tilde{\dot{\boldsymbol{C}}}|_{(\tau,\sigma)} = \frac{\mathcal{Q}(C(\tau,\sigma))}{\hat{r}(\tau,\sigma)^2} \star \tilde{\dot{\boldsymbol{C}}}(\tau, \sigma) \tag{54}$$

where

$$\dot{\boldsymbol{C}}(\tau, \sigma) = C_\star(\partial_\tau|_{(\tau,\sigma)}) \in \Gamma(\boldsymbol{C}, T\boldsymbol{\mathcal{M}}) \tag{55}$$

Here (53) is an equation over $\Gamma\Lambda^0\boldsymbol{\mathcal{B}}$ and (54) is an equation over $\Gamma(\boldsymbol{C}, T\boldsymbol{\mathcal{M}})$. In terms of the component maps $\hat{t}_\sigma(\tau) = \hat{t}(\tau, \sigma)$ and $\hat{r}_\sigma(\tau) = \hat{r}(\tau, \sigma)$, (54) gives the ordinary differential system

$$\ddot{\hat{t}}_\sigma(\tau) = \frac{\mathcal{Q}(\hat{t}_\sigma(\tau), \hat{r}_\sigma(\tau))}{\hat{r}_\sigma(\tau)^2}\,\dot{\hat{r}}_\sigma(\tau) \quad \text{and} \quad \ddot{\hat{r}}_\sigma(\tau) = \frac{\mathcal{Q}(\hat{t}_\sigma(\tau), \hat{r}_\sigma(\tau))}{\hat{r}_\sigma(\tau)^2}\,\dot{\hat{t}}_\sigma(\tau) \tag{56}$$

where $\dot{} = d/d\tau$.
Maxwell's equations yield on $\boldsymbol{\mathcal{M}}$

$$d\mathcal{Q} = \sum_{i=1}^{N(p)} \boldsymbol{C}_{[i]}^{-1\star}(\boldsymbol{\mathcal{J}}) \tag{57}$$

For a spherically symmetric charge distribution, the integral representation (47) reduces to

$$\int_{\partial \boldsymbol{S}} \mathcal{Q} = \int_{\boldsymbol{C}^{-1}(\boldsymbol{S})} \boldsymbol{\mathcal{J}} \tag{58}$$

where $\boldsymbol{S} \subset \boldsymbol{\mathcal{M}}$ is a curve and $\partial \boldsymbol{S}$ are its end points.
For $(t, r) \in \boldsymbol{\mathcal{M}}$ let $\boldsymbol{S}(t, r) = \{(t, r') \in \boldsymbol{\mathcal{M}} \,|\, 0 < r' < r\}$ then

$$\mathcal{Q}(t, r) - \mathcal{Q}(t, 0) = \int_{\boldsymbol{C}^{-1}(\boldsymbol{S}(t,r))} \boldsymbol{\mathcal{J}}$$

$S(t, r)$ represents a spherically symmetric ball of radius r at time t. Since $\boldsymbol{\mathcal{J}}$ is closed in regular domains $\mathcal{Q}(t, 0) = \mathcal{Q}_0$ must be independent of t and hence

$$\mathcal{Q}(t, r) = \int_{\boldsymbol{C}^{-1}(\boldsymbol{S}(t,r))} \boldsymbol{\mathcal{J}} + \mathcal{Q}_0 \tag{59}$$

If $\mathcal{Q}_0$ is non zero one has a point charge fixed at the centre of the ball. For currents that are smooth in regular domains $\mathcal{Q}_0 = 0$. Since (59) involves a field E at time t one must express (56) as a system of o.d.e's with evolution

parameter t. Since $\dot{\boldsymbol{C}}$ is required to be unit future timelike, $\hat{t}_\sigma$ is strictly increasing so set

$$\check{\tau}_\sigma = (\hat{t}_\sigma)^{-1} \qquad \text{and} \qquad \check{r}_\sigma = \hat{r}_\sigma \circ \check{\tau}_\sigma \tag{60}$$

Then, with $\prime = d/dt$

$$\begin{aligned} \dot{\hat{t}}_\sigma(\check{\tau}_\sigma(t)) &= \frac{1}{\check{\tau}'_\sigma(t)}\,, \quad \ddot{\hat{t}}_\sigma(\check{\tau}_\sigma(t)) = -\frac{\check{\tau}''_\sigma(t)}{(\check{\tau}'_\sigma(t))^3}\,, \\ \dot{\hat{r}}_\sigma(\check{\tau}_\sigma(t)) &= \frac{\check{r}'_\sigma(t)}{\check{\tau}'_\sigma(t)} \quad \text{and} \quad \ddot{\hat{r}}_\sigma(\check{\tau}_\sigma(t)) = \frac{\check{r}''_\sigma(t)}{(\check{\tau}'_\sigma(t))^2} - \frac{\check{r}'_\sigma(t)\check{\tau}''_\sigma(t)}{(\check{\tau}'_\sigma(t))^3} \end{aligned} \tag{61}$$

and substituting (61) into (56) yields ordinary differential equations for $\check{\tau}_\sigma(t)$ and $\check{r}_\sigma(t)$. These equations involve $\mathcal{Q}$ so must be solved in conjunction with (59). To express (59) in terms of $\check{\tau}_\sigma(t)$ and $\check{r}_\sigma(t)$ observe that

$$\begin{aligned} \boldsymbol{C}^{-1}(\boldsymbol{S}(t,\check{r}_\sigma(t))) &= \left\{(\tau,\sigma')\middle|\hat{t}_{\sigma'}(\tau) = t \quad \text{and} \quad \hat{r}_{\sigma'}(\tau) < \check{r}_\sigma(t)\right\} \\ &= \left\{(\check{\tau}(t),\sigma')\middle|\check{r}_{\sigma'}(t) < \check{r}_\sigma(t)\right\} \end{aligned}$$

and, since $\underline{\boldsymbol{\pi}}$ is injective on the set $\left\{(\check{\tau}(t),\sigma')\middle|\check{r}_{\sigma'}(t) < \check{r}_\sigma(t)\right\}$

$$\begin{aligned} \int_{\boldsymbol{C}^{-1}(\boldsymbol{S}(t,\check{r}_\sigma(t)))} \boldsymbol{\mathcal{J}} &= \int_{\{(\check{\tau}(t),\sigma')|\check{r}_{\sigma'}(t)<\check{r}_\sigma(t)\}} \boldsymbol{\mathcal{J}} = \int_{\underline{\boldsymbol{\pi}}\{(\check{\tau}(t),\sigma')|\check{r}_{\sigma'}(t)<\check{r}_\sigma(t)\}} \underline{\boldsymbol{\pi}}^{-1\star}(\boldsymbol{\mathcal{J}}) \\ &= \int_{\{(\sigma')|\check{r}_{\sigma'}(t)<\check{r}_\sigma(t)\}} \underline{\boldsymbol{\mathcal{J}}} \end{aligned}$$

Therefore

$$\mathcal{Q}(t,\check{r}_\sigma(t)) = \int_{\{\sigma'|\check{r}_{\sigma'}(t)<\check{r}_\sigma(t)\}} \underline{\boldsymbol{\mathcal{J}}} + \mathcal{Q}_0 \tag{62}$$

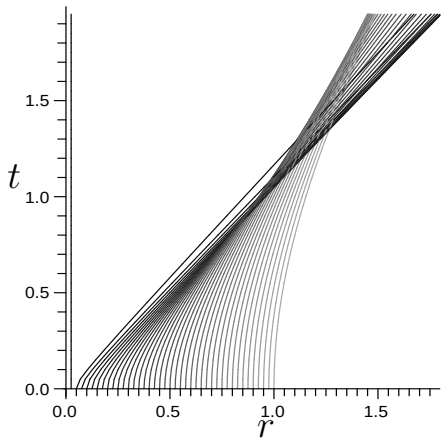


Figure 2: History of a spherically symmetric Gaussian ball of charge evolving from rest. The discretisation used $\sigma_i = i/40, i = 1 \ldots 40$, with $Q_i = 0.05\exp(-(5\sigma_i)^2)$, $\mathcal{Q}_0 = 0$, $t_\lambda = \lambda/20, \lambda = 0 \ldots 40$. $\check{\tau}_i(t_0) = 0$, $\check{\tau}_i'(t_0) = 1$, $\check{r}_i(t_0) = \sigma_i$ and $\check{r}_i'(t_0) = 0$. Evidence for the multi-component nature of the evolution is clearly visible as charge initially closer to the centre overtakes more slowly moving charge in the expanding ball. The curve furthest left, at $r = 0.025$, is vertical since it corresponds to the innermost discretised shell inside of which there is no charge.

Equations (56), (61) and (62) can now be integrated numerically by discretising σ and t. Discretise I by choosing $\sigma_0' < \sigma_1 < \sigma_1' < \sigma_2 < \sigma_2' < \cdots < \sigma_m < \sigma_m' \in I$ with $\sigma_0' = \inf(I)$ and $\sigma_m' = \sup(I)$. Let

$$Q_i = \int_{\sigma_{i-1}'}^{\sigma_i'} \underline{\boldsymbol{\mathcal{J}}}$$

Furthermore, since $\mathcal{Q}(t, r)$ will change with t, the o.d.e system generated from (56) and (61) will be integrated numerically in a series of time bands given by $t_0 < t_1 < \cdots < t_\lambda < \cdots < t_{\max}$. This yields the curves $\check{\tau}_i(t) \approx \check{\tau}_{\sigma_i}(t)$ and $\check{r}_i(t) \approx \check{r}_{\sigma_i}(t)$ for some initial conditions for $\check{\tau}_i(t_0)$, $\check{\tau}_i'(t_0)$, $\check{r}_i(t_0)$ and $\check{r}_i'(t_0)$, where for each t_λ and σ_i, $\mathcal{Q}(t, \check{r}(t))$ for the time interval $t_\lambda < t < t_{\lambda+1}$ is replaced by

$$\mathcal{Q}_i(t_\lambda) = \mathcal{Q}_0 + \sum_{\{j|\check{r}_j(t_\lambda) < \check{r}_i(t_\lambda)\}} Q_j$$

10 Conclusion

A formalism has been established for the description of the motion of electric charge under the influence of both external and self electromagnetic fields. The laws of classical covariant electrodynamics have been expressed in terms of a flow map between a structured body-time manifold $\mathcal{B}$ and Minkowski spacetime. By assuming that this map is not necessarily either surjective or injective, distinct domains in spacetime may be associated with possibly more than one pre-image in the body-time manifold. These pre-images in turn give rise to a complex collection of electric currents that determine the structure of the flow map via Maxwell's equations. The total proper charge density is a dynamic scalar related to the Jacobian of the flow map and a Lagrangian measure on a 3−dimensional body manifold on $\mathcal{B}$.

A simple example of a non-trivial flow map is explicitly constructed corresponding to the plane symmetric motion of charge in a prescribed constant laboratory electric field. It is also demonstrated how Maxwell's equations are treated in the presence of a prescribed source corresponding to this non-trivial flow map. Finally a fully coupled system is considered in terms of the evolution of a spherically symmetric ball of charge from rest with an initially smooth gaussian distribution of charge. The evolution is calculated numerically by discretisising the coupled equations of motion and Maxwell's equations. The results of this simulation (figure 2) indicate that the integral curves $\boldsymbol{C}_\sigma$ cross and that the initial crossing occurs within the charge distribution. As expected the ball of charge explodes outward but with some of the inner spheres of charge overlapping the outer ones. An interesting feature of these solutions is that although F is continuous across regions in spacetime where $\Delta = 0$ (and hence $\rho = \infty$) it is not in general differentiable. This is a general property of solutions where the sources can change discontinuously during the evolution of the coupled system.

The techniques established here have immediate application in accelerator science particularly in devices where charged bunches with large laboratory charge densities in ultra-relativistic motion are demanded [9]. They extend

naturally to multi-component continua such as plasmas where the phenomena of "wave breaking" in wake-field accelerators and bubble regimes may benefit from an analysis in terms of relativistic flow maps with properties analogous to those presented in this paper.

Acknowledgements The authors are grateful to the Cockcroft Institute for making this work possible, to the EPSRC for a Portfolio award and to the EC for a Framework 6 (FP6-2003-NEST-A) award. RWT is also grateful to EPSRC for a Springboard Fellowship.

References

[1] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Pergamon, Oxford, 1962).

[2] J. D. Jackson, *Classical Electrodynamics*, 3rd ed., (Wiley, 1998).

[3] F. Rohrlich, *Classical Charged Particles* (Addison-Wesley, Redwood City, 1990).

[4] P. A. M. Dirac, Proc. Roy. Soc. London A **167**, 148 (1938).

[5] C. Teitelboim, D. Villarroel and C. G. van Weert, Riv. Nuovo Cim. **3**, 9, (1980).

[6] R. C. Davidson, *Physics of Nonneutral Plasmas*, (Addison-Wesley, 1990. Reissued by World Scientific, 2001).

[7] R. C. Davidson, *Physics of Intense Charged Particle Beams*, (World Scientific, 2001).

[8] D A Burton, J Gratus and R W Tucker, Asymptotic Analysis of Ultra-relativistic Charge, Ann. Physics, 2006 (To Appear)

[9] D A Burton, J Gratus and R W Tucker, Phys. Rev. E, 2006, (Submitted)

δ- and Quasi-Regularity for Polynomial Ideals[1]

Marcus Hausdorf [1] Mehdi Sahbi [2] Werner M. Seiler [1]

[1] Computational Mathematics Group, Universität Kassel
34109 Kassel, Germany
[hausdorf,seiler]@mathematik.uni-kassel.de
[2] Institut für Algorithmen und Kognitive Systeme, Universität Karlsruhe
76128 Karlsruhe, Germany
mehdi_sahbi@yahoo.de

Abstract

We consider the effective construction of δ-regular coordinates for polynomial ideals. Special attention is given to quasi-stable ideals, i. e. monomial ideals possessing a Pommaret basis. Finally, we show that δ-regularity for an ideal $\mathcal{I}$ is equivalent to quasi-regularity for $\mathcal{P}/\mathcal{I}$ (in the sense of Serre).

Keywords: polynomial ideal, Pommaret basis, δ-regularity, quasi-regularity

1 Introduction

Involutive bases [3, 5, 9] are a special kind of Gröbner bases [1] with additional combinatorial properties. The underlying ideas originated in the theory of differential equations. In particular, Pommaret bases are closely related to the involution analysis of symbols in the formal theory of differential equations [11]. It is a well-known problem that a polynomial ideal $\mathcal{I}$ possesses only in suitable, so-called δ-regular, coordinates a Pommaret basis (note, however, that generic coordinates are δ-regular). In [10] it is

[1] This work has been partially supported by GIFT (NEST project no. 5006 in the Sixth Framework Programme of the European Union)

shown that coordinates regular in this sense are very useful for a number of applications; e. g. regular sequences or Noether normalisations of $\mathcal{P}/\mathcal{I}$ take a particularly simple form.

The traditional approach to obtain δ-regular coordinates consists of applying a random transformation (see e. g. [13] for a discussion in the context of differential equations). This method has at least two disadvantages. While random coordinates are δ-regular with probability 1, they still may be singular. More importantly, random transformations usually destroy any sparsity present in a basis of the ideal $\mathcal{I}$ making any subsequent computation much more expensive.

In [8] we presented a deterministic solution for the related problem of δ-regularity in partial differential equations based on a comparison of the Janet and Pommaret multiplicative variables. Here we first adapt this solution to polynomial ideals. Then we show that our criterion for singular coordinates is closely related to the algebraic theory of a class of monomial ideals studied by Bermejo and Gimenez [2]. Finally, we relate the theory of Pommaret bases to Serre's dual version of the Cartan test (see the letter by Serre appended to [7]). We prove that Serre's notion of quasi-regular coordinates for the factor ring $\mathcal{P}/\mathcal{I}$ coincides with δ-regularity for $\mathcal{I}$.

2 Involutive Bases

Identifying the Abelian monoid $(\mathbb{N}_0^n, +)$ with the set of terms x^μ in a polynomial ring $\mathcal{P} = \Bbbk[x_1, \ldots, x_n]$ over a field $\Bbbk$, we have the usual divisibility relation: $\mu|\nu$ if $\nu \in C(\mu) := \mu + \mathbb{N}_0^n$. An *involutive division* is a rule L (satisfying certain conditions, see [9] for details) restricting this relation by assigning to each member μ of every finite subset $\mathcal{N} \in \mathbb{N}_0^n$ a set $N_{L,\mathcal{N}}$ of allowed (*multiplicative*) indices, resulting in restricted *involutive* cones $C_{L,\mathcal{N}}(\mu) := \mu + \{\nu \mid \nu_i = 0 \,\text{for}\, i \notin N\}$. For this new relation, we write $\mu|_{L,\mathcal{N}}\nu$ if $\nu \in C_{L,\mathcal{N}}(\mu)$ (μ *involutively divides* ν). In this article, we only need the following two involutive divisions which we denote by P and J, respectively:

***Pommaret division*:**

$$N_{P,\mathcal{N}}(\mu) := \{i \mid i \leq \operatorname{cls}\mu\}, \text{where}\, \operatorname{cls}\mu := \min\{i \mid \mu_i \neq 0\}.^2$$

***Janet division*:**

$$N_{J,\mathcal{N}}(\mu) := \{i \mid \mu_i \geq \nu_i \,\text{for all}\, \nu \in \mathcal{N} \,\text{with}\, \mu_j = \nu_j \,\text{for}\, j > i\}.$$

The *(involutive) span* $\langle\mathcal{N}\rangle$ (resp. $\langle\mathcal{N}\rangle_L$) of $\mathcal{N}$ is the union of the (involutive) cones of its elements. $\mathcal{N}$ is *(involutively) autoreduced*, if no member is contained in the (involutive) cone of another element. A finite subset $\hat{\mathcal{N}} \subset \langle\mathcal{N}\rangle$ is a *weak involutive basis* of $\langle\mathcal{N}\rangle$, if $\langle\hat{\mathcal{N}}\rangle_L = \langle\mathcal{N}\rangle$, and a *(strong) involutive basis*, if furthermore $\hat{\mathcal{N}}$ is autoreduced. We refer to $\hat{\mathcal{N}}$ as a *(weak/strong) involutive completion* of the set $\mathcal{N}$, if $\mathcal{N} \subseteq \hat{\mathcal{N}}$. One can show that to every finite set $\mathcal{N}$ there exists a Janet basis of $\langle\mathcal{N}\rangle$, but not necessarily a Pommaret basis. By contrast, a basis minimal among all the Pommaret bases of $\mathcal{N}$ is unique, whereas the same does not hold for the Janet division.

While the definitions of the Pommaret and Janet division, respectively, look very different, the two divisions are in fact closely related, as the following result demonstrates.

Proposition 2.1 ([4])**.** *Let the finite set $\mathcal{N} \subset \mathbb{N}_0^n$ be involutively autoreduced with respect to the Pommaret division. Then $N_P(\nu) \subseteq N_{J,\mathcal{N}}(\nu)$ for all $\nu \in \mathcal{N}$.*

An involutive basis $\mathcal{N}$ leads via the involutive cones to a disjoint decomposition of the ideal $\mathcal{I} = \langle\mathcal{N}\rangle$ as a $\Bbbk$-linear space (a so-called *Stanley decomposition* [12]). For many applications it is also of interest to decompose the complement $\mathcal{I}^c := \mathbb{N}_0^n \setminus \mathcal{I}$; both Janet and Pommaret bases induce such *complementary Stanely decompositions*. In the latter case, we have the following result.

[2] This is independent of the set $\mathcal{N}$, so we will drop the reference to it in the sequel.

Proposition 2.2 ([10]). *The monoid ideal $\mathcal{I} \subseteq \mathbb{N}_0^n$ possesses a weak Pommaret basis of degree q, if and only if the sets $\bar{\mathcal{N}}_0 = \{\nu \in \mathcal{I}^c \mid |\nu| < q\}$ and $\bar{\mathcal{N}}_1 = \{\nu \in \mathcal{I}^c \mid |\nu| = q\}$ define the complementary decomposition*

$$\mathcal{I}^c = \bar{\mathcal{N}}_0 \cup \bigcup_{\nu \in \bar{\mathcal{N}}_1} C_P(\nu) \; . \tag{1}$$

The notion of an involutive basis can now be easily lifted to polynomial ideals. Choosing a term order $\prec$ determines for each $f \in \mathcal{P}$ its *leading term* $\mathrm{lt}_\prec f$ with *leading exponent vector* $\mathrm{le}_\prec f$. Let $\mathcal{F} \subset \mathcal{P}$ be a finite set. Then we assign to each element $f \in \mathcal{F}$ the multiplicative variables

$$X_{L,\mathcal{F},\prec}(f) = \{x_i \mid i \in N_{L,\mathrm{le}_\prec \mathcal{F}}(\mathrm{le}_\prec f)\} \; ; \tag{2}$$

the involutive span of $\mathcal{F}$ is then the set

$$\langle \mathcal{F} \rangle_{L,\prec} = \sum_{f \in \mathcal{F}} \Bbbk[X_{L,\mathcal{F},\prec}(f)] \cdot f \subseteq \langle \mathcal{F} \rangle \; . \tag{3}$$

A polynomial $g \in \mathcal{P}$ is *involutively reducible* with respect to $\mathcal{F}$, if it contains a term x^μ such that $\mathrm{le}_\prec f \,|_{L,\mathrm{le}_\prec \mathcal{F}}\, \mu$ for some $f \in \mathcal{F}$; g is *involutively head reducible*, if $x^\mu = \mathrm{lt}_\prec g$ in the previous definition. The set $\mathcal{F}$ is *involutively autoreduced*, if no polynomial $f \in \mathcal{F}$ contains a term x^μ such that another polynomial $f' \in \mathcal{F} \setminus \{f\}$ exists with $\mathrm{le}_\prec f' \,|_{L,\mathrm{le}_\prec \mathcal{F}}\, \mu$; the definition of *involutively head autoreduced* is similar. A finite set $\mathcal{H} \subset \mathcal{P}$ is a *weak involutive basis* of $\mathcal{I}$ for an involutive division L if $\mathrm{le}_\prec \mathcal{H}$ is a weak involutive basis of $\mathrm{le}_\prec \mathcal{I}$; it is a *(strong) involutive basis*, if $\mathrm{le}_\prec \mathcal{H}$ is a strong involutive basis of $\mathrm{le}_\prec \mathcal{I}$ and no two elements of $\mathcal{H}$ have the same leading exponents. As above, any finite set $\mathcal{F}$ can be *completed* to a Janet basis of $\langle \mathcal{F} \rangle$, while this is not necessarily true for Pommaret bases.

For the remainder of the article, all ideals $\mathcal{I}$ considered will be homogeneous. If $\mathcal{M}$ is a graded $\mathcal{P}$-module, we write $\mathcal{M}_q$ for the homogeneous component of degree q and $\mathcal{M}_{\geq q} := \bigoplus_{q' \geq q} \mathcal{M}_{q'}$ for the truncated module (similar for $\mathcal{M}_{<q}$). As usual, we call $\mathcal{I}^{\mathrm{sat}} := \mathcal{I} : \langle x_1, \ldots, x_n \rangle$ the saturation of $\mathcal{I}$.

Remark 2.3. If a Pommaret basis $\mathcal{H}$ of the ideal $\mathcal{I}$ exists, then a number of important invariants of the factor algebra $\mathcal{P}/\mathcal{I}$ can be immediately read off of $\mathcal{H}$ [10]:

- If $\deg \mathcal{H} := \max_{h\in\mathcal{H}} \deg h = q$, then the *dimension* D of the algebra $\mathcal{P}/\mathcal{I}$ is given by $D = \min\{i \mid \langle \mathcal{H}, x_1, \ldots, x_i\rangle_q = \mathcal{P}_q\}$ and $\{x_1, \ldots, x_D\}$ is a maximal independent set modulo $\mathcal{I}$ (in fact, the complementary Stanley decomposition of Proposition 2.2 yields at once the whole *Hilbert series* of $\mathcal{P}/\mathcal{I}$).
- If $\operatorname{cls}\mathcal{H} := \min_{h\in\mathcal{H}} \operatorname{cls} h = d$ (whith $\operatorname{cls} h := \operatorname{cls} \operatorname{le}_\prec h$), then $\operatorname{depth} \mathcal{P}/\mathcal{I} = d - 1$ and $(x_1, \ldots, x_{d-1})$ is a maximal regular sequence for $\mathcal{P}/\mathcal{I}$ (combined with the result above, this observation yields a simple proof of the well-known Hironaka criterion for Cohen-Macaulay rings).
- If $\prec$ is the *degree reverse lexicographic order*[3], then $\deg \mathcal{H}$ equals the *Castelnuovo-Mumford regularity* of $\mathcal{I}$ (this is a consequence of the interesting syzygy theory of Pommaret bases leading to a free resolution of minimal length).
- The isomorphism $\mathcal{P}/\mathcal{I} \cong \bigoplus_{\nu\in\bar{\mathcal{N}}_0} \Bbbk \cdot x^\nu \oplus \bigoplus_{\nu\in\bar{\mathcal{N}}_1} \Bbbk[x_1, \ldots, x_{\operatorname{cls}\nu}] \cdot x^\nu$ as $\Bbbk$-linear spaces is a *Rees decomposition* of $\mathcal{P}/\mathcal{I}$, where the sets $\bar{\mathcal{N}}_0$ and $\bar{\mathcal{N}}_1$ are defined for the Pommaret basis of $\operatorname{le}_\prec \mathcal{H}$ as in Proposition 2.2.

Together with Theorem 3.11 below, the first two items show that if a Pommaret basis exists, then the chosen coordinates are particularly adapted to the ideal $\mathcal{I}$ and considerably simplify the analysis of the algebra $\mathcal{P}/\mathcal{I}$. In the next section we will see how one can systematically construct such coordinates for any ideal $\mathcal{I}$.

3 δ-Regularity and Systems of Parameters

An ideal $\mathcal{I} \subseteq \mathcal{P}$ may be interpreted as an ideal in the symmetric algebra $S\mathcal{V}$ over an n-dimensional $\Bbbk$-linear space $\mathcal{V} \cong \mathcal{P}_1$ after having chosen a basis $(x_1, \ldots, x_n)$ of $\mathcal{V}$. As we will show now, the existence of a Pommaret basis for $\mathcal{I}$ depends only on this choice.

[3] Note that we use the ordering $x_n > \ldots > x_1$ on the variables.

Definition 3.1. The variables $\mathbf{x} = (x_1, \ldots, x_n)$ are *δ-regular* for the ideal $\mathcal{I} \subseteq \mathcal{P}$ and the term order $\prec$, if $\mathcal{I}$ possesses a Pommaret basis for $\prec$.

As in practice one defines an ideal $\mathcal{I} \subseteq \mathcal{P}$ by some finite generating set $\mathcal{F} \subset \mathcal{I}$, we introduce a concept of δ-regularity for such sets. Assume that $\mathcal{F}$ is involutively head autoreduced with respect to an involutive division L. We call the total number of multiplicative variables of its elements its *involutive size* and denote it by

$$|\mathcal{F}|_{L,\prec} = \sum_{f \in \mathcal{F}} |X_{L,\prec,\mathcal{F}}(f)| \,. \tag{4}$$

Let $\tilde{\mathbf{x}} = A\mathbf{x}$ be a linear change of coordinates with a regular matrix $A \in \mathbb{k}^{n \times n}$, i. e. a change of basis in the vector space $\mathcal{V}$. It transforms each polynomial $f \in \mathcal{P}$ into a polynomial $\tilde{f} \in \tilde{\mathcal{P}} = \mathbb{k}[\tilde{x}_1, \ldots, \tilde{x}_n]$ of the same degree. Thus $\mathcal{F}$ is transformed into a set $\tilde{\mathcal{F}} \subset \tilde{\mathcal{P}}$ which generally is no longer involutively head autoreduced.[4] Performing an involutive head autoreduction yields a set $\tilde{\mathcal{F}}^{\triangle}$. The leading exponents of $\tilde{\mathcal{F}}^{\triangle}$ may be very different from those of $\mathcal{F}$ and thus $|\mathcal{F}|_{L,\prec}$ may differ from $|\tilde{\mathcal{F}}^{\triangle}|_{L,\prec}$.

Definition 3.2. Let the finite set $\mathcal{F} \subset \mathcal{P}$ be involutively head autoreduced with respect to the Pommaret division. The coordinates $\mathbf{x}$ are *δ-regular* for $\mathcal{F}$, if after any linear change of coordinates $\tilde{\mathbf{x}} = A\mathbf{x}$ the inequality $|\mathcal{F}|_{P,\prec} \geq |\tilde{\mathcal{F}}^{\triangle}|_{P,\prec}$ holds.

Note that generally δ-regularity of variables $\mathbf{x}$ for a set $\mathcal{F}$ according to Definition 3.2 and for the ideal $\mathcal{I} = \langle \mathcal{F} \rangle$ according to Definition 3.1 are independent properties.

Example 3.3. One of the simplest instances where the definitions differ is not for an ideal but for a submodule of the free $\mathbb{k}[x, y]$-module with basis $\{\mathbf{e}_1, \mathbf{e}_2\}$. Consider the set $\mathcal{F} = \{y^2\mathbf{e}_1, xy\mathbf{e}_1 + \mathbf{e}_2, x\mathbf{e}_2\}$ and any term order for which $xy\mathbf{e}_1 \succ \mathbf{e}_2$. The used coordinates are not δ-regular for $\mathcal{F}$, as any transformation of the form $x = \bar{x} + a\bar{y}$ with $a \neq 0$ will increase the

[4] We consider here the involutive division and the term order as being defined on the exponent vectors. Thus after the transformation we can still use the same division and order as before.

involutive size. Nevertheless, the used coordinates are δ-regular for the submodule $\langle\mathcal{F}\rangle$. Indeed, adding the generator $y\mathbf{e}_2$ (the S-"polynomial" of the first two generators) makes $\mathcal{F}$ to a reduced Gröbner basis which is simultaneously a minimal Pommaret basis. Examples of this type are critical for the algorithmic determination of Pommaret bases: although a finite basis exist, some completion algorithms may loop infinitely in such a situation, as they try to construct a Pommaret basis for $\langle \mathrm{le}_\prec\mathcal{F}\rangle$ as an intermediate step.

Proposition 3.4. *Let $\mathcal{H}$ be a Pommaret basis of an ideal $\mathcal{I} \subseteq \mathcal{P}$. Then the given coordinates $\mathbf{x}$ are δ-regular for $\mathcal{H}$.*

Most coordinates are δ-regular for a given set $\mathcal{F}$. Choosing an arbitrary reference coordinate system, we may identify every system of coordinates with the regular matrix $A \in \mathbb{k}^{n\times n}$ defining the linear transformation from our reference system to it.

Proposition 3.5. *The coordinate systems that are δ-singular for a given finite involutively head autoreduced set $\mathcal{F} \subset \mathcal{P}$ form a Zariski closed set in $\mathbb{k}^{n\times n}$.*

Proof. We perform first a linear coordinate transformation with an undetermined matrix $A = (a_{ij}) \in \mathbb{k}^{n\times n}$, i. e. we treat its entries as parameters. This obviously leads to a δ-regular coordinate system, as each polynomial in $\tilde{\mathcal{F}}^{\triangle}$ will get its maximally possible class. δ-singular coordinates are defined by the vanishing of certain (leading) coefficients. Since these coefficients are polynomials in the entries a_{ij} of A, the set of all δ-singular coordinate systems can be described as the zero set of an ideal of $\mathbb{k}[a_{11}, \ldots, a_{nn}]$. □

Theorem 3.6. *Let the finite set $\mathcal{F} \subset \mathcal{P}$ be involutively head autoreduced for the Pommaret division and a class respecting term order*[5] *$\prec$. Furthermore assume that the underlying field $\mathbb{k}$ is infinite. If $|\mathcal{F}|_{J,\prec} > |\mathcal{F}|_{P,\prec}$, then the coordinates $\mathbf{x}$ are δ-singular for $\mathcal{F}$.*

[5] This means that for $\deg t_1 = \deg t_2$ and $\mathrm{cls}\, t_1 < \mathrm{cls}\, t_2$ we always have $t_1 \prec t_2$. The degree reverse lexicographic order has this property.

Proof. By Proposition 2.1, we have $X_P(f) \subseteq X_{J,\mathcal{F}}(f)$ for all $f \in \mathcal{F}$. Assume that for a polynomial $h \in \mathcal{F}$ the strict inclusion $X_P(h) \subset X_{J,\mathcal{F}}(h)$ holds. Thus at least one variable $x_\ell \in X_{J,\mathcal{F}}(h)$ with $\ell > k = \operatorname{cls} h$ exists. We perform the linear change of variables $x_i = \tilde{x}_i$ for $i \neq k$ and $x_k = \tilde{x}_k + a\tilde{x}_\ell$ with a yet arbitrary parameter $a \in \Bbbk \setminus \{0\}$. This induces the following transformation of the terms[6]:

$$x^\mu = \sum_{j=0}^{\mu_k} \binom{\mu_k}{j} a^j \tilde{x}^{\mu - j_k + j_\ell} \ . \tag{5}$$

Let $\operatorname{le}_\prec h = \mu$. Thus $\mu = [0, \ldots, 0, \mu_k, \ldots, \mu_n]$ with $\mu_k > 0$. Consider the multi index $\nu = \mu - (\mu_k)_k + (\mu_k)_\ell$; obviously, $\operatorname{cls} \nu > k$. Applying our transformation to h leads to a polynomial $\tilde{h}$ containing the term $\tilde{x}^\nu$. Note that ν cannot be an element of $\operatorname{le}_\prec \mathcal{F}$. Indeed, if it was, it would be an element of the same set $(\mu_{\ell+1}, \ldots, \mu_n)$ as μ. But this contradicts our assumption that ℓ is multiplicative for the multi index μ with respect to the Janet division, as by construction $\nu_\ell > \mu_\ell$.

Transforming all polynomials $f \in \mathcal{F}$ yields the set $\tilde{\mathcal{F}}$ on which we perform an involutive head autoreduction in order to obtain the set $\tilde{\mathcal{F}}^\triangle$. Since we assume that the gound field $\Bbbk$ is infinite, we can always choose the parameter a such that after the transformation each polynomial $\tilde{f} \in \tilde{\mathcal{F}}$ has at least the same class as the corresponding polynomial $f \in \mathcal{F}$, as our term order respects classes. This is a simple consequence of (5): cancellations of terms may occur only, if the parameter a is a zero of some polynomial (possibly one for each member of $\mathcal{F}$) with a degree not higher than $\deg \mathcal{F}$. By the definition of the Pommaret division, if $\operatorname{le}_\prec f_2 \,|_P \operatorname{le}_\prec f_1$, then $\operatorname{cls} \operatorname{le}_\prec f_2 \geq \operatorname{cls} \operatorname{le}_\prec f_1$. Hence even after the involutive head autoreduction the involutive size of $\tilde{\mathcal{F}}^\triangle$ cannot be smaller than that of $\mathcal{F}$.

Consider again the polynomial h. The leading term of the transformed polynomial $\tilde{h}$ must be greater than or equal to $\tilde{x}^\nu$. Thus its class is greater than k. This remains true even after an involutive head autoreduction with all those polynomials $\tilde{f} \in \tilde{\mathcal{F}}$ that are of class greater than k, as $x^\nu \notin \operatorname{lt}_\prec \mathcal{F}$.

[6] j_k means the multiindex with entry j at position k and zero elsewhere.

Hence the only possibility to obtain a leading term of class less than or equal to k consists of an involutive reduction with respect to a polynomial $\tilde{f} \in \tilde{\mathcal{F}}$ with $\operatorname{cls} f \leq k$. But this implies that $\operatorname{cls} \operatorname{le}_{\prec} \tilde{f} > k$. So we may conclude that after the transformation we have at least one polynomial more whose class is greater than k. So the coordinates $\mathbf{x}$ cannot be δ-regular. □

Corollary 3.7. *If the coordinates* $\mathbf{x}$ *are* δ*-regular for the finite Pommaret head autoreduced set* $\mathcal{F}$*, then* $\langle \mathcal{F} \rangle_{J,\prec} = \langle \mathcal{F} \rangle_{P,\prec}$ *for any class respecting term order* $\prec$*.*

It is important to note that this corollary provides us only with a necessary but *not* with a sufficient criterion for δ-regularity. In other words, even if the Janet and the Pommaret size are equal for a given set $\mathcal{F} \subset \mathcal{P}$, this fact does not imply that the used coordinates are δ-regular for $\mathcal{F}$.

Example 3.8. Let $\mathcal{F} = \{\underline{z^2} + y^2 - 2x^2,\ \underline{xz} + xy,\ \underline{yz} + y^2 + x^2\}$. The underlined terms are the leaders with respect to the degree reverse lexicographic order. One easily checks that the Janet and the Pommaret division yield the same multiplicative variables. If we perform the transformation $\tilde{x} = z$, $\tilde{y} = y + z$ and $\tilde{z} = x$, then we obtain after an autoreduction the set $\tilde{\mathcal{F}}^{\triangle} = \{\underline{\tilde{y}^2},\ \underline{\tilde{y}\tilde{z}},\ \underline{\tilde{z}^2} - \tilde{y}\tilde{x}\}$. Again the Janet and the Pommaret division yield the same multiplicative variables, but $|\tilde{\mathcal{F}}^{\triangle}|_{P,\prec} > |\mathcal{F}|_{P,\prec}$. Thus the coordinates (x, y, z) are not δ-regular for $\mathcal{F}$.

The explanation of this phenomenon is very simple. Our criterion depends only on the leading terms of the set $\mathcal{F}$; in other words, it analyses the monomial ideal $\langle \operatorname{lt}_{\prec} \mathcal{F} \rangle$. In Example 3.8 $\langle \operatorname{lt}_{\prec} \mathcal{F} \rangle = \langle xz, yz, z^2 \rangle$ and one easily verifies that the used generating set is already a Pommaret basis. However, for $\mathcal{I} = \langle \mathcal{F} \rangle$ the leading ideal is $\operatorname{lt}_{\prec} \mathcal{I} = \langle x^3, xz, yz, z^2 \rangle$ (one obtains a Janet basis for $\mathcal{I}$ by adding x^3 to $\mathcal{F}$) and obviously it does not possess a Pommaret basis, as such a basis would have to contain all monomials $x^3 y^k$ with $k \in \mathbb{N}$ (or we exploit our criterion noting that y is a Janet but not a Pommaret multiplicative variable for x^3). We have here just the opposite situation to Example 3.3: there $\operatorname{lt}_{\prec} \mathcal{I}$ had a Pommaret basis but $\langle \operatorname{lt}_{\prec} \mathcal{F} \rangle$ not.

Example 3.9. Even in the case that $\langle \mathrm{lt}_\prec \mathcal{F} \rangle = \langle \mathrm{lt}_\prec \mathcal{I} \rangle$ (i. e., $\mathcal{F}$ is a Gröbner basis), the notions of δ-regularity for a set $\mathcal{F}$ and for an ideal $\mathcal{I}$ generated by it are not equivalent. Consider the monomial ideal spanned by $\{x^2, y^2, z^2\}$. A Pommaret basis is obtained by adding $\{y^2z, x^2y, x^2z, x^2yz\}$ and thus the coordinates are δ-regular for the ideal. But for the subset $\{z^2, y^2, x^2, x^2z, x^2y\}$, the Janet and the Pommaret sizes differ (9 vs. 8), so the coordinates are not δ-regular for this set. This shows that in general for an ideal and δ-regular coordinates there may exist ideal bases for which the coordinates are not δ-regular. Whether or not such a situation is encountered during the completion to an involutive basis cannot be recognised a priori and depends also for instance on the choice of the selection term order in the completion algorithm.

Theorem 3.10. *In suitably chosen coordinates* $\mathbf{x}$ *every ideal* $\mathcal{I} \subseteq \mathcal{P}$ *has a Pommaret basis.*

Proof. We only sketch a proof here, as a rigorous demonstration requires the detailed formulation of a completion algorithm for the construction of an involutive basis ([5, 6]), which we omit here for lack of space.

Each iteration of the completion algorithm consists of selecting an element from the current basis, multiplying it with one of its non-multiplicative variables, performing an involutive reduction and then adding the result (if it is different from zero) to the basis. Only if this action has enlarged the ideal spanned by the leading terms, a test for δ-regularity and, if necessary, a coordinate change according to the proof of Theorem 3.6 are carried out. By regarding all the bases in a fixed reference coordinate system, the ideals spanned by the leading terms form an ascending chain that eventually becomes stationary at $\mathrm{lt}_\prec \mathcal{I}$. On the other hand, if the ideal spanned by the leading terms remains unchanged (which it does especially after we have reached $\mathrm{lt}_\prec \mathcal{I}$), we have performed one step in the monomial completion of this ideal, which has to terminate under the assumption that, after suitable transformations, we are in δ-regular coordinates. □

The search for δ-regular coordinates corresponds to putting $\mathcal{I}$ in Noether position, as the following result shows:

Theorem 3.11. *Under the assumptions of Remark 2.3 (in particular, if the coordinates* $\mathbf{x}$ *are* δ*-regular for* $\mathcal{I}$*), the restriction of the canonical projection* $\pi : \mathcal{P} \to \mathcal{P}/\mathcal{I}$ *to* $\Bbbk[x_1, \ldots, x_D]$ *is a Noether normalisation of* $\mathcal{P}/\mathcal{I}$ *(or equivalently,* $x_1, \ldots, x_D$ *form a homogeneous system of parameters).*

Proof. $\{x_1, \ldots, x_D\}$ is a maximal independent set modulo $\mathcal{I}$, so the restriction of the projection π to $\Bbbk[x_1, \ldots, x_D]$ is injective. Proposition 2.2 gives the complementary decomposition for $\mathrm{le}_\prec\mathcal{I}$ which is defined by a finite set $\mathcal{N} \subset \mathbb{N}_0^n$. As for each generator in $\mathcal{N}$ the associated multiplicative indices form a subset of $\{1, \ldots, D\}$ and since the complement of $\mathrm{lt}_\prec\mathcal{I}$ is a basis of $\mathcal{P}/\mathcal{I}$ as a vector space over $\Bbbk$, the finite set $\{\pi(x^\nu) \mid \nu \in \mathcal{N}\}$ generates $\mathcal{P}/\mathcal{I}$ as a $\Bbbk[x_1, \ldots, x_D]$-module. □

The converse of this theorem is in general not true: even if the variables are chosen such that $\Bbbk[x_1, \ldots, x_D]$ defines a Noether normalisation of $\mathcal{P}/\mathcal{I}$, this is not sufficient to conclude that $\mathcal{I}$ possesses a Pommaret basis. In the next section, we will show (in the case that $\mathcal{I}$ is a monomial ideal) that the existence of a Pommaret basis is equivalent to a stronger property.

4 Quasi-stable Ideals

While a polynomial ideal $\mathcal{I}$ always possesses a Pommaret basis (after a suitable coordinate transformation), the same is not true for a monomial ideal, as the transformed ideal is in general no longer monomial. Hence we give the class of monomial ideal possessing a Pommaret basis a special name.

Definition 4.1. A monomial ideal $\mathcal{I} \subseteq \mathcal{P}$ is *quasi-stable*, if it has a Pommaret basis.

Remark 4.2. Recall that a (possibly infinite) set $\mathcal{N} \subseteq \mathbb{N}_0^n$ is called *stable*, if for each multi index $\nu \in \mathcal{N}$ all multi indices $\nu - 1_k + 1_j$ with $k = \mathrm{cls}\,\nu < j \leq n$ are also contained in $\mathcal{N}$. A monomial ideal $\mathcal{I} \subseteq \mathcal{P}$ is stable, if the exponent vectors of the monomials contained in it form a stable set. If $\mathcal{I}$ is a quasi-stable ideal in the sense of the definition above, then one can

easily show that for a sufficiently high degree q the truncated ideal $\mathcal{I}_{\geq q}$ is stable (one may e. g. take $q = \deg \mathcal{H}$ with $\mathcal{H}$ the Pommaret basis of $\mathcal{I}$).

Proposition 4.3. *Let $\mathcal{I} \subseteq \mathcal{P}$ be a monomial ideal with $\dim \mathcal{P}/\mathcal{I} = D$. Then the following five statements are equivalent.*

(i) *$\mathcal{I}$ is quasi-stable.*
(ii) *The variable x_1 is not a zero divisor for $\mathcal{P}/\mathcal{I}^{\mathrm{sat}}$ and for all $1 \leq j < D$ the variable x_{j+1} is not a zero divisor for $\mathcal{P}/\langle \mathcal{I}, x_1, \ldots, x_j \rangle^{\mathrm{sat}}$.*
(iii) *We have $\mathcal{I} : \langle x_1 \rangle^{\infty} \subseteq \mathcal{I} : \langle x_2 \rangle^{\infty} \subseteq \cdots \subseteq \mathcal{I} : \langle x_D \rangle^{\infty}$ and for all $D < j \leq n$ an exponent $k_j \geq 1$ exists such that $x_j^{k_j} \in \mathcal{I}$.*
(iv) *For all $1 \leq j \leq n$ the equality $\mathcal{I} : \langle x_j \rangle^{\infty} = \mathcal{I} : \langle x_j, \ldots, x_n \rangle^{\infty}$ holds.*
(v) *For every associated prime ideal $\mathfrak{p} \in \mathrm{Ass}\,(\mathcal{P}/\mathcal{I})$ an integer $1 \leq j \leq n$ exists such that $\mathfrak{p} = \langle x_j, \ldots, x_n \rangle$.*

Proof. The equivalence of the statements (ii)–(v) was proven by Bermejo and Gimenez [2, Proposition 3.2] who called ideals satisfying one of these conditions *monomial ideals of nested type*.[7] We now show that this concept is identical with quasi-stability by proving the equivalence of (i) and (iii).

Assume first that the ideal $\mathcal{I}$ is quasi-stable with Pommaret basis $\mathcal{H}$. The existence of a term $x_j^{k_j} \in \mathcal{I}$ for all $D < j \leq n$ follows then immediately from Remark 2.3. Consider a term $x^{\mu} \in \mathcal{I} : \langle x_k \rangle^{\infty} \setminus \mathcal{I}$ for some $1 \leq k \leq n$. There exists an integer ℓ such that $x_k^{\ell} x^{\mu} \in \mathcal{I}$ and hence a generator $x^{\nu} \in \mathcal{H}$ such that $x^{\nu} \,|_P\, x_k^{\ell} x^{\mu}$. If $\mathrm{cls}\,\nu > k$, then ν would also be an involutive divisor of μ contradicting the assumption $x^{\mu} \notin \mathcal{I}$. Thus we find $\mathrm{cls}\,\nu \leq k$ and $\nu_k > \mu_k$. Next we consider for arbitrary exponents $m > 0$ the terms $x_{k+1}^m x^{\nu} \in \mathcal{I}$. For each m a generator $x^{\rho^{(m)}} \in \mathcal{H}$ exists which involutively divides $x_{k+1}^m x^{\nu}$. By the same reasoning as above, $\mathrm{cls}\, x^{\rho^{(m)}} > k+1$ is not possible for an involutively autoreduced basis $\mathcal{H}$ yielding the estimate $\mathrm{cls}\,\nu \leq \mathrm{cls}\, x^{\rho^{(m)}} \leq k+1$.

We claim now that there exists an integer m_0 such that $\rho^{(m)} = \rho^{(m_0)}$ for all $m \geq m_0$ and $\mathrm{cls}\, x^{\rho^{(m_0)}} = k+1$. Indeed, if $\mathrm{cls}\, x^{\rho^{(m)}} < k+1$, then we must have $\rho_{k+1}^{(m)} = v_{k+1} + m$, since x_{k+1} is not multiplicative for $x^{\rho^{(m)}}$. Hence $x^{\rho^{(m)}}$

[7] One must revert the ordering of the variables in order to recover the statements in [2].

cannot be an involutive divisor of $x_{k+1}^{m+1}x^\nu$ and $\rho^{(m+1)} \notin \{\rho^{(1)}, \ldots, \rho^{(m)}\}$. As the Pommaret basis $\mathcal{H}$ is a finite set, $\operatorname{cls} x^{\rho^{(m_0)}} = k+1$ for some value $m_0 > 0$. But then x_{k+1} is multiplicative for $x^{\rho^{(m_0)}}$ and thus $x^{\rho^{(m_0)}}$ is trivially an involutive divisor of $x_{k+1}^m x^\nu$ for all values $m \geq m_0$.

By construction, the generator $x^{\rho^{(m_0)}}$ is also an involutive divisor of $x_{k+1}^{m_0}x^\mu$, as x_k is multiplicative for it. Hence this term must lie in $\mathcal{I}$ and consequently $x^\mu \in \mathcal{I} : \langle x_{k+1}\rangle^\infty$. Thus we can conclude that $\mathcal{I} : \langle x_k\rangle^\infty \subseteq \mathcal{I} : \langle x_{k+1}\rangle^\infty$. This proves (iii).

For the converse assume that (iii) holds and let $\mathcal{B}$ be the minimal basis of the ideal $\mathcal{I}$. Let $x^\mu \in \mathcal{B}$ be an arbitrary term of class k. Then $x^\mu/x_k \in \mathcal{I} : \langle x_k\rangle^\infty$. By assumption, this means that also $x^\mu/x_k \in \mathcal{I} : \langle x_\ell\rangle^\infty$ for any nonmultiplicative index ℓ. Hence for each term $x^\mu \in \mathcal{B}$ and for each value $\operatorname{cls}(x^\mu) < \ell \leq n$ there exists an integer $q_{\mu,\ell}$ such that $x_\ell^{q_{\mu,\ell}}x^\mu/x_k \notin \mathcal{I}$ but $x_\ell^{q_{\mu,\ell}+1}x^\mu/x_k \in \mathcal{I}$. For the values $1 \leq \ell \leq \operatorname{cls} x^\mu$ we set $q_{\mu,\ell} = 0$. Observe that if $x^\nu \in \mathcal{B}$ is a minimal generator dividing $x_\ell^{q_{\mu,\ell}+1}x^\mu/x_k$, then $x^\nu \prec_{\text{invlex}} x^\mu$, since $\operatorname{cls}(x^\nu) \geq \operatorname{cls}(x^\mu)$ and $\nu_k < \mu_k$.

Consider now the set

$$\mathcal{H} = \left\{ x^{\mu+\rho} \mid x^\mu \in \mathcal{B},\ \forall 1 \leq \ell \leq n : 0 \leq \rho_\ell \leq q_{\mu,\ell} \right\}. \tag{6}$$

We claim that it is a weak involutive completion of $\mathcal{B}$ and thus a weak Pommaret basis of $\mathcal{I}$. In order to prove this assertion, we must show that each term $x^\lambda \in \mathcal{I}$ lies in the involutive cone of a member of $\mathcal{H}$.

As x^λ is assumed to be an element of $\mathcal{I}$, we can factor it as $x^\lambda = x^{\sigma^{(1)}}x^{\rho^{(1)}}x^{\mu^{(1)}}$ where $x^{\mu^{(1)}} \in \mathcal{B}$ is a minimal generator, $x^{\sigma^{(1)}}$ contains only multiplicative variables for $x^{\mu^{(1)}}$ and $x^{\rho^{(1)}}$ only non-multiplicative ones. If $x^{\mu^{(1)}+\rho^{(1)}} \in \mathcal{H}$, then we are done, as obviously $\operatorname{cls}\left(x^{\mu^{(1)}+\rho^{(1)}}\right) = \operatorname{cls}\left(x^{\mu^{(1)}}\right)$ and hence all variables contained in $x^{\sigma^{(1)}}$ are multiplicative for $x^{\mu^{(1)}+\rho^{(1)}}$, too.

Otherwise there exists a non-multiplicative variables x_ℓ such that $\rho_\ell^{(1)} > q_{\mu^{(1)},\ell}$. Any minimal generator $x^{\mu^{(2)}} \in \mathcal{H}$ dividing $x_\ell^{q_{\mu^{(1)},\ell}+1}x^{\mu^{(1)}}/x_k$ is also a divisor of x^λ and we find a second factorisation $x^\lambda = x^{\sigma^{(2)}}x^{\rho^{(2)}}x^{\mu^{(2)}}$ where again $x^{\sigma^{(2)}}$ consists only of multiplicative and $x^{\rho^{(2)}}$ only of non-multiplicative

variables for $x^{\mu^{(2)}}$. If $x^{\mu^{(2)}+\rho^{(2)}} \in \mathcal{H}$, then we are done by the same argument as above; otherwise we iterate.

