Tutorials, Schools, and Workshops in the Mathematical Sciences

Maria Ulan
Stanislav Hronek
Editors

Groups, Invariants, Integrals, and Mathematical Physics

The Wisła 20-21 Winter School and Workshop

Birkhäuser

Birkhäuser

Tutorials, Schools, and Workshops in the Mathematical Sciences

This series will serve as a resource for the publication of results and developments presented at summer or winter schools, workshops, tutorials, and seminars. Written in an accessible style, they present important and emerging topics in scientific research for PhD students and researchers. Filling a gap between traditional lecture notes, proceedings, and standard textbooks, the titles included in this series present material from the forefront of research.

Manuscripts are solicited by the editorial boards of each volume and then reviewed by a minimum of three peer reviewers to ensure the highest standards of scientific literature.

Baltic Institute of Mathematics Collection

Since 2012, the Baltic Institute of Mathematics has been organizing scientific events such as conferences, workshops, seminars, and schools. The main goal of these events is to facilitate networking between young and senior researchers to discuss the state-of-the-art and outstanding new results in mathematics, physics, and computer science, and their use in industry. The volumes in this collection are based on the Institute's events and are written in an accessible, pedagogical style for graduate students and early-career researchers.

Maria Ulan • Stanislav Hronek
Editors

Groups, Invariants, Integrals, and Mathematical Physics

The Wisła 20-21 Winter School and Workshop

Birkhäuser

Editors
Maria Ulan
Baltic Institute of Mathematics
Warszawa, Poland

Stanislav Hronek
Theoretical Physics and Astrophysics
Masaryk University
Brno, Czech Republic

ISSN 2522-0969 ISSN 2522-0977 (electronic)
Tutorials, Schools, and Workshops in the Mathematical Sciences
ISBN 978-3-031-25665-3 ISBN 978-3-031-25666-0 (eBook)
https://doi.org/10.1007/978-3-031-25666-0

This book is published under the imprint Birkhäuser, www.birkhauser-science.com by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

COVID-19 pandemic has driven most researchers online. Staying connected with community has been really important during the lockdown. *The Wisła 20-21 Winter School & Workshop: Groups, Invariants, Integrals, and Mathematical Physics* was organized online by the Baltic Institute of Mathematics. Even though it was a virtual event, participants were given an opportunity to interact with their colleagues and well-known researchers in the field. This book is a summary of selected and carefully reviewed lecture notes and contributions. The reader is expected to have some basic knowledge of differential geometry and category theory.

The school was devoted to differential invariants, moving frames, and Poisson algebras. There were four series of main lectures, given by Valentin Lychagin, Eivind Schneider, Peter J. Olver, and Vladimir Roubtsov, respectively:

- Differential contra algebraic invariants
- Differential invariants of Lie pseudogroups
- The Theory and Applications of Moving Frames
- Poisson algebras

It is our pleasure to share these lectures, given by experts in their fields, with an audience who were not fortunate to participate. Chapter “Differential Invariants in Algebra” presents lecture notes on differential invariants with a focus on Lie groups, pseudogroups, and their orbit spaces. Poisson structures in algebra and in geometry are discussed in chapter “Lectures on Poisson Algebras”. There are many research papers and lecture notes on moving frames, we refer to the Peter Olver’s webpage[1] as a good source for the interested reader.

The workshop was focused on the intersection of differential geometry, differential equations, and category theory. Contributions are written in a pedagogical style while simultaneously bringing to attention recent advances made by their authors. Chapter “Some Remarks on Multisymplectic and Variational Nature of Monge-Ampère Equations in Dimension Four” is focused on multisymplectic and

[1] https://www-users.cse.umn.edu/~olver/.

variational nature of Monge-Ampère equations in dimension four. The problem of the integrability of fifth-order equations admitting a Lie symmetry algebra is addressed in chapter "Generalized Solvable Structures Associated to Symmetry Algebras Isomorphic to $\mathfrak{gl}(2,\mathbb{R}) \ltimes \mathbb{R}$". Applications of van Kampen theorem for groupoids to computation of homotopy types of striped surfaces are discussed in chapter "Fundamental Groupoids and Homotopy Types of Non-compact Surfaces". Finally, chapter "A Geometric Framework to Compare Classical Field Theories" presents a geometric framework to compare classical systems of PDEs in the category of smooth manifolds.

We hope that this book will give you, dear reader, a good entry point, and that it will aid with motivation and competence to dive deeper into the world of differential geometry and mathematical physics.

Warszawa, Poland — Maria Ulan
Brno, Czech Republic — Stanislav Hronek
January 2022

Acknowledgements

Baltic Institute of Mathematics would like to thank the keynote speakers Prof. Valentin Lychagin and Prof. Peter J. Olver for their inspiring lectures, and Prof. Olav Arnfinn Laudal for kindliness, expertise, and continuous support. Prof. Jerzy Buzek, MEP deserves a special thank you for the honorary patronage of the Futurum 2020s Initiative that helps to organize and provide excellent meeting facilities during the Wisła 20-21 Winter School & Workshop, even despite the coronavirus pandemic.

The organizers would like to thank speakers of the lectures and all participants for a great, friendly atmosphere and strong motivation to learn. Extra acknowledgments go to the authors for contributing and presenting their research.

The editors would like to thank the reviewers for valuable comments and suggestions on preliminary drafts of the book, and the Birkhäuser Mathematics, Springer Nature crew, especially Chris Tominich for assistance.

Contents

Contributors

Lukas Silvester Barth Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany

Valentin Lychagin Department of Mathematics and Statistics, UiT the Arctic University of Norway, Tromsø, Norway

Sergiy Maksymenko Institute of Mathematics of NAS of Ukraine, Kyiv, Ukraine

Concepción Muriel Department of Mathematics, University of Cádiz, Campus Río San Pedro, Cádiz, Spain

Oleksii Nikitchenko Kyiv Academic University, Kyiv, Ukraine

Michael Roop Baltic Institute of Mathematics, Warsaw, Poland

Vladimir Rubtsov LAREMA UMR 6093, CNRS and Université d'Angers, Angers Cedex, France

IGAP (Institute of Geometry and Physics), Trieste, Italy

Adrián Ruiz Department of Mathematics, University of Cádiz, Campus Río San Pedro, Cádiz, Spain

Radek Suchánek Department of Mathematics and Statistics, Masaryk University, Brno, Czech Republic

LAREMA UMR 6093, CNRS, Université d'Angers, Angers Cedex, France

Differential Invariants in Algebra

Valentin Lychagin and Michael Roop

Abstract In these lectures, we discuss two approaches to studying orbit spaces of algebraic Lie groups. Due to algebraic approach orbit space, or quotient, is an algebraic manifold, while from the differential viewpoint a quotient is a differential equation. The main goal of these lectures is to show that the differential approach gives us a better understanding of structure of invariants and orbit spaces. We illustrate this on classical equivalence problems, such as SL—classification of binary and ternary forms, and affine classification of algebraic plane curves.

1 Introduction

The concept of an invariant appears whenever it comes to any kind of a classification problem. In these lectures, we would like to explain basic concepts of the invariant theory and show its applications to algebraic problems, such as SL-classification of binary and ternary forms, and affine classification of algebraic plane curves. It seems helpful to us to recommend books [1, 2] and references therein to the interested reader.

The origin of the invariant theory goes back to the middle of the nineteenth century and has not only mathematical motivation, such as affine classification of quadratic forms, finding canonical forms for equations of conics and quadrics, obtained in works of Euler, Lagrange, Cauchy, Gauss, but also a physical one (finding principal axes of inertia, investigation of planets' motion).

The first results on SL-classification of binary forms go back to 1841 and belong to Boole, who observed that discriminants of binary forms are invariant under linear

V. Lychagin
Department of Mathematics and Statistics, UiT the Arctic University of Norway, Tromsø, Norway
e-mail: vly000@post.uissst.no

M. Roop (✉)
Baltic Institute of Mathematics, Warsaw, Poland
e-mail: m.roop@baltinmat.eu

M. Ulan, S. Hronek (eds.), *Groups, Invariants, Integrals, and Mathematical Physics*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-031-25666-0_1

transformations with determinant equal to 1. Later, in 1845, Cayley constructed invariants using the technique of hyperdeterminants developed by Cayley himself [3, 4]. In 1849, Aronhold provided a systematic study of ternary forms of degree 3, and 2 years later he gave a general formulation of invariant theory for algebraic forms. He also obtained differential equations for invariants of algebraic forms, that were also obtained by Cayley for binary forms in 1852, which led to a series of works [5–8] known as memoirs upon quantics.

In 1863, Aronhold observed that the number of rationally independent absolute invariants equals the difference between the number of coefficients of the form and the number of coefficients in a linear transformation (in modern terms, the difference between the dimension of the space of forms and the dimension of the group) [9]. In 1861, Clebsch, using results of Aronhold, developed symbolic methods of finding invariants of algebraic forms [10]. These methods were later developed by Gordan and rapidly became popular.

In 1856, Cayley and Sylvester showed that binary forms of degrees up to four have a finite number of so-called *irreducible covariants*. Covariant is a polynomial in x, y, and coefficients of the form, invariant under the transformations of the group (e.g. of SL_2 transformations). Irreducibility means that such covariants cannot be expressed as rational functions of covariants of lower degree [11]. This became the origin of the finiteness problem for generating set of invariants.

Gordan was the first who proved the finiteness of a number of covariants for the binary form of arbitrary degree (Gordan's theorem) [12], and his method allowed to construct a complete system of irreducible covariants for binary forms of degrees 5 and 6. Later, Sylvester discovered the same result for the case of a binary form of degree 12. In 1880, von Gall constructed a complete system of covariants for a binary form of degree 8, and 8 years later for that of degree 7, which turned out to be more complicated than the case of degree 8 [13, 14]. Binary forms of degree 7 were also elaborated by Dixmier and Lazard [15]. Hammond provided the proof for the case of binary seventhics [16].

Finally, in 1890, Hilbert gave a complete proof of Gordan's result for the case of arbitrary n-ary forms of an arbitrary degree [17].

While solving the problem of constructing a complete system of irreducible invariants and covariants, the very notion of an *invariant* was changing. The theory of *differential invariants* was developed by Halphen in 1878 in his thesis [18] and was later generalized by Norwegian mathematician Sophus Lie, who showed that all previous results of invariant theory are particular cases of more general theory of invariants of continuous transformation groups [19, 20]. Lie did not use symbolic methods of Aronhold and Clebsch, that hardly could be extended to the cases of binary forms of higher degrees due to their dramatic bulkiness.

In the context of modern invariant theory and simultaneously in the context of these lectures, it is worth mentioning such results as Rosenlicht [21] and global Lie-Tresse theorems [22], that justified the appearance of rational differential invariants in classification problems and paved a way for solving algebraic equivalence problems using differential-geometric techniques [23, 24]. This will be the core point of the present lectures.

The paper is organized as follows. In Sect. 2, we start with $\mathrm{SL}_2(\mathbb{C})$ classification of binary forms and explain how to get rational differential invariants using the observation that binary forms are solutions of the Euler equation. In Sect. 3, we give a general introduction to modern invariant theory together with discussion of Rosenlicht and Lie-Tresse theorems and explanation how the last can be used to find smooth solutions to PDEs, as well as those with singularities. Sect. 4 is devoted to affine classification of algebraic plane curves. The last Sect. 5 concerns the problem of $\mathrm{SL}_3(\mathbb{C})$-classification of ternary forms using results obtained in the previous sections.

All essential computations for this paper were performed in Maple with the DifferentialGeometry package created by I. Anderson and his team [25], and the first author is grateful to him for the very first introduction to the package.

2 Invariants of Binary Forms

In this section, we study SL_2—invariants of binary n—forms. We show the difference between algebraic and differential approaches and the power of differential one in finding invariants.

2.1 *Algebraic Point of View*

Binary form of degree n is a homogeneous polynomial on $\mathbb{C}^2$

$$\phi_b = \sum_{i=0}^{n} b_{i,n-i} \frac{x^i}{i!} \frac{y^{n-i}}{(n-i)!}, \quad b_{i,n-i} \in \mathbb{C}. \tag{1}$$

The space of all binary forms of degree n is $\mathcal{B}_n \simeq \mathbb{C}^{n+1}$. The action of the Lie group

$$\mathrm{SL}_2(\mathbb{C}) = \{A \in \mathrm{Mat}_{2\times 2}(\mathbb{C}) \mid \det(A) = 1\}$$

on $\mathcal{B}_n$ is defined by the following way:

$$A\colon \mathcal{B}_n \ni \phi_b \mapsto A\phi_b = \phi_b \circ A^{-1} \in \mathcal{B}_n. \tag{2}$$

This action induces the action on coefficients $b_{i,n-i}$. Due to algebraic approach, where we believe that the quotient is an algebraic manifold, to describe the quotient space $\mathcal{B}_n/\mathrm{SL}_2(\mathbb{C})$ one needs to find polynomials $I(b) = I(b_{0,n}, \ldots, b_{n,0})$ invariant under the action (2). Such functions are called *algebraic invariants*.

Theorem 1.1 (Gordan-Hilbert, [12, 17]) *The algebra of polynomial* SL_2—*invariants of binary* n*-forms is finitely generated, and the quotient space is an affine, algebraic manifold.*

However, the problem of finding generators of this algebra and syzygies in this algebra turned out to be specific for every n. For instance, the case of $n = 3$ was elaborated by Bool in 1841, who observed that the discriminant of the cubic is an invariant. This became the origin of the classical invariant theory. Results regarding the case of $n = 4$ belong to Bool, Cayley and Eisinsteine (1840–1850) [3, 4, 26, 27]. For quintic ($n = 5$), the invariants were found by Sylvester and Hilbert (see, for example, [26, 27]). They are dramatically huge to write down explicitly, the invariant of degree 18 found by Hermite contains 848 terms! The main problem is that there is no general approach in the classical invariant theory. This motivates us to develop a differential approach [23, 24].

2.2 *Differential Point of View*

The key idea underlying the differential approach is to identify $\mathcal{B}_n$ with the space of smooth solutions to Euler equation

$$x f_x + y f_y = n f. \tag{3}$$

It is worth mentioning that class of solutions to (3) includes not only binary n-forms, but also other homogeneous functions of degree n. Thus, solving the problem for all solutions to (3) we at the same time solve the problem of SL_2-equivalence of binary forms.

Equation (3) defines a smooth submanifold $\mathcal{E}_1$ in the space of 1-jets $\mathbf{J}^1 = J^1\left(\mathbb{C}^2\right)$ of functions on $\mathbb{C}^2$:

$$\mathcal{E}_1 = \{x u_{10} + y u_{01} = n u_{00}\} \subset \mathbf{J}^1.$$

Solutions of (3) are special type surfaces $L_f \subset \mathcal{E}_1$

$$L_f = \left\{u_{00} = f(x, y),\ u_{10} = f_x,\ u_{01} = f_y\right\} \subset \mathcal{E}_1.$$

It is often reasonable to consider not only Eq. (3), but also a collection of its differential consequences up to some order k, i.e. a prolongation $\mathcal{E}_k \subset \mathbf{J}^k$. The space $\mathbf{J}^k$ is a space of k-jets of smooth functions on $\mathbb{C}^2$:

$$\mathbf{J}^k = \left\{[f]_p^k \mid p \in \mathbb{C}^2,\ f \in C^\infty\left(\mathbb{C}^2\right)\right\},$$

where $[f]_p^k$ is the equivalence class of functions, whose Taylor polynomials of the length k at the point $p \in \mathbb{C}^2$ are the same (values and all derivatives up to order k at the point p coincide). The space of k-jets is equipped with canonical coordinates $(x, y, u_{00}, \ldots, u_{ij}, \ldots)$, $0 \le i + j \le k$, $\dim\left(\mathbf{J}^k\right) = \binom{k+2}{2} + 2$, and

$$u_{ij}\left([f]_p^k\right) = \frac{\partial^{i+j} f}{\partial x^i \partial y^j}(p).$$

The action $A\colon \mathbb{C}^2 \to \mathbb{C}^2$ of the group SL_2 can be prolonged to $\mathbf{J}^k$ by the natural way

$$A^{(k)}\colon \mathbf{J}^k \to \mathbf{J}^k, \quad A^{(k)}\left([f]_p^k\right) = [Af]_{Ap}^k.$$

Moreover, if

$$L_f^{(k)} = \left\{u_{ij} = \frac{\partial^{i+j} f}{\partial x^i \partial y^j},\ 0 \le i + j \le k\right\}$$

is a graph of the k-jet of function f, then

$$A^{(k)}\left(L_f^{(k)}\right) = L_{Af}^{(k)}.$$

Let us now put $k = n$ and let $\mathcal{E}_n \subset \mathbf{J}^n$ be the $(n-1)$-prolongation of the Euler equation together with $u_{ij} = 0$:

$$\mathcal{E}_n = \left\{ \frac{d^{k+l}}{dx^k dy^l}\left(xu_{10} + yu_{01} - nu_{00}\right) = 0,\ 0 \le k + l \le n - 1,\ u_{ij} = 0, \right.$$
$$\left. n + 1 \le i + j \right\}.$$

One can show that $\dim \mathcal{E}_n = n + 3$. The prolongations $A^{(n)}$ of group elements $A \in \mathrm{SL}_2$ preserve the submanifold $\mathcal{E}_n$ and therefore define the action $A^{(n)}\colon \mathcal{E}_n \to \mathcal{E}_n$. Since $L_\phi^{(n)} \subset \mathcal{E}_n$, any binary n-form can be considered as a solution to $\mathcal{E}_n$. The property $A^{(n)}\left(L_\phi^{(n)}\right) = L_{A\phi}^{(n)}$ shows that the group $\mathrm{SL}_2(\mathbb{C}^2)$ is a symmetry group of the Euler equation.

A rational function $I \in C^\infty(\mathcal{E}^k)$ is said to be a *rational differential* SL_2*-invariant of order* k, or simply *differential invariant*, if $I \circ A^{(k)} = I$, for all $A \in \mathrm{SL}_2(\mathbb{C})$.

As we shall see further, the Lie-Tresse theorem states that the algebra of rational differential SL_2-invariants of order $\le n$ on the Euler equation $\mathcal{E}_n$ gives us realization of the quotient $\mathcal{E}_n/\mathrm{SL}_2(\mathbb{C})$ as a new differential equation of order 3, and $\mathrm{SL}_2(\mathbb{C})$-orbits of binary n-forms correspond to solutions of this equation.

The following observations will be important for us.

- the plane $\mathbb{C}^2$ is the affine space, i.e. a space with the standard translation of vectors (trivial connection) and distinguished point $\mathbf{0}$
- the plane $\mathbb{C}^2$ is the symplectic space, equipped with the structure form $\Omega = dx \wedge dy$
- the group $\mathrm{SL}_2(\mathbb{C})$ preserves these both affine and symplectic structures, and the point $\mathbf{0}$.

As we shall see further, these structures will allow us to equip the set of differential $\mathrm{SL}_2(\mathbb{C})$-invariants with additional structures and will give us explicit methods of finding invariants.

2.3 Relations Between Algebraic and Differential Invariants

One can easily see that due to (1)

$$b_{i,n-i} = \frac{\partial^n \phi_b}{\partial x^i \partial y^{n-i}}.$$

Therefore, the function $I(b_{n,0}, \dots, b_{0,n})$ is an $\mathrm{SL}_2(\mathbb{C})$-invariant if and only if $I(u_{n0}, \dots, u_{0n})$ is a differential $\mathrm{SL}_2(\mathbb{C})$-invariant of order n. Thus, algebraic $\mathrm{SL}_2(\mathbb{C})$-invariants of binary n-forms are differential invariants of the form $I(u_{0n}, \dots, u_{n0})$ and finding differential invariants we simultaneously find also algebraic ones.

2.4 Lie Equation

Since the Lie group $\mathrm{SL}_2(C)$ is connected, the condition $I \circ A^{(k)} = I$ can be written in an infinitesimal form:

$$X^{(k)}(I) = 0, \quad X \in \mathfrak{sl}_2, \tag{4}$$

where $X^{(k)}$ is the kth prolongation of the vector field $X \in \mathfrak{sl}_2$, and Eq. (4) is called *Lie equation*. The Lie algebra $\mathfrak{sl}_2$ is generated by vector fields

$$\mathfrak{sl}_2 = \langle X_+ = x\partial_y,\ X_- = y\partial_x,\ X_0 = x\partial_x - y\partial_y \rangle$$

with commutators

$$[X_+, X_-] = X_0, \quad [X_0, X_+] = 2X_+, \quad [X_0, X_-] = -2X_-. \tag{5}$$

Due to Lie algebra structure (5), condition $X_0^{(k)}(I) = 0$ is not independent, and Lie equation (4) becomes

$$X_+^{(k)}(I) = 0, \quad X_-^{(k)}(I) = 0.$$

This equation also appeared in Hilbert's lectures [26].

Following some empirical observations, according to which the number of functionally independent invariants equals the codimension of the regular orbit (we shall explain this strictly by means of the Rosenlicht theorem in the forthcoming sections), let us now compute the numbers of functionally independent algebraic and differential invariants.

Since

$$\dim(\mathbf{J}^k) = \frac{(k+1)(k+2)}{2} + 2,$$

the number of independent differential invariants of kth order on $\mathbf{J}^k$ equals

$$\dim(\mathbf{J}^k) - \dim(\mathfrak{sl}_2) = \frac{k(k+3)}{2}.$$

Since $\dim(\mathcal{E}_n) = n+3$, the number of differential invariants of binary n-forms equals $\dim(\mathcal{E}_n) - 3 = n$, and the number of independent algebraic invariants of binary n-forms equals $\dim(\mathbb{C}^{n+1}) - 3 = n+1-3 = n-2$.

This discussion is true for the case $n \geq 3$, when the Lie algebra of the stabilizer of the form is trivial. In the case $n = 2$ its dimension equals 1, and therefore there is only one invariant in this case, which is the discriminant.

2.5 Resultants and Discriminants

Here, we will repeat the Boole's result on the SL_2-invariance of the discriminant of binary forms.

Any binary n-form can be represented as a product of linear functions I_i^ϕ, $i = 1, \ldots, n$:

$$\phi = \prod_{i=1}^{n} I_i^\phi.$$

Obviously, functions I_i^ϕ are defined up to multipliers λ_i: $I_i^\phi \mapsto \lambda_i I_i^\phi$, where $\prod_{i=1}^{n} \lambda_i = 1$. Let $\psi \in \mathcal{B}_n$ be another binary form, $\psi = \prod_{i=1}^{m} I_i^\psi$. Then, one can define *resultant* between forms ϕ and ψ by the following way:

$$\mathrm{Res}(\phi, \psi) = \prod_{i,j} [I_i^{\phi}, I_j^{\psi}],$$

where $[I_i^{\phi}, I_j^{\psi}]$ is the Poisson bracket associated with the symplectic form $\Omega = dx \wedge dy$.

The function

$$\mathrm{Discr}(\phi) = \mathrm{Res}(\phi_x, \phi_y),$$

is called *discriminant*.

Remark that here (x, y) are canonical coordinates of the vector space $\mathbb{C}^2$, i.e. $\Omega = dx \wedge dy$ in these coordinates.

Let us collect basic properties of discriminants and resultants.

1. $\mathrm{Res}(\phi, \psi)$ does not depend on scalings $I_i^{\phi} \mapsto \alpha_i I_i^{\phi}$, $I_i^{\psi} \mapsto \beta_i I_i^{\psi}$
2. $\mathrm{Res}(\phi, \psi)$ is a polynomial in coefficients of ϕ, ψ of degree $(n + m)$
3. $\mathrm{Res}(\phi, \psi)$ is an $\mathrm{SL}_2(\mathbb{C})$-invariant: $\mathrm{Res}(A\phi, A\psi) = \mathrm{Res}(\phi, \psi)$
4. $\mathrm{Discr}(\phi)$ is a polynomial $\mathrm{SL}_2(\mathbb{C})$-invariant of degree $(2n - 2)$.

Using discriminants and resultants one gets algebraic invariants from differential ones.

Example Consider the following binary form of degree 3:

$$\phi_3(x, y) = x^3 + a_1 x^2 y + a_2 x y^2 + a_3 y^3 \tag{6}$$

1. The discriminant $\mathrm{Discr}(\phi)$ of cubic (6)

$$J_1 = \mathrm{Discr}(\phi) = 12a_1^3 a_3 - 3a_1^2 a_2^2 - 54a_1 a_2 a_3 + 12a_2^3 + 81a_3^3$$

   is a polynomial $\mathrm{SL}_2(\mathbb{C})$-invariant of order 4. This illustrates the property 4.
2. Let us take the differential SL_2-invariant $u_{20}u_{02} - u_{11}^2$ and restrict it on the cubic (6). We get the following quadric

$$\phi_2(x, y) = 4(3a_2 - a_1^2)x^2 + 4(9a_3 - a_1 a_2)xy + 4(3a_1 a_3 - a_2^2)y^2.$$

   Taking its discriminant, we get the polynomial invariant $J_2 = -16J_1$. This illustrates how one can get polynomial invariants from differential ones.

2.6 *Operations and Structures on Invariants*

2.6.1 Monoid Structure

Any function $\phi \in C^\infty(\mathbf{J}^k)$ generates a differential operator by the following way:

$$\widehat{\phi}\colon C^\infty(\mathbb{C}^2) \to C^\infty(\mathbb{C}^2),$$

or in coordinates

$$\widehat{\phi}\colon f(x,y) \mapsto \phi\left(x, y, f, f_x, f_y, \ldots\right),$$

if $\phi = \phi(x, y, u_{00}, u_{10}, u_{01}, \ldots)$. Then, condition for ϕ to be an $\mathrm{SL}_2(\mathbb{C})$-invariant reads

$$A \circ \widehat{\phi} = \widehat{\phi} \circ A, \quad A \in \mathrm{SL}_2(\mathbb{C}).$$

Now we can introduce an operation $*$ of composition for invariants by the following way:

$$\widehat{\phi * \psi} = \widehat{\phi} \circ \widehat{\psi}.$$

Exams

$$u_{00} * \psi = \psi, \quad u_{10} * \psi = \frac{d\psi}{dx}, \quad u_{01} * \psi = \frac{d\psi}{dy}, \quad u_{ij} * \psi = \frac{d^{i+j}\psi}{dx^i dy^j},$$

$$(u_{20}u_{02} - u_{11}^2) * \psi = \frac{d^2\psi}{dx^2}\frac{d^2\psi}{dy^2} - \left(\frac{d^2\psi}{dxdy}\right)^2,$$

where

$$\frac{d}{dx} = \frac{\partial}{\partial x} + \sum_{i,j=0} u_{i+1,j}\frac{\partial}{\partial u_{ij}}, \quad \frac{d}{dy} = \frac{\partial}{\partial y} + \sum_{i,j=0} u_{i,j+1}\frac{\partial}{\partial u_{ij}}$$

are total derivatives.

Note that the composition of differential invariants of orders k and l is a differential invariant of order $(k+l)$, and composition with u_{00} gives us the same invariant. This means that the composition operation endows the set of differential $\mathrm{SL}_2(\mathbb{C})$-invariants with a monoid structure.

Theorem 1.2 *The set of differential* $\mathrm{SL}_2(\mathbb{C})$*-invariants is a monoid with unit* u_{00}.

Example The differential $\mathrm{SL}_2(\mathbb{C})$-invariants of order 1 are

$$\phi = F(u_{00}, xu_{10} + yu_{01}).$$

Let ψ be another invariant of order k. Then,

$$\phi * \psi = F\left(\psi, x\frac{d\psi}{dx} + y\frac{d\psi}{dy}\right)$$

is a differential invariant of order $(k+1)$.

2.6.2 Poisson Structure

Recall that the symplectic form $\Omega = dx \wedge dy$ is SL_2-invariant. Define the Poisson bracket for functions on jet spaces by the following way:

$$\widehat{d}\phi \wedge \widehat{d}\psi = [\phi, \psi]\Omega,$$

where $\widehat{d}f = \frac{df}{dx}dx + \frac{df}{dy}dy$ is the total differential, $f \in C^\infty(\mathbf{J}^k)$. As we shall see below, $\widehat{d}$ is an invariant operator. Then, we get

$$[\phi, \psi] = \frac{d\phi}{dx}\frac{d\psi}{dy} - \frac{d\phi}{dy}\frac{d\psi}{dx},$$

and if ϕ and ψ are differential SL_2-invariants, then $[\phi, \psi]$ is a differential invariant too.

Theorem 1.3 *The algebra of* SL_2*-invariants is a Poisson algebra.*

Example Let us take two differential $\mathrm{SL}_2(\mathbb{C})$-invariants: $J_1 = u_{00}$ and $J_2 = u_{20}u_{02} - u_{11}^2$. Taking the Poisson bracket between them we get a differential $\mathrm{SL}_2(\mathbb{C})$-invariant of the third order:

$$J_3 = [J_1, J_2] = u_{01}(2u_{11}u_{21} - u_{02}u_{30} - u_{20}u_{12}) + u_{10}(u_{02}u_{21} + u_{20}u_{03} - 2u_{11}u_{12}).$$

As en exercise, we propose to check it to the reader.

2.6.3 Invariant Frame

Taking the kth term in the Taylor decomposition of a function $f(x, y)$, we get symmetric differential forms

$$d_k f = \sum_{i=0}^{k} \frac{\partial^k f}{\partial x^i \partial y^{k-i}} \frac{dx^i}{i!} \frac{dy^{k-i}}{(k-i)!}, \quad k = 1, 2, \ldots$$

We shall see later on that these tensors are defined by the affine connection, which is in our case the trivial connection. Therefore, they are invariants of the affine transformations, i.e.

$$d_k(Af) = A(d_k f), \quad A \in \mathrm{SL}_2(\mathbb{C}).$$

Let us define tensors Θ_k on jet spaces by the following way:

$$\Theta_k = \sum_{i=0}^{k} u_{i,k-i} \frac{dx^i}{i!} \frac{dy^{k-i}}{(k-i)!}.$$

Then, $d_k f = \Theta_k|_{L_f^k}$, and Θ_k are SL_2-invariants.

On the space $\mathbf{J}^2$ we have the following SL_2-invariant tensors:

$$\begin{aligned}
\Theta_1 &= u_{10} dx + u_{01} dy, \\
\Theta_2 &= u_{20} \frac{dx^2}{2} + u_{11} dx dy + u_{02} \frac{dy^2}{2}, \\
\Omega &= dx \wedge dy.
\end{aligned}$$

As we shall see further, the Lie-Tresse theorem states that the algebra of differential invariants is a differential algebra, and we now turn the algebra of invariants into the differential algebra by introducing the invariant derivations

$$\nabla_i = A_i \frac{d}{dx} + B_i \frac{d}{dy}, \quad i = 1, 2,$$

where A_i and B_i are functions on $\mathbf{J}^2$, satisfying the conditions:

$$\nabla_1 \rfloor \Omega = \Theta_1, \quad \nabla_2 \rfloor \Theta_2 = \Theta_1.$$

Direct computations give us the following result:

$$\nabla_1 = u_{01} \frac{d}{dx} - u_{10} \frac{d}{dy}, \tag{7}$$

$$\nabla_2 = \frac{2(u_{02}u_{10} - u_{11}u_{01})}{\Delta_2}\frac{d}{dx} + \frac{2(u_{20}u_{01} - u_{11}u_{10})}{\Delta_2}\frac{d}{dy}, \tag{8}$$

where $\Delta_2 = u_{20}u_{02} - u_{11}^2$.

Their bracket is

$$[\nabla_1, \nabla_2] = A\nabla_1 + B\nabla_2,$$

where A and B are differential SL_2-invariants of order 3, and

$$A|_{\mathcal{E}_3} = \frac{2(2-n)}{n-1}, \quad B|_{\mathcal{E}_3} = 0.$$

Theorem 1.4 *Let ϕ be a differential SL_2-invariant of order $\leq k$. Then, $\nabla_1(\phi)$ and $\nabla_2(\phi)$ are differential SL_2-invariants of order $\leq k+1$.*

This means that the algebra of differential SL_2-invariants equipped with invariant derivations ∇_1 and ∇_2 becomes a differential algebra. Summarizing all above discussion, we have:

Theorem 1.5 *The algebra of differential SL_2-invariants is a*

- *monoid with unit u_{00}*
- *Poisson algebra*
- *differential algebra*

We can see that the differential viewpoint allows us to endow the set of invariants with much more interesting structures comparing with those we had in the algebraic situation.

2.7 Invariant Coframe

Let us now construct the dual frame $\langle\omega_1, \omega_2\rangle$, which is an SL_2-invariant coframe, where $\omega_i = a_i dx + b_i dy$ and coefficients a_i, b_i are such that $\omega_i(\nabla_j) = \delta_{ij}$.

Simple computations give us

$$\omega_1 = \frac{u_{20}u_{01} - u_{11}u_{10}}{J_{21}}dx - \frac{u_{02}u_{10} - u_{11}u_{01}}{J_{21}}dy,$$

$$\omega_2 = \frac{\Delta_2}{2J_{21}}(u_{10}dx + u_{01}dy),$$

where

$$J_{21} = u_{01}^2 u_{20} - 2u_{10}u_{01}u_{11} + u_{10}^2 u_{02}$$

is an SL_2-invariant of order 2, called *flex invariant* [28].

The original coframe $\langle dx, dy\rangle$ is expressed in terms of $\langle \omega_1, \omega_2\rangle$ as

$$dx = u_{01}\omega_1 + \frac{2(u_{02}u_{10} - u_{11}u_{01})}{\Delta_2}\omega_2,$$
$$dy = -u_{10}\omega_1 + \frac{2(u_{20}u_{01} - u_{11}u_{10})}{\Delta_2}\omega_2.$$

And finally we are able to write down the invariant tensors Θ_k in the form

$$\Theta_k = \sum_{i=0}^{k} I_{i,k-i}\frac{\omega_1^i\omega_2^{k-i}}{i!(k-i)!}.$$

Since Θ_k are invariants, $\omega_{1,2}$ are invariants, we get:

Theorem 1.6 *Functions $I_{i,j}$ are SL_2-invariants of order $(i+j)$, and any rational differential invariant is a rational function of them.*

Exams

- $k = 0$

  The only invariant of the zeroth order is $I_{0,0} = u_{00}$.
- $k = 1$

$$\Theta_1 = \frac{2J_{21}}{\Delta_2}\omega_2.$$

- $k = 2$

$$\Theta_2 = \frac{J_{21}}{2}\omega_1^2 + \frac{2J_{21}}{\Delta_2}\omega_2^2.$$

- $k = 3$

$$I_{3,0} = -\frac{1}{6}u_{03}u_{10}^3 + \frac{1}{2}u_{12}u_{01}u_{10}^2 - \frac{1}{2}u_{21}u_{01}^2u_{10} + \frac{1}{6}u_{01}^3u_{30},$$

$$\begin{aligned} I_{1,2} = \Delta_2^{-2}((2u_{11}^2u_{30} - 4u_{11}u_{20}u_{21} + 2u_{12}u_{20}^2)u_{01}^3 + 2u_{10}(u_{21}u_{11}^2 -\\ - 2u_{02}u_{30}u_{11} + u_{20}(2u_{21}u_{02} - u_{03}u_{20}))u_{01}^2 + 2u_{10}^2(u_{02}^2u_{30} -\\ - 2u_{02}u_{12}u_{20} + 2u_{03}u_{11}u_{20} - u_{11}^2u_{12})u_{01}\\ - 2u_{10}^3(u_{02}^2u_{21} - 2u_{02}u_{11}u_{12} + u_{03}u_{11}^2)),\end{aligned}$$

(continued)

$$I_{2,1} = \Delta_2^{-1}((-u_{11}u_{30} + u_{20}u_{21})u_{01}^3 + u_{10}(u_{02}u_{30} + u_{11}u_{21} - 2u_{12}u_{20})u_{01}^2 -$$
$$- u_{10}^2(2u_{21}u_{02} - u_{03}u_{20} - u_{11}u_{12})u_{01} + u_{10}^3(u_{02}u_{12} - u_{03}u_{11})),$$
$$\begin{aligned} I_{0,3} = \Delta_2^{-3}\Bigg(& \frac{u_{03}}{3}(u_{01}u_{20} - u_{10}u_{11})^3 \\ & + 2(u_{01}u_{11} - u_{02}u_{10})(u_{01}u_{20} - u_{10}u_{11}) \cdot \\ & \cdot (u_{01}u_{11}u_{21} - u_{01}u_{12}u_{20} - u_{02}u_{10}u_{21} + u_{10}u_{11}u_{12}) - \\ & - \frac{4u_{30}}{3}(u_{01}u_{11} - u_{02}u_{10})^3\Bigg). \end{aligned}$$

2.8 Weights

Consider the vector field $V = x\partial_x + y\partial_y$. Its flow is the scale transformations on the plane $\mathbb{C}^2$, and its ∞-th prolongation is

$$V_* = x\partial_x + y\partial_y - \sum_{k=1} k \sum_{i=1}^{k} u_{i,k-i}\partial_{u_{i,k-i}}.$$

The vector field V, as well as V_* commutes with the $\mathrm{SL}_2(\mathbb{C})$-action and therefore for every SL_2-invariant I the function $V_*(I)$ is invariant too.