According to the observation made above, the sequence $x^{\mu^{(1)}}, x^{\mu^{(2)}}, \ldots$ of minimal generators constructed this way is strictly descending with respect to the inverse lexicographic order. However, the minimal basis $\mathcal{B}$ is a finite set and thus the iteration cannot go on infinitely. As the iteration only stops, if there exists an involutive cone containing x^λ, the involutive span of $\mathcal{H}$ is indeed $\mathcal{I}$ and thus $\mathcal{I}$ quasi-stable. □

Note that we actually proved that (iii) may be replaced by the equivalent statement $\mathcal{I} : \langle x_1 \rangle^\infty \subseteq \mathcal{I} : \langle x_2 \rangle^\infty \subseteq \cdots \subseteq \mathcal{I} : \langle x_n \rangle^\infty$ requiring no a priori knowledge of D (the dimension D arises then obviously as the smallest value k such that $\mathcal{I} : \langle x_k \rangle^\infty = \mathcal{P}$). In this formulation it is straightforward to verify (iii) effectively: bases of the colon ideals $\mathcal{I} : \langle x_k \rangle^\infty$ are obtained by setting $x_k = 1$ in a basis of $\mathcal{I}$ and for monomial ideals it is trivial to check inclusion.

For the sequel, we make the convention that the formal expression $\mathcal{I} : x_0^\infty$ equals $\mathcal{I}$. The following technical results will be useful later.

Lemma 4.4. *For a quasi-stable ideal $\mathcal{I}$ and for all $0 \leq i \leq n$, we have:*

(i) *$\mathcal{I} : \langle x_i \rangle^\infty$ can be minimally generated by elements having class at least $i + 1$.*

(ii) *$(\mathcal{I} : \langle x_i \rangle^\infty) : \langle x_j \rangle^\infty = \mathcal{I} : \langle x_j \rangle^\infty$ for all $i < j \leq n$.*

Proof. No element of a minimal basis of $\mathcal{H}$ of $\mathcal{I} : \langle x \rangle_i^\infty$ can depend on x_j. Now assume that $x^\nu \in \mathcal{H}$ satisfies $\operatorname{cls} \nu = \ell < k$. Then $x_j^m x^\nu$ is a minimal generator of $\mathcal{I}$ for some suitable exponent $m \in \mathbb{N}_0$. This in turn implies that $x_j^m x^\nu / x_\ell^{\nu_\ell} \in \mathcal{I} : \langle x_\ell \rangle^\infty \subseteq \mathcal{I} : \langle x_j \rangle^\infty$ and hence $x^\nu / x_\ell^{\nu_\ell} \in \mathcal{I} : \langle x_j \rangle^\infty$ which contradicts our assumption that x^ν was a minimal generator. This proves Part (i); Part (ii) follows directly from the definition of the saturation and Proposition 4.3 (iii). □

From the next proposition it follows that for a *monomial* set $\mathcal{H}$, equality of the Pommaret and the Janet size entails quasi-stability of the ideal $\langle \mathcal{H} \rangle$;

thus in this case a converse to Theorem 3.6 can be obtained.

Proposition 4.5. *Let $\mathcal{I} \subseteq \mathcal{P}$ be a monomial ideal and $\mathcal{H}$ a finite, Pommaret autoreduced monomial basis of it. If I is not quasi-stable, then $|\mathcal{H}|_J > |\mathcal{H}|_P$, i. e. for some generator in $\mathcal{H}$ a variable exists which is Janet but not Pommaret multiplicative.*

Proof. By Proposition 2.1 we have $|\mathcal{H}|_J \geq |\mathcal{H}|_P$. As $\mathcal{I}$ is not quasi-stable, there exists a minimal value k such that $\mathcal{I} : \langle x_k \rangle^\infty \nsubseteq \mathcal{I} : \langle x_{k+1} \rangle^\infty$. Let x^μ be a minimal generator of $\mathcal{I} : \langle x_k \rangle^\infty$ which is not contained in $\mathcal{I} : \langle x_{k+1} \rangle^\infty$. Then for a suitable exponent $m \in \mathbb{N}_0$ the term $x^{\bar{\mu}} = x_k^m x^\mu$ is a minimal generator of $\mathcal{I}$ and hence contained in $\mathcal{H}$.

We claim now that $\mathcal{H}$ contains a generator for which x_{k+1} is Janet but not Pommaret multiplicative. If $x_{k+1} \in X_{J,\mathcal{H}}(x^{\bar{\mu}})$, then we are done, since according to Lemma 4.4 (i) $\operatorname{cls} \bar{\mu} = k$ and hence $x_{k+1} \notin X_P(x^{\bar{\mu}})$. Otherwise $\mathcal{H}$ contains a term x^ν such that $\nu_\ell = \mu_\ell$ for $k+1 < \ell \leq n$ and $\nu_{k+1} > \mu_{k+1}$. If several generators with this property exist in $\mathcal{H}$, we choose one for which ν_{k+1} takes a maximal value so that we have $x_{k+1} \in X_{J,\mathcal{H}}(x^\nu)$ by definition of the Janet division. If $\operatorname{cls} \nu < k+1$, we are again done, since in this case $x_{k+1} \notin X_P(x^\nu)$. Finally assume that $\operatorname{cls} \nu = k+1$ and consider the term $x^\rho = x^\nu / x_{k+1}^{\nu_{k+1}}$. Obviously, $x^\rho \in \mathcal{I} : \langle x_{k+1} \rangle^\infty$ contradicting our assumption $x^\mu \notin \mathcal{I} : \langle x_{k+1} \rangle^\infty$ since $x^\rho \mid x^\mu$. Hence this case cannot arise. □

From Condition (v) in Proposition 4.3 we see that (modulo a permutation of the variables) quasi-stable ideals are precisely those monomial ideals with a single minimal associated prime ideal. Going one step further, one can even read off a primary decomposition from the ascending chain of ideals in Condition (iii).

As above, let D denote the dimension of $\mathcal{I}$. We restrict ourselves to the case that x_1 (and hence each variable) occurs in some minimal generator of $\mathcal{I}$; otherwise, we change $\mathcal{P}$ accordingly. $\mathcal{I}$ contains pure powers of exactly the variables $x_{D+1}, \ldots, x_n$; thus $\mathcal{I} : \langle x_D \rangle^\infty$ is $\langle x_{D+1}, \ldots, x_n \rangle$-primary. For $1 \leq i \leq D$, let $s_i := \min\{s \mid \mathcal{I} : \langle x_i \rangle^s = \mathcal{I} : \langle x_i \rangle^{s+1}\}$; this is just the the maximal x_i-degree of a minimal generating set of $\mathcal{I}$. Furthermore, let $\mathcal{S}_i :=$

$\mathcal{I}+\langle x_{i+1}^{s_{i+1}},\ldots,x_D^{s_D}\rangle$ and $\mathfrak{q}_i := \mathcal{S}_i : \langle x_i\rangle^\infty = \mathcal{I} : \langle x_i\rangle^\infty + \langle x_{i+1}^{s_{i+1}},\ldots,x_D^{s_D}\rangle$ for $0 \leq i \leq D$. Because of Lemma 4.4, $\mathfrak{q}_i$ is $\langle x_{i+1},\ldots,x_n\rangle$-primary. By repeatedly applying the well-known identity $\mathcal{I} = \mathcal{I} + \langle x^s\rangle \cap \mathcal{I} : \langle x\rangle^s$ (provided that $\mathcal{I} : \langle x\rangle^s = \mathcal{I} : \langle x\rangle^\infty$) and Lemma 4.4 (ii), we can decompose each ideal $\mathcal{I} : \langle x_i\rangle^\infty$ $(0 \leq i \leq D)$ as:

$$\begin{aligned}
\mathcal{I} : \langle x_i\rangle^\infty =& \big(\mathcal{I} : \langle x_i\rangle^\infty + \langle x_{i+1}^{s_{i+1}}\rangle\big) \cap \underbrace{(\mathcal{I} : \langle x_i\rangle^\infty) : \langle x_{i+1}\rangle^\infty}_{=\mathcal{I}:\langle x_{i+1}\rangle^\infty} \\
=& \big(\mathcal{I} : \langle x_i\rangle^\infty + \langle x_{i+1}^{s_{i+1}}, x_{i+2}^{s_{i+2}}\rangle\big) \cap \underbrace{\big(\mathcal{I} : \langle x_i\rangle^\infty + \langle x_{i+1}^{s_{i+1}}\rangle\big) : \langle x_{i+2}\rangle^\infty}_{=\mathcal{I}:\langle x_{i+2}\rangle^\infty + \langle x_{i+1}^{s_{i+1}}\rangle} \\
& \cap \mathcal{I} : \langle x_{i+1}\rangle^\infty \\
& \ldots \\
=& \mathfrak{q}_i \cap \big(\mathcal{I} : \langle x_D\rangle^\infty + \langle x_{i+1}^{s_{i+1}},\ldots,x_{D-1}^{s_{D-1}}\rangle\big) \cap \ldots \cap \mathcal{I} : \langle x_{i+1}\rangle^\infty
\end{aligned} \tag{7}$$

Because of the quasi-stability of $\mathcal{I}$, the last ideal in this decomposition is always contained in all the preceding ones except $\mathfrak{q}_i$, so these can be dropped. Since $\mathfrak{q}_D = \mathcal{I} : \langle x_D\rangle^\infty$, we thus get a primary decomposition $\mathcal{I} = \bigcap_{i=0}^D \mathfrak{q}_i$, where the radicals of the primary components are pairwise different.

Proposition 4.6. *In the decomposition $\mathcal{I} = \bigcap_{i=0}^D \mathfrak{q}_i$, the primary component $\mathfrak{q}_j$ is redundant if and only if $\mathcal{I} : \langle x_k\rangle^\infty = \mathcal{I} : \langle x_{k+1}\rangle^\infty$.*

Proof. The above construction immediately yields $\mathcal{I} : x_k^\infty = \bigcap_{i=k}^D \mathfrak{q}_i$. An elementary computation involving sums and intersection of ideals furthermore shows that $\mathcal{S}_k = \bigcap_{i=0}^k \mathfrak{q}_i$. Therefore, $\mathcal{I} = \mathcal{S}_{k-1} \cap \mathcal{I} : x_k^\infty$ (we set $\mathcal{S}_{-1} = \mathcal{P}$). From that, we see at once that $\mathfrak{q}_k$ is redundant if $\mathcal{I} : x_k^\infty = \mathcal{I} : x_{k+1}^\infty$. For the other direction, assume that $\mathcal{I} : x_k^\infty \subsetneq \mathcal{I} : x_{k+1}^\infty$. For $k = 0$, this immediately implies that $\mathfrak{q}_0$ cannot be redundant. For $k > 0$, take a minimal generator m of $\mathcal{I} : \langle x_{k+1}\rangle^\infty$ (having class at least $k+2$) which is not in $\mathcal{I} : \langle x_k\rangle^\infty$ and consider the monomial $x_k^{s_k} m$. It is obviously contained both in $\mathcal{S}_{k-1}$ and in $\mathcal{I} : \langle x_{k+1}\rangle^\infty$, but not in $\mathcal{I} : \langle x_k\rangle$ and thus also not in $\mathfrak{q}_k$ (since from $\mathcal{I} : \langle x_k\rangle^\infty = \mathfrak{q}_k \cap \mathcal{I} : \langle x_{k+1}\rangle^\infty$ we have that $\mathfrak{q}_k \cap (\mathcal{I} : \langle x_{k+1}\rangle^\infty \setminus \mathcal{I} : \langle x_k\rangle^\infty) = \varnothing$); therefore, $\mathfrak{q}_k \not\supseteq \mathfrak{q}_0 \cap \ldots \cap \mathfrak{q}_{k-1} \cap \mathfrak{q}_{k+1} \cap \ldots \cap \mathfrak{q}_D$. □

The following result on the Nother normalisation of a quasi-stable ideal is also due to Bermejo and Gimenez [2, Proposition 3.6].

Corollary 4.7. *Let $\mathcal{I} \subseteq \mathcal{P}$ be a monomial ideal with $\dim \mathcal{P}/\mathcal{I} = D$. Furthermore, let $\mathcal{I} = \mathfrak{q}_1 \cap \cdots \cap \mathfrak{q}_r$ be an irredundant monomial primary decomposition with $D_j = \dim \mathcal{P}/\mathfrak{q}_j$ for $1 \leq j \leq r$. The ideal $\mathcal{I}$ is quasi-stable, if and only if $\Bbbk[x_1, \ldots, x_D]$ defines a Noether normalisation of $\mathcal{P}/\mathcal{I}$ and $\Bbbk[x_1, \ldots, x_{D_j}]$ one of $\mathcal{P}/\mathfrak{q}_j$ for each primary component $\mathfrak{q}_j$.*

Proof. By assumption, each ideal $\mathfrak{q}_j$ is a monomial primary ideal. This implies that $\Bbbk[x_1, \ldots, x_{D_j}]$ defines a Noether normalisation of $\mathcal{P}/\mathfrak{q}_j$, if and only if the associated prime ideal is $\sqrt{\mathfrak{q}_j} = \langle x_{D_j+1}, \ldots, x_n \rangle$. Now the assertion follows from Condition (v) in Proposition 4.3. □

5 δ-Regularity vs. Quasi-Regularity

Definition 5.1 ([7])**.** A linear form $v = a_1 x_1 + \ldots a_n x_n \in \mathcal{P}_1$ is called *quasi-regular* at degree q for the $\mathcal{P}$-module $\mathcal{M}$, if $v \cdot m = 0$ entails $m \in \mathcal{M}_{<q}$. A finite sequence $(v_1, \ldots, v_k)$ of linear forms in $\mathcal{P}$ is *quasi-regular* at degree q for $\mathcal{M}$, if each v_i is quasi-regular at degree q for the factor module $\mathcal{M}/\langle v_1, \ldots, v_{i-1} \rangle \mathcal{M}$.

This generalisation of the classical notion of a regular sequence appears in the dual formulation of Cartan's test for an involutive polynomial module due to Serre (see his letter appended to [7]). Recall that a polynomial module $\mathcal{M}$ is involutive at a degree q_0, if no minimal generator of the Koszul homology $H_\bullet(\mathcal{M})$ has a symmetric degree greater than or equal to q_0.

Theorem 5.2 (Dual Cartan Test [7])**.** *Let $\mathcal{M}$ be a polynomial module finitely generated in degree less than $q > 0$. The module $\mathcal{M}$ is involutive at degree q, if and only if for generic coordinates $\{x_1, \ldots, x_n\}$ the maps*

$$\mu_k : \mathcal{M}_r/\langle x_1, \ldots, x_{k-1} \rangle \mathcal{M}_{r-1} \longrightarrow \mathcal{M}_{r+1}/\langle x_1, \ldots, x_{k-1} \rangle \mathcal{M}_r \qquad (8)$$

induced by the multiplication with x_k are injective for all $r \geq q$ and $1 \leq k \leq n$, i. e. if and only if $(x_1, \ldots, x_n)$ is a quasi-regular sequence at degree q for $\mathcal{M}$.

The goal of this section is to show that the notion of δ-regularity for an ideal $\mathcal{I}$ as discussed in Section 3 is equivalent to the above introduced concept of quasi-regularity for the polynomial module $\mathcal{P}/\mathcal{I}$.

Lemma 5.3. *Let $\mathcal{I} \subseteq \mathcal{P}$ be a homogeneous ideal and $\prec$ the degree reverse lexicographic order. The sequence $(x_1, \ldots, x_n)$ is quasi-regular at degree q for the module $\mathcal{M} = \mathcal{P}/\mathcal{I}$, if and only if it is quasi-regular at degree q for $\mathcal{M}' = \mathcal{P}/\mathrm{lt}_\prec \mathcal{I}$.*

Proof. Let $\mathcal{G}$ be a Gröbner basis of $\mathcal{I}$ for $\prec$, so that the normal form with respect to $\mathcal{G}$ defines an isomorphism between the vector spaces $\mathcal{M}$ and $\mathcal{M}'$. One direction is trivial, as an obvious necessary condition for $m = [f] \in \mathcal{M}$ to satisfy $x_1 \cdot m = 0$ is that $x_1 \cdot [\mathrm{lt}_\prec f] = 0$ in $\mathcal{M}'$. Hence quasi-regularity of x_1 for $\mathcal{M}'$ implies quasi-regularity of x_1 for $\mathcal{M}$ and by iteration the same holds true for the whole sequence.

For the converse let $r \geq q$ be an arbitrary degree. We may choose for the vector space $\mathcal{M}_r$ a basis where each member is represented by a monomial, i. e. the representatives simultaneously induce a basis of $\mathcal{M}'_r$. Let x^μ be one of these monomials. As x_1 is quasi-regular for $\mathcal{M}$, we have $x_1 \cdot [x^\mu] \neq 0$ in $\mathcal{M}$. Suppose that $x_1 \cdot [x^\mu] = 0$ in $\mathcal{M}'$ so that x_1 is not quasi-regular for $\mathcal{M}'$. Thus $x^{\mu+1_1} \in \mathrm{lt}_\prec \mathcal{I}$ and $\mathcal{G}$ contains a polynomial g with $\mathrm{lt}_\prec g \mid x^{\mu+1_1}$. Because of the assumption $x^\mu \notin \mathrm{lt}_\prec \mathcal{I}$, we must have $\mathrm{cls}(\mathrm{lt}_\prec g) = 1$. By definition of the reverse lexicographic order, this implies that every term in g is of class 1. Iteration of this argument shows that the normal form of $x^{\mu+1_1}$ with respect to $\mathcal{G}$ is divisible by x_1, i. e. it can be written as $x_1 f$ with $f \in \mathcal{P}_r$ and $\mathrm{lt}_\prec f \prec x^\mu$. Consider now the polynomial $\bar{f} = x^\mu - f \in \mathcal{P}_r \setminus \{0\}$. As it consists entirely of terms not contained in $\mathrm{lt}_\prec \mathcal{I}$, we have $[\bar{f}] \neq 0$ in $\mathcal{M}_r$. However, by construction $x_1 \cdot [\bar{f}] = 0$ contradicting the injectivity of multiplication by x_1 on $\mathcal{M}_r$.

For the remaining elements of the sequence $(x_1, \ldots, x_n)$ we note for each $1 \leq k < n$ the isomorphism $\mathcal{M}^{(k)} = \mathcal{M}/\langle x_1, \ldots, x_k \rangle \mathcal{M} \cong \mathcal{P}^{(k)}/\mathcal{I}^{(k)}$ where $\mathcal{P}^{(k)} = \Bbbk[x_{k+1}, \ldots, x_n]$ and $\mathcal{I}^{(k)} = \mathcal{I} \cap \mathcal{P}^{(k)}$. It implies that we may iterate the arguments above so that indeed quasi-regularity of $(x_1, \ldots, x_n)$ for $\mathcal{M}'$ is equivalent to quasi-regularity for $\mathcal{M}'$. □

Theorem 5.4. *The coordinates* $\mathbf{x}$ *are* δ*-regular for the homogeneous ideal* $\mathcal{I} \subseteq \mathcal{P}$ *in the sense that* $\mathcal{I}$ *possesses a Pommaret basis* $\mathcal{H}$ *for the degree reverse lexicographic term order with* $\deg \mathcal{H} = q$*, if and only if the sequence* $(x_1, \ldots, x_n)$ *is quasi-regular for the factor algebra* $\mathcal{M} = \mathcal{P}/\mathcal{I}$ *at degree* q *but not at any lower degree.*

Proof. By the definition of a Pommaret basis and by Lemma 5.3, it suffices to consider monomial ideals $\mathcal{I}$. Assume first that the basis $\{x_1, \ldots, x_n\}$ is δ-regular. By Proposition 2.2, the leading terms $\mathrm{lt}_{\prec}\mathcal{H}$ induce a complementary decomposition of $\mathcal{M}$ where all generators are of degree $q = \deg \mathcal{H}$ or less. Thus, if $\mathcal{M}_q \neq 0$ (otherwise there is nothing to show), then we can choose a vector space basis of it as part of the complementary decomposition and the variable x_1 is multiplicative for all its members. But this observation immediately implies that multiplication with x_1 is injective from degree q on, so that x_1 is quasi-regular for $\mathcal{M}$ at degree q.

For the remaining elements of the basis $\{x_1, \ldots, x_n\}$ we proceed as in the proof of Lemma 5.3 and use the isomorphism $\mathcal{M}^{(k)} \cong \mathcal{P}^{(k)}/\mathcal{I}^{(k)}$. A Pommaret basis of $\mathcal{I}^{(k)}$ is obtained by setting $x_1 = \cdots = x_k = 0$ in the subset $\mathcal{H}^{(k)} = \{h \in \mathcal{H} \mid \mathrm{cls}\, h > k\}$. Thus we can again iterate for each $1 < k \leq n$ the argument above so that indeed $(x_1, \ldots, x_n)$ is a quasi-regular sequence for $\mathcal{M}$ at degree q.

For the converse, we first show that quasi-regularity of $(x_1, \ldots, x_n)$ implies the existence of a Rees decomposition for $\mathcal{M}$. Exploiting again the isomorphism $\mathcal{M}^{(k)} \cong \mathcal{P}^{(k)}/\mathcal{I}^{(k)}$, one easily sees that a vector space basis of $\mathcal{M}_q^{(k)}$ is induced by all terms $x^\mu \notin \mathcal{I}$ with $|\mu| = q$ and $\mathrm{cls}\, \mu \geq k$. By the definition of quasi-regularity, multiplication with x_k is injective on $\mathcal{M}^{(k)}$, hence we take $\{x_1, \ldots, x_k\}$ as multiplicative variables for such a term (which is exactly the assignment used in the Rees decomposition induced by a Pommaret basis according to Remark 2.3).

We claim now that this assignment yields a Rees decomposition of $\mathcal{M}_{\geq q}$ (and hence one of $\mathcal{M}$, since we only have to add all terms $x^\mu \notin \mathcal{I}$ with $|\mu| < q$ without any multiplicative variables). The only thing to prove is that our decomposition covers all of $\mathcal{M}_{\geq q}$. If $x^\mu \notin \mathcal{I}$ is an arbitrary term

with $|\mu| = q+1$ and $\operatorname{cls}\mu = k$, then we can write $x^\mu = x_k \cdot x^{\mu-1_k}$. Obviously, $x^\mu \notin \mathcal{I}$ implies $x^{\mu-1_k} \notin \mathcal{I}$ and $\operatorname{cls}(\mu - 1_k) \geq k$ so that x_k is multiplicative for it. Hence all of $\mathcal{M}_{q+1}$ is covered and an easy induction shows that we have indeed a decomposition of $\mathcal{M}_{\geq q}$.

Proposition 2.2 entails now that $\mathcal{I}$ has a *weak* Pommaret basis of degree q. As the autoreduction of a weak basis to a strong one can only decrease the degree, $\mathcal{I}$ has a strong Pommaret basis of degree at most q. However, if the degree of the basis actually decreased, then, by the converse statement already proven, $(x_1, \ldots, x_n)$ would be a quasi-regular sequence for $\mathcal{M}$ at a lower degree than q contradicting our assumptions.

The same "reverse" argument shows that if $\mathcal{I}$ has a Pommaret basis of degree q, then the sequence $(x_1, \ldots, x_n)$ cannot be quasi-regular for $\mathcal{M}$ at a lower degree, as otherwise a Pommaret basis of lower degree would exist which is not possible by the uniqueness of strong Pommaret bases. □

For monomial ideals $\mathcal{I} \subseteq \mathcal{P}$ a much stronger statement is possible. Using again the isomorphism $\mathcal{M}^{(k)} \cong \mathcal{P}^{(k)}/\mathcal{I}^{(k)}$, we may identify elements of $\mathcal{M}^{(k)}$ with linear combinations of the terms $x_\nu \notin \mathcal{I}$ satisfying $\operatorname{cls} x_\nu > k$. Then we obtain the following simple relationship between the Pommaret basis of $\mathcal{I}$ and the kernels of the maps μ_k appearing in Theorem 5.2.

Proposition 5.5. *Let $x_1, \ldots, x_n$ be δ-regular coordinates for the quasi-stable ideal $\mathcal{I}$. Furthermore, let $\mathcal{H}$ be the Pommaret basis of $\mathcal{I}$ and set $\mathcal{H}_k = \{x_\nu \in \mathcal{H} \mid \operatorname{cls}\nu = k\}$ for any $1 \leq k \leq n$. Then the set $\{x_{\mu-1_k} \mid x^\mu \in \mathcal{H}_k\}$ is a basis of* $\ker \mu_k$.

Proof. Assume that $x_\nu \in \mathcal{H}_k$. Then $x_{\nu-1_k} \notin \mathcal{I}$, as otherwise the Pommaret basis $\mathcal{H}$ was not involutively autoreduced, and hence we find $x_{\nu-1_k} \in \ker \mu_k$.

Conversely, suppose that $x_\nu \in \ker \mu_k$. Obviously, this implies $x_{\nu+1_k} \in \mathcal{I}$ and the Pommaret basis $\mathcal{H}$ must contain an involutive divisor of $x_{\nu+1_k}$. If this divisor was not $x_{\nu+1_k}$ itself, the term x_ν would have to be an element of $\mathcal{I}$ which is obviously not possible. Since $x_\nu \in \ker \mu_k$ entails $\operatorname{cls}(\nu+1_k) = k$, we thus find $x_{\nu+1_k} \in \mathcal{H}_k$. □

Remark 5.6. These results also lead to a simple proof of the characterisation (ii) of a quasi-stable ideal in Proposition 4.3. If $\mathcal{I}$ is quasi-stable, then the coordinates $x_1, \ldots, x_n$ are δ-regular for it, hence by Theorem 5.4 they form a quasi-regular sequence for $\mathcal{P}/\mathcal{I}$ at a suitably chosen degree. By Proposition 4.3, Condition (iv), we have that $\mathcal{I}^{\mathrm{sat}} = \mathcal{I} : \langle x_1 \rangle^{\infty}$ and hence multiplication by x_1 is injective on $\mathcal{P}/\mathcal{I}^{\mathrm{sat}}$. As obviously $\mathcal{P}/\langle \mathcal{I}, x_1, \ldots, x_j \rangle^{\mathrm{sat}} \cong \mathcal{P}^{(j)}/(\mathcal{I}^{(j)})^{\mathrm{sat}}$, we can apply the same argument also for all $1 \leq j < D$.

Conversely, if $x_1 \cdot f \in \mathcal{I}$ for a polynomial $f \in \mathcal{P} \setminus \mathcal{I}$, then $f \in \mathcal{I}^{\mathrm{sat}} \setminus \mathcal{I}$ and hence $\deg f < \operatorname{sat} \mathcal{I}$. Thus x_1 is quasi-regular for $\mathcal{P}/\mathcal{I}$ at the degree $\operatorname{sat} \mathcal{I}$. Using again the isomorphisms $\mathcal{P}/\langle \mathcal{I}, x_1, \ldots, x_j \rangle^{\mathrm{sat}} \cong \mathcal{P}^{(j)}/(\mathcal{I}^{(j)})^{\mathrm{sat}}$, we can apply the same argument for all $1 \leq j < D$, so that $(x_1, \ldots, x_D)$ is a quasi-regular sequence for $\mathcal{P}/\mathcal{I}$ at a sufficiently high degree.

The characterisation (ii) of Proposition 4.3 obviously implies that the set $\{x_1, \ldots, x_D\}$ is maximally independent modulo $\mathcal{I}^{\mathrm{sat}}$.
Hence $\dim \mathcal{P}/\langle \mathcal{I}^{\mathrm{sat}}, x_1, \ldots, x_D \rangle = 0$ entailing that $(x_1, \ldots, x_n)$ is a quasi-regular sequence for the algebra $\mathcal{P}/\mathcal{I}$ at a sufficiently high degree. By Theorem 5.4, the ideal $\mathcal{I}$ is thus quasi-stable.

References

[1] Th. Becker and V. Weispfenning. *Gröbner Bases*. Graduate Texts in Mathematics 141. Springer-Verlag, New York, 1993.

[2] I. Bermejo and P. Gimenez. Saturation and Castelnuovo-Mumford regularity. *J. Alg.*, to appear, 2005.

[3] J. Calmet, M. Hausdorf, and W.M. Seiler. A constructive introduction to involution. In R. Akerkar, editor, *Proc. Int. Symp. Applications of Computer Algebra — ISACA 2000*, pages 33–50. Allied Publishers, New Delhi, 2001.

[4] V.P. Gerdt. On the relation between Pommaret and Janet bases. In V.G. Ghanza, E.W. Mayr, and E.V. Vorozhtsov, editors, *Computer Algebra in Scientific Computing — CASC 2000*, pages 167–182. Springer-Verlag, Berlin, 2000.

[5] V.P. Gerdt and Yu.A. Blinkov. Involutive bases of polynomial ideals. *Math. Comp. Simul.*, 45:519–542, 1998.

[6] V.P. Gerdt and Yu.A. Blinkov. Minimal involutive bases. *Math. Comp. Simul.*, 45:543–560, 1998.

[7] V.W. Guillemin and S. Sternberg. An algebraic model of transitive differential geometry. *Bull. Amer. Math. Soc.*, 70:16–47, 1964.

[8] M. Hausdorf and W.M. Seiler. An efficient algebraic algorithm for the geometric completion to involution. *Appl. Alg. Eng. Comm. Comp.*, 13:163–207, 2002.

[9] W.M. Seiler. A combinatorial approach to involution and δ-regularity I: Involutive bases in polynomial algebras of solvable type. Preprint Universität Mannheim, 2002.

[10] W.M. Seiler. A combinatorial approach to involution and δ-regularity II: Structure analysis of polynomial modules with Pommaret bases. Preprint Universität Mannheim, 2002.

[11] W.M. Seiler. *Involution — The Formal Theory of Differential Equations and its Applications in Computer Algebra and Numerical Analysis*. Algorithms and Computation in Mathematics. Springer-Verlag, Berlin/Heidelberg, Manuscript accepted for publication.

[12] R.P. Stanley. Hilbert functions of graded algebras. *Adv. Math.*, 28:57–83, 1978.

[13] H.D. Wahlquist. Monte Carlo calculation of Cartan characters. In N. Breton, R. Capovilla, and T. Matos, editors, *Aspects of General Relativity and Mathematical Physics*, pages 168–174. Centro de Investigacion y de Estudios Avanzados de IPN, Mexico City, 1993.

Theory of Laminated Wave Turbulence: Open Questions

Elena Kartashova

J. Kepler University, Linz, Austria, RISC
lena@risc.uni-linz.ac.at

Abstract

Theory of laminated wave turbulence developed recently allows to study at the same base both types of turbulent wave systems - statistical and discrete. The main novelty of the theory is the following: discrete effects do appear not only in the long-wave part of the spectral domain (corresponding to small wave numbers) but all through the spectra. One of the new problems put forth by the theory of laminated wave turbulence is construction of fast algorithms for computations in integers of order 10^{12} and more. In this paper we present the mathematical basics of this new theory in terms of integer points of resonant manifolds and formulate its open questions, both theoretical and computational, in order to attract "pure"mathematicians and specialists in computer algebra to work on the subject.

Keywords: Nonlinear waves, wave turbulence, discrete wave systems, irrational and transcendental algebraic equations, computations in integers

1 Introduction

The roots of the theory of nonlinear dispersive waves date back to hydrodynamics of the 19th century. It was observed, both experimentally and theoretically, that, under certain circumstances, the dissipative effects in nonlinear waves become less important then the dispersive ones. In this way, balance between nonlinearity and dispersion gives rise to formation of stable patterns (solitons, cnoidal waves, etc.). Driven by applications in plasma physics, these phenomena were widely studied, both analytically

and numerically, starting from the middle of the 20th century. The main mathematical break-through in the theory of nonlinear evolutionary PDEs was the discovery of the phenomenon of their integrability that became the starting point of the modern theory of integrable systems with Korteweg-de Vries equation being the first instance in which integrability appeared. But most evolutionary PDEs are not integrable, of course. As a powerful tool for numerical simulations, the method of kinetic equation has been developed in 1960-th and applied to many different types of dispersive evolutionary PDEs. The wave kinetic equation is approximately equivalent to the initial nonlinear PDE: it is an averaged equation imposed on a certain set of correlation functions and it is in fact one limiting case of the quantum Bose-Einstein equation while the Boltzman kinetic equation is its other limit. Some statistical assumptions have been used in order to obtain kinetic equations; the limit of their applicability then is a very complicated problem which should be solved separately for each specific equation.

The role of the nonlinear dispersive PDEs in the theoretical physics is so important that the notion of dispersion is used for "physical" classification of the equations in partial variables. On the other hand, the only mentioning of the notion "dispersion relation" in mathematical literature we have found in the book of V.I.Arnold [1] who writes about important physical principles and concepts such as energy, variational principle, the Lagrangian theory, dispersion relations, the Hamiltonian formalism, etc. which gave a rise for the development of large areas in mathematics (theory of Fourier series and integrals, functional analysis, algebraic geometry and many others). But he also could not find place for it in the consequent mathematical presentation of the theory of PDEs and the words "dispersion relation" appear only in the introduction.

In our paper we present wave turbulence theory as a base of the "physical" classification of PDEs, trying to avoid as much as possible specific physical jargon and give a "pure" mathematician a possibility to follow its general ideas and results. We show that the main mathematical object

of the wave turbulence theory is an algebraic system of equations called resonant manifolds. We also present here the model of the laminated wave turbulence that includes classical statistical results on the turbulence as well as the results on the discrete wave systems. It is shown that discrete characteristics of the wave systems can be described in terms of integer points on the rational manifolds which is the main novelty of the theory of laminated turbulence. Some applications of this theory for explanation of important physical effects are given. A few open mathematical and numerical problems are formulated at the end. Our purpose is to attract pure mathematicians to work on this subject.

2 General Notions

For the completeness of presentation we began this section with a very brief sketch of the traditional mathematical approach to the classification of PDEs.

2.1 Mathematical Classification

Well-known mathematical classification of PDEs is based on **the form of equations** and can be briefly presented as follows. For a bivariate PDE of the second order

$$a\psi_{xx} + b\psi_{xy} + c\psi_{yy} = F(x, y, \psi, \psi_x, \psi_y)$$

its characteristic equation is written as

$$\frac{\boldsymbol{d}x}{\boldsymbol{d}y} = \frac{b}{2a} \pm \frac{1}{2a}\sqrt{b^2 - 4ac}$$

and three types of PDEs are defined:

- $b^2 < 4ac$, elliptic PDE: $\psi_{xx} + \psi_{yy} = 0$
- $b^2 > 4ac$, hyperbolic PDE: $\psi_{xx} - \psi_{yy} - x\psi_x = 0$

- $b^2 = 4ac$, parabolic PDE: $\psi_{xx} - 2xy\psi_y - \psi = 0$

Each type of PDE demands then special type of initial/boundary conditions for the problem to be well-posed. "Bad"example of Tricomi equation $y\psi_{xx}+\psi_{yy} = 0$ shows immediately incompleteness of this classification even for second order PDEs because a PDE can change its type depending, for instance, on the initial conditions. This classification can be generalized to PDEs of more variables but not to PDEs of higher order.

2.2 Physical Classification

Physical classification of PDEs is based on **the form of solution** and is almost not known to pure mathematicians. In this case, a PDE is regarded in the very general form, without any restrictions on the number of variables or the order of equation. On the other hand, the necessary preliminary step in this classification is the division of all the variables into two groups - time- and space-like variables. This division originated from the special relativity theory where time and three-dimensional space are treated together as a single four-dimensional Minkowski space. In Minkowski space a metrics allowing to compute an interval s along a curve between two events is defined analogously to distance in Euclidean space:

$$\boldsymbol{d}s^2 = \boldsymbol{d}x^2 + \boldsymbol{d}y^2 + \boldsymbol{d}z^2 - c^2\boldsymbol{d}t^2$$

where c is speed of light, x, y, z and t denote respectively space and time variables. Notice that though in mathematical classification all variables are treated equally, obviously its results can be used in any applications only after similar division of variables have been done.

Suppose now that linear PDE with constant coefficients has a wave-like solution

$$\psi(x,t) = A\exp i[kx - \omega t] \quad \text{or} \quad \psi(x,t) = A\sin(kx - \omega t)$$

with amplitude A, wave-number k and wave frequency ω. Then the substitution of $\partial_t = -i\omega, \quad \partial_x = ik$ transforms LPDE into a **polynomial**

on ω and k, for instance:

$$\psi_t + \alpha\psi_x + \beta\psi_{xxx} = 0 \quad \Rightarrow \ \omega(k) = \alpha k - \beta k^3,$$

$$\psi_{tt} + \alpha^2\phi_{xxxx} = 0 \quad \Rightarrow \ \omega^2(k) = \alpha^2 k^4,$$

$$\psi_{tttt} - \alpha^2\psi_{xx} + \beta^2\psi = 0 \quad \Rightarrow \ \omega^4(k) = \alpha^2 k^2 + \beta^2$$

where α and β are constants.

Definition Real-valued function $\omega = \omega(k) : \ \mathrm{d}^2\omega/\mathrm{d}k^2 \neq 0$ is called *dispersion relation* or *dispersion function*. A linear PDE with wave-like solutions are called evolutionary dispersive LPDE. A nonlinear PDE with dispersive linear part are called evolutionary dispersive NPDE.

This way all PDEs are divided into two classes - dispersive and non-dispersive [2]. This classification is not complementary to a standard mathematical one. For instance, though hyperbolic PDEs normally do not have dispersive wave solutions, the hyperbolic equation $\psi_{tt} - \alpha^2\psi_{xx} + \beta^2\psi = 0$ has them. Given dispersion relation allows to re-construct corresponding linear PDE. All definitions above could be easily reformulated for a case of more space variables, namely $x_1, x_2, ..., x_n$. Linear part of the initial PDE takes then form

$$P(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_n})$$

and correspondingly dispersion relation can be computed from

$$P(-i\omega, ik_1, ..., ik_n) = 0$$

with the polynomial P. In this case we will have not a wave number k but a *wave vector* $\vec{k} = (k_1, ..., k_n)$ and the condition of non-zero second derivative of the dispersion function takes a matrix form:

$$|\frac{\partial^2\omega}{\partial k_i \partial k_j}| \neq 0.$$

2.3 Perturbation technique

Perturbation or asymptotic methods (see, for instance, [3]) are much in use in physics and are dealing with equations having some small parameter $\varepsilon > 0$. To understand the results presentated in the next Section one needs to have some clear idea about the perturbation technique and this is the reason why we give here a simple algebraic example of its application. The main idea of a perturbation method is very straightforward - an unknown solution, depending on ε, is written out in a form of infinite series on different powers of ε and coefficients in front of any power of ε are computed consequently. Let us take an algebraic equation

$$x^2 - (3 - 2\varepsilon)x + 2 + \varepsilon = 0 \tag{1}$$

and try to find its asymptotic solutions. If $\varepsilon = 0$ we get

$$x^2 - 3x + 2 = 0 \tag{2}$$

with roots $x = 1$ and $x = 2$. Eq.(1) is called *perturbed* and Eq.(2) - *unperturbed.* Natural suggestion is that the solutions of perturbed equation differ only a little bit from the solutions of unperturbed one. Let us look for solutions of Eq.(1) in the form

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + ...$$

where x_0 is a solution of Eq.(2), i.e. $x_0 = 1$ or $x_0 = 2$. Substituting this infinite series into Eq.(1), collecting all the terms with the same degree of ε and consequent equaling to zero all coefficients in front of different powers of ε leads to an algebraic system of equations

$$\begin{cases} \varepsilon^0 : x_0^2 - 3x_0 + 2 = 0, \\ \varepsilon^1 : 2x_0x_1 - 3x_1 - 2x_0 + 1 = 0, \\ \varepsilon^2 : 2x_0x_2 + x_1^2 - 3x_2 - 2x_1 = 0, \\ ... \end{cases} \tag{3}$$

with solutions

$$x_0 = 1, \quad x_1 = -1, \quad x_2 = 3, ... \quad \text{and} \quad x = 1 - \varepsilon + 3\varepsilon^2 + ...;$$

$$x_0 = 2, \quad x_1 = 3, \quad x_2 = -3, ... \quad \text{and} \quad x = 2 + 3\varepsilon - 3\varepsilon^2 + ...$$

Notice that exact solutions of Eq.(1) are

$$x = \frac{1}{2}[3 + 2\varepsilon \pm \sqrt{1 + 8\varepsilon + 4\varepsilon^2}]$$

and the use of binomial representation for the expression under the square root

$$(1 + 8\varepsilon + 4\varepsilon^2)^{\frac{1}{2}} = 1 + (8\varepsilon + 4\varepsilon^2) + \frac{\frac{1}{2}(\frac{-1}{2})}{2!}(8\varepsilon + 4\varepsilon^2)^2 + ... = 1 + 4\varepsilon - 6\varepsilon^2 + ...$$

gives finally

$$x = \frac{1}{2}(3 + 2\varepsilon - 1 - 4\varepsilon + 6\varepsilon^2 + ...) = 1 - \varepsilon + 3\varepsilon^2 + ...$$

and

$$x = \frac{1}{2}(3 + 2\varepsilon + 1 + 4\varepsilon - 6\varepsilon^2 + ...) = 2 + 3\varepsilon - 3\varepsilon^2 + ...$$

as before. This example was chosen because the exact solution in this case is known and can be compared to the asymptotic one. The same approach is used for partial differential equations, also in the cases when exact solutions are not known. The only difference would be more elaborated computations resulting in some system of ordinary differential equations instead of Eqs.(3) (see next Section).

3 Wave Turbulence Theory

3.1 Wave Resonances

Now we are going to introduce the notion of the wave resonance which is the mile-stone for the whole theory of evolutionary dispersive NPDE and

therefore for the wave turbulence theory. Let us consider first a linear oscillator driven by a small force

$$x_{tt} + p^2 x = \varepsilon e^{i\Omega t}.$$

Here p is eigenfrequency of the system, Ω is frequency of the driving force and $\varepsilon > 0$ is a small parameter. Deviation of this system from equilibrium is small (of order ε), if there is no resonance between the frequency of the driving force $\varepsilon e^{i\Omega t}$ and an eigenfrequency of the system. If these frequencies coincide then the amplitude of oscillator grows linearly with the time and this situation is called *resonance* in physics. Mathematically it means *existence of unbounded solutions.*

Let us now regard a (weakly) nonlinear PDE of the form

$$L(\psi) = \varepsilon N(\psi) \tag{4}$$

where L is an arbitrary linear dispersive operator and N is an arbitrary nonlinear operator. Any two solutions of $L(\psi) = 0$ can be written out as

$$A_1 \exp i[\vec{k}_1\vec{x} - \omega(\vec{k}_1)t] \quad \text{and} \quad A_2 \exp i[\vec{k}_2\vec{x} - \omega(\vec{k}_2)t]$$

with constant amplitudes A_1, A_2. Intuitively natural expectation is that solutions of weakly nonlinear PDE will have the same form as linear waves but perhaps with amplitudes depending on time. Taking into account that nonlinearity is small, each amplitude is regarded as *a slow-varying function of time*, that is $A_j = A_j(t/\varepsilon)$. Standard notation is $A_j = A_j(T)$ where $T = t/\varepsilon$ is called *slow time.* Since wave energy is by definition proportional to amplitude's square A_j^2 it means that in case of nonlinear PDE *waves exchange their energy.* This effect is also described as "waves are interacting with each other"or "there exists energy transfer through the wave spectrum"or similar.

Unlike linear waves for which their linear combination was also solution of $L(\psi) = 0$, it is not the case for nonlinear waves. Indeed, substitution of two linear waves into the operator $\epsilon N(\psi)$ generates terms of the form $\exp i[(\vec{k}_1 + \vec{k}_2)\vec{x} - [\omega(\vec{k}_1) + \omega(\vec{k}_2)]t$ which play the role of a small driving

force for the linear wave system similar to the case of linear oscillator above. This driving force gives a small effect on a wave system till resonance occurs, i.e. till the wave number and the wave frequency of the driving force does not coincide with some wave number and some frequency of eigenfunction:

$$\begin{cases} \omega_1 + \omega_2 = \omega_3, \\ \vec{k}_1 + \vec{k}_2 = \vec{k}_3 \end{cases} \tag{5}$$

where notation $\omega_i = \omega(\vec{k}_i)$ is used. This system describes so-called *resonance conditions* or *resonance manifold*.

The perturbation technique described above produces the equations for the amplitudes of resonantly interacting waves $A_j = A_j(T)$. Let us demonstrate it taking as example barotropic vorticity equation (BVE) on a sphere

$$\frac{\partial \triangle\psi}{\partial t} + 2\frac{\partial \psi}{\partial \lambda} + \varepsilon J(\psi, \triangle\psi) = 0 \tag{6}$$

where

$$\triangle\psi = \frac{\partial^2\psi}{\partial\phi^2} + \frac{1}{\cos^2\phi}\frac{\partial^2\psi}{\partial\lambda^2} - \tan\phi\frac{\partial\psi}{\partial\phi} \quad \text{and} \quad J(a,b) = \frac{1}{cos\phi}(\frac{\partial a}{\partial\lambda}\frac{\partial b}{\partial\phi} - \frac{\partial a}{\partial\phi}\frac{\partial b}{\partial\lambda}).$$

The linear part of spherical BVE has wave solutions in the form

$$AP_n^m(\sin\phi)\exp i[m\lambda + \frac{2m}{n(n+1)}t],$$

where A is constant wave amplitude, $\omega = -2m/[n(n+1)]$ and $P_n^m(x)$ is the associated Legendre function of degree n and order m.

One of the reasons to choose this equation as an example is following. Till now a linear wave was supposed to have much more simple form, namely, $A\exp i[\vec{k}\vec{x} - \omega(\vec{k})t]$ without any additional factor of a functional form. For a physicist it is intuitively clear that if the factor is some *oscillatory function of only space variables* then we will still have a wave of a sort " "but it would be difficult to include it in an overall definition.

We seem to be left at present with the looser idea that whenever oscillations in space are coupled with oscillation in time through a dispersion relation, we expect the typical effects of dispersive waves"[2]. By the way, most physically important dispersive equations have the waves of this form.

Now let us keep in mind that a wave is something more complicated then just a sin but still smooth and periodic, and let us look where perturbation method will lead us. An approximate solution has a form

$$\psi = \psi_0(\lambda, \phi, t, T) + \varepsilon\psi_1(\lambda, \phi, t, T) + \varepsilon^2\psi_2(\lambda, \phi, t, T) + ...$$

where $T = t/\varepsilon$ is the slow time and the zero approximation ψ_0 is given as a sum of three linear waves:

$$\psi_0(\lambda, \phi, t, T) = \sum_{k=1}^{3} A_k(T) P^{(k)} \exp(i\theta_k) \tag{7}$$

with notations $P^{(k)} = P_{n_k}^{m_k}$ and $\theta_k = m_k\lambda - \omega_k t$. Then

$$\begin{cases} \varepsilon^0: \quad \partial\triangle\psi_0/\partial t + 2\partial\psi_0/\partial\lambda = 0, \\ \varepsilon^1: \quad \partial\triangle\psi_1/\partial t + 2\partial\psi_1/\partial\lambda = -J(\psi_0, \triangle\psi_0) - \partial\triangle\psi_0/\partial T, \\ \varepsilon^2: ..... \end{cases} \tag{8}$$

and

$$\frac{\partial\triangle\psi_0}{\partial\lambda} = i\sum_{k=1}^{3} P^{(k)} m_k N_k [A_k \exp(i\theta_k) - A_k^* \exp(-i\theta_k)];$$

$$\frac{\partial\triangle\psi_0}{\partial\phi} = -\sum_{k=1}^{3} N_k \frac{d}{d\phi} P^{(k)} \cos\phi [A_k \exp(i\theta_k) + A_k^* \exp(-i\theta_k)];$$

$$\begin{aligned} J(\psi_0, \triangle\psi_0) = &- i\sum_{j,k=1}^{3} N_k m_j P^{(j)} \frac{d}{d\phi} P^{(k)} [A_j \exp(i\theta_j) - A_j^* \exp(-i\theta_j)] \\ &\cdot [A_k \exp(i\theta_k) + A_k^* \exp(-i\theta_k)] + i\sum_{j,k=1}^{3} N_k m_k P^{(k)} \frac{d}{d\phi} P^{(j)} \\ &\cdot [A_j \exp(i\theta_j) + A_j^* \exp(-i\theta_j)][A_k \exp(i\theta_k) - A_k^* \exp(-i\theta_k)]; \end{aligned}$$

$$\frac{\partial\triangle\psi_0}{\partial T} = \sum_{k=1}^{3} P^{(k)} N_k [\frac{dA_k}{dT} \exp(i\theta_k) + \frac{dA_k^*}{dT} \exp(-i\theta_k)].$$

leads to the condition of unbounded growth of the left hand in the form

$$J(\psi_0, \triangle\psi_0) = \frac{\partial\triangle\psi_0}{\partial T}$$

with resonance conditions $\theta_j + \theta_k = \theta_i \quad \forall\ j, k, i = 1, 2, 3$. Let us fix some specific resonance condition, say, $\theta_1 + \theta_2 = \theta_3$, then

$$\begin{aligned} \frac{\partial\triangle\psi_0}{\partial T} &\cong -N_3 P^{(3)} [\frac{dA_3}{dT} \exp(i\theta_3) + \frac{dA_3^*}{dT} \exp(-i\theta_3)], \\ J(\psi_0, \triangle\psi_0) &\cong -i(N_1 - N_2)(m_2 P^{(2)} \frac{d}{d\phi} P^{(1)}) - m_1 P^{(1)} \frac{d}{d\phi} P^{(2)} \cdot \\ &\quad A_1 A_2 \exp[i(\theta_1 + \theta_2)] - A_1^* A_2^* \exp[-i(\theta_1 + \theta_2)], \end{aligned}$$

where notation $\cong$ means that only those terms are written out which can generate chosen resonance. Let us substitute these expressions into the coefficient by ε^1, i.e. into the equation

$$\frac{\partial\triangle\psi_1}{\partial t} + 2\frac{\partial\psi_1}{\partial\lambda} = -J(\psi_0, \triangle\psi_0) - \frac{\partial\triangle\psi_0}{\partial T},$$

multiply both parts of it by

$$P^{(3)} \sin\phi[A_3 \exp(i\theta_3) + A_3^* \exp(-i\theta_3)]$$

and integrate all over the sphere with $t \to \infty$. As a result following equation can be obtained:

$$N_3 \frac{dA_3}{dT} = 2iZ(N_2 - N_1)A_1 A_2,$$

where

$$Z = \int_{-\pi/2}^{\pi/2} [m_2 P^{(2)} \frac{d}{d\phi} P^{(1)} - m_1 P^{(1)} \frac{d}{d\phi} P^{(2)}] \frac{d}{d\phi} P^{(3)} d\phi.$$

The same procedure obviously provides the analogous equations for A_2 and A_3 while fixing corresponding resonance conditions:

$$\begin{aligned} N_1 \frac{dA_1}{dT} &= -2iZ(N_2 - N_3)A_3 A_2^*, \\ N_2 \frac{dA_2}{dT} &= -2iZ(N_3 - N_1)A_1^* A_3, \end{aligned} \tag{9}$$

In general, the simplest system of equations on the amplitudes of three resonantly interacting waves is often regarded in the form

$$\begin{cases} \dot{A}_1 = \alpha_1 A_3 A_2^*, \\ \dot{A}_2 = \alpha_2 A_1^* A_3, \\ \dot{A}_3 = \alpha_3 A_1 A_2\,, \end{cases} \tag{10}$$

and is refereed to as a 3-wave system (keeping in mind that analogous system has to be written out for A_i^*). Coefficients α_i depend on the initial NPDE. Similar system of equations can be obtained for 4-wave interactions, with the products of three different amplitudes on the right hand, and so on.

3.2 Zakharov-Kolmogorov energy spectra

The idea that a dispersive wave system contains many resonances and wave interactions are stochastic led to the statistical theory of wave turbulence. This theory is well developed [4] and widely used in oceanology and plasma physics describing a lot of turbulent transport phenomena. Avoiding the language of Hamiltonian systems, correlators of a wave field, etc., one can formulate its main results in the following way. Any nonlinearity in Eq.(4) can be written out as

$$\Sigma_i \frac{V_{(12..i)}\delta(\vec{k}_1+\vec{k}_2+...+\vec{k}_i)}{\delta(\omega_1+\omega_2+...+\omega_i)} A_1 A_2 \cdots A_i \qquad (11)$$

where δ is Dirac delta-function and $V_{(12..i)}$ is a vertex coefficient. This presentation, together with some additional statistical suggestions, is used then to construct a wave kinetic equation, with corresponding vertex coefficients and delta-functions in the under-integral expression, of the form

$$\dot{A}_1 = \int |V_{(123)}|^2 \delta(\omega_1-\omega_2-\omega_3)\delta(\vec{k}_1-\vec{k}_2-\vec{k}_3)(A_2A_3-A_1A_2-A_1A_3)\mathbf{d}\vec{k}_2\mathbf{d}\vec{k}_3$$

for 3-waves interactions, and similar for i-waves interactions. One of the most important discoveries in the statistical wave turbulence theory are stationary exact solutions of the kinetic equations first found in [5]. These solutions are now called Zakharov-Kolmogorov (ZK) energy spectra and they describe energy cascade in the wave field. In other words, energy of the wave with wave vector $\vec{k}$ is proportional to k^α with $\alpha < 0$ and magnitude of α depends on the specific of the wave system. Discovery of ZK spectra played tremendous role in the wave turbulence theory and till the works of last decade (e.g. [6]-[8]) it was not realized that some turbulent effects are not due to the statistical properties of a wave field and are not described by kinetic equations or ZK energy spectra.