We say that invariant I has *weight* $w(I) \in \mathbb{Z}$, if

$$L_{V_*}(I) = w(I)I,$$

where L_{V_*} is the Lie derivative along the vector field V_*.

Example

$$w(u_{ij}) = -(i + j), \quad w(x) = 1, \quad w(\Delta_2) = -4.$$

Since tensors Θ_k are invariants of affine transformations, $w(\Theta_k) = 0$. Moreover, $w(\omega_1) = 2$, $w(\omega_2) = 0$, and therefore $w(I_{i,j}) = -2i$.

Weights can be used to find rational $\mathrm{GL}_2(\mathbb{C})$-invariants from polynomial $\mathrm{SL}_2(\mathbb{C})$-invariants using the following observation.

Lemma 1.1 *Rational* $\mathrm{GL}_2(\mathbb{C})$*-invariants (algebraic or differential) have the form*

$$I = \frac{P}{Q},$$

where P and Q are polynomial $\mathrm{SL}_2(\mathbb{C})$*-invariants (algebraic or differential) of the same weight.*

We leave the proof of this lemma to the reader as an exercise.

2.9 Invariants of Binary Forms for $n = 2, 3, 4$

Recall that $\mathcal{B}_n \simeq \mathbb{C}^{n+1}$, and the dimension of the group $\mathrm{SL}_2(\mathbb{C})$ equals 3, therefore general orbits have dimension 3 and codimension $(n-2)$, when $n \geq 3$.

An orbit $\mathrm{SL}_2(\mathbb{C})\phi$ is said to be *regular*, if the corresponding point on the quotient $\mathbb{C}^{n+1}/\mathrm{SL}_2(\mathbb{C})$ is smooth, i.e. there exist $(n-2)$ independent (in a neighborhood of the point) rational invariants $I_1, \ldots, I_{n-2}$, such that the orbit is given by equations $I_1 = c_1, \ldots, I_{n-2} = c_{n-2}$, where c_i are constants. Independence means that $dI_1 \wedge \ldots \wedge dI_{n-2} \neq 0$ in the neighborhood of the orbit. Thus $I_1, \ldots, I_{n-2}$ are regarded as local coordinates on the quotient, and $c_1, \ldots, c_{n-2}$ are coordinates of the orbit. The Rosenlicht theorem states that all other rational invariants are rational functions of $I_1, \ldots, I_{n-2}$.

For quadrics ($n = 2$) we have only one differential invariant $\Delta_2 = u_{20}u_{02} - u_{11}^2$. Recall that by replacing u_{ij} with b_{ij} we get algebraic invariants.

For cubics ($n = 3$) we need only $\dim\left(\mathbb{C}^4/\mathrm{SL}_2(\mathbb{C})\right) = 1$ algebraic invariant, which is the discriminant Δ_3 of the cubic, and $\dim(\mathcal{E}_3/\mathrm{SL}_2(\mathbb{C})) = 3$ independent rational differential invariants, which are

$$J_1 = \Delta_2 = u_{02}u_{20} - u_{11}^2, \quad J_2 = \nabla_1(\Delta_2), \quad J_3 = \Delta_2\nabla_2(u_{00}). \tag{9}$$

Let us restrict differential invariants (9) to the cubic ϕ. We get three functions $J_1^\phi, J_2^\phi, J_3^\phi$ on a plane, namely, binary forms of degrees 2, 3, 4, therefore, there is one polynomial relation between them:

$$(J_1^\phi)^5 + (J_2^\phi)^2(J_1^\phi)^2 - 16\Delta_3(\phi)(J_3^\phi)^2 = 0, \tag{10}$$

where $\Delta_3(\phi) = \mathrm{Discr}(\phi)$ is the discriminant of the cubic.

Syzygy (10) can be obtained in Maple using the following code:

```
restart;
with(DifferentialGeometry):with(Groebner):
```

```
DifferentialGeometry:-Preferences("JetNotation",
    "JetNotation2"):
with( JetCalculus ):
DGsetup( [x, y], [u], M, 4):
Delta2:=u[0,2]*u[2,0]-u[1,1]^2:
Define invariant derivations according to (7)-(8)
nabla1:=f->u[0,1]*TotalDiff(f,x)-u[1,0]*TotalDiff(f,y):
nabla2:=f->2*(u[0,2]*u[1,0]-u[1,1]*u[0,1])/
           Delta2*TotalDiff(f,x)+2*(u[2,0]*u[0,1]-
           u[1,1]*u[1,0])/Delta2*TotalDiff(f,y):
Let phi be a binary 3-form
phi:=add(b[i,3-i]*x^i/(i!)*y^(3-i)/(3-i)!,i=0..3):
First invariant (Hessian)
J1:=u[0,2]*u[2,0]-u[1,1]^2:
Second invariant
J2:=nabla1(J1):
Third invariant
J3:=simplify(Delta2*nabla2(u[0,0])):
Restricting invariants to the cubic
Restr:=(f1,f2)->eval(f1,{u[0,0]=f2,
u[0,1]=diff(f2,y),
u[1,0]=diff(f2,x),
u[2,0]=diff(f2,x$2),
u[0,2]=diff(f2,y$2),
u[1,1]=diff(f2,[x,y]),
u[3,0]=diff(f2,x$3),
u[2,1]=diff(f2,[x,x,y]),
u[1,2]=diff(f2,[x,y,y]),
u[0,3]=diff(f2,y$3)}):
Restriction of J1 to the cubic
J1phi:=Restr(J1,phi):
Restriction of J2 to the cubic
J2phi:=Restr(J2,phi):
Restriction of J3 to the cubic
J3phi:=Restr(J3,phi):
Finding syzygy
syz1:=Basis([J1phi-Z0, J2phi-Z2, J3phi-Z3],
  plex(x, y, Z0, Z2, Z3))[1]:
```

Removing the restriction to the cubic ϕ from (10), we get a differential equation of the third order:

$$\left\{(J_1)^5 + (J_2)^2(J_1)^2 - 16\Delta_3(\phi)(J_3)^2 = 0\right\} \subset \mathbf{J}^3. \tag{11}$$

Thus we have the following criterion of $\mathrm{SL}_2(\mathbb{C})$-equivalence of binary 3-forms:

Theorem 1.7 *Let ϕ be a regular binary 3-form ($\Delta_3(\phi) \neq 0$). Then,* $\mathrm{SL}_2(\mathbb{C})$*-orbit of ϕ consists of solutions to the third order differential equation* (11) *together with $\mathcal{E}_3$.*

For quartics ($n = 4$) we take the following differential invariants

$$J_0 = u_{00}, \quad J_2 = \Delta_2 = u_{02}u_{20} - u_{11}^2, \quad J_3 = -\nabla_1(J_2).$$

Again, if we restrict these invariants to a regular quartic ϕ, we will obtain quartics J_0^ϕ, J_2^ϕ, J_3^ϕ on the plane, and the polynomial relation between them is

$$9(J_3^\phi)^2 + 16(J_2^\phi)^3 + 144\alpha(J_0^\phi)^2 J_2^\phi + 864\delta(J_0^\phi)^3 = 0, \tag{12}$$

where

$$\alpha = 4b_{13}b_{31} - b_{40}b_{04} - 3b_{22}^2$$

is the Hankel apolar, and

$$\delta = b_{22}b_{40}b_{04} - b_{04}b_{31}^2 - b_{40}b_{13}^2 + 2b_{13}b_{22}b_{31} - b_{22}^3$$

is the Hankel determinant.

Relation (12) can be obtained by means of the same Maple code as we used for cubics.

Removing the restriction to the quartic ϕ from (12), we get a differential equation of the third order:

$$\left\{9(J_3)^2 + 16(J_2)^3 + 144\alpha(J_0)^2 J_2 + 864\delta(J_0)^3 = 0\right\} \subset \mathbf{J}^3. \tag{13}$$

Thus we have a similar theorem for quartics:

Theorem 1.8 *Let ϕ be a regular binary 4-form. Then,* $\mathrm{SL}_2(\mathbb{C})$*-orbit of ϕ consists of solutions to the third order differential equation* (13) *together with $\mathcal{E}_4$.*

3 Quotients

This section gives a general introduction into the structure of quotients of algebraic manifolds and equations under the action of algebraic groups. The main results are given by the Rosenlicht and the Lie-Tresse theorems.

3.1 Rosenlicht Theorem

Let Ω be a set with an action of a group G:

$$G \times \Omega \to \Omega, \quad g \times \omega \mapsto g\omega,$$

Then, the set G/Ω of all G-orbits is called *quotient*:

$$\Omega/G = \bigcup_{\omega \in \Omega} \{G\omega\}.$$

Remark 1.1 The projection $\pi : \Omega \to \Omega/G$ allows us to identify functions on the quotient Ω/G with functions on Ω that are G-invariants, i.e. $f \circ g = f$.

Let Ω be a topological space, G be a topological group and let G-action be continuous. Then, the quotient Ω/G is naturally a topological space, that is, a subset $U \subset \Omega/G$ is said to be open if and only if the preimage $\pi^{-1}(U) \subset \Omega$ is open.

Remark 1.2 In general, we cannot guarantee that the quotient Ω/G shall inherit topological properties (e.g. the Hausdorff condition) of Ω.

Exams

1. Let $\Omega = \mathbb{R}^2$, $G = \mathrm{SL}_2(\mathbb{R})$, and $\mathrm{SL}_2(\mathbb{R}) \times \mathbb{R}^2 \to \mathbb{R}^2$ be the natural action. Then,

$$\mathbb{R}^2/\mathrm{SL}_2(\mathbb{R}) = \mathbf{0} \cup \bigstar,$$

where $\mathbf{0} = \mathrm{SL}_2(\mathbb{R})(0)$ is the orbit of the origin, $0 \in \mathbb{R}^2$, and $\bigstar$ is the orbit of any nonzero point. This is an example of the famous *Sierpinski topological space*, consisting of two points, one of which $\mathbf{0}$ is closed, but another one $\bigstar$ is open.
2. Let $\Omega = \mathbb{R}^2$, $G = \mathbb{R}^* = \mathbb{R} \setminus 0$, and $\mathbb{R}^* \times \mathbb{R}^2 \to \mathbb{R}^2$ be the natural action. Then,

$$\mathbb{R}^2/\mathbb{R}^* = \mathbf{0} \cup \mathbb{R}P^1,$$

where $\mathbb{R}P^1$ is the projective 1-dimensional space.

If Ω is a smooth manifold and G is a Lie group, then we have no way to determine whether the quotient Ω/G is also a smooth manifold, except for the case when G-action is free and proper.

Let G be an algebraic manifold (an irreducible variety without singularities over a field of zero characteristic), G be an algebraic group, and $G \times \Omega \to \Omega$ be an algebraic action. By $\mathcal{F}(\Omega)$ we denote the field of rational functions on Ω and by $\mathcal{F}(\Omega)^G \subset \mathcal{F}(\Omega)$ the field of rational G-invariants. An orbit $G\omega \subset \Omega$ (as well as the point ω) is said to be *regular*, if there are $m = \mathrm{codim}(G\omega)$ G-invariants $x_1, \ldots, x_m$, such that their differentials are linear independent at the points of the orbit.

Let $\Omega_0 = \Omega \setminus \mathrm{Sing}$ be the set of all regular points and $Q(\Omega) = \Omega_0/G$ be the set of all regular orbits.

Theorem 1.9 (Rosenlicht, [1, 21]) *The set Ω_0 is open and dense in Ω in the Zariski topology.*

Invariants $x_1, \ldots, x_m$ can be considered as local coordinates on the quotient $Q(\Omega)$ in the neighborhood of the point $G\omega \in Q(\Omega)$. On intersections of charts these coordinates are related by rational functions, which means that $Q(\Omega)$ is an algebraic manifold of the dimension $m = \mathrm{codim}(G\omega)$. Thus we have the rational map $\pi : \Omega_0 \to Q(\Omega)$ of algebraic manifolds, which gives us a field isomorphism $\mathcal{F}(\Omega)^G = \pi^*(\mathcal{F}(Q(\Omega)))$.

It is essential that the Rosenlicht's theorem is valid only for algebraic manifolds. Indeed, following the algebraic case, let Ω be a smooth manifold, and G be a Lie group. An orbit $G\omega$ (as the point ω itself) is said to be *regular*, if there are $m = \mathrm{codim}(G\omega)$ smooth independent (in the above sense) invariants. Again, let $\Omega_{\mathrm{reg}} \subset \Omega$ be the set of regular points, then the quotient Ω_{reg}/G is a smooth manifold, and the projection $\pi : \Omega_{\mathrm{reg}} \to \Omega_{\mathrm{reg}}/G$ gives us an isomorphism of algebras $C^\infty(\Omega_{\mathrm{reg}})^G$ and $C^\infty(\Omega_{\mathrm{reg}}/G)$, $\pi^*\left(C^\infty(\Omega_{\mathrm{reg}}/G)\right) = C^\infty(\Omega_{\mathrm{reg}})^G$. In contrast to the algebraic case we could not guarantee that Ω_{reg} is dense in Ω.

Let, again, Ω be an algebraic manifold, and let $\mathfrak{g}$ be a Lie subalgebra of the Lie algebra of vector fields on Ω. The Lie algebra $\mathfrak{g}$ is said to be *algebraic* if there exists an algebraic action of the algebraic group G, such that $\mathfrak{g}$ coincides with the image of the Lie algebra $\mathrm{Lie}(G)$ under this action. By an *algebraic closure* of the Lie algebra $\mathfrak{g}$ we mean an intersection of all algebraic Lie algebras, containing $\mathfrak{g}$.

Exams

1. $\Omega = \mathbb{R}$, the Lie algebra

$$\mathfrak{g} = \mathfrak{sl}_2 = \langle \partial_x, x\partial_x, x^2\partial_x \rangle$$

is algebraic.

2. $\Omega = \mathbb{R}^2$, and the Lie algebra

$$\mathfrak{g} = \langle x\partial_x + \lambda y \partial_y \rangle$$

(continued)

is algebraic if $\lambda \in \mathbb{Q}$. In the case $\lambda \notin \mathbb{Q}$ the closure is $\tilde{\mathfrak{g}} = \langle x\partial_x, y\partial_y \rangle$.

3. $\Omega = S^1 \times S^1$ — torus, the Lie algebra

$$\mathfrak{g} = \langle \partial_\phi + \lambda\partial_\psi \rangle$$

is algebraic if $\lambda \in \mathbb{Q}$. In the case $\lambda \notin \mathbb{Q}$ the closure is $\tilde{\mathfrak{g}} = \langle \partial_\phi, \partial_\psi \rangle$.

It turns out that the Rosenlicht theorem is also valid for algebraic Lie algebras, or for algebraic closure in the case of general Lie algebras.

Indeed, let $\mathfrak{g}$ be a Lie algebra of vector fields on an algebraic manifold Ω and let $\tilde{\mathfrak{g}}$ be its algebraic closure. Then, the field $\mathcal{F}(\Omega)^{\mathfrak{g}}$ of rational $\mathfrak{g}$-invariants has a transcendence degree equal to the codimension of $\tilde{\mathfrak{g}}$-orbits that is the dimension of the quotient $Q(\Omega)$.

3.2 *Algebraicity in Jet Geometry*

Let $\pi : E(\pi) \to M$ be a smooth bundle over a manifold M and let $\pi_k : \mathbf{J}^k \to M$ be the bundle of sections of k-jets.

The manifold $\mathbf{J}^k$ is equipped with the Cartan distribution, which in canonical jet coordinates (x, u^j_σ) is given by differential 1-forms

$$\varkappa^j_\sigma = du^j_\sigma - \sum_i u^j_{\sigma i} dx_i. \tag{14}$$

The Lie-Bäklund theorem [29, 30] states that types of Lie transformations, i.e. local diffeomorphisms of $\mathbf{J}^k$ preserving the Cartan distribution (14), are determined by the dimension of π, namely, they are prolongations of

- the pseudogroup $\mathrm{Cont}(\pi)$ of local *contact transformations* of $\mathbf{J}^1$, in the case $\dim \pi = 1$;
- the pseudogroup $\mathrm{Point}(\pi)$ of local *point transformations* of $\mathbf{J}^0$, i.e. local diffeomorphisms of $\mathbf{J}^0$, in the case $\dim \pi > 1$.

Moreover, it is known that

- all bundles $\pi_{k,k-1} : \mathbf{J}^k \to \mathbf{J}^{k-1}$ are affine bundles for $k \geq 2$, when $\dim \pi \geq 2$, and for $k \geq 3$, when $\dim \pi = 1$;
- prolongations of pseudogroups in canonical jet coordinates (x, u^j_σ) are given by rational in u^j_σ functions.

Therefore,

- in the case $\dim \pi \geq 2$ the fibres $\mathbf{J}_\theta^{k,0}$ of the projections $\pi_{k,0}\colon \mathbf{J}^k \to \mathbf{J}^0$ at points $\theta \in \mathbf{J}^0$ are algebraic manifolds, and the stationary subgroup $\mathrm{Point}_\theta(\pi) \subset \mathrm{Point}(\pi)$ gives us birational isomorphisms of the manifold;
- in the case $\dim \pi = 1$ the fibres $\mathbf{J}_\theta^{k,1}$ of the projections $\pi_{k,1}\colon \mathbf{J}^k \to \mathbf{J}^1$ at points $\theta \in \mathbf{J}^1$ are algebraic manifolds, and the stationary subgroup $\mathrm{Cont}_\theta(\pi) \subset \mathrm{Cont}(\pi)$ gives us birational isomorphisms of the manifold.

3.3 Algebraic Differential Equations

A differential equation $\mathcal{E}_k \subset \mathbf{J}^k$ is said to be *algebraic*, if fibres $\mathcal{E}_{k,\theta}$ of the projections $\pi_{k,0}\colon \mathcal{E}_k \to \mathbf{J}^0$, when $\dim \pi \geq 2$, or $\pi_{k,1}\colon \mathcal{E}_k \to \mathbf{J}^1$, when $\dim \pi = 1$, are algebraic manifolds.

Remark 1.3 If $\mathcal{E}_k$ is algebraic and formally integrable, then the prolongations $\mathcal{E}_k^{(l)} = \mathcal{E}_{k+l} \subset \mathbf{J}^{k+l}$ are algebraic too.

By a symmetry algebra of algebraic differential equations we mean one of the following:

- for $\dim \pi \geq 2$, a Lie algebra $\mathrm{sym}(\mathcal{E}_k)$ of point symmetries (point vector fields), which is transitive on $\mathbf{J}^0$, and stationary subalgebras $\mathrm{sym}_\theta(\mathcal{E}_k)$, $\theta \in \mathbf{J}^0$, produce actions of algebraic Lie algebras on algebraic manifolds $\mathcal{E}_{l,\theta}$, for all $l \geq k$;
- for $\dim \pi = 1$, a Lie algebra $\mathrm{sym}(\mathcal{E}_k)$ of contact symmetries (contact vector fields), which is transitive on $\mathbf{J}^1$, and stationary subalgebras $\mathrm{sym}_\theta(\mathcal{E}_k)$, $\theta \in \mathbf{J}^1$, produce actions of algebraic Lie algebras on algebraic manifolds $\mathcal{E}_{l,\theta}$, for all $l \geq k$.

Let $\mathcal{E}_k$ be a formally integrable algebraic differential equation, $\mathcal{E}_l$ be its $(l-k)$-prolongation, and $\mathfrak{g}$ be its algebraic symmetry Lie algebra. Then, all the $\mathcal{E}_l$ are algebraic manifolds, and we have a tower of algebraic bundles:

$$\mathcal{E}_k \longleftarrow \mathcal{E}_{k+1} \longleftarrow \cdots \longleftarrow \mathcal{E}_l \longleftarrow \mathcal{E}_{l+1} \longleftarrow \cdots .$$

A point $\theta \in \mathcal{E}_l$ (a $\mathfrak{g}$-orbit) is said to be *strongly regular*, if it is regular and its projection to $\mathcal{E}_{l-i}$ for all $i = 1, \ldots, l-k$ is regular too.

Let $\mathcal{E}_l^0 \subset \mathcal{E}_l$ be the set of all strongly regular points and $Q_l(\mathcal{E})$ be the set of all regular $\mathfrak{g}$-orbits. Then, due to the Rosenlicht's theorem, $Q_l(\mathcal{E})$ are algebraic manifolds, and projections $\varkappa_l\colon \mathcal{E}_l^0 \to Q_l(\mathcal{E})$ are rational maps, such that $\varkappa_l^*(\mathcal{F}(Q_l(\mathcal{E}))) = \mathcal{F}(\mathcal{E}_l^0)^{\mathfrak{g}}$, where $\mathcal{F}(Q_l(\mathcal{E}))$ is the field of rational functions on $Q_l(\mathcal{E})$, and $\mathcal{F}(\mathcal{E}_l^0)^{\mathfrak{g}}$ is the field of rational $\mathfrak{g}$-invariant functions (*rational differential invariants*).

Since the $\mathfrak{g}$-action preserves the Cartan distribution $C(\mathcal{E}_l)$, projections $\varkappa_l$ define distributions on the quotients $Q_l(\mathcal{E})$. Finally, we have the tower of algebraic bundles of the quotients

$$Q_k(\mathcal{E}) \stackrel{\pi_{k+1,k}}{\longleftarrow} Q_{k+1}(\mathcal{E}) \longleftarrow \cdots \longleftarrow Q_l(\mathcal{E}) \stackrel{\pi_{l+1,l}}{\longleftarrow} Q_{l+1}(\mathcal{E}) \longleftarrow \cdots, \qquad (15)$$

such that $(\pi_{l+1,l})_*(C(Q_{l+1}(\mathcal{E}))) = C(Q_l(\mathcal{E}))$ for $l \geq k$.

Locally, sequence (15) has the same structure as for some equation F, which is called a *quotient PDE*.

3.4 Lie-Tresse Theorem

First, we discuss Lie-Tresse derivatives, which are necessary for description of quotient PDEs.

Let $\omega \in \Omega^1(\mathbf{J}^k)$ be a differential 1-form on the space of k-jets and let C_k be the Cartan distribution. Then, the class

$$\omega^h = \pi^*_{k+1,k}(\omega) \mod \operatorname{Ann}(C_{k+1})$$

is called a *horizontal part* of ω. In the canonical jet coordinates (x, u^j_σ) we have

$$\omega = \sum_{i=1}^{n} a_i dx_i + \sum_{\substack{j \leq m \\ |\sigma| \leq k}} b^j_\sigma du^j_\sigma,$$

and its horizontal part is

$$\omega^h = \sum_{\substack{j \leq m \\ |\sigma| \leq k \\ i \leq n}} \left(a_i + b^j_\sigma u^j_{\sigma i}\right) dx_i,$$

where $n = \dim M, m = \dim \pi$.

Applying this construction to the differential df of the function $f \in C^\infty(\mathbf{J}^k)$ we get a *total differential* $\widehat{d} f = (df)^h$. In canonical coordinates it is

$$\widehat{d} f = \sum_{i=1}^{n} \frac{df}{dx_i} dx_i, \quad \frac{d}{dx_i} = \frac{\partial}{\partial x_i} + \sum_{j,\sigma} u^j_{\sigma i} \frac{\partial}{\partial u^j_\sigma}.$$

It is worth mentioning that the operation of taking the horizontal part as well as total differentials are invariant with respect to point and contact transformations.

Functions $f_1, \ldots, f_n \in C^\infty(\mathbb{J}^k)$ are said to be *in general position* in some domain D if

$$\widehat{d} f_1 \wedge \ldots \wedge \widehat{d} f_n \neq 0 \text{ in } D. \qquad (16)$$

Given fixed $f_1, \dots, f_n$ satisfying (16) one has the following decomposition for $f \in C^\infty(\mathbf{J}^k)$ in D:

$$\widehat{d} f = \sum_{i=1}^{n} F_i \widehat{d} f_i,$$

where F_i are smooth functions in the domain $\pi_{k+1,k}^{-1}(D) \subset \mathbf{J}^{k+1}$, called *Tresse derivatives* and denoted by $\frac{df}{df_i}$.

Theorem 1.10 *Let $f_1, \dots, f_n$ be $\mathfrak{g}$-invariants of order $\leq k$ in general position. Then, for any $\mathfrak{g}$-invariant f of order $\leq k$ the Tresse derivatives $\frac{df}{df_i}$ are $\mathfrak{g}$-invariants of order $\leq k+1$.*

Example Consider the action of the Lie group of translations on a plane. Its Lie algebra is

$$\mathfrak{g} = \langle \partial_x, \partial_y \rangle.$$

Let us take its invariants $f_1 = u_{00}$, $f_2 = u_{10}$, $f = u_{01}$. Then, the Tresse derivatives are of the form

$$\frac{d}{df_1} = \frac{u_{11}}{u_{10}u_{11} - u_{01}u_{20}} \frac{d}{dx} + \frac{u_{20}}{u_{01}u_{20} - u_{10}u_{11}} \frac{d}{dy},$$
$$\frac{d}{df_2} = \frac{u_{01}}{u_{01}u_{20} - u_{10}u_{11}} \frac{d}{dx} + \frac{u_{10}}{u_{10}u_{11} - u_{01}u_{20}} \frac{d}{dy}.$$

Applying them to the differential invariant $f = u_{01}$ of the first order, we get two more invariants of the second order:

$$J_1 = \frac{df}{df_1} = \frac{u_{20}u_{02} - u_{11}^2}{u_{10}u_{20} - u_{10}u_{11}}, \quad J_2 = \frac{df}{df_2} = \frac{u_{01}u_{11} - u_{02}u_{10}}{u_{01}u_{20} - u_{10}u_{11}}.$$

The following statement known as the *global Lie-Tresse theorem* [22] gives the conditions of finiteness for a generating set of invariants of a pseudogroup action on a differential equation:

Theorem 1.11 (Kruglikov, Lychagin) *Let $\mathcal{E}_k \subset \mathbf{J}^k$ be an algebraic formally integrable differential equation and let $\mathfrak{g}$ be its algebraic symmetry Lie algebra. Then, there exist rational differential $\mathfrak{g}$-invariants $a_1, \dots, a_n, b^1, \dots, b^N$ of order $\leq l$, such that the field of rational $\mathfrak{g}$-invariants is generated by rational functions of these functions and Tresse derivatives $\frac{d^{|\alpha|}b^j}{da^\alpha}$.*

Local version of this result goes back to S. Lie and A. Tresse.

Remark 1.4

1. In contrast to algebraic invariants, where we have only algebraic operations, in the case of differential invariants we have more operations. Namely, the Tresse derivatives give us new differential invariants.
2. The algebra of differential invariants is not freely generated, there are relations between invariants, called *syzygies*. The syzygies provide us with new differential equations, called *quotient equations*.
3. From the geometrical viewpoint, the Lie-Tresse theorem states that there is a level l and a domain $D \subset Q_l(\mathcal{E})$, where invariants $a_1, \dots, a_n, b^1, \dots, b^N$ serve as local coordinates, and the preimage of D in the tower

$$Q_l(\mathcal{E}) \xleftarrow{\pi_{l+1,l}} Q_{l+1}(\mathcal{E}) \longleftarrow \cdots \longleftarrow Q_r(\mathcal{E}) \xleftarrow{\pi_{r+1,r}} Q_{r+1}(\mathcal{E}) \longleftarrow \cdots \tag{17}$$

is an infinitely prolonged differential equation given by the syzygy. For this reason we call the quotient tower (17) an *algebraic diffiety*.

3.5 Integrability via Quotients

Here we discuss the importance of above constructions for integrability of differential equations. First, let us summarize the relations between differential equations and their quotients:

1. Let L be a solution to a differential equation $\mathcal{E}$ (in the sense of integral manifolds of the Cartan distribution) and let $a_i|_L, b^j|_L$ be the values of differential invariants on the solution L. Then, we have $b^j|_L = B^j(a|_L)$, and functions B^j are exactly solutions to the quotient differential equations.
2. Let $b^j = B^j(a)$ be a solution to a quotient PDE. Then, adding differential constraints $b^j - B^j(a) = 0$ we get a finite type equation $\mathcal{E} \cap \left\{b^j - B^j(a) = 0\right\}$ with solutions being a $\mathfrak{g}$-orbit of a solution to $\mathcal{E}$. This gives us a method of finding compatible constraints to be added to the original system of PDEs, which reduces the integration of the PDE to the integration of a completely integrable Cartan distribution having the same symmetry algebra. This is essential for finding smooth solutions, as well as those with singularities [31, 32].
3. Symmetries of quotient PDEs are Bäcklund-type transformations for the equation $\mathcal{E}$.

Let us now illustrate this on examples. As an exercise, we recommend the reader to do the computations for these examples.

Exams

1. Invariants of the Lie algebra $\mathfrak{g} = \langle \partial_x \rangle$ of x-translations on the line $\Omega = \mathbb{R}$ are generated by

$$\langle a = u_0, b = u_1 \rangle$$

and Tresse derivative

$$\frac{d}{da} = u_1^{-1} \frac{d}{dx}.$$

Then, for the x-invariant ODE of the third order $F(u_0, u_1, u_2, u_3) = 0$ the quotient equation is of order 2 and has the form

$$F\left(a, b, b\frac{db}{da}, b^2\frac{d^2b}{da^2}\right) = 0.$$

This is a standard reduction of order for ODEs of the form $F(u_0, u_1, u_2, u_3) = 0$.

Let us now choose other Lie-Tresse coordinates:

$$\langle a = u_2, b^1 = u_0, b^2 = u_1 \rangle$$

and Tresse derivative

$$\frac{d}{da} = u_3^{-1} \frac{d}{dx}.$$

In this case, the quotient equation for $F(u_0, u_1, u_2, u_3) = 0$ is a system of ODEs:

$$F\left(b^1, b^2, a, a\left(\frac{db^2}{da}\right)^{-1}\right) = 0, \quad a\frac{db^1}{da} - b^2\frac{db^2}{da} = 0.$$

2. Invariants of the Lie algebra $\mathfrak{g} = \langle \partial_x, x\partial_x \rangle$ of affine transformations of the line $\Omega = \mathbb{R}$ are

$$\left\langle u_0, \frac{u_2}{u_1^2}, \frac{u_3}{u_1^3}, \frac{u_4}{u_1^4}, \ldots \right\rangle.$$

Let us take

(continued)

$$\left\langle a = u_0, b = \frac{u_2}{u_1^2} \right\rangle$$

and consider a $\mathfrak{g}$-invariant equation

$$F\left(u_0, \frac{u_2}{u_1^2}, \frac{u_3}{u_1^3}, \frac{u_4}{u_1^4}\right) = 0.$$

Its quotient will be

$$F\left(a, b, \frac{db}{da} + 2b^2, \frac{d^2b}{da^2} + 6b\frac{db}{da} + 6b^3\right) = 0.$$

3. Invariants of the Lie algebra $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{R}) = \langle \partial_x, x\partial_x, x^2\partial_x \rangle$ on the line $\Omega = \mathbb{R}$ are

$$\left\langle u_0, \frac{u_3}{u_1^3} - \frac{3u_2^2}{2u_1^4}, \frac{u_4}{u_1^4} - 6\frac{u_2u_3}{u_1^5} + 6\frac{u_2^3}{u_1^6}, \dots \right\rangle.$$

Let us take

$$\left\langle a = u_0, b = \frac{u_3}{u_1^3} - \frac{3u_2^2}{2u_1^4} \right\rangle$$

and consider a $\mathfrak{g}$-invariant equation

$$F\left(u_0, \frac{u_3}{u_1^3} - \frac{3u_2^2}{2u_1^4}, \frac{u_4}{u_1^4} - 6\frac{u_2u_3}{u_1^5} + 6\frac{u_2^3}{u_1^6}\right) = 0.$$

Its quotient will be

$$F\left(a, b, \frac{db}{da}\right) = 0.$$

4. Invariants of the Lie algebra $\mathfrak{g} = \langle \partial_x, \partial_y \rangle$ on the plane $\Omega = \mathbb{R}^2$ are

$$\langle u_{00}, u_{10}, u_{01}, u_{20}, u_{11}, u_{02} \dots \rangle .$$

Let us take

$$\left\langle a_1 = u_{10}, a_2 = u_{01}, b^1 = u_{00}, b^2 = u_{11} \right\rangle$$

(continued)

as Lie-Tresse coordinates. Then, assuming $b^1 = B^1(a_1, a_2)$, $b^2 = B^2(a_1, a_2)$, we have

$$B^1_{a_1} = \delta^{-1}(u_{10}u_{02} - u_{01}u_{11}), \quad B^1_{a_2} = \delta^{-1}(u_{01}u_{20} - u_{10}u_{11}),$$

$$B^2_{a_1} = \delta^{-1}(u_{02}u_{21} - u_{11}u_{12}), \quad B^2_{a_2} = \delta^{-1}(u_{20}u_{12} - u_{11}u_{21}),$$

where $\delta = u_{20}u_{02} - u_{11}^2$ is the Hessian. The syzygies

$$\begin{aligned}
0 &= -B^1_{a_2a_2}B^2B^1_{a_1a_1} + B^2(B^1_{a_1a_2})^2 - B^1_{a_1a_2},\\
0 &= a_1B^1_{a_1a_1} + a_2B^1_{a_1a_2} - B^1_{a_1},\\
0 &= a_1B^2B^1_{a_1a_1}B^1_{a_1a_2} + a_2B^2(B^1_{a_1a_2})^2 - B^2B^1_{a_1a_1}B^1_{a_2} - a_2B^1_{a_1a_2}
\end{aligned}$$

are quotient PDEs for the equation $u_{11} = B^2(u_{10}, u_{01})$.

In particular, equation $u_{11} = 0$ is *self-dual*, it coincides with its quotient.

Remark 1.5

1. If an ODE of order k admits a solvable symmetry Lie algebra $\mathfrak{g}$, and $\dim \mathfrak{g} = k$, then the integration can be done explicitly using the Lie-Bianchi theorem. If the Lie algebra $\mathfrak{g}$ is not solvable, but still $\dim \mathfrak{g} = k$, then the integration can be done by means of model equations [33].
2. If $\dim \mathfrak{g} = k - 1$, the integration splits into the integration of the first order quotient equation and integration of $(k - 1)$ order equation with the same symmetry algebra $\mathfrak{g}$. Continuing, we reduce the integration to the integration to a series of quotients.

4 Algebraic Plane Curves

This section is devoted to finding affine invariants for algebraic plane curves using affine connections.

4.1 Connections and Affine Structures

The motivation to study connections goes back to classical mechanics, when one needs to define acceleration. If we consider a vector field Y on a manifold M as the field of velocities, then we should be able to compare tangent vectors at different points of the manifold. Let $x(t)$ be a path on the manifold M and assume that we have linear isomorphisms $\lambda(t)\colon T_{x(t)}M \to T_{x(0)}M$ of tangent spaces. Then, taking images $Y(t) = \lambda(t)\left(Y_{x(t)}\right) \in T_{x(0)}M$ of vectors $Y(t) \in T_{x(t)}M$, we get the velocity of variation of the vector field along the path $x(t)$:

$$\left.\frac{dY(t)}{dt}\right|_{t=0} \in T_{x(0)}M. \tag{18}$$

Let $x(t)$ be the trajectory of another vector field X on the manifold M. Then, taking derivatives (18) at points of M, we get a vector field $\nabla_X Y$ on M. Assuming that the map $X \times Y \to \nabla_X Y$ is $C^\infty(M)$-linear in X, we obtain the notion of a *covariant derivative*.