3.3 Small Divisors Problem

In order to use presentation (11) one has to check whether so defined nonlinearity is finite. This problem is known as *the small divisors problem*

and its solution depends on whether wave vectors have real or integer coordinates.

Wave systems with *continuous spectra* were studied by Kolmogorov, Arnold and Moser [9]-[11] and main results of KAM-theory can be briefly formulated as follows. If dispersion function ω is defined on real-valued wave vectors and the ratio $\alpha_{ij} = \omega_i/\omega_j$ *is not a rational number* for any two wave vectors $\vec{k}_i$ and $\vec{k}_j$, then

- (1C) Wave system is decomposed into disjoint invariant sets (KAM tori) carrying quasi-periodic motions;
- (2C) If the size of the wave system tends to infinity, (1C) does not contradict ergodicity, random phase approximation can be assumed, kinetic equations and ZK energy spectra describe the wave system properly;
- (3C) Union of invariant tori has positive Liouville measure and $\mathbb{Q}$ has measure 0, there exclusion of the waves with rational ratio of their dispersions is supposed to be not very important.

Wave systems with *discrete spectra* demonstrate [6]-[8] substantially different energetic behavior:

- (1D) Wave system is decomposed into disjoint discrete classes carrying periodic motions or empty; for any two waves with wave vectors $\vec{k}_i$ and $\vec{k}_j$ belonging to the same class, the ratio $\alpha_{ij} = \omega_i/\omega_j$ *is a rational number*;
- (2D) Energetic behavior of the wave system does not depend on its size, is not stochastic and is described by a few isolated periodic processes governed by Sys.(10);
- (3D) In many wave systems (for instance, with rational dispersion function, as spherical Rossby waves in the earth atmosphere) *KZ spectra do not exist* and *discrete classes describe the total energy transfer.*

3.4 Laminated Wave Turbulence

The results formulated in the previous section gave rise to the model of laminated wave turbulence [12] which includes two co-existing layers of turbulence in a wave system - continuous and discrete layers, each demonstrating specific energetic behavior. In other words, KAM-theory describes the wave systems leaving some "holes"in the wave spectra which are "full-filled"in the theory of laminated turbulence.

Continuous layer, with its kinetic equation, energy cascades, ZK spectra, etc. is well-studied while the existence of the discrete layer was realized quite recently. In order to understand which manifestations of the discrete layer are to be expected in numerical or laboratory experiments, let us regard an example with dispersion function $\omega = 1/\sqrt{m^2+n^2}$. First of all, it is important to realize: the fact that the ratio $\alpha_{ij} = \omega_i/\omega_j$ is a rational number *does not imply* that dispersion function ω is a rational function. Indeed, for $\omega = 1/\sqrt{m^2+n^2}$ we have

$$\omega : \mathbb{Z} \times \mathbb{Z} \to \mathbb{R}$$

and for wave vectors $\vec{k}_1 = (2,1)$ and $\vec{k}_2 = (9,18)$, the ratio $\omega_1/\omega_2 = 1/9$ is rational number though ω is irrational function of integer variables. Decomposition of all discrete waves into disjoint discrete classes Cl_q in this case has the form

$$\{\vec{k}_i = (m_i, n_i)\} \in Cl_q \text{ if } |\vec{k}_i| = \gamma_i\sqrt{q},\ \forall\ i = 1, 2, ...$$

where γ_i is some integer and q is *the same square-free integer* for all wave vectors of the class Cl_q. The equation $\omega_1 + \omega_2 = \omega_3$ has solutions only if

$$\exists \tilde{q} :\ \vec{k}_1,\ \vec{k}_2,\ \vec{k}_3 \in Cl_{\tilde{q}}.$$

This is *necessary condition*, not sufficient. The use of this necessary condition allows to cut back substantially computation time needed to find solution of irrational equations in integers. Namely, one has to construct

classes first and afterwards look for the solutions among the waves belonging to the same class. In this way, instead of solving *the irrational equation on 6 variables*

$$1/\sqrt{m_1^2+n_1^2}+1/\sqrt{m_2^2+n_2^2}=1/\sqrt{m_3^2+n_3^2} \tag{12}$$

it is enough to solve *the rational equation on 3 variables*

$$1/\gamma_1+1/\gamma_2=1/\gamma_3. \tag{13}$$

There are few different reasons for a class $Cl_{\tilde{q}}$ to be empty:

- Index $\tilde{q}$ has no representation as a sum of two squares of integer numbers. Example: Cl_3 .
  All these indices are described by Euler's theorem: *an integer can be represented as a sum of two squares if and only if its prime factorization contains every prime factor $p \equiv 4u+3$ in an even degree.*

- Index $\tilde{q}$ can be represented as a sum of two squares but Eq.(13) is not satisfied in the physically relevant finite domain $|m_i|, |n_i| \leq D \ \forall i = 1,2,3$. Example: Cl_{500009} in the domain $D = 1000$.
  The number $500009 = 325^2 + 628^2$ and the minimal possible form of Eq.(13) (that is, with minimal weights) is
  $$\frac{1}{2\sqrt{500009}}+\frac{1}{2\sqrt{500009}}=\frac{1}{\sqrt{500009}}$$
  which obviously can not be satisfied with all $m_i, n_i \leq 1000$.

- Index $\tilde{q}$ can be represented as a sum of two squares and Eq.(13) is satisfied but not linear equation $\vec{k}_1+\vec{k}_2=\vec{k}_3$, that is the second equation of Sys.(5), in finite domain. Example: Cl_{465881} in the domain $D = 1000$.
  The number $465881 = 59^2 + 680^2 = 316^2 + 605^2$ and Eq.(13) takes the form
  $$\frac{1}{2\sqrt{465881}}+\frac{1}{2\sqrt{465881}}=\frac{1}{\sqrt{465881}} \ \Rightarrow \ \frac{1}{2}+\frac{1}{2}=1.$$

There are reasons to suggest that in some cases the linear equation renders the solution set empty not only in a finite domain but generally, for example for Cl_2, but at present we do not posses a rigorous prove of it. Obviously, each non-empty class has infinite number of elements due to the existence of proportional vectors so that $\vec{k}_1 = (2,1)$ with the norm $|\vec{k}_1| = \sqrt{5}$ and its proportional $\vec{k}_2 = (9,18)$ with the norm $|\vec{k}_2| = 9\sqrt{5}$ belong to the same class Cl_5. On the other hand, not all the elements of a class are parts of some solution which means that *not all waves take part in resonant interactions*. Complete analysis of the properties of the classes is an important necessary step before performing any computer simulations, an example of the exhaustive analysis can be in [13] for the dispersion function $\omega = (m^2 + n^2)^{1/4}$ and resonance conditions

$$\begin{cases} (m_1^2 + n_1^2)^{1/4} + (m_2^2 + n_2^2)^{1/4} = (m_3^2 + n_3^2)^{1/4} + (m_4^2 + n_4^2)^{1/4} \\ m_1 + m_2 = m_3 + m_4, \quad n_1 + n_2 = n_3 + n_4. \end{cases}$$

Let us come back to physical interpretation of these results. Resonantly interacting waves will change their amplitudes according to Sys.(10). In this case the role of ZK spectra $k^\alpha,\ \alpha < 0,$ is played by the interaction coefficient $Z \sim k^\alpha,\ \alpha > 0$ (see Fig.1). Non-interacting waves will have constant amplitudes (they are not shown in Fig.1). In the next Section we demonstrate some examples of different wave systems whose behavior is explained by the theory of laminated turbulence.

4 Examples

- **Ex.1** Turbulence of capillary waves (dispersion function $\omega^2 = k^3$, three-wave interactions) was studied in [14] in the frame of simplified dynamical equations for the potential flow of an ideal incompressible fluid. Coexistence of ZK energy spectra and a set of discrete waves with constant amplitudes was clearly demonstrated. The reason why

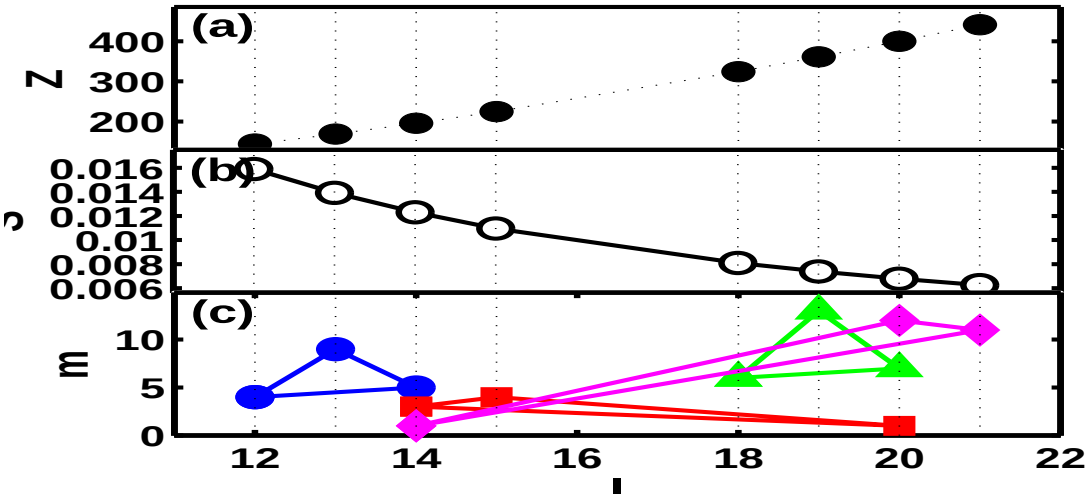


Figure 1: Two layers of turbulence are shown symbolically. Low panel: 2D-domain in spectral space, nodes of the integer lattice are connected with the lines which correspond to 3-wave resonant interactions of the discrete layer. Middle panel: ZK energy spectrum $k^{-3/2}$ with "the holes"in the nodes of the integer lattice. Upper panel: Interaction coefficient $Z \sim k^{3/2}$ in the nodes of the integer lattice.

in this case the amplitudes are constant is following: equation

$$k_1^{3/2} + k_2^{3/2} = k_3^{3/2}$$

has no integer solutions [8]. It means that there exist no three-wave resonant interactions among discrete capillary waves, they take no part in the energy transfer through the wave spectrum and just keep their energy at the low enough level of nonlinearity.

- **Ex.2** Similar numerical simulations [15] with gravity waves on the surface of deep ideal incompressible fluid (dispersion function $\omega^4 = k^2$, four-wave interactions) show again coexistence of ZK energy spectra and a set of discrete waves. But in this case waves amplitudes are not constant any more, discrete waves do exchange their energy and in fact play major role in the energy transfer due to the fact that equation

$$k_1^{1/2} + k_2^{1/2} = k_3^{1/2} + k_4^{1/2}$$

has many non-trivial integer solutions [16] (it is important in this case that 2-dimensional waves are regraded, i.e. $k = |\vec{k}| = \sqrt{m^2 + n^2}$ with integer m, n).

- **Ex.3** Some recurrent patterns were found in different atmospheric data sets (rawindsonde time series of zonal wind, atmospheric angular momentum, atmospheric pressure, etc.) These large-scale quasi-periodic patterns appear repeatedly at fixed geographic locations, have periods 10-100 days and are called intra-seasonal oscillations in the Earth atmosphere. In [17] Eq.(6) (dispersion function $\omega = -2m/n(n+1)$, three-wave interactions) is studied which is classically regarded as a basic model of climate variability in the Earth atmosphere. It is shown that a possible explanation of the intra-seasonal oscillations can be done in terms of a few specific, resonantly interacting triads of planetary waves, isolated from the system of all the rest planetary waves.

  Remark In contrast to the first two examples, in this case *only discrete layer of turbulence exists.* Indeed, while $\omega = -2m/n(n+1)$ is a rational function, any ratio $\omega(m_i, n_i)/\omega(m_j, n_j)$ is a rational number and KAM-theory is not applicable.

- **Ex.4** A very challenging idea indeed is to use the theory of laminated turbulence to explain so-called anomalous energy transport in tokamaks. Turbulent processes responsible for these effects are usually described as H- and L-modes and ELMs (high, low and edge localized modes consequently). Interpretation of the known experimental results in terms of non-resonant (H), resonant (L) and resonant with small non-zero resonance width (ELM) modes gives immediately a lot of interesting results. In this case *non-resonant* discrete waves are of major interest because they will keep their energy as in **Ex.1**, for a substantial period of time. This approach allows to get two kind of results: 1) to describe the set of the boundary conditions providing no resonances at all - say, if the ratio of the sides in the rectangular

domain is 2/7, no exact resonances exist; or 2) to compute explicitly all the characteristics of the non-resonant waves (wave numbers, frequencies, etc.) for given boundary conditions. Some preliminary results are presented in [18], in the frame of Hasegawa-Mima equation in a plane rectangular domain with zero boundary conditions (dispersion function $\omega = 1/\sqrt{n^2+m^2}$, three-wave interactions).

It is important to remember that a choice of initial and/or boundary conditions for a specific PDE might lead to a substantially different form of dispersion function and consequently to the qualitatively different behavior of the wave system. For instance, **Ex.3** and **Ex.4** are described by the same Eq.(6) regarded on a sphere (*rational dispersion function, only discrete layer of turbulence exists*) and in a rectangular (*irrational dispersion function, both layers exist*) respectively. For some equations, a special choice of boundary conditions leads to transcendental dispersion functions.

5 Open Questions

We have seen that the main algebraic object of the wave turbulence theory is the equation

$$\omega(m_1,n_1)+\omega(m_2,n_2)+...+\omega(m_s,n_s)=0 \tag{14}$$

where dispersion function ω is a solution of a dispersive evolutionary LPDE with $|\frac{\partial^2\omega}{\partial k_i \partial k_j}| \neq 0$. So defined class of dispersion functions includes rational, irrational or transcendental function, for instance

$$\omega(k)=\alpha k-\beta k^3,\ \ \omega^4(k)=\alpha^2k^2+\beta^2,\ \ \omega=m/(k+1),\ \ \omega=\tanh\alpha k,\ \cdots$$

where α and β are constants and $k=\sqrt{m^2+n^2}$. Continuous layer of the wave turbulence, that is, with $m_i, n_i \in \mathbb{R}$, is well studied. On the contrary, there are still a lot of unanswered questions concerning the discrete layer of turbulence, $m_i, n_i \in \mathbb{Z}$, and we formulated here just a few of them.

Eq.(14) can be regarded as *a summation rule* for the rational points of the manifold given by ω. These manifolds have very special structure -

namely, they can be transformed into a one-parametric family of simpler manifolds, namely (12) into (13). This situation is general enough, the definition of classes can be generalized for a given $c \in \mathbb{N}, c \neq 0, 1, -1$ considering algebraic numbers $k^{1/c}, k \in \mathbb{N}$ and their unique representation

$$k_c = \gamma q^{1/c}, \gamma \in \mathbb{Z}$$

where q is a product

$$q = p_1^{e_1} p_2^{e_2} ... p_n^{e_n},$$

while $p_1, ...p_n$ are all different primes and the powers $e_1, ...e_n \in \mathbb{N}$ are all smaller than c. Then algebraic numbers with the same q form the class Cl_q and the following statement holds:

The equation $a_1 k_1 + a_2 k_2 ... + a_n k_n = 0, a_i \in \mathbb{Z}$ *where each* $k_i = \gamma_i q_i^{1/c}$ *belongs to some class* $q_i \in q_1, q_2 ... q_l, \quad l < n$ *with ... is equivalent to a system*

$$\begin{cases} a_{q_1,1}\gamma_{q_1,1} + a_{q_1,2}\gamma_{q_1,2} + ... + a_{q_1,n_1}\gamma_{q_1,n_1} = 0 \\ a_{q_2,1}\gamma_{q_2,1} + a_{q_2,2}\gamma_{q_2,2} + ... + a_{q_2,n_2}\gamma_{q_2,n_2} = 0 \\ ... \\ a_{q_l,1}\gamma_{q_l,1} + a_{q_l,2}\gamma_{q_l,2} + ... + a_{q_l,nl}\gamma_{q_l,nl} = 0 \end{cases} \tag{15}$$

The questions are: what is the geometry underlying this parametrization? What is known about these sort of manifolds? What other properties of the resonance manifold are defined by a given summation rule? What additional information about these manifolds gives us the fact that they have many (often infinitely many) integer points?

Another group of questions concerns transcendental dispersion functions. All our examples were constructed for rational and irrational dispersion functions and the theoretical results were based on some classical theorems on the linear independence of some sets of algebraic numbers. In the case of a transcendental dispersion function like $\omega = \tanh \alpha k$ similar reasoning

can be carried out using the theorem on the linear independence of the exponents but it is not done yet. The question about special functions in this context is completely unexplored though very important. For instance, a dispersion function for capillary waves in a circle domain is described by Bessel function. Any results on their resonant interactions will shed some light on the nature of Faraday instability.

One of the most interesting questions about the resonance manifolds would be to study their invariants, i.e. some new function $f = f(m, n, \omega)$ such that $\omega_1 + \omega_2 = \omega_3$ implies $f_1 + f_2 = f_3$. An example of this sort of analysis is given in [19] for 3-wave interactions of drift waves with $\omega = \alpha x/(1 + y^2)$ but for real-valued wave vectors, $x, y \in \mathbb{R}$. Existence of the invariants is important because it is directly connected with the integrability of corresponding nonlinear PDE. Coming back to the physical language this means that the wave system possesses some additional conservation law.

6 Summary

We formulated above some interesting questions of the theory of laminated wave turbulence in the terms of integer points on the resonant manifolds. A very important task would be to develop fast algorithms to compute these integer points. The parametrization property allows to construct specific algorithms for a given dispersion function as it was done in [16] for 4-wave interactions of gravity waves, $\omega = (m^2 + n^2)^{1/4}$. The work on the generic algorithm for a dispersion function $\omega = \omega(k)$ ($k = \sqrt{m^2 + n^2}$ or $k = \sqrt{m^2 + n^2 + l^2}$ with integer $m,\ n,\ l$) is on the way [13] but it does not cover even simple cases of rational dispersion functions like $\omega = m/n^2$ not mentioning transcendental dispersion functions.

ACKNOWLEDGMENTS

Author acknowledges support of the Austrian Science Foundation (FWF) under projects SFB F013/F1304. Author expresses special gratitude to the anonymous referee for the valuable remarks.

References

[1] Arnold, V.I. *Lectures on Partial Differential Equations*, Springer Series: Universitext, 2004

[2] Whitham, G.B. *Linear and Nonlinear Waves*, Wiley Series in Pure and Applied Mathematics, 1999

[3] Holmes, M.A. *Introduction to Perturbation Methods*, Springer-Verlag, New York, 1995

[4] Zakharov, V.E. and L'vov, V.S. and Falkovich, G. *Kolmogorov Spectra of Turbulence*, Springer Series in Nonlinear Dynamics, 1992

[5] Zakharov, V.E. and Filonenko, N.N.: Weak turbulence of capillary waves. *Zh. Prikl. Mekh. Tekh. Phys.* 4(5), pp.62-67, 1967

[6] Kartashova, E.A.: Partitioning of ensembles of weakly interacting dispersing waves in resonators into disjoint classes, *Physica D*, 46, 43-56, 1990

[7] Kartashova, E.A.: Weakly nonlinear theory of finite-size effects in resonators, *Phys. Rev. Letters*, 72, 2013-2016, 1994

[8] Kartashova, E.A.: Wave resonances in systems with discrete spectra, In: Zakharov, V.E. (Ed.), *Nonlinear Waves and Weak Turbulence*, Series Advances in the Mathematical Sciences, AMS, 95-129, 1998

[9] Kolmogorov, A.N.: On the conservation of conditionally periodic motions for a small change in Hamiltons function, *Dokl. Akad. Nauk SSSR*, 98, 527, 1954. English translation in: *Lecture notes in Physics 93*, Springer, 1979

[10] Arnold, V.I.: Proof of a theorem by A.N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian, *Russian Math. Surveys*, 18, 9, 1963

[11] Moser, J.: On invariant curves of area preserving mappings of an annulus, *Nachr. Akad. Wiss. Gtt., Math. Phys. Kl.*, 1-20, 1962

[12] Kartashova, E.A.: A model of laminated wave turbulence, *JETP Letters*, 83(7), 341-345, 2006

[13] Kartashova E., Kartashov A.: Laminated Wave Turbulence: Generic Algorithms I, *E-print* arXiv.org:math-ph/0609020 . Submitted to *J. Comp. Phys.*, 2006

[14] Pushkarev, A.N. and Zakharov, V.E.: Turbulence of capillary waves - theory and numerical simulations, *Physica D*, 135, 98-116, 2000

[15] Zakharov, V.E. and Korotkevich, A.O. and Pushkarev, A.N. and Dyachenko, A.I.: Mesoscopic wave turbulence, *JETP Letters*, 82(8), 487-490, 2005
[16] Kartashova, E.: Fast Computation Algorithm for Discrete Resonances among Gravity Waves, To appear in *J. Low Temp. Phys.*. *E-print* arXiv.org:nlin/0605067, 2006
[17] Kartashova, E. and L'vov, V.: Large-scale variability in the Earth atmosphere. *E-print* arXiv.org:nlin/0606058. Shortened version submitted to *Phys. Rev. Letters*, 2006
[18] Kartashova, E.: Kinetic equation and Clipping - two limits of wave turbulence theory, *E-print* arXiv.org:math-ph/0509006, 2005
[19] Balk, A.M. and Nazarenko, S.V. and Zakharov, V.E.: New invariant for drift turbulence, *Phys. Letters A*, 152(5-6), 276-280, 1991

Invariant Form of BK-factorization and its Applications

Kartashova E. [1] Rudenko O. [2]

[1] J. Kepler University, RISC, lena@risc.uni-linz.ac.at
[2] Weizmann Institute of Science, Dep. Chem. Physics, oleksii@weizmann.ac.il

Abstract

We present invariant formulation of the Beals-Kartashova factorization procedure (BK-factorization) which allows to factorize simultaneously classes operators equivalent to the initial one under gauge transformations. We compare two factorization methods - Hensel descent and Beals-Kartashova factorization procedure (BK-factorization) aiming to emphasize the constructiveness of BK-factorization. We also show the possibility to use the same procedure for the construction of the approximate factorization of LPDE in the case when corresponding LPDO is not exactly factorizable and point out the problems to solve while factoring approximately noncommutative polynomials.

Keywords: linear partial differential operator, Hensel descent, BK-factorization, invariant transformations, approximate factorization

1 Introduction

Factorization of ordinary and partial linear differential operators (LODOs and LPDOs) is a very well-studied problem and a lot of pure existence theorems are known. For LODOs it is proven that a factorization is unique up to factor permutation while for LPDOs even uniqueness is not true any more and in fact parametric families of factorizations can be constructed for a given LPDO as will be demonstrated below.
First **constructive method** of factoring second order hyperbolic LPDO

in the form

$$\partial_x\partial_y + a\partial_x + b\partial_y + c, \tag{1}$$

belong to Laplace who formulated it in terms of invariants $\hat{a} = c - ab - a_x$ and $\hat{b} = c - ab - b_y$ now called Laplace invariants. An operator (1) is factorizable if at least one of its Laplace invariants is equal to zero. Various algorithms are known now for factoring LPDOs over different differential fields beginning with the simplest field of rational functions [1].

Recently two papers ([2], [3]) on factoring arbitrary order LPDOs have been published. In [2] a modification of well-known Hensel lifting algorithm (see, for instance, [4]) is presented and **sufficient conditions** for the existence of intersection of principal ideals are given. These results are applied then to re-formulate the factorization formulae for second and third order operators from the ring $D = \mathbb{Q}(x, y)[\partial_x, \partial_y]$ obtained by Miller (1932) in terms of principal intersections.

In [3] **necessary and sufficient conditions** are given for factoring of bivariate LPDOs of arbitrary order with coefficients being arbitrary smooth functions. In [5] it was shown that this procedure called now BK-factorization principally can not be generalized on the case of more than two variables. In was also shown that conditions of factorization found in [3] are invariants under gauge transformation and classical Laplace invariants are particular case of this generalized invariants.

In this paper we re-formulate BK-factorization in more suitable for applications invariant form and illustrate it with a few examples, give a sample of the symbolical implementation of this method in MATHEMATICA and also discuss some possibilities to use this method for approximate factorization of LPDOs.

2 Hensel descent and BK-factorization

Hensel descent, the latest known to us new method published before BK-factorization, has been considered as constructive. There also exist an opinion that BK-factorization is a minor generalization of the Hensel de-

scent. To clarify the matter, we begin this paper with comparative analysis of these two methods in order to show merits and draw-backs of each of them.

- In [3] factorization of a bivariate LPDO is looked for in the form

$$A_n = \sum_{j+k\le n} a_{jk}\partial_x^j\partial_y^k = (p_1\partial_x + p_2\partial_y + p_3)\Big(\sum_{j+k<n} p_{jk}\partial_x^k\partial_y^j\Big) \qquad (2)$$

  and in [2] - operator of $m \ge 2$ independent variables is regarded and for $m = 2$ factorization is looked for in the form

$$A_n = \sum_{j+k\le n} a_{jk}\partial_x^j\partial_y^k = \Big(\sum_{j+k\le l} p_{jk}\partial_x^k\partial_y^j\Big)\Big(\sum_{j+k\le r} p_{jk}\partial_x^k\partial_y^j\Big) \text{ with } l+r = n. \qquad (3)$$

- In [3] coefficients a_{jk} are arbitrary smooth functions, for instance trigonometric functions; in [2] conditions for reducibility of an operator are studied when "coefficients are from a universal field of zero characteristic" , while "studying factorization algorithms we will assume that the input operators are from the ring $\mathbb{Q}(x_1, ..., x_m)[\partial_1, ..., \partial_m]$" This suggestion is necessary:"From now on the coefficients of a given second-order operator are assumed to be from the base field $\mathbb{Q}(x, y)$. This is necessary if the goal is to obtain constructive answers allowing to factorize large classes of operator" ([2], Sec.3);"In this section we study third-order LPDOs from the ring $D = \mathbb{Q}(x, y)[\partial_x, \partial_y]$." ([2], Sec.4).

- In [3] it was shown that in generic case factorization can be constructed explicitly and algebraically, while in [2] (Sec.5) it is concluded that "the factorization problem for second- and third-order differential operators in two variables has been shown to require the solution of a partial Riccati equation, which in turn requires to solve a general first-order ODE and possibly ordinary Riccati equation. The bottleneck for designing a factorization algorithm for a LPDO is

general first-order ODE which make the full problem intractable at present because in general there are no solution algorithm available"

- In [5] it is pointed out that BK-factorization procedure has to be modified in some way (presently unknown to authors) in order to proceed with operators of more than 2 independent variables, while in [2] (Sec.5) it is written that "some of the results described in this article may be generalized to any number of independent variables"
- In [5] it shown that BK-factorization procedure gives rise to construction of the whole class of generalized invariants particular case of them being classical Laplace invariants. This leads to a possibility to factorize simultaneously the whole class of operators equivalent under gauge transformation (see next Section) while Hensel descent is used for factoring of a one specific operator.

We summarize all this in the Table below.

Property / Method	BK-factorization	Hensel descent
Order of operator	n	n
Coefficients of operator	arbitrary smooth functions	rational functions
Number of variables	2	possibly > 2
Conditions	necessary and sufficient	sufficient
Form of factors	of order 1 and $(n-1)$	of order k and $(n-k)$
Formulated as	explicit formulae	ideals intersection

In the next Sections we demonstrate some other interesting properties of BK-factorization - first of all, that it has invariant form and can be used therefore to factorize simultaneously the whole classes of equivalent LPDOs. Second, the use of this invariant form of BK-factorization for construction of approximate factorization for LPDEs to be solved numerically.

3 Invariant Formulation

We present here briefly main ideas presented in [3], [5] beginning with the definition of equivalent operators.

Definition. *The operators A, $\tilde{A}$ are called equivalent if there is a gauge transformation that takes one to the other:*

$$\tilde{A}g = e^{-\varphi}A(e^{\varphi}g).$$

BK-factorization is then pure algebraic procedure which allows to to construct explicitly a factorization of an arbitrary order LPDO A in the form

$$A := \sum_{j+k\le n} a_{jk}\partial_x^j\partial_y^k = L \circ \sum_{j+k\le(n-1)} p_{jk}\partial_x^j\partial_y^k$$

with first-order operator $L = \partial_x - \omega\partial_y + p$ where ω is **an arbitrary simple root** of the characteristic polynomial

$$\mathcal{P}(t) = \sum_{k=0}^{n} a_{n-k,k}t^{n-k}, \quad \mathcal{P}(\omega) = 0. \tag{4}$$

Factorization is possible then for each simple root $\tilde{\omega}$ of (4) **iff**

for $n = 2$ $\Rightarrow$ $l_2 = 0$,

for $n = 3$ $\Rightarrow$ $l_3 = 0$, & $l_{31} = 0$,

for $n = 4$ $\Rightarrow$ $l_4 = 0$, & $l_{41} = 0$, & $l_{42} = 0$,

and so on. All functions l_2, l_3, l_{31}, l_4 l_{41}, $l_{42}, ...$ are explicit functions of a_{ij} and $\tilde{\omega}$.

Theorem. *All $l_2, l_3, l_{31}, ...$ are* **invariants** *under gauge transformations.*

Definition. *Invariants $l_2, l_3, l_{31}, ...$ are called generalized invariants of a bivariate operator of arbitrary order.*

In particular case of the operator (1) its generalized invariants coincide with Laplace invariants.

Corollary. *If an operator A is factorizable, then all operators equivalent to it, are also factorizable.*

As the first step of BK-factorization, coefficients p_{ij} are computed as solutions of some system of algebraic equations. At the second step, equality to zero of all generalized invariants $l_{ij} = 0$ has to be checked so that

no differential equations are to be solved in generic case. Generic case corresponds to a simple root of characteristic polynomial, and each simple root generates corresponding factorization. Moreover, putting some restrictions on the coefficients of the initial LPDO $a_{i,j}$ as functions of x and y, one can describe *all* factorizable operators in a given class of functions (see Example 5.3 in [3]). The same keeps true for all operators equivalent to a given one. Equivalent operators are easy to compute:

$$e^{-\varphi}\partial_x e^{\varphi} = \partial_x + \varphi_x, \quad e^{-\varphi}\partial_y e^{\varphi} = \partial_y + \varphi_y,$$

$$e^{-\varphi}\partial_x\partial_y e^{\varphi} = e^{-\varphi}\partial_x e^{\varphi} e^{-\varphi}\partial_y e^{\varphi} = (\partial_x + \varphi_x) \circ (\partial_y + \varphi_y)$$

and so on. Some examples:

- $A_1 = \partial_x\partial_y + x\partial_x + 1 = \partial_x(\partial_y + x), \quad l_2(A_1) = 1 - 1 - 0 = 0;$
- $A_2 = \partial_x\partial_y + x\partial_x + \partial_y + x + 1, \quad A_2 = e^{-x} A_1 e^{x}; \quad l_2(A_2) = (x+1) - 1 - x = 0;$
- $A_3 = \partial_x\partial_y + 2x\partial_x + (y+1)\partial_y + 2(xy + x + 1), \quad A_3 = e^{-xy} A_2 e^{xy};$ $l_2(A_3) = 2(x + 1 + xy) - 2 - 2x(y+1) = 0;$
- $A_4 = \partial_x\partial_y + x\partial_x + (\cos x + 1)\partial_y + x\cos x + x + 1, \quad A_4 = e^{-\sin x} A_2 e^{\sin x};$ $l_2(A_4) = 0.$

Generic case which can be treated pure algebraically by BK-factorization corresponds to **a simple root of characteristic polynomial**. Each multiple root leads to necessity of solving some Ricatti equation(s) (RE). If appeared RE happens to be solvable, such a root generates a parametric family of factorizations for a given operator. For instance, well-known Landau operator

$$\partial^3_{xxx} + x\partial^3_{xxy} + 2\partial_{xx} + (2x+2)\partial^2_{xy} + \partial_x + (2+x)\partial_y$$

has characteristic polynomial with one distinct root $\omega_1 = -x$ and one double root $\omega_{2,3} = 0$. Factorization then has form

$$(\partial_x + r)(\partial_x - r + 2)(\partial_x + x\partial_y)$$

where r is a solution of Ricatti equation

$$1 - 2r + \partial_x(r) + r^2 = 0$$

which is easily solvable:

$$r = 1 + \frac{1}{x + Y(y)}$$

with arbitrary smooth function $Y(y)$ of one variable y so that factorization has form

$$A = (\partial_x + 1 + \frac{1}{x + Y(y)})(\partial_x + 1 - \frac{1}{x + Y(y)})(\partial_x + x\partial_y).$$

Notice that to factorize **an ordinary differential operator** it is always necessary to solve some RE. Nevertheless, just formal application of BK-factorization will produce all the linear factors in the case when corresponding RE are solvable. For instance, the factorization has been constructed in [6]

$$x\partial_{xxx} + (x^2 - 1)\partial_{xx} - x\partial_x + \frac{2}{x^2} - 1 = (\partial_x + \frac{x^2 - 1}{x})(x\partial_x - \sqrt{2})(\partial_x + \frac{\sqrt{2} - 1}{x}).$$

while both RE appearing at the intermediate steps are solvable.

These two last examples show the main difference between factorizing of ordinary and partial differential operators - LODO has always unique factorization while LPDO may have many. An interesting question here would be to compute the exact number of all possible factorizations of a given LPDO into all linear factors (its upper bound is, of course, trivial: $n!$). A really challenging task in this context would be to describe some additional conditions on the coefficients of an initial operator which lead to solvable RE.

4 Left and Right Factors

Factorization of an operator is the first step on the way of solving corresponding equation. But for solution we need **right** factors and BK-factorization constructs **left** factors which are easy to construct. On the

other hand, the existence of a certain right factor of a LPDO is equivalent to the existence of a corresponding left factor of the *transpose* of that operator. Moreover taking transposes is trivial algebraically, so there is also nothing lost from the point of view of algorithmic computation. In our paper [3] we just used the transpose in one example assuming that the operation is well known. Apparently, it is not, and some new works appeared quite recently on the construction of a complicated new right-factor algorithms (i.e. [7] and others) for bivariate operators of order 2 and 3 though our left-factor algorithm gives the explicit formulae for arbitrary order n. This is the reason why we decided to include the explicit formula for the adjoint in our present text.

Definition. *The transpose A^t of an operator $A = \sum a_\alpha \partial^\alpha,\ \partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$. is defined as*

$$A^t u = \sum (-1)^{|\alpha|} \partial^\alpha (a_\alpha u).$$

and the identity

$$\partial^\gamma (uv) = \sum \binom{\gamma}{\alpha} \partial^\alpha u \, \partial^{\gamma-\alpha} v$$

implies that

$$A^t = \sum (-1)^{|\alpha+\beta|} \binom{\alpha+\beta}{\alpha} (\partial^\beta a_{\alpha+\beta}) \partial^\alpha.$$

Now the coefficients are

$$A^t = \sum \widetilde{a}_\alpha \partial^\alpha,$$

$$\widetilde{a}_\alpha = \sum (-1)^{|\alpha+\beta|} \binom{\alpha+\beta}{\alpha} \partial^\beta (a_{\alpha+\beta}).$$

with a standard convention for binomial coefficients in several variables, e.g. in two variables

$$\binom{\alpha}{\beta} = \binom{(\alpha_1, \alpha_2)}{(\beta_1, \beta_2)} = \binom{\alpha_1}{\beta_1} \binom{\alpha_2}{\beta_2}.$$

In particular, for order 2 in two variables the coefficients are

$$\widetilde{a}_{jk} = a_{jk}, \quad j+k = 2; \widetilde{a}_{10} = -a_{10}+2\partial_x a_{20}+\partial_y a_{11}, \widetilde{a}_{01} = -a_{01}+\partial_x a_{11}+2\partial_y a_{02},$$

$$\widetilde{a}_{00} = a_{00} - \partial_x a_{10} - \partial_y a_{01} + \partial_x^2 a_{20} + \partial_x \partial_x a_{11} + \partial_y^2 a_{02}.$$

For instance, the operator

$$\partial_{xx} - \partial_{yy} + y\partial_x + x\partial_y + \frac{1}{4}(y^2 - x^2) - 1 \tag{5}$$

is factorizable as

$$\left[\partial_x + \partial_y + \tfrac{1}{2}(y - x)\right] \left[...\right]$$

and its transpose A_1^t is factorizable then as

$$\left[...\right] \left[\partial_x - \partial_y + \tfrac{1}{2}(y + x)\right].$$

Implementation of the BK-factorization for bivariate operators of order $n \leq 4$ is therefore quite straightforward and has been done in MATHEMATICA while all roots of characteristic polynomial are known in radicals. For instance, for the operator (5) with 2 simple roots we get one factorization

$$\left[\partial_x - \partial_y + \tfrac{1}{2}(y + x)\right] \left[\partial_x + \partial_y + \tfrac{1}{2}(y - x)\right].$$

corresponding to the first root while in the case of the second root, generalized invariant is equal to 2.

If $n \geq 5$ the problem is generally not solvable in radicals and very simple example of non-solvable case is: $x^5 - 4x - 2 = 0$. Thus, to find solutions in radical for $n > 4$ one needs some constructive procedure of finding solvable Galois group but this lies beyond the scope of the present paper.

5 Approximate Factorization

An interesting possible application of the invariant form of BK-factorization is to use it for construction of approximate factorization of a given LPDE,

in the case when exact factorization of corresponding LPDO does not exists. Indeed, as a results of BK-factorization one gets

(1) **factorization coefficients** $\{p_{ij}^{(i)}\}$ for the i-th factorization of a given operator, and

(2) **generalized invariants** $l_2^{(i)} p_{ij}^{(i)}$, with all $p_{ij}^{(i)}, l_{(k_j)}^{(i)}$ being explicit functions of the coefficients of initial operator a_{ij}.

In numerical simulations coefficients a_{ij} of the equation are always given with some non-zero accuracy, say $\varepsilon > 0$, which means that it is enough to construct an approximate factorization in the following sense. One has to find restrictions on the coefficients a_{ij} of an initial LPDO which provide $|l_{k_j}^{(j)}| < \varepsilon$ with a given accuracy $0 < \varepsilon << 1$. Many different strategies are possible here, we just give a brief sketch of two approaches we are working on right now:

5.1 Quantifier Elimination

We illustrate this idea on the simple example of a hyperbolic operator $\partial_{xx} - \partial_{yy} + a_{10}\partial_x + a_{01}\partial_y + a_{00}$ with linear polynomial coefficients.

What we have is:

$$a_{00}(x, y) = b_3 x + b_2 y + b_1,$$

$$a_{10}(x, y) = c_3 x + c_2 y + c_1,$$

$$a_{01}(x, y) = d_3 x + d_2 y + d_1;$$

a function constructed from general invariants

$$\mathcal{R} = \frac{s_3 - s_2}{2} + \frac{(s_3 x + s_2 y + s_1)^2}{4}$$

with $s_i = c_i - d_i$.

What we need is:

To find some function(s) $F = F(a_{ij})$ such that if $F(a_{i,j}) = 0$, then

$$-\varepsilon < a_{00} - \mathcal{R} < \varepsilon, \quad \text{for some constant} \quad 0 < \varepsilon << 1,$$

i.e. to find some conditions on the initial polynomials which provide that function $\mathcal{R}$ differs not too much from one these polynomials, namely a_{00}.

Notice that simple symmetry considerations allowed us to reduce number of variables needed for CAD calculations. Initially we had 9 variables $b_3, b_2, b_1, c_3, c_2, c_1, d_3, d_2, d_1$ but in fact it is enough to regard only 6 variables $s_1, s_2, s_3, b_1, b_2, b_3$. Nevertheless, the computation time may become crucial while using this approach due to the substantial number of variables. On the other hand, this approach allows us to work generally on the operator level including initial and/or boundary conditions first at some later stage.

5.2 Auxiliary Operator

Another approach is to construct a new auxiliary operator with coefficients $\widetilde{a}_{ij} = f(x,y))a_{ij}$ for all or for some of the coefficients a_{ij} of the initial operator, keep invariants (almost) equal to zero and find function(s) $f(x,y)$ minimizing the differences between the coefficients of initial and new operators. In this way an auxiliary operator is constructed which can be regarded as an approximate operator for the initial operator. Of course, it does not mean that solutions of the initial and approximate operators will be also close but simple properties of linear operators show that it is necessary (but not sufficient!) step on the way of construction of a good approximate solution of a given LPDE - in the case of a well-posed problem, of course. In particularly, it means that one has to introduce proper metrics in the space of operators **and** in the space of solutions. Choice of the both metrics and of a function f will depend on (1) coefficients of the initial operator; (2) class of functions in which we are looking for a solution; (3) initial and/or boundary conditions.

To demonstrate all this let us regard two different un-factorizable modifications of the operator (5):

$$A = \partial_{xx} - \partial_{yy} + y\partial_x + x\partial_y + \frac{1}{2}(y^2 - x^2) - 1 \tag{6}$$

with $l_2(A) = \frac{1}{4}(y^2 - x^2)$ and

$$B = \partial_{xx} - \partial_{yy} + \sin y\ \partial_x + \cos x\ \partial_y + \frac{1}{2}(\sin^2 y - \cos^2 x) \tag{7}$$

with $l_2(B) = \frac{1}{2}(\cos y - \sin x)$ (see Fig.1). One can see immediately that $l_2(B)$ is a bounded function of two variables and $l_2(A)$ is an unbounded. This means that quite different choice of function f is needed for these two cases in order to minimize the invariants. Influence of initial/boundary conditions is now also very clear - for instance, best approximation of $l_2(B)$ can be obtained in the narrow belts of the lines parallel to one of the coordinate axis while for $l_2(A)$ these directions are in no way special.

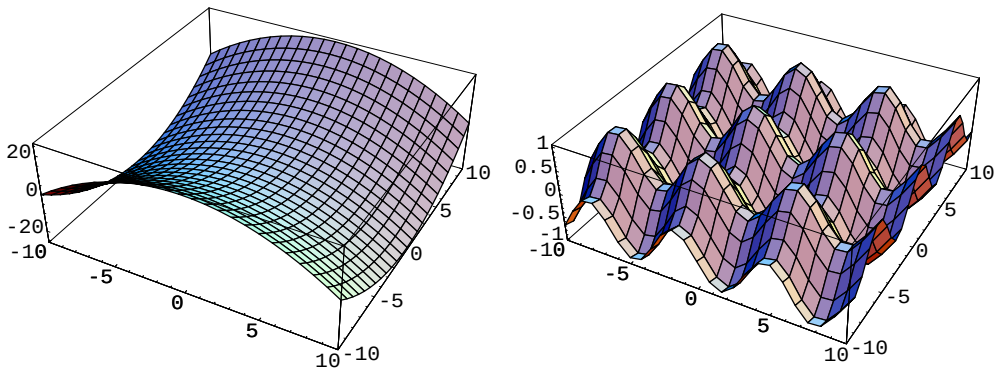


Figure 1: Invariant $l_2(A) = \frac{1}{4}(y^2 - x^2)$ (left) and invariant $l_2(B) = \frac{1}{2}(\cos y - \sin x)$ (right), in the domain $-10 \le x, y \le 10$

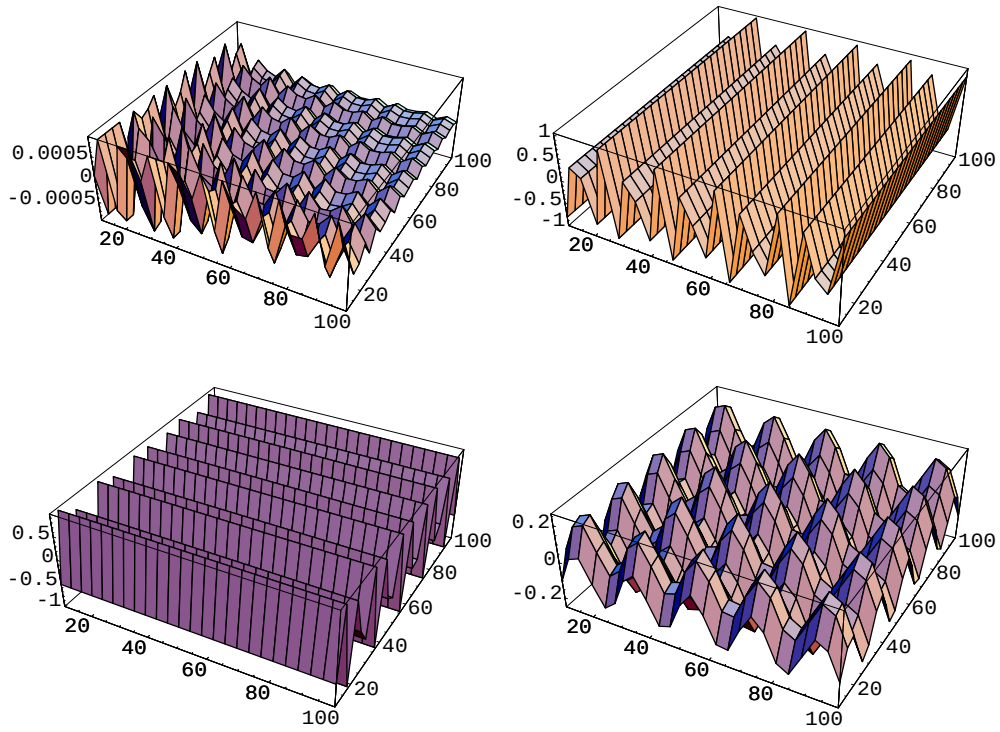


Figure 2: Upper panel: $l_2(\tilde{B})$ (left) and $a_{10} - \tilde{a}_{10}$ (right); lower panel: $a_{01} - \tilde{a}_{01}$ (left) and $a_{00} - \tilde{a}_{00}$ (right); in the domain $10 \leq x, y \leq 100$

To construct a sample of such an approximate factorization for the operator (7) we just suppose intuitively that auxiliary operator $\tilde{B}$ is "good" if its coefficients differ from the coefficients of (7) not much, and its invariant is small. Our MATHEMATICA implementation of the BK-factorization includes simple graphic functions to display the differences between all the parameters of the initial and auxiliary operators. A choice of the function $f(x, y) = \sin \frac{1}{xy}$ gives an auxiliary operator $\tilde{B}$ of the form

$$\begin{aligned} \tilde{B} &= \partial_{xx} - \partial_{yy} + \sin y \sin \frac{1}{xy}\, \partial_x + \cos x \sin \frac{1}{xy}\, \partial_y \\ &\quad + \frac{1}{2}(\sin^2 y - \cos^2 x) \sin \frac{1}{xy}. \end{aligned} \tag{8}$$

It is demonstrated at the Fig.2 that for $10 \leq x, y \leq 100$ operator $\tilde{B}$ gives good enough approximation and correspondingly approximate factorization of the initial operator B has form

$$B \sim \Big[\frac{1}{2}\Big(-\cos x \sin\frac{1}{xy} + \sin\frac{1}{xy}\sin y\Big) + \partial_x + \partial_y\Big] \cdot \Big[\frac{1}{2}\Big(\cos x \sin\frac{1}{xy} + \sin\frac{1}{xy}\sin y\Big) + \partial_x - \partial_y\Big]$$

with $|l_2(\tilde{B})| \sim 5 \cdot 10^{-4}$. On the other hand, in the domain $0.001 \leq x, y \leq 1$ qualitatively different approximation is needed while in this domain $|l_2(\tilde{B})| \sim 10^2$ (see Fig. 3).

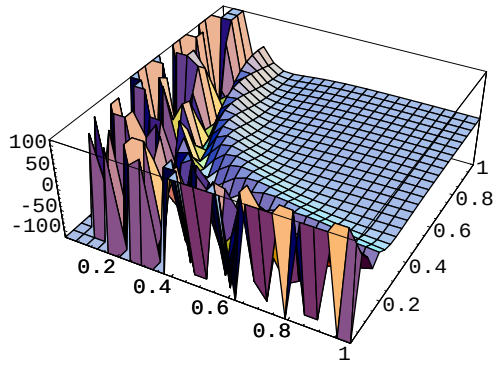


Figure 3: Invariant $l_2(\tilde{B})$ in the domain $0.0001 \leq x, y \leq 1$

Obviously, if we get enough approximate factorizations of the given LPDE with different solvable first-order factors we can write out explicitly general solution of the initial LPDE. Otherwise, one gets a chain of the linear first-order equations

$$A_{i0,n}\psi_0 = 0, \;\; A_{i0,n-1}\psi_1 = \psi_0, \;\; ....$$

to be solved numerically which is a great numerical simplification, of course, specially for higher order LPDEs. On the other hand, while performing

numerical simulations, one has to take into account a lot of other factors, first of all, initial and boundary conditions. It would be a nontrivial task to include them into the exact formulae given by BK-factorization. In order to estimate usefulness of this approach from numerical point of view we still have to answer all the questions concerning computation time, stability, computation error, etc. For instance, coming back to the example of approximate factorization given in the previous section, one have to estimate what is numerically more reasonable for a given set on initial and boundary conditions - to solve numerically the system of equations

$$\begin{cases} [\frac{1}{2}(\cos x \sin \frac{1}{xy} + \sin \frac{1}{xy} \sin y) + \partial_x - \partial_y] \circ \psi_0 = 0 \\ [\frac{1}{2}(-\cos x \sin \frac{1}{xy} + \sin \frac{1}{xy} \sin y) + \partial_x + \partial_y] \circ \psi_1 = \psi_0 \end{cases}$$

or one equation $B \circ \psi = 0$. Some answers can be given by the method presented in [9] where a symbolic approach is used to generate automatically finite difference schemes for LPDEs and to check their von Neumann stability. Some preliminary steps to be taken in this direction might be following: (1) to take a non-factorizable but solvable operator, for instance, $A_1 = \partial_x \partial_y + x\partial_x + 2$, then LPDE $A_1(\psi) = 0$, has general solution

$$\psi = -\partial_x \Big(X(x) e^{-xy} + \int e^{x(y'-y)} Y(y') dy' \Big)$$

with two arbitrary functions $X(x)$ and $Y(y)$; (2) to construct its approximate factorization $\widetilde{A}_1 = L_1 \circ L_2$; (3) to get computational schemes using [9] - for A_1 and $\widetilde{A}_1$; (4) compute both numerically; (5) to compare results for A_1 and $\widetilde{A}_1$ with the general solution for some classes of initial data and for a fixed choice of computational scheme.

6 Summary

We presented here invariant formulation of BK-factorization and formulated some ideas about using it for approximate factorization of LPDOs, i.e. non-commutative polynomials. The great number of results is known

on factorization of commutative polynomials (see, for instance, [10] and others) where measure of the difference between factorizable and non-factorizable polynomials can be introduced as a function of the coefficients of a polynomial. It is not the case for non-commutative polynomials the reason being that infinitesimal changes in the coefficients can change drastically the solution of the corresponding LPDE, as well as changes of the initial and boundary conditions. The problem of defining a reasonable measure for non-commutative polynomials is under the study.

ACKNOWLEDGMENTS

E.K. acknowledges the support of the Austrian Science Foundation (FWF) under projects SFB F013/F1304 and Prof. Langer and Prof. Engl for valuable discussions. O.R. acknowledges the support of the US-Israel Binational Science Foundation. Authors are grateful to the unanimous referee whose remarks have helped us to formulate our results more clearly.

References

[1] Beke, E. Die Irreducibilitt der homogenen linearen Differentialgleichungen, *Math. Annalen* 45, pp.278-300 (1894); Bronstein, M. An improved algorithm for factoring linear ordinary differential operators, in: J. von zur Gathen (Ed.), *Proc. ISSAC'94*, pp.336-340 (1994); and many others

[2] Grigoriev, D. and Schwarz, F. Factoring and Solving Linear Partial Differential Equations. *Computing* **73**, pp.179-197 (2004)

[3] Beals, R. and Kartashova, E.A. Constructively factoring linear partial differential operators in two variables. *TMPh* **145** (2), pp.1510-1523 (2005)

[4] Kaltofen, E. Factorization of polynomials, in: Buchberger, B., Collins, G.E. and Loos, R.(Eds.), *Computing supplementum* 4, pp.95-113 (1982)

[5] Kartashova, E.A. A Hierarchy of Generalized Invariants for Linear Partial Differential Operators. *TMPh*, 147(3), pp. 839-846 (2006)

[6] Kartashova, E. BK-factorization and Darboux-Laplace transformations, in: Arabnia, H.R. and Gravvanis, G.A. (Eds.), *Proc. CSC'05 (Proceedings of The 2005 International Conference on Scientific Computing)*, pp. 144-150, CSREA Press, USA (2005)

[7] Diaz, G.A.L. Symbolic Methods for Factoring Linear Differential Operators. *PhD Thesis*, Refs.: Prof. Winkler, F. and Prof. Pfalzgraf, J., RISC, Austria

(2006). Text on-line: http://apache2.risc.uni-linz.ac.at/publications/download/risc-2871/Thesis.pdf

[8] Kartashova, E., McCallum, S., Winkler, F. *Work in progress*, unpublished.