Let M be a smooth manifold and let $\mathcal{D}(M)$ be the module of vector fields on M. Then, the *covariant derivative* is a map

$$\nabla_X\colon \mathcal{D}(M) \to \mathcal{D}(M), \quad X \in \mathcal{D}(M),$$

satisfying conditions

1. $\nabla_{X_1+X_2} = \nabla_{X_1} + \nabla_{X_2}$
2. $\nabla_{fX} = f\nabla_X,\ f \in C^\infty(M)$,
3. $\nabla_X(Y_1 + Y_2) = \nabla_X(Y_1) + \nabla_X(Y_2)$
4. $\nabla_X(fY) = X(f)Y + f\nabla_X(Y)$,

where $X_i, Y_i, X, Y \in \mathcal{D}(M)$, $f \in C^\infty(M)$. Any affine (linear) connection on a manifold M is defined by its covariant derivative.

Let ∇ and $\tilde{\nabla}$ be two affine connections, then the difference $\Gamma_X = \nabla_X - \tilde{\nabla}_X\colon \mathcal{D}(M) \to \mathcal{D}(M)$ is a linear operator, $\Gamma_X \in \mathrm{End}(\mathcal{D}(M))$, i.e. a map $X \mapsto \Gamma_X$ is $\mathbb{R}$-linear, and $\Gamma_X(fY) = f\Gamma_X(Y)$. In other words, $\Gamma \in \mathrm{End}(\mathcal{D}(M)) \otimes \Omega^1(M)$ is an $\mathrm{End}(\mathcal{D}(M))$-valued differential one-form on M, called *connection form*, and finding connection on a manifold is equivalent to finding a connection form.

Let $M = \mathbb{R}^n$ with coordinates $(x_1, \ldots, x_n)$ be a real vector space. Consider M as an affine space with standard identifications of tangent spaces at different points, we come to the covariant derivatives

$$\nabla^s_{\partial_i}(\partial_j) = 0,$$

and any other connection has the form

$$\nabla_{\partial_i}(\partial_j) = \sum_k \Gamma^k_{ij}\partial_k,$$

where now and further on $\partial_i = \partial_{x_i}$, $d_i = dx_i$, Γ^k_{ij} are Christoffel symbols.

The *torsion tensor* T of a connection ∇ is

$$T(X, Y) = \nabla_X(Y) - \nabla_Y(X) - [X, Y],$$

which is a skew-symmetric tensor with values in vector fields, i.e. $T \in \mathcal{D}(M) \otimes \Omega^2(M)$. In coordinates, it has the form

$$T = \sum_{i,j,k} (\Gamma_{ij}^k - \Gamma_{ji}^k)\partial_k \otimes d_i \wedge d_j.$$

The connection is called *torsion-free*, if $T = 0$, i.e. $\Gamma_{ij}^k = \Gamma_{ji}^k$.

The *curvature tensor* C of a connection ∇ is

$$C \in \mathrm{End}(\mathcal{D}(M)) \otimes \Omega^2(M), \quad C(X, Y)(Z) = [\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z,$$

where $C(X, Y) \in \mathrm{End}(\mathcal{D}(M))$. In coordinates it has the form

$$C = \sum_{i,j,k,l} C_{jkl}^i \partial_i \otimes d_j \otimes d_k \wedge d_l,$$

where coefficients C_{lij}^k are related to Christoffel symbols by the following way:

$$C_{lij}^k = \frac{\partial \Gamma_{lj}^i}{\partial x_k} - \frac{\partial \Gamma_{kj}^i}{\partial x_l} + \sum_m (\Gamma_{lj}^m \Gamma_{km}^i - \Gamma_{kj}^m \Gamma_{lm}^i).$$

The torsion-free connection is said to be *flat*, if $C = 0$.

Let (M, g) be a pseudo-Riemannian manifold with a pseudo-metric tensor g. Then, there exists a unique torsion-free connection, called *Levi-Civita connection*, such that

$$g(\nabla_X Y, Z) + g(Y, \nabla_X Z) = X(g(Y, Z)), \quad X, Y, Z \in \mathcal{D}(M).$$

This relation means that $\nabla_X(g) = 0$ for all vector fields X. Christoffel symbols are related to metric g as follows:

$$\Gamma_{ij}^k = \frac{1}{2} \sum_l g^{kl} \left(\frac{\partial g_{il}}{\partial x_j} + \frac{\partial g_{jl}}{\partial x_i} - \frac{\partial g_{ij}}{\partial x_l} \right),$$

where $g_{ij} = g(\partial_i, \partial_j)$ and $\|g^{ij}\| = \|g_{ij}\|^{-1}$.

Let $\mathcal{T}_p^q(M) = (\mathcal{D}(M))^{\otimes p} \otimes (\Omega^1(M))^{\otimes q}$ be the module of p-contravariant and q-covariant tensors on the manifold M and let

$$\mathcal{T}(M) = \oplus_{p,q} \mathcal{T}_p^q(M)$$

be the bigraded tensor algebra. Then, any affine connection ∇ on the manifold M defines a derivation d_∇ of degree $(1, 1)$ in this algebra by the following way. On functions its action is $d_\nabla(f) = df$. Define this derivation on vector fields:

$$d_\nabla \colon \mathcal{D}(M) \to \mathcal{D}(M) \otimes \Omega^1(M), \quad \langle d_\nabla(X), Y\rangle = \nabla_Y(X).$$

In coordinates we have

$$d_\nabla(\partial_i) = \sum_{j,k} \Gamma^k_{ij} \partial_k \otimes d_j.$$

Then, we define this derivation on 1-forms:

$$d_\nabla \colon \Omega^1(M) \to \Omega^1(M) \otimes \Omega^1(M), \quad d_\nabla(\omega)(Y, X) = X(\omega(Y)) - \omega(\nabla_X(Y)).$$

In coordinates we have

$$d_\nabla(d_k) = -\sum_{i,j} \Gamma^k_{ij} d_j \otimes d_i$$

The action of d_∇ on higher order tensors is expanded by means of the Leibnitz rule:

$$d_\nabla(\theta_1 \otimes \theta_2) = d_\nabla(\theta_1) \otimes \theta_2 + \theta_1 \otimes d_\nabla(\theta_2).$$

We will use these constructions to get invariant symmetric tensors that will provide us with affine invariants on a plane.

4.2 Symmetric Tensors

Let $\Sigma^k(M) \subset (\Omega^1(M))^{\otimes k}$ be the module of symmetric tensors. Then,

$$\Sigma^*(M) = \oplus_{k\geq 0} \Sigma^k(M)$$

is a commutative algebra with the symmetric product. The derivation d_∇ defines a derivation of degree 1 in this algebra

$$d^s_\nabla \colon \Sigma^*(M) \to \Sigma^{*+1}(M),$$

where

$$d^s_\nabla \colon \Sigma^k(M) \xrightarrow{d_\nabla} \Sigma^k(M) \otimes \Omega^1(M) \xrightarrow{\mathrm{Sym}} \Sigma^{k+1}(M).$$

The derivation $\Sigma^k(M)$ allows to define higher order differentials $\theta_k(f)$ of functions $f \in C^\infty(M)$:

$$\Sigma^k(M) \ni \theta_k(f) = (d^s_\nabla)^k(f) \tag{19}$$

Exams Consider torsion-free connection ∇. Then, we have

$$\theta_1(f) = df = \sum_k \partial_k(f) d_k,$$

$$\theta_2(f) = \sum_{i,j} \left(\partial_{ij}(f) - \sum_k \Gamma_{ij}^k \partial_k(f) \right) d_i \cdot d_j.$$

4.3 Affine Invariants

Let us consider affine invariants of the plane. The affine Lie algebra

$$\mathfrak{aff}_2 = \langle \partial_x, \partial_y, x\partial_x, x\partial_y, y\partial_x, y\partial_y \rangle$$

acts transitively on $\mathbb{R}^2$, and therefore $\mathbf{J}^k/\mathfrak{aff}_2 = \mathbf{J}_0^k/\mathfrak{gl}_2$, where

$$\mathfrak{gl}_2 = \langle x\partial_x, x\partial_y, y\partial_x, y\partial_y \rangle.$$

The group of affine transformations preserves the trivial connection ∇^s, therefore due to construction (19) symmetric tensors

$$\Theta_k = \sum_{i=0}^{k} u_{i,k-i} \frac{dx^i}{i!} \frac{dy^{k-i}}{(k-i)!}$$

are invariants of affine transformations.

Similar to Sect.2, we construct an invariant frame ∇_1, ∇_2

$$\nabla_i = A_i \frac{d}{dx} + B_i \frac{d}{dy},$$

such that

$$2\nabla_1 \rfloor \Theta_2 = \Theta_1, \quad \Theta_2(\nabla_1, \nabla_2) = 0, \quad \Theta_2(\nabla_1, \nabla_1) = \Theta_2(\nabla_2, \nabla_2).$$

Then, we get

$$\nabla_1 = \frac{u_{02}u_{10} - u_{11}u_{01}}{u_{20}u_{02} - u_{11}^2} \frac{d}{dx} + \frac{u_{20}u_{01} - u_{11}u_{10}}{u_{20}u_{02} - u_{11}^2} \frac{d}{dy},$$

$$\nabla_2 = \frac{1}{\sqrt{u_{20}u_{02} - u_{11}^2}} \left(-u_{01}\frac{d}{dx} + u_{10}\frac{d}{dy} \right),$$

Note that the function $I_0 = \Theta_0 = u_{00}$ is an affine invariant of order zero, and therefore the function

$$I_2 = \nabla_1(I_0) = \Theta_1(\nabla_1) = 2\Theta_2(\nabla_1, \nabla_1) = \|\nabla_1\|^2 = \frac{u_{01}^2 u_{20} - 2u_{10}u_{01}u_{11} + u_{10}^2 u_{02}}{u_{20}u_{02} - u_{11}^2}$$

is a second order differential affine invariant.

The dual coframe $\langle \omega_1, \omega_2 \rangle$ consists of horizontal 1-forms, such that $\omega_i(\nabla_j) = \delta_{ij}$, and has the form

$$\omega_1 = \frac{1}{I_2}(u_{10}dx + u_{01}dy),$$
$$\omega_2 = \frac{1}{I_2\sqrt{u_{20}u_{02} - u_{11}^2}} \left((u_{11}u_{10} - u_{01}u_{20})dx + (u_{10}u_{02} - u_{11}u_{01})dy\right),$$

and we also get an affine invariant volume form

$$\omega_1 \wedge \omega_2 = \frac{\sqrt{u_{20}u_{02} - u_{11}^2}}{I_2} dx \wedge dy.$$

Summarizing above discussion, we observe that any regular function f defines the following geometric structures associated with the affine geometry on $\mathbb{R}^2$

- pseudo-Riemannian structure $\Theta_2(f)$, that gives all Riemannian invariants [34],
- symplectic structure $(\omega_1 \wedge \omega_2)(f)$,
- cubic form $\Theta_3(f)$ and Wagner connection [35],

and others.

Writing down symmetric tensors Θ_k in terms of invariant coframe, we get

$$\Theta_k = \sum_{i=0}^{k} I_{i,k-i} \frac{\omega_1^i}{i!} \frac{\omega_2^{k-i}}{(k-i)!},$$

which gives us rational affine invariants (perhaps one should take squares to get rid of square roots) $I_0 = u_{00}$,

$$I_2 = \frac{u_{01}^2 u_{20} - 2u_{10}u_{01}u_{11} + u_{10}^2 u_{02}}{u_{20}u_{02} - u_{11}^2}, \tag{20}$$

and $I_{i,k-i}$.

Since $\dim \mathbf{J}_0^k = \binom{k+2}{2}$ and $\dim(\mathfrak{gl}_2) = 4$ we observe that functions $I_0, I_2, I_{i,k-i}, 3 \le i \le k$ generate the field of rational affine differential invariants of order k.

4.4 Invariants of Algebraic Curves

A plane algebraic curve is given by equation

$$P_k(x, y) = 0,$$

where $P_k(x, y)$ is an irreducible polynomial of degree k, which is defined up to a multiplier $P_k \mapsto \lambda P_k$, $\lambda \neq 0$. This action is generated by an infinitely prolonged vector field $u_{00}\partial_{u_{00}}$:

$$\gamma = \sum_{ij} u_{ij} \frac{\partial}{\partial u_{ij}}.$$

An invariant I is said to be of weight $w(I)$, if and only if

$$\gamma(I) = w(I)I.$$

Affine invariants of zero weight are affine invariants of algebraic plane curves. Since $w(I_0) = w(I_2) = w(I_{i,j}) = 1$, one can choose

$$\mathfrak{a}_2 = \frac{I_2}{I_0}, \quad \mathfrak{a}_{ij} = \frac{I_{ij}}{I_0}$$

as a generating set of rational affine invariants of algebraic plane curves.

Remark 1.6 An algebraic plane curve is defined by its k-th jet at the point $\mathbf{0}$, and therefore values

$$\mathfrak{a}_2(P_k)(0), \quad \mathfrak{a}_{ij}(P_k)(0)$$

define the curve (completely over $\mathbb{C}$ and up to $\pm$ over $\mathbb{R}$).

To find rational invariants (without square roots of the Hessian) we will use the coframe given by total differentials of invariants $I_0 = u_{00}$ and $I_2 = (u_{01}^2 u_{20} - 2u_{10}u_{01}u_{11} + u_{10}^2 u_{02})(u_{20}u_{02} - u_{11}^2)^{-1}$:

$$\omega_1 = \widehat{d}u_{00} = \Theta_1,$$
$$\omega_2 = \widehat{d}I_2,$$

and the Tresse frame as follows:

$$\tau_1 = A_{11}\frac{d}{dx} + A_{12}\frac{d}{dy},$$

$$\tau_2 = A_{21}\frac{d}{dx} + A_{22}\frac{d}{dy},$$

where

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} u_{10} & \frac{dI_2}{dx} \\ u_{01} & \frac{dI_2}{dy} \end{pmatrix}^{-1}.$$

Expressing the original coframe $\langle dx, dy\rangle$, we get

$$\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} u_{10} & u_{01} \\ \frac{dI_2}{dx} & \frac{dI_2}{dy} \end{pmatrix}^{-1} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}.$$

Again, expression for symmetric tensors Θ_k in terms of the Tresse coframe

$$\Theta_k = \sum_{i=0}^{k} I_{i,k-i}\frac{\omega_1^i}{i!}\frac{\omega_2^{k-i}}{(k-i)!}, \tag{21}$$

gives us affine invariants $I_{i,k-i}$ of the weight $(1-k)$, and we get

Theorem 1.12 *Rational affine differential invariants are rational functions of invariants I_{ij} given by* (21).

For algebraic curves, we have

Theorem 1.13 *Rational affine differential invariants of algebraic curves are rational functions of invariants $I_{ij} I_0^{i+j-1}$.*

5 Invariants of Ternary Forms

In this section, we discuss the $\mathrm{SL}_3(\mathbb{C})$-classification problem for ternary forms of an arbitrary degree n, similar to the case of binary forms considered in Sect. 2.

Ternary forms of degree n are homogeneous polynomials on $\mathbb{C}^3$ of the form

$$\mathcal{T}_n \ni \phi_b = \sum_{i+j+k=n} b_{i,j,k}\frac{x^i}{i!}\frac{y^j}{j!}\frac{z^k}{k!}. \tag{22}$$

The action of the Lie group

$$\mathrm{SL}_3(\mathbb{C}) = \{A \in \mathrm{Mat}_{3\times 3}(\mathbb{C}) \mid \det(A) = 1\}$$

on $\mathcal{T}_n$ is defined by the following way:

$$A\colon \mathcal{T}_n \ni \phi_b \mapsto A\phi_b = \phi_b \circ A^{-1} \in \mathcal{T}_n. \tag{23}$$

The corresponding Lie algebra $\mathfrak{sl}_3$ consists of vector fields:

$$X_1 = x\partial_x - y\partial_y, \quad X_2 = x\partial_x - z\partial_z, \quad X_3 = y\partial_x, \quad X_4 = z\partial_x,$$

$$X_5 = x\partial_y, \quad X_6 = z\partial_y, \quad X_7 = x\partial_z, \quad X_8 = y\partial_z.$$

Similar to the case of binary forms, we consider (22) as smooth solutions to the Euler equation:

$$xf_x + yf_y + zf_z = nf. \tag{24}$$

Equation (24) defines a smooth manifold in the space of 1-jets of functions on $\mathbb{C}^3$:

$$\mathcal{E}_1 = \{xu_{100} + yu_{010} + zu_{001} = nu_{000}\} \subset \mathbf{J}^1.$$

As in the previous sections, we will use the notation $\mathcal{E}_k$ for the collection of all prolongations of (24) to the space $\mathbf{J}^k$ up to order k.

The action $A\colon \mathbb{C}^3 \to \mathbb{C}^3$ of the group SL_3 can be prolonged to $\mathbf{J}^k$ by the natural way

$$A^{(k)}\colon \mathbf{J}^k \to \mathbf{J}^k, \quad A^{(k)}\left([f]_p^k\right) = [Af]_{Ap}^k.$$

A rational function $I \in C^\infty(\mathcal{E}_k)$ is said to be a *differential* SL_3*-invariant of order* k, if $I \circ A^{(k)} = I$, for all $A \in \mathrm{SL}_3(\mathbb{C})$.

Using the results of Sect. 4 we define $\mathrm{SL}_3(\mathbb{C})$-invariant symmetric tensors:

$$\Theta_m = \sum_{i+j+k=m} u_{ijk}\frac{dx^i}{i!}\frac{dy^j}{j!}\frac{dz^k}{k!}. \tag{25}$$

To construct an invariant coframe we will need an inverse of Θ_2:

$$\begin{aligned}\Theta_2^{-1} = \frac{2}{A}((u_{002}u_{020} - u_{011}^2)\partial_x\partial_x - 2(u_{002}u_{110} - u_{011}u_{101})\partial_x\partial_y + \\ + 2(u_{011}u_{110} - u_{020}u_{101})\partial_x\partial_z - 2(u_{011}u_{200} - u_{101}u_{110})\partial_y\partial_z + \\ + (u_{002}u_{200} - u_{101}^2)\partial_y\partial_y + (u_{020}u_{200} - u_{110}^2)\partial_z\partial_z),\end{aligned}$$

where

$$A = u_{002}u_{020}u_{200} - u_{002}u_{110}^2 - u_{011}^2u_{200} + 2u_{011}u_{101}u_{110} - u_{020}u_{101}^2$$

is a differential $SL_3(\mathbb{C})$-invariant of order 2.

As the first invariant form ω_1, we take

$$\omega_1 = \Theta_1 = u_{100}dx + u_{010}dy + u_{001}dz.$$

The second invariant form will be the total differential of the invariant A

$$\omega_2 = \frac{dA}{dx}dx + \frac{dA}{dy}dy + \frac{dA}{dz}dz = A_1dx + A_2dy + A_3dz,$$

where

$$\begin{aligned} A_1 = {} & u_{002}u_{020}u_{300} - 2u_{002}u_{110}u_{210} + u_{002}u_{120}u_{200} - u_{011}^2u_{300} + \\ & + 2u_{011}u_{101}u_{210} + 2u_{011}u_{110}u_{201} - 2u_{011}u_{111}u_{200} - 2u_{020}u_{101}u_{201} + \\ & + u_{020}u_{102}u_{200} - u_{101}^2u_{120} + 2u_{101}u_{110}u_{111} - u_{102}u_{110}^2 \end{aligned}$$

$$\begin{aligned} A_2 = {} & u_{002}u_{020}u_{210} + u_{002}u_{030}u_{200} - 2u_{002}u_{110}u_{120} - u_{011}^2u_{210} - 2u_{011}u_{021}u_{200} + \\ & + 2u_{011}u_{101}u_{120} + 2u_{011}u_{110}u_{111} + u_{012}u_{020}u_{200} - u_{012}u_{110}^2 - 2u_{020}u_{101}u_{111} + \\ & + 2u_{021}u_{101}u_{110} - u_{030}u_{101}^2 \end{aligned}$$

$$\begin{aligned} A_3 = {} & u_{002}u_{020}u_{201} + u_{002}u_{021}u_{200} - 2u_{002}u_{110}u_{111} + u_{003}u_{020}u_{200} - \\ & - u_{003}u_{110}^2 - u_{011}^2u_{201} - 2u_{011}u_{012}u_{200} + 2u_{011}u_{101}u_{111} + 2u_{011}u_{102}u_{110} + \\ & + 2u_{012}u_{101}u_{110} - 2u_{020}u_{101}u_{102} - u_{021}u_{101}^2. \end{aligned}$$

The third invariant form $\omega_3 = F_1dx + F_2dy + F_3dz$ is found from the conditions of orthogonality to ω_2 and Θ_1 in the sense of Θ_2:

$$\Theta_2^{-1}(\omega_2, \omega_3) = 0, \quad \Theta_2^{-1}(\Theta_1, \omega_3) = 0,$$

which define the form ω_3 up to a multiplier:

$$F_1 = F_3\frac{(u_{001}u_{110} - u_{010}u_{101})A_1 + (-u_{001}u_{200} + u_{100}u_{101})A_2 + (u_{010}u_{200} - u_{100}u_{110})A_3}{(u_{001}u_{011} - u_{002}u_{010})A_1 + (-u_{001}u_{101} + u_{002}u_{100})A_2 + (u_{010}u_{101} - u_{011}u_{100})A_3},$$

$$F_2 = F_3\frac{(u_{001}u_{020} - u_{010}u_{011})A_1 + (-u_{001}u_{110} + u_{011}u_{100})A_2 + (u_{010}u_{110} - u_{020}u_{100})A_3}{(u_{001}u_{011} - u_{002}u_{010})A_1 + (-u_{001}u_{101} + u_{002}u_{100})A_2 + (u_{010}u_{101} - u_{011}u_{100})A_3}.$$

We put F_3 equal to the denominator in the above expressions:

$$F_3 = (u_{001}u_{011} - u_{002}u_{010})A_1 + (-u_{001}u_{101} + u_{002}u_{100})A_2$$

$$+ (u_{010}u_{101} - u_{011}u_{100})A_3.$$

One can check that in this case the form ω_3 will be invariant.

Now that we have constructed an invariant coframe $\langle\omega_1, \omega_2, \omega_3\rangle$, we are able to construct an invariant frame $\langle\nabla_1, \nabla_2, \nabla_3\rangle$ dual to $\langle\omega_1, \omega_2, \omega_3\rangle$:

$$\omega_i(\nabla_j) = \delta_{ij}.$$

And finally we are able to express the original coframe $\langle dx, dy, dz\rangle$ in terms of an invariant one:

$$\begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} u_{100} & u_{010} & u_{001} \\ \frac{dA}{dx} & \frac{dA}{dy} & \frac{dA}{dz} \\ F_1 & F_2 & F_3 \end{pmatrix}^{-1} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}.$$

Therefore tensors (25) are written by the following way:

$$\Theta_m = \sum_{i+j+k=m} I_{ijk} \frac{\omega_1^i}{i!} \frac{\omega_2^j}{j!} \frac{\omega_3^k}{k!}.$$

Theorem 1.14 *Functions I_{ijk} are SL_3-invariants of order $(i + j + k)$, and any rational differential invariant is a rational function of them.*

However, explicit expressions for invariants $I_{i,j,k}$ look bulky and straightforward computations work slowly in the case of ternary forms. To this reason, to find a generating set of invariants, we will use the Lie-Tresse theorem. Namely, we take five third-order independent invariants

$$J_1 = u_{00}, \quad J_2 = A, \quad J_3 = \nabla_1(J_2), \quad J_4 = \nabla_2(J_2), \quad J_5 = \nabla_3(J_2). \tag{26}$$

Since $\dim \mathcal{E}_3 = 13$, $\dim \mathfrak{sl}_3 = 8$, then we need five differential invariants to separate regular orbits. According to the global Lie-Tresse theorem, all other rational differential invariants can be found from (26) by applying invariant derivations ∇_i.

Theorem 1.15 *The field of rational $\mathfrak{sl}_3$-invariants is generated by (26) and invariant derivations ∇_i. They separate regular orbits.*

If we restrict (26) to the ternary form of degree n, we will get five functions on a three-dimensional space, therefore, there are 2 relations between them:

$$F_1(J_1^\phi, J_2^\phi, J_3^\phi, J_4^\phi, J_5^\phi) = 0, \quad F_2(J_1^\phi, J_2^\phi, J_3^\phi, J_4^\phi, J_5^\phi) = 0. \tag{27}$$

To write out syzygies (27) explicitly, one can use the similar Maple code as we used in Sect. 2 for cubics.

Theorem 1.16 *Let ϕ be a regular ternary form of degree n. Then,* $\mathrm{SL}_3(\mathbb{C})$*-orbit of ϕ consists of solutions to a quotient PDE*

$$F_1(J_1, J_2, J_3, J_4, J_5) = 0, \quad F_2(J_1, J_2, J_3, J_4, J_5) = 0.$$

together with $\mathcal{E}_n$.

Acknowledgments This work was partially supported by the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS" (project 19-7-1-13-3).

References

1. Vinberg, E., Popov, V.: Invariant theory. Itogi Nauki i Tekhniki. Ser. Sovrem. Probl. Mat. Fund. Napr. **55**, 137–309 (1989)
2. Alekseev, V.G.: Theory of Rational Invariants of Binary Forms (in Russian). Yuriev (1899)
3. Cayley, A.: On linear transformations. Camb. Dublin Math. J. **4** (1844)
4. Cayley, A.: Mémoire sur les hyperdéterminantes. Crelle's J. **30** (1845)
5. Cayley, A.: Seven memoirs upon quantics. Philos. Trans. **144,146,148,149,151** (1854–1861)
6. Cayley, A.: 8th memoir upon quantics. Philos. Trans. **157** (1867)
7. Cayley, A.: 9th memoir upon quantics. Philos. Trans. **161** (1871)
8. Cayley, A.: 10th memoir upon quantics. Philos. Trans. **169** (1878)
9. Aronhold, S.: Ueber ein fundamentale Begründung der Invariantentheorie. Crelle's J. **62** (1863)
10. Clebsch, A.: Ueber simultane Integration linearer partieller Differentialgleichungen. Crelle's J. **65** (1866)
11. Sylvester, J.J.: Détermination d'une limite supérieure au nombre total des invariants et covariants irréducibles des formes binaires. C. R. **86** (1878)
12. Gordan, P.: Beweis dass jede Covariante und Invariante einer binären Form eine ganze Function mit numerischen Coefficienten einer endlichen Anzahl solcher Formen 1st. Crelle's J. **69** (1869)
13. von Gall, A.F.: Das vollständige Formensystem einer binären Form achter Ordnung. Math. Ann. **17**(1), 31–51 (1880)
14. von Gall, A.F.: Das vollständige Formensystem einer binären Form 7ter Ordnung. Math. Ann. **31**(3), 318–336 (1888)
15. Dixmier, J., Lazard, D.: Minimum number of fundamental invariants for the binary form of degree 7. J. Symbol. Comput. **6**(1), 113–115 (1988)
16. Hammond, J.: A simple proof of the existence of irreducible invariants of degrees 20 and 30 for the binary seventhic. Math. Ann. **36**, 255 (1890)
17. Hilbert, D.: Über die Endlichkleit des Invariantsystems für binäre Grundformen. Math. Ann. **33**, 223–226 (1889)
18. Halphen, G.H.: Sur les invariants différentiels. Paris, France: Gauthier-Villars (1878)
19. Lie, S.: Vorlesungen über Differentialgleichungen mit bekannten infinitesimalen Transformationen. Lepzig (1891)
20. Lie, S.: Vorlesungen über continuirliche Gruppen. Lepzig (1893)
21. Rosenlicht, M.: Some basic theorems on algebraic groups. Am. J. Math. **78**, 401–443 (1956)
22. Kruglikov, B., Lychagin, V.: Global Lie-Tresse theorem. Selecta Math. **22**, 1357–1411 (2016)
23. Lychagin, V., Bibikov, P.: Differential contra algebraic invariants: applications to classical algebraic problems. Lobachevskii J. Math. **37**(1), 36–49 (2016)
24. Lychagin, V., Bibikov, P.: Projective classification of binary and ternary forms. J. Geom. Phys. **61**(10), 1914–1927 (2011)

25. Anderson, I., Torre, C.G.: The differential geometry package. Downloads. Paper 4. http://digitalcommons.usu.edu/dg_downloads/4 (2016)
26. Hilbert, D.: Theory of Algebraic Invariants. Cambridge University Press (1993)
27. Schur, I., Grunsky, H.: Vorlesungen über Invariantentheorie. Die Grundlehren der mathematischen Wissenschaften, vol. 143. Springer, Berlin (1968)
28. Goldberg, V., Lychagin, V.: Geodesic webs on a two-dimensional manifold and Euler equations. Acta. Appl. Math. **109**, 5–17 (2010)
29. Vinogradov, A.M., Krasilshchik, I.S. (eds.): Symmetries and Conservation Laws for Differential Equations of Mathematical Physics. Factorial, Moscow (1997)
30. Krasilshchik, I.S., Lychagin, V.V., Vinogradov, A.M.: Geometry of Jet Spaces and Nonlinear Partial Differential Equations. Gordon and Breach Science Publishers (1986)
31. Schneider, E.: Solutions of second-order PDEs with first-order quotients. Lobachevskii J. Math. **41**(12), 2491–2509 (2020)
32. Lychagin, V., Roop, M.: Shock waves in Euler flows of gases. Lobachevskii J. Math. **41**(12), 2466–2472 (2020)
33. Lychagin, V.: Symmetries and integrals. In: Ulan, M., Schneider, E. (eds.) Differential Geometry, Differential Equations, and Mathematical Physics, pp. 73–121. Birkhauser, Cham (2021)
34. Lychagin, V., Yumaguzhin, V.: Invariants in relativity theory. Lobachevskii J. Math. **36**(3), 298–312 (2015)
35. Lychagin, V., Yumaguzhin, V.: On equivalence of third order linear differential operators on two-dimensional manifolds. J. Geom. Phys. **146**, 1–18 (2019)

Lectures on Poisson Algebras

Vladimir Rubtsov and Radek Suchánek

1 Introduction

The notion of a Poisson algebra was probably introduced in the first time by A.M. Vinogradov and J. S. Krasil'shchik in 1975 under the name "canonical algebra" and by J. Braconnier in his short note "Algèbres de Poisson" (Comptes rendus Ac.Sci) in 1977.

It was a natural "algebraic interpretation" of the notion of Poisson structure and Poisson brackets appeared in nineteenth century in the framework of Classical Mechanics.

Nowadays, Poisson algebras have proved to have a very rich mathematical structure (see the beautiful short article of Y. Kosmann-Schwarzbach in "Encyclopedia of Mathematics" [1].

There are many good books and lecture notes devoted to Poisson structures and much more less sources concerning the algebraic side of the story. We should mention the extensive book of C. Laurent-Gengoux, A. Pichereau and P. Vanhaecke "Poisson structures" which covers also many algebraic aspects and structures associated with the notion of a Poisson variety [2].

V. Rubtsov
LAREMA UMR 6093, CNRS and Université d'Angers, Angers Cedex, France

IGAP (Institute of Geometry and Physics), Trieste, Italy
e-mail: vladimir.roubtsov@univ-angers.fr

R. Suchánek (✉)
Department of Mathematics and Statistics, Masaryk University, Brno, Czech Republic

LAREMA UMR 6093, CNRS, Université d'Angers, Angers Cedex, France

M. Ulan, S. Hronek (eds.), *Groups, Invariants, Integrals, and Mathematical Physics*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-031-25666-0_2

These notes are slightly extended reproduction of the mini-course of the same title (5 lectures) given by the first author during the virtual (online) Winter School and Workshop “Wisla 20–21” and handled and written by the second.

The virtual character of lectures has imposed few constraints and has defined a specific choice of material and chosen subjects. The authors had to pass between Scylla of unknown audience level and tantalizing Charybdis of their will to exit out of the standard set of content in numerous existed lecture notes about Poisson structures in algebra and in geometry.

This contradiction can probably explain some strange “jumps” of a difficulty level between various chapters of our notes. We hope to come back to these notes and to extend, to improve, or even, to write a new comprehensive book covering the subject. Right now we were on a serious time (and the Covid pandemic) limits pressure and it was also a reason of some non-homogeneous choice of the lecture notes matters and its details.

There are (almost) no “new results” in the lectures. The only exception is a content of the last brief chapter where we provide our classification results (joint with A. Odesskii and V. Sokolov) on the “low rank” double quadratic Poisson brackets. We had kept this chapter almost in it’s specific form of a computer presentation.

The lectures are based on many various well- and less known sources, which we had tried to carefully quot in the main body of the text. But we need to specify most influential: the beautiful paper of A.M. Vinogradov and J.S. Krasil’shchik [3] “What is the hamiltonian formalism?”, lectures of P. Cartier “Some fundamental techniques in the theory of integrable systems” [4], a short book of K. H. Bhaskara and K. Viswanath “Poisson algebras and Poisson manifolds” [5].

2 Motivation

2.1 Lagrangian and Hamiltonian Mechanics

A fundamental discovery of Lagrange: a *Lagrangian* function $L\colon TM \to \mathbb{R}$

$$L = T - V\ ,$$

describes the motion of a particle. The equation of motion is described by the Euler-Lagrange equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) = \frac{\partial L}{\partial q^i}\ .$$

Here $(\bar{q}) = (q^1, \dots, q^n)$ are coordinates on M, $(\bar{q}, \dot{\bar{q}}) = (q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n)$ are the induced coordinates on the tangent bundle TM. The above Euler-Lagrange equation follows from the variational principle

$$\delta \int L(\bar{q}, \dot{\bar{q}}) \mathrm{d}t = 0 ,$$

where δ (in this context) refers to the variational derivative. Using the Legendre transformation, one can pass from the tangent bundle to the cotangent bundle T^*M with the coordinates $(\bar{q}, \bar{p})$, and reformulate the Euler-Lagrange equations equivalently by the Hamiltons equations

$$\dot{\bar{q}} = \frac{\partial H}{\partial \bar{p}}, \quad \dot{\bar{p}} = -\frac{\partial H}{\partial \bar{q}} \tag{1}$$

The link between Lagrangian and Hamiltonian is given by

$$H(\bar{q}, \bar{p}) =< \bar{p}, \dot{\bar{q}} > -L(\bar{q}, \dot{\bar{q}}) .$$

2.2 Hamiltonian Mechanics and Poisson Brackets

Instead of starting with the Lagrangian L function, one can start with the Hamiltonian function $H\colon T^*M \to \mathbb{R}$ and build the mechanics independently of the Lagrangian. We now go to the simpler setup of flat spaces, i.e. consider $M \cong \mathbb{R}^N$. We will get back to the setup of smooth manifolds later on.

Let F, G be two differentiable functions on $\mathbb{R}^{2n}$ with coordinates $(\bar{q}, \bar{p})$, where $\bar{p} = (p_1, \dots, p_n), \bar{q} = (q_1, \dots, q_n)$. Introduce a new functions as the result of the following skew-symmetric operation on F and G.

$$\{F, G\} := \sum_{k=1}^{n} (\frac{\partial F}{\partial q_k} \frac{\partial G}{\partial p_k} - \frac{\partial G}{\partial q_k} \frac{\partial F}{\partial p_k}) = \frac{\partial F}{\partial \bar{q}} \frac{\partial G}{\partial \bar{p}} - \frac{\partial G}{\partial \bar{q}} \frac{\partial F}{\partial \bar{p}} \in C^\infty(\mathbb{R}^{2n}) . \tag{2}$$

This operations allows to write the Hamilton's equations (1) associated with a Hamiltonian function $H = H(\bar{q}, \bar{p})$ in the following manner (putting $\bar{q}$ and $\bar{p}$ on an equal footing):

$$\dot{\bar{q}} = \{\bar{q}, H\} , \quad \dot{\bar{p}} = \{\bar{p}, H\} . \tag{3}$$

Indeed we have

$$\{q_i, H\} = \sum_{k=1}^{n} (\frac{\partial q_i}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial H}{\partial q_k} \frac{\partial q_i}{\partial p_k}) = \frac{\partial H}{\partial p_i} ,$$

$$\{p_i, H\} = \sum_{k=1}^{n} (\frac{\partial p_i}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial H}{\partial q_k} \frac{\partial p_i}{\partial p_k}) = -\frac{\partial H}{\partial q_i} ,$$

since $\frac{\partial q_i}{\partial q_k} = \frac{\partial p_i}{\partial p_k} = \delta^i_k$ and $\frac{\partial q_i}{\partial p_k} = \frac{\partial p_i}{\partial q_k} = 0$.

Note that for each solution $\{(q(t), p(t)\}$ of the Hamilton system (3) and for any $F \in C^\infty(\mathbb{R}^{2n})$

$$\begin{aligned}\frac{d}{dt}F(\bar{q}(t), \bar{p}(t)) &:= \sum_{k=1}^{n}(\frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial p_i}\dot{p}_i) \\ &= \sum_{k=1}^{n}(\frac{\partial F}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial H}{\partial q_i}) \\ &= \{F, H\}(\bar{q}(t), \bar{p}(t)) \ .\end{aligned}$$

In particular, for the Hamiltonian function H,

$$\frac{\mathrm{d}}{\mathrm{d}t}H(\bar{q}(t), \bar{p}(t)) = \{H, H\}(\bar{q}(t), \bar{p}(t)) = 0 \ ,$$

which is the *conservation law* for H (i.e. $H(\bar{q}(t), \bar{p}(t))$ is constant along the trajectories). D. Poisson had observed that if $\{F, H\}$ and $\{G, H\}$ vanish, then $\{\{F, G\}, H\}$ vanishes (Poisson bracket of two constants of motion is again a constant of motion). This statement can be explained abstractly with the help of the *Jacobi identity*

$$\{\{F, G\}, H\} + \{\{G, H\}, F\} + \{\{H, F\}, G\} = 0 \ , \tag{4}$$

since if $\{F, H\} = \{G, H\} = 0$, then the above equality gives $\{\{F, G\}, H\} = 0$. Now let π_{ij} be arbitrary smooth functions on $\mathbb{R}^d$, $x_1, \ldots, x_d$ coordinates on $\mathbb{R}^d$ such that $\pi_{ij} = -\pi_{ji}$. Define

$$\{F, G\} := \sum_{i,j=1}^{d} \pi_{ij}\frac{\partial F}{\partial x_i}\frac{\partial G}{\partial x_j} \ . \tag{5}$$

Then (5) satisfies the Jacobi identity (4) if and only if

$$\sum_{k=1}^{d}(\pi_{lk}\frac{\partial \pi_{ij}}{\partial x_l} + \pi_{li}\frac{\partial \pi_{jk}}{\partial x_l} + \pi_{lj}\frac{\partial \pi_{ki}}{\partial x_l}) = 0 \ . \tag{6}$$

Skew-symmetry and the Jacobi identity imply that the operation $\{\cdot, \cdot\}$ defines a *Lie algebra* structure on $C^\infty(\mathbb{R}^d)$. Moreover this structure is compatible with the product operation $\cdot$ on $C^\infty(\mathbb{R}^d)$, $U \subset \mathbb{R}^d$, meaning that it satisfies the *Leibniz rule*

$$\{FG, H\} = F\{G, H\} + \{F, H\}G \ .$$

3 Poisson Algebras

There are two ways to think about a Poisson structure (given by a Poisson bracket) on a manifold M (algebraic, smooth, analytic, etc.).

A Geometric Viewpoint To each function H (smooth, holomorphic, etc.) on M, one can associate a vector field X_H, where H is the Hamiltonian in the mechanical interpretation. The vector field is given by

$$X_H := \{-, H\} . \tag{7}$$

An Algebraic Viewpoint Consider a vector space (typically infinite dimensional) $\mathcal{A}$, with two algebra structures:

1. Structure of an associative and commutative algebra with a unit.
2. A Lie algebra structure.

The Poisson structure: commutative structure + Lie algebra structure + compatibility conditions.

Extracting the algebraic aspects of the above ideas, one is lead to the following definition.

Definition 2.1 (Poisson Algebra) Let $(\mathcal{A}, \cdot)$ be an associative, unital and commutative $\mathbb{K}$-algebra ($\operatorname{char} \mathbb{K} = 0$), with the unit denoted 1. Let

$$\{-, -\} \colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$$

be a Lie bracket defining a $\mathbb{K}$-Lie algebra structure on $\mathcal{A}$. Then $\mathcal{A}$ is called *a Poisson algebra* if the operations $\cdot$ and $\{-, -\}$ are compatible in the following sense

$$\{a \cdot b, c\} = a \cdot \{b, c\} + \{a, c\} \cdot b . \tag{8}$$

The Lie bracket is then called *a Poisson bracket.*

When the context is clear, we will usually omit the $\cdot$ symbol of the associative operation in $\mathcal{A}$.

Remark 2.1 Note that the Lie bracket $\{-, -\}$ in the above definition is not necessarily given by the commutator. In fact, if it is given by the commutator, i.e. $\{a, b\} = a \cdot b - b \cdot a$, then $\{a, b\} = 0$ for all $a, b \in \mathcal{A}$, since the $\cdot$ is assumed to be a commutative operation on $\mathcal{A}$.

3.1 Subalgebras and Ideals

Definition 2.2 (Subalgebras) A vector subspaces $\mathcal{B} \subset \mathcal{A}$ is *a Poisson subalgebra* in a Poisson algebra $(\mathcal{A}, \{-,-\}, \cdot)$ if $\mathcal{B} \cdot \mathcal{B} \subset \mathcal{B}$ and $\{\mathcal{B}, \mathcal{B}\} \subset \mathcal{B}$.

Definition 2.3 (Ideals) A vector subspaces $\mathcal{J} \subset \mathcal{A}$ is *a Poisson ideal* of a Poisson algebra $(\mathcal{A}, \{-,-\}, \cdot)$ if $\mathcal{J} \cdot \mathcal{A} \subset \mathcal{J}$ and $\{\mathcal{J}, \mathcal{A}\} \subset \mathcal{J}$.

Note 2.1 In the first definition, $\mathcal{B}$ is itself a Poisson algebra with operations given by restriction of operations in $\mathcal{A}$. The inclusion map $\iota \colon \mathcal{B} \subset \mathcal{A}$ is a Poisson algebra morphism. In the second definition, the quotient algebra $\mathcal{A}/\mathcal{J}$ inherits a Poisson bracket from $\mathcal{A}$ by the requirement that the canonical projection on the quotient $p \colon \mathcal{A} \to \mathcal{A}/\mathcal{J}$ is a Poisson morphism.

Remark 2.2 For a fixed field $\mathbb{K}$ ($\operatorname{char} \mathbb{K} = 0$), $\mathrm{Poiss}_{\mathbb{K}}$ denotes a category, whose objects are the Poisson $\mathbb{K}$-algebras, and whose morphisms are $\mathbb{K}$-morphisms of Poisson algebras. Group object in $\mathrm{Poiss}_{\mathbb{K}}$ is the Poisson-Lie group $(G, \cdot, \{-,-\})$ s.t. the bracket on G is *multiplicative*, i.e. the multiplication $\cdot G \times G \to G$ is a Poisson morphism, where $G \times G$ is considered with a Poisson bracket given by $\{-,-\}_{G\times G} := \{-,-\}_G \oplus \{-,-\}_G$.

Definition 2.4 (Prime) A Poisson ideal $\mathcal{P} \subset \mathcal{A}$ is *Poisson prime*, if for all Poisson ideals $\mathcal{I}, \mathcal{J} \subset \mathcal{A}$

$$\mathcal{I}\mathcal{J} \subset \mathcal{P} \Rightarrow \mathcal{I} \subset \mathcal{P} \text{ or } \mathcal{J} \subset \mathcal{P}$$

Remark 2.3 If $\mathcal{A}$ is *Noetherian*,[1] then the above definition is equivalent to $\mathcal{P}$ being both prime ideal and Poisson ideal.

Definition 2.5 (Spectrum) The *Poisson spectrum of* $\mathcal{A}$, denoted $\mathrm{PSpec}(\mathcal{A})$, is the set of Poisson prime ideals of $\mathcal{A}$.

Definition 2.6 (Maximal Ideal) A maximal ideal $\mathcal{M} \subset \mathcal{A}$ is a *Poisson maximal ideal* if it is a Poisson ideal.

Remark 2.4 A Poisson maximal ideal is not the same thing as a maximal Poisson ideal.

Definition 2.7 (Core) The *Poisson core* of an ideal $\mathcal{I} \subset \mathcal{A}$, denoted $\mathrm{PC}(\mathcal{I})$, is the largest Poisson ideal of $\mathcal{A}$ contained in $\mathcal{I}$.

Remark 2.5 If $\mathcal{I}$ is a prime ideal of $\mathcal{A}$, then $\mathrm{PC}(\mathcal{I})$ is Poisson prime.

Definition 2.8 (Primitive Ideal) A Poisson ideal $\mathcal{P}$ of $\mathcal{A}$ is a *Poisson primitive* if $\mathcal{P}$ is the Poisson core of some maximal ideal $\mathcal{M}$, i.e. $\mathrm{PC}(\mathcal{M}) = \mathcal{P}$

Remark 2.6 Each Poisson primitive ideal is Poisson prime.

[1] Given an increasing sequence of ideals in $\mathcal{A}$, $\mathcal{I}_1 \subseteq \mathcal{I}_2 \subseteq \ldots$, there always exists $k \in \mathbb{N}$ such that $\mathcal{I}_k = \mathcal{I}_{k+n}$ for all $n \in \mathbb{N}$.

Definition 2.9 (Localization) Let R be a commutative ring and $S \subset R$ a multiplicative subset.[2] A *localization of* R *at* S, denoted $S^{-1}R$, is the commutative ring of equivalence classes in $R \times S$ defined by

$$(\tilde{r}, \tilde{s}) \in [r, s] \iff \exists k \in S : (\tilde{r}s - r\tilde{s})k = 0 .$$

One usually denotes the equivalence class $[r, s]$ by $\frac{r}{s}$ or by rs^{-1} (this reflects the similarity with the construction of the field of fractions). The ring operations are given by

$$\frac{\tilde{r}}{\tilde{s}} + \frac{r}{s} := \frac{\tilde{r}s + r\tilde{s}}{\tilde{s}s}$$
$$\frac{\tilde{r}}{\tilde{s}} \cdot \frac{r}{s} := \frac{\tilde{r}r}{\tilde{s}s}$$

Exercise 2.1 Show that the above construction indeed yields a ring structure on the set $S^{-1}R$. Describe the identity elements with respect to both ring operations.

If $S \subset \mathcal{A}$ is a multiplicatively closed subset of a Poisson algebra $\mathcal{A}$, then the localization $S^{-1}\mathcal{A}$ is also a Poisson algebra with the brackets

$$\{as^{-1}, bt^{-1}\} := (st\{a, b\} - sb\{a, t\} - at\{s, b\} + ab\{s, t\})(s^2t^2)^{-1},$$

for all $a, b \in \mathcal{A}$ and all $s, t, \in S$. For any Poisson ideal $\mathcal{P} \subset \mathcal{A}$, the localization $S^{-1}\mathcal{P}$ is a Poisson ideal in $S^{-1}\mathcal{A}$. For two multiplicative sets $S, T \subset \mathcal{A}$ such that $S \subset T$, there is a Poisson morphism $: S^{-1}\mathcal{A} \to T^{-1}\mathcal{A}$.

Exercise 2.2 Let $s, t \in \mathcal{A}$ and consider the multiplicative subsets of $\mathcal{A}$

$$S := \{s^n \mid n \in \mathbb{N}\} \qquad\qquad T := \{t^n \mid n \in \mathbb{N}\} ,$$

where $s, t \in \mathcal{A}$ are such that $t = su$ for appropriate $u \in \mathcal{A}$. Show that the map $\varphi_{s,t} \colon S^{-1}\mathcal{A} \to T^{-1}\mathcal{A}$ defined by $\varphi_{s,t}(\frac{a}{s^m}) := \frac{au^n}{t^n}$ is independent of u and is a Poisson morphism.

Example Any Poisson maximal ideal is maximal Poisson, but the converse is not true. A counter-example is the Weyl-Poisson algebra $\mathbb{C}[x, y]$, with $\{x, y\} = 1$. This is a simple Poisson algebra but not a simple associative algebra (consider the trivial ideal $\{0\}$).

[2] $s_1, s_2 \in S \Rightarrow s_1s_2 \in S$.

3.2 Morphisms and Derivations

Definition 2.10 (Morphisms and Isomorphisms) Let $(\mathcal{A}_i, \{-,-\}_i, \cdot_i)_{i=1,2}$ be two Poisson algebras over $\mathbb{K}$ and let $\varphi\colon \mathcal{A}_1 \to \mathcal{A}_2$ be a $\mathbb{K}$-linear map satisfying for all $a, b \in \mathcal{A}_1$

1. $\varphi(a \cdot_1 b) = \varphi(a) \cdot_2 \varphi(b)$,
2. $\varphi(\{a, b\}_1) = \{\varphi(a), \varphi(b)\}_2$

Then φ is called *a morphism of Poisson algebras*. If φ is a bijective morphism of Poisson algebras s.t. the inverse map $\varphi^{-1}\colon \mathcal{A}_2 \to \mathcal{A}_1$ is also a Poisson algebra morphism, then φ is called *an isomorphism of Poisson algebras*.

Definition 2.11 (Derivation) A $\mathbb{K}$-linear map $\varphi\colon \mathcal{A} \to \mathcal{A}$ is called *a* $\mathbb{K}$*-derivation* on $\mathcal{A}$ if

$$\varphi(a \cdot b) = a\varphi(b) + \varphi(a) \cdot b\ . \tag{9}$$

The set of all derivations of $\mathcal{A}$ is denoted by $\mathcal{D}(\mathcal{A})$, $\mathcal{D}^1(\mathcal{A})$ or $\mathcal{D}(\mathcal{A}, \mathcal{A})$. A bilinear map $\varphi\colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ is called *a bi-derivation* if

$$\begin{aligned} \varphi(ab, c) &= a\varphi(b, c) + b\varphi(a, c), \\ \varphi(a, bc) &= \varphi(a, b)c + b\varphi(a, c) \end{aligned} \tag{10}$$

Note 2.2 Directly from the definition, $\mathcal{D}(\mathcal{A}) \subset \mathrm{End}_{\mathbb{K}}(\mathcal{A}) = \mathrm{Hom}_{\mathbb{K}}(\mathcal{A}, \mathcal{A})$. The $\mathbb{K}$-linear map, given for every $b \in \mathcal{A}$, $a \mapsto \{a, b\}$ is a derivation on $\mathcal{A}$.

Note 2.3 The biderivation associated with derivation on $\mathcal{A}$ given by a Poisson bracket is always skew-symmetric.

From now on, we will often abbreviate the notation and refer to a Poisson algebra $(\mathcal{A}, \{-,-\}, \cdot)$ simply by $\mathcal{A}$, if the context is clear. Also, all the vector spaces, algebras (Poisson, Lie), and the corresponding linear maps (morphisms, derivations), will be considered over a general field $\mathbb{K}$, with $\mathrm{char}\,\mathbb{K} = 0$, unless stated otherwise.

4 Hamiltonian Derivations and Casimirs

Definition 2.12 (Hamiltonian Derivations and Casimirs) Let $\mathcal{A}$ be a Poisson algebra and $a \in \mathcal{A}$. The derivation $X_a := \{-, a\} \in \mathcal{D}(\mathcal{A})$ is called a Hamiltonian derivation with a Hamiltonian (or a Hamilton function) a associated to X_a. We denote by

$$\mathrm{Ham}(\mathcal{A}) := \{X_a \mid a \in \mathcal{A}\} \tag{11}$$

the set of all Hamiltonian derivations. We have a linear map $\mathcal{A} \to \mathrm{Ham}(\mathcal{A})$, $a \mapsto X_a$. Let $a \in \mathcal{A}$ be s.t. $X_a = 0$, then a is called a *Casimir element*. We denote by

$$\mathrm{Cas}(\mathcal{A}) := \{a \in \mathcal{A} | X_a = 0\} \tag{12}$$

the set of all Casimir elements.

Definition 2.13 The (Poisson) *center of* $\mathcal{A}$ is $\mathrm{Z}_P(\mathcal{A}) := \{a \in \mathcal{A} \mid \{a, b\} = 0 \text{ for all } b \in \mathcal{A}\}$. We have $\mathrm{Cas}(\mathcal{A}) = \mathrm{Z}_P(\mathcal{A})$.

Definition 2.14 *Augmentation* of an (associative) $\mathbb{K}$-algebra $\mathcal{A}$ is a $\mathbb{K}$-algebra homomorphism $\alpha \colon \mathcal{A} \to \mathbb{K}$. The pair $(\mathcal{A}, \alpha)$ is called an *augmented algebra*. The kernel $\ker \alpha$ is called the augmentation ideal of $\mathcal{A}$.

Lemma 2.1 *For arbitrary monic polynomial* $p(x) \in \mathrm{Cas}[x]$ *and arbitrary* $b \in \mathcal{A}$

$$\{p(x), b\} = p'(x)\{x, b\} \, , \tag{13}$$

where $p'(x) = \frac{\partial p(x)}{\partial x}$.

Proof By induction. Suppose $\tau := \deg p = 0$. Then $p(x) = c_0$, where $c_0 \in \mathrm{Cas}(\mathcal{A})$, and hence for all $b \in \mathcal{A}$

$$\{p(x), b\} = \{c_0, b\} = 0 = p'(x)\{x, b\} \, .$$

Let the induction hypothesis hold for monic polynomials of deg $\leq \tau - 1$ and suppose $\tau > 0$. Then $p(x) = x^\tau + c_{\tau-1} x\tau - 1 + \ldots + c_1 x + c_0$, where all $c_i \in \mathrm{Cas}(\mathcal{A})$. We can rewrite $p(x) = x^\tau + c_{\tau-1} q(x)$, where $q(x)$ is monic and $\deg q = \tau - 1$. Then

$$\begin{aligned}
\{p(x), b\} &= \{x x^{\tau-1}, b\} + c_{\tau-1}\{q(x), b\} \\
&= x^{\tau-1}\{x, b\} + x\{x^{\tau-1}, b\} + c_{\tau-1} q'(x)\{x, b\} \\
&= (x^{\tau-1} + x(\tau - 1)x^{\tau-2} + c_{\tau-1} q'(x))\{x, b\} \\
&= p'(x)\{x, b\} \, .
\end{aligned}$$

□

Proposition 2.1 *Let* $\mathcal{A}$ *be a Poisson algebra.*

1. $\mathrm{Cas}(\mathcal{A})$ *is a subalgebra of* $(\mathcal{A})$. *Moreover,* $\mathbb{K}$ *can be naturally identified with a subset of* $\mathrm{Cas}(\mathcal{A})$.
2. *If* $\mathcal{A}$ *is an integral domain (i.e. has no zero divisors), then* $\mathrm{Cas}(\mathcal{A})$ *is integrally closed in* $\mathcal{A}$.
3. *Not every representation of* $\mathcal{A}$ *on* $\mathrm{End}(\mathcal{A})$ *defines an* $\mathcal{A}$*-module structure on* $\mathrm{Ham}(\mathcal{A}) \subset \mathrm{End}(\mathcal{A})$.
4. $\mathrm{Ham}(\mathcal{A})$ *is a* $\mathrm{Cas}(\mathcal{A})$*-module.*
5. *The map* $\varphi \colon \mathcal{A} \to \mathcal{D}(\mathcal{A})$*, defined by* $H \mapsto -X_H$ *is a morphism of Lie algebras.*

6. *There is a short exact sequence of Lie algebras*

$$0 \to \mathrm{Cas}(\mathcal{A}) \to \mathcal{A} \to \mathrm{Ham}(\mathcal{A}) \to 0\,.$$

Proof

1. That $\mathrm{Cas}(\mathcal{A})$ is a subalgebra of $\mathcal{A}$ is obvious. Now we show that scalars correspond to Casimir elements. Notice that since $\mathcal{A}$ is an algebra with the unit, 1, and because $\mathbb{K}$ has an action on $\mathcal{A}$ (denoted also by juxtaposition), there is a natural injective morphism $\iota\colon \mathbb{K} \to \mathcal{A}$, given by $k \mapsto k1$. This means that we can identify $\mathbb{K}$ with $\iota(\mathbb{K}) = \mathbb{K}1 \subset \mathcal{A}$. We will simply write $\iota(k) = k$.[3] Now consider arbitrary $a \in \mathcal{A}$ and $k \in \mathbb{K}$. Then we have

$$\{a, k1\} = \{a, k\}1 + k\{a, 1\},$$

which implies that $k\{a, 1\} = 0$ (otherwise $\{a, k1\} \neq \{a, k\}$) and hence $\{a, 1\} = 0$.[4] But if $\{a, 1\} = 0$ then the $\mathbb{K}$-bilinearity of $\{-, -\}$ implies $k\{a, 1\} = \{a, k\} = 0$. So arbitrary $k \in \mathbb{K} \cdot 1$ satisfies $\{a, k\} = 0$ for all $a \in \mathcal{A}$, hence $\mathbb{K}1 \subset \mathrm{Cas}(\mathcal{A})$.
2. Suppose a is an integral element over $\mathrm{Cas}(\mathcal{A})$, i.e. $a \in \mathcal{A}$ and there is a monic polynomial $p(x) \in \mathrm{Cas}(\mathcal{A})[x]$ s.t. $p(a) = 0$. We need to show that $a \in \mathrm{Cas}(\mathcal{A})$. Without loss of generality, suppose p is the smallest degree polynomial s.t. $p(a) = 0$. If $\deg p = 1$, then $p(x) = x - a$. Hence $a \in \mathbb{K} \subset \mathrm{Cas}(\mathcal{A})$ by the previous statement. Suppose $\tau := \deg p > 1$. Using (13), for any $b \in \mathcal{A}$ we have

$$0 = \{p(a), b\} = p'(a)\{a, b\}\,.$$

Now $p'(a) \neq 0$, otherwise there exists $k \in \mathbb{K}$ s.t. $\frac{p'}{k}$ is a monic polynomial, which satisfies $\frac{p'}{k}(a) = 0$ and $\deg(\frac{p'}{k}) = \deg p' = \tau - 1$. But this would be a contradiction with the assumption of p being the smallest degree monic polynomial with the property $p(a) = 0$. Because $\mathcal{A}$ has no zero divisors, $\{a, b\} = 0$. As $b \in \mathcal{A}$ is arbitrary, this shows that $a \in \mathrm{Cas}(\mathcal{A})$.
3. Consider the left action $l\colon \mathcal{A} \to \mathrm{End}(\mathrm{End}(\mathcal{A}))$, $x \mapsto l_x$, where $l_x(a)$. Suppose l defines an A-module structure on $\mathrm{Ham}(\mathcal{A})$. Then for any $x \in \mathcal{A}$ and arbitrary $X_a \in \mathrm{Ham}(\mathcal{A})$, we have $l_x X_a \in \mathrm{Ham}(\mathcal{A})$. This means there is $d \in \mathcal{A}$ s.t. for every $b \in \mathcal{A} : l_x X_a(b) = X_d(b)$. But this is not true, in general, since

$$l_x X_a(b) = x\{b, a\} = \{b, xa\} - \{b, x\}a\,,$$

while

$$X_d(b) = \{b, d\}\,.$$

[3] In this notation, ka can be interpreted equivalently as action of $\mathbb{K}$ on $\mathcal{A}$, or as a multiplication in $\mathcal{A}$ after identifying $\mathbb{K}$ with $\iota(\mathbb{K})$. Of course we have $k1 = 1k = k$.

[4] Since for $a \neq 0 : ka = 0 \Rightarrow k = 0$.

In other words,

$$(l_x X_a - X_d)(b) = 0 \iff d = xa \text{ and } \{b, x\}a = 0\ , \tag{14}$$

for every $a, b, x \in \mathcal{A}$, which is obviously not always the case. Notice though, if we assume that $x \in \mathrm{Cas}(\mathcal{A})$, then $\{b, x\} = 0$ for every $b \in \mathcal{A}$ and thus (14) can always be satisfied. This shows that the action of $\mathcal{A}$ on $\mathrm{End}(\mathcal{A})$, if restricted to $\mathrm{Cas}(\mathcal{A})$, defines an action on $\mathrm{Ham}(\mathcal{A})$, i.e. statement 4 follows.

4. See the proof of statement 3.
5. By definition, $\varphi(a)(c) = -X_a(c) = -\{c, a\}$. We want to show that $\varphi(\{a, b\}) = [\varphi(a), \varphi(b)]$, where

$$[\varphi(a), \varphi(b)] = [X_a, X_b] := X_a \circ X_b - X_b \circ X_a\ .$$

From the Jacobi identity (4), we have for arbitrary $a, b, c \in \mathcal{A}$

$$-\{c, \{a, b\}\} = \{\{c, b\}, a\} - \{\{c, a\}, b\} = X_a(X_b(c)) - X_b(X_a(c))\ .$$

Since $\varphi(\{a, b\}) = -X_{\{a,b\}} = -\{c, \{a, b\}\}$, this shows that φ is a Lie algebra morphism.

6. The first map is an inclusion, $\iota\colon\ \mathrm{Cas}(\mathcal{A}) \to \mathcal{A}$, which is obviously a Lie algebra morphism. The second map is $\varphi\colon \mathcal{A} \to \mathrm{Ham}(\mathcal{A})$, and is given by $a \mapsto -X_a$. The previous statement shows that it is a Lie algebra morphism. Now consider $a \in \mathcal{A}$. Then $X_a = 0$ iff $\forall b \in \mathcal{A} : X_a(b) = \{b, a\} = 0$. Thus $\ker\varphi = \mathrm{im}\,\iota$, meaning that the short sequence

$$0 \xrightarrow{0} \mathrm{Cas}(\mathcal{A}) \xrightarrow{\iota} \mathcal{A} \xrightarrow{\varphi} \mathrm{Ham}(\mathcal{A}) \xrightarrow{0} 0$$

is exact.

$\square$

Remark 2.7 Corollary of the fifth statement of the above proposition is that $\mathrm{Ham}(\mathcal{A}) \subset \mathcal{D}(\mathcal{A})$ is a Lie subalgebra.

4.1 Exterior Algebra of a Commutative Algebra

In the following constructions, $\mathcal{A}$ is always commutative, associative, unital $\mathbb{K}$-algebra with the unit 1.

Module of Kahler Differentials of $\mathcal{A}$ Let $\Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A})$ be a $\mathbb{K}$-vector space generated by elements of the form $b\underline{\mathrm{d}}(a)$, for all $a, b \in \mathcal{A}$, satisfying relations

$$\underline{\mathrm{d}}(ab) = \underline{\mathrm{d}}(a)b + a\underline{\mathrm{d}}(b)\ .$$

where $\underline{\mathrm{d}}\colon \mathcal{A} \to \Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A})$ is an $\mathcal{A}$-linear map called the *universal derivation*. More precisely, for any derivation $\partial \in \mathcal{D}(\mathcal{A})$, there is a unique $\mathcal{A}$-linear map $\tilde{\partial}\colon \Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A}) \to \mathcal{A}$, such that the following diagram commutes

$$\begin{tikzcd} A \arrow[r, "\underline{\mathrm{d}}"] \arrow[dr, "\partial"'] & \Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A}) \arrow[d, dashed, "\exists!\tilde{\partial}"] \\ & \mathcal{A} \end{tikzcd} \tag{15}$$

The $\mathcal{A}$-module structure on $\Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A})$ is given by multiplication from the left. Using the universal property of $(\Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A}), \underline{\mathrm{d}})$, we have a canonical $\mathcal{A}$-module isomorphism

$$\mathcal{D}(\mathcal{A}) \cong \mathrm{Hom}_{\mathcal{A}}(\Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A}), \mathcal{A}) \ . \tag{16}$$

This $\mathcal{A}$-isomorphism can be understood as duality between $\mathcal{D}(\mathcal{A})$ and $\Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A})$. Moreover, it can be extended to polyderivations and exterior forms on the algebra $\mathcal{A}$.

Exterior Algebra of $\mathcal{A}$ Let $\Omega^\bullet(\mathcal{A})$ be the (graded) exterior algebra of the vector space of Kahler differentials of $\mathcal{A}$

$$\Omega^\bullet(\mathcal{A}) := \oplus_{k\geq 0}\Omega^k(\mathcal{A}) := \oplus_{k\geq 0}\Lambda^k(\Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A})),$$

together with the universal derivative $\underline{\mathrm{d}}\colon \mathcal{A} \to \Omega^1(\mathcal{A})$, which extends to

$$\underline{\mathrm{d}}\colon \Omega^k(\mathcal{A}) \to \Omega^{k+1}(\mathcal{A})$$

for all k. The universal derivative satisfies all the properties of the exterior differential on forms. In particular, $\underline{\mathrm{d}}$ is a boundary operator

$$\underline{\mathrm{d}}^2 = 0 \ ,$$

and acts with respect to the wedge product as

$$\underline{\mathrm{d}}(\omega_1 \wedge \omega_2) = \underline{\mathrm{d}}\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge \underline{\mathrm{d}}\omega_2 \ ,$$

where $\omega_1 \in \Omega^k(\mathcal{A})$ and $\omega_2 \in \Omega^\bullet(\mathcal{A})$ are arbitrary. The $\mathcal{A}$-duality (16) between $\mathcal{D}(\mathcal{A})$ and $\Omega_{\mathcal{A}/\mathbb{K}}$ can be written as

$$\mathcal{D}(\mathcal{A}) \times \Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A}) \to \mathcal{A} \ ,$$

$$< x, \underline{\mathrm{d}}a >:= (\underline{\mathrm{d}}a)(X) := X(a) \in \mathcal{A} \ ,$$

and may be extended to the duality

$$\Lambda^k(\mathcal{D}(\mathcal{A})) \cong \mathrm{Hom}_{\mathcal{A}}(\Omega^k(\mathcal{A}), \mathcal{A}) \ .$$

Moreover, one can define the symmetric algebra of derivations as

$$S(\mathcal{D}(\mathcal{A})) := \mathrm{Hom}_{\mathcal{A}}(S(\Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A}), \mathcal{A}) ,$$

which, as was already shown, is a graded Poisson algebra of degree 1.

Take $k = 2$, then we have

$$\Lambda^2(\mathcal{D}(\mathcal{A})) \cong \mathrm{Hom}_{\mathcal{A}}(\Omega^2(\mathcal{A}), \mathcal{A}) . \tag{17}$$

Suppose now that $\mathcal{A}$ is a Poisson algebra equipped with a Poisson bracket

$$\{-,-\}\colon \mathcal{A} \times \mathcal{A} \to \mathcal{A} .$$

Then the above duality (17) provides an element $\pi \in \Lambda^2(\mathcal{D}(\mathcal{A}))$ by the formula

$$< \pi, \underline{\mathrm{d}}a \wedge \underline{\mathrm{d}}b >= \{a, b\} , \tag{18}$$

for all $a, b \in \mathcal{A}$. Then π is called a *Poisson biderivation*. The Jacobi identity for the Poisson bracket $\{-,-\}$ imposes an important constraint on the biderivation π, which we will discuss below (see (22)).

5 Homology and Cohomology

We will discuss two types of (co)homologies that can be defined for an algebra $\mathcal{A}$. Firstly we define (co)homology of commutative, associative, unitary algebras (Hochschild), then for Poisson algebras (Lichnerowicz-Poisson).

5.1 Hochschild (Co)Homology

We use the following notation $\otimes = \otimes_{\mathbb{K}}$ and $\mathcal{A}^{\otimes k} = \underbrace{\mathcal{A} \otimes \ldots \otimes \mathcal{A}}_{k\text{-times}}$. We define $\mathcal{A}^{\otimes 0} := \mathbb{K}$.

Hochschild Chain Complex Let $\mathcal{A}$ be a commutative, associative, unitary $\mathbb{K}$-algebra, and let $\mathcal{M}$ be an $\mathcal{A}$-bimodule.[5] Then there is the following chain complex of $\mathcal{A}$-bimodules $\mathcal{M} \otimes \mathcal{A}^{\otimes k}$, $k \in \mathbb{N}$

$$0 \xleftarrow{0} \mathcal{M} \xleftarrow{d_1} \mathcal{M} \otimes \mathcal{A} \xleftarrow{d_2} \mathcal{M} \otimes \mathcal{A}^2 \xleftarrow{d_3} \ldots \tag{19}$$

[5] That is, $\mathcal{A}$ acts on $\mathcal{M}$ from left and right.

The boundary operators d_k, $k \in \mathbb{N}$, $k > 0$ are defined via the "face maps" ∂_i as

$$d_k := \sum_{i=0}^{k} (-1)^i \partial_i \ ,$$

where the maps ∂_i are

$$\partial_i(m \otimes a_1 \otimes \ldots \otimes a_k) := \begin{cases} ma_1 \otimes \ldots \otimes a_k, & \text{if } i = 0 \\ m \otimes a_1 \otimes \ldots \otimes a_i a_{i+1} \otimes \ldots \otimes a_k, & \text{if } 0 < i < k \\ a_k m \otimes a_1 \otimes \ldots \otimes a_{k-1}, & \text{if } i = k \end{cases}$$

for all $m \in \mathbf{M}$ and $a_i \in \mathcal{A}$. This is the *Hochschild chain complex*.

Exercise 2.3 Check that the above defined boundary maps d_k are

1. $\mathbb{K}$-multilinear,
2. well-defined,
3. satisfy ${d_k}^2 = 0$

for all $k \in \mathbb{N}_0$.

Definition 2.15 (Hochschild Homology) The homology of the chain complex (19), denoted $HH_*(\mathcal{A}, \mathcal{M})$, is called the Hochschild homology of $\mathcal{A}$ with coefficients in $\mathcal{M}$. The k-th Hochschild homology $\mathbb{K}$-module is

$$HH_k(\mathcal{A}, \mathcal{M}) := \ker d_k / im_{d_{k+1}} \ .$$

Hochschild Cochain Complex Using the $\mathrm{Hom}_{\mathbb{K}}$-functor, we can construct the Hochschild cochain complex

$$0 \xrightarrow{0} \mathcal{M} \xrightarrow{d^1} \mathrm{Hom}_{\mathbb{K}}(\mathcal{A}, \mathcal{M}) \xrightarrow{d^2} \mathrm{Hom}_{\mathbb{K}}(\mathcal{A}^{\otimes 2}, \mathcal{M}) \xrightarrow{d^3} \ldots \tag{20}$$

The coboundary operators d^k are defined as

$$d^k := \sum_{i=0}^{k} (-1)^i \partial^i \ ,$$

where the maps ∂^i are defined for $f \in \mathrm{Hom}_{\mathbb{K}}(\mathcal{A}^{\otimes k}, \mathcal{M})$

$$(\partial^i f)(a_0 \otimes a_1 \otimes \ldots \otimes a_k) := \begin{cases} a_0 f(a_1 \otimes \ldots \otimes a_k), & \text{if } i = 0 \\ f(a_0 \otimes a_1 \otimes \ldots \otimes a_i a_{i+1} \otimes \ldots \otimes a_k), & \text{if } 0 < i < k \\ f(a_0 \otimes a_1 \otimes \ldots \otimes a_{k-1}) a_k, & \text{if } i = k \end{cases}$$

where $a_i \in \mathcal{A}$ for all i. This is the *Hochschild cochain complex*.

Definition 2.16 (Hochschild Cohomology) The cohomology of the cochain complex (20), denoted $HH^*(\mathcal{A}, \mathcal{M})$, is called the Hochschild cohomology of $\mathcal{A}$ with coefficients in $\mathcal{M}$. The k-th Hochschild cohomology $\mathbb{K}$-module is

$$HH^k(\mathcal{A}, \mathcal{M}) := \ker d^k / \operatorname{im} d^{k-1} \ .$$

Proposition 2.2 *The* 0*-th Hochschild homology of* $\mathcal{A}$ *with coefficients in* $\mathcal{M}$ *satisfy*

$$HH_0(\mathcal{A}, \mathcal{M}) = \mathcal{M}/[\mathcal{M}, \mathcal{A}] \ ,$$

where $[\mathcal{A}, \mathcal{M}] = \{am - ma |\ a \in \mathcal{A}, m \in \mathcal{M}\}$. *In particular*

$$HH_0(\mathcal{A}, \mathcal{A}) = \mathcal{A}/[\mathcal{A}, \mathcal{A}] \ .$$

The 0*-th Hochschild cohomology of* $\mathcal{A}$ *with coefficients in* $\mathcal{M}$ *satisfy*

$$HH^0(\mathcal{M}, \mathcal{A}) = \{m \in M |\ am = ma,\ \forall a \in \mathcal{A}\} \ .$$

In particular,

$$HH^0(\mathcal{A}, \mathcal{A}) = \mathrm{Z}(A) \ ,$$

where $\mathrm{Z}(A)$ *is the center of* $\mathcal{A}$.