[9] Levandovskyy, V. A Symbolic Approach to Finite Difference Schemes and their von Neumann Stability.*Contributed talk* at SFB Status Seminar, Strobl, Austria. Text online: http://www.risc.uni-linz.ac.at (2006)

[10] Gao, S., Kaltofen, E., May, J., Yang, Z., Zhi., L. Approximate factorization of multivariate Polynomials via differential eqautions, in: Jaime Gutierrez (Ed.) *Proc. ISSAC'04*, pp.167-174 (2004)

Sufficient Set of Integrability Conditions of an Orthonomic System[1]

Extended abstract

Michal Marvan

Mathematical Institute of the Silesian University at Opava
Na Rybníčku 1, 746 01 Opava, Czech Republic

Given a system of PDE, Riquier theory starts with answering the question: "Which coefficients of the Taylor expansion of a solution can be chosen arbitrarily (parametric derivatives) and which are then uniquely determined by the system (principal derivatives)?" The answer is given in two parts: *prolongation*, which says how to assign values to the principal derivatives, and *completion*, which says how to cope with possible conflicts.

Current developments are driven by computer algebra applications, such as symmetry analysis. Existing installations can deal with systems containing hundreds of equations, which means that efficiency is an important issue. In [4] we deal with the problem of finding a small sufficient subset of integrability conditions – a subset of integrability conditions which when satisfied imply satisfaction of all.

To be more specific, let symbols u^k_μ denote derivatives, where u^k is a dependent variable and μ stands for a monomial over the set $\mathcal{X}$ of independent variables. Let $\prec$ be a ranking of the set of derivatives, i.e., $u^k \prec u^k_x$, and $u^k_\mu \prec u^k_\nu \Rightarrow u^k_{x\mu} \prec u^k_{x\nu}$ for all $x \in \mathcal{X}$ and monomials μ, ν.

Consider a system Σ of PDE resolved with respect to certain derivatives,

$$\Sigma = \{u^k_\mu = \Phi^k_\mu\},$$

[1] The author acknowledges the support from GAČR under grant 201/04/0538 and from the MŠMT under project MSM4781305904. Special thanks are due to V.P. Gerdt for encouragement and enlightening discussions.

where Φ^k_μ are functions of the other derivatives (including zeroth-order derivatives u^k) and independent variables x_i. Let $\operatorname{dom}\Sigma$ denote the set of all derivatives u^k_μ on the l.h.s. of Σ. The system Σ is said to be *orthonomic* with respect to the ranking $\prec$ if it is *algebraically reduced*, meaning that the r.h.s. Φ^k_μ do not depend on derivatives from $\operatorname{dom}\Sigma$; *one-to-one*, meaning that for every $u^k_\mu \in \operatorname{dom}\Sigma$ the system contains exactly one equation with u^k_μ on its l.h.s.; and *normal*, meaning that every $u^k_\mu \in \operatorname{dom}\Sigma$ is $\prec$-maximal in its equation.

Henceforth Σ is assumed to be orthonomic. We consider the infinite prolongation

$$\Sigma^\infty = \{u^k_{\mu\sigma} = D_\sigma \Phi^k_\mu\}$$

where σ is an arbitrary monomial over $\mathcal{X}$ and $D_x = \partial/\partial x + \sum_{k,\mu} u^k_{\mu x}\, \partial/\partial u^k_\mu$ is the usual total derivative. Derivatives from $\operatorname{dom}\Sigma^\infty$ are said to be *principal*, the others are said to be *parametric*.

In general, Σ^∞ is not orthonomic even if Σ is so. Deviations from being one-to-one are expressed by *integrability conditions*: all r.h.s for one and the same principal derivative have to reduce (see below) to one and the same expression. The above-mentioned completion procedure consists in augmenting the system with all its integrability conditions.

It is well known that it suffices to check a finite subset of integrability conditions. Classical Janet's method as well as Reid's method of minimal and supplementary conditions still leave considerable redundancy. In polynomial elimination theory (theory of Gröbner bases), substantial progress was due to exploatation of syzygy methods (Boulier [1]; Gebauer and Möller [3]; Carboara, Kreuzer and Robbiano [2]); to Riquier theory, syzygies were applied by Rust [6].

To introduce the results of [4], we first answer the question: "What is reduction?" An orthonomic Σ^∞ determines a submanifold $\mathcal{E}_{\Sigma^\infty}$ in the jet space J^∞ in a natural way (parametric derivatives become coordinates on $\mathcal{E}_{\Sigma^\infty}$). Then reduction is the restriction operator $S = |_{\mathcal{E}_{\Sigma^\infty}}$. But what if Σ^∞ is not orthonomic? We simply choose a subsystem Σ' of Σ^∞ that is orthonomic and satisfies $\operatorname{dom}\Sigma' = \operatorname{dom}\Sigma^\infty$; then reduction is $S = |_{\mathcal{E}_{\Sigma'}}$.

There are usually plenty of ways to choose Σ'. A purely theoretical construction is possible by transfinite induction in the well-ordered set $\operatorname{dom}\Sigma^\infty$. It is also quite easy to write a reduction algorithm working on a compute-on-demand basis. There are essentially two rules to be followed when computing the reduction Su^k_μ of a principal derivative u^k_μ: (a) if the system contains an equation $u^k_\mu = \Phi^k_\mu$, then $Su^k_\mu := \Phi^k_\mu$; (b) if we already know $Su^k_{\mu/x}$ for some independent variable x, then $Su^k_\mu := SD_xSu^k_{\mu/x}$.
Henceforth we fix an orthonomic subsystem Σ' of Σ^∞ such that $\operatorname{dom}\Sigma' = \operatorname{dom}\Sigma^\infty$. Let S denote the reduction $|_{\mathcal{E}_{\Sigma'}}$. For every derivative $u^k_\mu \in \operatorname{dom}\Sigma^\infty$ we introduce the set of all independent variables x such that $u^k_{\mu/x} \in \operatorname{dom}\Sigma^\infty$ and denote it by $\mathcal{X}^k_\mu$. Elements of this set can be interpreted as "directions" leading to u^k_μ from another principal derivative.
If in Σ there is an equation of the form $u^k_\mu = \Phi^k_\mu$ such that the set $\mathcal{X}^k_\mu$ is nonempty and $x \in \mathcal{X}^k_\mu$, then $\Phi^k_\mu = SD_xSu^k_{\mu/x}$ is an *integrability condition of the first kind* at the point u^k_μ, meaning that prolongation by rule (a) and rule (b) coincide. If $x, y \in \mathcal{X}^k_\mu$, then $SD_xSu^k_{\mu/x} = SD_ySu^k_{\mu/y}$ is an *integrability condition of the second kind* at the point u^k_μ, meaning that prolongations by rule (b) coincide for any two directions.

Construction 1. (sufficient set of integrability conditions) For every principal derivative u^k_μ, let $\sim$ denote the symmetric relation in the set $\mathcal{X}^k_\mu$ given by

$$x \sim y \Leftrightarrow u^k_{\mu/xy} \in \operatorname{dom}\Sigma^\infty.$$

Let $\approx$ denote the equivalence relation obtained as the reflexive and transitive closure of the relation $\sim$.
For every $u^k_\mu \in \operatorname{dom}\Sigma$ with nonempty $\mathcal{X}^k_\mu$, choose one integrability condition of the first kind $\Phi^k_\mu = SD_xSu^k_{\mu/x}$, where $x \in \mathcal{X}^k_\mu$ is arbitrary.
For every $u^k_\mu \in \operatorname{dom}\Sigma^\infty$ such that the set $\mathcal{X}^k_\mu$ contains s equivalence classes $[x_1]_\approx, \ldots, [x_s]_\approx$ with $s > 1$, choose arbitrary representatives $x_1, \ldots, x_s$ of these classes and consider integrability condition of the second kind in the form of a chain of equations

$$SD_{x_1}Su^k_{\mu/x_1} = SD_{x_2}Su^k_{\mu/x_2} = \ldots = SD_{x_s}Su^k_{\mu/x_s}.$$

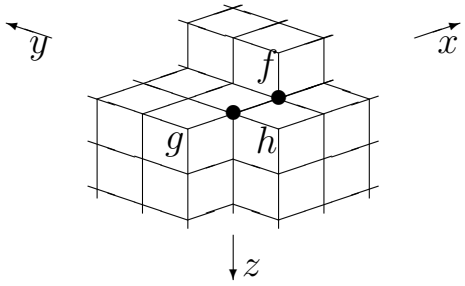
Theorem 2. *Suppose that an orthonomic system Σ satisfies some sufficient set of integrability conditions as above. Then the system Σ^∞ is equivalent to the system Σ' and the manifolds $\mathcal{E}_{\Sigma^\infty}$ and $\mathcal{E}_{\Sigma'}$ coincide.*

An integrability condition of the second kind is said to be *trivial* whenever $x \approx y$. Trivial integrability condition is redundant and may be omitted. In particular, a nontrivial integrability condition of the second kind at a point u^k_μ exists iff there are at least two distinct equivalence classes in $\mathcal{X}^k_\mu/\!\approx$. Any such point is the least common multiple $\operatorname{lcm}(\alpha, \beta)$ with $u^k_\alpha, u^k_\beta \in \operatorname{dom}\Sigma$ (cardinal point), but not the converse.
It is easy to write an algorithm to find all cardinal derivatives and compute the relation $\approx$ to select nontrivial among them. In the old and recent literature we were unable to find any example (unless very special ones, such as linear systems with constant coefficients) with a result better than that given by our algorithm. This leads us to the conjecture that the set of integrability conditions found by the algorithm is irredundant.
Given an orthonomic system Σ, let P_Σ denote the poset of nontrivial cardinal derivatives under the ordering by divisibility. Minimal cardinal points belong to P_Σ and harbour minimal integrability conditions in Reid's [5] sense. In generic case, P_Σ is an antichain. For every number $n > 2$ of independent variables, examples exist of P_Σ being a chain of length $n - 1$.

Proposition 3. *The poset P_Σ does not contain a chain of length n, where n is the number of independent variables.*

For every fixed k, derivatives u^k_μ can be visualized as points in the n-dimensional grid $\mathbb{N}^n \subset \mathbb{R}^n$. Integrability conditions of the second kind then have a rather transparent geometric interpretation, at least in low dimensions.

Example 4. Considering a system of equations $u_{xxy} = f$, $u_{yz} = g$, $u_{xz} = h$, nontrivial integrability conditions (of the second kind only) are visualized by bold dots on the following geometric scheme:



Of them, u_{xyz} is minimal and u_{xxyz} is supplementary in Reid's [5] sense.

References

[1] F. Boulier, An optimization of Seidenberg's elimination algorithm in differential algebra, *Math. Comput. Simulation* **42** (1996) 439–448.

[2] M. Caboara, M. Kreuzer and L. Robbiano, Efficiently computing minimal sets of critical pairs, *J. Symb. Comp.* **38** (2004) 1169–1190.

[3] R. Gebauer and M. Möller, On an instalation of Buchberger's algorithm, *J. Symb. Comp.* **6** (1988) 275–286.

[4] M. Marvan, Sufficient set of integrability conditions of an orthonomic system, preprint, `www.arxiv.org/abs/nlin.SI/0605009`.

[5] G.J. Reid, Algorithms for reducing a system of PDEs to standard form, determining the dimension of its solution space and calculating its Taylor series solution, *Eur. J. Appl. Math.* **2** (1991) 293–318.

[6] C.J. Rust, Rankings of derivatives for elimination algorithms, and formal solvability of analytic PDE, Ph.D. dissertation, University of Chicago, 1998.

Non-integrability criteria by means of the Galois groups of Variational Equations: recent results and works in progress[1]

Extended abstract

Juan J. Morales-Ruiz

Departament de Matemàtica Aplicada II
Universitat Politècnica de Catalunya
Edifici Omega, Campus Nord
c/ Jordi Girona, 1-3, E-08034 Barcelona, Spain
E-mail: Juan.Morales-Ruiz@upc.edu

Given a dynamical system,

$$\dot{z} = X(z), \tag{1}$$

and a particular integral curve of it, $z = z(t)$, at the end of the XIX century Poincaré proposed the variational equation along $z = z(t)$,

$$\dot{\xi} = \frac{\partial X}{\partial x}(z(t))\xi, \tag{2}$$

as the fundamental tool to study the behavior of (1) in a neighborhood $z(t)$ [5]. The equation (2) gives the linear part of the flow of (1) along $z = z(t)$. So, we have the following *General Principle*:

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

If we assume that the dynamical system (1) is "integrable" in any reasonable sense, then it is natural to conjecture that the linearized differential equation (2) must be also "integrable".

It seems clear that in order to convert this principle in a true conjecture it is necessary to clarify what kind of "integrability " is considered for equations (1) and (2).

As (2) is a linear differential equation, it is natural to consider the integrability of this equation in the context of the Galois theory of linear differential equations [6]. In order to do that, for technical reasons it is convenient to move to the complex analytical category, i.e., all the equations are complex analytical and defined over complex analytical spaces, fiber bundles, etc. Then we can define the integrability of a linear differential equation: a linear differential equation is integrable if it can be solved by a combination of algebraic functions, quadratures, and exponential of quadratures (the classical terminology is that the general solution is obtained by Liouvillian extensions over the field of coefficients). Moreover, the equation (2) is integrable if and only if it has a Galois group with a solvable identity component.

For complex analytical hamiltonian systems the *General Principle* works very well and in a joint work with Ramis we obtained the following result [3] (see also [2]), that in some sense can be considered as a generalization of a result by Ziglin in 1982 [7].

Theorem 1 *Assume that a complex analytical hamiltonian system is completely integrable by means of meromorphic first integrals in a neighborhood of the integral curve $z = z(t)$. Then the identity component of the Galois group of the (2) is a commutative group.*

We remark that this theorem is a typical version of several possible theorems and it has been applied by several authors to the study of the non-integrability of a wide range of systems (see [4] for concrete references):

a) N-body problems, problems with homogeneous potentials and cosmological models.

b) Some physical problems .

c) Other mechanical problems (rigid body, spring–pendulum,...).

d) Systems with some chaotic behaviour (splitting of asymptotic surfaces).

One of the essential points in the proof of the above theorem is the following lemma (this lemma is called in the reference [1] the *Key Lemma*):

Key Lemma ([3], see also [2]) *Let f be a meromorphic first integral of the dynamical system (1). Then the Galois group of (2) has a non-trivial rational invariant .*

Theorem 1 has been generalized recently to higher order variational equations (the solutions of these equations are the quadratic, cubic, etc. contributions to the flow of the Hamiltonian system along the particular solution $z = z(t)$) [4].

In this talk I will try to point out some recent works and works in progress about applications of Theorem 1 or some extensions based on the *General Principle* and the *Key Lemma*: applications to celestial mechanics and homogeneous potentials, non-holonomic dynamics, variational equations over invariant manifolds, newtonian homogeneous fields (see the work by Maciejewsky and Przybylska in this conference), connections with Lie symmetries, connection with the Malgrange Galois theory, etc.

References

[1] M. Audin, *"Les systèmes Hamiltoniens et leur intégrabilité"*, Cours Spécialisés, Collection SMF **8** Société Mathematique de France, Marseille 2001.

[2] J.J. Morales-Ruiz, *"Differential Galois Theory and Non-Integrability of Hamiltonian systems"*, Progress in Mathematics **179**, Birkhäuser 1999.

[3] J.J. Morales-Ruiz, J.P. Ramis, Galoisian obstructions to integrability of Hamiltonian systems, *Methods and Applications of Analysis* **8** (2001) 33–96.

[4] J.J. Morales-Ruiz, C, Simó, J.P. Ramis,Integrability of Hamiltonian Systems and Differential Galois Groups of Higher Variational Equations, to appair in *Ann. Sc. Ec. Norm. Sup.* .
[5] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste,* Vol. I. Gauthiers-Villars, Paris, 1892.
[6] M. van der Put and M. Singer, *Galois Theory of Linear Differential Equations*, Springer, Berlin 2003.
[7] S.L. Ziglin, Branching of solutions and non-existence of first integrals in Hamiltonian mechanics I, *Funct. Anal. Appl.* **16** (1982) 181–189.

Reduction Properties of Ordinary Differential Equations of Maximal Symmetry

Moyo S [1] Leach P G L [2]

[1] Durban University of Technology, Department of Mathematics
P O Box 953, Steve Biko Campus, Durban 4000, South Africa
moyos@dut.ac.za
[2] University of KwaZulu-Natal
School of Mathematical Sciences, Durban 4041
South Africa, leachp@ukzn.ac.za

Abstract

We study the ordinary differential equation $yy^{(n+1)} + \alpha y'y^{(n)} = 0$ and show that this equation is always integrable for a certain value of α. We also note that $\alpha = 0$ is a special case for which this equation has a nonlocal symmetry which enables one to reduce it to an equation of maximal symmetry. Different features of the differential equation and its intrinsic connection to the $sl(2, R)$ subalgebra are illustrated including the connection to integrating factors. In this paper, we look at the reduction properties of these equations from an algebraic point of view.

Keywords: Maximal symmetry, Integrating factor.

1 Introduction

Our original motivation comes from the Ermakov-Pinney equation [1, 2] which in its simplest form is

$$w'' + \frac{K}{w^3} = 0, \tag{1}$$

where K is a constant. In theoretical discussions the sign of the constant K is immaterial and in fact it is often rescaled to unity. In practical applications it would be negative to avoid 'collapse into the origin' due to its interpretation as the square of angular momentum [3, 4].

The general form of (1), *videlicet*

$$\ddot{\rho} + \omega^2(t)\rho = \frac{1}{\rho^3} \tag{2}$$

occurs in the study of the time-dependent linear oscillator, be it the classical or the quantal problem, as the differential equation which determines the time-dependent rescaling of the space variable and the definition of 'new time'. In this context we mention the references [5, 6, 7, 8, 9, 10, 11].

The same form is one of the basic equations of Ermakov systems which, with their generalisations, have been the subject of an extensive literature over the last thirty years. Some references are [12, 13, 14, 15, 16, 17].

Another origin of (1) – of particular interest in this work – is as an integral of the third-order equation of maximal symmetry which in its elemental form is

$$y''' = 0. \tag{3}$$

The integration of this equation, which is a feature of the calculation of the symmetries of all linear ordinary differential equations of maximal symmetry [18], by means of an integrating factor gives a variety of results depending upon the integrating factor used. Including the one relevant to (1) some obvious integrating factors give

$$\begin{array}{rcl} 1.y''' = 0 & \longrightarrow & I_3 = y'' \\ x.y''' = 0 & & I_2 = xy'' - y' \\ \frac{1}{2}x^2.y''' = 0 & & I_1 = \frac{1}{2}x^2y'' - xy' + y \\ y''.y''' = 0 & & J = \frac{1}{2}y''^2 \quad \text{ie} \quad \frac{1}{2}I_3^2 \\ y.y''' = 0 & & y''y^3 + K = 0 \end{array} \tag{4}$$

and the last of these is to (1) when the integral is interpreted as an equation. (The numbering of the fundamental first integrals follows the convention given in Flessas *et al* [19, 20].)

To illustrate the point on integrating factors we consider the equation of maximal symmetry (3) which has seven Lie point symmetries. These are

$$\begin{aligned} G_1 &= \partial_y \\ G_2 &= x\partial_y \\ G_3 &= x^2\partial_y + 2x\partial_{y'} \\ G_4 &= y\partial_y + y'\partial_{y'} \\ G_5 &= \partial_x \\ G_6 &= x\partial_x + y\partial_y \\ G_7 &= x^2\partial_x + 2xy\partial_y. \end{aligned}$$

The algebra is $3A_1, \{sl(2,R) \oplus_s A_1\}$ and $3A_1$. The autonomous integrating factors for (3) are y'' and y as mentioned above. We list the symmetries and algebra when each of the integrating factors is treated as an equation and as a function.

When we multiply $y''' = 0$ by the integrating factor y'' we obtain $y''y''' = 0$. Integrating this expression gives $\frac{1}{2}y''^2 = k$, where k is a constant of integration. This may be rewritten as $y'' = k$ without loss of generality. This may be treated as the function y'' or as an equation $y'' - k = 0$ in which k is a parameter. We then have three cases for which we list the symmetries as follows:

$y'' = 0$	$y'' = k$	y''
$G_1 = \partial_y$	$G_1 = \partial_y$	$G_1 = \partial_y$
$G_2 = x\partial_y$	$G_2 = x\partial_y$	$G_2 = x\partial_y$
$G_3 = y\partial_y$	$G_3 = (\frac{1}{2}x^2k - y)\partial_y$	$G_3 = \partial_x$
$G_4 = \partial_x$	$G_4 = \partial_x + 2xk\partial_y$	$G_4 = x\partial_x + 2y\partial_y$
$G_5 = x\partial_x$	$G_5 = x\partial_x + x^2k\partial_y$	
$G_6 = x^2\partial_x + xy\partial_y$	$G_6 = x^2\partial_x + (xy + \frac{1}{2}x^3k)\partial_y$	
$G_7 = y\partial_x$	$G_7 = (y - \frac{3}{2}x^2k)\partial_x - x^3k^2\partial_y$	
$G_8 = xy\partial_x + y^2\partial_y$	$G_8 = (xy - \frac{1}{2}x^3k)\partial_x + (y^2 - \frac{1}{4}x^4k^2)\partial_y.$	

(5)

Remark

- The algebra of the symmetries listed in columns one and two is $sl(3,R) : 2A_1 \oplus_s \{sl(2,R) \oplus A_1\} \oplus 2A_1$ whereas that for the third column is $A^1_{4,9} : A_2 \oplus_s 2A_1$.
- This is a clear indication of the distinction of the algebraic properties between first integrals, that is the function of column three, and configurational invariants, that the equations of columns one and two.

If y is used an the integrating factor we obtain $yy''' = 0$. Integrating this equation gives

$$yy'' - \tfrac{1}{2}y'^2 = k \tag{6}$$

which can be written as

$$(y^{1/2})'' = \frac{k}{(y^{1/2})^3} \tag{7}$$

and is the simplest form of the Ermakov-Pinney equation [1, 2]. As before we write down the point symmetries corresponding to the three cases of the differential equation $u'' = k/u^3$ where $u = y^{1/2}$. We have the following:

$$\begin{array}{lll}
u'' = 0 & u'' = k/u^3 & u''u^3 \\
G_1 = \partial_u & G_1 = \partial_u & G_1 = \partial_u \\
G_2 = x\partial_u & G_2 = 2x\partial_x + u\partial_u & G_2 = 2x\partial_x + u\partial_u \\
G_3 = u\partial_u & G_3 = x^2\partial_x + xu\partial_u & G_3 = x^2\partial_x + xu\partial_u. \\
G_4 = \partial_x & & \\
G_5 = x\partial_x & & \\
G_6 = x^2\partial_x + xu\partial_u & & \\
G_7 = u\partial_x & & \\
G_8 = xu\partial_x + u^2\partial_u & &
\end{array} \tag{8}$$

The transformation of $yy'' - \frac{1}{2}y'^2 = k$ to $u'' = k/u^3$ does not make a difference in terms of the symmetries as we just have a point transformation in this case.

The characteristic feature of the Ermakov-Pinney equation, (1), is that it possesses the three-element algebra of Lie point symmetries $sl(2, R)$ which in itself is characteristic of all scalar ordinary differential equations of maximal symmetry. One of the purposes of this paper is to investigate higher-order analogues of the Ermakov-Pinney equation. The basic criterion is algebraic.

In investigations of the Emden-Fowler equation [21, 22, 23, 24] the existence of a Lie point symmetry for certain indices[1] is intimately connected [25, 26, 27] with the solution of the fourth-order equation

$$2yy'''' + 5y'y''' + \alpha y''' = 0, \tag{9}$$

where α is a parameter which occurs in both the Emden-Fowler equation and in the symmetry [26, 27, 28, 29, 30, 31]. The determination of the first integral associated with the symmetry is not possible in closed-form for nonzero α. When $\alpha = 0$, the Lie symmetry becomes a Noether symmetry and the associated integral follows directly from an application of Noether's theorem [32]. Equation (9) possesses just two Lie point symmetries. However, in the case that $\alpha = 0$, *ie*, when the equation has the form

$$2yy'''' + 5y'y''' = 0, \tag{10}$$

there are the three Lie point symmetries

$$\Gamma_1 = \partial_x, \quad \Gamma_2 = x\partial_x \quad \text{and} \quad \Gamma_3 = y\partial_y. \tag{11}$$

Reduction by Γ_1 leads to a third-order equation which also has three Lie point symmetries. Two of these are the descendants of Γ_2 and Γ_3 as one would expect since Γ_1 is the normal subgroup in both cases. The third symmetry of the reduced equation,

$$\Lambda_4 = 2u^2\partial_u + uv\partial_v \tag{12}$$

[1] For instance when the Emden-Fowler equation of index two given by $y'' = f(x)y^2$ gives rise to (9). See [25] for a detailed treatment of this equation.

(the variables of the reduced equation are $u = y$ and $v = y'$) is a hidden symmetry of Type II [33, 34, 35] and has its origin in the nonlocal symmetry of the fourth-order equation,

$$\Gamma_4 = 3\left(\int y\mathrm{d}x\right)\partial_x + 2y^2\partial_y. \tag{13}$$

When the symmetry (12) is used to reduce the third-order equation to a second-order equation, the resulting equation is of maximal symmetry and so linear when expressed in the correct coordinates [25]. Without this hidden symmetry the reduction by the three symmetries given in (11) leads to an Abel's equation of the second kind of particularly daunting aspect.

Equation (10) was used by Euler *et al* [36] as an example in their study of the integrability properties of equations of the form

$$y^{(n+1)} = h\left(y, y^{(n)}\right) y' \tag{14}$$

and they showed that the equation could be reduced to $\mathrm{d}^4Y/\mathrm{d}X^4 = 0$, that is the fourth-order equation of maximal symmetry, by means of a complex sequence of nonlocal transformations which by most curious happenstance included the very Emden-Fowler equation from which it arose.

In this paper we draw together various features to which we have alluded above to make a coherent study. We commence with first integrals which possess three Lie point symmetries with the algebra $sl(2, R)$ and have a structure resembling that of (1). The associated differential equation is of the form of (14) with an explicit form of the function h, that is the imposition of the algebraic constraint provides a precise definition of the associated differential equation. Equation (10) does not fit into this structure. Equation (10) is a particular case of the two-parameter family of differential equations,

$$yy^{(n+1)} + \alpha y'y^{(n)} = 0, \tag{15}$$

and we make a study of the point symmetries for general values of the parameter α using the parameter-testing facility of Program LIE [37, 38]. We find that there is a value of the parameter, α, for which (15) is always integrable. We see that (10) and its useful nonlocal symmetry is peculiar.

2 Higher-order analogues of the Ermakov-Pinney equation

As we indicated above, there are several approaches which may be taken. Here we assume that the integral has the same structure as (1) and possesses the Lie algebra $sl(2,R)$ of point symmetries.

Suppose that there exists an integral of the form

$$I = y^{(n)}y^{\alpha}, \tag{16}$$

when α is a parameter to be determined, with the associated $(n+1)$th-order equation

$$yy^{(n+1)} + \alpha y'y^{(n)} = 0 \tag{17}$$

such that the integral, I, has the $sl(2,R)$ symmetries appropriate [18] to the nth-order equation of maximal symmetry, $y^{(n)} = 0$, *ie*

$$\begin{aligned} \Gamma_1 &= \partial_x \\ \Gamma_2 &= x\partial_x + \tfrac{1}{2}(n-1)y\partial_y \\ \Gamma_3 &= x^2\partial_x + (n-1)xy\partial_y. \end{aligned} \tag{18}$$

The first symmetry, Γ_1, is implied by the autonomy of (16). Since the Lie Bracket, $[\Gamma_1, \Gamma_3]_{LB} = 2\Gamma_1$, we need use only Γ_1 and Γ_3. The $(n+1)$th extension of Γ_3 is

$$\begin{aligned} \Gamma_3^{[n+1]} &= x^2\partial_x + (n-1)xy\partial_y + [(n-1)y + (n-3)xy']\,\partial_{y'} \\ &\quad + [2(n-2)y' + (n-5)xy'']\,\partial_{y''} + [3(n-3)y'' + (n-7)xy''']\,\partial_{y'''} \\ &\quad + \ldots + [n-(2n+1)]\,xy^{(n)}\partial_{y^{(n)}} - \left[(n+1)y^{(n)} + (n+3)xy^{(n+1)}\right] \\ &\quad \cdot\partial_{y^{(n+1)}}. \end{aligned} \tag{19}$$

Since both integral and equation are autonomous, (20) may be split into an x-free part and an x-dependent part. The former does not contain any operators of relevance to the integral and so gives zero automatically. In the case of the latter we obtain

$$-(n+1)xy^{(n)}y^{\alpha} + (n-1)\alpha xyy^{(n)}y^{\alpha-1} \tag{20}$$

which is zero provided the parameter α takes the value $(n+1)/(n-1)$. Evidently $n \neq 1$, *ie* there does not exist a first-order Ermakov-Pinney equation.

We recall that the Ermakov-Pinney equation has the form

$$w'' = \frac{K}{w^3} \tag{21}$$

in the notation adopted above whereas its primitive form is

$$yy'' - \tfrac{1}{2}y'^2 = \tfrac{1}{2}K \tag{22}$$

as the direct integral of $y''' = 0$ in association with the integrating factor y. The transformation is $y = \frac{1}{2}w^2$. Does a similar property persist at the higher order?

Consider the general equation (17) constrained to possess the representation of the $sl(2,R)$ subalgebra given in (18)[2],

$$yy^{(n+1)} + \frac{n+1}{n-1}y'y^{(n)} = 0, \tag{23}$$

which is the derivative of the Ermakov-Pinney equation

$$y^{(n)} + \frac{K}{y^{(n+1)/(n-1)}} = 0. \tag{24}$$

For the purposes of this treatment equation (24) defines the general Ermakov-Pinney equation for $n > 1$. The property holds for $n = 2$. For $n = 3$ the Ermakov-Pinney equation is

$$y''' + \frac{K}{y^2} = 0 \tag{25}$$

and the corresponding fourth-order equation is

$$yy'''' + 2y'y''' = 0. \tag{26}$$

[2] Equation (23) is not the most general nth order ordinary differential equation invariant under $sl(2,R)$. See [39] for a detailed treatment of this question.

To see if this equation has a simple form we set $y = w^\alpha$, where α is a parameter to be determined. The fourth-order equation becomes, after a modicum of simplification,

$$\begin{aligned} w \;\; & {}^3w'''' + 2(3\alpha - 1)w^2w'w''' + 3(\alpha - 1)w^2w''^2 \\ + \;\; & 12(\alpha - 1)^2ww'^2w'' + 3(\alpha - 1)^2(\alpha - 2)w'^4 = 0. \end{aligned} \tag{27}$$

It is evident that the original form of the equation is the simplest available under this class of transformations. In the case of $n = 2$ the ability to reduce the nonlinear equation obtained by differentiation of the second-order Ermakov-Pinney equation to the third-order equation of maximal symmetry was accidental and not an intrinsic property of Ermakov-Pinney equations.

3 The structure of Euler *et al*

A second approach to the investigation of equations of the structure of the Ermakov-Pinney equation is to begin from the structure treated by Euler *et al* [36]. The model equation which they treated had the general form

$$y^{(n+1)} = h\left(y, y^{(n)}\right) y'. \tag{28}$$

We impose an $sl(2, R)$ algebraic structure on this equation. We take the structure to be

$$\begin{aligned} \Gamma_1 &= \partial_x \\ \Gamma_2 &= x\partial_x + my\partial_y \\ \Gamma_3 &= x^2\partial_x + 2mxy\partial_y, \end{aligned} \tag{29}$$

where the parameter m is at our disposal.

The structure assumed for (28) makes the possession of Γ_1 automatic. The action of the $(n+1)$th extension of Γ_2 leads to

$$my\frac{\partial h}{\partial y} + (m - n)y^{(n)}\frac{\partial h}{\partial y^{(n)}} = -nh \tag{30}$$

$$\frac{\mathrm{d}y}{my} = \frac{\mathrm{d}y^{(n)}}{(m - n)y^{(n)}} = \frac{\mathrm{d}h}{-nh} \tag{31}$$

and from the associated Lagrange's system, (31), we find that the characteristics of the first-order linear partial differential equation, (30), are

$$u = hy^{n/m} \quad \text{and} \quad v = \frac{y^{(n)}}{y^{(m-n)/m}} \tag{32}$$

so that the form of (28) invariant under the actions of Γ_1 and Γ_2 is

$$y^{n/m}y^{(n+1)} = g\left(\frac{y^{(n)}}{y^{(m-n)/m}}\right)y'. \tag{33}$$

We now turn to Γ_3. The $(n+1)$th extension is

$$\begin{aligned}
\Gamma_3^{[n+1]} &= x^2\partial_x + 2mxy\partial_y + 2\left[my + (m-1)xy'\right]\partial_{y'} \\
&+ 2\left[(2m-1)y' + (m-2)xy''\right]\partial_{y''} + \ldots \\
&+ 2\left\{\left[nm - \tfrac{1}{2}n(n-1)\right]y^{(n-1)} + (m-n)xy^{(n)}\right\}\partial_{y^{(n)}} \\
&+ 2\left\{\left[(n+1)m - \tfrac{1}{2}n(n+1)\right]y^{(n)} + (m-n-1)xy^{(n+1)}\right\} \\
&\cdot \partial_{y^{(n+1)}}.
\end{aligned}$$

When this is applied to (33), there is no need to consider the part which has x as coefficient since the actions of Γ_1 and Γ_2 have already done that. The effective part of the operator gives

$$\begin{aligned}
&y^{n/m}\left[(n+1)m - \tfrac{1}{2}n(n+1)\right]y^{(n)} \\
&= gmy + \frac{g'}{y^{(m-n)/m}}\left[nm - \tfrac{1}{2}n(n-1)\right]y^{(n-1)}y'.
\end{aligned} \tag{34}$$

However, $y^{(n-1)}$ is not permitted. The coefficient of $y^{(n-1)}$ must be zero, *ie* $m = \frac{1}{2}(n-1)$. With this restriction on the value of m we find from (34) that

$$g = -\frac{n+1}{n-1}y\frac{n+1}{n-1}y^{(n)}$$

and the equation is specifically (23). Thus the $sl(2,R)$ equation in combination with the constraint of Euler *et al* is unique at all orders.

4 Conclusion

We have studied the differential equation $yy^{(n+1)} + \alpha y' y^{(n)} = 0$ and shown that this equation is always integrable for a certain value of α. The value of $\alpha = 0$ is a special case for which this equation has a nonlocal symmetry which enables one to reduce it to an equation of maximal symmetry. Different features of the differential equation and its intrinsic connection to the $sl(2, R)$ subalgebra are illustrated including the connection to integrating factors. We have shown how the integrating factors, for example, of the third ordinary differential equation give different symmetry properties depending on which integrating factor is used. It is important to mention that if y is an integrating factor of $y^{(n)} = 0$, then the integral obtained by using this integrating factor always has the $sl(2, R)$ subalgebra.

In the case that $n = 1$ in (20) the representation of $sl(2, R)$, (18), takes a form for which there is no second-order equation. There is an equation of the third order which has these symmetries. Indeed the third-order equation with a double set of these symmetries, one in x and the other in y, is the Kummer-Schwarz equation which is a nonlinear representation of a third-order equation of maximal symmetry when one admits contact symmetries.

There are still some additional questions which still to be answered like why is there a difference in reduction properties in reduced equations when moving from $6th, 5th, 4th, 3rd$ and $2nd$ order and a rigorous look into the reduction properties of these equations from an algebraic point of view.

References

[1] Ermakov, V.: Second-order differential equations. Conditions of complete integrability *Universita Izvestia Kiev Series III*, **9**, 1-25 (trans Harin AO), 1880.

[2] Pinney, E.: The nonlinear differential equation $y''(x) + p(x)y + cy^{-3} = 0$, *Proceedings of the American Mathematical Society*, **1**, 681, 1950.

[3] Whittaker, E.T.: *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Dover, New York, 1944.

[4] Camiz, P., Geradi, A., Marchioro, C., Presutti, E. and Scacciatelli, E.: Exact solu-

tion of a time-dependent quantal harmonic oscillator with a singular perturbation, *Journal of Mathematical Physics*, **12**, 2040-2043, 1971.

[5] Lewis, H. Ralph Jr.: Classical and quantum systems with time-dependent harmonic oscillator-type Hamiltonians, *Physical Review Letters*, **18**, 510-512, 1967.

[6] Lewis, H. Ralph Jr.: Motion of a time-dependent harmonic oscillator and of a charged particle in a time-dependent axially symmetric electromagnetic field, *Physical Review*, **172**, 1313-1315, 1968.

[7] Lewis, H. Ralph Jr.: Class of exact invariants for classical and quantum time-dependent harmonic oscillators, *kept Journal of Mathematical Physics*, **9**, 1976-1986, 1968.

[8] Lewis, H. Ralph Jr. and Riesenfeld, W.B.: An exact quantum theory of the time-dependent harmonic oscillator and of a charged particle in a time-dependent electromagnetic field, *Journal of Mathematical Physics*, **10**, 1458-1473, 1969.

[9] Leach, P.G.L.: On a Direct method for the Determination of an Exact Invariant for the Time-Dependent Harmonic Oscillator, *Journal of the Australian Mathematical Society, Series B*, **20**, 97-105, 1977.

[10] Leach, P.G.L.: Invariants and Wave-Functions for some Time-Dependent Harmonic Oscillator-Type Hamiltonians, *Journal of Mathematical Physics*, **18**, 1902-1907, 1977.

[11] Leach, P.G.L.: Berry's phase and wave functions for time-dependent Hamiltonian systems, *Journal of Physics A: Mathematical and General*, **23**, 2695-2699, 1990.

[12] Athorne, C., Rogers, C., Ramgulam, U. and Osbaldestin, A.: On linearization of the Ermakov system, *Physics Letters A*, **143**, 207-212, 1990.

[13] Govinder, K.S., Athorne, C. and Leach, P.G.L.: The algebraic structure of generalized Ermakov systems in three dimensions *Journal of Physics A: Mathematical and General* **26** 4035-4046, 1993.

[14] Ray, J.R.: Nonlinear superposition law for generalized Ermakov systems, *Physics Letters A*, **78**, 4-6, 1980.

[15] Ray, J.R. and Reid, J.L.: Noether's theorem and Ermakov systems for nonlinear equations of motion, Il Nuovo Cimento, **59A**, 134-140, 1980.

[16] Ray, J.R. and Reid, J.L.: Ermakov systems, velocity dependent potentials and nonlinear superposition, *Journal of Mathematical Physics*, **22**, 91-95, 1981.

[17] Sarlet, W.: Further generalizations of Ray-Reid systems, *Physics Letters*, **82A**, 161-164, 1981.

[18] Mahomed, F.M. and Leach P.G.L.: Symmetry Lie algebras of nth order ordinary differential equations, *Journal of Mathematical Analysis and Applications*, **151**, 80-107, 1990.

[19] Flessas, G.P., Govinder, K.S. and Leach, P.G.L.:Remarks on the symmetry Lie algebras of first integrals of scalar third order ordinary differential equations with maximal symmetry, *Bulletin of the Greek Mathematical Society*, **36**, 63-79, 1994.

[20] Flessas, G.P., Govinder, K.S. and Leach, P.G.L.: Characterisation of the algebraic properties of first integrals of scalar ordinary differential equations of maximal symmetry, *Journal of Mathematical Analysis and Applications*, **212**, 349-374, 1997.

[21] Lane, I.J. Homer.: On the theoretical temperature of the Sun under the hypothesis of a gaseous mass maintaining its volume by its internal heat and depending on the laws of gases known to terrestrial experiment, *American Journal of the Sciences and Arts*, **4**, 57-74, 1869-70.

[22] Emden, R.: *Gaskugeln, Anwedungen der mechanishen Warmentheorie auf Kosmologie und meteorologische Probleme*, BG Teubner, Leipzig, 1987.

[23] Fowler, R.H.: The form near infinity of real, continuous solutions of a certain differential equation of the second order *Quarterly Journal of Mathematics* , **45**, 289-350, 1914.

[24] Fowler, R. H.: Further studies of Emden's and similar differential equations, *Quarterly Journal of Mathematics (Oxford)*, **2**, 259-288, 1931.

[25] Leach, P.G.L. and Govinder, K.S.: Hidden symmetries and integration of the generalized Emden-Fowler equation of index two, *Proceedings 14th IMACS World Congress on Computational and Applied Mathematics*, Ames, W.F. (Eds.), Georgia Institute of Technology, Atlanta, 300-303, 1994.

[26] Leach, P.G.L.: An exact invariant for a class of time-dependent anharmonic oscillators with cubic anharmonicity, *Journal of Mathematical Physics*, **22**, 465-470, 1981.

[27] Leach, P.G.L.: First integrals for the modified Emden equation $\ddot{q} + \alpha(t)\dot{q} + q^n = 0$, *Journal of Mathematical Physics*, **26**, 2510-2514, 1985.

[28] Leach, P.G.L. and Maharaj, S.D.: A first integral for a class of time-dependent anharmonic oscillators with multiple anharmonicities, *Journal of Mathematical Physics*, **33**, 2023-2030, 1992.

[29] Leach, P.G.L., Maartens, R. and Maharaj, S.D.:Self-similar solutions of the generalized Emden-Fowler equation, *International Journal of Nonlinear Mechanics*, **27**, 575-582, 1992.

[30] Maharaj, S.D., Leach, P.G.L. and Maartens, R.: Shear-free spherically symmetric solutions with conformal symmetry *General Relativity and Gravitation*, **23**, 261-267, 1991.

[31] Mellin, C. M., Mahomed, F.M. and Leach, P.G.L.: Solution of generalized Emden-Fowler equations with two symmetries, *International Journal of Nonlinear Mechanics*, **29**, 529-538, 1994.

[32] Noether, E.: Invariante Variationsprobleme, *Königlich Gesellschaft der Wissenschaften Göttingen Nachrichten Mathematik-physik Klasse*, **2**, 235-267, 1918.

[33] Abraham, S.B. and Guo, A.: Hidden symmetries associated with the projective group of nonlinear first order ordinary differential equations, *Journal of Physics A: Mathematical and General*, **25**, 5597-5608, 1992.

[34] Abraham, S.B.: Hidden symmetries and linearization of the modified Painlevé-Ince

equation, *Journal of Mathematical Physics*, **34**, 4809-4816, 1993.

[35] Abraham, S.B. and Leach, P.G.L.: Hidden symmetries of nonlinear ordinary differential equations, *Exploiting symmetry in Applied and Numerical Analysis*, Allgower,E., Georg, K. and Miranda, R. (Eds.), Lectures in Applied Mathematics, **29** , AMS, Providence, 1-10, 1993.

[36] Euler, N. and Leach, P.G.L.: First integrals and reduction of a class of nonlinear higher order ordinary differential equations, *Journal of Mathematical Analysis and Applications*, **287**, 337-347, 2003.

[37] Head, A.K.: `LIE`, a PC program for Lie analysis of differential equations, *kept computer Physics Communications*, **77**, 241-248, 1993.

[38] Sherring, J., Head, A.K. and Prince, G.E.: Dimsym and LIE: symmetry determining packages, *Mathematical and Computer Modelling*, **25**, 153-164, 1997.

[39] Leach, P.G.L. and Govinder, K.S.: On the uniqueness of the Schwarzian and linearisation by nonlocal contact transformation, *Journal of Mathematical Analysis and Application*, **235**, 84-107, 1999.

Integrability of homogeneous systems. Results and problems[1]

Andrzej J. Maciejewski [1] Maria Przybylska [2,3]

[1] Institute of Astronomy, University of Zielona Góra, Podgórna 50, PL-65–246 Zielona Góra, Poland maciejka@astro.ia.uz.zgora.pl
[2] Institut Fourier, UMR 5582 du CNRS, Université de Grenoble I, 100 rue des Maths, BP 74, 38402 Saint-Martin d'Hères Cedex, France
[3] Toruń Centre for Astronomy, N. Copernicus University, Gagarina 11, PL-87–100 Toruń, Poland, Maria.Przybylska@astri.uni.torun.pl

Abstract

We report our recent results concerning the integrability of two classes of equations: natural Hamiltonian systems with homogeneous potentials and Newton's equations with homogeneous right-hand sides. Our approach combines the methods of differential Galois theory and a certain kind of the global analysis of the investigated systems. We show that among the methods known till now, our approach gives the strongest necessary conditions for the integrability.

Keywords: integrability, non-integrability criteria, differential Galois group, hypergeometric equation, Hamiltonian equations, Newton equations

1 Introduction

During the last two decades analytical methods used to study differential equations have been considerably developed. In the integrability theory, after the epoch of the Painlevé analysis, rigorous methods of proving the non-integrability appeared. Now, it is commonly accepted (but in the

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

whole generality not proved) that a 'generic' system of differential equations is not integrable. Having this fact in mind, all efforts to find integrable systems or to prove the non-integrability of the given system, can be considered as an exotic curiosity. However, it is not like that. First of all, integrable systems are interesting as such, not only from the mathematical point of view. Furthermore, many physical systems depend on parameters. For such systems there is an interesting question how to find values of parameters for which the system is integrable. Obviously, numerical methods cannot answer this question. On the contrary, applying analytical methods we can distinguish a low dimensional set in the parameters' space such that for the parameters from this set all necessary conditions for the integrability are satisfied. If the distinguished set consists of a finite number of points, then we can apply a direct method for searching first integrals. It is also possible to apply numerical methods to check if for prescribed parameters' values the system behaves regularly. Moreover, even if we know that the considered system is not integrable, it is of great importance to gain information about the reason and mechanisms of the non-integrability.

The meaning of the integrability depends on the context, or, more precisely, on a class of the investigated systems. All accepted definitions of the integrability follow the same philosophy: a system is integrable iff it possesses such a number of first integrals (and/or other tensor invariants) that it is solvable by quadratures. For example, the integrability of a Hamiltonian system in the Liouville sense implies that solutions of Hamilton's equations can be expressed by quadratures.

Generally, only few methods for proving the non-integrability are known. Let us mention here three of them.

a) The Lagutynskii-Levelt method, see [1, 2] This is an algebraic version of the Kovalevskaya analysis. It can be used for proving the non-existence of one or more polynomial or rational first integrals of general homogeneous or qusi-homogeneous polynomial systems. In the cited references it was formulated in the form applicable for ho-

mogeneous systems. In [3] we generalised this simple but effective method for quasi-homogeneous systems.

b) The Ziglin method [4, 5] is applied to Hamiltonian systems. It can be used for proving the non-integrability in the Liouville sense. It is less known that it can also be applied for proving the non-existence of a certain number of additional meromorphic first integrals which do not necessarily commute.

c) The Morales-Ramis method is, in some sense, a purely algebraic version of the Ziglin method. It applies for proving the meromorphic non-integrability in the Liouville sense. In [6] we extended the application of the Morales-Ramis approach for proving the non-integrability in the non-commutative sense.

All the above methods are based on a linearisation of the investigated system around a particular solution. Among them the most powerfull is the Morales-Ramis method.

The aim of this paper is to show the main results of the integrability studies of two classes of polynomial systems applying the Morales-Ramis theory. The first of them are natural Hamiltonian systems

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{q} = \boldsymbol{p}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p} = -V'(\boldsymbol{q}), \qquad V'(\boldsymbol{q}) = \operatorname{grad} V(\boldsymbol{q}), \tag{1}$$

with Hamilton function

$$H = \frac{1}{2}\sum_{i=1}^{n} p_i^2 + V(\boldsymbol{q}), \tag{2}$$

where $V = V(\boldsymbol{q}) \in \mathbb{C}[\boldsymbol{q}]$ is a homogeneous polynomial of degree $k > 2$. The second class is a direct generalisation of (1), namely, the class of autonomous Newton equations

$$\ddot{\boldsymbol{q}} = -\boldsymbol{F}(\boldsymbol{q}), \qquad \boldsymbol{q} = (q_1, \ldots, q_n)^T \in \mathbb{C}^n, \tag{3}$$

with homogeneous polynomials $F_i = F_i(q_1, \ldots, q_n)$ of degree $k - 1 > 1$. Obviously, the Hamilton equations (1) are a special case of the Newton equations with $\boldsymbol{F} = V'$.

For Hamiltonian systems of the form (1) the Morales-Ramis theory is directly applicable. As a matter of fact, J.J. Morales-Ruiz and J.-P. Ramis formulated a theorem, see [7, 8], which gives the necessary conditions for the integrability of such systems. Thus, our aim was to improve this result. We observed that it is important to take into account globally all the implications of the Morales-Ramis theorem applied for all accessible particular solutions. As an outcome, we obtain a 'finiteness' theorem. Roughly speaking: for a fixed k among Hamiltonian systems (1) there is at most a finite number of the ones satisfying the necessary conditions for the integrability.

For the Newton equations, the Morales-Ramis theory is not applicable directly, as such systems are not Hamiltonian. Nevertheless, we show that it is possible to use the differential Galois theory for the integrability study of such systems. Then we follow the ideas which appeared to be so fruitfull in our investigations of Hamiltonian systems.

The idea of this paper is to give a guide-book presenting main ideas, results and open problems. An interested reader can find all the details in cited publications. We present also completely new results contained in papers which are under preparations.

2 Sketch of the applied theory

All the methods mentioned in the introduction arose as developments of an idea due to Weierstrass and Kovalevskaya that solutions of an integrable system considered as functions of the complex time should be single valued. Of course, in applications, the independent variable, or the time is real, so the solutions are functions of a real variable. Nevertheless, to understand all properties of functions of a real variable it is fruitfull to investigate them in a complex domain.

Thus let us consider a complex analytic manifold M^n, a holomorphic vector field $\boldsymbol{v}$ on M^n and the following dynamical system generated by this field

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x} = \boldsymbol{v}(\boldsymbol{x}), \qquad \boldsymbol{x} \in M^n, \quad t \in \mathbb{C}. \tag{4}$$

All the mentioned methods are based on the linearisation of the original system around a particular non-equilibrium solution $\boldsymbol{\varphi}(t)$ of (4), thus we introduce variational variables $\boldsymbol{\xi}$ as

$$\boldsymbol{x}(t) = \boldsymbol{\varphi}(t) + \boldsymbol{\xi}(t). \tag{5}$$

The linearisation of equation (4) along $\boldsymbol{\varphi}(t)$ has the form

$$\dot{\boldsymbol{\xi}} = A(t)\boldsymbol{\xi}, \qquad A(t) = \frac{\partial \boldsymbol{v}}{\partial \boldsymbol{x}}(\boldsymbol{\varphi}(t)), \qquad \boldsymbol{\xi} \in T_\Gamma M^n. \tag{6}$$

The particular solution $\boldsymbol{\varphi}(t)$ defines a Riemann surface Γ with t as a local coordinate. We can consider the entries of matrix A as elements of field $\mathcal{K} := \mathcal{M}(\Gamma)$ of meromorphic functions on Γ. This field with differentiation with respect to t as a derivation is a differential field. Only constant functions from $\mathcal{K}$ have a vanishing derivative, so the subfield of constants of $\mathcal{K}$ is $\mathbb{C}$.

It is obvious that solutions of (6) are not necessarily elements of $\mathcal{K}^n$. The fundamental theorem of the differential Galois theory guarantees that there exists a differential field $\mathcal{F} \supset \mathcal{K}$ such that it contains n linearly independent (over $\mathbb{C}$) solutions of (6). The smallest differential extension $\mathcal{F} \supset \mathcal{K}$ with this property is called the Picard-Vessiot extension. A group $\mathcal{G}$ of differential automorphisms of $\mathcal{F}$ which does not change $\mathcal{K}$ is called the differential Galois group of equation (6). It can be shown that $\mathcal{G}$ is a linear algebraic group. Thus, it is a union of disjoint connected components. One of them containing the identity is called the identity component $\mathcal{G}^\circ$.

But for us more important is the fact that the differential Galois group of variational equations has a close connection with the integrability of variational equations (6) as well as (4).

Now we explain why the differential Galois group of variational equations is important in a study of the integrability. At first, we introduce a few definitions. Let us consider a holomorphic function F defined in a certain connected neighbourhood of solution $\boldsymbol{\varphi}(t)$. In this neighbourhood we have the expansion

$$F(\boldsymbol{\varphi}(t)+\boldsymbol{\xi}) = F_m(\boldsymbol{\xi}) + O(\|\boldsymbol{\xi}\|^{m+1}), \qquad F_m \neq 0. \tag{7}$$

Then the leading term f of F is the lowest order term of the above expansion i.e., $f(\boldsymbol{\xi}) := F_m(\boldsymbol{\xi})$. Note that $f(\boldsymbol{\xi})$ is a homogeneous polynomial of variables $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ of degree m; its coefficients are polynomials in $\boldsymbol{\varphi}(t)$. If F is a meromorphic function, then it can be written as $F = P/Q$ for certain holomorphic functions P and Q. Then the leading term f of F is defined as $f = p/q$, where p and q are leading terms of P and Q, respectively. In this case $f(\boldsymbol{\xi})$ is a homogeneous rational function of $\boldsymbol{\xi}$.