Proposition 2.3 *Let* $\mathcal{A}$ *be a commutative* $\mathbb{K}$*-algebra and* $\mathcal{M}$ *an* $\mathcal{A}$*-bimodule. Then the* 1*st Hochschild homology satisfies*

$$HH_1(\mathcal{A}, \mathcal{M}) \cong M \otimes_{\mathcal{A}} \Omega_{\mathcal{A}/\mathbb{K}}(\mathcal{A}) \ .$$

Denote by $\mathcal{D}(\mathcal{A}, \mathcal{M})$ the space of $\mathbb{K}$-linear functions $f : \mathcal{A} \to \mathcal{M}$ such that

$$f(ab) = af(b) + f(a)b \ .$$

Denote by $\mathcal{PD}(\mathcal{A}, \mathcal{M})$ the space of $\mathbb{K}$-linear functions $f_m : \mathcal{A} \to \mathcal{M}$, which are given by

$$f_m(a) = ma - am \ .$$

Exercise 2.4 Check that the above defined f_m satisfies $f_m \in \mathcal{D}(\mathcal{A}, \mathcal{M})$.

Proposition 2.4 $HH^1(\mathcal{A}, \mathcal{M}) = \mathcal{D}(\mathcal{A}, \mathcal{M})/\mathcal{PD}(\mathcal{A}, \mathcal{M})$.

5.2 Lichnerowicz-Poisson Cohomology

Let $\mathcal{A}$ be a commutative Poisson algebra with the Poisson bracket $\{-,-\}$ and consider $\Lambda(\mathcal{D}(\mathcal{A})) = \bigoplus_{k\geq 0} \Lambda^k(\mathcal{D}(\mathcal{A}))$ defined as follows. For $k = 0$, we define $\Lambda^0(\mathcal{D}(\mathcal{A})) = \mathcal{A}$. An element $X \in \Lambda^k(\mathcal{D}(\mathcal{A})), k > 0$, is a multilinear, antisymmetric mapping

$$X : \mathcal{A}^k \to \mathcal{A}$$

and the mapping $\partial_X : \mathcal{A} \to \mathcal{A}$ given by

$$\partial_X(a) := X(a, a_1, \ldots, a_{k-1})$$

is a derivation.

Schouten-Nijenhuis Bracket Let $X, Y \in \Lambda(\mathcal{D}(\mathcal{A}))$ be decomposable,

$$X = x_1 \wedge \ldots \wedge x_k \qquad\qquad Y = y_1 \wedge \ldots \wedge y_l\ ,$$

where all $x_i, y_j \in \mathcal{D}(\mathcal{A})$. The Schouten-Nijenhuis bracket $[\![-,-]\!]$ is given by

$$[\![X, Y]\!] := \sum_{i,j} (-1)^{i+j} [x_i, y_j] x_1 \wedge \ldots \wedge \hat{x_i} \wedge \ldots \wedge x_k \wedge y_1 \wedge \ldots \wedge \hat{y_j} \wedge \ldots \wedge y_l\ ,$$

where $[-,-]$ is the commutator of differential operators, and the above definition is extended linearly on the whole $\Lambda(\mathcal{D}(\mathcal{A}))$.

Exercise 2.5 Show that the Schouten-Nijenhuis bracket $[\![-,-]\!]$ satisfies for all $P, Q \in \Lambda(\mathcal{D}(\mathcal{A}))$

$$[\![P, Q]\!] = (-1)^{\deg P \deg Q} [\![Q, P]\!] \tag{21}$$

Poisson Operator Let $\pi \in \Lambda^2(\mathcal{D}(\mathcal{A}))$ be the *Poisson biderivation*, defined by (18). Then the Schouten-Nijenhuis bracket gives an element $[\![\pi, \pi]\!] \in \Lambda^3(\mathcal{D}(\mathcal{A}))$.

Proposition 2.5 *The Jacobi identity for the Poisson bracket* (45) *implies the triviality of the* 3*-derivation:*

$$[\![\pi, \pi]\!] = 0\ . \tag{22}$$

Exercise 2.6 Prove the above proposition.

Equation (22) is sometimes called *Poisson Master equation*. It can be interpreted as a nilpotency condition (${\delta_\pi}^2 = 0$) for the operator $\delta_\pi : \Lambda(\mathcal{D}(\mathcal{A})) \to \Lambda(\mathcal{D}(\mathcal{A}))$, defined by

$$\delta_\pi X = [\![\pi, X]\!] \,.$$

The operator δ_π is called *Lichnerowicz-Poisson operator* and leads to the following notion of cohomology.

Definition 2.17 (Lichnerowicz-Poisson Cohomology) The cohomology of the chain complex

$$\ldots \xrightarrow{\delta_\pi} \Lambda^{k-1}(\mathcal{D}(\mathcal{A})) \xrightarrow{\delta_\pi} \Lambda^{k}(\mathcal{D}(\mathcal{A})) \xrightarrow{\delta_\pi} \Lambda^{k+1}(\mathcal{D}(\mathcal{A})) \xrightarrow{\delta_\pi} \ldots \,,$$

is called the Lichnerowicz-Poisson cohomology, denoted $HP^*(\mathcal{A})$.

In terms of the Poisson bracket, the Lichnerowicz-Poisson operator can be rewritten as

$$(\delta_\pi X)\,(a_0, a_1, \ldots, a_k) = \sum_i (-1)^i \{a_i, X(a_0, a_1, \ldots, \hat{a_i}, \ldots, a_k)\} \tag{23}$$

$$+ \sum_{0 \le i < j} X(\{a_i, a_j\}, a_1, \ldots, \hat{a_i}, \ldots, \hat{a_j}, \ldots, a_k) \,. \tag{24}$$

5.3 *Low-Dimensional Poisson Cohomology*

$k = 0$: the operator $\delta_\pi : \mathcal{A} \to \mathcal{D}(\mathcal{A})) \cong \Lambda^1(\mathcal{D}(\mathcal{A}))$ acts as $a \mapsto \partial_a = \{-, a\}$, where $a \in \mathcal{A}$ can be seen as a Hamiltonian for a Hamiltonian derivation $\{-, a\}$. We have

$$HP^0 \cong \mathrm{Cas}(\mathcal{A})$$

since $\delta_\pi(a)(b) = 0$ for all $b \in \mathcal{A}$ if and only if $a \in \mathrm{Cas}(\mathcal{A})$. Elements of $HP^0 \cong \mathrm{Cas}(\mathcal{A})$ are called Hamiltonians with zero dynamics.

$k = 1$: 1-coboundary is a derivation $X \in \mathcal{D}(\mathcal{A})$ which is a Hamiltonian derivation ∂_a for some element $a \in \mathcal{A}$. 1-cocycle is an element $X \in \mathcal{D}(\mathcal{A})$ s.t.

$$\delta_\pi(X) = 0 \Rightarrow [\![\pi, X]\!] = 0 \,.$$

Derivations $X \in \mathcal{D}(\mathcal{A})$ which satisfies $\delta_\pi(X) = 0$ are called *Poisson, or canonical, derivations*. The set of all such elements is denoted $\mathrm{Can}(\mathcal{A})$. The equation $[\![\pi, X]\!] = 0$ represents a conservation of the Poisson structure along X. If we denote $L_X\pi = [\![\pi, X]\!]$, one can easily show that

$$< L_X\pi, \underline{\mathrm{d}}a \wedge \underline{\mathrm{d}}b >= L_X < \pi, \underline{\mathrm{d}}a \wedge \underline{\mathrm{d}}b > \,.$$

Theorem 2.1 (Basic Theorem of Classical Mechanics)

$$HP^1(\mathcal{A}) \cong \mathrm{Can}(\mathcal{A})/\,\mathrm{Ham}(\mathcal{A}) ,$$

where $\mathrm{Ham}(\mathcal{A})$ *is given by* (11).

Example (Hamiltonian Derivations on Polynomial Algebra [3]) Consider $\mathcal{A} = \mathbb{K}[x_1, x_2]$ with the Poisson bracket $\{x_1, x_2\} = x_2$. Then $\partial \in \mathrm{Ham}(\mathcal{A}) \subset \mathcal{D}(\mathcal{A})$ if

$$\partial = \alpha_1 \frac{\partial}{\partial x_1} + \alpha_2 \frac{\partial}{\partial x_2} \tag{25}$$

and exists $h \in \mathcal{A}$ such that

$$\partial = \partial_h = \{-, h\} .$$

We will compute the coefficients α_i so that the above equation holds. We have

$$\{x_1, h\} = \frac{\partial x_1}{\partial x_1}\frac{\partial h}{\partial x_2} - \frac{\partial x_1}{\partial x_2}\frac{\partial h}{\partial x_1} = \frac{\partial h}{\partial x_2} .$$

Similarly we get

$$\{x_2, h\} = -\frac{\partial h}{\partial x_1} .$$

For $f \in \mathcal{A}$, the bracket with h is

$$\{f, h\} = \frac{\partial f}{\partial x_1}\{x_1, h\} + \frac{\partial f}{\partial x_2}\{x_2, h\} = \frac{\partial f_1}{\partial x_1}\frac{\partial h}{\partial x_2} - \frac{\partial f_1}{\partial x_2}\frac{\partial h}{\partial x_1} .$$

From (25) we have

$$\partial(f) = \alpha_1 \frac{\partial f}{\partial x_1} + \alpha_2 \frac{\partial f}{\partial x_2} ,$$

so to have $\partial = \partial_h$, one must have

$$\alpha_1 = \frac{\partial h}{\partial x_2} , \qquad \alpha_2 = -\frac{\partial h}{\partial x_1} .$$

This means that the set of Hamiltonian derivations on $\mathcal{A}$ is given by

$$\mathrm{Ham}(\mathcal{A}) = \{\frac{\partial h}{\partial x_2}\frac{\partial}{\partial x_1} - \frac{\partial h}{\partial x_1}\frac{\partial}{\partial x_2} |\ h \in \mathcal{A}\} .$$

Exercise 2.7 Consider $\mathcal{A}$ from the previous example. Show that

$$\mathrm{Can}(\mathcal{A}) / \mathrm{Ham}(\mathcal{A}) \cong \mathbb{K}$$

Indication: check that if $\partial = \alpha_1 \frac{\partial}{\partial x_1} + \alpha_2 \frac{\partial}{\partial x_2} \in \mathrm{Can}(\mathcal{A})$, then

$$\alpha_1 = c + x_2 \frac{\partial h}{\partial x_2}$$
$$\alpha_2 = -x_2 \frac{\partial h}{\partial x_1}$$

where $c \in \mathbb{K}$.

Example Consider the set

$$\mathrm{Sym}(\mathrm{Hom}_{\mathbb{K}}(\mathcal{D}^p(\mathcal{A}), \mathcal{A})) = \{\varphi\colon \mathcal{D}^p(\mathcal{A}) \to \mathcal{A} \mid \varphi \text{ multilinear}\} .$$

An element of the above set is not only $\mathbb{K}$-linear in all p-arguments but also a $\mathbb{K}$-derivation in every argument. Recall that $\mathcal{D}(\mathcal{A})$ is the $\mathcal{A}$-module of $\mathbb{K}$-derivations of $\mathcal{A}$.

5.3.1 Compatible Poisson Structures

As a useful application of Lichnerowicz-Poisson cohomology, consider the following description of *compatible Poisson structures*.

Definition 2.18 Two Poisson structures given by biderivations π and θ are compatible if for all $\lambda \in \mathbb{K}$, $\pi + \lambda\theta$ is again a Poisson structure.

As an immediate consequence of the above definition, we have the following proposition.

Proposition 2.6 *If two Poisson structures $\pi, \theta \in \Lambda^2 (\mathcal{D}(\mathcal{A}))$ are compatible, then*

$$[\![\pi, \theta]\!] = 0$$

Equivalently, if π and θ are compatible, then they are closed with respect to coderivations in the following sense

$$\delta_\pi \theta = \delta_\theta \pi = 0 .$$

Exercise 2.8 Prove the Proposition 2.6.

Notice that if $[\![\pi, \theta]\!] = 0$ then for all $\lambda \in \mathbb{K}$

$$[\![\pi + \lambda\theta, \pi + \lambda\theta]\!] = 0 \ .$$

Let us fix a biderivation π. Consider the *Lie derivative* of π along $X \in \mathcal{D}(\mathcal{A})$

$$L_X\pi := [\![\pi, X]\!] \in \Lambda^2 (\mathcal{D}(\mathcal{A})) \ .$$

Let π be a Poisson biderivation (i.e. it corresponds to a Poisson structure). Then $L_X\pi$ is called an *infinitesimal deformation of the Poisson structure*. Now consider a deformation of π of the form

$$\pi \mapsto \pi + L_X\pi \ .$$

By assumption, we have $[\![\pi, \pi]\!] = 0$ since π yields a Poisson structure. If we assume that the deformation also yields a Poisson structure, then the graded antisymmetry of the bracket (see (21) yields

$$0 = [\![\pi + L_X\pi, \pi + L_X\pi]\!] = 2[\![\pi, L_X\pi]\!] + [\![L_X\pi, L_X\pi]\!] \ .$$

Let us denote $\gamma := L_X\pi$ and recall that $[\![\theta, X]\!] = \delta_\theta(X)$. Then the above can be written as

$$\delta_\pi\gamma + \frac{1}{2}[\![\gamma, \gamma]\!] = 0 \ .$$

This equation is called *Maurer-Cartan equation* (for a differential graded algebra $\Lambda (\mathcal{D}(\mathcal{A}))$).

5.3.2 Interpretation of $HP^2(\mathcal{A})$

Let $\pi \in \Lambda^2 (\mathcal{D}(\mathcal{A}))$ and consider the image of $\delta_\pi : \Lambda^1 (\mathcal{D}(\mathcal{A})) \to \Lambda^2 (\mathcal{D}(\mathcal{A}))$, given by

$$B_\pi{}^2 = \{\theta \in \Lambda^2 (\mathcal{D}(\mathcal{A})) \mid \theta = [\![\pi, X]\!] \text{ for some } X \in \mathcal{D}(\mathcal{A})\} \ ,$$

The kernel of $\delta_\pi : \Lambda^2 (\mathcal{D}(\mathcal{A})) \to \Lambda^3 (\mathcal{D}(\mathcal{A}))$ is

$$Z_\pi{}^2 = \{\theta \in \Lambda^2 (\mathcal{D}(\mathcal{A})) \mid [\![\pi, \theta]\!] = 0]\} \ .$$

By definition

$$HP^2(\mathcal{A}) = Z_\pi{}^2 / B_\pi{}^2 \ .$$

We have the following proposition, which yields an interpretation of the second Lichnerowicz-Poisson cohomology.

Proposition 2.7 ([6]) *If* $\pi \in \Lambda^2(\mathcal{D}(\mathcal{A}))$ *is a Poisson biderivation* ($[\![\pi,\pi]\!] = 0$) *and* $HP^2(\mathcal{A}) = 0$*, then the set of all structures compatible with* π

$$\mathrm{Comm}(\pi) := \{\theta \in \Lambda^2(\mathcal{D}(\mathcal{A})) \mid [\![\theta,\theta]\!] = [\![\pi,\theta]\!] = 0\} \tag{26}$$

is a set of infinitesimal deformations of π *along* $X \in \mathcal{D}(\mathcal{A})$*. That is,* $\theta = L_X\pi$ *such that*

$$L_X{}^2(\theta) = L_Y\pi$$

for some $Y \in \mathcal{D}(\mathcal{A})$*, where* $L_X{}^2 = L_X \circ L_X$*.*

Proof Let $\theta \in \mathrm{Comm}(\pi)$, then $\delta_\pi\theta = 0$ so

$$\mathrm{Comm}(\pi) \subset \ker\delta_\pi \ .$$

If $\theta = \delta_\pi X$ then $0 = [\theta] \in HP^2(\mathcal{A})$. Suppose $\theta \in \mathrm{Comm}(\pi)$ and $HP^2(\mathcal{A}) = 0$. Then $\theta = \delta_\pi X = [\![\pi, X]\!] = L_X\pi$. But at the same time $[\![\theta,\theta]\!] = [\![L_X\pi, L_X\pi]\!] = 0$. This is equivalent to $(L_X \circ L_X)(\theta) \in \ker\delta_\pi$ since

$$L_X(L_X\pi) = L_X([\![\pi, X]\!]) = [\![[\![\pi, X]\!], X]\!] = [\![\theta, X]\!] \ ,$$

and

$$\delta_\pi([\![\theta, X]\!]) = \delta_\pi(\delta_\theta X) = -\delta_\theta(\delta_\pi X) = [\![\delta_\pi X, \delta_\pi X]\!] = 0 \ .$$

Hence if $HP^2(\mathcal{A}) = 0$ then there is $Y \in \mathcal{D}(\mathcal{A})$ such that $L_X{}^2\theta = L_Y(\pi) = \delta_\pi(Y)$. □

From the above, we obtain a mapping

$$\tau : \ \mathrm{Comm}(\pi) \to HP^2(\mathcal{A})$$

and

$$\tau(\theta) = 0 \iff \theta = \delta_\pi(X)$$

for some $X \in \mathcal{D}(\mathcal{A})$. The image of τ is described by the following proposition.

Proposition 2.8 *Let* $\partial \in HP^2(\mathcal{A})$*. Then*

$$\partial \in \operatorname{im}\tau \iff \partial = \Delta + \delta_\pi(\mathcal{D}(\mathcal{A})) \ ,$$

where the representative $\Delta \in \Lambda^2(\mathcal{D}(\mathcal{A}))$ *of the class* ∂ *satisfies*

1. $[\![\partial, \partial]\!] = [\![\Delta, \Delta]\!] = \partial_\pi(\alpha)$*, where* $\alpha \in \Lambda^2(\mathcal{D}(\mathcal{A}))$*,*
2. there is $X \in \mathcal{D}(\mathcal{A})$ *such that* $\alpha + 2L_X(\Delta) - {L_X}^2(\pi) = \ker \delta_\pi$.

Proof Let $\Delta \in \Lambda^2(\mathcal{D}(\mathcal{A}))$ be such that

$$\partial = \Delta + \delta_\pi(\mathcal{D}(\mathcal{A})) \in HP^2(\mathcal{A}) .$$

If $\partial \in \operatorname{im}\tau$, then there is a $\theta \in \operatorname{Comm}(\pi)$ (see (26)) such that

$$\tau(\theta) = \Delta + \delta_\pi(\mathcal{D}(\mathcal{A})) .$$

Because

$$0 = [\![\partial, \partial]\!] \Rightarrow [\![\Delta, \Delta]\!] = \partial_\pi(\alpha) \in HP^3(\mathcal{A})$$

and $\tau[\![\theta, \theta]\!] = 0$, we have

$$\begin{aligned}
[\![\tau(\theta), \tau(\theta)]\!] &= [\![\Delta + \delta_\pi(X), \Delta + \delta_\pi(X)]\!] \\
&= [\![\Delta, \Delta]\!] + [\![\Delta, \delta_\pi(X)]\!] + [\![\delta_\pi(X), \Delta]\!] + [\![\delta_\pi(X), \delta_\pi(X)]\!] \\
&= \delta_\pi(\alpha) - [\![\Delta, L_X\pi]\!] - [\![L_X\pi, \Delta]\!] - \delta_\pi {L_X}^2(\pi) \\
&= \delta_\pi\left(\alpha - {L_X}^2(\pi)\right) - 2[\![L_X\pi, \Delta]\!] \\
&= \delta_\pi\left(\alpha - {L_X}^2(\pi)\right) - 2[\![[\![\pi, X]\!], \Delta]\!] \\
&= \delta_\pi\left(\alpha + 2L_X(\Delta) - {L_X}^2(\pi)\right) = 0 ,
\end{aligned}$$

since $[\![\tau(\theta), \tau(\theta)]\!] = 0$. □

5.4 Poisson Homology

The following construction is due to Brylinski [7] and [8]. Let $\mathcal{A}$ be a commutative $\mathbb{K}$-algebra and $\Omega(\mathcal{A}) = \Lambda(\Omega_{\mathcal{A}/\mathbb{K}})$ be the exterior algebra of $\mathcal{A}$. The boundary morphism $d_\pi : \Omega^k(\mathcal{A}) \to \Omega^{k-1}(\mathcal{A})$ is defined by the "homotopy-like" formula

$$d_\pi\omega = (i_\pi \circ \underline{\mathrm{d}})(\omega) - (\underline{\mathrm{d}} \circ i_\pi)(\omega) , \tag{27}$$

where i_π is the contraction operator with respect to π, i.e $.i_\pi := < \pi, - >$. It is straightforward to check that this operator satisfies $d_\pi \circ d_\pi = 0$. In coordinates, we can describe d_π on a decomposable $\omega = a_0, \underline{\mathrm{d}}a_1 \wedge \ldots \wedge a_k$ by the general formula

$$d_\pi \omega = \sum_{l=1}^{k+1} (-1)^{l+1} \{a_0, a_i\} \wedge \underline{d}a_1 \wedge \ldots \wedge \hat{\underline{d}a_l} \wedge \ldots \wedge \underline{d}a_{k+1}$$
$$- \sum_{i<j} (-1)^{i+j} \underline{d}\left(\{a_i, a_j\}\right) \wedge \underline{d}a_1 \wedge \ldots \wedge \hat{\underline{d}a_i} \wedge \ldots \wedge \hat{\underline{d}a_j} \wedge \ldots \wedge \underline{d}a_{k+1} \ .$$

By definition,

$$d_\pi (a_0 da_1) = i_\pi (da_0 \wedge da_1) = \{a_0, a_1\} \ .$$

Note that the Jacobi identity for the triple a_0, a_1, a_2

$$\{a_0, \{a_1, a_2\}\} + \{a_1, \{a_2, a_O\}\} + \{a_2, \{a_0, a_1\}\} = 0$$

implies $(d_\pi)^2 = 0$.

5.5 Duality

Let π be a *symplectic* (or *non-degenerate*) Poisson structure on $\mathcal{A}$, meaning that the Hamiltonian map

$$\Gamma_\pi : \Omega^1_{\mathcal{A}/\mathbb{K}} \to \mathcal{D}(\mathcal{A})$$

defined by

$$\Gamma_\pi(\alpha) = < \pi, \alpha > \in \mathcal{D}(\mathcal{A})$$

is an isomorphism, and there exists the inverse ${\Gamma_\pi}^{-1} \colon \mathcal{D}(\mathcal{A}) \to \Omega^1(\mathcal{A})$. The inverse is given by

$${\Gamma_\pi}^{-1}(\partial) = \alpha_\partial \text{ such that } < \pi, \alpha_\partial > = \partial \ .$$

In this case, one can check that

$$\delta_\pi = \Gamma_\pi \circ \underline{d} \circ {\Gamma_\pi}^{-1} \ ,$$

where δ_π is the Lichnerowicz-Poisson operator. The above equation is symbolical. It can be described more precisely by the following commutative diagram

$$\begin{CD} \Omega^1(\mathcal{A}) @>{\Gamma_\pi}>> \Lambda^1\,(\mathcal{D}(\mathcal{A})) \\ @V{\underline{\mathrm{d}}}VV @VV{\delta_\pi}V \\ \Omega^2(\mathcal{A}) @>{\Lambda^2\Gamma_\pi}>> \Lambda^2\,(\mathcal{D}(\mathcal{A})) \end{CD}$$

which holds for any Poisson biderivation π.

Proposition 2.9 *For general k, the following diagram commutes*

$$\begin{CD} \Omega^k(\mathcal{A}) @>{\Lambda^k\Gamma_\pi}>> \Lambda^k\,(\mathcal{D}(\mathcal{A})) \\ @V{\underline{\mathrm{d}}}VV @VV{\delta_\pi}V \\ \Omega^{k+1}(\mathcal{A}) @>{\Lambda^{k+1}\Gamma_\pi}>> \Lambda^{k+1}\,(\mathcal{D}(\mathcal{A})) \end{CD}$$

Proof To check the commutativity of the diagram, consider decomposable $\omega = a_0\underline{\mathrm{d}}a_1 \wedge \ldots \wedge \underline{\mathrm{d}}a_k \in \Omega^k(\mathcal{A})$. Then

$$\Lambda^{k+1}\Gamma_\pi\,(\underline{\mathrm{d}}\omega) = \Lambda^{k+1}\Gamma_\pi\,(\underline{\mathrm{d}}a_0 \wedge \underline{\mathrm{d}}a_1 \wedge \ldots \wedge \underline{\mathrm{d}}a_k) = \partial_{a_0} \wedge \partial_{a_1} \wedge \ldots \wedge \partial_{a_k}\ ,$$

and

$$\begin{aligned} \delta_\pi\,(\Lambda^k\Gamma_\pi\,(\omega)) &= \delta_\pi\,(a_0\partial_{a_1} \wedge \ldots \wedge \partial_{a_k}) \\ &= -[\![a_0\partial_{a_1} \wedge \ldots \wedge \partial_{a_k}, \pi]\!] \\ &= -[\![a_0, \pi]\!] \wedge \partial_{a_1} \wedge \ldots \wedge \partial_{a_k} \\ &= \partial_{a_0} \wedge \partial_{a_1} \wedge \ldots \wedge \partial_{a_k}\ . \end{aligned}$$

Since the above can be extended linearly, the diagram commutes. □

Remark 2.8 If $\mathcal{A} = C^\infty(M)$ is the algebra of smooth function on a smooth manifold M, then it was shown in [7] that

$$HP_k(\mathcal{A}) \cong HP^{2n-k}(\mathcal{A})\ .$$

Moreover, by the Proposition 2.9, the latter group is isomorphic with the de Rham cohomology

$$HP^{2n-k}(\mathcal{A}) = H^{2n-k}_{DR}(M)\ .$$

In the following lemma, we use the notion of a *graded commutator* in a graded algebra

$$[\partial_1, \partial_2] := \partial_1 \circ \partial_2 - (-1)^{\deg \partial_1 \deg \partial_2}\partial_2 \circ \partial_1\ .$$

We have $\deg d_\pi = -1$, $\deg \underline{\mathrm{d}} = 1$, $\deg i_\pi = -2$, $L_\pi = 1$

Lemma 2.2 *The operator $d_\pi : \Omega^k(\mathcal{A}) \to \Omega^{k-1}(\mathcal{A})$ commutes in graded sense with $\underline{\mathrm{d}}$ and i_π, i.e.*

$$[d_\pi, \underline{\mathrm{d}}] = d_\pi \circ \underline{\mathrm{d}} + \underline{\mathrm{d}} \circ d_\pi = 0 ,$$
$$[d_\pi, i_\pi] = d_\pi \circ i_\pi - i_\pi \circ d_\pi = 0 .$$

Proof To prove the first equation, we use the definition of d_π (see (27)) and that $\underline{\mathrm{d}} \circ \underline{\mathrm{d}} = 0$. Then

$$(i_\pi \circ \underline{\mathrm{d}} - \underline{\mathrm{d}} \circ i_\pi) \circ \underline{\mathrm{d}} + \underline{\mathrm{d}} \circ (i_\pi \circ \underline{\mathrm{d}} - \underline{\mathrm{d}} \circ i_\pi) = -\underline{\mathrm{d}} \circ i_\pi \circ \underline{\mathrm{d}} + \underline{\mathrm{d}} \circ i_\pi \circ \underline{\mathrm{d}} = 0$$

Hence $[d_\pi, \underline{\mathrm{d}}] = 0$. Similarly for the second equation

$$[d_\pi, i_\pi] = [L_\pi, i_\pi] = i_{[\pi,\pi]} = 0 .$$

□

Remark 2.9 When $\mathcal{A} = C^\infty(M)$, where M is a smooth Poisson manifold, we will see all the above, and more general, identities in the later section as well.

Example (0th Poisson Homology) For the kernel of $d^0_\pi : \Omega^0(\mathcal{A}) \cong \mathcal{A} \to 0$ we obviously have $\ker d^0_\pi \cong \mathcal{A}$. For the image of $d^1_\pi : \Omega^1(\mathcal{A}) \to \Omega^0(\mathcal{A}) \cong \mathcal{A}$ we have $\operatorname{im} d^1_\pi = \{\mathcal{A}, \mathcal{A}\}$. This is because $\Omega^1(\mathcal{A})$ consists of element $a_0 da_1$, where $a_0, a_1 \in \mathcal{A}$ are arbitrary, and

$$d^1_\pi(a_0 da_1) = \{a_0, a_1\} \in \{\mathcal{A}, \mathcal{A}\} .$$

Hence

$$HP_0(\mathcal{A}) \cong \mathcal{A}/\{\mathcal{A}, \mathcal{A}\}$$

There is no simple interpretation for higher homology groups, $k > 0$.

6 Polynomial Poisson Algebras

Let $\mathcal{A} = \mathbb{C}[x_1, \dots, x_n]$ be the polynomial algebra over complex numbers. If x_i is a generator of $\mathcal{A}$, then $\partial_i := \frac{\partial}{\partial x_i} \in \mathcal{D}(\mathcal{B})$. The vector operator $\nabla := (\partial_1, \dots, \partial_n)$ is called *gradient* One can define the *Jacobian matrix* of n elements $f_1, \dots, f_n \in \mathcal{A}$ as

$$\mathrm{Jac}(f_1, \ldots, f_n) := \left(\frac{\partial f_i}{\partial x_j} \right).$$

The determinant det(Jac($f_1, \ldots, f_n$)) is called *Jacobian*. It is clear the i-th row of $\mathrm{Jac}(f_1, \ldots, f_n)$ is $\nabla(f_i)$.

6.1 Nambu-Jacobi-Poisson Algebras

Now we fix $f_1, \ldots, f_{n-2} \in \mathcal{A}$ and define the following bilinear operation $\{-,-\} \colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$, which yields a Poisson algebra structure on the polynomial algebra $\mathcal{A}$, called *Nambu-Poisson-Jacobi structure*.

Definition 2.19 The *Nambu-Jacobi-Poisson bracket* of $F, G \in \mathcal{A}$ is

$$\{F, G\} := \det \mathrm{Jac}(F, G, f_1, \ldots, f_{n-2}) \in \mathcal{A} . \tag{28}$$

When $\mathcal{A} = \mathbb{C}[x_1, x_2, x_3]$, there is only one f determining the Nambu-Jacobi-Poisson bracket, and we will denote the bracket by $\{-,-\}_f$. We proceed with the following elementary lemma.

Lemma 2.3 *For* $1 \leq i \leq n$ *and* $f_1, \ldots, f_{i-1}, f_{i+1}, \ldots, f_n \in \mathcal{A}$*, the operation* $D \colon \mathcal{A} \to \mathcal{A}$ *given by* $D(g) := \mathrm{Jac}(f_1, \ldots, f_{i-1}, g, f_{i+1}, \ldots, f_n)$ *is a derivation of* $\mathcal{A}$*.*

Theorem 2.2 *The bracket* (28) *is a Poisson bracket on* $\mathcal{A}$*.*

Proof To prove this theorem for any n, we observe that skew symmetry is evident from the skew-symmetry of det, the Leibniz rule follows from the Lemma 2.3, so the only non-trivial statement is, as usual, the Jacobi identity. But this follows from the Fundamental identity for the Nambu bracket $\{f_1, \ldots, f_n\} := \mathrm{Jac}(f_1, \ldots, f_n)$ (see (29) below). □

An interesting property of the algebraic bracket structure on the polynomial algebra is given by the following theorem.

Proposition 2.10 *If* $f_1, \ldots, f_{n-2}$ *are algebraically dependent over* $\mathbb{C}$*, then* $\{-,-\} = 0$*.*

Proof The proof is left as an easy exercise for the reader. □

Remark 2.10 If we consider F, G as rational functions, the result of bracket (28) is still a polynomial. Consider a quotient by (convenient) ideal

$$\mathcal{B} := \mathbb{C}[x_1, \ldots, x_n] / < p_1, \ldots, p_k > ,$$

where $p_i \in \mathcal{A}$. Then the Theorem 2.2 still holds, i.e. the bracket (28) yields a Poisson algebra structure on $\mathcal{B}$. This is valid in an even more general setup of power series rings [9].

The Nambu-Jacobi-Poisson bracket is a special case of a $(n-m)$-ary operation

$$\{F_1, \dots, F_{n-m}\} := \lambda \det \mathrm{Jac}(F_1, \dots, F_{n-m}, f_1, \dots, f_m) \in \mathcal{A} ,$$

where $\lambda, F_i \in \mathcal{A}$ for all i. The above multibracket $\{-, \dots, -\} \colon \mathcal{A}^{\otimes(n-m)} \to \mathcal{A}$ satisfies for every permutation σ the antisymmetry condition

$$\{F_1, \dots, F_{n-m}\} = (-1)^{\sigma}\{F_{\sigma(1)}, \dots, F_{\sigma(n-m)} .$$

It also satisfies the Leibniz rule in every argument

$$\{hF_1, \dots, F_{n-m}\} = F_1\{h, \dots, F_{n-m}\} + h\{F_1, \dots, F_{n-m}\} ,$$

and the *Fundamental identity*

$$\begin{aligned} \sum_k \{G_1, \dots, G_{k-1}, \{F_1, \dots, F_{n-m-1}, G_k\}, G_{k+1}, \dots, G_{n-m}\} = \\ = \{F_1, \dots, F_{n-m-1}, \{G_1, \dots, G_{n-m}\}\} , \end{aligned} \tag{29}$$

which is a generalization of the Jacobi identity (the Jacobi identity and Nambu-Jacobi-Poisson bracket is restored in the case $n - m = 2$). For more details about the Nambu structures and their generalizations, see [9].

Consider now a $(n-2) \times n$ matrix over $\mathcal{A}$

$$M = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n-2,1} & \dots & 0a_{n-2,n} \end{pmatrix} .$$

Suppose $i \neq j$ and denote by $\hat{M}_{ij}$ the matrix given by deleting the i-th and j-th column (thus the result being $(n-2) \times (n-2)$ matrix. If we choose

$$M = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_{n-2}}{\partial x_1} & \dots & \frac{\partial f_{n-2}}{\partial x_n} \end{pmatrix} ,$$

then for the generators of $\mathcal{A}$ we have

$$\{x_i, x_j\} = (-1)^{i+j-1} \det \hat{M}_{ij}$$

and we define $\{x_i, x_i\} = 0$.

Example Let $n = 3$, $f = \frac{1}{3}({x_1}^3 + {x_2}^3 + {x_3}^3) + \tau x_1 x_2 x_3$, where $\tau \in \mathbb{C}$. Then $\{x_i, x_j\} = \mathrm{Jac}(x_i, x_j, f) = \tau x_i x_j + {x_k}^2$, where (i, j, k) is a permutation of $\{1, 2, 3\}$.

Example (Sklyanin Elliptic Poisson Brackets) Let $n = 4$ and consider $f_1 = q_1(x_1, x_2, x_3, x_4)$, $f_2 = q_2(x_1, x_2, x_3, x_4)$, where q_1, q_2 are *quadratic* polynomials. Choosing $q_1 = {x_1}^2 + {x_2}^2 + {x_3}^2 + {x_4}^2$ and $q_2 = \alpha {x_2}^2 + \beta {x_3}^2 + \gamma {x_4}^2$ such that $\alpha\beta\gamma + \alpha + \beta + \gamma \neq 0$, we obtain the original Sklyanin-Poisson structure [9–11].

6.2 Poisson-Calabi-Yau Algebra

This is a Jacobian algebra $\mathcal{A} = \mathbb{C}[x_1, x_2, x_3]$ with $f = -{x_1}^2 x_3$ [12]. The Nambu-Jacobi-Poisson bracket (28) is

$$\{x_1, x_2\}_f = -{x_1}^2, \qquad \{x_2, x_3\}_f = -2x_1x_3, \qquad \{x_1, x_3\}_f = 0\,.$$

It is interesting that the algebra $\Omega^1(\mathcal{A}) = \mathcal{A} < \mathrm{d}x_1, \mathrm{d}x_2, \mathrm{d}x_3 >$ (think of a free algebra over A) is also a Poisson algebra with

$$\begin{aligned}
\{\mathrm{d}x_1, \mathrm{d}x_2\}_\Omega &= \mathrm{d}\{x_1, x_2\}_f = -2x_1\mathrm{d}x_1\,,\\
\{\mathrm{d}x_2, \mathrm{d}x_3\}_\Omega &= \mathrm{d}\{x_2, x_3\}_f = -2x_3\mathrm{d}x_1 - 2x_1\mathrm{d}x_3\,,\\
\{\mathrm{d}x_1, \mathrm{d}x_3\}_\Omega &= \mathrm{d}\{x_1, x_3\}_f = 0\,.
\end{aligned}$$

There is a corresponding sequence [13]

$$\begin{tikzcd}
0 \arrow[r] & \Omega^0(\mathcal{A}) \arrow[r, "\underline{\mathrm{d}}"] \arrow[d, "\cong"] & \Omega^1(\mathcal{A}) \arrow[r, "\underline{\mathrm{d}}"] \arrow[d, "\cong"] & \Omega^2(\mathcal{A}) \arrow[r, "\underline{\mathrm{d}}"] \arrow[d, "\cong"] & \Omega^3(\mathcal{A}) \arrow[r] \arrow[d, "\cong"] & 0\\
0 & \Lambda^0((\mathcal{D}(\mathcal{A})) \arrow[l] & \Lambda^1((\mathcal{D}(\mathcal{A})) \arrow[l, "d_\pi"'] & \Lambda^2((\mathcal{D}(\mathcal{A})) \arrow[l, "d_\pi"'] & \Lambda^3((\mathcal{D}(\mathcal{A})) \arrow[l, "d_\pi"'] & 0 \arrow[l]
\end{tikzcd}$$

where $\underline{\mathrm{d}}$ is the universal derivative (see (15)) and d_π is given by (27). The isomorphism are coming from the duality (16).

6.2.1 Low-Dimensional Cohomology of the PCY Algebra

Considering the above sequence, we firstly notice that $\Lambda^0((\mathcal{D}(\mathcal{A})) \cong \mathcal{A}$ and $\Lambda^1((\mathcal{D}(\mathcal{A})) \cong \mathcal{D}(\mathcal{A})$, so that the first map amounts to mapping $\mathcal{A} \to \mathcal{D}(\mathcal{A})$, and $\delta_\pi(a) = [\![\pi, a]\!] = \partial_a \in \mathrm{Ham}(\mathcal{A})$ is a Hamiltonian derivation. Thus $\delta_\pi(a) = 0$ iff a is a Casimir element (see def. (12)) of $\mathcal{A}$ and we have

$$HP^0(\mathcal{A}) \cong \mathrm{Cas}(\mathcal{A}) = < {x_1}^2 x_3 >_{\mathcal{A}} ,$$

where $\mathrm{Cas}(\mathcal{A})$ are the Cassimir elements of $\mathcal{A}$. For the first cohomology, consider $\partial \in \mathcal{D}(\mathcal{A})$, we have $\delta_\pi(\partial) = [\![\pi, \partial]\!] = -[\![\partial, \pi]\!] = \mathcal{L}_\partial \pi$. So we see that $\partial \in \ker \delta_\pi$ iff $\mathcal{L}_\partial \pi = 0$, meaning that π is invariant with respect to ∂. We have already met these operators in the section in which we computed the low-dimensional Poisson cohomology for more general algebras: the set of such operators is denoted $\mathrm{Can}(\mathcal{A})$ and ∂ is called Poisson canonical. The first cohomology is

$$HP^1(\mathcal{A}) = \mathrm{Can}(\mathcal{A}) / \mathrm{Ham}(\mathcal{A}) .$$

6.3 *Dual Poisson Complex*

Consider the chain complex

$$\Lambda^3((\mathcal{D}(\mathcal{A})) \xrightarrow{d_\pi} \Lambda^2((\mathcal{D}(\mathcal{A})) \xrightarrow{d_\pi} \Lambda^1((\mathcal{D}(\mathcal{A})) \xrightarrow{d_\pi} \Lambda^0((\mathcal{D}(\mathcal{A})) .$$

This complex was introduced by Brylinski in [7]. It has highly non-trivial (Poisson) homology. For example, the lowest homology is

$$HP_0(\mathcal{A}) \cong \mathcal{A} / < \{a, b\}_f \,|\, a, b \in \mathcal{A} >_{\mathcal{A}} ,$$

since $\Lambda^0((\mathcal{D}(\mathcal{A})) = \mathcal{A}$ and the image of $\delta_\pi : \Lambda^1((\mathcal{D}(\mathcal{A})) \to \Lambda^0((\mathcal{D}(\mathcal{A}))$ is given by $\delta_\pi(adb) = \{a, b\}_f$. Following the definition of the Nambu-Jacobi-Poisson bracket on $\mathcal{A}$ we get

$$\{a, b\}_f = \det \begin{pmatrix} \partial_1 a & \partial_2 a & \partial_3 a \\ \partial_1 b & \partial_2 b & \partial_3 b \\ \partial_1 f & \partial_2 f & \partial_3 f \end{pmatrix}$$

Writing $\nabla a = (\partial_1 a, \partial_2 a, \partial_3 a)$, we can express $\{a, b\}_f$ as

$$\{a, b\}_f = \nabla f \cdot (\nabla a \times \nabla b) ,$$

where $\cdot$ is the dot product and $\times$ is the vector product. More details on Poisson (co)homology of the Dual Poisson complex can be found in [2].

Exercise 2.9 Describe all differentials in the complex above in terms of vector analysis operations: ∇, curl, $\times$, $(-,-)$, div.

Generalized SPDUNR Poisson Algebra The following is the generalized Sklyanin-Painlevé-Dubrovin-Ugaglia-Nelson-Regge Poisson algebra:

$$\mathcal{A}_f = (\mathbb{C}[x_1, x_1, x_3], \{-,-\}_f) ,$$

where $\{-,-\}_f$ is the Jacobian Poisson-Nambu structure on $\mathbb{C}^3$ (see (28)), and $F, G \in \mathbb{C}[x_1, x_1, x_3]$. Let M_f be the zero locus of

$$f = x_1x_1x_3 + \sum_{i=1}^{3} a_i {x_i}^3 - \sum_{i=1}^{3} \epsilon_i {x_i}^2 + \sum_{i=1}^{3} c_i x_i + \omega ,$$

where $\epsilon_i \in \{0, 1\}$ and $a_i, c_i, \omega \in \mathbb{C}$. The bracket is given by

$$\{f, x_i\}_f := 0, \text{for } i = 1, 2, 3 ,$$

and

$$\{x_1, x_2\}_f := x_1x_2 + 3a_3{x_3}^2 - 2\epsilon_3 x_3 + c_3 ,$$

the result being cyclic in $(1, 2, 3)$ for other x_i, x_j. For a generic set of constraints, the bracket is nowhere vanishing on M_f.

7 Graded Poisson Algebras

Let $\mathcal{A}$ be a Poisson algebra. We shall suppose that $\mathcal{A}$ is an associative *graded algebra*, that is, $\mathcal{A}$ contains a set of vector subspaces $(\mathcal{A}^k)_{k\in\mathbb{N}_0}$ s.t. $A = \bigoplus_{k\in\mathbb{N}_0} \mathcal{A}^k$ and $\mathcal{A}^k \cdot \mathcal{A}^l \subset \mathcal{A}^{k+l}$ for all $k, l \in \mathbb{N}_0$. Moreover, we assume $\mathcal{A}_0 := \mathbb{K}$.

Definition 2.20 Let $d \in \mathbb{N}_0$ be arbitrary. $\mathcal{A}$ is called a *graded Poisson algebra of degree* d if $\forall a \in \mathcal{A}^k, b \in \mathcal{A}^l : \{a, b\} \in \mathcal{A}^{k+l-d}$ (for $n < 0$ define $A^n = 0$).

The graded Poisson algebras can be constructed from non-commutative, associative, unital algebras $\mathcal{U}$, which are *filtered*:

$$\mathcal{U} = \bigcup_{k\in\mathbb{N}_0} \mathcal{U}_k, \text{ where } 1 \in \mathbb{K} = \mathcal{U}_0 \subset \mathcal{U}_1 \subset \cdots \subset \mathcal{U}_k \subset \ldots .$$

To every such algebra $\mathcal{U}$, we can define the associated graded algebra $\mathcal{S} = \mathrm{gr}(\mathcal{U})$, $\mathcal{S} = \bigoplus_{k\in\mathbb{N}_0} \mathcal{S}_k$, where $\mathcal{S}_k := \mathcal{U}_k / U_{k-1}$, $k \geq 1$, and $\mathcal{S}_0 := \mathcal{U}_0 = \mathbb{K}$. We denote by $\mathrm{gr}_k : \mathcal{U}_k \to \mathcal{S}_k$ the canonical projections. Then $\mathcal{U}$ is a graded Poisson algebra of degree $d \geq 1$ if $uv - vu \in \mathcal{U}_{k+l-d}$, for all $u \in \mathcal{U}_k$, $v \in \mathcal{U}_l$ (define $\mathcal{U}_k = 0$ for $k < 0$).

Proposition 2.11 *The associated graded algebra $\mathcal{S}$ given by the graded algebra $\mathcal{U}$ of degree d is a graded Poisson algebra of degree d.*

Proof Since the canonical projection $\mathrm{gr}_k : \mathcal{U}_k \to S_k$ is a surjection, to each $a \in \mathcal{S}$ exists $k \in \mathbb{N}_0$ so that $a \in \mathcal{S}_k$, thus there exists $u \in \mathcal{U}_k$ s.t. $a = \mathrm{gr}_k(u)$. Let $b \in \mathcal{S}_l$ and $v \in \mathcal{U}_l$ s.t. $b = gr_l(v)$. The product in $\mathcal{S}$ is defined by

$$ab = gr_{k+l}(uv) \in \mathcal{S}_{k+l} \ .$$

The product is well-defined, because if we pick different representatives, say $\tilde{a} = a + x, x \in \mathcal{U}_{k-1}$ and $\tilde{b} = b + y, y \in \mathcal{U}_{l-1}$, then

$$\tilde{a}\tilde{b} = (a+x)(b+y) = ab + \underbrace{ay + xb + xy}_{\in \mathcal{U}_{k+l-1}} = ab \ .$$

The unit in $\mathcal{S}$ is the same as in $\mathcal{U}$, and the associativity of $\mathcal{S}$ is also inherited from the associativity of $\mathcal{U}$. So we see that $\mathcal{S}$ is a graded algebra. The Lie bracket is

$$\{a, b\} = \{\mathrm{gr}_k(u), \mathrm{gr}_l(v)\} := gr_{k+l-d}(uv - vu)$$

and is of degree d. By a similar argument as for the product, $\{-, -\}$ is well-defined on $\mathcal{S}$. Since $ab \in \mathcal{S}_{k+l}$ and $ba \in \mathcal{S}_{l+k} = \mathcal{S}_{k+l}$, there exist $u, u' \in \mathcal{U}_{k+l}$ s.t. $ab = \mathrm{gr}_{k+l}(u)$, $ba = \mathrm{gr}_{k+l}(u')$ and thus $ab - ba = \mathrm{gr}_{k+l}(u - u') = 0$. This shows that the product in $\mathcal{S}$ is commutative, hence $\mathcal{S}$ is Poisson graded of degree d. □

Module Structures on $\mathcal{A}$ Let $\mathcal{A}$ be a commutative, associative $\mathbb{K}$-algebra with the unit 1. $\mathrm{End}(\mathcal{A}) = \mathrm{Hom}(\mathcal{A}, \mathcal{A})$ has two $\mathcal{A}$-module structures, *left* and *right*: $\forall a, x \in \mathcal{A}$

$$l_a\varphi(x) := a\varphi(x), \qquad r_a\varphi(x) := \varphi(ax) \ . \tag{30}$$

Lemma 2.4 *For arbitrary $a \in A$, denote*

$$\delta_a := r_a - l_a. \tag{31}$$

Then δ_a satisfies the Leibniz rule

$$\delta_a(\varphi \circ \psi) = \delta_a\varphi \circ \psi + \varphi \circ \delta_a\psi \ ,$$

that is, $\delta_a \in \mathcal{D}(\mathrm{End}(\mathcal{A})) \subset \mathrm{Hom}(\mathrm{End}(\mathcal{A}), \mathrm{End}(\mathcal{A}))$.

Proof For arbitrary $\varphi, \psi \in \mathrm{End}(\mathcal{A})$ and $a, u \in \mathcal{A}$, using the definition (31), we have

$$\delta_a(\varphi \circ \psi)(u) = \delta_a(\varphi(\psi(u))) = \varphi(\psi(au)) - a\varphi(\psi(u)) \ .$$

On the other hand

$$
\begin{aligned}
(\delta_a\varphi \circ \psi + \varphi \circ \delta_a \psi)(u) &= \delta_a \varphi(\psi(u)) + \varphi(\delta_a \psi(u)) \\
&= \varphi(a\psi(u)) - a\varphi(\psi(u)) + \varphi(\psi(au) - a\psi(u)) \\
&= \varphi(\psi(au)) - a\varphi(\psi(u)) \, .
\end{aligned}
$$

□

Lemma 2.5 $[\delta_a, \delta_b] = 0$ *for all* $a, b \in \mathcal{A}$, *where* δ_a *is given by* (31).

Proof The proof is a straightforward computation and we leave it to the reader as an exercise. □

7.1 Algebra of Differential Operators

Definition 2.21 For all $k \in \mathbb{N}_0$, we define

$$
\mathrm{Diff}_k(\mathcal{A}) := \bigcap_{\substack{a_i \in \mathcal{A} \\ 0 \le i \le k}} \ker(\delta_{a_0} \circ \ldots \delta_{a_i})
$$

and

$$
\mathrm{Diff}_*(\mathcal{A}) := \bigcup_{k \ge 0} \mathrm{Diff}_k(\mathcal{A}) \, ,
$$

which is an abelian group under the addition $+$. Then $\mathrm{Diff}_*(\mathcal{A})$ inherits the two $\mathcal{A}$-module structures (30) of $\mathrm{End}(\mathcal{A})$. We will write $\mathrm{Diff}_*^{(+)}(\mathcal{A})$ to emphasize the bimodule structure. The elements of $\mathrm{Diff}_k(\mathcal{A})$ will be called differential operators of order $\le k$ on a commutative algebra $\mathcal{A}$.

Note 2.4 Directly from the above definition we have that for all $k \in \mathbb{N}_0$: $\mathrm{Diff}_{k-1}(\mathcal{A}) \subset \mathrm{Diff}_k(\mathcal{A})$.

Remark 2.11 One can generalize the above definition to the case $\varphi \colon P \to Q$, where P and Q are projective, finitely generated $\mathcal{A}$-modules ($0 \le k$) and φ is an $\mathcal{A}$-module homomorphism. Then

$$
\mathrm{Diff}_k(P, Q) := \{\varphi \colon P \to Q \mid \delta_{a_0} \circ \cdots \circ \delta_{a_k}(\varphi) = 0 \text{ for all } a_0, \ldots, a_k \in \mathcal{A}\} \, .
$$

In this notation, $\mathrm{Diff}_k(\mathcal{A}) = \mathrm{Diff}_k(\mathcal{A}, \mathcal{A})$.

Example (Lie Algebra Structure on $\mathrm{Diff}_*^{(+)}(\mathcal{A})$*)* Consider $\mathrm{End}_{\mathbb{K}}(\mathcal{A})$, equipped with a Lie algebra structure given by the commutator

$$[\varphi, \psi] = \varphi \circ \psi - \psi \circ \varphi \, .$$

Using the derivation property from Lemma 2.4, we have

$$\delta_a[\varphi, \psi] = [\delta_a\varphi, \psi] + [\varphi, \delta_a\psi] \tag{32}$$

for all $a \in \mathcal{A}$ and $\varphi \in \mathrm{End}_{\mathbb{K}}(\mathcal{A})$. Hence δ_a acts as a derivation on the commutator. Suppose that $\varphi, \psi \in \mathrm{Diff}_1^{(+)}(\mathcal{A})$, then from (32) we get

$$\delta_b \circ \delta_a[\varphi, \psi] = [\delta_a\varphi, \delta_b\psi] + [\delta_b\varphi, \delta_a\psi] \, ,$$

which does not have to vanish. Hence $[\varphi, \psi] \notin \mathrm{Diff}_1^{(+)}(\mathcal{A})$, meaning that $\mathrm{Diff}_1^{(+)}(\mathcal{A})$ is not a Lie algebra. Applying δ once again yields

$$\delta_c \circ \delta_b \circ \delta_a[\varphi, \psi] = 0 \, .$$

Thus $[\varphi, \psi] \in \mathrm{Diff}_2^{(+)}(\mathcal{A})$. Proceeding in a similar fashion one can show that the composition of $\varphi \in \mathrm{Diff}_k^{(+)}(\mathcal{A})$ and $\psi \in \mathrm{Diff}_l^{(+)}(\mathcal{A})$ is of order $\leq k + l$, that is $\varphi \circ \psi \in \mathrm{Diff}_{k+l}^{(+)}(\mathcal{A})$ and the filtered bimodule $\mathrm{Diff}_*^{(+)}(\mathcal{A})$ is a Lie algebra.

Remark 2.12 The $\mathcal{A}$-bimodule $\mathrm{Diff}_*^{(+)}(P, P)$ is also filtered, since

$$\delta(\varphi \circ \psi) = \delta(\varphi) \circ \psi + \varphi \circ \delta\psi$$

and so the composition of a differential operator of degree $\leq k$ with a differential operator of degree $\leq l$ results in a differential operator of degree $\leq k + l$

$$\mathrm{Diff}_k^{(+)}(P, P) \otimes_{\mathbb{K}} \mathrm{Diff}_l^{(+)}(P, P) \to \mathrm{Diff}_{k+l}^{(+)}(P, P) \, .$$

Example Let $\varphi \in \mathrm{End}_{\mathbb{K}}(\mathcal{A})$, and recall that $\mathcal{A}$ is associative, commutative and unital algebra over $\mathbb{K}$.

- $\mathrm{Diff}_0(\mathcal{A}) = \bigcap_{a \in \mathcal{A}} \ker \delta_a$. Using the definition (31), we have

$$\delta_a\varphi(u) = \varphi(au) - a\varphi(u) \, .$$

(continued)

So $\varphi \in \mathrm{Diff}_0(\mathcal{A})$ iff $\varphi(au) = a\varphi(u)$ for all $a, u \in \mathcal{A}$. Choosing $u = 1$ and writing $a = a1$ gives $\varphi(a1) = a\varphi(1)$, meaning that φ is completely determined by its value on the unit element of $\mathcal{A}$, which gives

$$\mathrm{Diff}_0(\mathcal{A}) = \mathrm{End}_{\mathcal{A}}(\mathcal{A}) = \mathcal{A} .$$

Note that the above case is rather special. If we consider the case of $\mathcal{A}$-modules P, Q, then we get

$$\mathrm{Diff}_0(P, Q) = \mathrm{End}_{\mathcal{A}}(P, Q) .$$

- $\mathrm{Diff}_1(\mathcal{A}) = \bigcap_{a,b\in\mathcal{A}} \ker \delta_b \circ \delta_a$, where

$$\delta_b \circ \delta_a \varphi(u) = \varphi(bau) - b\varphi(au) - a\varphi(bu) + ba\varphi(u) . \tag{33}$$

- $\mathrm{Diff}_2(\mathcal{A}) = \bigcap_{a,b,c\in\mathcal{A}} \ker \delta_c \circ \delta_b \circ \delta_a$, where

$$\delta_c \circ \delta_b \circ \delta_a \varphi(u) = \varphi(bau) - b\varphi(au) - a\varphi(bu) + ba\varphi(u) .$$

The goal of the following example is to demonstrate that the above given algebraic definition of differential operators on a commutative algebra $\mathcal{A}$ fits with the standard picture of differential operators on functions.

Example Let $\mathcal{A} = C^\infty(\mathbb{R})$ be the algebra of smooth functions of one real variable, the algebra binary operation given by multiplication of functions. Take $\varphi = \partial_x := \frac{\partial}{\partial x}$. Then

$$\delta_a \partial_x(u) = 0 \iff \partial_x(au) = a\partial_x u . \tag{34}$$

which is not the case for all $a, u \in C^\infty(\mathbb{R})$. As expected (since $\partial_x \notin \mathcal{A}$), the above implies $\partial_x \notin \mathrm{Diff}_0(\mathcal{A})$. On the other hand, using (33),

$$\delta_b \circ \delta_a \varphi(u) = \partial_x(bau) - b\partial_x(au) - a\partial_x(bu) + ba\partial_x u = 0$$

is satisfied for all $a, b, u \in C^\infty(\mathbb{R})$, thus $\partial_x \in \mathrm{Diff}_1(C^\infty(\mathbb{R}))$. Similarly, if we take $f\partial_x$, where $f \in C^\infty(\mathbb{R})$, then

(continued)

$$\delta_b \circ \delta_a f \partial_x(u) = f \partial_x(bau) - bf \partial_x(au) - af \partial_x(bu) + baf \partial_x u = 0 ,$$

and so $f \partial_x \in \mathrm{Diff}_1(C^\infty(\mathbb{R}))$. Finally, consider $\partial_x^2 := \partial_x \circ \partial_x$. Then

$$\delta_b \circ \delta_a \partial_x^2 = \delta_b(\delta_a \partial_x \circ \partial_x + \partial_x \circ \delta_a \partial_x) = \delta_a \partial_x \circ \delta_b \partial_x + \delta_b \partial_x \circ \delta_a \partial_x ,$$

which is not vanishing for all $a, b \in C^\infty(\mathbb{R})$. Thus $\partial_x^2 \notin \mathrm{Diff}_1(C^\infty(\mathbb{R}))$. We can easily check that

$$\delta_c \delta_b \circ \delta_a f \partial_x^2 = 0 ,$$

so $f \partial_x^2 \in \mathrm{Diff}_2(C^\infty(\mathbb{R}))$. One can show that $f \partial_x^i \in \mathrm{Diff}_i(C^\infty(\mathbb{R}))$, for all $i \in \mathbb{N}_0$ and $f \in C^\infty(\mathbb{R})$, where $\partial_x^i := \underbrace{\partial_x \circ \ldots \circ \partial_x}_{i\text{-times}}$. Since we have the sequence of inclusions

$$\mathrm{Diff}_0 \mathcal{A} \hookrightarrow \mathrm{Diff}_1 \mathcal{A} \hookrightarrow \ldots \hookrightarrow \mathrm{Diff}_k \mathcal{A} \hookrightarrow \ldots .$$

Altogether we get

$$D_k := \sum_{i=0}^{k} f_i \partial_x^i \in \mathrm{Diff}_k(C^\infty(\mathbb{R})) .$$

Definition 2.22 Consider the factor space

$$\mathrm{Smbl}_k(\mathcal{A}) := \mathrm{Diff}_k^{(+)}(\mathcal{A}) / \mathrm{Diff}_{k-1}^{(+)}(\mathcal{A}) .$$

The *symbols algebra of* $\mathcal{A}$ is

$$\mathrm{Smbl}_*(\mathcal{A}) := \bigoplus_{k \in \mathbb{N}_0} \mathrm{Smbl}_k(\mathcal{A}) ,$$

with the graded algebra structure

$$\mathrm{Smbl}_k \cdot \mathrm{Smbl}_l \subset \mathrm{Smbl}_{k+l} \{\mathrm{Smbl}_k, \mathrm{Smbl}_l\} \subset \mathrm{Smbl}_{k+l-1} .$$

Remark 2.13 Recall that to any filtered algebra, we can associate a graded commutative algebra. This is precisely the case of the symbol algebra $\mathrm{Smbl}_*(\mathcal{A})$ with respect to the filtered algebra $\mathrm{Diff}_*^{(+)}(\mathcal{A})$.

Proposition 2.12 *The symbol algebra* $\mathrm{Smbl}_*(\mathcal{A})$ *is a Poisson graded algebra of degree* 1.

Exercise 2.10

1. $\mathrm{Diff}_k^{(+)}(\mathcal{A})$ is an $\mathcal{A}$-submodule in $\mathrm{End}(\mathcal{A})$.
2. $\mathrm{Diff}_*(\mathcal{A})$ is a filtered subalgebra in $\mathrm{End}(\mathcal{A})$.
3. $\mathrm{Diff}_0(\mathcal{A}) = \mathcal{A}$ and $\mathrm{Diff}_1(\mathcal{A}) = \mathcal{D}(\mathcal{A}) \oplus \mathcal{A}$.
4. $\mathrm{Smbl}_0(\mathcal{A}) = \mathcal{A}$ and $\mathrm{Smbl}_1(\mathcal{A}) = \mathcal{D}(\mathcal{A})$.

8 Intermezzo: Tensor, Symmetric and Exterior Algebras

We will now review the constructions of tensor, symmetric, and exterior algebras over a finite dimensional $\mathbb{K}$-vector space V. We will also discuss the situation when V is a Lie algebra, which leads to a Poisson structure.

8.1 Tensor Algebra of a Vector Space

Tensor algebra of a $\mathbb{K}$ vector space, denoted by $T(V)$ is

$$T(V) := \bigoplus_{k=0}^{\infty} T^k(V) ,$$

where

$$T^k(V) := \otimes^k V = V \otimes \ldots \otimes V .$$

The space $T^k(V)$ consists of $\mathbb{K}$-multilinear mappings

$$\tau \colon V^* \times \ldots \times V^* \to \mathbb{K} ,$$

where V^* is the vector space dual to V. by definition, $T^0(V) = \mathbb{K}$. Also, $T^1(V) = V$. The algebra product in $T(V)$ is given by the canonical isomorphism defined by the tensor product

$$T^k(V) \otimes T^l(V) \to T^{k+l}(V) .$$

The tensor algebra satisfies the following universal property: for every $\mathbb{K}$-algebra $\mathcal{A}$ and arbitrary linear map $\psi \colon V \to \mathcal{A}$, there is a uniquely given linear map $\tilde{\psi} \colon T(V) \to \mathcal{A}$ such that the following diagram commutes

$$\begin{tikzcd} V \arrow[r, "\iota"] \arrow[dr, "\psi"'] & T(V) \arrow[d, dashed, "\exists!\tilde{\psi}"] \\ & \mathcal{A} \end{tikzcd}$$

That is $\tilde{\psi} \circ \iota = \psi$, where ι is the canonical embedding of V into $T(V)$.

8.2 Symmetric Algebra of a Vector Space

Let V be a $\mathbb{K}$-vector space, $T(V)$ its tensor algebra. Consider

$$\mathcal{J} :=< u \otimes v - v \otimes u \mid u, v \in V > , \tag{35}$$

which is a two-sided ideal in $T(V)$. Then the quotient algebra $S(V) := T(V)/\mathcal{J}$, called the *symmetric algebra of* V, is an associative and commutative algebra, satisfying the following universal property. For any associative, commutative and unital algebra $\mathcal{A}$ and every linear mapping $\psi : V \to \mathcal{A}$, there is precisely one unital algebra homomorphism $\tilde{\psi}$ such that the following diagram commutes

$$\begin{tikzcd} V \arrow[r, "\iota"] \arrow[dr, "\psi"'] & S(V) \arrow[d, dashed, "\exists!\tilde{\psi}"] \\ & \mathrm{End}(V) \end{tikzcd}$$

where ϵ is the canonical embedding of V into $S(V)$, given by the composition of the canonical embedding $V \to T(V)$ and the canonical quotient projection $T(V) \to S(V)$.

Remark 2.14 Let us mention some useful properties about the above defined algebras.

- $S(V)$ is a free, associative, commutative and unital algebra on $n = \dim V$ generators. The product is given as follows. To avoid confusion, we will denote the classes in $S(V)$ with bracket notation. Let $[s] \in S^k(V), [t] \in S^l(V)$. Then the product is

  $$[s][t] := [s \otimes t] \in S^{k+l}(V) .$$

  This product is well defined. To check this, consider different representatives of the equivalence classes $[\tilde{s}] = [s], [\tilde{t}] = [t]$. This means there exist $j_s, j_t \in \mathcal{J}$ such that $s = \tilde{s} + j_s$ and $t = \tilde{t} + j_t$. Then

  $$[s \otimes t] = [(\tilde{s} + j_s) \otimes \tilde{t} + j_t)] = [\tilde{s} \otimes \tilde{t} + \tilde{s} \otimes j_t + j_s \otimes \tilde{t} + j_s \otimes j_t] = [\tilde{s} \otimes \tilde{t}] ,$$

where the last equality follows from the fact that $\tilde{s} \otimes j_t + j_s \otimes \tilde{t} + j_s \otimes j_t \in \mathcal{J}$, since $\mathcal{J}$ is a two-sided ideal in $T(V)$.

- The ideal $\mathcal{J}$ is homogeneous, meaning that the factor algebra $S(V)$ inherits the grading of $T(V)$. Thus we have $S(V) = \oplus_{k=0}^{\infty} S^k(V)$, where $S^k(V) := T^k(V)/\mathcal{J}^k$, where $\mathcal{J}^k = \mathcal{J} \cap T^k(V)$.
- Consider the canonical projection $\mathrm{pr}_k \colon T^k(V) \to S^k(V)$. We can restrict pr_k to the linear subspace of symmetric k-tensors

$$\mathrm{pr}_k |_{\mathrm{Sym}_k(V)} \colon \ \mathrm{Sym}_k(V) \to S^k(V) \ .$$

Because we always assume $\operatorname{char} \mathbb{K} = O$, the above map can be inverted, yielding a graded vector space isomorphism

$$\mathrm{Sym}(V) = \oplus_{k=0}^{\infty} \mathrm{Sym}_k(V) \cong \oplus_{k=0}^{\infty} S^k(V) = S(V)$$

Although the space of symmetric tensors and the symmetric algebra are isomorphic as a graded $\mathbb{K}$-vector spaces, it does not make sense to speak about isomorphism of algebras, since $\mathrm{Sym}(V)$ does not posses algebra structure in the sense that the tensor product of two symmetric tensors does not have to be a symmetric tensor. In $\operatorname{char} \mathbb{K} > 0$, we even lose the graded vector spaces isomorphism.
- $\mathrm{Sym}(V)$ is a linear subspace of $T(V)$.
- $S(V)$ is a not a subalgebra of $T(V)$.

Symmetrization Let $\mathrm{s}_k \colon T^k(V) \to \mathrm{Sym}^k(V)$ be the symmetrization map, given for arbitrary $t \in T^k(V)$ by

$$\mathrm{s}_k(t) := \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} \sigma \cdot t \ , \tag{36}$$

where $\mathfrak{S}_k$ is the permutation group on k-elements and $\sigma \cdot t$ is the action of the permutation group on k-tensors, given on $t = t_1 \otimes \ldots \otimes t_k$ by $\sigma \cdot t := t_{\sigma(1)} \otimes \ldots \otimes t_{\sigma(k)}$ (and we extend $\cdot$ on general $t \in S_k$ by $\mathbb{K}$-linearity).

Proposition 2.13 *The ideal* (35) *is a graded ideal* $\mathcal{J} = \oplus_{k=0}^{\infty} \mathcal{J}_k$, *where* $\mathcal{J}_k := J \cap T^k(V)$. *Moreover,* $\ker \mathrm{s}_k = \mathcal{J}_k$ *and* $T^k(V) = \mathcal{J}_k \oplus \mathrm{Sym}^k(V) \cong \mathcal{J}_k \oplus S^k(V)$

Let $\{e_1, \ldots, e_n\}$ be a basis of V. The map (36) gives an isomorphism

$$S(V) \cong \mathbb{K}[x_1, \ldots, x_n]$$

such that

$$\mathrm{s}_k(e_{i_1} \otimes \ldots \otimes e_{i_k}) = x_{i_1} \ldots x_{i_k} \ .$$

The special case of PBW theorem says that the monoms $\{e_1^{i_1}, \ldots, e_n^{i_1}\}$ form a basis of $S(V)$ as a $\mathbb{K}$-vector space and $S(V) \cong \mathbb{K}[e_1, \ldots, e_n]$.

Example Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The symmetric algebra $S(\mathfrak{g}^*)$ is isomorphic to $\mathbb{K}[\mathfrak{g}^*]$, the polynomial algebra with $\dim \mathfrak{g}^*$ variables. The bracket on $\mathfrak{g}^*$ makes $\mathbb{K}[\mathfrak{g}^*]$ a (commutative) Poisson algebra.

A simpler version of this construction is $\mathcal{A} = S(V^*) = \mathbb{K}[V^*]$, where V is a finite dimensional vector space equipped with a skew-symmetric bilinear form $B\colon V \times V \to \mathbb{K}$, which provides a Poisson brackets on $\mathbb{K}[V^*]$. For instance, let $\dim V = 2$. Consider $X, Y \in V$ linearly independent, so that $V =< X, Y >$. Then $\mathbb{K}[V^*] \cong \mathbb{K}[X, Y]$ and $\{X, Y\} = B(X, Y) := 1$. In this case, the pair (V, B) is a symplectic plane.

8.3 Exterior Algebra of a Vector Space

Let V be a n-dimensional $\mathbb{K}$-vector space. The *exterior algebra* of V, denoted $\Lambda(V)$, is defined as a graded subspace in the tensor algebra $T(V)$, formed by completely antisymmetric tensors. Recall that $t \in T^k(V)$ is *completely antisymmetric* (or *alternating*, or *completely skew-symmetric*) if

$$\sigma \cdot t := t_{\sigma(1)} \otimes \ldots \otimes t_{\sigma(k)} = \mathrm{sgn}(\sigma) t_1 \otimes \ldots \otimes t_k$$

for all k-permutations σ (and we extend $\cdot$ on general $t \in S_k$ by $\mathbb{K}$-linearity). Then $\Lambda^k(V) \subset T^k(V)$ is formed by all such k-tensors. Note that $\Lambda^1(V) = V$ and for $k = 0$ we define $\Lambda^0(V) = \mathbb{K}$. The whole algebra is given by the $\mathbb{K}$-vector space

$$\Lambda(V) := \oplus_{k \geq 0} \Lambda^k(V) ,$$

with the product $\wedge$, called *wedge product*,

$$\Lambda^k(V) \otimes \Lambda^l(V) \to \Lambda^{k+l}(V)$$

given by

$$\omega_1 \wedge \omega_2 := \frac{(k+l)!}{k!l!} \mathcal{A}lt(\omega_1 \otimes \omega_2) , \tag{37}$$

where

$$\mathcal{A}lt\colon T^k(V) \to \Lambda^k(V) ,$$

is a projection on the subspace of alternating tensors, called *alternating map*, defined for arbitrary $t \in T^k(V)$ as

$$\mathcal{A}lt(t) := \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} \operatorname{sgn}(\sigma)\sigma \cdot t \, .$$

Example Consider $t \in T^2(V)$, given by $t = v \otimes w - w \otimes v$. Then $\mathcal{A}lt(t) = \frac{1}{2}(v \otimes w - w \otimes v - w \otimes v + v \otimes w = t$, i.e. t is already an element of $\Lambda^2(V)$.

The wedge product satisfies

$$\omega_1 \wedge \omega_2 = (-1)^{pq} \omega_2 \wedge \omega_1,$$

$$\omega_1 \wedge (\omega_2 \wedge \omega_3) = (\omega_1 \wedge \omega_2) \wedge \omega_3$$

meaning that $\Lambda(V)$ is a graded, associative $\mathbb{K}$-algebra with the unit 1.

Similarly as in the case of symmetric algebra, we can define $\Lambda(V)$ as a quotient of the tensor algebra. Consider a two-sided ideal $\mathcal{J} \subset T(V)$, given by

$$\mathcal{J} :=< t_1 \otimes t_2 + t_2 \otimes t_1 \mid t_1, t_2 \in T(V) >$$

Then we define $\Lambda(V) := T(V)/\mathcal{J}$. Moreover, $\Lambda(V)$ satisfies the following universal property. For every associative, unital algebra $\mathcal{A}$ and any $\mathbb{K}$-linear map $\varphi \colon V \to \mathcal{A}$, such that $j(v)^2 = 0$ for all $v \in V$, there is a unique algebra homomorphism $\tilde{\varphi} \colon \Lambda(V) \to \mathcal{A}$, such that the following diagram commutes

$$\begin{tikzcd} V \arrow[r, "\iota"] \arrow[dr, "\varphi"'] & \Lambda(V) \arrow[d, dashed, "\exists! \tilde{\varphi}"] \\ & \mathcal{A} \end{tikzcd}$$

Using the universal property, one can show that the two above construction of $\Lambda(V)$, either as a subspace of alternating tensors or the quotient algebra, are isomorphic (in a unique way).

8.4 Poisson Structure on a Symmetric Algebra $S(\mathfrak{g})$

Theorem 2.3 *$S(\mathfrak{g})$ is a graded Poisson algebra of degree* 1 *or of degree* 2.