One can prove that if F is a meromorphic (holomorphic) first integral of equation (4), then its leading term f is a rational (polynomial) first integral of variational equations (6). Moreover, if system (4) has $k \geq 2$ functionally independent meromorphic first integrals $F_1, \ldots, F_k$, then by the Ziglin Lemma [4, 9, 10] one can assume that the leading terms of $F_1, \ldots, F_k$, are functionally independent rational first integrals of (6).

Additionally, if $\mathcal{G} \subset \mathrm{GL}(n, \mathbb{C})$ is the differential Galois group of (6), and f is its rational first integral, then $f(g(\boldsymbol{\xi})) = f(\boldsymbol{\xi})$ for every $g \in \mathcal{G}$ [11, 7] (see also [9]). This means that f is a rational invariant of group $\mathcal{G}$. Thus we have a correspondence between the first integrals of the system (4) and invariants of $\mathcal{G}$.

Lemma 1. *If equation* (4) *has k functionally independent first integrals which are meromorphic in a connected neighbourhood of a non-equilibrium solution $\boldsymbol{\varphi}(t)$, then the differential Galois group $\mathcal{G}$ of the variational equations along $\boldsymbol{\varphi}(t)$ has k functionally independent rational invariants.*

As mentioned, a differential Galois group is an algebraic group, thus in particular it is a Lie group, and one can consider its Lie algebra.

This Lie algebra reflects only the properties of the identity component of the group. It is easy to show that if a Lie group has an invariant, then also its Lie algebra has an integral. Let us explain what the last expression means. Let $\mathfrak{g} \subset \mathrm{GL}(n, \mathbb{C})$ denotes the Lie algebra of $\mathcal{G}$. Then an element $Y \in \mathfrak{g}$ can be considered as a linear vector field: $\boldsymbol{x} \mapsto Y(\boldsymbol{x}) := Y\boldsymbol{x}$, for $\boldsymbol{x} \in \mathbb{C}^n$. We say that $f \in \mathbb{C}(\boldsymbol{x})$ is an integral of $\mathfrak{g}$, iff $Y(f)(\boldsymbol{x}) = \mathrm{d}f(\boldsymbol{x}) \cdot Y(\boldsymbol{x}) = 0$, for all $Y \in \mathfrak{g}$.

Proposition 1. *If $f_1, \ldots, f_k \in \mathbb{C}(\boldsymbol{x})$ are algebraically independent invariants of an algebraic group $\mathcal{G} \subset \mathrm{GL}(n, \mathbb{C})$, then they are algebraically independent first integrals of the Lie algebra $\mathfrak{g}$ of $\mathcal{G}$.*

The above facts are the starting points for applications of differential Galois methods to integrability studies. If the considered system is Hamiltonian, then we have additional constrains. First of all, the differential Galois group of variational equations is a subgroup of the symplectic group. Secondly, commutation of first integrals imposed by the Liouville integrability implies commutation of variational first integrals. Using all these facts Morales and Ramis proved the following theorem [7, 11].

Theorem 1. *Assume that a Hamiltonian system is meromorphically integrable in the Liouville sense in a neighbourhood of a phase curve Γ, and that variational equations along Γ are Fuchsian. Then the identity component of the differential Galois group of the variational equations is Abelian.*

It can happen that a considered system satisfies all conditions of the above theorem, but nevertheless it is not integrable. It is nothing strange as this theorem gives only necessary conditions for the integrability, for examples of such systems see e.g [7, 12, 13]. Such examples show a need of stronger necessary conditions for the integrability. As far as we know, there is only one method which gives effectively such conditions. It was developed by C. Simo, J.J Morales and J.-P. Ramis [7, 14] and is based on higher order variational equations. The simplest way to derive the higher order variational equations is following. Instead of (5) we introduce a

formal series

$$\boldsymbol{x} = \varphi(t) + \sum_{i=1}^{\infty} \varepsilon^i \boldsymbol{\xi}_i, \tag{8}$$

where ε is a formal parameter. Inserting the above series into both sides of (4) and comparing terms of the same order with respect to ε we obtain an infinite chain of linear non-homogeneous equations of the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\xi}_k = A(t)\boldsymbol{\xi}_k + \boldsymbol{X}_k(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{k-1}), \quad \boldsymbol{X}_1 \equiv 0, \qquad k = 1, 2, \ldots. \tag{9}$$

The first of the above equations coincides with (6). Although for $k > 1$ equations (9) are non-homogeneous, there exists an appropriate framework allowing to define the differential Galois group for such equations, for details see [7, 14]. The following theorem was announced in [7], and in [14] the reader can find its proof with the detailed theoretical exposition.

Theorem 2. *Assume that a Hamiltonian system is meromorphically integrable in the Liouville sense in a neighbourhood of the analytic phase curve* Γ*, and variational equations along* Γ *are Fuchsian. Then the identity component of the differential Galois group of* k*-th variational equations along* Γ *is Abelian for all* $k \geq 1$*.*

For applications of the above theorem see [7, 14, 12, 15, 13].

3 Particular solutions and variational equations

Let $\boldsymbol{F} = (F_1, \ldots, F_n) \in \mathbb{C}[\boldsymbol{q}]^n$, and all F_i be homogeneous polynomials of the same degree. Assume that for a non-zero $\boldsymbol{d} \in \mathbb{C}^n$, $\boldsymbol{F}(\boldsymbol{d})$ is parallel to $\boldsymbol{d}$. Then the same property has $\widetilde{\boldsymbol{d}} = \gamma\boldsymbol{d}$ for an arbitrary $\gamma \in \mathbb{C}^\star$. Thus such vectors lie on a line. This line is called a Darboux point (or a radial direction) of the force F provided that $\boldsymbol{F}$ does not vanish on it. If $\boldsymbol{F} = V'$ for a homogeneous potential $V \in \mathbb{C}[\boldsymbol{q}]$ we talk about a Darboux point of the potential. In other words, a Darboux point is a point $[d_1 : \cdots : d_n]$ in projective space $\mathbb{CP}^{n-1}$ such that $\boldsymbol{F}(\boldsymbol{d}) = \gamma\boldsymbol{d}$ for a certain $\gamma \in \mathbb{C}^\star$. In

practice Darboux points of a force are given by solutions of the non-linear equation

$$\boldsymbol{F}(\boldsymbol{d}) = \boldsymbol{d}, \tag{10}$$

however, we must remember that different solutions of the above equation can give the same Darboux point.

Darboux points are important because they give particular solutions of Newton's equations (3) (and also for Hamilton's equations (1) when $\boldsymbol{F} = V'$). In fact, if $\boldsymbol{d}$ is a Darboux point satisfying (10), then

$$\boldsymbol{q}(t) = \varphi(t)\boldsymbol{d}, \qquad \boldsymbol{p}(t) := \dot{\boldsymbol{q}}(t) = \dot{\varphi}(t)\boldsymbol{d}, \tag{11}$$

is a solution of (3) provided $\varphi(t)$ satisfies the equation

$$\ddot{\varphi} = -\varphi^{k-1}. \tag{12}$$

This equation determines a family of elliptic (for $k = 3$) or hyperelliptic curves (for $k > 3$)

$$\dot{\varphi}^2 = \frac{2}{k}\left(\varepsilon - \varphi^k\right), \tag{13}$$

depending on a parameter ε.

The variational equations along solution (11) have the form

$$\ddot{\boldsymbol{x}} = -\varphi(t)^{k-2}\boldsymbol{F}'(\boldsymbol{d})\boldsymbol{x}, \tag{14}$$

where $\boldsymbol{F}'(\boldsymbol{d})$ is the Jacobi matrix calculated at the Darboux point $\boldsymbol{d}$. In the case of Hamiltonian systems the Jacobi matrix of $\boldsymbol{F}$ is the Hessian V'' of the potential V.

Let us assume that the Jacobi matrix is diagonalizable, thus in a certain basis equations (14) take the simple form

$$\ddot{\eta}_i = -\lambda_i\varphi(t)^{k-2}\eta_i, \qquad i = 1, \ldots, n, \tag{15}$$

where $(\lambda_1, \ldots, \lambda_n)$ are eigenvalues of $\boldsymbol{F}'(\boldsymbol{d})$. By homogeneity of $\boldsymbol{F}$ one of eigenvalues, e.g. λ_n, is $\lambda_n = k - 1$.

Since variational equations split into a system of independent second order variational equations, their differential Galois group is a direct product of differential Galois groups for these particular equations

$$\mathcal{G} = \mathcal{G}_1 \times \cdots \times \mathcal{G}_n. \tag{16}$$

Let us note that by the absence in (15) terms proportional to the first order derivative $\dot{\eta}_i$ every $\mathcal{G}_i$ is an algebraic subgroup of $\mathrm{Sp}(2, \mathbb{C})$ for $i = 1, \ldots, n$ and $\mathcal{G} \subset \mathrm{Sp}(2n, \mathbb{C})$. Here we underline that this statement is valid for Hamilton's as well as for Newton's equations. Changing the independent variable $t \longrightarrow z := \varphi(t)^k/\varepsilon$, as it was proposed by Yoshida [16], we transform every equation (15) into a very particular linear equation with rational coefficients

$$z(1-z)\eta_i'' + \left(\frac{k-1}{k} - \frac{3k-2}{2k}z\right)\eta_i' + \frac{\lambda_i}{2k}\eta_i = 0, \qquad i = 1, \ldots, n. \tag{17}$$

One can recognise immediately that the above equation is the Gauss hypergeometric equation [17]

$$z(1-z)\eta_i'' + [c - (a+b+1)z]\eta_i' - ab\eta_i = 0, \tag{18}$$

with parameters

$$a + b = \frac{k-2}{2k}, \qquad ab = -\frac{\lambda_i}{2k}, \qquad c = 1 - \frac{1}{k}. \tag{19}$$

Properties of this equation are characterised by the differences of exponents at singularities $z = 0, 1, \infty$, that are

$$\lambda = 1 - c = \frac{1}{k}, \quad \mu = c - a - b = \frac{1}{2}, \quad \nu = a - b = \frac{1}{2k}\sqrt{(k-2)^2 - 8k\lambda_i}, \tag{20}$$

respectively. Transformation to the hypergeometric equation is crucial for further considerations, because for this equation its differential Galois group was analysed in detail by Kimura [18]. However, to make this fact useful, we should know a relation between differential Galois groups of

equations (15) and (18). It can be shown, see e.g. [7, 10], that the identity components of these groups are isomorphic.

The well known Kimura theorem [18] specifies all cases when all solutions of the hypergeometric equation (18) are Liouvillian. In these cases, by the Lie-Kolchin theorem the identity component of the differential Galois group is solvable. Kimura gave the complete list of hypergeometric equations with Liouvillian solutions in the form of a table with admissible differences of exponents. Application of this result to our hypergeometric equation with differences of exponents given by (20) yields the following result.

Lemma 2. *The identity component $\mathcal{G}(\lambda)^\circ$ of the differential Galois group of equation* (18) *with parameters a, b and c given by* (19) *is solvable if and only if (k, λ) belong to the following list*

1. $\left(k, p + \frac{k}{2}p(p-1)\right)$,
2. $\left(k, \frac{1}{2}\left[\frac{k-1}{k} + p(p+1)k\right]\right)$,
3. $\left(3, -\frac{1}{24} + \frac{1}{6}(1+3p)^2\right)$,
4. $\left(3, -\frac{1}{24} + \frac{3}{32}(1+4p)^2\right)$,
5. $\left(3, -\frac{1}{24} + \frac{3}{50}(1+5p)^2\right)$,
6. $\left(3, -\frac{1}{24} + \frac{3}{50}(2+5p)^2\right)$,
7. $\left(4, -\frac{1}{8} + \frac{2}{9}(1+3p)^2\right)$,
8. $\left(5, -\frac{9}{40} + \frac{5}{18}(1+3p)^2\right)$,
9. $\left(5, -\frac{9}{40} + \frac{1}{10}(2+5p)^2\right)$.

Here p is an integer.

4 Hamiltonian systems with homogeneous potentials. Results and open problems

For Hamiltonian systems we have a standard definition of the integrability in the Liouville sense and one can use directly the main theorem of the Morales-Ramis theory, namely Theorem 1. Considering systems with homogeneous potentials we divide all potentials of the same degree

into equivalence classes. Two potentials $V(\boldsymbol{q})$ and $V_A(\boldsymbol{q})$ are equivalent iff there exists a matrix $A \in \mathrm{PO}(n,\mathbb{C}) \subset \mathrm{GL}(n,\mathbb{C})$ such that $V_A(\boldsymbol{q}) := V(A\boldsymbol{q})$. Here by $\mathrm{PO}(n,\mathbb{C})$ we denote a group of $n \times n$ complex matrices A such that $AA^T = \alpha E$ where $\alpha \in \mathbb{C}^\star$ and E is the identity matrix.

For Hamiltonian systems with homogeneous potentials necessary integrability conditions obtained from the analysis of differential Galois group of the hypergeometric equation were formulated by Morales and Ramis in [7, 8]. They are expressed by means of eigenvalues $\lambda_1, \ldots, \lambda_n$ of the Hessian V'' of potential V calculated at a Darboux point.

Theorem 3. *If Hamiltonian system* (1) *with a polynomial homogeneous potential* $V(\boldsymbol{q})$ *of degree* $k > 2$ *is meromorphically integrable in the Liouville sense, then at a Darboux point values of* (k, λ_i) *for* $i = 1, \ldots, n$ *belong to the list given in Lemma 2.*

One can observe that the last eigenvalue $\lambda_n = k - 1$ does not give any restriction for the integrability. The remaining eigenvalues yield the strongest known necessary integrability conditions. But still, for potentials depending on some parameters, generally these conditions distinguish only infinite families of parameters for which the system can be integrable.

Theorem 3 was a starting point of our analysis. It is clear that the more Darboux points we have, the more constrains for the integrability follow from this theorem. Thus, it is reasonable to consider simultaneously all Darboux points for a given potential. With each Darboux point $\boldsymbol{d}$ we have associated $n-1$ non-trivial eigenvalues $\{\lambda_1, \ldots, \lambda_{n-1}\}$ of Hessian $V''(\boldsymbol{d})$. Our basic observation was that these eigenvalues taken at different Darboux points are not arbitrary.

We investigated in detail the case of two degrees of freedoms. In this case a homogeneous potential of degree k has at most k Darboux points (except the case of radial potential). For a Darboux point $\boldsymbol{d}_i$ there exists only one non-trivial eigenvalue of $V'(\boldsymbol{d}_i)$ which we denoted by λ_i. In [13] we proved the following theorems.

Theorem 4. *Assume that a homogeneous polynomial potential* $V(q_1, q_2)$ *of degree* $k > 2$

- *has* $0 < l \leq k$ *simple Darboux points with corresponding eigenvalues* λ_i, *and*
- *if* $(q_2 \pm \mathrm{i}q_1)$ *is a linear factor of* V, *then it has multiplicity 1.*

Then the shifted eigenvalues $\Lambda_i = \lambda_i - 1$ *satisfy the following relation*

$$\sum_{i=1}^{l} \frac{1}{\Lambda_i} = -1. \tag{21}$$

Theorem 5. *Let a homogeneous polynomial potential* $V(q_1, q_2)$ *of degree* $k > 2$

- *have* $0 < l \leq k$ *simple Darboux points, and*
- *if* $(q_2 \pm \mathrm{i}q_1)$ *is a linear factor of* V, *then its multiplicity* $r_\pm \neq k/2$.

Then

$$\sum_{i=1}^{l} \frac{1}{\Lambda_i} = -1 - \theta_{r_+,2}\frac{r_+}{k-2r_+} - \theta_{r_-,2}\frac{r_-}{k-2r_-}. \tag{22}$$

In (22) we used the step function

$$\theta_{x,y} := \begin{cases} 0 & \text{for} \quad x < y, \\ 1 & \text{for} \quad x \geq y. \end{cases} \tag{23}$$

These relations were obtained from the Global Residue Theorem [19] applied to an appropriate meromorphic one-form on $\mathbb{CP}^1$.

Let us explain why the above theorems are important. If the considered system is integrable, then all $\Lambda_i = \lambda_i - 1$ belong to sets of rational numbers determined by Lemma 2. In [13] we proved that there is at most a finite number of such Λ_i satisfying the conditions of Theorem 3 and relation (21) or (22) . As for a given set of λ_i there is at most a finite number of non-equivalent potentials which take these eigenvalues at Darboux points, we have the following result.

Theorem 6. *For a fixed $k > 2$, among homogeneous potentials of degree k satisfying the assumptions of Theorem 5, at most a finite number of non-equivalent potentials is integrable.*

A weak point of the above theorem is that we have to assume a kind of genericity of the potential, i.e. that Darboux points are simple. But the integrability is not a generic property, so we cannot discard non-generic potentials. Another type of the degeneracy appears when the potential does not have the maximal number of Darboux points. The lowest degenerations of this type are however tractable. A given potential does not have the maximal number of Darboux points iff it has a factor $(\alpha q_1 + \beta q_2)$, $\alpha^2 + \beta^2 \neq 0$ with multiplicity greater than one. When this multiplicity is two, we have the following.

Theorem 7. *Assume that a homogeneous potential $V \in \mathbb{C}[q_1, q_2]$ of degree $k > 2$ has a linear factor $(\alpha q_1 + \beta q_2)$, $\alpha^2 + \beta^2 \neq 0$ with the multiplicity 2. Then, Hamilton's equations* (1) *does not admit an additional rational first integral.*

For details see [13].

There exist homogeneous potentials without Darboux points. For them Theorem 3 is not applicable. In spite of the fact that such potentials are very special, we cannot discard them from investigations. In fact, using the direct method we have found one new, according to our knowledge, integrable potential [20]

$$V = (q_2 - \mathrm{i}q_1)^2(q_2 + \mathrm{i}q_1)^5, \qquad \mathrm{i} = \sqrt{-1}, \tag{24}$$

with an additional first integral quartic in momenta

$$\begin{aligned} F =& 4(p_1 - \mathrm{i}p_2)^4 z_1 + 4(p_1 - \mathrm{i}p_2)^3(p_1 + \mathrm{i}p_2)z_2 - 4(p_1 - \mathrm{i}p_2)^2 z_1^2 z_2^6 \\ & - 8\left(p_1^2 + p_2^2\right) z_1 z_2^7 - (p_1 + \mathrm{i}p_2)^2 z_2^8 - 4z_1^3 z_2^{12}, \quad z_1 = q_2 - \mathrm{i}q_1, \\ & z_2 = q_2 + \mathrm{i}q_1. \end{aligned} \tag{25}$$

In the case of potentials possessing only one Darboux point when the only non-trivial $\Lambda = \lambda - 1$ belongs to one of the appropriate sets from Lemma 2,

the additional integrability conditions can be only obtained from an analysis of higher order variational equations along the considered particular solution.

If a potential possesses a multiple Darboux point, then at this point always $\Lambda = 0$ and there is no relation of the type (21). In such a case, we can apply the higher order variational equations corresponding to the particular solution generated by this multiple Darboux point. It is interesting that an application of the higher order variational equations in this case is quite easy – namely, in calculations only the first terms of solutions of variational equations play a role. The calculations up to order $k = 11$ suggest the hypothesis that the only integrable potentials with multiple Darboux points exist for k even and they are radial potentials $V = (q_1^2 + q_2^2)^{k/2}$.

The described procedure yields the opportunity to classify all integrable polynomial homogeneous potentials of a given degree k possessing at least one Darboux point. In [15] it was shown that there is no another meromorphically integrable Hamiltonian systems with homogeneous potentials of degree 3 except these already known and collected in [21, 22]. For potentials with $k = 4$ the integrability problem was definitively solved for all potentials except one one-parameter family

$$V_2 = \frac{1-\alpha}{2} q_1^2 \left(q_1 + \mathrm{i} q_2\right)^2 + \frac{1}{4}\left(q_1^2 + q_2^2\right)^2, \quad \alpha \in \mathbb{C}^\star. \tag{26}$$

In this case, using the Morales-Ramis theory, one can only prove that the necessary condition of the integrability for V_2 is

$$\alpha \in \{p + 2p(p-1) \,|\, p \in \mathbb{Z}\} \cup \left\{\frac{3}{8} + 2p(p+1) \,|\, p \in \mathbb{Z}\right\} \cup \left\{-\frac{1}{8} + \tfrac{2}{9}(1+3p)^2 \,|\, p \in \mathbb{Z}\right\}.$$

For all details see [13].

For $k = 5$ the situation becomes more complicated. For the first time appear potentials with a non-maximal number Darboux points and these

factorisable by term $(q_2 \pm \mathrm{i}q_1)$ with the multiplicity greater than one. Application of the higher order variational equations is more complicated because our particular solution becomes a hyperelliptic curve. Thus variational equations no more transform into the product of Lamé equations. The absence of logarithmic terms in solutions of higher order variational equations still yields very strong obstructions, but it is not clear if its presence implies that the identity component of a differential Galois group of higher order variational equations is not Abelian. Under the assumption that this hypothesis is valid, one can solve the integrability problem for all potentials possessing at least one Darboux point except

$$V_{14} = \frac{1}{15}(q_2 - \mathrm{i}q_1)^2(3q_2^3 + 6\mathrm{i}q_1q_2^2 - 9q_1^2q_2 - 2\mathrm{i}q_1^3), \tag{27}$$

corresponding to $(\Lambda_1, \Lambda_2, \Lambda_3) = (-1, -1, -1)$ [23].

Starting from $k = 6$ at first sight the problem of seeking the sets of admissible Λ_i seems to simplify. Indeed, now $\lambda_i = \Lambda_i + 1$ can belong only to the first or the second family of rational numbers given in Lemma 2. One can observe that for any $k \geq 6$ there are two admissible families of Λ_i

$$\mathcal{I}_k^{(I)} = \{-1, -1, \underbrace{k-2, \ldots, k-2}_{k-2 \text{ times}}\}, \quad \mathcal{I}_k^{(II)} = \left\{-\frac{k+1}{2k}, \underbrace{k+1, \ldots, k+1}_{k-1 \text{ times}}\right\}. \tag{28}$$

The reconstruction of the potential made for $\mathcal{I}_k^{(II)}$ yields the known integrable potential with binomial coefficients introduced in paper [24]

$$V_k = \sum_{i=0}^{[k/2]} 2^{-2i}\binom{k-i}{i} q_1^{2i} q_2^{k-2i} = \frac{1}{r}\left[\left(\frac{r+q_2}{2}\right)^{k+1} + (-1)^k\left(\frac{r-q_2}{2}\right)^{k+1}\right]. \tag{29}$$

Here $[x]$ denotes the integer part of x and $r = \sqrt{q_1^2 + q_2^2}$.

Potentials corresponding to $\mathcal{I}_k^{(I)}$ are give by polynomial solutions of the second order differential equation

$$v''(z) + \frac{1-k}{z-\alpha}v'(z) + k(k-1)\frac{\alpha}{(z-\alpha)(z^2+1)}v(z) = 0, \tag{30}$$

as $V(q_1, q_2) = q_1^k v(q_2/q_1)$. Here $\alpha \in \mathbb{C}$ is a free parameter. One can observe that the case $\alpha = 0$ corresponds to the separable in the Cartesian coordinates potential $V(q_1, q_2) = q_1^k + aq_2^k$, where $a \in \mathbb{C}^\star$.

But, unfortunately, families $\mathcal{I}_k^{(I)}$ and $\mathcal{I}_k^{(II)}$ are not the only ones satisfying relation (21). For $k \in \{14, 17, 19, 26, 32, 34, 37, 71, \ldots\}$ other sets of $\{\Lambda_i\}$ satisfying (21) exist. The similar situation takes place for the generalised relation (22) for potentials factorisable by $(q_2 \pm \mathrm{i}q_1)^{r_\pm}$. Except for a family which exists for all k, we have found 'sporadic' sets of Λ_i satisfying (22) which exist for a certain k and $r_\pm$.

Obviously, for a fixed k one can try to repeat all calculations similar to those made for $k = 3, 4, 5$, but it seems that there is no chance to obtain a general integrability result valid for any $k > 5$ without other stronger necessary conditions for the integrability.

The natural question appears about the extension of this analysis for Hamiltonian systems with a greater number of degrees of freedom. Theorem 3 is formulated for any number of degrees of freedom, thus only the extensions of relations (21) and (22) are necessary. Such relations connecting $n - 1$ non-trivial eigenvalues of the Hessian matrix calculated at all Darboux points really exist and are now under consideration.

5 Newton homogeneous equations. Results and open problems

We did not find in the literature works devoted to a systematic study of the integrability of Newton's equations. Thus, we start our investigation from finding a proper notion of the integrability of such systems. Since for the vector field

$$\boldsymbol{v}(\boldsymbol{q}, \boldsymbol{p}) := \sum_{i=1}^{n} \left(p_i \frac{\partial}{\partial q_i} - F_i(\boldsymbol{q}) \frac{\partial}{\partial p_i} \right), \qquad \boldsymbol{p} := \dot{\boldsymbol{q}}, \tag{31}$$

defined by the right hand sides of Newton's equation (3) we have

$$\operatorname{div}(\boldsymbol{v}) := \sum_{i=1}^{n} \frac{\partial v_i}{\partial x_i} = 0, \qquad \boldsymbol{x} = (\boldsymbol{q}, \boldsymbol{p}), \tag{32}$$

thus

$$\mu = \mathrm{d}q_1 \wedge \cdots \wedge \mathrm{d}q_n \wedge \mathrm{d}p_1 \wedge \cdots \wedge \mathrm{d}p_n, \tag{33}$$

is an invariant $2n$-form. According to the Jacobi theorem [25], if a system of m first order autonomous differential equations has $m-2$ functionally independent first integrals and an invariant m-form (the Jacobi Last Multiplier), then it is integrable by quadratures. Thus, we adopt the following definition.

Definition 1. We say that a system of n Newton's equations of the form (3) is integrable in the Jacobi sense iff it admits $2n-2$ functionally independent first integrals.

It is convenient to divide the Newton equations into equivalence classes. We say that forces $\boldsymbol{F}$ and $\widetilde{\boldsymbol{F}}$ are equivalent if there exists a non-singular matrix $A \in \mathrm{GL}(n,\mathbb{C})$ such that $\widetilde{\boldsymbol{F}}(\boldsymbol{q}) = A^{-1}\boldsymbol{F}(A\boldsymbol{q})$. Obviously, if Newton's equations with force $\boldsymbol{F}$ are integrable in the Jacobi sense, then Newton's equations with forces equivalent to $\boldsymbol{F}$ are also integrable in the Jacobi sense. It should be mentioned that the set of all Newton's systems with homogeneous forces of degree $k-1$ contains all Hamiltonian systems with homogeneous potentials of degree k. Hence analysing Newton's equations we have to distinguish clearly the non-Hamiltonian class of such equations.

Our main effort was to find the necessary conditions for the integrability in the Jacobi sense. More precisely, we wanted to find such conditions applying the main ideas of the Morales-Ramis theory.

As was shown in Section 3 for the Newton homogeneous equations particular solutions related to Darboux points also exist. If Jacobi matrix $\boldsymbol{F}'(\boldsymbol{d})$ at a Darboux point $\boldsymbol{d}$ is diagonalisable, then variational equations have the form of a product of second order equations, and their differential Galois group is a direct product

$$\mathcal{G} = \mathcal{G}_1 \times \cdots \times \mathcal{G}_n \subset \mathrm{Sp}(2n,\mathbb{C}), \tag{34}$$

with $\mathcal{G}_i \subset \mathrm{Sp}(2,\mathbb{C})$, for $i = 1,\ldots,n$. It should be remarked here that, generally, a differential Galois group of variational equations along a particular

solution is a subgroup of $\mathrm{GL}(2n,\mathbb{C})$. Only the specific form of particular solutions corresponding to Darboux points guarantees that it is a subgroup of $\mathrm{Sp}(2n,\mathbb{C})$.

The starting point is Lemma 1 which applied to the considered system implies that if it is integrable in the Jacobi sense, then the differential Galois group of the variational equations possesses $2n-2$ rational invariants. To find explicit obstructions for the integrability from this implication it is useful to consider the Lie algebra of the differential Galois group. Decomposition (34) implies that the Lie algebra $\mathfrak{g}$ of $\mathcal{G}$ is a direct sum

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_n, \qquad \mathfrak{g}_i \subset \mathrm{sp}(2,\mathbb{C}). \tag{35}$$

Now, we can apply Proposition 1. Considering first integrals of a Lie subalgebra $\mathfrak{g}$ of $\mathrm{sp}(2,\mathbb{C})$ it is convenient to use the following facts. An element Y of Lie algebra $\mathrm{sp}(2n,\mathbb{C})$, considered as a linear vector field, is a global Hamiltonian vector field with a quadratic polynomials homogeneous Hamiltonian $H : \mathbb{C}^{2n} \to \mathbb{C}$. In this way we can identify $\mathrm{sp}(2n,\mathbb{C})$ with a $\mathbb{C}$-linear vector space of homogeneous polynomials of degree 2 with the canonical Poisson bracket as the Lie bracket. Thus for a Lie algebra $\mathfrak{g} \subset \mathrm{sp}(2n,\mathbb{C})$, a rational function is a first integral of $\mathfrak{g}$, iff $\{H,f\} = 0$, for all $H \in \mathfrak{g}$. In the considered case Lie algebra $\mathfrak{g}$ is a direct sum and this allows us to show the following.

Theorem 8. *Assume that the Newton system* (3) *with polynomial homogeneous right-hand sides of degree $l \geq 2$ is integrable in the Jacobi sense. Then if $\boldsymbol{d}$ is a Darboux point such that $\boldsymbol{F}'(\boldsymbol{d})$ is semi-simple, then the identity component of the differential Galois group of variational equations along a particular solution defined by $\boldsymbol{d}$ is Abelian.*

Using the above theorem and Lemma 2 we obtained a result similar to Theorem 3.

Theorem 9. *Assume that the Newton system* (3) *with polynomial homogeneous right-hand sides of degree $k-1 \geq 2$ is integrable in the Jacobi sense. Then if $\boldsymbol{d}$ is a Darboux point such that $\boldsymbol{F}'(\boldsymbol{d})$ is semi-simple and*

$\lambda_1, \ldots, \lambda_n$ are the corresponding eigenvalues, then (k, λ_i) for $i = 1, \ldots, n$ belong to the list from Lemma 2.

It is remarkable that the necessary conditions for the Liouville integrability of Hamiltonian systems with a homogeneous potential and the integrability in the Jacobi sense are the same.

Then we focused our attention on the case $n = 2$. Our aim was to perform a global analysis similar to that we did for Hamiltonian systems. The obtained result is the following.

Theorem 10. *Let us assume that $\boldsymbol{F} = (F_1, F_2)$ is a homogeneous polynomial force of degree $k - 1$, and moreover*

a) $\boldsymbol{F}$ admits $0 < l \leq k$ simple Darboux points $\boldsymbol{d}_i$ with non-trivial eigenvalues λ_i of $\boldsymbol{F}'(\boldsymbol{d}_i)$, for $i = 1, \ldots, l$;

b) if polynomials $G(q_1, q_2) := F_2(q_1, q_2) - q_2 F_1(q_1, q_2)$ and $F_1(q_1, q_2)$ have a common linear factor $(q_2 - s_i q_1)$, then

$$\mathrm{mult}(G, q_2 - s_i q_1) = \mathrm{mult}(F_1, q_2 - s_i q_1).$$

Under the above assumptions $\Lambda_i := \lambda_i - 1$ satisfy

$$\sum_{i=1}^{l} \frac{1}{\Lambda_i} = -1. \tag{36}$$

We proved this theorem as in the Hamiltonian case, i.e. we applied the Global Residue Theorem to an appropriately defined differential one-form ω on $\mathbb{CP}^1$. A Darboux point $\boldsymbol{d}_i$ is a simple pole of this form with residue $1/\Lambda_i$. In a case when $\mathrm{mult}(G, q_2 - s_i q_1) \neq \mathrm{mult}(F_1, q_2 - s_i q_1)$ relation among Λ_i becomes more involved. The differential form ω has additional poles at s_i with non-vanishing residues which depend on coefficients of (F_1, F_2). For these residues we have no restrictions and this leads to considerable complications.

We performed a complete analysis for the non-factorisable Newton equations of homogeneity degree 2, 3 and 4, except for certain cases of forces

which do not admit Darboux points. Among this class we also found one superintegrable system with three first integrals

$$\begin{aligned} &\ddot{q}_1 = -q_2^{k-1}, \qquad \ddot{q}_2 = 0, \\ &I_1 = p_2, \qquad I_2 = \frac{1}{k+1} q_2^{k+1} + p_2(q_2 p_1 - q_1 p_2), \qquad I_3 = p_1 p_2 + \frac{1}{k} q_2^k. \end{aligned} \tag{37}$$

Let us note that almost all obtained integrable systems are Hamiltonian. This is related to the fact that most of the integrable systems possess first integrals quadratic in velocities. If the quadratic form of velocities in this first integral is non-degenerated, then system is Hamiltonian and this first integral becomes the Hamilton function. All details of the integrability analysis of Newton's equations will be published in [26].

References

[1] Lagutinskii, M. N.: *Partial Algebraic Integrals*, Kharkov, 1908, in Russian.

[2] Moulin Ollagnier, J., Nowicki, A., and Strelcyn, J.-M.: On the non-existence of constants of derivations: the proof of a theorem of Jouanolou and its development. *Bull. Sci. Math.*, 119(3), 195–233, 1995.

[3] Maciejewski, A. J. and Przybylska, M.: Normal forms and Kovalevskaya-Lagutinskii-Levelt method, in preparation.

[4] Ziglin, S. L.: Branching of solutions and non-existence of first integrals in Hamiltonian mechanics. I. *Functional Anal. Appl.*, 16, 181–189, 1982.

[5] Ziglin. S. L.: Branching of solutions and non-existence of first integrals in Hamiltonian mechanics. II. *Functional Anal. Appl.*, 17, 6–17, 1983.

[6] Maciejewski, A. J. and Przybylska, M.,: Differential Galois obstructions for non-commutative integrability, *Nonlinearity*, submitted.

[7] Morales Ruiz, J. J.: *Differential Galois theory and non-integrability of Hamiltonian systems*, volume 179 of *Progress in Mathematics*, Birkhäuser Verlag, Basel, 1999.

[8] Morales-Ruiz, J. J. and Ramis, J. P.: A note on the non-integrability of some Hamiltonian systems with a homogeneous potential, *Methods Appl. Anal.*, 8(1), 113–120, 2001.

[9] Audin, M.: *Les systèmes hamiltoniens et leur intégrabilité*, Cours Spécialisés 8, Collection SMF, SMF et EDP Sciences, Paris, 2001.

[10] Baider, A., Churchill, R. C., Rod, D. L., and Singer, M. F.: On the infinitesimal geometry of integrable systems, in *Mechanics day (Waterloo, ON, 1992)*, volume 7 of *Fields Inst. Commun.*, Amer. Math. Soc., Providence, RI, 5–56, 1996.

[11] Morales-Ruiz, J. J. and Ramis, J. P.: Galoisian obstructions to integrability of Hamiltonian systems. I, *Methods Appl. Anal.*, 8(1), 33–95, 2001.
[12] Maciejewski, A. J., Przybylska, M., and Weil, J.-A.: Non-integrability of the generalized spring-pendulum problem, *J. Phys. A, Math. Gen.*, 37(7), 2579–2597, 2004.
[13] Maciejewski, A. J. and Przybylska, M., Darboux points and integrability of Hamiltonian systems with homogeneous polynomial potential, *J. Math. Phys.*, 46(6), 062901, 33 pp., 2005.
[14] Morales-Ruiz, J. J., Ramis, J. P., and Simó, C.: Integrability of Hamiltonian systems and differential Galois groups of higher variational equations, 2006, submitted to *Ann. Sci. École Norm. Sup.*, available at http://www.maia.ub.es/dsg/2005/0506simo.pdf.
[15] Maciejewski, A. J. and Przybylska, M.: All meromorphically integrable 2D Hamiltonian systems with homogeneous potential of degree 3, *Phys. Lett. A*, 327(5-6), 461–473, 2004.
[16] Yoshida, H.: A criterion for the nonexistence of an additional integral in Hamiltonian systems with a homogeneous potential, *Phys. D*, 29(1-2), 128–142, 1987.
[17] Whittaker, E. T. and Watson, G. N.: *A Course of Modern Analysis*, Cambridge University Press, London, 1935.
[18] Kimura, T.: On Fuchsian differential equations reducible to hypergeometric equations by linear transformations, *Funkcial. Ekvac.*, 13, 213–232, 1970.
[19] Miranda, R.: *Algebraic curves and Riemann surfaces*, American Mathematical Society, Providence, RI, 1995.
[20] Nakagawa, K., Maciejewski, A. J., and Przybylska, M.: New integrable Hamiltonian system with quartic in momenta first integral, *Phys. Lett. A*, 343(1-3), 171–173, 2005.
[21] Hietarinta, J.: A search for integrable two-dimensional Hamiltonian systems with polynomial potential, *Phys. Lett. A*, 96(6), 273–278, 1983.
[22] Hietarinta, J.: Direct methods for the search of the second invariant, *Phys. Rep.*, 147(2), 87–154, 1987.
[23] Przybylska, M.: Darboux points and integrability of Hamiltonian systems with homogeneous polynomial potential II, 2006, in preparation.
[24] Ramani, A., Dorizzi, B., and Grammaticos, B.: Painlevé conjecture revisited, *Phys. Rev. Lett.*, 49(21), 1539–1541, 1982.
[25] Arnold, V. I., Kozlov, V. V., and Neishtadt, A. I.: *Dynamical systems. III*, volume 3 of *Encyclopaedia of Mathematical Sciences*, Springer-Verlag, Berlin-New York, 1988.
[26] Przybylska, M.: Galoisian integrability analysis for homogeneous Newton equations, 2006, in preparation.

From Homological Perturbation to Spectral Sequences: a Case Study[1]

A. Romero

Universidad de La Rioja, Departamento de Matemáticas y Computación,
ana.romero@dmc.unirioja.es

Abstract

In this paper, a program computing spectral sequences is reported. The theoretical algorithm supporting this program is based on effective homology and homological perturbation techniques. We illustrate the fundamental ideas of this algorithm by means of an example related to the famous Serre spectral sequence.

Keywords: Symbolic computation, spectral sequences, Serre spectral sequence, constructive algebraic topology, homological perturbation.

1 Introduction

Homological methods are important in the field of formal integrability of PDE systems. In particular, Spencer cohomology is a tool which can be used to determine the involutivity of a system (see [1] and [2]). Recently, Sergeraert has developed some programs computing the Koszul homology of polynomial ideals[2], a notion closely connected to Spencer cohomology. This program is based on his theory of effective homology and, in particular, uses intensively homological perturbation techniques.

In parallel, the author of the present paper has obtained algorithms computing spectral sequences of filtered chain complexes, even when the

[1] Partially supported by Comunidad Autónoma de La Rioja, project ANGI-2005/19, and by Ministerio de Educación y Ciencia, project MTM2006-06513 and by NEST-Adventure contract 5006 (GIFT).

[2] A pdf presentation for several talks about this subject can be found in
`www-fourier.ujf-grenoble.fr/~sergerar/Papers/Koszul.pdf`

original chain complex is not of finite type, on condition that this chain complex is an object *with effective homology*. The aim of this paper is to illustrate, by means of a significant example, the corresponding computer program which has been written down to implement this algorithm, in order to facilitate its further application to reach more knowledge in the case of the Koszul homology and the Spencer cohomology.

The organization of the paper is the following. In Section 2 we introduce an example of application of the Serre spectral sequence, showing the non-constructive nature of this spectral sequence and comparing this method with the effective homology one. Then, in the next section, some necessary definitions and results about spectral sequences and effective homology are explained. Sections 4 and 5 show two important examples of application of the effective homology technique, including as a particular case the example presented before. In Section 6 we explain how the effective homology method can also be used to compute spectral sequences, and we illustrate it again by means of our particular case study. Finally, the paper ends with a section of conclusions and further work.

2 An example of spectral sequence

A Spectral Sequence is a family of "pages" $\{E^r_{p,q}, d^r_{p,q}\}$ of differential bigraded modules, each page being made of the homology groups of the preceding one. If we know the stage r in the spectral sequence (E^r, d^r) we can build the bigraded module at the stage $r+1$, E^{r+1}, but this cannot define the next differential d^{r+1} which therefore must be independently defined too.

One of the first examples of spectral sequence is due to Serre (using previous work of Leray), involving the fibrations $G \hookrightarrow E \to B$, where G is the *fiber* space, B the *base* space and E the *total* space. The three spaces were initially topological spaces, but this notion of fibration can be generalized to many other situations, in particular the case where B is a simplicial set and F is a simplicial group. The total space E can be considered as a *twisted product* of B and G, and the underlying twisting

operator $\tau : B \to G$ explains how the twisted product $E = B \times_\tau G$ is different from the trivial product $B \times G$. The definition of the associated spectral sequence is given by the following theorem, and it can be found in [3].

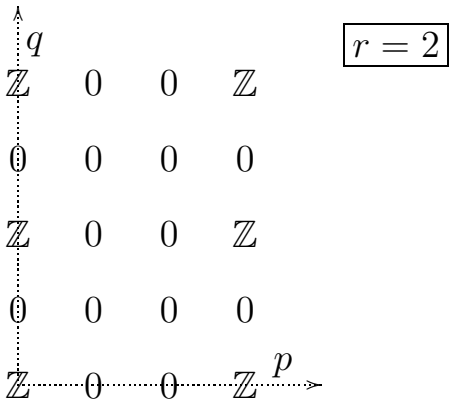
Theorem. *Let $G \hookrightarrow E \to B$ be a fibration with a base space B simply connected. Then a first quadrant spectral sequence $\{E^r_{p,q}, d^r_{p,q}\}_{r\geq 2}$ is defined with $E^2_{p,q} = H_p(B, H_q(G))$ and $E^r_{p,q} \Rightarrow H_{p+q}(E)$.*

It is frequently thought this spectral sequence is a process allowing one to compute the groups $H_*(E)$ when the groups $H_*(B)$ and $H_*(G)$ are known. But in general this is false, because the differentials $d^r_{p,q}$ are unknown and in many cases we do not have the necessary information to compute them. And even if we know all the differentials $d^r_{p,q}$ and we can reach the limit groups $E^\infty_{p,q}$, we must deal with a (not always solvable) extension problem to determine the homology groups $H_*(E)$. This means that the Serre spectral sequence is not an algorithm that allows us to compute the homology groups of the total space of the fibration, but in fact it is a (rich and interesting) set of relations between the groups $H_*(G)$, $H_*(E)$ and $H_*(B)$. Moreover, we must emphasize here that in many cases this spectral sequence can not be determined. To illustrate this non-constructive nature, we include here one of the initial examples of Serre, considering the beginning of his calculations.

The computation of sphere homotopy groups is known as a difficult problem in algebraic topology. It is not hard to prove that $\pi_n(S^k) = 0$ for $n < k$ and $\pi_k(S^k) = \mathbb{Z}$. Furthermore, in 1937 Freudenthal proved that $\pi_4(S^2) = \mathbb{Z}_2$, and at the beginning of the fifties Serre computed many sphere homotopy groups, being his famous spectral sequence the main tool to obtain these calculations. In particular, Serre proved $\pi_6(S^3)$ has 12 elements, but he was unable to choose between the two possible options $\mathbb{Z}_{12}$ and $\mathbb{Z}_2 + \mathbb{Z}_6$.

For instance, how can we use the Serre spectral sequence to determine the homotopy groups of S^3? It is well-known that $\pi_n(S^3) = 0$ for $n < 3$ and $\pi_3(S^3) = \mathbb{Z}$. To compute $\pi_4(S^3)$, we consider a fibration $F_2 \hookrightarrow X_4 \to S^3$,

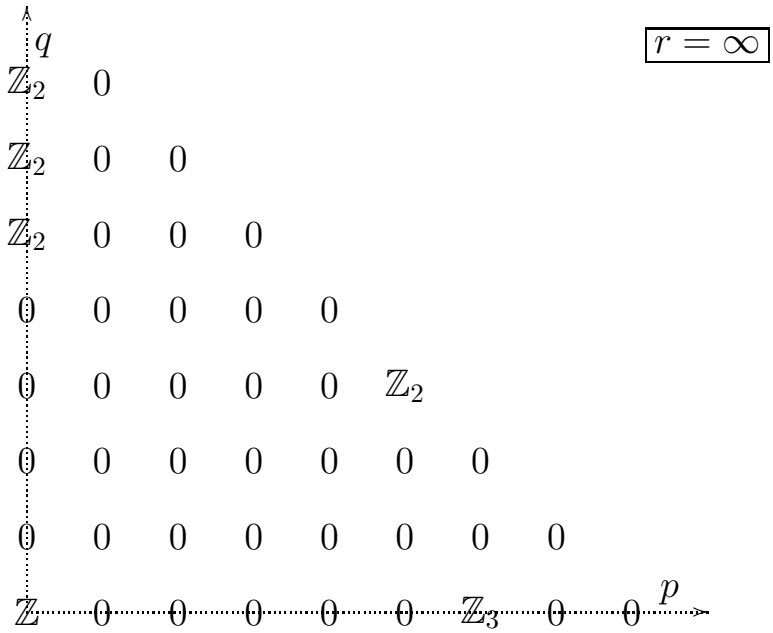
where $F_2 = K(\mathbb{Z}, 2)$ is an Eilenberg-MacLane space. The beginning of the spectral sequence (that is, the groups $E^2_{p,q}$) is determined by means of the well-known homology groups of S^3 and F_2; the result is shown in the next figure.



We observe that all the arrows $d^2_{p,q} : E^2_{p,q} \to E^2_{p-2,q+1}$ are necessarily null so that the groups $E^3_{p,q}$ are equal to the corresponding $E^2_{p,q}$. But problems arise when trying to determine the differential maps $d^3_{p,q}$. The arrow $d^3_{3,0}$ must be an isomorphism, but to know the arrows $d^3_{3,2q}$ some other (extra) information than which is given by the spectral sequence itself is necessary. In this particular case, a specific tool (the multiplicative structure of the cohomology) gives the solution, the arrow $d^3_{3,2q} : \mathbb{Z} \to \mathbb{Z}$ is the multiplication by $q + 1$. This implies the $E^3_{3,2q}$ die and $E^r_{0,2q} = \mathbb{Z}_q$ for $4 \leq r \leq \infty$ and $q \geq 2$. In this way, the Serre spectral sequence entirely gives the homology groups $H_0(X_4) = \mathbb{Z}$, $H_{2n}(X_4) = \mathbb{Z}_n$ for $n \geq 2$ and the other $H_n(X_4)$ are null. In particular, the Hurewicz theorem and the long exact sequence of homotopy imply that $\pi_4(S^3) = \pi_4(X_4) = H_4(X_4) = \mathbb{Z}_2$, a result known by Freudenthal.

Then, a new fibration $F_3 \hookrightarrow X_5 \to X_4$ is considered to determine $\pi_5(S^3)$, where $F_3 = K(\mathbb{Z}_2, 3)$ is chosen because $\pi_4(X_4) = \mathbb{Z}_2$. In this case Serre was also able to obtain all the necessary ingredients to compute the maps $d^r_{p,q}$ which play an important role in the beginning of the associated spectral sequence. The main tool (extra information) are the multiplicative structure in cohomology and more generally the module structure with respect to the Steenrod algebra $\mathcal{A}_2$. The final groups $E^\infty_{p,q}$ (with $p + q \leq 8$) of this

spectral sequence are showed in the following figure.



Again the Hurewicz theorem and the long homotopy exact sequence imply $\pi_5(S^3) = \pi_5(X_4) = \pi_5(X_5) = H_5(X_5) = \mathbb{Z}_2$; it was the first important result obtained by Serre.

These two examples illustrate the fact that the computation of the Serre spectral sequence is not an easy task and in some situations some other information than which is given by the spectral sequence itself is needed. In other cases, the computation of the Serre spectral sequence is in fact *not possible*, since some differentials $d^r_{p,q}$ can not be determined by any other mean (we do not have the necessary extra information). Therefore, as we have said, the Serre spectral sequence is not an algorithm that allows us to compute $H_*(E)$ if $H_*(B)$ and $H_*(G)$ are known.

On the contrary, the method based on the notion of *object with effective homology* provides real algorithms for the computation of homology groups of many complicated spaces. In particular, this technique can be applied to compute the homology groups of the total space of fibrations when the base and fiber spaces are objects with effective homology, replacing in this way the Serre spectral sequence.

Based on the effective homology method, the Kenzo system [4] was developed. Kenzo is a Common Lisp program devoted to Symbolic Computation in Algebraic Topology that works with rich and complex algebraic structures (chain complexes, differential graded algebras, simplicial sets, simplicial groups, morphisms between these objects, etc). It implements the effective homology method for the computation of homology groups of different spaces. As an example, we show in the following lines how this program can be used to compute the homology groups of the space X_5 introduced in this section.

First of all, the object X_5 must be built, and we can do it by means of the following instructions. We do not explain the Lisp functions that appear here but most of them are self-explanatory.

```
> (setf s3 (sphere 3))
[K1 Simplicial-Set]
> (setf f3 (z-whitehead s3 (chml-clss s3 3)))
[K37 Fibration K1 -> K25]
> (setf x4 (fibration-total f3))
[K43 Simplicial-Set]
> (setf f4 (z2-whitehead x4 (chml-clss x4 4)))
[K292 Fibration K43 -> K278]
> (setf x5 (fibration-total f4))
[K298 Simplicial-Set]
```

The result of the last evaluation is the object K_{298}, which is an instance of the class `Simplicial-Set`, and is located through the symbol `x5`. We can ask for the effective homology of X_5:

```
> (efhm x5)
[K608 Homotopy-Equivalence K298 <= K598 => K594]
```

We will see in the following section what a homotopy equivalence is. The homology groups of X_5 are then easily computable using its effective homology, for example, in degrees 5 and 6 the known results $H_5(X_5) = \mathbb{Z}_2$ and $H_6(X_5) = \mathbb{Z}_6$ are obtained.

```
> (homology x5 5)
Homology in dimension 5 :
Component Z/2Z
---done---
> (homology x5 6)
Homology in dimension 6 :
Component Z/6Z
---done---
```

But although the effective homology method allows us to compute the homology groups of the total space of a fibration (in our example, X_5), the structure of the Serre spectral sequence can also give useful informations about the involved construction, sometimes even more interesting than the final homology groups. In fact, both techniques can be combined and it can be seen the effective homology method can also be applied to obtain an algorithm that compute, as a by-product, the relevant spectral sequence (with the whole set of its components).

This algorithm combining both spectral sequence and effective homology methods has been concretely implemented as an extension of the Kenzo program, allowing the user to compute in an easy way spectral sequences associated with filtered complexes, and as a particular case, Serre spectral sequences. For example, as we will see in Section 6, with these programs all the ingredients of the spectral sequence associated with the fibration $F_3 \hookrightarrow X_5 \rightarrow X_4$ introduced in this section (groups, differential maps, convergence...) are easily obtained by means of simple instructions, without needing any extra information.

3 Preliminaries

In this section, some necessary concepts about spectral sequences and effective homology are presented. First of all, we include here some basic definitions and results of Algebraic Topology that can be found, for instance, in [5].

Definition 1. A *chain complex* is a pair (C, d) where $C = \{C_n\}_{n\in\mathbb{Z}}$ is a graded Abelian group and $d = \{d_n : C_n \to C_{n-1}\}_{n\in\mathbb{Z}}$ (*the differential map*) is a graded group homomorphism of degree -1 such that $d_{n-1}d_n = 0\ \forall n \in \mathbb{Z}$. The graded *homology group* of the chain complex C is $H(C) = \{H_n(C)\}_{n\in\mathbb{N}}$, with

$$H_n(C) = \text{Ker } d_n/\text{Im } d_{n+1}$$

A *chain complex homomorphism* $f : (A, d_A) \to (B, d_B)$ between two chain complexes (A, d_A) and (B, d_B) is a graded group homomorphism (degree 0) such that $fd_A = d_Bf$.

Definition 2. A *filtration* F of a chain complex (C, d) is a family of subchain complexes $F_pC \subset C$ such that

$$\cdots \subset F_{p-1}C_n \subset F_pC_n \subset F_{p+1}C_n \subset \cdots \quad \forall n \in \mathbb{Z}$$

Note 1. A filtration F on C induces a filtration on the graded homology group $H(C)$. Let $i_p : F_pC \hookrightarrow C$ the p-injection, then $F_p(H(C)) = H(i_p)(H(F_p(C)))$.

Definition 3. A filtration F of a chain complex C is said to be *bounded* if for each degree n there are integers $s = s(n) < t = t(n)$ such that $F_sC_n = 0$ and $F_tC_n = C_n$.

Definition 4. A $\mathbb{Z}$-*bigraded module* is a family of $\mathbb{Z}$-modules $E = \{E_{p,q}\}_{p,q\in\mathbb{Z}}$. A *differential* $d : E \to E$ of bidegree $(-r, r-1)$ is a family of homomorphisms of $\mathbb{Z}$-modules $d_{p,q} : E_{p,q} \to E_{p-r,q+r-1}$ for each $p, q \in \mathbb{Z}$, with $d_{p,q} \circ d_{p+r,q-r+1} = 0$. The *homology* of E under this differential is the bigraded module $H(E) \equiv H(E, d) = \{H_{p,q}(E)\}_{p,q\in\mathbb{Z}}$ with $H_{p,q}(E) =$ Ker $d_{p,q}/$ Im $d_{p+r,q-r+1}$

Definition 5. A *spectral sequence* $E = \{E^r, d^r\}$ is a family of $\mathbb{Z}$-bigraded modules E^1, E^2,..., each provided with a differential $d^r = \{d^r_{p,q}\}$ of bidegree $(-r, r-1)$ and with isomorphisms $H(E^r, d^r) \cong E^{r+1}$, $r = 1, 2, \ldots$

Definition 6. A spectral sequence $E = \{E^r, d^r\}$ is said to be *convergent* if for every $p, q \in \mathbb{Z}$ there exists $r_{p,q} \in \mathbb{N}$ such that $d^r_{p,q} = 0 = d^r_{p+q,q-r+1}$ for all $r \geq r_{p,q}$.

If $E = \{E^r, d^r\}$ is convergent, then $E^r_{p,q} = E^{r_{p,q}}_{p,q}$ $\forall r \geq r_{p,q}$. We define $E^\infty_{p,q} = E^{r_{p,q}}_{p,q}$, which can be seen as the "limit" of the groups $E^r_{p,q}$ when $r \to \infty$.

Definition 7. A spectral sequence (E^r, d^r) is said to *converge* to a graded module H (denoted by $E^1 \Rightarrow H$) if there is a filtration F of H and for each p isomorphisms $E^\infty_p \cong F_pH/F_{p-1}H$ of graded modules.

Theorem 1. (Theorem 3.1, Chapter XI, in [5]) *Each filtration F of a chain complex (C, d) determines a spectral sequence (E^r, d^r), defined by*

$$E^r_{p,q} = \frac{Z^r_{p,q} \cup F_{p-1}C_{p+q}}{dZ^{r-1}_{p+r-1,q-r+2} \cup F_{p-1}C_{p+q}}$$

where $Z^r_{p,q}$ is the submodule $[a|\ a \in F_pC_{p+q}, d(a) \in F_{p-r}C_{p+q-1}]$, and the differential map $d^r : E^r_{p,q} \to E^r_{p-r,q+r-1}$ is the homomorphism induced on these subquotients by the differential on C, $d : C \to C$.
If F is bounded, $E^1 \Rightarrow H(C)$; more explicitly,

$$E^\infty_{p,q} \cong F_p(H_{p+q}C)/F_{p-1}(H_{p+q}C)$$

(with $F_p(HC)$ induced by the filtration F, as explained in Note 1).

On the other hand, we also include here some definitions and fundamental ideas about the effective homology method. More details can be found in [6].

Definition 8. A *reduction* $\rho : D \Rightarrow C$ between two chain complexes is a triple (f, g, h) where

a) the components $f : D \to C$ and $g : C \to D$ are chain complex morphisms;

b) the component $h : D \to D$ is a graded group homomorphism of degree +1;

c) the following relations are satisfied
$fg = \mathrm{id}_C$; $gf + d_Dh + hd_D = \mathrm{id}_D$; $fh = 0$; $hg = 0$; $hh = 0$

Remark 1. These relations express that D is the direct sum of C and a contractible (acyclic) complex. This decomposition is simply $D = \text{Ker } f \oplus \text{Im } g$, with $\text{Im } g \cong C$ and $H_*(\text{Ker } f) = 0$. In particular, this implies that the graded homology groups $H_*(D)$ and $H_*(C)$ are canonically isomorphic.

Definition 9. A *(strong chain) equivalence* between the complexes C and E (denoted by $C \Longleftrightarrow E$) is a triple (D, ρ, ρ') where D is a chain complex, and ρ and ρ' are reductions from D over C and E respectively:

$$C \overset{\rho}{\Longleftarrow} D \overset{\rho'}{\Longrightarrow} E$$

Note 2. An effective chain complex is essentially a free chain complex C where each group C_n is finitely generated, and there is an algorithm that returns a $\mathbb{Z}$-base in each degree n (for details, see [6]).

Definition 10. An *object with effective homology* is a triple (X, HC, ε) where HC is an effective chain complex and ε is an equivalence between a free chain complex canonically associated with X and HC.

Note 3. It is important to understand that in general the HC component of an object with effective homology is *not* made of the homology groups of X; this component HC is a free $\mathbb{Z}$-chain complex of finite type, in general with a non-null differential, allowing to *compute* the homology groups of X; the justification is the equivalence ε.

In this way, the notion of object with effective homology makes it possible to compute homology groups of complicated spaces by means of homology groups of effective complexes (which can easily be obtained using some elementary operations).

The next theorem is a very useful tool that will be considered to obtain the effective homology of several spaces. In particular, it is one of the main ingredients for the proof of the effective homology version of the Serre spectral sequence, explained in Section 5. The general idea of this theorem is that given a reduction, if we *perturb* the *big* complex then it

is possible to perturb the *small* one so that we obtain a new reduction between the perturbed complexes. A reference where this theorem can be found is [7].

Theorem 2 (Basic Perturbation Lemma, BPL)**.** *Let $\rho = (f, g, h)$ be a reduction $\rho : C \Rightarrow D$ and δ a* perturbation *of d_C, that is, an operator defined on C of degree -1 satisfying the relation $(d_C + \delta) \circ (d_C + \delta) = 0$. Furthermore, the composite function $h \circ \delta$ is assumed* locally nilpotent*, that is, $\forall x \in C$, there exists $n \in \mathbb{N}$ such that $(h \circ \delta)^n x = 0$. Then a new reduction $\rho' : C' \Rightarrow D'$, $\rho' = (f', g', h')$, can be constructed where:*

a) C' is the chain complex obtained from C by replacing the old differential d_C by $(d_C + \delta)$,

b) the new chain complex D' is obtained from the chain complex D by replacing the old differential d_D by $(d_D + \bar{\delta})$, with $\bar{\delta} = f \circ \delta \circ \phi \circ g = f \circ \psi \circ \delta \circ g$,

c) $f' = f \circ \psi = f \circ (Id - \delta \circ \phi \circ h)$,

d) $g' = \phi \circ g$,

e) $h' = \phi \circ h = h \circ \psi$,

where the operators ϕ and ψ are defined by

$$\phi = \sum_{i=0}^{\infty} (-1)^i (h \circ \delta)^i; \quad \psi = \sum_{i=0}^{\infty} (-1)^i (\delta \circ h)^i = Id - \delta \circ \phi \circ h,$$

(the series are convergent thanks to the locally nilpotency of $h \circ \delta$)

In Sections 4 and 5, we present two important examples of application of this theorem.

4 Effective homology of a bicomplex

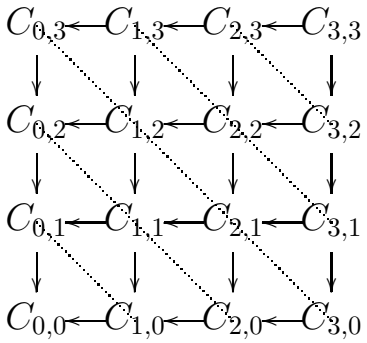
The computation of the effective homology of a bicomplex (double complex) is a very simple example where the BPL can be applied. First of all, let us recall the definition of a bicomplex.

Definition 11. A *bicomplex* (or *double complex*) is a bigraded module $C = \{C_{p,q}\}_{p,q\in\mathbb{Z}}$ provided with morphisms $d'_{p,q} : C_{p,q} \to C_{p-1,q}$ and $d''_{p,q} : C_{p,q} \to C_{p,q-1}$ satisfying $d'd' = 0, d''d'' = 0$ and $d'd'' + d''d' = 0$. Then, we define the *totalization* $(T(C), d)$ of the bicomplex C as the chain complex given by

$$T_n(C) = \bigoplus_{p+q=n} C_{p,q}$$

and differential map $d = d' + d''$.

This notion is easy to understand by means of the following diagram, where the horizontal arrows are the maps $d'_{p,q}$ and the vertical arrows are the differentials $d''_{p,q}$. The totalization is represented by the diagonals.



From now on, we consider C to be a *first quadrant* bicomplex, that is, such that $C_{p,q} = 0$ if $p < 0$ or $q < 0$.

The identity $d''d'' = 0$ implies that for a fixed $i \in \mathbb{N}$ the column $C^i = \{C_{i,n}\}_{n\in\mathbb{N}}$ is a chain complex, so it makes sense to look for the relation between the homologies of the columns C^i and that of the totalization $T(C)$. Let us suppose that the columns C^i are objects with effective homology, in particular such that there exist reductions $C^i \Rightarrow HC^i$ with HC^i an effective complex for all $i \in \mathbb{N}$. Then we are going to construct a new effective complex HC which provides us the effective homology of the totalization $T(C)$.

As a first step, we build a chain complex $(T(C), d')$ totalization of the bicomplex C, but where only the vertical arrows are considered. Using the reductions of each C^i over HC^i, it is easy to construct a reduction of

$(T(C), d')$ over a new chain complex $(T(HC), \bar{d}')$ which is the totalization of a new double complex with columns HC^i (and where all the horizontal arrows are null). It is clear that in each degree the component $T_i(HC)$ is a sum of finite type groups, so that the chain complex $T(HC)$ is an effective complex.

The reduction $(T(C), d') \Rightarrow (T(HC), \bar{d}')$ is the first ingredient for the application of the BPL. Then, we also need a perturbation of the differential d', which is defined by the horizontal arrows, $\delta = d''$. It is not difficult to see that the composition $h \circ \delta$ is locally nilpotent, so that the conditions of the BPL are satisfied. In this way, we deduce a reduction from the complex $(T(C), d)$ (the initial one, where now all the arrows are considered) over a finite type complex, obtaining the looked-for effective homology of $(T(C), d)$.

A natural generalization of double complexes are multicomplexes, where, in addition to horizontal and vertical arrows, morphisms $d^r_{p,q} : C_{p,q} \to C_{p-r,q+r-1}$ are considered for each $r \in \mathbb{N}$. The totalization is obtained in the same way, with differential map defined as the sum of all the components, $d = \sum d^r$.

Again, if for each column C^i there exists a reduction $C^i \Rightarrow HC^i$ where the HC^i are effective complexes, then using the BPL as before it is possible to construct a new effective complex HC and a reduction $T(C) \Rightarrow HC$ that provides us the effective homology of the multicomplex C.

5 Effective homology of a fibration

Given a fibration

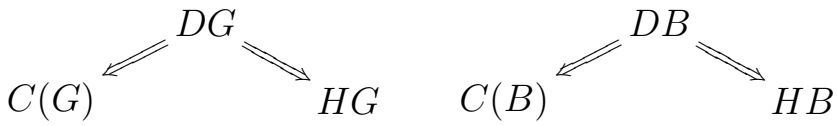
$$G \hookrightarrow E \to B$$

with fiber G and base B, where G and B are objects with effective homology, in this section we explain how to determine the effective homology of the total space $E = B \times_\tau G$.

From now on, all the chain complexes canonically associated with simplicial sets are *normalized*, that is, only the non-degenerate n-simplices of

X are considered to be generators of $C_n(X)$.

Let us suppose there exist two homotopy equivalences



where HG and HB are effective complexes. How can we obtain a new equivalence between $C(B \times_\tau G)$ and an effective chain complex?

The starting point is the Eilenberg-Zilber reduction $C(B \times G) \Rightarrow C(B) \otimes C(G)$ (see [8]), that relates the cartesian product of two simplicial sets with the tensorial product of the associated chain complexes. In our case, we must also take account of the torsion τ, that does not change the underlying graded group, only the differential is modified (by a perturbation $\delta(b, g) = (\partial_0 b, \partial_0 g \cdot \tau(b)) - (\partial_0 b, \partial_0 g)$). We could try to apply the BPL; for this, the nilpotency condition must be satisfied.

In both chain complexes $C(B \times G)$ and $C(B) \otimes C(G)$ it is possible to define the following filtrations. First of all, $C(B \times G)$ is filtered through the degeneracy degree with respect to the base space: a generator $(x_n, y_n) \in C_n(B \times G)$ has a filtration degree less or equal to q if $\exists \bar{x}_q \in B_q$ such that $x_n = \eta_{i_{n-q}} \cdots \eta_{i_1} \bar{x}_q$. On the other hand, the filtration on $C(B) \otimes C(G)$ is defined through the dimension of the base component,

$$F_p(C(B) \otimes C(G)) = \bigoplus_{m \leq p} C(B)_m \otimes C(F)$$

It is not difficult to see that the three operators involved in the Eilenberg-Zilber reduction (that is, the three components f, g and h) are compatible with these filtrations. On the contrary, the perturbation δ decreases the filtration degree on $C(B \times G)$ by one unit, so that the composition $h \circ \delta$ is locally nilpotent and the hypothesis of the BPL are satisfied. In this way, a new reduction $C(B \times_\tau G) \Rightarrow C(B) \otimes_t C(G)$ is obtained, where the symbol $\otimes_t$ represents a twisted (perturbed) tensor product, induced by τ.

On the other hand, with the effective homologies of B and G, it is easy to build a new equivalence

$$\begin{array}{ccc} & DB \otimes DG & \\ \swarrow & & \searrow \\ C(B) \otimes C(G) & & HB \otimes HG \end{array}$$

Let us consider now the necessary perturbation $\bar{\delta}$ of $C(B) \otimes C(G)$ to obtain the twisted cartesian product $C(B) \otimes_t C(G)$ (this perturbation $\bar{\delta}$ has been obtained when applying the BPL to the Eilenberg-Zilber reduction). If the base space B is 1-reduced then it can be seen that $\bar{\delta}$ decreases the filtration degree at least by 2. This perturbation can be transferred to the top chain complex $DB \otimes DG$, obtaining a twisted product $DB \otimes_t DG$, modified by a perturbation on $DB \otimes DG$ with the same property about the filtration degree. Finally, the homotopy operator of the reduction $DB \otimes DG \Rightarrow HB \otimes HG$ increases the filtration degree at most by one, and therefore the Basic Perturbation Lemma can be applied to the right reduction and an equivalence is obtained as follows.

$$\begin{array}{ccc} & DB \otimes_t DG & \\ \swarrow & & \searrow \\ C(B) \otimes_t C(G) & & HB \otimes_t HG \end{array}$$

The chain complex $HB \otimes_t HG$ is an effective complex, so that the composition of the two equivalences

$$\begin{array}{ccccc} & \overset{Id}{\swarrow} C(B \times_\tau G) & \searrow & \swarrow DB \otimes_t DG & \searrow \\ C(B \times_\tau G) & & C(B) \otimes_t C(G) & & HB \otimes_t HG \end{array}$$

is the effective homology of $B \times_\tau G$.

We consider now our particular example X_5 introduced in Section 2, which is the total space of the fibration $F_3 \hookrightarrow X_5 \to X_4$, where $F_3 = K(\mathbb{Z}_2, 3)$. The object X_4 is again the total space of a fibration $F_2 \hookrightarrow X_4 \to S^3$ with $F_2 = K(\mathbb{Z}, 2)$.

To compute the effective homology of X_5, we need the effective homologies of the fiber and base spaces, F_3 and X_4. First, the simplicial group $F_3 = K(\mathbb{Z}_2, 3)$ is of finite type and therefore its effective homology is trivial. To compute the effective homology of X_4, which is the total space of another fibration, we must apply again the same method, so that the effective homologies of F_2 and S^3 are necessary. On one hand, the simplicial set S^3 is already of finite type and therefore its effective homology presents no problem. And finally, the difficult part is the computation of the effective homology of $F_2 = K(\mathbb{Z}, 2)$.

The computation of the effective homology of Eilenberg-MacLane spaces $K(\pi, n)$'s is in general a difficult problem (especially in what regards to the algorithmic complexity), but in the case $\pi = \mathbb{Z}$ it can be solved as follows. First, the space $K(\mathbb{Z}, 0)$ is considered to be the simplicial group with all the components equal to $\mathbb{Z}$ and with the identity map as faces and degeneracies. Then, the simplicial group $K(\mathbb{Z}, n)$ can be recursively defined as the classifying space of $K(\mathbb{Z}, n-1)$, that is, $K(\mathbb{Z}, n) = \mathcal{W}(K(\mathbb{Z}, n-1))$. In our case, $F_2 = K(\mathbb{Z}, 2) = \mathcal{W}(K(\mathbb{Z}, 1)) = \mathcal{W}(\mathcal{W}(K(\mathbb{Z}, 0)))$.

It is well known $K(\mathbb{Z}, 1)$ has the homotopy type of the circle S^1. Moreover, although we are not going to give the details about this construction, it can be seen that there exist a mechanism (similar to that of the fibration) for the computation of the effective homology of the classifying space of a simplicial group with effective homology. Therefore we can apply this method to compute the effective homology of F_2.

In this way, we have the necessary ingredients to obtain the effective homology of X_4, and recursively, that of X_5. With this effective homology we can easily compute, as we have seen in Section 2, the homology groups of this space.

As we have showed in this section, the effective homology method applied to a fibration $G \hookrightarrow E \to B$ (with base space B 1-reduced) gives in particular to its user an algorithm to compute the homology groups of the total space E, replacing in this way the Serre spectral sequence technique. But anyway, even if the homology groups of E are known, this spectral

sequence has a great interest by itself and therefore it also would be interesting to compute the whole set of its elements. As we see in the next section, the effective homology method can also be useful for this task.

6 Computing spectral sequences

The next theorem combines both spectral sequence and effective homology concepts and is the main result that allows us to use the effective homology method to compute spectral sequences of filtered complexes. The proof is straightforward and is not included here.

Theorem 3. *Let C be a filtered chain complex with effective homology (HC, ε), with $\varepsilon = (D, \rho, \rho')$, $\rho = (f, g, h)$, and $\rho' = (f', g', h')$. Let us suppose that filtrations are also defined on the chain complexes HC and D. If the maps f, f', g, and g' are morphisms of filtered complexes (i.e., they are compatible with the filtrations) and both homotopies h and h' have order $\leq t$ (i.e. they increase the filtration degree at most by t), then the spectral sequences of the complexes C and HC are isomorphic for $r > t$:*

$$E(C)^r_{p,q} \cong E(HC)^r_{p,q} \quad \forall r > t$$

This theorem provides us an algorithm to compute spectral sequences of (complicated) filtered complexes with effective homology. If a filtered complex is effective, then its spectral sequence (that of Theorem 1) can be computed by means of elementary operations with matrices (in a similar way to the computation of homology groups); otherwise, the effective homology is needed to compute the $E^r_{p,q}$ by means of an analogous spectral sequence deduced of an appropriate filtration on the associated effective complex, which is isomorphic to the spectral sequence of the initial complex after some level r. In particular, we can apply this result to compute the Serre spectral sequence, as we explain in the following paragraph.

The Serre spectral sequence associated with a fibration $G \hookrightarrow E \to B$ can be defined as the spectral sequence of the total space E, with the natural filtration of cartesian products. The space E is not effective in most

situations, so in general it is not possible to compute directly its spectral sequence. However, as we have seen in Section 5, provided that the spaces B and G are spaces with effective homology (and B is 1-reduced) we can also build the effective homology of the total space E, which allows us to determine the homology groups of E. Moreover, the natural filtration of tensor products can be defined on the effective complex and we have already seen that all the homotopies involved in the equivalence have order ≤ 1. Applying Theorem 3, the spectral sequence of E and that of the effective complex are isomorphic after level $r = 2$, and in this way we can compute the Serre spectral sequence associated with the fibration by means of the spectral sequence of an effective complex (which can easily be computed).

Using these results, we have developed a set of programs enhancing the Kenzo system that allow computations of spectral sequences of filtered complexes when the effective homology of this complex is available. The programs determine not only the groups, but also the differential maps d^r in the spectral sequence, as well as the stage r on which the convergence has been reached and the filtration of the homology groups by the spectral sequence. As a particular case, the computation of Serre spectral sequences where the base and the fiber spaces are objects with effective homology is possible.

We consider again the example X_5 introduced in Section 2, total space of the fibration $F_3 \hookrightarrow X_5 \rightarrow X_4$. As we have seen, the effective homology of X_5 can be determined by means of the effective homology of F_3 and X_4, and in fact the Kenzo program implements this computation and uses it to determine the homology groups $H_*(X_5)$. This effective homology is also necessary to compute the corresponding Serre spectral sequence with our new programs.

First of all, the space X_5 and its effective equivalent object must be filtered with the natural filtrations of cartesian products and tensor products respectively, as follows.

```
>(change-chcm-to-fltrchcm x5 fbrt-flin '(fbrt-flin))
[K298 Filtered-Simplicial-Set]
>(change-chcm-to-fltrchcm (rbcc (efhm x5)) tnpr-flin
'(tnpr-flin))
[K594 Filtered-Chain-Complex]
```

Then, the whole set of the Serre spectral sequence can easily be obtained. We show here the computation of some groups.

```
>(print-spct-sqn-cmpns x5 2 6 0)
Spectral sequence E^2_{6,0}
Component Z/3Z
>(print-spct-sqn-cmpns x5 4 8 0)
Spectral sequence E^4_{8,0}
Component Z/4Z
>(print-spct-sqn-cmpns x5 4 4 3)
Spectral sequence E^4_{4,3}
Component Z/2Z
```

The differential maps $d^r_{p,q}$ can also be determined for every r. For example, $d^4_{8,0} : E^4_{8,0} = \mathbb{Z}_4 \to E^4_{4,3} = \mathbb{Z}_2$ sends the generator of $E^4_{8,0} = \mathbb{Z}_4$ to the generator of $E^4_{4,3} = \mathbb{Z}_2$.

```
>(spct-sqn-dffr x5 4 8 0 '(1))
(1)
```

The convergence level of the spectral sequence for $p + q = 8$ is $r = 9$.

```
>(spct-sqn-cnvg-level x5 8)
9
```

And finally, we can determine the filtration of the homology groups by the spectral sequence. For instance, for $H_6(X_5) \equiv H_6 = \mathbb{Z}_6$, we obtain $F_0H_6 = F_1H_6 = \ldots = F_5H_6 = \mathbb{Z}_2, F_6H_6 = H_6 = \mathbb{Z}_6$.

```
>(homology-fltr x5 6 0)
Filtration F_0 H_6
Component Z/2Z
>(homology-fltr x5 6 5)
Filtration F_5 H_6
Component Z/2Z
>(homology-fltr x5 6 6)
Filtration F_6 H_6
Component Z/6Z
```

7 Conclusions and further work

In this paper, a program computing spectral sequences of filtered complexes has been presented. It is based on the effective homology method and in particular allows its user to compute the Serre spectral sequence associated with a fibration where the base and fiber space are objects with effective homology. For a better understanding of the fundamental ideas on which this program is based, we have considered a particular example of application.

At this point, new goals appear. First of all, one of our next aims is the application of our program to the computation of the Koszul homology and the Spencer cohomology ([2]), enriching in this way the program developed by Sergeraert. On the other hand, we are planning to extend our programs to the case of spectral sequences which are not necessarily associated with a filtered complex. Concretely, we focus on the Bousfield-Kan spectral sequence [9], used to compute homotopy groups of simplicial sets.

References

[1] Guillemin, V.W. and Sternberg, S.: An algebraic model of transitive differential geometry, *Bull. Amer. Math. Soc.*, 70, 16-47, 1964.

[2] Lambe, L. A. and Seiler, W. M.: Differential Equations, Spencer Cohomology, and Computing Resolutions, *Georgian Math. J.*, 9, 723-772, 2002.

[3] Serre, J.P.: Homologie singulière des espaces fibrés, *Ann. of Math*, 54, 425-505, 1951.

[4] Dousson, X., Rubio, J., Sergeraert, F. and Siret, Y.: *The Kenzo program*, Institut Fourier, Grenoble, 1999.
`http://www-fourier.ujf-grenoble.fr/~sergerar/Kenzo/`
[5] Mac Lane, S.: *Homology*, Springer, 1994.
[6] Rubio, J. and Sergeraert, F.: Constructive Algebraic Topology, *Bulletin des Sciences Mathématiques*, 126, 389-412, 2002.
[7] Brown, R.: The twisted Eilenberg–Zilber theorem, *Celebrazioni Arch. Secolo XX Simp. Top.*, 34-37, 1967.
[8] Eilenberg, S. and Zilber, J.A.: On products of complexes, *Am. J. Math.*, 75, 200-204, 1953.
[9] Bousfield, A. K. and Kan, D. M.: *Homotopy limits, completions and localizations, Lecture Notes in Mathematics*, 304, Springer-Verlag, 1972.

Mayer-Vietoris Trees of Monomial Ideals[1]

Eduardo Sáenz de Cabezón

Universidad de La Rioja, Dpto. Matemáticas y Computación,
eduardo.saenz-de-cabezon@dmc.unirioja.es

Abstract

Homological Algebra provides tools and concepts that are at the merging of important problems in Formal Theory of Differential Equations and Commutative Algebra. As was pointed in [1], using a result by Serre, ̀the knowledge of the vanishing of a certain Tor gives some computational insight in the completion of general systems of partial differential equations and is also useful for the concrete determination of formal power series of solutions. A good way to compute Tor is to use Koszul homology of ideals of the polynomial ring, a central issue in Commutative Algebra strongly related to the most important homological invariants of these rings, such as minimal resolutions, Betti numbers, Hilbert function, Castelnuovo-Mumford regularity, etc. Using tools coming from Algebraic Topology, namely the Mayer-Vietoris sequences, we compute the Koszul homology of ideals of the polynomial ring. We focus on the computations for monomial ideals and the results achieved can be used to compute the Koszul homology of general polynomial ideals making use of Gröbner basis techniques and Homological Perturbation.

Keywords: Koszul homology, monomial ideals, minimal free resolutions, Mayer-Vietoris trees.

1 Introduction

Given a differential system, we can associate to it a symbol comodule N which is dual to a module M over the polynomial algebra. A criterion for an involutive symbol in terms of this module M is provided by the Cartan Test, see the details of all these considerations in [1]. A result by Serre

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

relates the Cartan criterion to the vanishing of certain Tor modules of the module M dual to the symbol comodule N or equivalently, to the vanishing of certain $Cotor$ modules of the symbol comodule itself.

For computing $Cotor$ one can use Spencer Cohomology and dually, for computing Tor one uses Koszul homology. In the sequel we will focus on the Koszul homology side. Being this homology bigraded, the interpretation of involution in terms of it makes extensive use of this bigrading: the degree of involution is the symmetric degree at which the Koszul homology vanishes i.e. the Castelnuovo-Mumford regularity. Some references to Formal Theory and the use of Spencer cohomology are [2],[3] or [4].

Thus, our goals include the *complete* computation of the Koszul homology of ideals of the polynomial ring, where *complete* means we will compute not only the dimensions of the modules (Betti numbers) or even the graded or multigraded version of these dimensions (graded and multigraded Betti numbers), but also an explicit set of generators for each of the modules. We will also be able to read all the information about the ideal that comes from the computation of Koszul homology: minimal resolution, Hilbert function and Castelnuovo-Mumford regularity. The main tool we will use in our strategy will be a short exact sequence of complexes, which is an analogue to the Mayer-Vietoris short exact sequence from Algebraic Topology. Applying recursively these sequences will lead us to the actual computations.

In the first section of this paper, the terminolgy and main definitions are given. In the second section, the Mayer-Vietoris tree of a monomial ideal is introduced and used to compute homological invariants of the ideal from its tree. The third section describes an algorithm for computing Mayer-Vietoris trees, and read the homological information in them. Finally, the last section is devoted to examples and conclusions.

2 Basic Terminology and Definitions

Let $\mathbf{k}$ be a field and $R = \mathbf{k}[x_1, \ldots, x_n]$ the ring of polynomials in n variables over $\mathbf{k}$.

2.1 The Koszul Complex and Koszul Homology

Let $\mathcal{V}$ be a n-dimensional $\mathbf{k}$-vector space. Let $S\mathcal{V}$ and $\wedge\mathcal{V}$ be the Symmetric and Exterior algebras of $\mathcal{V}$ respectively. We consider the basis of $\mathcal{V}$ given by $\{x_1, \ldots, x_n\}$; then we can identify $S\mathcal{V}$ and R and consider the following complex

$$\mathbb{K} : 0 \to R \otimes \wedge^n \mathcal{V} \xrightarrow{\partial} R \otimes \wedge^{n-1} \mathcal{V} \xrightarrow{\partial} \cdots R \otimes \wedge^1 \mathcal{V} \xrightarrow{\partial} R \otimes \wedge^0 \mathcal{V} \to \mathbf{k} \to 0$$

Any element of $R \otimes \wedge^i \mathcal{V}$ can be written in two different ways: First, as a $\mathbf{k}$-linear combination of elements of the form $x_1^{\mu_1} \ldots x_n^{\mu_n} \otimes x_1^{j_1} \wedge \cdots \wedge x_n^{j_n}$, where the μ_k are integers and the j_k are either 0 or 1, and exactly i of them are 1's. In this case, the differentials ∂ are given by the rule

$$\begin{aligned}\partial(x_1^{\mu_1} \ldots x_n^{\mu_n} \otimes x_1^{j_1} \wedge \cdots \wedge x_n^{j_n}) = &\sum_{j_k=1} (-1)^{\sigma(k)+1} x_k \cdot x_1^{\mu_1} \ldots x_n^{\mu_n} \\ &\otimes x_1^{j_1} \wedge \cdots \wedge x_k^{j_k - 1} \wedge \cdots \wedge x_n^{j_n}\end{aligned}$$

where $\sigma(k)$ is the position of k in the set $\{k | j_k = 1\}$.

Alternatively, elements of $R \otimes \wedge^i \mathcal{V}$ can be expressed as $\mathbf{k}$-linear combinations of elements of the form $x_1^{\mu_1} \ldots x_n^{\mu_n} \otimes x_{j_1} \wedge \cdots \wedge x_{j_i}$ with $j_1 < \cdots j_i \in \binom{n}{i}$;then, the differentials have the form

$$\begin{aligned}\partial(x_1^{\mu_1} \ldots x_n^{\mu_n} \otimes x_{j_1} \wedge \cdots \wedge x_{j_n}) = &\sum_{k=1}^{i} (-1)^{k+1} x_{j_k} \cdot x_1^{\mu_1} \ldots x_n^{\mu_n} \\ &\otimes x_{j_1} \wedge \cdots \wedge \widehat{x_{j_k}} \wedge \cdots \wedge x_{j_n}\end{aligned} \tag{1}$$

This differential verifies $\partial^2 = 0$ and makes $\mathbb{K}$ a complex, which is called the *Koszul complex*. This complex is a minimal free resolution of $\mathbf{k} = R/\mathfrak{m}$, where $\mathfrak{m} = \langle x_1, \ldots, x_n \rangle$, the maximal ideal in R.

Given a graded module $\mathcal{M}$, its *Koszul complex* $(\mathbb{K}(\mathcal{M}), \partial)$ is the tensor product complex $\mathcal{M} \otimes_R \mathbb{K}$:

$$\mathbb{K}(\mathcal{M}) : 0 \to \mathcal{M} \otimes \wedge^n \mathcal{V} \xrightarrow{\partial} \mathcal{M} \otimes \wedge^{n-1} \mathcal{V} \xrightarrow{\partial} \cdots \mathcal{M} \otimes \wedge^1 \mathcal{V} \xrightarrow{\partial} \mathcal{M} \otimes \wedge^0 \mathcal{V} \to \mathbf{k} \to 0$$

This complex is no longer acyclic, and we define the *Koszul homology* of $\mathcal{M}$ as the homology of $\mathbb{K}(\mathcal{M})$.

Grading, bigrading and multigrading of the Koszul Complex

Consider an element of $R \otimes \wedge\mathcal{V}$ of the form $x^\mu \otimes x^J$ where $x^\mu = x_1^{\mu_1} \dots x_n^{\mu_n}$ and $J = x_1^{j_1} \wedge \cdots \wedge x_n^{j_n}$ as before. We say that the *total degree* of $x^\mu \otimes x^J$ is $\mu_1 + \cdots + \mu_n + j_1 + \cdots + j_n$ and that the *total multidegree* of $x^\mu \otimes x^J$ is $(\mu_1 + j_1, \cdots, \mu_n + j_n)$. Equivalently, if J is given in the form $J = j_1 < \cdots < j_i$ then the total degree of $x^\mu \otimes x^J$ is $\mu_1 + \cdots + \mu_n + i$ and the total multidegree is $(\mu_1 + [1 \in J], \dots, \mu_n + [n \in J])$ where $[i \in J]$ equals 1 if i is in J and 0 otherwise.

It is clear that for these elements, the Koszul differential preserves both the total degree and total multidegree. Thus, we can consider the following (multi)gradings in $\mathbb{K}$ and $\mathbb{K}(\mathcal{M})$:

- With respect to the total degree,we have

$$\mathbb{K} = \bigoplus_{d\in\mathbb{N}} \mathbb{K}_d \qquad \textit{and} \qquad \mathbb{K}(\mathcal{M}) = \bigoplus_{d\in\mathbb{N}} \mathbb{K}_d(\mathcal{M})$$

  where

$$\mathbb{K}_d : 0 \to R_{d-n} \otimes \wedge^n\mathcal{V} \xrightarrow{\partial} R_{d-n+1} \otimes \wedge^{n-1}\mathcal{V} \xrightarrow{\partial} \cdots R_{d-1} \otimes \wedge^1\mathcal{V}$$
$$\xrightarrow{\partial} R_d \otimes \wedge 0\mathcal{V} \to \mathbf{k} \to 0$$

  and similarly for $\mathbb{K}_d(\mathcal{M})$. Here, R_l denotes the polynomials of degree l and in $\mathbb{K}(\mathcal{M})$ we have that $\mathcal{M}_l$ is the degree l component of $\mathcal{M}$. Because of this grading in $\mathbb{K}$ and $\mathbb{K}(\mathcal{M})$, the homologies of them are also graded:

$$H_*(\mathbb{K}) = \bigoplus_{d\in\mathbb{N}} H_*(\mathbb{K}_d) \qquad \textit{and} \qquad H_*(\mathbb{K}(\mathcal{M})) = \bigoplus_{d\in\mathbb{N}} H_*(\mathbb{K}_d(\mathcal{M}))$$

  And then, for each homological dimension p we have $H_p(\mathbb{K}) = \bigoplus_{d\in\mathbb{N}} H_p(\mathbb{K}_d) = \bigoplus_{q+p=d} H_{q,p}(\mathbb{K})$; so we have a bigrading and we denote by $H_{q,p}(\mathbb{K})$ and $H_{q,p}(\mathbb{K}(\mathcal{M}))$ the respective homology modules at $R_q \otimes \wedge^p\mathcal{V}$ and $M_q \otimes \wedge^p\mathcal{V}$. We say that q is the *symmetric degree* of $H_{q,p}(\mathbb{K})$ or $H_{q,p}(\mathbb{K}(\mathcal{M}))$ and p is its *exterior degree*.

- With respect to the total multidegree, we have

$$\mathbb{K} = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} \mathbb{K}_{\mathbf{a}} \qquad and \qquad \mathbb{K}(\mathcal{M}) = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} \mathbb{K}_{\mathbf{a}}(\mathcal{M})$$

where for every $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{N}^n$ with $a_{j_1}, \ldots, a_{j_l} \neq 0$

$$\mathbb{K}_{\mathbf{a}} : 0 \to R_{\mathbf{a}-(0..1_{j_1}..1_{j_l}..0)} \otimes \wedge^l \mathcal{V} \xrightarrow{\partial} \bigoplus_{k=1,\ldots,l} R_{\mathbf{a}-(0..1_{j_1}..0_{j_k}..1_{j_l}..0)} \otimes \wedge^{l-1} \mathcal{V}$$

$$\cdots \xrightarrow{\partial} \bigoplus_{k=1,\ldots,l} R_{\mathbf{a}-(0,\ldots,1_{j_k},\ldots,0)} \otimes \wedge^1 \mathcal{V} \xrightarrow{\partial} R_{\mathbf{a}} \otimes \wedge^0 \mathcal{V} \longrightarrow \mathbf{k} \to 0$$

and similarly for $\mathbb{K}_{\mathbf{a}}(\mathcal{M})$. Here, R_μ denotes the set of polynomials of multidegree μ. In this case, we have that the homologies of $\mathbb{K}$ and $\mathbb{K}(\mathcal{M})$ are also multigraded:

$$H_*(\mathbb{K}) = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} H_*(\mathbb{K}_{\mathbf{a}}) = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} H_{\mathbf{a}}(\mathbb{K})$$

$$H_p(\mathbb{K}) = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} H_p(\mathbb{K}_{\mathbf{a}}) = H_{\mathbf{a},p}(\mathbb{K})$$

$$H_*(\mathbb{K}(\mathcal{M}) = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} H_*(\mathbb{K}_{\mathbf{a}}(\mathcal{M})) = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} H_{\mathbf{a}}(\mathbb{K}(\mathcal{M}))$$

$$H_p(\mathbb{K}(\mathcal{M})) = \bigoplus_{\mathbf{a}\in\mathbb{N}^n} H_p(\mathbb{K}_{\mathbf{a}}(\mathcal{M})) = H_{\mathbf{a},p}(\mathbb{K}(\mathcal{M}))$$

2.2 Koszul Complex and Tor

From the definitions, it is clear that we can identify the Koszul homology modules with $Tor^R_\bullet(\mathcal{M}, \mathbf{k})$: We have a resolution of $\mathbf{k}$ (the Koszul complex) to which we have applied the functor $\mathcal{M}\otimes -$. The homology of the resulting complex is by definition $Tor^R_\bullet(\mathcal{M}, \mathbf{k})$.

Another way of computing the Koszul homology of $\mathcal{M}$ would start with a resolution $\mathbb{P} : \cdots \to P_i \xrightarrow{\delta_i} P_{i-1} \to \cdots \to P_0$ of $\mathcal{M}$ and then compute the

homology of $\mathbb{P} \otimes \mathbf{k}$. This homology is independent of the chosen resolution of $\mathcal{M}$ or of $\mathbf{k}$. If $\mathbb{P}$ is minimal,the differentials are given by matrices with polynomial coefficients, none of which is a nonzero constant. Thus, tensoring with $\mathbf{k}$ yields the zero differential everywhere and then the number of generators of each P_i equals $Tor_i^R(\mathcal{M}, \mathbf{k})$ and the dimension of the i-th Koszul homology of $\mathcal{M}$. If $\mathbb{P}$ is not minimal one could either minimalize it with some standard procedure (see for example [5]) or compute the homology of the resulting resolution $\mathbb{P} \otimes \mathbf{k}$. The Tor modules inherit the gradings, bigradings and multigradings we have seen before just considering the corresponding gradings in the homology modules used to define Tor.

2.3 Monomial Ideals and Multigrading

Definition 2.1. A monomial in R is a product $x^{\mathbf{a}} = x_1^{a_1} \cdots x_n^{a_n}$ with $a_i \geq 0 \, \forall i$. We say that $\mathbf{a} \in \mathbb{N}_0^n$ is the multidegree of $x^{\mathbf{a}}$. An ideal $I \subset R$ is called a monomial ideal if it is generated by monomials.

A monomial ideal is uniquely determined by its monomials, i.e. two monomial ideals are the same if and only if they contain the same monomials. An important result for monomial ideals is the so called Dickson's Lemma, which states that all monomial ideals of R are finitely generated, see for example [6] for a proof. Moreover, this minimal set of monomial generators is unique.

2.4 Mayer-Vietoris sequence of Koszul Complexes

The Mayer-Vietoris sequence is a standard tool in Algebraic Topology (see for example [7]) used to explore the relations between the homology groups of two spaces A and B, their intersection $A \cap B$ and their union $A \cup B$. Making some assumptions, we can define a short exact sequence

$$0 \to A \cap B \to A \oplus B \to A \cup B \to 0$$

which induces a long exact sequence in homology, which is called the *Mayer-Vietoris sequence*:

$$\cdots \to H_n(A \cap B) \to H_n(A) \oplus H_n(B) \to H_n(A \cup B) \xrightarrow{\Delta} H_{n-1}(A \cap B) \to \cdots$$

We can construct an analogue of it in the following way: consider two monomial ideals A and B, minimally generated by $\{a_1, \ldots, a_{r_A}\}$ and $\{b_1, \ldots, b_{r_B}\}$ respectively. Then we have that the ideal $A \cap B$ is generated by the set $\{lcm(a, b)\}$ where a is a generator of A and b is a generator of B; and in the place of $A \cup B$ we install $A + B$, the ideal generated by all the generators of A and B.

Given a monomial ideal I minimaly generated by $\{m_1, \ldots, m_r\}$ we define the (recursive) Mayer-Vietoris exact sequence of I in the following way:

For each $1 \leq s \leq r$ denote $I_s := \langle m_1, \ldots m_s \rangle$, $\tilde{I}_s := I_{s-1} \cap \langle m_s \rangle = \langle m_{1,s}, \ldots, m_{s-1,s} \rangle$, where $m_{i,j}$ denotes $lcm(m_i, m_j)$. Then, for each s we have the following exact sequence of ideals:

$$0 \to \tilde{I}_s \to I_{s-1} \oplus \langle m_s \rangle \to I_s \to 0$$

and the following short exact sequence of Koszul complexes:

$$0 \longrightarrow \mathbb{K}(\tilde{I}_s) \xrightarrow{\mathfrak{i}} \mathbb{K}(I_{s-1}) \oplus \mathbb{K}(\langle m_s \rangle) \xrightarrow{\mathfrak{j}} \mathbb{K}(I_s) \longrightarrow 0,$$

the maps given by

$$\mathfrak{i}(\gamma) = (\gamma, -\gamma), \qquad \mathfrak{j}(\eta, \eta') = \eta + \eta'$$

for $\gamma \in \mathbb{K}(\tilde{I}_s)$, $\eta \in \mathbb{K}(I_{s-1})$ and $\eta' \in \mathbb{K}(\langle m_s \rangle)$.

These sequences induce a long exact sequence in Koszul homology for each s, the set of all of them is what we call the (recursive) *Mayer-Vietoris Sequence* of I.

Since the differential ∂ of $\mathbb{K}(I)$ preserves multidegree, and denoting by $\mathbb{K}_{\mathbf{a}}(I)$ the multidegree $\mathbf{a}$ piece of $\mathbb{K}(I)$, we have a multigraded version of the sequence:

$$0 \longrightarrow \mathbb{K}_{\mathbf{a}}(\tilde{I}_s) \longrightarrow \mathbb{K}_{\mathbf{a}}(I_{s-1}) \oplus \mathbb{K}_{\mathbf{a}}(\langle m_s \rangle) \longrightarrow \mathbb{K}(I_s)_{\mathbf{a}} \longrightarrow 0$$

and the corresponding multigraded version of the Mayer-Vietoris Sequence of I:

$$\cdots \longrightarrow H_{i+1}(\mathbb{K}_{\mathbf{a}}(I_s)) \xrightarrow{\Delta} H_i(\mathbb{K}_{\mathbf{a}}(\tilde{I}_s) \longrightarrow$$

$$H_i(\mathbb{K}_{\mathbf{a}}(I_{s-1}) \oplus \mathbb{K}_{\mathbf{a}}(\langle m_s \rangle)) \longrightarrow H_i(\mathbb{K}_{\mathbf{a}}(I_s)) \xrightarrow{\Delta} \cdots \tag{2}$$

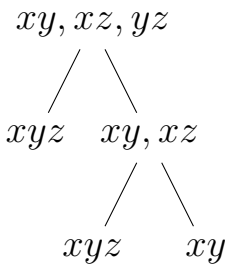
3 The Mayer-Vietoris Tree of I.

Using recursively these exact sequences for every $\mathbf{a} \in \mathbb{N}^n$ we could compute the Koszul homology of $I = \langle m_1, \ldots, m_r \rangle$. When making use of them to compute $H_*(\mathbb{K}_{\mathbf{a}}(I))$, we need $H_*(\mathbb{K}_{\mathbf{a}}(\tilde{I}_r))$ and $H_*(\mathbb{K}_{\mathbf{a}}(I_{r-1}))$; and for each of these two computations, one needs the corresponding smaller ideals. Note that the size of the involved ideals decreases until they are generated by only one monomial, in which case the Koszul homology is trivial.

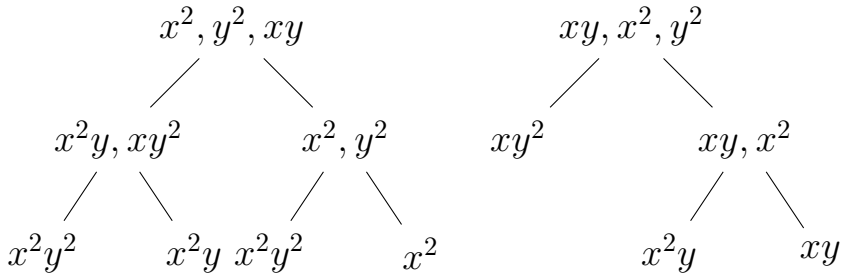
The involved ideals can be displayed as a tree, the root of which is I and every node J has as *children* $\tilde{J}$ on the left and J' on the right (if J is generated by s monomials, $\tilde{J}$ denotes $\tilde{J}_s$ and J' denotes J_{s-1}). This is what we call the **Mayer-Vietoris Tree** of a monomial ideal I, and we will denote it $MVT(I)$. Note that the $MVT(I)$ depends on the ordering in which the monomials are given, thus, for a given ideal I we have several (and eventually different) trees depending on the order of its generators.

Remark 3.1. Observe that if we have the monomials in our ideal J sorted in a way such that m_s has the biggest exponent in some of the variables, then the generators in $\tilde{J}$ will have the same exponent in this variable. Keeping the nodes ordered in such a way gives us small trees, as the left branch will have a length of at most the number of variables plus one.

Example 3.2. The monomial ideal $I = \langle xy, xz, yz \rangle$, has the following Mayer-Vietoris tree, if we change the order of the generating monomials, we obtain isomorphic trees (i.e. isomorphic as graphs and with isomorphic ideals in the corresponding nodes)



On the other side, for $I = \langle x^2, y^2, xy \rangle$ we obtain two different trees according to the ordering of the monomials:



Since by definition $MVT(I)$ is a *complete tree* i.e. every *father* has exactly two *children*, we can assign position indices to every node, in the following way: I has position 1 and if J has position p then $\tilde{J}$ has position $2p$ and J' has position $2p + 1$. We will denote this $MVT_1(I) = I$, $MVT_p(I) = J$, $MVT_{2p}(I) = \tilde{J}$, $MVT_{2p+1}(I) = J'$. These indices will be very useful for efficiently reading the information hidden in $MVT(I)$.

Remark 3.3. Two famous objects that express the combinatorial structure of monomial ideals based on the least common multiples of their generators are the Taylor resolution [8] and the lcm-lattice [9]. For a monomial ideal $I = \langle m_1, \ldots, m_r \rangle$ the first one is a complex supported in elements of the form m_U for all subset U of $\{1, \ldots, r\}$ i.e. $2^r - 1$ elements, and the lcm-latice contains all the different least common multiples (lcm) of generators of the ideal, in general, the lcm-lattice is smaller than the Taylor complex. The Mayer-Vietoris tree is in general smaller, because it takes into account divisibility between lcm's of generators. In the last example, the Taylor resolution is suported on 7 elements, the lcm-lattice contains 6 different

monomials, and the (minimal) Mayer-Vietoris tree has only 5 different monomials.

3.1 $MVT(I)$ and Koszul homology computations

Let J be a node in $MVT(I)$ and denote by $md(H_i\mathbb{K}((J)))$ (or $md_i(J)$ for short) the set of multidegrees in which $H_i(\mathbb{K}(J))$ is different from zero. We have that all the $md_i(I)$ are present as exponents of generators in the nodes of the Mayer-Vietoris tree. In the following lemmata J is a node of $MVT(I)$ minimally generated by $\{m_1, \ldots, m_s\}$.

Lemma 3.4. $H_0(\mathbb{K}(J)) = H_0(\mathbb{K}(J')) \oplus \langle m_s \rangle$

Proof: In homological dimension 0 we have the following exact sequence in homology

$$\cdots \longrightarrow H_0(\mathbb{K}_{\mathbf{a}}(\tilde{J})) \longrightarrow H_0(\mathbb{K}_{\mathbf{a}}(J') \oplus \mathbb{K}_{\mathbf{a}}(\langle m_s \rangle)) \longrightarrow H_0(\mathbb{K}_{\mathbf{a}}(J)) \longrightarrow 0$$

but if x^μ is such that $\mu \in md_0(J' \oplus \langle m_s \rangle)$ then $x^\mu = m_j$ for some j. Now, if $\mu \in md_0(\tilde{J})$ then $x^\mu = m_{is}$ for some i. Then $m_j = m_{is}$, and we have that $m_j | m_i$, $m_j | m_s$ which is a contradiction. Then the exact sequence in homology is of the form

$$0 \longrightarrow H_0(\mathbb{K}_{\mathbf{a}}(J') \oplus \mathbb{K}_{\mathbf{a}}(\langle m_s \rangle)) \longrightarrow H_0(\mathbb{K}_{\mathbf{a}}(J)) \longrightarrow 0$$

and hence the result. $\square$

Lemma 3.5. $\forall i > 0$, $md_i(J) \subseteq md_{i-1}(\tilde{J}) \bigcup md_i(J')$

Proof: Let $\mathbf{a} \in \mathbb{N}_0^n$ such that $\mathbf{a} \notin md_{i-1}(\tilde{J}) \bigcup md_i(J')$ then the exact sequence in homology

$$\cdot\cdot \longrightarrow H_i(\mathbb{K}_{\mathbf{a}}(J')) \longrightarrow H_i(\mathbb{K}_{\mathbf{a}}(J)) \xrightarrow{\Delta} H_{i-1}(\mathbb{K}_{\mathbf{a}}(\tilde{J})) \longrightarrow H_{i-1}(\mathbb{K}_{\mathbf{a}}(J')) \longrightarrow \cdot\cdot$$

is of the form

$$0 \longrightarrow H_i(\mathbb{K}_{\mathbf{a}}(J)) \longrightarrow 0$$

and then $H_i(\mathbb{K}_{\mathbf{a}}(J)) = 0$. $\square$

Proposition 3.6. *If $H_i(\mathbb{K}_{\mathbf{a}}(I)) \neq 0$ for some i, then $x^{\mathbf{a}}$ is a generator of some $J \in MVT(I)$*

Proof: The result comes from the preceding lemmata. □

From this proposition we have that all the multidegrees of Koszul generators of I appear in $MVT(I)$. We could have some multidegrees appearing in the Mayer-Vietoris tree that do not correspond to Koszul generators. However, there are some of them that will always have a corresponding generator. Among the nodes in $MVT(I)$ we call *relevant nodes* to those in even position or in position 1. This is because for all other nodes in odd position, the corresponding generators have already appeared in a *relevant node*.

Lemma 3.7. *If $x^{\mathbf{a}}$ appears only once as a generator of a relevant node J in $MVT(I)$ then there exists exactly one generator in $H_*(\mathbb{K}(I))$ which has multidegree $\mathbf{a}$.*

Proof: Let $x^{\mathbf{a}}$ be a generator of I (i.e. of the relevant node in position 1). Then $x^{\mathbf{a}}$ appears only in this relevant node in $MVT(I)$ and it is clear that there is a generator of $H_0(\mathbb{K}(I))$ that has multidegree $\mathbf{a}$, namely $x^{\mathbf{a}}$ itself.