Proof

Degree 1 Let $\mathfrak{g}$ be a $\mathbb{K}$-Lie algebra. We will show that the Lie bracket extends in a unique way to a Poisson bracket of degree 1 on $S(\mathfrak{g})$. Let $g \in \mathfrak{g}$ be arbitrary. Consider the endomorphism[6] $\mathrm{ad}_g \equiv \delta_g \in \mathrm{End}_{\mathbb{K}}(\mathfrak{g})$, given by

$$\mathrm{ad}_g(h) = [g, h] \, , h \in \mathfrak{g} \, .$$

Then ad_g can be extended in a unique way to the whole $S(\mathfrak{g})$ to a derivation on $S(\mathfrak{g})$ as follows. Consider a *decomposable* $t \in S(\mathfrak{g})$, i.e. t can be written as $t = t_1 \cdot \ldots \cdot t_k$, where $\cdot$ denotes the (symmetric) product in $S(\mathfrak{g})$ and $t_1, \ldots, t_k \in \mathfrak{g}$. Then

$$\mathrm{ad}_g(t) := \sum_{i=1}^{k} t_1 \cdot \ldots \cdot t_{i-1} \cdot ad_x(t_i) \cdot t_{i+1} \cdot \ldots \cdot t_k \, .$$

A general element of $S(\mathfrak{g})$ is a $\mathbb{K}$-linear combination of decomposables, so the above definition of the bracket can be extended linearly to the whole $S(\mathfrak{g})$. This extension is also denoted ad_g. For arbitrary $t, u \in S(\mathfrak{g})$ we have

$$\mathrm{ad}_g(t \cdot u) = \mathrm{ad}_g(t) \cdot u + t \cdot \mathrm{ad}_g(u) \, .$$

Degree 2 It is sufficient to define a skew-symmetric bilinear form

$$\Omega \colon \mathfrak{g} \times \mathfrak{g} \to \mathbb{K} \, .$$

Then there is a unique extension of $\Omega(g, -) \colon \mathfrak{g} \to \mathbb{K}$ to a derivation

$$\partial \colon S(\mathfrak{g}) \to \mathcal{D}(S(\mathfrak{g})) \, ,$$

which can be extended to a degree 2 derivation on $S(\mathfrak{g})$.

□

Example Let Ω be non-degenerate, skew-symmetric 2-form on $\mathfrak{g} = \mathbb{R}^{2m}$, that is, we consider $\mathfrak{g}$ to be abelian (the Lie bracket is zero). Then there is a basis, called *symplectic basis*, $e_1, \ldots, e_n$ $(n = 2m)$, such that the matrix of Ω is written in this basis as

(continued)

[6] Notice that using the action of $\mathfrak{g}$ on itself given by multiplication from the left, we can identify elements in $\mathfrak{g}$ as endomorphisms of $\mathfrak{g}$, $h \mapsto l_h$. Then we have $\delta_g(l_h)(u) = h(gu) - gh(u) = (ad_g h)u$, thus $\delta_g|_{\mathfrak{g}} \equiv \mathrm{ad}_g$.

$$\begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix} ,$$

where I_n is the identity $m \times m$ matrix. Then for each $i \in \{1, \dots, m\}$, the brackets of basis elements satisfy $\{e_i, e_{i+m}\} = -\{e_{i+m}, e_i\} = 1$ and all other combinations of basis elements have zero brackets. Using this choice of basis, we can identify $S(\mathfrak{g}) \cong \mathbb{K}[x_1, \dots, x_{2m}]$. Note that this is a degree 2 bracket on $\mathfrak{g}$, and differs from the Lie bracket $[-, -]_{\mathfrak{g}}$, which we assumed is trivial. Let $m = 1$, so $\mathfrak{g} = \mathbb{R}^2$ and we consider coordinates x, y. Consider $P, Q \in \mathbb{K}[x, y]$ given by

$$P(x, y) = \sum_{i,j} a_{ij} x^i y^j , \qquad Q(x, y) = \sum_{p,q} b_{pq} x^p y^q .$$

The bracket is

$$\{P, Q\} = \sum_{i,j,p,q} a_{ij} b_{pq} \{x^i y^j, x^p y^q\} = \sum_{i,j,p,q} a_{ij} b_{pq} (jp - iq) x^{i+p-1} y^{q+j-1}.$$

If we pick $P = xy$, $Q = 4x^2$ and choose the bracket as

$$\{P, Q\} = \frac{\partial P}{\partial x} \frac{\partial Q}{\partial y} - \frac{\partial Q}{\partial x} \frac{\partial P}{\partial y} = -8x^2 ,$$

with the general formula being

$$\{P, Q\} = \sum_{i=1}^{m} \frac{\partial P}{\partial x_i} \frac{\partial Q}{\partial x_{i+m}} - \frac{\partial Q}{\partial x_{i+m}} \frac{\partial P}{\partial x_i}$$

then this gives a bracket of degree 2 on $\mathbb{K}[x_1, \dots, x_{2m}]$. The corresponding degree 2 filtered Poisson algebra is $\mathcal{A}_{2m}(\mathfrak{g})$. The algebra $\mathcal{A}_{2m}(\mathfrak{g})$ has $2m$ generators (x_i, y_i), $1 \le i \le m$ and relations

$$x_i y_i - y_i x_i = 1 , 1 \le i \le m$$
$$x_i y_j - y_j x_i = y_i y_j - y_j y_i = x_i x_j - x_j x_i = 0 , i \ne j .$$

$\mathcal{A}_{2m}(\mathfrak{g})$ is called the $2m$-th Weyl algebra.

Let us conclude this section with the following remark, which puts the above algebraic constructions in the context of smooth manifolds.

Remark 2.15 The constructions of tensor algebra, symmetric algebra, and exterior algebra can be extended to the case of smooth manifolds by considering the tangent (or cotangent) space at a given point. This leads to the notion of *tensor fields* as sections of the bundle $\otimes TM \to M$, *symmetric tensor fields* as sections of the bundle $S(TM) \to M$, and *differential forms* as sections of the bundle $\Lambda(TM) \to M$. We will speak more about these objects in chapter concerning the differential calculus on Poisson manifolds.

9 Universal Enveloping and PBW Theorem

Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{K}$. We assume that there is an associative algebra $\mathcal{A}$ such that $\mathfrak{g}$ can be embedded in $\mathcal{A}$ (so that one can multiply elements of $\mathfrak{g}$), and the Lie bracket of $\mathfrak{g}$ is given by the commutator in $\mathcal{A}$, i.e. for all $x, y \in \mathfrak{g}$

$$[x, y] = xy - yx \ ,$$

where the product on the right-hand side is the product in $\mathcal{A}$. Let us denote by $\{e_i\}_{1\leq i\leq n}$ a basis of $\mathfrak{g}$ as a vector space. The Lie structure is defined by the *structure constants* c_{ij}^k

$$[e_i, e_j] = c_{ij}^k e_k.$$

We consider $S \subset \mathfrak{g} \times \mathfrak{g}$, a set of pairs, defined as follows

$$S = \{(e_i e_j, e_j e_i + \sum_{k=1}^{n} c_{ij}^k e_k) \in \mathfrak{g} \times \mathfrak{g} \mid i > j\} \ .$$

It is easy to verify that there is no ambiguity in the presentation of the triple product $e_i e_j e_k$ with $i > j > k$, since

$$(e_i e_j)e_k = (\sum_{l=1}^{n} c_{ij}^l e_l + e_j e_i)e_k = \sum_{l=1}^{n} c_{ij}^l e_l e_k + e_j e_i e_k \ ,$$

$$e_i(e_j e_k) = e_i(\sum_{m=1}^{n} c_{jk}^m e_m + e_k e_j) = \sum_{m=1}^{n} c_{jk}^m e_i e_m + e_i e_k e_j \ .$$

Thus the difference vanishes due to the Jacobi identity (4)

$$(e_i e_j)e_k - e_i(e_j e_k) = \sum_{l=1}^{n}\sum_{m=1}^{n}(c_{ij}^m c_{km}^l + c_{jk}^m c_{im}^l + c_{ki}^m c_{jm}^l)e_l = 0 \ .$$

9.1 Universal Enveloping Algebra

Let $\mathfrak{g}$ be a n-dimensional ($n < \infty$) Lie algebra, $\{e_i\}_{1\leq i\leq n}$ a basis of $\mathfrak{g}$. Let $T(\mathfrak{g}) = \oplus_i T^i(\mathfrak{g})$ be the tensor algebra of the underlying vector space of $\mathfrak{g}$.

Proposition 2.14 *$T(\mathfrak{g})$ is an associative algebra with respect to the tensor product $\otimes\colon T^p(\mathfrak{g}) \times T^q(\mathfrak{g}) \to T^{p+q}(\mathfrak{g})$*

$$(x_1 \otimes \ldots x_p, y_1 \otimes \ldots y_q) \mapsto x_1 \otimes \ldots x_p \otimes y_1 \otimes \ldots y_q\ .$$

Remark 2.16 $T(\mathfrak{g})$ is generated as $T(\mathfrak{g}) = \mathbb{K} < e_1, \ldots, e_n >$ (in the algebra sense). Put differently, the tensor algebra is isomorphic (as a $\mathbb{K}$-algebra) to a free, associative, non-commutative $\mathbb{K}$-algebra on n-generators.

Consider a subspace $\mathcal{I} \subset T(\mathfrak{g})$, generated as

$$\mathcal{I} :=< t_1 \otimes ([x, y] - x \otimes y + y \otimes x) \otimes t_2 \mid t_1, t_2 \in T(\mathfrak{g}), x, y \in \mathfrak{g} >\ . \tag{38}$$

Note that $[x, y] - x \otimes y + y \otimes x \in \mathfrak{g} \oplus T^2(\mathfrak{g})$.

Proposition 2.15 *$\mathcal{I}$ is a (two-sided) ideal in $T(\mathfrak{g})$, i.e. $\forall j \in \mathcal{I}, t \in T(\mathfrak{g}) : j \otimes t, t \otimes j \in \mathcal{I}$.*

Proof The proof is obvious from the form of generators of $\mathcal{I}$. □

Remark 2.17 Let $\mathcal{A}$ be an associative algebra. We will denote by $\mathcal{A}_{\mathrm{Lie}}$ the corresponding Lie algebra $\mathcal{A}_{\mathrm{Lie}}$, with the Lie bracket given by the commutator $[x, y] := xy - yx$, for all $x, y \in \mathcal{A}$.

Definition 2.23 (Universal Enveloping Algebra) Let $\mathfrak{g}$ be a Lie algebra. The factor space

$$\mathcal{U}(\mathfrak{g}) := T(\mathfrak{g})/\mathcal{I}\ ,$$

where $\mathcal{I}$ is given by (38), is called the universal enveloping algebra of $\mathfrak{g}$.

Proposition 2.16 *$\mathcal{U}(\mathfrak{g})$ is an associative algebra. Moreover, $\mathcal{U}(\mathfrak{g})$ can be endowed with a Lie algebra structure, which is compatible with the Lie algebra structure on $\mathfrak{g}$.*

Proof Let $u_i = [t_i] \in \mathcal{U}(\mathfrak{g}), i = 1, 2$ and define $u_1 \cdot u_2. := [t_1 \otimes t_2]$. There is a canonical embedding of the field $\mathbb{K}$ and the Lie algebra $\mathfrak{g}$ in $\mathcal{U}(\mathfrak{g})$. Consider the canonical embedding ι of $\mathfrak{g}$ in $T(\mathfrak{g})$, and the canonical projection $\mathrm{pr}\colon T(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})$. Then the embedding of $\mathfrak{g}$ is given by the composition $\epsilon := \mathrm{pr} \circ \iota$,

$$\epsilon := \mathrm{pr} \circ \iota\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g})\ , \tag{39}$$

$x \mapsto u_x := x + J$, $J \in \mathcal{I}$ (and in the same way we can embed $\mathbb{K} \to \mathcal{U}(\mathfrak{g})$). We shall identify x and u_x. Let $x, y \in \mathfrak{g}$. Then $[x, y] - x \otimes y + y \otimes x \in \mathcal{I}$ and we have

$$\epsilon([x, y] - x \otimes y + y \otimes x) = u_{[x,y]} - u_x u_y + u_y u_x \overset{\text{notation}}{=} [x, y] - xy + yx = 0\,.$$

Hence $[x, y] = xy - yx$ in $\mathcal{U}(\mathfrak{g})$. This means that the Lie algebra structure on $\mathcal{U}(\mathfrak{g})$, which is given by the commutator, can be restricted to $\mathfrak{g}$ and the canonical embedding $\epsilon\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g})_{\mathrm{Lie}}$ is a homomorphism of Lie algebras, i.e.

$$\epsilon([x, y]) = [\epsilon(x), \epsilon(y)] = \epsilon(x)\epsilon(y) - \epsilon(y)\epsilon(x)$$

for all $x, y \in \mathfrak{g}$. □

Exercise 2.11 Show that the product in $\mathcal{U}(\mathfrak{g})$ is well-defined.

Universal Property of $\mathcal{U}(\mathfrak{g})$ Let $\epsilon\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ be the canonical Lie algebra homomorphism (39). The universal enveloping algebra satisfies the following universal property. For any associative, unital algebra $\mathcal{A}$ (over the same field as $\mathfrak{g}$) and any Lie algebra homomorphism $\psi\colon \mathfrak{g} \to \mathcal{A}_{\mathrm{Lie}}$, there is a unique associative, unital $\mathbb{K}$-algebra homomorphism $\tilde{\psi}\colon \mathcal{U}(\mathfrak{g})_{\mathrm{Lie}} \to \mathcal{A}_{\mathrm{Lie}}$ s.t. $\psi = \tilde{\psi} \circ \epsilon$, i.e. the following diagram commutes

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\epsilon"] \arrow[dr, "\psi"'] & \mathcal{U}(\mathfrak{g})_{\mathrm{Lie}} \arrow[d, dashed, "\exists!\tilde{\psi}"] \\ & \mathcal{A}_{\mathrm{Lie}} \end{tikzcd} \tag{40}$$

Remark 2.18 For $\mathfrak{g} = 0$ we have $\mathcal{U}(\mathfrak{g}) = \mathbb{K}$. For $\mathfrak{g}$ abelian (i.e. the bracket of $\mathfrak{g}$ is trivial), we have $\mathcal{U}(\mathfrak{g}) = \mathbb{K}[\mathfrak{g}]$, where $\mathbb{K}[\mathfrak{g}]$ is a polynomial ring over $\mathfrak{g}$.[7]

9.2 Poincaré-Birkhoff-Witt (PBW) Theorem

Theorem 2.4 (PBW Theorem) *Let $\mathfrak{g}$ be a $\mathbb{K}$-Lie algebra and $\{e_i\}_{1\le i\le n}$ be a basis of $\mathfrak{g}$ as a $\mathbb{K}$-vector space. Let $(\mathcal{U}(\mathfrak{g}), \epsilon)$ be the universal enveloping algebra. Then $\{1, \epsilon(e_i)\}_{1\le i\le n}$ is a basis of $\mathcal{U}(\mathfrak{g})$ as a $\mathbb{K}$-algebra and the canonical embedding $\epsilon\colon \mathfrak{g} \to \mathcal{U}(\mathfrak{g})$ is injective.*

Remark 2.19 We can rephrase the above theorem as follows. Let $\{e_i\}_{1\le i\le n}$ be a basis of $\mathfrak{g}$. For $k \in \mathbb{N}_0$ consider an index set $I_k = \{i_1, \dots i_k\}$, $I_0 = \emptyset$. We say that I_k is *increased* if $i_1 \le \cdots \le i_k$. Denote by $E_{I_k} \in \mathcal{U}(\mathfrak{g})$ the element $E_{I_k} = e_{i_1} \dots e_{i_k}$ and $E_{I_0} = 1$. Then the following set is a basis for $\mathcal{U}(\mathfrak{g})$ as a $\mathbb{K}$-vector space

---

[7] Every element of $\mathfrak{g}$ serves as a variable. One can think of $\mathbb{K}[\mathfrak{g}]$ as $\mathbb{K}[g_1, g_2, \dots]$ for all $g_i \in \mathfrak{g}$.

$$\{E_{I_k} \mid I_k \text{ increasing}, k \in \mathbb{N}_0\} .$$

In other words, elements of the form $e_1^{\alpha_1} \dots e_k^{\alpha_k}$, where $\alpha_1 \leq \dots \leq \alpha_n$, are linearly independent and generate $\mathcal{U}(\mathfrak{g})$.

Proposition 2.17 *$\mathcal{U}(\mathfrak{g})$ is a filtered algebra. The filtration is given by a degree of elements $x \in \mathcal{U}(\mathfrak{g})$, which is defined*

- *either as the degree of the polynomial, which describes x (after a choice of a base by PBW),*
- *or by the set $\mathcal{U}_p(\mathfrak{g})$, which is the image of $\oplus_{i=1}^p T^i(\mathfrak{g})$ under the canonical projection* $\mathrm{pr}\colon T(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g})$, *i.e. $\mathcal{U}_p(\mathfrak{g})$ is the set of elements of degree $\leq p$.*

Proof ***(Sketch of the Proof of the PBW Theorem)*** We will consider the case when $\mathfrak{g}$ comes equipped with a faithful representation.[8]

Consider the following algebra homomorphism

$$\tilde{\Delta}\colon T(\mathfrak{g}) \to T(\mathfrak{g}) \otimes T(\mathfrak{g}) ,$$

such that for all $t \in \mathfrak{g}$, $\tilde{\Delta}$ satisfies

$$\tilde{\Delta}(t) = t \otimes 1 + 1 \otimes t \ , t \in \mathfrak{g} \ .$$

Notice that $1 \otimes T(\mathfrak{g})$ and $T(\mathfrak{g}) \otimes 1$, commute with respect to the algebraic structure of $T(\mathfrak{g})$. □

Lemma 2.6 *The algebra homomorphism $\tilde{\Delta}$ descends to a map $\Delta\colon \mathcal{U}(\mathfrak{g}) \to \mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g})$ (Δ is a coproduct with respect to the Hopf algebra structure on $\mathcal{U}(\mathfrak{g})$).*

Proof of the Lemma The map Δ is given by the following commutative diagram

$$\begin{tikzcd}
\mathfrak{g} \arrow[r, "\epsilon"] \arrow[dr, "\iota"'] & \mathcal{U}(\mathfrak{g}) \arrow[r, dashed, "\Delta"] \arrow[d, "\tilde{\iota}"] & \mathcal{U}(\mathfrak{g}) \otimes \mathcal{U}(\mathfrak{g}) \\
& T(\mathfrak{g}) \arrow[r, "\tilde{\Delta}"] & T(\mathfrak{g}) \otimes T(\mathfrak{g}) \arrow[u, "\mathrm{pr} \otimes \mathrm{pr}"']
\end{tikzcd}$$

where ι and ϵ are the canonical embeddings of $\mathfrak{g}$ into the corresponding tensor algebra $T(\mathfrak{g})$ and the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$, respectively. The map $\tilde{\iota}$ is given by the universal property of $\mathcal{U}(\mathfrak{g})$ (consider $T(\mathfrak{g})$ with trivial bracket), and pr is the canonical quotient projection. All tensor products are considered over $\mathbb{K}$ (the field over which we consider $\mathfrak{g}$). □

[8] The kernel of the representation, as a homomorphism from $\mathfrak{g}$, is trivial. Example: $\mathfrak{g}$ acts faithfully on $C^\infty(\mathfrak{g})$. Another example: $\mathfrak{g}$ is a matrix algebra ($n \times n$ matrices), then $\mathfrak{g}$ acts faithfully on $\mathbb{R}^n$.

Since the basis of $\mathfrak{g}$ gives rise to the basis of $S(\mathfrak{g})$, another formulation of the PBW theorem is that $S(\mathfrak{g})$ gives rise to basis of $\mathcal{U}(\mathfrak{g})$. Let $\mathrm{Sym}(\mathfrak{g}) := \oplus_{k=0}^{\infty} \mathrm{Sym}^k(\mathfrak{g})$ be the space of symmetric tensors, where $t \in \mathrm{Sym}^k(\mathfrak{g})$ if

$$t(u_{\sigma(1)}, \ldots, u_{\sigma(k)}) = t(u_1, \ldots, u_k)$$

for all k-permutations σ. Clearly we have an injection $\mathrm{Sym}(\mathfrak{g}) \overset{j}{\hookrightarrow} T(\mathfrak{g})$. Too prove the PBW theorem, we prove that the composition

$$\mathrm{Sym}(\mathfrak{g}) \xrightarrow{j} T(\mathfrak{g}) \xrightarrow{\mathrm{pr}} \mathcal{U}(\mathfrak{g})$$

is injective.

Induction on k

1. Suppose $k = 1$. Then we can use the existence of faithful representation of $\mathfrak{g}$ on some vector space V. This amounts to the existence of an injective $\mathbb{K}$-algebra homomorphism $\rho\colon \mathfrak{g} \to \mathrm{End}(V)$. Note that $\mathrm{Sym}_{\leq 1}(\mathfrak{g}) = \mathfrak{g}$. Using the universal property of $\mathcal{U}(\mathfrak{g})$, we have the following commutative diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\epsilon"] \arrow[dr, "\rho"'] & \mathcal{U}(\mathfrak{g}) \arrow[d, dashed, "\tilde{\rho}"] \\ & \mathrm{End}(V) \end{tikzcd}$$

   Since $\rho = \tilde{\rho} \circ \epsilon$ and ρ is injective, ϵ is injective. Thus $\mathrm{pr} \circ \epsilon = \mathrm{pr} \circ j|_{\mathfrak{g}}$ is injective.
2. Take $\mathrm{Sym}_{\leq k}(\mathfrak{g}) := \oplus_{i=1}^{k} \mathrm{Sym}^i(\mathfrak{g})$ and suppose that the map $\mathrm{pr} \circ j$ is injective on $\mathrm{Sym}_{\leq k}(\mathfrak{g})$. Consider $f \in \mathrm{Sym}_{\leq k+1}(\mathfrak{g})$. We want to prove that if

$$(\mathrm{pr} \circ j)(f) = 0 \in \mathcal{U}(\mathfrak{g})$$

   then $f = 0$. Define

$$g := \tilde{\Delta}(f) - f \otimes 1 - 1 \otimes f\ .$$

   This implies that $g \in \mathrm{Sym}_{\leq k}(\mathfrak{g}) \otimes \mathrm{Sym}_{\leq k}(\mathfrak{g})$. Thus

$$(\mathrm{pr} \otimes \mathrm{pr}) \circ (j \otimes j)(g) = 0\ ,$$

   implies $g = 0$ by the induction hypothesis. But $g = 0$ is equivalent to

$$\tilde{\Delta}(f) = f \otimes 1 + 1 \otimes f\ .$$

   Now suppose $(\mathrm{pr} \otimes \mathrm{pr})\tilde{\Delta}(f) = 0$, which is equivalent to $(\mathrm{pr} \otimes \mathrm{pr})(f \otimes 1 + 1 \otimes f) = 0 \in \mathcal{U}(\mathfrak{g})$. This can happen if and only if $f = 0$. Hence $\mathrm{Sym}(\mathfrak{g})$ can be injectively embedded into $\mathcal{U}(\mathfrak{g})$.

Note 2.5 Note that the assumption $\operatorname{char}\mathbb{K} = 0$ is necessary in the above proof. To see that, suppose $\operatorname{char}\mathbb{K} = p$ and $\dim\mathfrak{g} > 1$. Then, using the binomial theorem, $\tilde{\Delta}(v^p) - v^p \otimes 1 - 1 \otimes v^p = (v \otimes 1 - 1 \otimes v)^p - v^p \otimes 1 - 1 \otimes v^p = 0$ for all $v \in \mathfrak{g}$.

Example Let $\dim\mathfrak{g} = 2$ and $\{e_1, e_2\}$ be a basis of $\mathfrak{g}$ such that $[e_1, e_2] = e_2$. Then we have the relation $e_1e_2 - e_2e_1 = e_2$ in $\mathcal{U}(\mathfrak{g})$, which means that $E = e_1^{i_1}e_2^{j_1}e_1^{i_2}e_2^{j_2}\cdots e_1^{i_k}e_2^{j_k}$ is a linear combination of $e_1^{\alpha_i}e_2^{\alpha_j}$. For example

$$
\begin{aligned}
E = e_1e_2^2e_1 &= e_1e_2e_2e_1 = e_1e_2(e_1e_2 - e_2) = e_1e_2e_1e_2 - e_1e_2^2 \\
&= e_1(e_1e_2 - e_2)e_2 - e_1e_2^2 = e_1^2e_2^2 - 2e_1e_2^2
\end{aligned}
$$

Example Consider the algebra $\mathfrak{g} = {}_2(\mathbb{C}) = \{A \in \operatorname{Mat}_2(\mathbb{C}) \mid \operatorname{tr} A = 0\}$, where tr is the trace of a matrix and the Lie bracket is given by the commutator. The Chevalley-Eilenberg basis of $\mathfrak{g}$ is given by the following matrices

$$
e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
$$

and the brackets are

$$
[e, f] = h, \qquad [h, e] = 2e, \qquad [h, f] = -2f \ .
$$

Then $\mathcal{U}(\mathfrak{sl}_2(\mathbb{C}))$ is the associative $\mathbb{C}$-algebra with generators e, f, h and relations

$$
ef - fe = h, \qquad he - eh = 2e, \qquad hf - fh = -2f \ .
$$

The basis of $\mathcal{U}(\mathfrak{sl}_2(\mathbb{C}))$ is given by monoms $e^\alpha h^\beta f^\gamma$, where $\alpha, \beta, \gamma \in \mathbb{N}_0$.

9.3 *Universal Enveloping and Differential Operators*

Consider a real, finite dimensional Lie algebra $\mathfrak{g}$ and the corresponding universal enveloping algebra $\mathcal{U}(\mathfrak{g})$. There is a natural filtration in $\mathcal{U}(\mathfrak{g})$ given by a sequence of subspaces $\{\mathcal{U}_k\}_{k\geq 0}$ s.t. $\mathcal{U}_k(\mathfrak{g}) \subset \mathcal{U}_{k+1}(\mathfrak{g})$ and $\mathcal{U}_k(\mathfrak{g})\cdot\mathcal{U}_l(\mathfrak{g}) \subset \mathcal{U}_{k+l}(\mathfrak{g})$. Moreover, if $\alpha \in \mathcal{U}_k(\mathfrak{g})$, $\beta \in \mathcal{U}_l(\mathfrak{g})$ then the commutator yields

$$
[\alpha, \beta] := \alpha\beta - \beta\alpha \in \mathcal{U}_{k+l-1}(\mathfrak{g}) \ . \tag{41}
$$

We define

$$\mathcal{A}_k := \mathcal{U}_k(\mathfrak{g})/\mathcal{U}_{k-1}(\mathfrak{g}) \ ,$$

and

$$\mathcal{A} := \bigoplus_{k\geq 0} \mathcal{A}_k \ .$$

The associative algebra structure of $\mathcal{U}(\mathfrak{g})$ induces a multiplication in $\mathcal{A}$, which is compatible with the order $\mathcal{A}_k \cdot \mathcal{A}_l \subset \mathcal{A}_{k+l}$. The bracket (41) implies commutativity of the multiplication

$$\mathcal{A}_k \cdot \mathcal{A}_l = \mathcal{A}_l \cdot \mathcal{A}_k \ ,$$

and defines a bracket operation

$$\{-,-\}\colon \mathcal{A}_k \times \mathcal{A}_l \to \mathcal{A}_{k+l-1}$$

on $\mathcal{A}$, making it an associative graded algebra.

Exercise 2.12 Verify that the above operations gives a Poisson algebra structure on $\mathcal{A}$.

Each component $\mathcal{A}_k$ of $\mathcal{A}$ is isomorphic to the space $\mathrm{Pol}_k(\mathfrak{g}^*)$ of degree k homogeneous polynomials on $\mathfrak{g}^*$ (the dual v. space of $\mathfrak{g}$). The algebra $\mathcal{A}$ is isomorphic to

$$\mathcal{A} \cong \mathrm{Pol}(\mathfrak{g}^*) := \bigoplus_{k\geq 0} \mathrm{Pol}_k(\mathfrak{g}^*) \ .$$

and the brackets are compatible with the brackets from $\mathfrak{g}^*$.

Now consider a smooth manifold M. Then the space $C^\infty(M)$ of smooth functions on M is an associative, unital, commutative algebra with respect to addition and multiplication of functions. Hence we can use the Definition 2.21 to define the filtered, associative algebra of differential operators

$$\mathrm{Diff}_*(M) := \bigcup_{k\geq 0} \mathrm{Diff}_k(M) \ ,$$

where $\mathrm{Diff}_k(M)$ is the set of differential operators of order $\leq k$ on M

$$\mathrm{Diff}_k(M) := \{\Delta\colon C^\infty(M) \to C^\infty(M) |\ \delta_{f_0} \circ \ldots \circ \delta_{f_k}(\Delta) = 0 \ , \forall f_i \in C^\infty(M)\} \ ,$$

and for $k = 0$ we define $\mathrm{Diff}_0(M) = C^\infty(M)$. Notice that $\delta_f(\Delta) = [\Delta, f]$ (see Lemma 2.4 for the definition of δ_f). Recall that the algebra structure of

$\mathrm{Diff}_*(M)$ is given by the composition of operators. In Example 32, we have seen that the commutator of operators defines a Lie algebra structure on $\mathrm{Diff}_*(M)$. The commutator acts with respect to filtered structure as follows. For $\delta \in \mathrm{Diff}_k$, $\nabla \in \mathrm{Diff}_l$ we have

$$[\delta, \nabla] = \delta \circ \nabla - \nabla \circ \delta \in \mathrm{Diff}_{m+k-1}(M) .$$

Remark 2.20 Since $\mathrm{Diff}_0(M) = C^\infty(M)$, we have $\mathrm{Diff}_1(M) \cong \mathrm{Diff}_0(M) \oplus \Gamma(M, TM)$, i.e. $\mathrm{Diff}_1(M)$ splits into functions and vector fields (seen as differential operators on $C^\infty(M)$.

Now we can define the Poisson algebra of degree 1 as in Definition 2.22

$$\mathrm{Smbl}(M) = \bigoplus_{k\geq 0} \mathrm{Smbl}_k ,$$

called the symbol algebra of $C^\infty(M)$, where $\mathrm{Smbl}_k = \mathrm{Diff}_k(M)/\,\mathrm{Diff}_{k-1}(M)$.

We shall now give another description of the Poisson algebra $\mathrm{Smbl}(M)$. Consider the space $\mathcal{P}_k(T^*M)$ of functions $p\colon T^*M \to \mathbb{R}$ such that $p \mapsto f(q, p)$ is, for every fixed $q \in M$, a degree k homogeneous polynomial on T_q^*M. Take

$$\mathcal{P}(T^*M) := \bigoplus_{k\geq 0} \mathcal{P}_k(T^*M) \subset C^\infty(T^*M) ,$$

which is the space of all fibre-wise polynomial functions on T^*M. Moreover, $\mathcal{P}(T^*M)$ is a Poisson subalgebra in $C^\infty(T^*M)$ (seen as a Poisson algebra with the standard Poisson brackets).

Theorem 2.5 *Poisson algebras* $\mathrm{Smbl}(M)$ *and* $\mathcal{P}(T^*M)$ *are isomorphic.*

Proof It is enough to prove this isomorphism on generators of $\mathrm{Smbl}_0(M)$ and $\mathrm{Smbl}_1(M)$ on one side, and on $\mathcal{P}_0(T^*M)$ and $\mathcal{P}_1(T^*M)$ on the other side. It is clear that $\mathrm{Smbl}_0(M) \cong \mathcal{P}_0(T^*M)$ since both of them are equal to $C^\infty(T^*M)$. Further, in Remark 2.20 we have seen that

$$\mathrm{Diff}_1(M) \cong \mathrm{Diff}_0(M) \oplus \Gamma(M, TM) ,$$

which gives

$$\mathrm{Smbl}_1(M) = \mathrm{Diff}_1(M)/\,\mathrm{Diff}_0(M) \cong \Gamma(M, TM) ,$$

Now consider a vector field $X \in \Gamma(M, TM)$ and define $f_X \in C^\infty(T^*M)$ by

$$f_X(q, p) :=< p, X_q >\in \mathbb{R} .$$

This gives us a bijection between $\Gamma(M, TM)$ and $\mathcal{P}_1(T^*M)$, since

$$< -, X_q >\colon T_q^*M \to \mathbb{R}$$

is a (homogeneous) 1-st order polynomial function (for every fixed $q \in M$). Moreover, for all vector fields $X, Y \in \Gamma(M, TM)$

$$f_{[X,Y]} := \{f_X, f_Y\} ,$$

so the bijection is an algebra homomorphism. For $F \in C^\infty(M)$, the function $f_F \in C^\infty(T^*M)$ is defined by

$$f_F(q, p) := F(q) .$$

Then the following identities hold for all $F, G \in C^\infty(M)$ and $X \in \Gamma(M, TM)$

$$\{f_X, f_F\} = f_{X(F)} ,$$
$$\{f_F, f_G\} = 0 .$$

This isomorphism can be extended to isomorphisms $\mathrm{Smbl}_k(M) \cong \mathcal{P}_k(T^*M)$ for all k and hence to isomorphism $\mathrm{Smbl}(M) \cong \mathcal{P}(T^*M)$. In other words, for all $k \geq 1$, the following sequence is exact

$$0 \longrightarrow \mathrm{Diff}_{k-1}(M) \xrightarrow{\ \iota\ } \mathrm{Diff}_k(M) \xrightarrow{\ \sigma\ } \mathcal{P}_k(T^*M) \longrightarrow 0 ,$$

where ι is the inclusion (given by the filtered structure of $\mathrm{Diff}_*(M)$) and σ is the *symbol map*. To each $\Delta \in \mathrm{Diff}_k(M)$, the symbol assigns a polynomial function $\sigma(\Delta)\colon T^*M \to \mathbb{R}$, which is fiberwise homogeneous of degree m. We will describe it in local coordinates $(\bar{q}, \bar{p})$ of T^*M, induced by the local coordinates $(\bar{q})$ on M. A differential operator $\Delta \in \mathrm{Diff}_k(M)$ has the form

$$\Delta = \sum_{|\alpha| \leq m} a_\alpha \partial^\alpha ,$$

where $\partial^\alpha := \partial_{q_1}^{\alpha_1} \circ \ldots \circ \partial_{q_k}^{\alpha_k}$, $\alpha_i \in \mathbb{N}$ and $a_\alpha \in C^\infty(M)$. Then, the symbol is

$$\sigma(\Delta(q, p)) = \sum_{|\alpha| = m} a_\alpha(q) p^\alpha ,$$

where $p^\alpha := p_1^{\alpha_1} \ldots p_k^{\alpha_k}$. Hence $\sigma(\Delta(q, p))$ is (for every q) a homogeneous polynomial function of degree m in the variable p. □

Remark 2.21 Let us describe briefly an *invariant* form of the symbol $\sigma(\Delta)$. Let $F \in C^\infty(M)$. Then $e^{tF}\Delta e^{-tF}$ is a differential operator of order $\leq k$ if $\Delta \in \mathrm{Diff}_k(M)$. Consider the following formal expression

$$e^{tF}\Delta e^{-tF} = \Delta + t[F, \Delta] + \frac{t^2}{2}[F, [F, \Delta]] + \ldots + \frac{t^k}{k!}[F, [F, \ldots, [F, \Delta]] \ldots] .$$

The above expression is finite, since the i-th term has degree $\leq i - k$. Then the symbol map gives the coefficient of the leading term of this expression, seen as a polynomial in t

$$\sigma(\Delta)(q, d_q\phi) = \frac{1}{k!}[F, [F, \ldots, [F, \Delta]] \ldots](q) .$$

Example Let $M = G$ be a Lie group with the Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$. Then $\mathcal{U}_k(\mathfrak{g}) \subset \mathrm{Diff}_k(G)$ as the left invariant differential operators of order $\leq k$ on G.

10 Poisson Manifolds

The following section will be focused on Poisson structure over smooth manifolds. An important example is the Poisson structure on the cotangent bundle of a smooth manifold.

10.1 Poisson Structure on the Cotangent Bundle

Let $F, G \in C^{\infty}(T^*M)$ be smooth functions on the cotangent bundle, and let (q^i, p_i) be local coordinates on T^*M. Define the *Poisson brackets* by

$$\{F, G\} = \sum_{i=1}^{n} (\frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i} - \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i}) . \tag{42}$$

This is a bilinear mapping (linear in both arguments F and G) which is *skew-symmetric*

$$\{F, G\} = -\{G, F\} \tag{43}$$

satisfies the *Leibniz rule*

$$\{F, GH\} = \{F, G\}H + G\{F, H\} , \tag{44}$$

and the *Jacobi identity*

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0 . \tag{45}$$

10.2 Poisson Manifolds

Generalizing the above idea of Poisson brackets on the cotangent bundle, we can give the following definition.

Definition 2.24 A Poisson bracket on a manifold M is a bilinear mapping

$$\{-,-\}\colon C^{\infty}(M) \times C^{\infty}(M) \to C^{\infty}(M) ,$$

such that (43), (44), and (45) are satisfied. A manifold M equipped with a Poisson bracket is called a Poisson manifold.