If $x^{\mathbf{a}}$ appears only as a generator of the node J in even position p, then exists $L \in MVT(I)$ such that $J = \tilde{L}$. Then we have the following exact sequence in homology (we know that $x^{\mathbf{a}}$ does not appear in L)

$$0 \longrightarrow H_1(\mathbb{K}_{\mathbf{a}}(L)) \longrightarrow H_0(\mathbb{K}_{\mathbf{a}}(J)) \longrightarrow 0$$

so we have that the connecting morphism is an isomorphism and we have exactly one generator in $H_1(\mathbb{K}(L))$ with multidegree $\mathbf{a}$. When iterating the process to compute $H(\mathbb{K}(I))$ we will always be in one of the following situations:

- If there exists $M \in MVT(I)$ such that $M' = L$ then we have the following exact sequence

  $$\cdot\cdot \longrightarrow H_1(\mathbb{K}_{\mathbf{a}}(M)) \longrightarrow H_1(\mathbb{K}_{\mathbf{a}}(L) \oplus \mathbb{K}_{\mathbf{a}}(\langle m_s \rangle)) \longrightarrow H_0(\mathbb{K}_{\mathbf{a}}(\tilde{M})) \longrightarrow 0$$

but in this case $\tilde{M}$ is a relevant node, and thus the sequence is just $0 \longrightarrow H_1(\mathbb{K}_{\mathbf{a}}(M)) \longrightarrow H_1(\mathbb{K}_{\mathbf{a}}(L) \oplus \mathbb{K}_{\mathbf{a}}(\langle m_s \rangle)) \longrightarrow 0$

- If there exists $M \in MVT(I)$ such that $\tilde{M} = L$ then we have the following exact sequence

$$\cdot\cdot \longrightarrow H_2(\mathbb{K}_{\mathbf{a}}(M)) \longrightarrow H_2(\mathbb{K}_{\mathbf{a}}(M') \oplus \mathbb{K}_{\mathbf{a}}(\langle m_s \rangle)) \longrightarrow H_1(\mathbb{K}_{\mathbf{a}}(L)) \longrightarrow 0$$

but in this case the sequence is just

$$0 \longrightarrow H_2(\mathbb{K}_{\mathbf{a}}(M)) \longrightarrow H_1(\mathbb{K}_{\mathbf{a}}(L)) \longrightarrow 0$$

Iterating this process, we have that there is one generator in some $H_i(\mathbb{K}(I))$ in multidegree $\mathbf{a}$. $\square$

With this proposition, collecting all the non repeated generators of the relevant nodes, we have a lower bound for the Betti numbers of I, let us call $\hat{\beta}_i(I)$ to these 'estimated' Betti numbers. On the other hand, if we collect *all* the generators in the relevant nodes, we have an upper bound for the Betti numbers. We denote $\bar{\beta}_i(I)$ these upper bounds. In some cases these are all the Koszul generators, and thus, the bounds for the Betti numbers are sharp.

Remark 3.8. Observe that what we obtain from the Mayer-Vietoris Tree of the ideal is not only bounds for the Betti numbers, but a subset of the multidegrees in which the Betti numbers are nonzero, and a superset of this. THis information of course provides the lower and upper bound for the Betti numbers, but is more complete. In fact this will be very useful for applications.

Proposition 3.9. *If I is a generic monomial ideal, then $\hat{\beta}_i(I) = \beta_i(I)\, \forall i$. If I is minimally resolved by its Taylor resolution, then $\overline{\beta_i}(I) = \beta_i(I)\, \forall i$.*

Proof:

- A monomial ideal $I = \langle m_1, \ldots, m_r \rangle$ is called *generic* [10] if whenever two distinct minimal generators m_i and m_j have the same positive

degree in some variable, there is a third generator m_k which strictly divides $lcm(m_i, m_j)$. For every monomial ideal $I = \langle m_1, \ldots, m_r \rangle$ its *Scarf Complex* of I, Δ_I is the collection of all subsets of $\{m_1, \ldots, m_r\}$ whose least common multiple is unique:

$$\Delta_I = \{\sigma \subseteq \{1, \ldots, r\} | m_\sigma = m_\tau \Rightarrow \sigma = \tau\}$$

If I is a generic monomial ideal, then the chain complex supported on the Scarf complex Δ_I minimally resolves R/I. Now, let c be a generator of the Koszul homology of I, let $\mathbf{a}$ be its multidegree. Then, $\mathbf{a}$ is the multidegree of a generator of a relevant node in $MVT(I)$ and thus it is the least common multiple of a set of minimal generators of I, lets call this set S; we have then that $m_S = x^{\mathbf{a}}$. Assume $\mathbf{a}$ appears in some other relevant node, then there exists another set T of minimal generators of I such that $m_T = x^{\mathbf{a}}$. If I is generic, we know that the Scarf complex minimally resolves I and then the multidegree of each generator of the Koszul homology of I corresponds to the least common multiple of exactly one set of minimal generators of I. Hence, we have a contradiction, and $x^{\mathbf{a}}$ appears only once in a relevant node of $MVT(I)$.

- If the Taylor Resolution of I is minimal, we know that we have a Koszul generator corresponding to every subset of minimal generators of I. These are exactly the generators appearing in the relevant nodes of $MVT(I)$ in this case. □

Of course if we only have non-repeated generators in the relevant nodes of $MVT(I)$ then $\overline{\beta_i}(I) = \hat{\beta}_i(I) = \beta_i(I) \, \forall i$. When we have repeated generators we still can obtain all the relevant information,including the Betti numbers, the Multigraded Betti numbers, and the actual generators of the homology modules, using the properties of the exact sequence, and other considerations (see [14] and later).

3.2 Mayer-Vietoris resolution of Monomial Ideals

This small section gives just a brief (and somewhat informal) description of the process of obtaining a (minimal) resolution of our ideal I using its Mayer-Vietoris tree. A usual technique to build resolutions in a recursive way is using *mapping cone resolutions* [11]. For them, one must use what is known as the *algebraic mapping cone* (see [7] for example) of a map between two chain complexes. The usual procedure is to have a map between two resolutions of ideals with fewer generators and from the mapping cone of this map, one constructs a resolution of an ideal with one more generator. Two problems show up when using such a construction: first, it's is not always easy to build the chain complex map between the two small resolutions; and second, the mapping cone of two minimal resolutions needs not to be minimal, so some minimization process must be performed later, and this can be rather inefficient. The first problem is solved using tecniques from *Effective Homology* [12]; the second can be treated in an efficient way using these techniques.

The starting point for the recursive process we want to implement is the short exact sequence of ideals associated to the Mayer-Vietoris sequence of an ideal I. We begin with the sequence

$$0 \longrightarrow \tilde{I}_2 \xrightarrow{\mathtt{i}} I_1 \oplus \langle m_2 \rangle \xrightarrow{\mathtt{j}} I_2 \longrightarrow 0$$

We want to use the following theorem from *effective homology*:

Theorem 3.10. *Let $(A, \mathtt{i}, \rho, B, \mathtt{j}, \sigma, C)$ be an effective short exact sequence of effective chain-complexes:*

$$0 \longrightarrow A \underset{\rho}{\overset{\mathtt{i}}{\rightleftarrows}} B \underset{\sigma}{\overset{\mathtt{j}}{\rightleftarrows}} C \longrightarrow 0$$

where $\mathtt{i}$ and $\mathtt{j}$ are chain complex morphisms; ρ and σ are graded module morphisms and the following relations hold: $id_A = \rho\mathtt{i}$, $id_B = \mathtt{i}\rho + \sigma\mathtt{j}$ and $id_C = \mathtt{j}\sigma$, then, an algorithm constructs a canonical reduction (see [12]) between $Cone(\mathtt{i})$ and C from the data.

The recursive application of this theorem builds a resolution of I supported on the Mayer-Vietoris tree of I. To use it, we need to transform our initial short exact sequence in an *effective* one, i.e. we need maps $\rho : I_{s-1} \oplus \langle m_s \rangle \to \tilde{I}_s$ and $\sigma : I_s \to I_{s-1} \oplus \langle m_s \rangle$ such that $\rho\mathfrak{i} = 1$, $\mathfrak{j}\sigma = 1$ and $\mathfrak{i}\rho + \sigma\mathfrak{j} = 1$, and it is not hard to define such maps for our monomial ideals, for each elements one just needs to be careful with the summands that lie in each of the ideals involved, or the differences among them. Another ingredient we need are *effective* resolutions of the leftmost ideals, i.e. we need explicit contracting homotopies for them. As the initial step works with ideals generated by one and two monomials, we know that the Taylor resolution of them is minimal, and moreover, an explicit contracting homotopy is known for it [13] thus, theorem 3.10 provides us an *effective* resolution of I_2 and we can go on with the process.

The use of effective short exact sequences allows us to overcome the main (algorithmic) dificulty when using this recursive procedure, namely the construction of the comparison maps $\mathfrak{i}$ [11]. This procedure gives us also a good recursive description of the differentials involved in the so constructed resolution. We know that $Cone(\mathfrak{i})_q = B_q \oplus A_{q-1}$ and the differential is given by $d_q^{Cone(\mathfrak{i})} = \begin{pmatrix} d_q^B & \mathfrak{i}_q \\ 0 & -d_{q-1}^A \end{pmatrix}$. Thus, if we keep minimality at each step, we know that the only possible part of the matrix which can be reduced is that corresponding to $\mathfrak{i}_q$, and the minimalization process is improved. Moreover, as we keep track of the multidegrees involved, only when the same multidegree appears in the resolutions of both $\tilde{I}_s$ and $I_{s_1} \oplus \langle m_s \rangle$ at the same homological dimension we can have some non-minimality on the resolution of I_s. This provides a new criterion for detecting multidegrees such that even if are repeted at some relevant nodes of $MTV(I)$, will contribute to the Koszul homology of I. Note that the size of the resolution constructed this way is given by the number of generators in the relevant nodes of the tree.

4 Reading the Mayer-Vietoris Tree

We know that the only multidegrees relevant to the homology computation of I are those in position 1 of $MVT(I)$, from which we obtain $H_0(\mathbb{K}(I))$ and those in even position in $MVT(I)$. The dimension of the homology to which (up to possible cancellations if we have long homology sequences) they contribute, can also be read from their position in the tree. If we asign a *dimension* to every node in $MVT(I)$ in the following way: $dim(MVT_1(I)) = 0$, if $dim(MVT_p(I)) = d$ then $dim(MVT_{2p}(I)) = d + 1$, $dim(MVT_{2p+1}(I)) = d$, then the generators of each relevant node contribute to the homology modules in the homological dimension given by the dimension of the node. To verify this just consider the Mayer-Vietoris sequence at each multidegree in the tree, and the fact that at the bottoms of the tree, one has always a node with two generators and two children with one genertor each, the right child contributes with its homology just in dimension zero (it is a generator of its father) and the left child contributes with one generator in dimension one, the recursive construction of the tree and the sequences yield the correspondence between dimensions.

If we compute $MVT(I)$ keeping track of the dimension of a node and the number of generators it has, then the (bounds of the) Betti numbers can be computed at once with the tree. Algorithm 1 computes the Mayer-Vietoris tree of I and every node is given by its position, dimension and generators:

On steps 10 and 11 the children of a given node are computed. The procedures $tilde(ideal)$ and $ideal'$ compute the corresponding new ideals and give the list of generators sorted. A good ordering would be as said before, putting in the last place a generator with highest exponent in some variable. Another strategy could be keeping the ideal generic as long as possible, etc. Note that the complexity of this algorithm depends strongly on the number of generators of I and has only a weak dependence on the number of variables (the neccesary divisibility tests depend on the number of variables); it is independent on the degrees of the generators involved.

Once the tree is built in this way, we can use the information in it reading

Algorithm 1 : Mayer-Vietoris Tree of a Monomial ideal I

Input : Minimal generatng set of $I = \langle m_1, \ldots m_r \rangle$
Output : $MVT(I)$ as a list of pairs (position, ideal)

```
1 if r = 1 then return {({1, 0, {m_1}})}
2    else
3    tree := {(1, 0, {m_1, ..., m_r})}
4    undone := tree
5    while undone ≠ ∅ do
6       node := first(undone)
7       undone := tail(undone)
8       ideal := ideal(node)
9       pos := position(node), dim := dimension(node)
10        append(tree, (2 * pos, dim + 1, tilde(ideal)))
11        append(tree, (1 + 2 * pos, dim, ideal'))
12        if number of generators(tilde(ideal)) > 1
           then append(undone, (2 * p, dim + 1, tilde(ideal)))
13        if number of generators(ideal') > 1
           then append(undone, (1 + 2 * p, dim, ideal'))
14    end while
15    return tree
16 end if
```

the tree upwards looking for the resolution, or reading directly from the output of the algorithm the bounds $\hat{\beta}_i$ and $\bar{\beta}_i$ for each i.

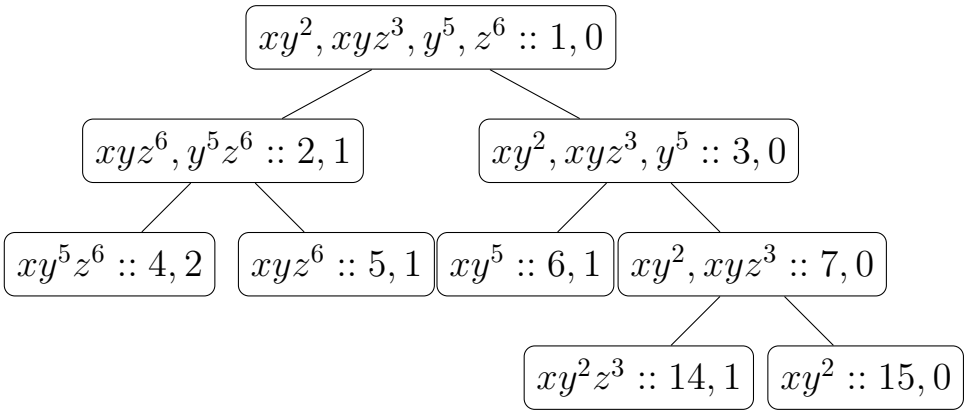
As the multidegree of the generators and the dimension(s) in which they (can) contribute to the homology are given in the tree, one can read from it the involutivity degree of the system our ideal is associated to. We proceed downwards from the highest possible degree. If we can assure this

multidegree is present in the Koszul homology, because it appears only once in some relevant node or because of some other considerations (see later) then we have our involutivity degree. If we find it does not contribute to the Koszul homology of the ideal, we go to the next multidegree and perform the same test until we find an element of the homology, and we are done.

5 Examples and Experiments

5.1 A first small example

Here we have an example of the computation of the Koszul homology of a monomial ideal in 3 variables with four generators. In this case, the ideal is not generic and still it's homology can be automatically computed from the Mayer-Vietoris tree. Let $I = \langle xy^2, xyz^3, y^5, z^6 \rangle$, then it's Mayer-Vietoris tree is



here the numbers denote the positions and dimensions of the nodes in the tree. According to our rules, we can read the Betti numbers of I from $MVT(I)$:

$$\beta_0(I) = 4,\ \beta_1(I) = 4,\ \beta_2(I) = 1$$

and looking at the generators in each position we also obtain the *multi-graded Betti numbers*:

$$\begin{array}{ll} \beta_0: & 120,\ 113,\ 050,\ 006 \\ \beta_1: & 116,\ 056,\ 150,\ 123 \\ \beta_2: & 156 \end{array}$$

If we look at the bigrading, we see that the dimensions f the bigraded homology modules of I are:

$$dim(H_{3,0}(I)) = 1, \quad dim(H_{5,0}(I)) = 2, \quad dim(H_{6,0}(I)) = 1$$

$$dim(H_{5,1}(I)) = 2, \quad dim(H_{7,1}(I)) = 1, \quad dim(H_{10,1}(I)) = 1$$

$$dim(H_{10,2}(I)) = 1$$

and thus, the *degree of involution* of I is 10.

In general some 'unpleasant' multidegrees, i.e. those that are repeated in some relevant nodes, will appear. In this case, we can still use some criteria to determine wether they contribute to the Koszul homology of the ideal or not, like the maps in the resolution we saw in the precedent section or the length of the associated Mayer-Vietoris sequence. We can also use *local* computations of the homology, using simplicial techniques for the computation of generators, as seen in [14]. These simplicial techniques include easy criteria for the vanishing of the homology at a given multidegree, or actual computation of the generators at each multidegree.

In the following table we show the results of computations[1] in some random examples, n is the number of variables g the number of generators, S the size of the minimal resolution and U the number of 'unpleasant' multidegrees in which we have to make some computations; the column CoCoA shows the time (in seconds) used by the CoCoA [15] command BettiDiagram in computing the Betti Diagram of I and the column Tree shows the time in which our algorithm, implemented in CoCoAL, returns the estimated Betti numbers of I. The question mark means that computations were stopped after five hours with no result.

[1] All computations were made on a Pentium IV processor (2.5GHz) running CoCoA 4.5 under Linux (Mandriva 2006).

n	g	S	U	CoCoA	Tree
5	19	235	20	3.61	2.44
6	34	713	10	18.89	5.59
30	12	3275	26	60.85	13.18
50	13	8191	0	800.92	54.01
70	12	4095	0	1809.0	161.81
100	10	1023	0	732.16	11.0
105	13	8191	0	?	100.74

5.2 Valla Ideals

In his paper [16], G. Valla studies the Betti numbers of some monomial ideals strongly related with fat points. These ideals are not stable and in [16], formulas for their Betti numbers are given. If we want to study the multigraded Betti numbers of these ideals or their Koszul homology, we can use their Mayer-Vietoris trees. The ideals Valla studies can be described as follows:

$$I_{a,b} = \langle x_1^{a+b-2j} J^j, x_1^{a-t} J^t \rangle, \quad j = 0, \ldots, b-1, \quad t = b, \ldots a$$

where a and b are positive integers such that $a \geq b$, and $J = \langle x_2, \ldots, x_n \rangle$.

These ideals grow very rapidly, and their Mayer-Vietoris trees compute very efficiently their multigraded Betti numbers. Here is a table with several examples. In the table, n is the number of variables, a and b define the corresponding Valla ideal, g is the number if generators, Min is the size of the minimal resolution, MVT is the size of the Mayer-Vietoris resolution, $Time$ shows the time (in seconds) that took to compute the Mayer-Vietoris tree of each ideal, and $CoCoA$ shows the time it took CoCoA to compute the BettiDiagram of each ideal.

Note that for all the examples in the table (and for all other examples computed so far by the author) the Mayer-Vietoris tree provides the exact (multigraded) Betti numbers. The question about this being a general result is only a conjecture.

n	a	b	g	Min	MVT	Time	CoCoA
3	3	2	10	31	31	0.13	0.02
3	6	4	28	97	97	0.45	0.05
4	3	2	20	111	111	0.32	0.04
4	6	4	84	545	545	3.83	0.27
6	3	2	56	1023	1023	3.08	0.53
6	6	4	462	10625	10625	123.03	91.26
8	3	2	120	7423	7423	21.73	23.86
8	6	4	1716	141569	141569	2906.49	?
10	3	2	220	47103	47103	167.10	?

5.3 Ideals from Reliability Theory

In their paper [17], Giglio and Wynn relate coherent systems, from Reliability Theory, with monomial ideals. In reliability computations, an inclusion-exclusion identity for the probability of failure of the system is used. In the commutative algebra side, this is strongly related to compute multigraded Betti numbers and Multigraded Hilbert series. In [17], the Scarf complex (see [10] for example) is used to improve this inclusion-exclusion identity, so that fewer terms are used. Instead of the Scarf complex, we can use Mayer-Vietoris trees for this improvements. In the case of generic ideals, both the Scarf complex and The Mayer-Vietoris trees will yield the same terms. In the non-generic case, we expect the trees will be closer to the minimal resolution than the non-generic Scarf complex. In fact this can be seen in the examples computed in [17]:

The first non-generic example in [17], corresponding to a binary network, is the ideal $I = \langle x_1x_6, x_1x_4x_7, x_2x_4x_6, x_1x_4x_5x_8, x_2x_7, x_3x_4x_5x_6, x_2x_5x_8, x_3x_5x_7, x_3x_8\rangle$ in $\mathbb{Q}[x_1, \ldots, x_8]$. The Scarf complex associated to I has size 103, the Mayer-Vietoris resolution has size 87, the same that the minimal resolution.

The following example, corresponding to a binary non-network, is the ideal in $\mathbb{Q}[x_1, \ldots, x_4]$ given by $I = \langle x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4\rangle$, the Scarf complex of which has size 19, while the Mayer-Vietoris and the minimal resolution have both size 17.

Finally, the last example, corresponding to a multistate system, is the ideal $I = \langle x_1^3x_2^2x_3^3x_4, x_1^2x_2^3x_3^3x_4, x_1^2x_3^2x_4^2, x_1x_2x_3^2x_4^2, x_2^2x_3^2x_4^2, x_1^3x_3x_4^3, x_1^2x_2x_3x_4^3,$ $x_1x_2^2x_3x_4^3, x_2^3x_3x_4^3\rangle$ in the ring $\mathbb{Q}[x_1, \ldots, x_4]$. In this case, the size of the Scarf complex is 31, while both Mayer-Vietoris and minimal resolutions have size 25.

In all these examples we see that the Mayer-Vietoris trees improve the computations made with the help of the Scarf complex, and thy even yield the actual (multigraded) Betti numbers for the examples in the literature.

6 Future work

Future work includes making a complete study about which monomial orderings are better to obtain smaller and/or better trees for a given ideal, and give caracterizations of families of ideals such that their Mayer-Vietoris trees have no repeated relevant generators. Also merging the Mayer-Vietoris tree with the different techniques presented in [14] and the criteria coming from the Mayer-Vietoris resolution and homology sequence, in order to implement an algorithm to compute the Koszul homology of monomial ideals. Finally, another issue is to fully describe and implement the transfer of the results in the monomial case to computations in the polynomial case. For this, the Homological Perturbation Lemma gives a good tool, and it's application to our case is described in [1] and [18].

References

[1] Lambe, L. and Seiler, W.: Differential Equations, Spencer Cohomology and Computing Resolutions, *G*eorgian Mathematical Journal, 9, 723-772, 2002.

[2] Chrastina, J.: *The Formal Theory of Differential Equations, Folia Mathematica 6*, Masaryk University, Brno, 1998.

[3] Quillen, D.: *Formal Properties of Over-Determined Systems of Linear Partial Differential Equations, PhD Thesis*, Harvard University, Cambridge, 1964.

[4] Spencer D., Overdetermined Systems of linear Partial Differential Equations, *Bull. Amer. Math. Soc.* 75, 179-239, 1969.

[5] Cox, D., Little, D. and O'Shea, D.: *Using Algebraic Geometry*, Springer Verlag, New York, 1998.

[6] Cox, D., Little, D. and O'Shea, D.: *Ideals, Varieties and Algorithms*, Springer Verlag, New York, 1996.

[7] Massey,W.: *A Basic Course in Algebraic Topology*, Springer Verlag, New York, 1991.

[8] Taylor, D.: *Ideals Generated by Monomials in an R-sequence, PhD. Thesis*, University of Chicago, 1960.

[9] Gasharov, V., Peeva, I. and Welker, V.: The lcm-lattice in Monomial Resolutions, *Math. Res. Lett.* 6, 521-532, 1999.

[10] Miller E. and Sturmfels, B.: *Combinatorial Commutative Algebra*, Springer Verlag, New York, 2004.

[11] Herzog, J. and Takayama, Y., Resolutions by Mapping Cones, *Homology, Homotopy and Applications*, 4-2, 277-294, 2002.

[12] Rubio, J. and Sergeraert, F.: Constructive Algebraic Topology, *Bulletin des Sciences Mathématiques*, 126, 389-412, 2002.

[13] Fröberg, R.: Some complex constructions with applications to Poincaré series, in *Séminaire d'Algèbre Paul Dubreil 31ème année, Paris 1977-1978*, Springer Verlag, Berlin, 272-284, 1979.

[14] Sáenz de Cabezón, E.: Computing Koszul homology for Monomial Ideals, in *Proceedings of the RWCA*, Basel, Available at `http://arxiv.org/abs/math.AC/0605325`, 2006.

[15] CoCoATeam, *CoCoA: a system for doing Computations in Commutative Algebra*, Available at `http://cocoa.dima.unige.it`

[16] Valla, G.: Betti Numbers of some Monomial Ideals, *Proc. of the AMS*, 133-1, 57-63, 2004.

[17] Giglio, B. and Wynn, H., Monomial Ideals and the Scarf Complex for Coherent Systems in Reliability Theory, *Annals of Statistics*, 32-3, 1289-1311, 2004.

[18] Sergeraert, F.: Koszul Complexes and Minimal resolutions, preprint available at `http://www-fourier.ujf-grenoble.fr/ sergerar/Papers/Koszul.pdf`

Normal forms of 3 degree of freedom Hamiltonian systems at equilibrium in the semisimple resonant case[1]

Jan A. Sanders

Vrije Universiteit, Faculty of Sciences, Department of Mathematics
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Abstract

An algorithm is given to compute a Stanley decomposition for the normal form of a three degree of freedom Hamiltonian at equilibrium in the semisimple resonant case. This algorithm is then applied to compute Stanley decompositions of the normal form of the first and second order resonances.

Keywords: Hamiltonian, normal form, Stanley decomposition

1 Introduction

We consider Hamiltonians at equilibrium with quadratic term

$$H^0 = \sum_{j=1}^{3} m_j x_j y_j,$$

where $x_j = q_j + ip_j$ and $y_j = q_j - ip_j$, and the q_j, p_j are the real canonical coordinates. We assume $m_j \in \mathbb{N}$, although it is straightforward to apply the results in the more general case $m_j \in \mathbb{Z}$. The signs are important in the nonsemisimple case, and of course, in the stability considerations. With these quadratic terms we speak of the semisimple resonant case.

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

For the asymptotic analysis of such resonances, see [SV85]. Most bibliographic references are taken form the second edition of [SV85], in preparation. We now pose the problem to find the description of a general element

$$H \in k[[x_1, y_1, x_2, y_2, x_3, y_3]]$$

such that $\{H^0, H\} = 0$ (see [Mur03, Section 4.5]). Since the flow of H^0 defines a compact Lie group (S^1) action on $T^*\mathbb{R}^3$, we know beforehand that H can be written as a function of a finite number of invariants of the flow of H^0, that is, as

$$H = \sum_{k=1}^{q} F_k(\alpha_1, \cdots, \alpha_{p_k})\beta_k,$$

where $\{H^0, \alpha_\iota\} = \{H^0, \beta_\iota\} = 0$ for all relevant ι. The $\alpha_\iota, \beta_\iota$ are monomials in the $x_1, \cdots, y_3$ variables and are to be determined explicitly. The F_k are completely arbitrary polynomials or formal power series. If it follows from

$$\sum_{k=1}^{q} F_k(\alpha_1, \cdots, \alpha_{p_k})\beta_k = 0$$

that all the F_k are identically zero, we say that we have obtained a **Stanley decomposition** of the normal form. While the existence of the Stanley decomposition follows from the Hilbert finiteness theorem, it is general not unique: both $F(x)$ and $c + G(x)x$ are Stanley decompositions of general functions in one variable x. Notice that the number of primary variables α_ι is in principle variable, contrary to the case of Hironaka decompositions. One can define the minimum number q in the Stanley decomposition as the **Stanley dimension**. In general one can only obtain upper estimates on this dimension by a smart choice of decomposition.
We show that if $M = m_1 + m_2 + m_3$, the Stanley dimension of the ring of invariants of H^0 is bounded by $1 + 2M$.
We do this by giving an algorithm to compute a Stanley decomposition, and we illustrate this by giving the explicit formulae for the genuine zeroth, first and second order resonances, that is, those resonances which have more

than one generator of degree ≤ 4, not counting complex conjugates and $x_j y_j$'s. These resonances are the most important ones from the point of view of the asymptotic approximation of the solutions.

2 The kernel of $\mathrm{ad}(H^0)$

First of all, we see immediately that the elements $\tau_j = x_j y_j$ all commute with H^0. We let $\mathcal{I} = k[[\tau_1, \tau_2, \tau_3]]$. In principle, we work with real Hamiltonians as they are given by a physical problem, but it is easier to work with complex coordinates, so we take the coefficients to be complex too. In practice, one can forget the reality condition and work over $\mathbb{C}$. In the end, the complex dimension will be the same as the real one, after applying the reality condition.
Any monomial in $\ker \mathrm{ad}(H^0)$ is an element of one of the spaces $\mathcal{I}[[y_1^{n_1} x_2^{n_2} x_3^{n_3}]]$, $\mathcal{I}[[x_1^{n_1} y_2^{n_2} x_3^{n_3}]]$, $\mathcal{I}[[x_1^{n_1} x_2^{n_2} y_3^{n_3}]]$, where $\mathbf{n} = (n_1, n_2, n_3)$ is a solution of $n_1 m_1 = n_2 m_2 + n_3 m_3$, $n_2 m_2 = n_1 m_1 + n_3 m_3$, $n_3 m_3 = n_1 m_1 + n_2 m_2$, respectively, and all the $n_j \geq 0$.
In the equation $n_1 m_1 = n_2 m_2 + n_3 m_3$ one cannot have a nontrivial solution of $n_1 = 0$, but if $n_1 > 0$, one can either have $n_2 = 0$ or $n_3 = 0$, but not both. We allow in the sequel n_2 to be zero, that is, we require $n_1 > 0$, $n_2 \geq 0$ and $n_3 > 0$.
We formulate this in general as follows. Consider the three equations

$$n_i m_i = n_{i^+} m_{i^+} + n_{i^{++}} m_{i^{++}}.$$

where the increment in the indices is in $\mathbb{Z}/3 = (1, 2, 3)$ (that is, $2^{++} \equiv 1$, etc.), where we allow n_{i^+} to be zero, but n_i and $n_{i^{++}}$ are strictly positive. We now solve for given $\mathbf{m}$ the equation $n_1 m_1 = n_2 m_2 + n_3 m_3$, and then apply a cyclic permutation to the indices of $\mathbf{m}$.
Suppose that $\gcd(m_2, m_3) = g_1 > 1$. In that case, assuming $\mathbf{m}$ is primitive, we may conclude that $g_1 | n_1$. Let $n_1 = g_1 \bar{n}_1$, $m_j = g_1 \bar{m}_j, j = 2, 3$. Then

$$\bar{n}_1 m_1 = n_2 \bar{m}_2 + n_3 \bar{m}_3, \quad \gcd(\bar{m}_2, \bar{m}_3) = 1.$$

By cyclic permutation, we assume now that $\gcd(\bar{m}_i, \bar{m}_j) = 1$, and we call $\bar{\mathbf{m}}$ the **reduced** resonance. Observe that the Stanley dimension of the ring of invariants is the same for a resonance and its reduction.
Obviously, keeping track of the divisions by the gcd's, one can reconstruct the solution of the original resonance problem from the reduced one. Observe that in terms of the coordinates, the division is equivalent to taking a root, and this is not a symplectic transformation.
Dropping the bars, we again consider $n_1 m_1 = n_2 m_2 + n_3 m_3$, but now we have $\gcd(m_2, m_3) = 1$.
If $m_1 = 1$, we are immediately done, since the solution is simply $n_1 = n_2 m_2 + n_3 m_3$, with arbitrary integers $n_2 \geq 0, n_3 > 0$.
So we assume $m_1 > 1$ and we calculate modm_1, keeping track of the positivity of our coefficients. Let $m_j = \bar{m}_j + k_j m_1, j = 2, 3$, with $0 < m_j < m_1$ since $\gcd(m_j, m_1) = 1$. Let $\tilde{m}_3 = m_1 - m_3$, so again $0 < \tilde{m}_3 < m_1$. For $q = 0, \cdots, m_1 - 1$ let

$$\begin{aligned} n_2 &= q\tilde{m}_3 + l_2 m_1 \\ n_3 &= q\bar{m}_2 + l_3 m_1 \end{aligned}$$

with the condition that if $q = 0$, then $l_3 > 0$. Then

$$\begin{aligned} n_1 m_1 &= (q\tilde{m}_3 + l_2 m_1) m_2 + (q\bar{m}_2 + l_3 m_1) m_3 \\ &= q\tilde{m}_3 m_2 + q\bar{m}_2 m_3 + m_1(l_2 m_2 + l_3 m_3) \\ &= q\tilde{m}_3(\bar{m}_2 + k_2 m_1) + q\bar{m}_2(\bar{m}_3 + k_3 m_1) + m_1(l_2 m_2 + l_3 m_3) \\ &= q\tilde{m}_3\bar{m}_2 + q\bar{m}_2\bar{m}_3 + m_1(q\tilde{m}_3 k_2 + q\bar{m}_2 k_3 + l_2 m_2 + l_3 m_3) \\ &= m_1(q(k_2\tilde{m}_3 + (1 + k_3)\bar{m}_2) + l_2 m_2 + l_3 m_3) \end{aligned}$$

or

$$n_1 = q(k_2\tilde{m}_3 + (1 + k_3)\bar{m}_2) + l_2 m_2 + l_3 m_3, \quad q = 0, \cdots, m_1 - 1.$$

This is the general solution of the equation $n_1 = n_2 m_2 + n_3 m_3$.
The solution is not necessarily giving us an irreducible monomial: it could be the product of several monomials in $\ker \mathrm{ad}(H^0)$. To analyze this we put

$$q\bar{m}_2 = \psi_2^q m_1 + \phi_2^q, 0 \leq \phi_2^q < m_1, \psi_2^q \geq 0$$

and

$$q\tilde{m}_3 = \psi_3^q m_1 + \phi_3^q, 0 \leq \phi_3^q < m_1, \psi_3^q \geq 0.$$

We now write $y_1^{n_1} x_2^{n_2} x_3^{n_3}$ as $\langle n_1, n_2, n_3 \rangle$. Then

$$\begin{aligned}
&\langle n_1, n_2, n_3 \rangle = \\
&= \langle q(k_2\tilde{m}_3 + (1+k_3)\bar{m}_2) + l_2 m_2 + l_3 m_3, q\tilde{m}_3 + l_2 m_1, q\bar{m}_2 + l_3 m_1 \rangle \\
&= \langle m_2, m_1, 0 \rangle^{l_2} \langle m_3, 0, m_1 \rangle^{l_3} \\
&\quad \cdot \langle q(k_2\tilde{m}_3 + (1+k_3)\bar{m}_2), \psi_3^q m_1 + \phi_3^q, \psi_2^q m_1 + \phi_2^q \rangle.
\end{aligned}$$

Let $\phi_1^q = q(k_2\tilde{m}_3 + (1+k_3)\bar{m}_2) - \psi_2^q m_3$. Then

$$\begin{aligned}
\phi_1^q &= q(k_2\tilde{m}_3 + (1+k_3)\bar{m}_2) - \psi_2^q m_3 \\
&= k_2 q\tilde{m}_3 + (1+k_3)(\psi_2^q m_1 + \phi_2^q) - \psi_2^q(\bar{m}_3 + k_3 m_1) \\
&= k_2 q\tilde{m}_3 + (1+k_3)\phi_2^q + \psi_2^q m_1 - \psi_2^q \bar{m}_3 \\
&= k_2 q\tilde{m}_3 + (1+k_3)\phi_2^q + \psi_2^q \tilde{m}_3 \geq 0.
\end{aligned}$$

We now write $\phi_1^q = \tilde{\psi}_3^q m_2 + \chi_1^q$, and we let $\hat{\psi}_3^q = \min(\tilde{\psi}_3^q, \psi_3^q)$. We have

$$\begin{aligned}
&\langle n_1, n_2, n_3 \rangle = \\
&= \langle m_2, m_1, 0 \rangle^{l_2 + \hat{\psi}_3} \langle m_3, 0, m_1 \rangle^{l_3 + \psi_2^q} \\
&\quad \cdot \langle (\tilde{\psi}_3^q - \hat{\psi}_3^q) m_2 + \chi_1^q, (\psi_3^q - \hat{\psi}_3^q) m_1 + \phi_3^q, \phi_2^q \rangle.
\end{aligned}$$

We define

$$\begin{aligned}
\alpha_\iota &= \langle m_{\iota^+}, m_\iota, 0 \rangle \\
\beta_\iota^0 &= \langle m_{\iota^{++}}, 0, m_\iota \rangle \\
\beta_\iota^q &= \langle (\tilde{\psi}_{\iota^{++}}^q - \hat{\psi}_{\iota^{++}}^q) m_{\iota^+} + \chi_\iota^q, (\psi_{\iota^{++}}^q - \hat{\psi}_{\iota^{++}}^q) m_\iota + \phi_{\iota^{++}}^q, \phi_{\iota^+}^q \rangle.
\end{aligned}$$

Thus

$$\langle n_1, n_2, n_3 \rangle = \alpha_1^{l_2'} (\beta_1^0)^{l_3'} \beta_1^q, \quad l_2', l_3' \in \mathbb{N}, q = 0, \cdots, m_1 - 1,$$

or, in other words, $\langle n_1, n_2, n_3\rangle \in \mathcal{I}[[\alpha_1, \beta_1^0]]\beta_1^q$. This means that $\mathcal{I}[[\alpha_1, \beta_1^0]]\beta_1^q$ is the solution space of the resonance problem. Notice that by construction these spaces have only 0 intersection.
Let $\mathcal{K}$ be defined as $\bigoplus_{\iota\in\mathbb{Z}/3}\mathcal{K}_\iota$, where

$$\mathcal{K}_\iota = \bigoplus_{q=0}^{m_\iota-1} \mathcal{I}[[\alpha_\iota, \beta_\iota^0]]\beta_\iota^q.$$

Then we have

Theorem 1 *Let $\bar{\mathcal{K}}$ denote the space of complex conjugates (that is, x_j and y_j interchanged) of the elements of $\mathcal{K}$. Then $\mathcal{I}\oplus\mathcal{K}\oplus\bar{\mathcal{K}}$ is a Stanley decomposition of the $m_1 : m_2 : m_3$-resonance.*

Corollary 1 *In each $\mathcal{K}_\iota$ there are m_ι direct summands. Therefore there are $M = m_1+m_2+m_3$ direct summands in $\mathcal{K}$. This enables us to estimate the Stanley dimension from above by $1+2M$.*

Remark 1 *The number of generators need not be minimal. In particular the β^q's can be generated by one or more elements. We conjecture that the $\beta^q, q = 1, \cdots, m_\iota - 1$, are generated as polynomials by at most two invariants. Furthermore, the β_ι^q's, are for $q > 0$ not algebraically independent of α_ι and β_ι^0. The relations among them constitute what we will call here the defining curve. Since the Stanley decomposition is the ring freely generated by the invariants divided out by the ideal of the defining curve, this gives us a description of the normal form that is independent of the choices made in writing down the Stanley decomposition.*

Remark 2 *The generating functions of the following resonances have been computed by A. Fekken [Fek86]. They are the Poincaré-Hilbert series of the Stanley decomposition and can be computed by computing the Molien series of the group action given by the flow of H^0, that is, by computing circle integrals (or residues).*

The 15 tables contain all the information to compute the Stanley decomposition for the lower order resonances.

ι	α	β^0
1	y_1x_2	y_1x_3
2	y_2x_3	x_1y_2
3	x_1y_3	x_2y_3

Table 1: The $1:1:1$-resonance ([FHPY02])

ι	α	β^0
1	$y_1^2x_2$	$y_1^2x_3$
2	y_2x_3	$x_1^2y_2$
3	$x_1^2y_3$	x_2y_3

Table 2: The $1:2:2$-resonance ([MMV81]). This is derived from the $1:1:1$-resonance by squaring x_1 and y_1.

ι	α	β^0
1	$y_1^3x_2$	$y_1^3x_3$
2	y_2x_3	$x_1^3y_2$
3	$x_1^3y_3$	x_2y_3

Table 3: The $1:3:3$-resonance.This is derived from the $1:1:1$-resonance by raising x_1 and y_1 to the third power.

ι	α	β^0	β^1
1	y_1x_2	$y_1^2x_3$	
2	$y_2^2x_3$	x_1y_2	
3	$x_1^2y_3$	$x_2^2y_3$	$x_1x_2y_3$

Table 4: The $1:1:2$-resonance ([vdAS79, vdA83]). The defining curve is $((\beta_3^1)^2 - \alpha_3\beta_3^0)$.

ι	α	β^0	β^1
1	$y_1^2x_2$	$y_1^4x_3$	
2	$y_2^2x_3$	$x_1^2y_2$	
3	$x_1^4y_3$	$x_2^2y_3$	$x_1^2x_2y_3$

Table 5: The $1:2:4$-resonance ([vdA83]). This is derived from the $1:1:2$-resonance by squaring x_1 and y_1.

ι	α	β^0	β^1
1	$y_1^3x_2$	$y_1^6x_3$	
2	$y_2^2x_3$	$x_1^3y_2$	
3	$x_1^6y_3$	$x_2^2y_3$	$x_1^3x_2y_3$

Table 6: The $1:3:6$-resonance. This is derived from the $1:1:2$-resonance by raising x_1 and y_1 to the third power.

ι	α	β^0	β^1	β^2
1	y_1x_2	$y_1^3x_3$		
2	$y_2^3x_3$	$x_1^3y_2$		
3	$x_1^3y_3$	$x_2^3y_3$	$x_1^2x_2y_3$	$x_1x_2^2x_3$

Table 7: The $1:1:3$-resonance The defining curve is $(\beta_3^1\beta_3^2-\alpha_3\beta_3^0, \beta_3^1\beta_3^1-\alpha_3\beta_3^2, \beta_3^2\beta_3^2-\beta_3^0\beta_3^1)$.

ι	α	β^0	β^1	β^2
1	$y_1^2x_2$	$y_1^6x_3$		
2	$y_2^3x_3$	$x_1^6y_2$		
3	$x_1^6y_3$	$x_2^3y_3$	$x_1^4x_2y_3$	$x_1^2x_2^2y_3$

Table 8: The $1:2:6$-resonance ([VdADW94]). This is derived from the $1:1:3$-resonance by squaring x_1 and y_1.

ι	α	β^0	β^1	β^2
1	$y_1^3x_2$	$y_1^9x_3$		
2	$y_2^3x_3$	$x_1^9y_2$		
3	$x_1^9y_3$	$x_2^3y_3$	$x_1^6x_2y_3$	$x_1^3x_2^2y_3$

Table 9: The $1:3:9$-resonance. This is derived from the $1:1:3$-resonance by raising x_1 and y_1 to the third power.

ι	α	β^0	β^1	β^2
1	$y_1^2x_2$	$y_1^3x_3$		
2	$y_2^3x_3^2$	$x_1^2y_2$	$x_1y_2^2x_3$	
3	$x_1^3y_3$	$x_2^3y_3^2$	$x_1x_2y_3$	$x_1^2x_2^2y_3^2$

Table 10: The $1:2:3$-resonance. The defining curve is $(\beta_2^1\beta_2^1 - \alpha_2\beta_2^0, (\beta_3^1)^3 - \alpha_3\beta_3^0)$.

ι	α	β^0	β^1	β^2
1	$y_1^2x_2$	$y_1^3x_3^2$		
2	$y_2^3x_3^4$	$x_1^2y_2$	$x_1y_2^2x_3^2$	
3	$x_1^3y_3^2$	$x_2^3y_3^4$	$x_1x_2y_3^2$	$x_1^2x_2^2y_3^4$

Table 11: The $2:4:3$-resonance ([vdA83, Kum75]). This is derived from the $1:2:3$-resonance by squaring x_3 and y_3.

ι	α	β^0	β^1	β^2	β^3	β^4
1	$y_1^2x_2$	$y_1^5x_3$				
2	$y_2^5x_3^2$	$x_1^2y_2$	$x_1y_2^3x_3$			
3	$x_1^5y_3$	$x_2^5y_3^2$	$x_1^3x_2y_3$	$x_1x_2^2y_3$	$x_1^4x_2^3y_3^2$	$x_1^2x_2^4y_3^2$

Table 12: The $1:2:5$-resonance ([VdADW94, HW96, Hal99]). The defining curve is $((\beta_2^1)^2 - \alpha_2\beta_2^0, \beta_3^3 - \beta_3^1\beta_3^2, \beta_3^4 - (\beta_3^2)^2, (\beta_3^2)^3 - \beta_3^0\beta_3^1, (\beta_3^1)^2 - \alpha_3\beta_3^2, \beta_3^1(\beta_3^2)^2 - \alpha_3\beta_3^0)$.

ι	α	β^0	β^1	β^2	β^3
1	$y_1^3x_2$	$y_1^4x_3$			
2	$y_2^4x_3^3$	$x_1^3y_2$	$x_1y_2^3x_3^2$	$x_1^2y_2^2x_3$	
3	$x_1^4y_3$	$x_2^4y_3^3$	$x_1x_2y_3$	$x_1^2x_2^2y_3^2$	$x_1^3x_2^3y_3^3$

Table 13: The $1:3:4$-resonance. The defining curve is $((\beta_2^2)^2 - \beta_2^0\beta_2^1, (\beta_2^1)^2 - \alpha_2\beta_2^2, \beta_2^1\beta_2^2 - \alpha_2\beta_2^0, (\beta_3^1)^4 - \alpha_3\beta_3^0)$.

ι	α	β^0	β^1	β^2	β^3	β^4
1	$y_1^3x_2$	$y_1^5x_3$				
2	$y_2^5x_3^3$	$x_1^2y_2$	$x_1^2y_2^4x_3^2$	$x_1y_2^2x_3$		
3	$x_1^5y_3$	$x_2^5y_3^3$	$x_1^2x_2y_3$	$x_1^4x_2^2y_3^2$	$x_1x_2^3y_3^2$	$x_1^3x_2^4y_3^3$

Table 14: The $1:3:5$-resonance. The defining curve is $(\beta_2^1 - (\beta_2^2)^2, (\beta_2^2)^3 - \alpha_2\beta_2^0), \beta_3^4 - \beta_3^1\beta_3^3, (\beta_3^1)^3 - \alpha_3\beta_3^3, (\beta_3^3)^2 - \beta_3^0\beta_3^1, (\beta_3^1)^2\beta_3^3 - \alpha_3\beta_3^0)$.

ι	α	β^0	β^1	β^2	β^3	β^4	β^5	β^6
1	$y_1^3x_2$	$y_1^7x_3$						
2	$y_2^7x_3^3$	$x_1^3y_2$	$x_1y_2^5x_3^2$	$x_1^2y_2^3x_3$				
3	$x_1^7y_3$	$x_2^7y_3^3$	$x_1^4x_2y_3$	$x_1x_2^2y_3$	$x_1^5x_2^3y_3^2$	$x_1^2x_2^4y_3^2$	$x_1^6x_2^5y_3^3$	$x_1^3x_2^6y_3^3$

Table 15: The $1:3:7$-resonance ([VH92]). The defining curve is $((\beta_2^1)^2 - \alpha_2\beta_2^2, (\beta_2^2)^2 - \beta_2^0\beta_2^1, \beta_2^1\beta_2^2 - \alpha_2\beta_2^0, \beta_3^3 - \beta_3^1\beta_3^2, \beta_3^4 - \beta_3^2\beta_3^2, \beta_3^5 - \beta_3^1\beta_3^2\beta_3^2, \beta_3^6 - \beta_3^2\beta_3^2\beta_3^2, \beta_3^1)^2 - \alpha_3\beta_3^2, (\beta_3^2)^4 - \beta_3^0\beta_3^1, \beta_3^1(\beta_3^2)^3 - \alpha_3\beta_3^0)$.

3 Concluding remarks

An obvious application of the given results is the computation of the non-semisimple case. Nilpotent terms in H^0 are possible whenever there is a $1:1$-subresonance and show up in the tables as quadratic terms of type $x_i y_j$. By computing the action of the nilpotent term on the other generators, one can then try to obtain the nonsemisimple normal form, see [Mur03].

References

[Fek86] Alex Fekken. On the resonant normal form of a fully resonant hamiltonian function. Technical report, Vrije Universiteit, Amsterdam, 1986. Rapport 317.

[FHPY02] S. Ferrer, H. Hanssmann, J. Palacián, and P. Yanguas. On perturbed oscillators in $1:1:1$ resonance: the case of axially symmetric cubic potentials. *J. Geom. Phys.*, 40:320–369, 2002.

[Hal99] G. Haller. *Chaos Near Resonance*. Springer, New York, 1999.

[HW96] G. Haller and S. Wiggins. Geometry and chaos near resonant equilibria of 3-DOF Hamiltonian systems. *Physica D*, 90:319–365, 1996.

[Kum75] M. Kummer. An interaction of three resonant modes in a nonlinear lattice. *Journal of Mathematical Analysis and Applications*, 52:64–104, 1975.

[MMV81] L. Martinet, P. Magnenat, and Ferdinand Verhulst. On the number of isolating integrals in resonant systems with 3 degrees of freedom. *Celestial Mech.*, 25(1):93–99, 1981.

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

[SV85] Jan A. Sanders and Ferdinand Verhulst. *Averaging methods in nonlinear dynamical systems*. Springer-Verlag, New York, 1985.

[vdA83] E. van der Aa. First order resonances in three-degrees-of-freedom systems. *Celestial Mechanics*, 31:163–191, 1983.

[VdADW94] E. Van der Aa and M. De Winkel. Hamiltonian systems in $1:2:\omega$–resonance ($\omega = 5$ or 6). *Int. J. Nonlin. Mech.*, 29:261–270, 1994.

[vdAS79] Els van der Aa and Jan A. Sanders. The $1:2:1$-resonance, its periodic orbits and integrals. In Ferdinand Verhulst, editor, *Asymptotic analysis*, volume 711 of *Lecture Notes in Mathematics*, pages 187–208. Springer-Verlag, Berlin, 1979. From theory to application.

[VH92] Ferdinand Verhulst and I. Hoveijn. Integrability and chaos in Hamiltonian normal forms. In *Geometry and analysis in nonlinear dynamics (Groningen, 1989)*, pages 114–134. Longman Sci. Tech., Harlow, 1992.

Evolutionary Systems for Slender Thermomechanical Structures[1]

D. Q. Cao [1 2] D. Liu [1] S. Preston [3] R. W. Tucker [1]

[1] Department of Physics, Lancaster University
Lancaster LA1 4YB, UK
r.tucker@lancaster.ac.uk
[2] School of Astronautics, Harbin Institute of Technology
P O Box 137, Harbin 150001, China
[3] Department of Mathematical Sciences, Portland State University
Portland, OR 97207, USA

Abstract

After reviewing thermomechanical constitutive theory in the context of a simple Cosserat rod model, we explore a simple thermomechanical rod model by generalizing the classical Kirchhoff constitutive relations. The mechanically induced dissipation of energy is distinguished from that produced by the coupling to a temperature field along the rod. The model involves relatively few parameters that may be estimated from experimental data. A further objective is to analyze the evolutionary properties of the governing system of partial differential equations that depend on the (extended Kirchhoff) free energy function and the constitutive relations compatible with the Clausius-Duhem relation appropriate for a slender rod. The evolution characteristics of the system are explored in specific numerical simulations. These indicate the presence of thermally induced damping of axial and torsional excitation modes under forced and free vibration.

Keywords: Cosserat rod, thermomechanics, nonlinear dynamics, constitutive relations, Clausius-Duhem, modelling, Kirchhoff, simulation

[1] Work supported by NEST-Adventure contract 5006 (GIFT).

1 Introduction

Many engineering components have one or more spatial dimensions much smaller than the other(s). Thus they qualify as *slender* structures. Rods, cables, chains, ropes, discs and shells all exhibit these characteristics. In this article emphasis is on structures that can be approximated by curves in space. The aim is to explore simple models, compatible with the laws of physics, that can be used to describe the thermomechanical behavior of rods or strings. Since all materials conduct heat to some extent their thermomechanical characteristics will affect their mechanical response to different types of loading. The behavior of this response is one of the motivations for what follows.

The detailed response of materials to mechanical and thermal loads depends to a great extent on the constitutive properties exhibited by the material and these in turn are constrained by the laws of thermodynamics. Such constraints leave a great deal of freedom in empirical choices of constitutive relation and guidance from experiment often becomes necessary. However this requires a particular theoretical framework or model to be effective.

The linear theory of 1-d thermomechanics can be found in [1]. An extensive theoretical investigation into the non-linear thermomechanics of rods was initiated in 1974 by Green, Naghdi and Wenner[2]. This has been subsequently refined [3] and generalized to include multi-phase structures [4].

The thermomechanics of beams with either flexure [6, 7, 8, 9] or axial deformation [10, 11] have been explored. The energy dissipation in a one-dimensional rod was investigated in [12] using the characteristic equation approach, while exponential stability for linear [13], semilinear [14] and nonlinear [13] one-dimensional rods has been studied using specific energy estimates. Boundary control laws [15] and uniform stabilization methods [16] of linear thermo-elastic beams have been designed using Lyapunov methods.

The thermomechanics of MEMS components has also attracted attention in more recent times [17, 18, 19, 20, 21]. For the transverse vibration of a double-ended tuning fork, Lauderdale and O'Reilly [22] proposed a rod-based model using the rod theory developed by Green, Naghdi and several of their coworkers [23, 24], which incorporates end masses, the anisotropy of silicon, finite deformation effects and thermal influences.

The simple Cosserat theory of rods has been extensively studied and found to have wide applicability [25]. In many circumstances the classical Kirchhoff constitutive relations offer a valuable approximation that can accommodate a large range of coupled non-linear mechanical behaviors in rods. However the inclusion of dissipation in these simple models remains somewhat ad-hoc. In particular when the coupling with heat is involved it is sometimes difficult to separate the effects of mechanical damping from thermally induced damping within the formalism.

Although in principle earlier formulations of the thermomechanics of rods offer many different approximations for modeling they appear to us as somewhat complicated. In this article, after reviewing thermomechanical constitutive theory in the context of the simple Cosserat model, we attempt to find a simple thermomechanical model that can be used to generalize the classical Kirchhoff constitutive relations used in Cosserat rod models. In this way we hope to distinguish the effects of mechanically induced dissipation of energy from that produced by the coupling to a temperature field along the rod. The model involves relatively few parameters that may be estimated from experimental data. A further objective is to analyze the evolutionary characteristics of the governing system of partial differential equations that depend on the (extended Kirchhoff) free energy function and the constitutive relations compatible with the Clausius-Duhem relation appropriate for a slender rod. The evolution characteristics of particular system are explored using specific numerical simulations. These indicate the presence of thermally induced damping of excitation modes under forced and free vibration.

Throughout this paper vectors are elements of Euclidean 3-space with its usual affine structure. They are denoted by lowercase, bold-face symbols, e.g., $\mathbf{u}$, $\mathbf{v}$; tensors are denoted by upper-case, bold-face symbols, e.g., $\mathbf{I}$, $\mathbf{J}$; matrices are denoted by upper-case, italic, bold-face symbols, e.g., $\boldsymbol{M}$, $\boldsymbol{K}$. The symbols $(\dot{})$ and $(')$ denote differentiation with respect to time t and arc-length parameter s, respectively. For any function f that depends explicitly on some variable x we sometimes write f_x for $\frac{\partial f}{\partial x}$. Repeated indices follow the Einstein summation convention.

2 Thermomechanical Cosserat Equations

2.1 The simple Cosserat rod

The general mathematical theory of non-linear elasticity is well established. The general theory of one-dimensional Cosserat continua derived as limits of three-dimensional continua can be consulted in Antman [25]. The theory is fundamentally formulated in the Lagrangian picture in which material elements of a rod are labeled by s. The dynamical evolution of a "slender structure" with mass density, $s \in [0, L_0] \mapsto \rho(s)$, and cross-sectional area, $s \in [0, L_0] \mapsto A(s)$, is governed by Newton's dynamical laws:

$$\rho A\, \ddot{\mathbf{r}} = \mathbf{n}' + \mathbf{f} \tag{1}$$

$$\partial_t(\, \mathbf{I}(\mathbf{w})) = \mathbf{m}' + \mathbf{r}'\mathbf{n} + \mathbf{l} \tag{2}$$

applied to a triad of ortho-normal vectors (directors):

$$s \in [0, L_0] \mapsto \{\mathbf{d_1}(s,t), \mathbf{d_2}(s,t), \mathbf{d_3}(s,t)\} \tag{3}$$

over the space-curve:

$$s \in [0, L_0] \mapsto \mathbf{r}(s,t) \tag{4}$$

at time t where $\mathbf{n}' = \partial_s \mathbf{n}, \equiv \mathbf{n}_s$ $\dot{\mathbf{r}} = \partial_t \mathbf{r}$, $\mathbf{f}$ and $\mathbf{l}$ denote external force and torque densities respectively and $s \in [0, L_0] \mapsto \mathbf{I}$ is a "slender structure" moment of inertia tensor.

In these field equations the *contact forces* $\mathbf{n}$ and *contact torques* $\mathbf{m}$ are determined by constitutive relations[2] involving the *strains* $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$. The strains are themselves defined in terms of the configuration variables $\mathbf{r}$ and $\mathbf{d}_k$ for $k = 1, 2, 3$ by the relations:

$$\mathbf{r}' = \mathbf{v}, \quad \mathbf{d}_k' = \mathbf{u}\,\mathbf{d}_k, \quad \dot{\mathbf{d}}_k = \mathbf{w}\,\mathbf{d}_k. \tag{5}$$

The latter ensures that the triad remains ortho-normal under evolution. One has

$$\mathbf{d}_i\mathbf{d}_i' = \mathbf{d}_i(\mathbf{u}\mathbf{d}_i) = \mathbf{u}(\mathbf{d}_i \cdot \mathbf{d}_i) - \mathbf{d}_i(\mathbf{d}_i \cdot \mathbf{u}) = 2\mathbf{u} \tag{6}$$

$$\mathbf{d}_i\dot{\mathbf{d}}_i = \mathbf{d}_i(\mathbf{w}\mathbf{d}_i) = \mathbf{w}(\mathbf{d}_i \cdot \mathbf{d}_i) - \mathbf{d}_i(\mathbf{d}_i \cdot \mathbf{w}) = 2\mathbf{w} \tag{7}$$

The last equation identifies

$$\mathbf{w} = \frac{1}{2}\sum_{k=1}^{3} \mathbf{d}_k\,\dot{\mathbf{d}}_k \tag{8}$$

with the local angular velocity vector of the director triad.

Here and in much of the following, the components Z_i of any Euclidean vector Z will be taken in the dynamical basis $\mathbf{d}_i$ (i.e. $Z_i = Z \cdot \mathbf{d}_i$), so $\mathbf{v}(s,t) = v_i(s,t)\mathbf{d}_i(s,t)$, $\mathbf{u}(s,t) = u_i(s,t)\mathbf{d}_i(s,t)$, $\mathbf{w}(s,t) = w_i(s,t)\mathbf{d}_i(s,t)$ with summation over 1 to 3.

Assume further that the structure possesses a positive temperature $T(s,t)$, positive internal energy $\epsilon(s,t)$ per unit reference mass, entropy density $\eta(s,t)$ per unit reference mass and a heat flux $q(s,t)$ directed along the rod. In addition there may be an applied heat power $h(s,t)$ per unit reference mass along the structure.

The local energy balance between these sources of power and the power expended by kinetic energy is expressed in the first law of thermodynamics:

[2] In general for non-uniform rods and non-stationary environments the constitutive relations (see below) and director components of the inertia tensor could also depend explicitly on s and t.

$$\rho A\,\dot{\epsilon} + \rho A\,\dot{\mathbf{r}}\cdot\ddot{\mathbf{r}} + \mathbf{w}\cdot(\mathbf{I}(\mathbf{w}))^{\cdot} = \mathbf{f}\cdot\dot{\mathbf{r}} + \mathbf{l}\cdot\mathbf{w} + \rho A\,h + \partial_s(\mathbf{n}\cdot\dot{\mathbf{r}} + \mathbf{m}\cdot\mathbf{w}) - A\partial_s q \quad (9)$$

Elimination of the mechanical power, using the above Cosserat equations of motion, yields:

$$\rho A\,\dot{\epsilon} = P + \rho A\,h - A\partial_s q \quad (10)$$

where $P = n_i\dot{v}_i + \mathbf{m}\cdot\mathbf{w}'$. Furthermore using the relations: $\mathbf{d}_i{}'\,\dot{\mathbf{d}}_i = \mathbf{uw}$, $\mathbf{w}' = \dot{\mathbf{u}} + \mathbf{uw}$ gives $P = n_i\dot{v}_i + m_i\dot{u}_i$ and hence

$$\rho A\,\dot{\epsilon} = n_i\dot{v}_i + m_i\dot{u}_i + \rho A\,h - A\partial_s q \quad (11)$$

The second law of thermodynamics will be adopted as a constraint on the constitutive relations to be discussed below. This is implemented in terms of the Clasius-Duhem inequality

$$\frac{d}{dt}\int_0^{L_0}\rho A\,\eta\,ds \;\geq\; \int_0^{L_0}\rho A\;\frac{h}{T}\,ds - \frac{Aq}{T}\Big|_0^{L_0} \quad (12)$$

For the purposes of this article its local form

$$\rho A\,\dot{\eta} \;\geq\; \rho A\,\frac{h}{T} - \partial_s(\frac{Aq}{T}) \quad (13)$$

will be exploited. By eliminating h using (11) this becomes, in terms of the free energy $\phi \equiv \epsilon - T\eta$,

$$n_i\dot{v}_i + m_i\dot{u}_i - \frac{Aq}{T}\partial_s T - \rho A\,(\dot{\phi} + \dot{T}\eta) \;\geq 0 \quad (14)$$

Thus in addition to six independent dynamical variables encoded in the fields $\mathbf{r}$, $\mathbf{d}_i$ one has four additional dynamical thermal fields η, q, T, ϕ (or ϵ) over the structure.

Now restrict to material constitutive relations of the form:

$$n_i = \hat{n}_i(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j), \qquad m_i = \hat{m}_i(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j)$$
$$\phi = \hat{\phi}(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j), \qquad \eta = \hat{\eta}(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j)$$
$$q = \hat{q}(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j)$$

constrained by (13). Inserting these in (13) gives:

$$(\hat{n}_i - \rho A\, \phi_{v_i})\dot{v}_i + (\hat{m}_i - \rho A\, \phi_{u_i})\dot{u}_i + \rho A\, (\hat{\phi}_T + \hat{\eta})\dot{T} + \rho A\, \hat{\phi}_{\dot{u}_i}\ddot{u}_i + \rho A\, \hat{\phi}_{\dot{v}_i}\ddot{v}_i + \rho A\, \hat{\phi}_{T_s}\dot{T}_s - \frac{A\hat{q}}{T}T' \ \geq 0 \tag{15}$$

Since $\ddot{u}_i, \ddot{v}_i, \dot{T}', \dot{T}$ are not arguments of the constitutive functions one must have $\hat{\eta} = -\hat{\phi}_T, \hat{\phi}_{\dot{u}_i} = \hat{\phi}_{\dot{v}_i} = 0, \hat{\phi}_{T'} = 0$. Thus $\phi = \hat{\phi}(u_i, v_i, T)$ and:

$$(\hat{n}_i - \rho A\, \phi_{v_i})\dot{v}_i + (\hat{m}_i - \rho A\, \phi_{u_i})\dot{u}_i - \frac{A\hat{q}}{T}T' \ \geq 0 \tag{16}$$

Suppose, for some fields α, β, the derivative arguments of the constitutive functions are scaled as follows:

$$T' \mapsto \beta\, T', \quad \dot{u}_i \mapsto \alpha\, \dot{u}_i, \quad \dot{v}_i \mapsto \alpha\, \dot{v}_i. \tag{17}$$

If one writes (16) as $\mathcal{Y}(\alpha, \beta) \geq 0$ then since $\mathcal{Y}(0,0) = 0$ one has generically $\partial_\alpha \mathcal{Y}|_{\alpha=\beta=0} = 0$, $\partial_\beta \mathcal{Y}|_{\alpha=\beta=0} = 0$ or, since these must hold for all $\dot{\boldsymbol{v}}_i, \dot{u}_i, T'$,

$$\hat{n}_i(u_j, v_j, T, 0, 0, 0) = \rho A\, \hat{\phi}_{v_i}(u_j, v_j, T), \tag{18}$$
$$\hat{m}_i(u_j, v_j, T, 0, 0, 0) = \rho A\, \hat{\phi}_{u_i}(u_j, v_j, T), \tag{19}$$
$$\hat{q}(u_j, v_j, T, 0, 0, 0) = 0. \tag{20}$$

Thus one may split off contact forces $\boldsymbol{n}^D$ and torques $\boldsymbol{m}^D$ (with $\hat{n}_i^D$ $(u_j, v_j, T, 0, 0, 0) = 0$ and $\hat{m}_i^D(u_j, v_j, T, 0, 0, 0) = 0$) responsible for dissipation as
follows:

$$\hat{n}_i(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = \rho A\, \hat{\phi}_{v_i}(u_j, v_j, T) + \hat{n}_i^D(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) \tag{21}$$

$$\hat{m}_i(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = \rho A\, \hat{\phi}_{u_i}(u_j, v_j, T) + \hat{m}_i^D(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) \tag{22}$$

and the final form of the Clausius-Duhem constraint on these functions becomes:

$$\begin{aligned}\hat{n}_i^D(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j)\,\dot{v}_i + \hat{m}_i^D(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j)\,\dot{u}_i \\ -\tfrac{AT'}{T}\,\hat{q}(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) \;\geq\; 0\end{aligned} \tag{23}$$

Simple relations compatible with this take the form:

$$\hat{n}_i^D(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = \mathcal{N}_{ik}(u_j, v_j, T, T', \dot{u}_j)\dot{v}_k \tag{24}$$

$$\hat{m}_i^D(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = \mathcal{M}_{ik}(u_j, v_j, T, T', \dot{v}_j)\dot{u}_k \tag{25}$$

for positive matrices $\mathcal{N}, \mathcal{M}$ with $\mathcal{N}_{ik}(u_j, v_j, T, 0, 0) = 0$ and $\mathcal{M}_{ik}(u_j, v_j, T, 0, 0) = 0$ and

$$\hat{q}(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = -\mathcal{K}(u_j, v_j, T, \dot{u}_j, \dot{v}_j)T' \tag{26}$$

for any positive function $\mathcal{K}$ with $\mathcal{K}(u_j, v_j, T, 0, 0) = 0$.

3 A closed system with particular constitutive relations

Consider the closed system of equations (1,2,11) with $\hat{n}_i, \hat{m}_i$ linear in dissipative variables. Thus assume constitutive relations (24), (25) with $\mathcal{N}_{ij}, \mathcal{M}_{ij}$ depending on variables $y = \{v_j, u_j, T\}$ only, but with $\hat{q}(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j)$ general at this point.

Calculating the time derivative in equation (2) with $\mathbf{I}(\mathbf{w})) = \rho\mathbf{I}_j^i w_i\,\mathbf{d}_j$:

$$\partial_t(\mathbf{I}(\mathbf{w})) = \rho\mathbf{I}_j^i\partial_t(w_i\,\mathbf{d}_j) = \rho\mathbf{I}_j^i(\dot{w}_i\,\mathbf{d}_j + w_i(\mathbf{w}\,\mathbf{d}_j)) \tag{27}$$

and the derivative $\mathbf{m}' = m_k'\,\mathbf{d}_k + m_k(\mathbf{u}\,\mathbf{d}_k)$ write (2) in the form

$$\rho\mathbf{I}_k^i\dot{w}_i = -\rho\mathbf{I}_j^i w_i(\mathbf{w}\,\mathbf{d}_j)_k + m_k' + m_j(\mathbf{u}\,\mathbf{d}_j)_k + (\mathbf{vn})_k + l_k,$$

Applying $\mathbf{I}_l^{-1\;k}$ yields

$$\begin{aligned}\rho\dot{w}_l \;=\; & -\rho\mathbf{I}_l^{-1\;k}\mathbf{I}_j^i w_i(\mathbf{w}\,\mathbf{d}_j)_k + \mathbf{I}_l^{-1\;k}m_k' + \mathbf{I}_l^{-1\;k}(\mathbf{um})_k \\ & +\mathbf{I}_l^{-1\;k}(\mathbf{vn})_k + \mathbf{I}_l^{-1\;k}l_k.\end{aligned} \tag{28}$$

Similarly, using $\dot{\mathbf{r}} = \dot{r}_i\,\mathbf{d}_i + r_i(\mathbf{w}\,\mathbf{d}_i)$ and $\ddot{\mathbf{r}} = \ddot{r}_i\,\mathbf{d}_i + 2\dot{r}_i(\mathbf{w}\,\mathbf{d}_i)$ $+ r_i(\dot{\mathbf{w}}\,\mathbf{d}_i) + r_i(\mathbf{w}(\mathbf{w}\,\mathbf{d}_i))$, Equation (1) becomes

$$\begin{aligned}\rho A\ddot{r}_k = &- \rho A[2\dot{r}_i(\mathbf{w}\,\mathbf{d}_i)_k + r_i(\dot{\mathbf{w}}\,\mathbf{d}_i)_k + r_i(\mathbf{w}(\mathbf{w}\,\mathbf{d}_i))_k] \\ &+ n'_k + n_i(\mathbf{u}\,\mathbf{d}_i)_k + f_k. \end{aligned} \tag{29}$$

In the second term on the right side $\dot{\mathbf{w}}\,\mathbf{d}_i = (\dot{w}_k\,\mathbf{d}_k + w_k\,\dot{\mathbf{d}}_k)\,\mathbf{d}_i =$ $(\dot{w}_k\,\mathbf{d}_k + w_k(\mathbf{w}\,\mathbf{d}_k))\,\mathbf{d}_i = (\dot{w}_k\,\mathbf{d}_k + (\mathbf{w}\mathbf{w}))\,\mathbf{d}_i =$ $\dot{w}_l(\,\mathbf{d}_l\,\mathbf{d}_i)$, and $(\dot{\mathbf{w}}\,\mathbf{d}_i)_k = \dot{w}_l(\,\mathbf{d}_l\,\mathbf{d}_i)_k = \dot{w}_l\epsilon_{lik}$. Substituting (29) into (29) yields

$$\begin{aligned}\rho A\ddot{r}_k = &- \rho A[2\dot{r}_i(\mathbf{w}\,\mathbf{d}_i)_k + r_i(\mathbf{w}(\mathbf{w}\,\mathbf{d}_i))_k] \\ &+ Ar_i\epsilon_{lik}\rho\mathbf{I}_l^{-1\,p}\,\mathbf{I}_j^i w_i(\mathbf{w}\,\mathbf{d}_j)_p \\ &- Ar_i\epsilon_{lik}\left[\mathbf{I}_l^{-1\,p}m'_p + \mathbf{I}_l^{-1\,p}m_j(\mathbf{u}\,\mathbf{d}_j)_p + \mathbf{I}_l^{-1\,p}(\mathbf{v}\mathbf{n})_p + \mathbf{I}_l^{-1\,p}l_p\right] \\ &+ n'_k + n_i(\mathbf{u}\,\mathbf{d}_i)_k + f_k, \end{aligned} \tag{30}$$

or

$$\begin{aligned}\rho A\ddot{r}_k = &- \rho A[2\dot{r}_i(\mathbf{w}\,\mathbf{d}_i)_k + r_i(\mathbf{w}(\mathbf{w}\,\mathbf{d}_i))_k] \\ &+ \rho Ar_i\epsilon_{lik}\mathbf{I}_l^{-1\,p}\mathbf{I}_j^i w_i(\mathbf{w}\,\mathbf{d}_j)_p - Ar_i\epsilon_{lik}\mathbf{I}_l^{-1\ p}m'_p + n'_k \\ &- Ar_i\epsilon_{lik}\mathbf{I}_l^{-1\ p}\,(m_j(\mathbf{u}\,\mathbf{d}_j)_p + n_j(\mathbf{v}\,\mathbf{d}_j)_p) - Ar_i\epsilon_{lik}\mathbf{I}_l^{-1\ p}l_p + f_k. \end{aligned} \tag{31}$$

We have separated here the kinematic terms from the constitutive terms of different order and from the terms containing the external forces.

Next substitute the constitutive relations (21)-(25) (using the notation $y = (u_j, v_j, T)$) for the strains and temperature variables:

$$\begin{cases} \hat{n}_i(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = \rho A\,\hat{\phi}_{v_i}(y) + \mathcal{N}_{ik}(y, T', \dot{u}_j)\dot{v}_k, \\ \hat{m}_i(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = \rho A\,\hat{\phi}_{u_i}(y) + \mathcal{M}_{ik}(y, T', \dot{v}_j)\dot{u}_k. \end{cases} \tag{32}$$

Derivatives with respect to s of these constitutive functions have the form

$$\left\{\begin{array}{ll}\partial_s\hat{n}_i(u_j,v_j,T,T',\dot{u}_j,\dot{v}_j) = & U_i(y,T',\dot{u}_j,\dot{v}_k,u',v') \\ & +\dfrac{\partial\mathcal{N}_{ik}}{\partial T'}T''\dot{v}_k+\dfrac{\partial\mathcal{N}_{ik}}{\partial\dot{u}_p}\dot{u}'_p\dot{v}_k+\mathcal{N}_{ik}\dot{v}'_k, \\ \partial_s\hat{m}_i(u_j,v_j,T,T',\dot{u}_j,\dot{v}_j) = & V_i(y,T',\dot{u}_j,\dot{v}_k,u',v') \\ & +\dfrac{\partial\mathcal{M}_{ik}}{\partial T'}T''\dot{u}_k+\dfrac{\partial\mathcal{M}_{ik}}{\partial\dot{v}_p}\dot{v}'_p\dot{u}_k+\mathcal{M}_{ik}\dot{u}'_k.\end{array}\right. \tag{33}$$

with highest derivatives explicitly displayed.
Inserting (32) and (33) into the balance equations (29) and (**??**) yields

$$\begin{aligned}\rho A\ddot{r}_k = & -\rho A[2\dot{r}_i(\mathbf{w}\,\mathbf{d}_i)_k+r_i(\mathbf{w}(\mathbf{w}\,\mathbf{d}_i))_k]+\rho Ar_d\epsilon_{ldk}\mathbf{I}_l^{-1\,p}\mathbf{I}^i_j w_i(\mathbf{w}\,\mathbf{d}_j)_p \\ & -Ar_d\epsilon_{ldk}\mathbf{I}_l^{-1\ i}\left(V_i(y,T',\dot{u}_j,\dot{v}_j,u',v')+\frac{\partial\mathcal{M}_{ij}}{\partial T'}T''\dot{u}_j\right. \\ & \left.+\frac{\partial\mathcal{M}_{ij}}{\partial\dot{v}_p}\dot{v}'_p\dot{u}_j+\mathcal{M}_{ij}\dot{u}'_j\right) \\ & +\left(U_k(y,T',\dot{u}_j,\dot{v}_j,u',v')+\frac{\partial\mathcal{N}_{kj}}{\partial T'}T''\dot{v}_j+\frac{\partial\mathcal{N}_{kj}}{\partial\dot{u}_p}\dot{u}'_p\dot{v}_j+\mathcal{N}_{kj}\dot{v}'_j\right) \\ & -Ar_d\epsilon_{ldk}\mathbf{I}_l^{-1\ p}\epsilon_{ijp}\,(u_i\hat{m}_j+v_i\hat{n}_j)-Ar_d\epsilon_{jdk}\mathbf{I}_j^{-1\,p}l_p+f_k,\end{aligned} \tag{34}$$

and

$$\begin{aligned}\rho\dot{w}_k = & -\rho\mathbf{I}_k^{-1\,p}\mathbf{I}^i_j w_i(\mathbf{w}\,\mathbf{d}_j)_p+\mathbf{I}_k^{-1\,i}l_i \\ & +\mathbf{I}_k^{-1\ i}\left(V_i(y,T',\dot{u}_j,\dot{v}_j,u',v')+\frac{\partial\mathcal{M}_{ij}}{\partial T'}T''\dot{u}_j+\frac{\partial\mathcal{M}_{ij}}{\partial\dot{v}_p}\dot{v}'_p\dot{u}_j+\mathcal{M}_{ij}\dot{u}'_j\right) \\ & +\mathbf{I}_k^{-1\,p}\epsilon_{ijp}(u_i\hat{m}_j+v_i\hat{n}_j).\end{aligned} \tag{35}$$

The energy equation (11) takes the form

$$\rho A\,\hat{\epsilon}_T\dot{T} = -\,\rho A\,(\hat{\epsilon}_{v_i}\dot{v}_i + \hat{\epsilon}_{u_j}\dot{u}_j) + (\rho A\,\hat{\phi}_{v_i}(u_j, v_j, T) + \mathcal{N}_{ik}(u_j, v_j, T)\dot{v}_k)\dot{v}_i$$
$$+\,(\rho A\,\hat{\phi}_{u_i}(u_j, v_j, T) + \mathcal{M}_{ik}(u_j, v_j, T)\dot{u}_k)\dot{u}_i + \rho A\,h - \rho A\partial_s\hat{q}, \quad (36)$$

with $\epsilon = \phi + T\eta = \phi - T\phi_T = \hat{\epsilon}(u, v, T)$. Equations (34,35) and (36) are written to explicitly display highest order derivatives.

One now has a closed system of equations for seven functions of two independent variables: 3 components $r_i(s,t)$ of the position vector, three angular variables ϕ_i defining the frame $\mathbf{d}_i(s,t)$ and the temperature $T(s,t)$. We assume that the forces $\boldsymbol{f}$ and momenta $\boldsymbol{l}$ may depend on the spatial gradients of these functions.

The initial value problem for the system (34,35,36) requires one to specify the initial values (at $t = 0$) of the variables $r_i, \dot{r}_i, \phi_i$, time derivatives of ϕ_i (i.e. equivalently w_i) and that of the temperature T. This system does not allow the direct application of existing results of existence and uniqueness of the Cauchy-Kovalevskaya type ([30]) since the hyperplane $t = 0$ is characteristic for the system due to the presence of third mixed derivatives $\dot{v}_i', \dot{u}_j'$ on the right side of equations (34,35) and of the second mixed derivatives $\dot{v}_i, \dot{u}_j$ on the right side of (36), see ([30]). Existing results for the characteristic Cauchy problem (see [29]) shows that even for the linearization of the system (34,35, 36) near the trivial solution $u = v = 0, T = T_0$ the characteristic Cauchy problem with initial data at $t = 0$ does not have a distributional solution for all initial data.

Notice that without the 3rd order terms (no dissipation) one would get a mixed hyperbolic-parabolic system of equations and that would guarantee a non-characteristic property of $t = 0$ if (as is usually true) the matrix of coefficients of the second order time derivatives in the left side of (34,35) is non-degenerate ([27, 28]).

Yet, one sees from the initial data given above that one can calculate values of all unknown functions and their derivatives on the right side of equations (34,35, 36) and then determine initial values of time derivatives on the left side from the equations. Iterating this procedure one can get

the values of all time derivatives of variables (r_i, ϕ_i, T) at the hyperplane $t = 0$. This allows one to construct a formal power series for the solutions. Thus the classical approach gives one hope for at least the local existence of a solution.

Probably, semigroup methods for studying the coupled hyperbolic-parabolic systems ([27, 28]) can be adopted for the study of the system (34,35, 36) with dissipation (see also [26]. The *essentially nonlinear* character of these equations for properly chosen constitutive relations may yield systems that are amenable to solution as Cauchy or mixed problems (cf the situation with the KdV equation).

3.1 Modeling with Extended Kirchhoff Constitutive Relations

To gain some insight into the influence of thermodynamic constraints on the constitutive assumptions for rods we take some guidance from the phenomenology associated with slender materials that exhibit linear mechanical behavior. In the absence of effects due to heat and dissipation the classic Kirchhoff constitutive relations [25] encompass such a phenomenology. It is natural to adopt a minimal extension of these relations, compatible with this phenomenology, by first assuming that the free energy $\phi = \varepsilon - T\eta$ takes the form

$$\begin{aligned}\rho A\hat{\phi}(u_j, v_j, T) = \ & \frac{1}{2}\hat{\mathbf{v}} \cdot \mathbf{K}(\hat{T})\hat{\mathbf{v}} + \frac{1}{2}\hat{\mathbf{u}} \cdot \mathbf{J}(\hat{T})\hat{\mathbf{u}} \\ & - \frac{\rho C_p A}{2T_0}\hat{T}^2 - \boldsymbol{\alpha}(\hat{T}) \cdot \hat{\mathbf{v}} - \boldsymbol{\beta}(\hat{T}) \cdot \hat{\mathbf{u}} \end{aligned} \tag{37}$$

where $\hat{T} = T - T_0$, $\hat{\mathbf{v}} = \mathbf{v} - \mathbf{v}_0$, $\hat{\mathbf{u}} = \mathbf{u} - \mathbf{u}_0$, the tensors $\mathbf{K}(\hat{T})$ and $\mathbf{J}(\hat{T})$ are determined in terms of ρ, the elastic moduli $E(\hat{T}), G(\hat{T})$, (all assumed independent of s) and the geometrical shape of the cross-section. The strains $\mathbf{v}_0, \mathbf{u}_0$ describe a static reference configuration at temperature T_0 and u_j and v_j are the components of $\mathbf{u}$ and $\mathbf{v}$ taken in the director basis $\mathbf{d}_j$, i.e., $\mathbf{u} = u_j\,\mathbf{d}_j$ and $\mathbf{v} = v_j\,\mathbf{d}_j$. The coupling of heat to the rod depends ultimately on its atomic structure and dynamic state and the

quantity ρC_p is the macroscopic heat capacity of the material. In this model we have additionally encoded the thermomechanical coupling into the temperature dependent body vectors $\boldsymbol{\alpha}(\hat{T})$ and $\boldsymbol{\beta}(\hat{T})$, These describe the thermal effects of the strains $\mathbf{u}$ and $\mathbf{v}$ in the internal energy. In general these coupling vectors may be expressed $\boldsymbol{\alpha}(\hat{T}) = \alpha_j(\hat{T}))\mathbf{d}_j$ and $\boldsymbol{\beta}(\hat{T}) = \beta_j(\hat{T}))\mathbf{d}_j$. with director components $\alpha_j(\hat{T})$ and $\beta_j(\hat{T})$ $(j = 1, 2, 3)$ assumed independent of s. Such couplings may model Cosserat rods with thermal properties inherited from transverse anisotropic media with chiral (spring-like) characteristics. The extended Kirchhoff constitutive relations now follow from (37) as

$$\begin{cases} \hat{\mathbf{n}}(u_j, v_j, T, \dot{u}_j, \dot{v}_j, T') = \mathbf{K}(\hat{T})(\mathbf{v} - \mathbf{v}_0) - \boldsymbol{\alpha}(\hat{T}) + \hat{\mathbf{n}}^D(u_j, v_j, T, \dot{u}_j, \dot{v}_j, T'), \\ \hat{\mathbf{m}}(u_j, v_j, T, \dot{u}_j, \dot{v}_j, T') = \mathbf{J}(\hat{T})(\mathbf{u} - \mathbf{u}_0) - \boldsymbol{\beta}(\hat{T}) + \hat{\mathbf{m}}^D(u_j, v_j, T, \dot{u}_j, \dot{v}_j, T'). \end{cases} \tag{38}$$

Thus the dynamical equations for our extended Kirchhoff rod with dissipation follow as

$$\begin{cases} \rho A\ddot{\mathbf{r}} = \left(\mathbf{K}(\hat{T})(\mathbf{v} - \mathbf{v}_0) - \boldsymbol{\alpha}(\hat{T}) + \hat{\mathbf{n}}^D\right)' + \mathbf{f} \\ \frac{\partial}{\partial t}(\mathbf{I}(\mathbf{w})) = \left(\mathbf{J}(\hat{T})(\mathbf{u} - \mathbf{u}_0) - \boldsymbol{\beta}(\hat{T}) + \hat{\mathbf{m}}^D\right)' \\ \qquad + \mathbf{v}\left(\mathbf{K}(\hat{T})(\mathbf{v} - \mathbf{v}_0) - \boldsymbol{\alpha}(\hat{T}) + \hat{\mathbf{n}}^D\right) \\ \qquad + \mathbf{l}\frac{\rho C_p A}{T_0}\dot{T}T + T\left((\boldsymbol{\alpha}_T(\hat{T}) \cdot \hat{\mathbf{v}})^{\cdot} + (\boldsymbol{\beta}_T(\hat{T}) \cdot \hat{\mathbf{u}})^{\cdot}\right) - \hat{n}_i^D\dot{v}_i - \hat{m}_i^D\dot{u}_i \\ \qquad = \rho Ah + A\left(\mathcal{K}T'\right)' + \frac{1}{2}\left(T\hat{\mathbf{v}} \cdot \mathbf{K}_T(\hat{T})\hat{\mathbf{v}} + T\hat{\mathbf{u}} \cdot \mathbf{J}_T(\hat{T})\hat{\mathbf{u}}\right)^{\cdot} \end{cases} \tag{39}$$

In linear theories for a homogeneous rod the classical Fourier heat law relating heat flux and temperature assumes $(\mathcal{K})$ is a constant, i.e. $\hat{q}(u_j, v_j, T, T', \dot{u}_j, \dot{v}_j) = k\,T'$, where the constant $k > 0$ is its thermal conductivity. We shall tentatively adopt this correspondence in these extended constitutive relations along with the simplifying assumption that we ex-

plore temperature regimes where $\mathbf{K}$ and $\mathbf{J}$ are independent of T. Furthermore to model mechanically induced dissipation one must adopt functional forms for $\hat{\mathbf{n}}^D, \hat{\mathbf{m}}^D$ compatible with the Clasius-Duhem inequality as discussed above. In this note we set $\hat{\mathbf{n}}^D = \hat{\mathbf{m}}^D = 0$ so that all observed damping can be attributed to thermal dissipation.

4 Analysis of Particular Dynamical Configurations

Consider a rod with $\mathbf{K}$ and $\mathbf{J}$ independent of temperature having both ends fixed and clamped but executing axial and torsional vibrations about an equilibrium state with the rod in a state of axial compression:

$$\mathbf{r}(s,t) = (s + z(s,t))\mathbf{e}_3, \quad s \in [0, \ell] \tag{40}$$

where, relative to a fixed global frame $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ in space

$$\begin{cases} \mathbf{d}_1(s,t) = \cos\varphi(s,t)\mathbf{e}_1 + \sin\varphi(s,t)\mathbf{e}_2 \\ \mathbf{d}_2(s,t) = -\sin\varphi(s,t)\mathbf{e}_1 + \cos\varphi(s,t)\mathbf{e}_2 \\ \mathbf{d}_3(s,t) = \mathbf{e}_3 \end{cases} \tag{41}$$

Then

$$\begin{cases} \mathbf{w}(s,t) = \dot{\varphi}(s,t)\mathbf{e}_3, \quad \mathbf{u}(s,t) = \varphi'(s,t)\mathbf{e}_3, \\ \mathbf{v}(s,t) = \mathbf{r}'(s,t) = (1 + z'(s,t))\mathbf{e}_3 \end{cases} \tag{42}$$

With $\hat{T} \equiv \vartheta$ and the choice $\boldsymbol{\alpha} = \hat{T}\hat{\alpha}_3\mathbf{d}_3$, $\boldsymbol{\beta} = \hat{T}\hat{\beta}_3\mathbf{d}_3$ (with $\hat{\alpha}_3, \hat{\beta}_3$ constant) here, the equations of motion become

$$\begin{cases} \rho A\ddot{z} = EAz'' - \mathcal{A}_3 EA\vartheta' + f_3 \\ I_{33}\ddot{\varphi} = J_{33}\varphi'' - \mathcal{B}_3 J_{33}\vartheta' + l_3 \\ \dfrac{\rho C_p A}{T_0}\dot{\vartheta}(T_0 + \vartheta) + (T_0 + \vartheta)(\hat{\alpha}_3\dot{z}' + \hat{\beta}_3\dot{\varphi}') = \rho A h + kA\vartheta'' \end{cases} \tag{43}$$

In terms of the linear coefficients of expansion $\mathcal{A}_j$ we have $\hat{\alpha}_j = EA\mathcal{A}_j$. We also write $\hat{\beta}_j = J_{jj}\mathcal{B}_j$ (no sum) for $j = 1, 2, 3$, so that for zero heat source h the system above simplifies to

$$\begin{cases} \rho A\ddot{z} - EAz'' + \mathcal{A}_3 EA\vartheta' = f_3, \\ I_{33}\ddot{\varphi} - J_{33}\varphi'' + \mathcal{B}_3 J_{33}\vartheta' = l_3, \\ \dfrac{\rho C_p A}{T_0}\dot{\vartheta} - \dfrac{kA}{T_0+\vartheta}\vartheta'' + \mathcal{A}_3 EA\dot{z}' + \mathcal{B}_3 J_{33}\dot{\varphi}' = 0. \end{cases} \tag{44}$$

The adopted initial conditions are:

$$\begin{cases} z(s,0) = z_0(s), & \dot{z}(s,0) = \dot{z}_0(s), \\ \varphi(s,0) = \varphi_0(s), & \dot{\varphi}(s,0) = \dot{\varphi}_0(s), \quad \vartheta(s,0) = \vartheta_0(s) \end{cases} \tag{45}$$

with boundary conditions

$$z(0,t) = z(\ell,t) = \varphi(0,t) = \varphi(\ell,t) = \vartheta'(0,t) = \vartheta'(\ell,t) = 0 \tag{46}$$

The above system has been analysed numerically and exhibits thermally induced damping of all modes in general. However for a rod at room temperature with the characteristics of copper although such damping is apparent it is small compared with mechanical damping. This is demonstrated in the following numerical analysis of (43) based on a uniform rod of reference length $\ell = 2.0$m and rectangular cross section, 0.10m by 0.04m. The mass density and Young's modulus are respectively $\rho = 2.010^3$[kg/m^3] and $E = 70.010^9$Pa (copper). It will be assumed that $G = \nu E$ with $\nu = 0.35$. The coefficient of linear thermal expansion, the specific heat capacity, and the thermal conductivity are taken as $\mathcal{A}_3 = 23.1\,10^{-6}$ [1/K], $C_p = 904$ [J/K/kg], and $k = 236$ [W/m/K], respectively with reference temperature $T_0 = 293.16$ [K]. When the rod is driven the external sources will be assumed harmonic in time.

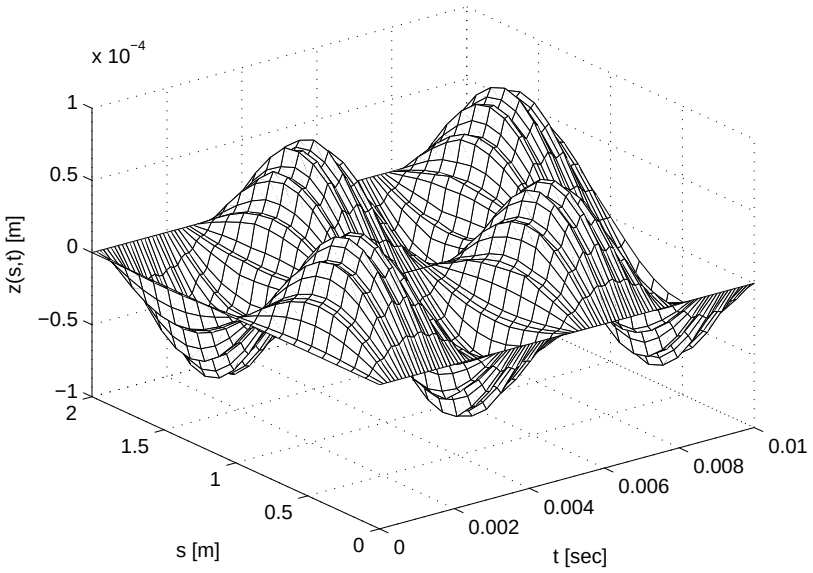


Figure 1: Axial displacement response to harmonic loading. Both ends of the rod are fixed in space and clamped.

The axial and torsional responses of the model described by (43) for a rod (with both ends fixed in space and clamped and with zero initial displacements and velocities) under the external loadings

$$f_3(s,t) = 6000\sin(\pi s)\sin(0.3\omega_0 t)\ \mathrm{N}\cdot\mathrm{m}^{-1}, \qquad \omega_0 = \frac{\pi}{2}\sqrt{\frac{E}{\rho\ell^2}}\ \mathrm{rad}\cdot\mathrm{sec}^{-1},$$
$$l_3(s,t) = 5000\sin(2\pi s)\sin(0.24\omega_0 t)\ \mathrm{N}, \qquad h(s,t) = 0$$

were obtained using Femlab/Matlab and are shown in Figure 1 and Figure 2, respectively.

The behavior of the temperature field of the rod is shown in Figure 3.

The weak thermomechanical coupling necessarily leads to relatively small thermally induced mechanical damping. In order to illustrate the nature of the asymptotic behavior of pure axial oscillations (i.e. φ =constant), we artificially amplify the coefficient of linear thermal expansion to $\mathcal{A}_3 = 46.210^{-6}$ [1/K] and the thermal conductivity to $k = 23610^4$ [W/m/K].

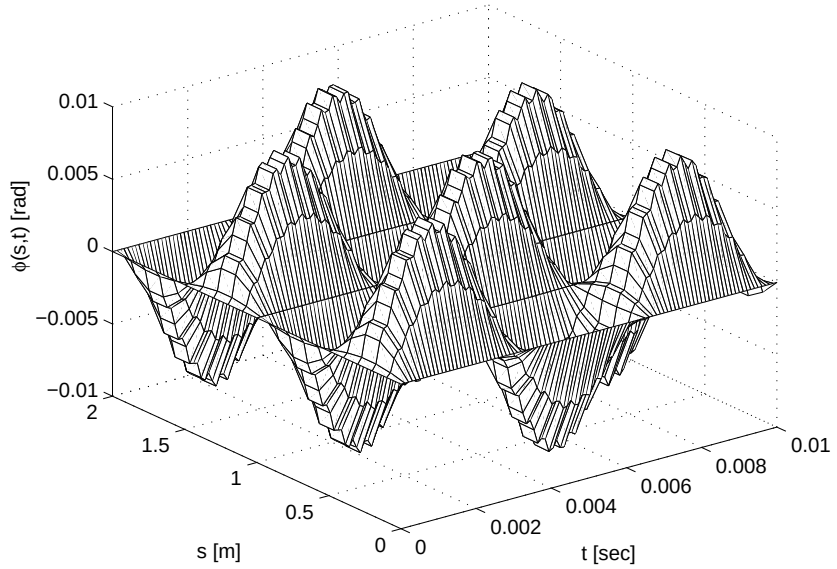


Figure 2: Torsional displacement response of the clamped-clamped rod to harmonic loads.

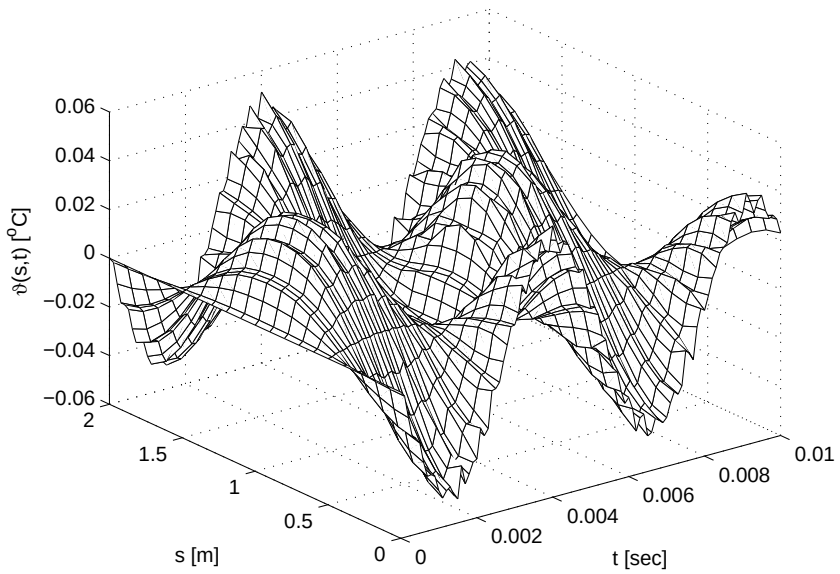


Figure 3: Temperature response of the clamped-clamped rod to harmonic loads.

With zero external force $f_3(s,t) = 0$ and heat source $h(s,t) = 0$, consider a purely axial response governed by

$$\begin{cases} \rho A\ddot{z} - EAz'' + \mathcal{A}_3 EA\vartheta' = 0, \\ \dfrac{\rho C_p A}{T_0}\dot{\vartheta} - \dfrac{kA}{T_0+\vartheta}\vartheta'' + \mathcal{A}_3 EA\dot{z}' = 0. \end{cases} \tag{47}$$

With the initial conditions

$$z(s,0) = 0.04\sin(\pi s), \quad \dot{z}(s,0) = 0, \quad \text{and} \quad \vartheta(s,0) = 0. \tag{48}$$

Figure 4 shows the evolution of axial motion. These and similar simulations indicate that the motion is damped. The associated temperature field (e.g. shown in Figure 5) behaves similarly.

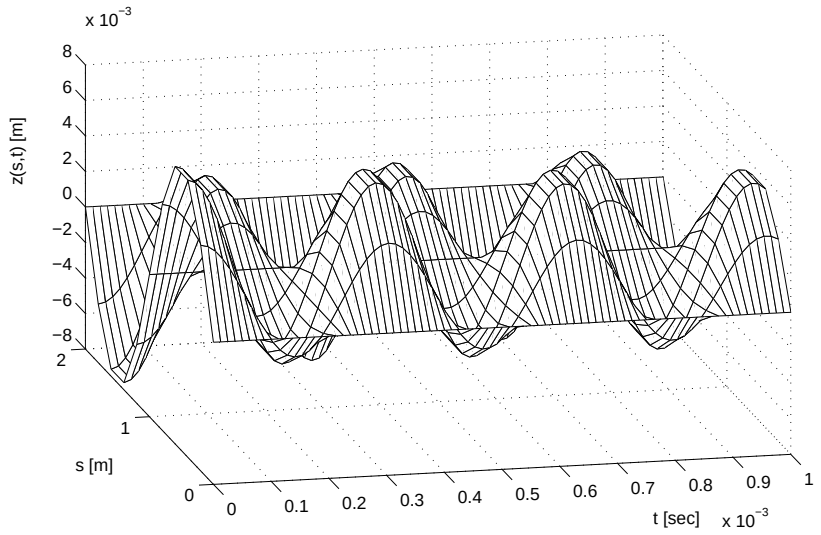


Figure 4: Transient responses of the clamped-clamped rod with initial conditions (48).

5 Summary and Concluding Remarks

The governing equations for a slender heat conducting rod have been derived based on the tenets of simple Cosserat modeling. The contact forces

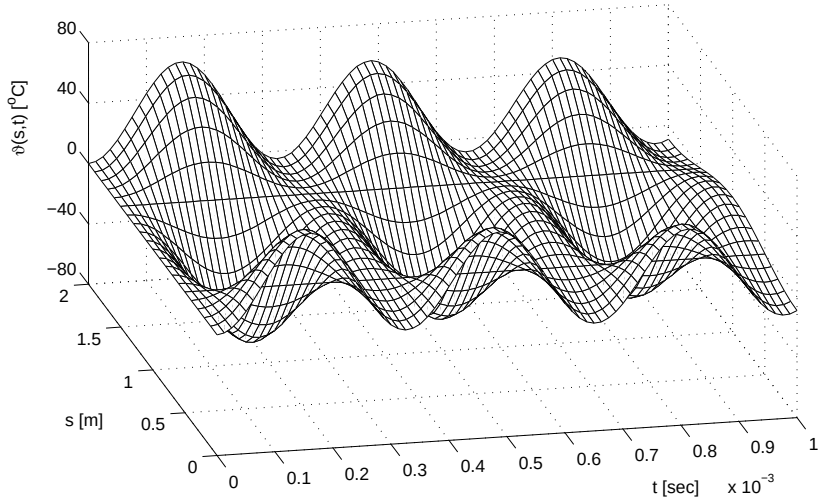


Figure 5: Temperature response of the clamped-clamped rod with initial conditions (48).

and torques, the free energy, entropy density and heat flux are dependent in general on the Cosserat strains arising from the deformation gradient along the rod, the temperature and temperature gradient and the temporal rates of change of the Cosserat strains. The constraints on these general constitutive dependencies, implied by adopting the Clausius-Duhem inequality appropriate for a rod, are made explicit. A large class of constitutive relations compatible with this inequality are discussed and the properties of the governing equations of motion explored for systems where the contact forces and torques are linearly dependent on the strain rates. Thermomechanical modeling of a rod is developed in terms of a free energy that extends the classic Kirchhoff constitutive relations when generalized to include heat conduction. The dissipative behavior of the model is explored by numerical simulation of the evolution of particular dynamic configurations involving axial, torsional vibrations in the presence of internal heat flow. Thermally induced damping is in evidence in these simulations. Although in some situations such effects may be small relative to mechanical viscoelastic and hysteretic damping these studies show that for realistic parameters thermal coupling with heat flow produces effective dissipation

and mechanical damping. There remain challenging problems to ascertain how the simple model here responds to other thermal loadings particularly those involving high frequency and impulsive sources since these are of relevance to many practical application in modern micro-technology.

Acknowledgements The authors are grateful to the EPSRC and a Framework 6 (FP6-2003-NEST-A) award for financial support for this research.

References

[1] B.A. Boley, J. H. Weiner, Theory of Thermal Stresses, Dover Publications, 1997.

[2] A. E. Green and P. M. Naghdi, M.L. Wenner, (1974), *On the theory of rods*, Proc. R. Soc. Lond. A **337** 485-507.

[3] A. E. Green and P. M. Naghdi, (1991), *A thermomechanical theory of a Cosserat point with application to composite materials* Q. J. Mech. Appl. Math. **44** 335-355.

[4] A. E. Green and P. M. Naghdi, (1995), *A unified procedure for construction of theories of deformable media. III. Mixtures of interacting continua*, Proc. R. Soc. Lond. A **448** 379-388.

[5] A.H. Nayfeh and D.T.Mook 1979 *Nonlinear Oscillations*. New York: Wiley-Interscience.

[6] K. T. Andrews, J. R. Fernandez and M. Shillor 2004 *Computational Mechanics* **33**, 225-234. A thermoviscoelastic beam with a tip body.

[7] G. Y. Wu 2005 *Journal of Sound and Vibration* **284**, 343-360. The analysis of dynamic instability and vibration motions of a pinned beam with transverse magnetic fields and thermal loads.

[8] C. V. Massalas and V. K. Kalpakidis 1984 *Letters in Applied Engineering Sciences* **22**, 459-465. Coupled thermoelastic vibrations of a Timoshenko beam.

[9] E. Manoach and P. Ribeiro 2004 *International Journal of Mechanical Sciecnces* **46**, 1589-1606. Coupled, thermoelastic, large amplitude vibrations of Timoshenko beams.

[10] S. K. R. Choudhuri 1972 *AIAA Journal* **10**, 531-533. Thermoelastic stress in a rod due to distributed time-dependent heat sources.

[11] D. Givoli and O. Rand 1995 *AIAA Journal* **33**, 776-778. Dynamic thermoelastic coupling effects in a rod.

[12] S. W. Hansen 1992 *Journal of Mathematical Analysis and Applications* **167**, 429-442. Exponential energy decay in a linear thermoelastic rod.

[13] R. Racke 2002 *Mathematical Mathethods in the Applied Sciences* **25**, 409-441. Thermoelsticity with second sound–exponential stability in linear and non-linear 1-d.

[14] H. Gao 2003 *Mathematical Methods in the Applied Sciences* **26**, 1255-1271. Global attractor for the semilinear thermoelastic problem.

[15] S. W. Hansen 1997 *Journal of Mathematical Analysis and Applications* **210**, 182-205. Boundary control of a linear thermoelastic beam.

[16] M. Grobbelaar-Van Dalsen 2003 *Mathematical Methods in the Applied Sciences* **26**, 1223-1240. Uniform stabilization of a one-dimensional hybrid thermo-elastic structure.

[17] R. Lifshitz and M. L. Roukes 2000 *Physical Review* B **61**(8), 5600-5609. Thermoelastic damping in micro- and nanomechanical systems.

[18] F. L. Guo and G. A. Rogerson 2003 *Mechanics Research Communications* **30**, 513-518. Thermoelastic effect on a micro-machined beam resonator.

[19] A. H. Nayfeh, M. I. Younis and E. M. Abdel-Rahman 2005 *Nonlinear Dynamics* **41**, 211-236. Reduced-order models for MEMS applicatons.

[20] O. Gottlieb and A. R. Champneys 2005 *IUTAM Symposium on Chaotic Dynamics and Control of Systems and Processes in Mechanics, Solid Mechanics and Its Applicaions* **Vol. 122**, 47-57. Global bifurcations of nonlinear thermoelastic microbeams subject to electrodynamic actuation.

[21] S. Lenci and G. Rega 2006 *Journal of Micromechanics and Microengineering* **16** 390-401. Control of pull-in dynamics in a nonlinear thermoelastic electrically actuated microbeam.

[22] T. A. Lauderdale and O. M. O'Reilly 2005 *International Journal of Solids and Structures* **42**, 6523-6549. Modeling MEMS resonators with rod-like components accounting for anisotropy, temperature, and strain dependencies.

[23] A. E. Green and P. M. Naghdi 1979 *International Journal of Solids and Structures* **15**, 829-853. On thermal effects in the theory of rods.

[24] A. E. Green and P. M. Naghdi 1995 *Proceedings of the Royal Society of London* **A 448**, 357-377. A unified procedure for the construction of theories of deformable media. II Generalized continua.

[25] S.S. Antman 1991 *Nonlinear Problems of Elasticity, Applied Mathematical Sciences* **107**. New York: Springer-Verlag.

[26] H.Pecher, 1980 *On Global Solutions of Third Order Partial Differential Equations*, J. of Math. Analysis and Applications, **73**, 278-299.

[27] S. Jiang and R. Racke 2000 *Evolution Equations in Thermoelasticity*. Boca Raton: Chapman & Hall/ CRC.

[28] S Zheng, 1995 *Nonlinear parabolic Equations and hyperbolic-parabolic couples systems*, Longman.

[29] L.Hormander, (1968), *On the characteristic Cauchy problem*, Ann. of Math., **88**, 341-370.

[30] R. Courant, D. Hilbert, *Methods of Mathematical Physics*, vol.II, Wiley, N.Y., 1962.

Author Index

ISBN 3-86644-035-9

www.uvka.de