Remark 2.22 Every $F \in C^{\infty}(M)$, the mapping $X_F\colon C^{\infty}(M) \to C^{\infty}(M)$, defined via the Poisson bracket as

$$X_F := \{F, -\} ,$$

is a *derivation* on the algebra $C^{\infty}(M)$. More precisely, X_F is a 1st order differential operator, and hence can be considered as a vector field on M. This field is called *a Hamiltonian vector field* with *Hamiltonian* F.

10.3 Hamiltonian Mapping

For every $F, G \in C^{\infty}(M)$, The Leibniz rule (44) gives

$$X_{FG} = FX_G + GX_F$$

and one can define a mapping

$$\mathcal{H}\colon T^*M \to TM$$

by $\mathcal{H}(\mathrm{d}F) := X_F$, or

$$< \mathrm{d}G, \mathcal{H}(\mathrm{d}F) >:= \{F, G\} ,$$

where the bracket on the left-hand side denotes the evaluation of the 1-form $\mathrm{d}G$ on the vector field $\mathcal{H}(\mathrm{d}F)$. The map $\mathcal{H}$ is given in local coordinates (x^i) on M by a matrix

$$H^{kl} := \{x^k, x^l\} .$$

Denoting $\partial_k := \frac{\partial}{\partial x^k}$, the Poisson brackets can then be written as

$$\{F, G\} = \sum_{k,l} H^{kl} \partial_k F \partial_l G \ ,$$

and the components of the vector field X_F are

$$X_F^k = \sum_l H^{lk} \partial_k F \ .$$

Suppose the manifold is the cotangent bundle T^*M and the local coordinates (x^i) are (q^i, p_i), $i = 1, \ldots, n$. Then, the Hamilton map is given by the matrix

$$H = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \ ,$$

where I_n is the identity matrix, $2n = \dim T^*M$. In coordinates, the skew-symmetry of the Poisson bracket reads as

$$H^{kl} = -H^{lk} \ ,$$

and the Jacobi identity is

$$\sum_i (H^{ki} \partial_i H^{lm} + H^{li} \partial_i H^{mk} + H^{mi} \partial_i H^{kl}) = 0 \ .$$

If the matrix $H = (H^{kl})$ is non-degenerate, the mapping $\mathcal{H}\colon T^*M \to TM$ is invertible. If we denote by (Ω_{kl}) the inverse of H, then one can define a 2-form

$$\underline{\Omega} = \frac{1}{2} \sum \Omega_{kl} \mathrm{d}x^k \wedge \mathrm{d}x^l \ ,$$

or in more intrinsic way

$$\underline{\Omega}_x(X, Y) := < \mathcal{H}_x^{-1}(X), Y > \ ,$$

where $x \in M$ and $X, Y \in T_x M$. The Jacobi identity implies that

$$\mathrm{d}\underline{\Omega} = -\frac{1}{6} \sum_{k,l,m} \Omega_{klm} \mathrm{d}x^k \wedge \mathrm{d}x^l \wedge \mathrm{d}x^m \ ,$$

where

$$\Omega_{klm} := \partial_k \Omega_{lm} + \partial_l \Omega_{mk} + \partial_m \Omega_{kl} \ .$$

The above formula for $\mathrm{d}\underline{\Omega}$ can be proved using the following expression for Ω_{klm}

$$\Omega_{klm} := \sum_{p,q,r} \Omega_{kp}\Omega_{lq}\Omega_{mr} \left(\sum_i \mathcal{H}^{pi}\partial_i\mathcal{H}^{qr} + \mathcal{H}^{qi}\partial_i\mathcal{H}^{rp} + \mathcal{H}^{ri}\partial_i\mathcal{H}^{pq} \right) .$$

10.4 Poisson Bracket on a Symplectic Manifold

Definition 2.25 A smooth manifold M, equipped with a closed 2-form Ω ($\mathrm{d}\Omega = 0$) which is non-degenerate, i.e. if $X \in T_x U$, $x \in M$ s.t.

$$\Omega(X, Y) = 0$$

is called a symplectic manifold.

For every symplectic manifold, one can define the Poisson bracket on the algebra $C^\infty(M)$ by

$$\{F, G\} =< \mathrm{d}G, \mathcal{H}(\mathrm{d}F) >= \Omega\left(\mathcal{H}(\mathrm{d}F), \mathcal{H}(\mathrm{d}G)\right) .$$

That is, each symplectic manifold is a Poisson manifold. Let us emphasize that the inverse statement is not true, i.e. there are Poisson manifolds which are not symplectic.

10.5 Examples of Poisson and Symplectic Manifolds

Example (Cotangent Bundle) Let M be a smooth manifold, $\dim M = n$. Then the cotangent bundle T^*M has dimension (as a smooth manifold) $\dim T^*M = 2n$. The cotangent bundle comes equipped with the *canonical* (or *tautological*) *Liouville* 1*-form*, which is defined as follows. Let $\pi : TM \to M$, and $\pi_* : T^*M \to M$ denote the projections from the tangent and cotangent bundle, respectively, and by $\mathrm{d}\pi_*$ the tangent map to π_*. Consider the following commutative diagram

$$\begin{tikzcd} T(T^*M) \arrow[r, dashed, "\mathrm{d}\pi_*"] \arrow[d, "\tilde{\pi}"'] & TM \arrow[d, "\pi"] \\ T^*M \arrow[r, "\pi_*"'] & M \end{tikzcd}$$

(continued)

where $\tilde{\pi}$ is the tangent bundle projection down to T^*M (seen as manifold). Then the Liouville 1-form $\rho \in \Omega^1(T^*M)$ is defined for arbitrary $X \in T(T^*M)$ as

$$\rho(X) := < \pi(X), \mathrm{d}\pi_*(X) > \ .$$

Then $\Omega := \mathrm{d}\rho$ is a symplectic form on T^*M. In canonical coordinates,

$$\Omega = \mathrm{d}\rho = \mathrm{d}(\sum_{i=1}^{n} p_i \mathrm{d}q^i) = \sum_i \mathrm{d}p_i \wedge \mathrm{d}q^i \ .$$

Poisson brackets given by this symplectic structure are given by (42).

Example (Sphere and Projective Space) Let $S^2 \subset \mathbb{R}^3$ be the unit 2-sphere, i.e. solution of the equation $x^2 + y^2 + z^2 = 1$. Then the 2-form

$$\Omega = x\mathrm{d}y \wedge \mathrm{d}z + y\mathrm{d}z \wedge \mathrm{d}x + z\mathrm{d}x \wedge \mathrm{d}y \ ,$$

where the values of x, y, z are assumed to satisfy the equation for S^2, defines a symplectic form on S^2. Note that one can vies the 2-sphere can be identified with the complex projective line $S^2 = \mathbb{P}^1(\mathbb{C})$).

Consider now $\mathbb{C}^n$ with the Hermitian form

$$< z, w >= \sum_{k=1}^{n} \bar{z}^k w^k \ .$$

Define a 1-form

$$\rho := i < \bar{z}, \mathrm{d}z >= i \sum_{k=1}^{n} \bar{z}^k \mathrm{d}z^k \ ,$$

where $i = \sqrt{-1}$, and the 2-form

$$\Omega = \mathrm{d}\rho = i\mathrm{d} < \bar{z}, \mathrm{d}z >= i \sum_{k=1}^{n} \mathrm{d}\bar{z}^k \wedge \mathrm{d}z^k \ .$$

The 2-form Ω is real, i.e. it can be written as

$$\Omega = -2 \sum_{k=1}^{n} \mathrm{d}x^k \wedge \mathrm{d}y^k \ ,$$

(continued)

for $x^k + iy^k = z^k$. Hence $(\mathbb{C}^n, \Omega)$ is a *real symplectic* manifold. Poisson brackets are

$$\{z^j, z^k\} = \{\bar{z}^j, \bar{z}^k\} = 0 \qquad\qquad \{z^j, \bar{z}^k\} = i\delta^{jk} .$$

If we fix a quadratic Hamiltonian

$$H(z) :=< z, z >= \sum_{k=1}^{n} |z^k|^2$$

on $\mathbb{C}^n$, then the level set $H(z) = 1$ is a $(2n-1)$-dimensional sphere S^{2n-1}. The Hamiltonian vector field X_H on $\mathbb{C}^n$ restricted to S^{2n-1} defines a vector field pX_H. Recall that $\mathrm{U}(1) \cong \mathbb{R}^*$ acts on S^{2n-1} by $(\varphi, z) \mapsto e^{-i\varphi}$ and the quotient by this action is the complex projective space, $S^{2n-1}/\mathbb{R}^* \cong \mathbb{P}(\mathbb{C}^n)$. The 2-form Ω induces a 2-form on S^{2n-1}, which is the lift of a unique 2-form on $\mathbb{P}(\mathbb{C}^n)$, making it a real symplectic manifold. More precisely

$$\Omega_{\mathbb{P}(\mathbb{C}^n)} = \phi^*(\Omega|_{S^{2n-1}}) ,$$

where $\phi\colon \mathbb{C}^n \setminus \{0\} \to S^{2n-1}$, $\phi(z) = |z|^{-1}z$.

More details about Poisson manifolds can be found, for example, in [2].

10.6 Poisson Manifolds and Lie Theory

Dual Space of a Lie Algebra Let $\mathfrak{g}$ be a finite dimensional Lie algebra over $\mathbb{R}$, say $\dim \mathfrak{g} = n$, and $\mathfrak{g}^* = \hom(\mathfrak{g}, \mathbb{R})$ the dual vector space. Consider the space of smooth functions $C^\infty(\mathfrak{g}^*)$ and observe that one can embed $\mathfrak{g}^* \to C^\infty(\mathfrak{g}^*)$ as a subspace of linear functions

$$\mathfrak{g}^* = \{F_X \in C^\infty(\mathfrak{g}^*) |\ F_X(\xi) = \xi(X), X \in \mathfrak{g}, \xi \in \mathfrak{g}^*\} .$$

There is a *unique Poisson bracket* on $\mathfrak{g}^*$ such that

$$\{F_X, F_Y\} = F_{[X,Y]} , \tag{46}$$

meaning that the assignment $X \mapsto F_X$ yields a Lie algebra homomorphism $\mathfrak{g} \to C^\infty(\mathfrak{g}^*)$.

Description in Coordinates—Brackets Let $[X^k, X^l] = c^{kl}_m X^m$, where $X^k \in \mathfrak{g}$ for all k. Then functions $\xi^1 = F_{X^1}, \dots \xi^n = F_{X^n}$ form a system of linear coordinates on $\mathfrak{g}^*$, and we can define a Poisson bracket on $\mathfrak{g}^*$ via the Poisson brackets on $C^\infty(\mathfrak{g}^*)$

$$\{F, G\} = \sum_{k,l,m} C^{kl}_m \frac{\partial F}{\partial \xi^k} \frac{\partial G}{\partial \xi^l} \xi^m$$

so that

$$\{\xi^k, \xi^l\} = C^{kl}_m \xi^m \ .$$

Description in Coordinates—Gradient For any $F \in C^\infty(\mathfrak{g}^*)$, its *gradient*

$$\nabla F \colon \mathfrak{g}^* \to \mathfrak{g}$$

is a $\mathfrak{g}$-valued function on $\mathfrak{g}^*$ defined by

$$< \nabla F(\xi), \eta >:= \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} F(\xi + t\eta) \ ,$$

where $\xi, \eta \in \mathfrak{g}^*$. In the coordinates

$$\nabla F = \sum_k \frac{\partial F}{\partial \xi^k} X^k \ ,$$

which implies

$$\{F, G\}(\xi) =< \xi, [\nabla F(\xi), \nabla G(\xi)] > \ .$$

G-invariant Mappings on the T^*G Let $\mathfrak{g} = \mathrm{Lie}(G)$ be a finite dimensional Lie algebra of a Lie group G. Every Lie group is parallelizable, i.e. there is a vector bundle isomorphism between T^*G and $G \times \mathfrak{g}^*$ as vector bundles over G. This is expressed by the following commutative diagram

$$\begin{tikzcd} T^*G \arrow[r, "\varphi"] \arrow[dr, "\pi"'] & G \times \mathfrak{g}^* \arrow[d, "\pi_1"] \\ & G \end{tikzcd}$$

where φ is the bundle isomorphism and π_1 is the projections on the 1st factor. We will denote by $\pi_2 \colon G \times \mathfrak{g}^* \to \mathfrak{g}^*$ the projection on the second factor. Let $p \in T^*G$. Then

$$\varphi(p) = (h, \xi) \tag{47}$$

for some $h \in G$ and $\xi \in \mathfrak{g}^*$. Now consider the left translation automorphism $l_g\colon G \to G$, given by $l_g(h) = gh$. The tangent map of l_g is denoted $(l_g)_*\colon TG \to TG$ and the cotangent map by $(l_g)^*\colon T^*G \to T^*G$. Notice that for $p \in T^*G$ we have

$$\left(\pi_2 \circ \varphi \circ (l_g)^*\right)(p) = \pi_2(g^{-1}h, \xi) = \xi \; . \tag{48}$$

This follows from the fact that $\mathfrak{g}^*$ can be identified with the space of left-invariant 1-forms on G, i.e. 1-forms η such that $(l_g)^*\eta = \eta$. Thus from (47) and (48), we have

$$\pi_2 \circ \varphi \circ (l_g)^* = \pi_2 \circ \varphi \; . \tag{49}$$

For the following considerations, we recall that any smooth map between manifolds, say $F\colon M \to N$, induces the corresponding algebra homomorphism in the opposite direction, $F^\bullet\colon C^\infty(N) \to C^\infty(M)$, which is given by the precomposition. Because $\mathfrak{g}^*$ is a smooth manifold (as a vector space with global coordinates), we have the homomorphism

$$(\pi_2 \circ \varphi)^\bullet\colon C^\infty(\mathfrak{g}^*) \to C^\infty(T^*G) \; .$$

We want to show that the algebra $C^\infty(\mathfrak{g}^*)$ can be identified with the G-invariant subspace of the algebra $C^\infty(T^*G)^G$, i.e.

$$C^\infty(\mathfrak{g}^*) \cong C^\infty(T^*G)^G \; ,$$

where

$$C^\infty(T^*G)^G = \{F \in C^\infty(T^*G)^G \,|\, \forall g \in G : F \circ (l_g)^* = F\} \; . \tag{50}$$

To see this, consider $F \in C^\infty(\mathfrak{g}^*)$ and denote

$$\tilde{F} := (\pi_2 \circ \varphi)^\bullet F = F \circ \pi_2 \circ \varphi \in C^\infty(T^*G) \; .$$

Then $\tilde{F}$ is G-invariant in the sense of (50). Indeed, using (48), we obtain

$$\tilde{F} \circ (l_g)^* = F \circ \pi_2 \circ \varphi \circ (l_g)^* = F \circ \pi_2 \circ \varphi = \tilde{F} \; ,$$

so we have an injection $C^\infty(\mathfrak{g}^*) \hookrightarrow C^\infty(T^*G)^G$. On the other hand, consider $\tilde{H} \in C^\infty(T^*G)^G$. Then we can define $H \in C^\infty(\mathfrak{g}^*)$

$$H(\xi) := \tilde{H}\left(\varphi^{-1}(e, \xi)\right), \xi \in \mathfrak{g}^* \; ,$$

where φ is the trivialization of T^*G as described in the above commutative diagram. The above definition of H is unambiguous due to G-invariance of $\tilde{H}$. This assignment $\tilde{H} \mapsto H$ is obviously injective, hence the bijection between $C^\infty(\mathfrak{g}^*)$ and $C^\infty(T^*G)^G$. The brackets are preserved as well

$$\{F, H\} \circ \pi_2 = \{F \circ \pi_2, H \circ \pi_2\}$$

for all $F, H \in C^\infty(\mathfrak{g}^*)$.

Remark 2.23 For a Lie group G, the cotangent bundle T^*G is a symplectic manifold with non-degenerate Poisson structure (since G is a smooth manifold).

Remark 2.24 $\mathfrak{g}^*$ does not have a structure of a Lie algebra, although it is a dual vector space to the Lie algebra $\mathfrak{g}$, it is not a coalgebra. On the other hand, $C^\infty(\mathfrak{g}^*)$ is a Lie algebra.

One More Example of a Symplectic Manifold Let G be a Lie group, $\mathfrak{g}$ its Lie algebra, and $\mathfrak{g}^*$ the corresponding dual vector space. The group G acts on $\mathfrak{g}$ by the *adjoint action* $\mathrm{Ad}\colon G \to \mathrm{Aut}(\mathfrak{g})$. At $g \in G$, the map

$$\mathrm{Ad}_g\colon \mathfrak{g} \to \mathfrak{g}$$

is the derivative at the identity element $e \in G$ of the conjugation map $G \to G$, given by $h \mapsto ghg^{-1}$. One can differentiate Ad at e, to obtain the *adjoint action of* $\mathfrak{g}$ *on* $\mathfrak{g}$, i.e the map

$$\mathrm{ad}\colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g})\ .$$

In fact, one can show that $\mathrm{ad}\colon \mathfrak{g} \to \mathcal{D}(\mathfrak{g})$, where $\mathcal{D}(\mathfrak{g})$ is the algebra of derivations on $\mathfrak{g}$, which is the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$. For $X, Y \in \mathfrak{g}$, the ad map acts as the Lie bracket

$$\mathrm{ad}_X(Y) = [X, Y]\ .$$

Using Ad, one can define the *coadjoint action of* G *on* $\mathfrak{g}^*$, denoted $\mathrm{Ad}^*\colon G \to \mathcal{A}ut(\mathfrak{g}^*)$, by

$$\mathrm{Ad}^*_g(\xi)(X) = \xi(\mathrm{Ad}_{g^{-1}}(X)),$$

where $\xi \in \mathfrak{g}^*$. Finally, using Ad^*, we can define the *coadjoint action of* $\mathfrak{g}$ *on* $\mathfrak{g}^*$, denoted $\mathrm{ad}^*\colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g}^*)$, as

$$(\mathrm{ad}^*_X(\xi)(Y)) := -\xi([X, Y])\ ,$$

$X, Y \in \mathfrak{g}$ and $\xi \in \mathfrak{g}^*$.

10.7 Symplectic Foliation on $\mathfrak{g}^*$

The brackets on $\mathfrak{g}^*$ are non-symplectic.

Symplectic Structure on the Coadjoint Orbit Now suppose that G is a connected Lie group and denote by $\mathcal{O} \subset \mathfrak{g}^*$ the G-orbit of the coadjoint action Ad^*. Then for all $\xi \in \mathcal{O}$, the tangent spaces of $\mathcal{O}$ at ξ is given by

$$T_\xi \mathcal{O} = \{\mathrm{ad}_X^*(\xi) \in \mathfrak{g}^* | X \in \mathfrak{g}\} .$$

One can define a symplectic structure on $\mathcal{O}$:

$$\Omega_\xi : T_\xi \mathcal{O} \times T_\xi \mathcal{O} \to \mathbb{R} ,$$

as follows. For arbitrary $(X, \xi), (Y, \xi) \in T_\xi \mathcal{O}$, define

$$\Omega_\xi((X, \xi), (Y, \xi)) := \xi(\mathrm{ad}_X(Y)) = \xi([X, Y]) . \tag{51}$$

This structure is usually attributed to Kostant [14], Kirillov [15, 16], and Souriau [17]. For brevity, we will refer to this symplectic structure as KKS.

10.7.1 Coadjoint Invariant Functions

A function $F \in C^\infty(g^*)$ such that

$$\{F, H\} = 0 \, \forall H \in C^\infty(g^*)$$

is called a *Casimir function*. Let F be a Casimir function. Then for arbitrary $X \in \mathfrak{g}$, the corresponding linear function[9] F_X satisfy

$$\{F, F_X\} = -\{F_X, F\} = -\xi_{F_X}(F) = \mathrm{ad}_X^*(F) = 0 ,$$

where $\xi_{F_X} = \{F_X, -\}$ is the Hamiltonian function corresponding to F_X. This means that the Casimir functions are *coadjoint-invariant functions*.

Remark 2.25 If G is a semisimple Lie group, then there is an open denset subset $U \subset \mathfrak{g}^*$, stable under G, and the orbits are separated by Casimir functions. This yields a foliation of $\mathfrak{g}^*$. Let us note that this foliation is not a smooth fibration and has a rather complicated structure.

Poisson Bracket on $\mathfrak{g}^*$ and $\mathcal{O}$ There is a link between the Poisson bracket on $\mathfrak{g}^*$ (46) and on the coadjoint orbit $\mathcal{O}$. Denote by $\iota \colon \mathcal{O} \to \mathfrak{g}^*$ the embedding of the

---

[9] This is the evaluation map $F_X : \mathfrak{g}^* \to \mathbb{R}$ (defined above), $F_X(\xi) = \xi(X)$.

orbit into $\mathfrak{g}^*$, and by $\iota^*\colon C^\infty(\mathfrak{g}^*) \to C^\infty(\mathcal{O})$ the corresponding algebra map. Then we have

$$\iota^*\{f, g\}_{C^\infty(\mathfrak{g}^*)} = \{\iota^* f, \iota^* g\}_{C^\infty(\mathcal{O})} .$$

The left-hand side is the Poisson bracket on $C^\infty(\mathfrak{g}^*)$, restricted to $\mathcal{O}$. On the right-hand side is the *symplectic Poisson bracket* on $C^\infty(\mathcal{O})$, given by the symplectic structure Ω on $\mathcal{O}$, given by (51). We will now describe two examples of this construction.

Example Let $G = \mathrm{SO}(3)$. The Lie algebra is $\mathfrak{g} = \mathfrak{so}(3)$ and it is isomorphic with its dual $\mathfrak{so}(3) \cong \mathfrak{so}^*(3)$ via the Killing form. Moreover, $\mathfrak{so}^*(3) \cong \mathbb{R}^3$. The orbits of the coadjoint action are described by the equation $x^2+y^2+z^2 = R^2$. Hence for $R = 0$ we have a 0-dimensional symplectic leaf (a e point). For $R > 0$ we have 2-dimensional symplectic leaves (the coadjoint orbits). On $\mathbb{R}^3$ we have the volume form $\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$, which if restricted to the orbits, yields

$$\Omega_R = \frac{1}{R^2}(x\mathrm{d}y \wedge \mathrm{d}z + y\mathrm{d}z \wedge \mathrm{d}x + z\mathrm{d}x \wedge \mathrm{d}y)$$

The Casimir functions on $\mathfrak{so}(R)$ are given by $F = x^2 + y^2 + z^2$. Hence the Casimir elements are parametrized by $R > 0$.

Example Consider $G = \mathrm{SL}_2(\mathbb{R})$, given by

$$\mathrm{SL}_2(\mathbb{R}) = \{A \in \mathrm{Mat}_2(\mathbb{R}) |\ \det A = 1\} ,$$

The Lie algebra is

$$\mathfrak{sl}_2(\mathbb{R}) = \{A \in \mathrm{Mat}_2(\mathbb{R}) | \operatorname{tr} A = 0\} \cong \mathbb{R}^3 .$$

The $\mathrm{SL}_2(\mathbb{R})$-orbits in $\mathbb{R}^3$ are cones and hyperboloids, with the symplectic structure given by KKS structure.

Remark 2.26 For a finite dimensional Lie algebra $\mathfrak{g}$ and the corresponding linear extension of the bracket to the symmetric algebra $S(\mathfrak{g})$, the Casimir elements can be identified with the $\mathfrak{g}$-invariant subspace of $S(\mathfrak{g})$

$$\mathrm{Cas}(\mathfrak{g}) := \mathrm{Z}(\mathcal{U}(\mathfrak{g})) \cong S(\mathfrak{g})^{\mathfrak{g}}$$

where $\mathrm{Z}(\mathcal{U}(\mathfrak{g}))$ denotes the center of the universal enveloping algebra of $\mathfrak{g}$ (see 40). For a reductive $\mathfrak{g}$

$$HP^k\,(S(\mathfrak{g})) \cong H^k_{CE}(\mathfrak{g}) \otimes_{\mathbb{K}} \mathrm{Cas}(\mathfrak{g})\ ,$$

where H_{CE} is the Chevalley-Eilenberg cohomology.

11 Differential Calculus on Poisson Manifolds

Let M be a smooth manifold. Then there is a natural isomorphism $P_m \cong \Gamma(M, S^m(TM))$, where $S^m(TM)$ is the m-symmetric power of the tangent bundle TM. Then the Poisson brackets on $C^\infty(T^*M)$ can be reformulated as Poisson brackets on symmetric tensor fields.

We also consider *m-vector fields* over M [18], i.e. sections of the m-th exterior bundle

$$\Lambda^m(TM) \to M\ ,$$

The set of all m-vector fields is denoted $\mathfrak{X}^m(M) := C^\infty(M, \Lambda^m(TM))$. For $m = 1$ we obtain the standard notion of vector fields, i.e. $\mathfrak{X}^1(M) = C^\infty(M, TM)$. Taking the *Whitney sum* of all the exterior powers of TM gives the exterior bundle

$$\Lambda(TM) := \oplus_{m\geq 0}\Lambda^m(TM) \to M\ .$$

Sections of this bundle are *multivector fields*. The set of all multivector fields, $\mathfrak{X}(M) := C^\infty(M, \Lambda(TM))$, is an algebra with respect to the exterior product (37). There is a dual construction leading to differential forms over M. If we start with T^*M instead of TM, we get *differential m-forms over M* (shortly just *m-forms*) as sections of the bundle

$$\Lambda^m(T^*M) \to M\ .$$

The set of all differential m-forms is denoted $\Omega^m(M) := C^\infty(M, \Lambda^m(T^*M))$. Note that m-vector fields are skew-symmetric polylinear functions on 1-forms. Taking sections of the bundle of all m-forms, $0 \leq m \leq \dim M$,

$$\Lambda(T^*M) := \oplus_m\Lambda^m(T^*M) \to M\ ,$$

we obtain the algebra (with respect to the exterior product) of differential forms over M

$$\Omega(M) = \oplus_{m\geq 0}\Omega^m(M)\ ,$$

Remark 2.27 Both algebras $\mathfrak{X}(M)$, and $\Omega(M)$, are C^∞-modules.

Now suppose that $(M, \{\cdot,\cdot\})$ is a Poisson manifold, $\mathcal{H}\colon T^*M \to TM$ the corresponding Hamiltonian mapping. One can consider a bivector field $\pi_* \in \mathfrak{X}^2(M)$, such that

$$\{F, G\} := \sum_{k,l} H^{kl}\partial_k F \partial_l G\ .$$

We would like to see the above studied Jacobi identity in a more invariant, intrinsic way using certain structure on the space of multivector fields $\mathfrak{X}(M)$. This structure is called a *Lie super-Schouten bracket*, or *Schouten-Nijenhuis bracket*, or shortly just *Schouten bracket*,

$$[\![-,-]\!]\colon \mathfrak{X}^p(M)\times\mathfrak{X}^q(M)\to\mathfrak{X}^{p+q-1}(M)\ .$$

We define it inductively using the following properties of multivector fields

1. $s\wedge t = (-1)^{|s||t|}t\wedge s$, where $|s|$ is the degree[10] of s,
2. $[\![s,t]\!] = (-1)^{|s|(|t|+1)+|t|}[\![t,s]\!]$ (graded commutativity),
3. $[\![s,t\wedge r]\!] = [\![s,t]\!]\wedge r + (-1)^{|t|(|s|-1)}t\wedge[\![s,r]\!]$,
4. $[\![s\wedge t,r]\!] = s\wedge[\![t,r]\!] + (-1)^{|t|(|r|-1)}[\![s,r]\!]\wedge t$.

For degrees ≤ 1 we further define

1. $|s| = |t| = 1 : [\![s,t]\!] := [s,t]$, where $[-,-]$ is the Lie bracket
2. $|s| = 1, |t| = 0 : [\![s,t]\!] := s(t) = \mathcal{L}_s(t)$, where $\mathcal{L}$ is the Lie derivative
3. $|s| = |t| = 0 : [\![s,t]\!] := 0$.

Multivector fields s and t are called decomposable if

$$s = s_1\wedge\ldots\wedge s_p \qquad\qquad t = t_1\wedge\ldots\wedge t_q\ ,$$

for some p, q, and where all $s_i, t_j \in \mathfrak{X}^1(M)$. In this case

$$[\![s,t]\!] = \sum_{k=1}^{p}\sum_{l=1}^{q}[s_k,t_l]\wedge s_1\wedge\ldots\wedge\hat{s_k}\wedge\ldots\wedge s_p\wedge t_1\wedge\ldots\wedge\hat{t_l}\wedge\ldots\wedge t_q\ .$$

There is at most one Schouten bracket satisfying the above properties.

Once we properly defined the space of multivectors on a manifold and the notion of Schouten bracket, we can introduce the notion of Poisson tensor.

[10] Meaning $s\in\mathfrak{X}^{|s|}(M)$.

Definition 2.26 (Poisson Tensor) Let M be a Poisson manifold with Poisson bracket $\{-,-\}$ and with the Schouten bracket $[\![-,-]\!]$. A 2-vector field $\pi \in \mathfrak{X}(M)$ is called a Poisson tensor, if for all $f, g \in C^\infty(M)$

$$\{f, g\} = [\![[\![\pi, f]\!], g]\!] .$$

Theorem 2.6 *The Poisson bracket* $\{-,-\}$ *satisfies the Jacobi identity iff* $[\![\pi, \pi]\!] = 0$.

11.1 Coordinate-Free Construction of the Schouten Bracket

Consider the exterior algebra $\Omega(M)$. Let $\omega \in \Omega^p(M)$ and $X \in \mathfrak{X}^1(M)$. We associate with X two operators on $\Omega(M)$

$$\begin{aligned} \iota_X &: \Omega^p(M) \to \Omega^{p-1}(M) && \text{(interior product)} , \\ \mathcal{L}_X &: \Omega^p(M) \to \Omega^p(M) && \text{(Lie derivative)} . \end{aligned}$$

The interior product is defined by

$$(\iota_X \omega)(X_2, \ldots, X_p) := \omega(X_1, X_2, \ldots, X_p) ,$$

where all $X_i \in \mathfrak{X}^1(M)$. The Lie derivative is defined as

$$\mathcal{L}_X \omega := \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \phi_t^*(\omega) ,$$

where ϕ_t^* denotes the pullback along the flow ϕ_t of X. This means

$$(\mathcal{L}_X \omega)(X_1, \ldots, X_p) = \mathcal{L}_X(\omega(X_1, \ldots, X_p)) - \sum_k \omega(X_1, \ldots, [X, X_k], \ldots, X_p) .$$

We denote by d the de Rham differential $\mathrm{d} \colon \Omega^p(M) \to \Omega^{p+1}(M)$, given by

$$\begin{aligned} \mathrm{d}\omega(X_1, \ldots, X_{p+1}) = {} & \sum_k \mathcal{L}_{X_k}(\omega(X_1, \ldots, \hat{X}_k, \ldots, X_{p+1})) \\ & + \sum_{k<l} (-1)^{k+l} \omega([X_k, X_l], X_1, \ldots, \hat{X}_k, \ldots, \hat{X}_l, \ldots, X_{p+1}) . \end{aligned}$$

Theorem 2.7 (Cartan Triple $(\iota_X, \mathcal{L}_X, \mathrm{d})$**)** *The following identities are always satisfied*

1. $\iota_X \circ \iota_Y + \iota_Y \circ \iota_X = 0$,
2. $[\mathcal{L}_X, \mathcal{L}_Y] = \mathcal{L}_{[X,Y]}$,

3. $[\mathcal{L}_X, \iota_Y] = \iota_{[X,Y]}$,
4. $[\mathcal{L}_X, \mathrm{d}] = 0$,
5. $d^2 = \mathrm{d} \circ \mathrm{d} = 0$,
6. $\mathcal{L}_X = \iota_x \circ \mathrm{d} + \mathrm{d} \circ \iota_X$.

We want now to define the analogue of this "differential calculus" for multivectors. For decomposable $t \in \mathfrak{X}^p(M)$, $t = X_1 \wedge \ldots \wedge X_p$, we define

$$\iota_t = \iota_{X_1} \circ \ldots \circ \iota_{X_p} .$$

The above can be extend by linearity for arbitrary $t \in \mathfrak{X}^p(M)$. This definition is correct because of the first property of the above theorem and yields an operator

$$\iota_t \colon \Omega^m(M) \to \Omega^{m-p}(M) ,$$

which further satisfies

$$\iota_{t \wedge s} = \iota_t \circ \iota_s .$$

To define the Lie derivative $\mathcal{L}_t$ along a multivector field, one can start with the notion of a *graded operator*

$$\partial \colon \Omega^m(M) \to \Omega^{m+r}(M).$$

In this case, ∂ is called *a graded operator of degree* $\deg \partial = |\partial| = r$. If in addition ∂ satisfies

$$\partial(\omega \wedge \theta) = \partial\omega \wedge \theta + (-1)^{mr} \omega \wedge \partial\theta,$$

where $\omega \in \Omega^m(M)$, then ∂ is called a *graded derivative of degree* r. For example, ι_x, $\mathcal{L}_x$, d are graded derivatives of the following degrees $|\iota_x| = -1$, $|\mathcal{L}_x| = 0$, $|\mathrm{d}| = 1$.

We recall that *graded brackets* of two operators are defined by

$$[\Delta, \nabla] = \Delta \circ \nabla - (-1)^{|\Delta||\nabla|} \nabla \circ \Delta .$$

Remark 2.28 If ∂ and D are graded operators of degrees $|\partial|$ and $|D|$, respectively, then $[\partial, D]$ is a graded derivation of degree $|\partial| + |D|$.

The property 6. of the previous theorem now reads

$$[\iota_X, \mathrm{d}] = \iota_X \circ \mathrm{d} + \mathrm{d} \circ \iota_X = \mathcal{L}_X$$

for a vector field X. Using the graded bracket we now define the Lie derivative along a multivector field t

$$\mathcal{L}_t := [\iota_t, \mathrm{d}] .$$

If $t \in \mathfrak{X}^p(M)$ then

$$\mathcal{L}_t \colon \Omega^m(M) \to \Omega^{p-m+1}(M) \,,$$

meaning that $\mathcal{L}_t$ is a graded derivation of degree $|\mathcal{L}_t| = m - 1$.

We can proceed with the definition of a bracket on the algebra of multivectors (denoted by the same symbol as the Schouten bracket)

$$[\![-,-]\!] \colon \mathfrak{X}^p(M) \times \mathfrak{X}^q(M) \to \mathfrak{X}^{p+q-1}(M) \,.$$

If $s \in \mathfrak{X}^p(M), t \in \mathfrak{X}^q(M)$, then there is a unique multivector field $[\![t,s]\!] \in \mathfrak{X}^{p+q-1}(M)$ satisfying

$$\iota_{[\![s,t]\!]} = [\mathcal{L}_s, \iota_t] \,,$$

where the right-hand side is given by the graded commutator, i.e.

$$[\mathcal{L}_s, \iota_t] = \mathcal{L}_s \circ \iota_t - (-1)^{(p-1)q} \iota_t \circ \mathcal{L}_s \,.$$

When the multivectors are decomposable as $s = s_1 \wedge \ldots \wedge s_p$ and $t = t_1 \wedge \ldots \wedge t_q$ then

$$[\![s,t]\!] = \sum_{k=1}^{p} \sum_{l=1}^{q} [s_k, t_l] \wedge s_1 \wedge \ldots \wedge \hat{s_k} \wedge \ldots \wedge s_p \wedge t_1 \wedge \ldots \wedge \hat{t_l} \wedge \ldots \wedge t_q \,.$$

This defines $[\![-,-]\!]$ on the whole $\mathfrak{X}(M)$ by bilinear extension. Moreover, it satisfies the *graded commutativity*

$$[\![s,t]\!] = -(-1)^{(|s|-1)(|t|-1)} [\![t,s]\!] \,,$$

as well as the analogy of the second property of the Theorem 2.7

$$[\mathcal{L}_s, \mathcal{L}_t] = \mathcal{L}_{[\![s,t]\!]} \,,$$

and also satisfies the *graded Jacobi identity*

$$0 = (-1)^{(|s|-1)(|r|-1)} [\![s, [\![t,r]\!]]\!] + (-1)^{(|t|-1)(|s|-1)} [\![t, [\![r,s]\!]]\!] + (-1)^{(|r|-1)(|t|-1)} [\![r, [\![s,t]\!]]\!] \,.$$

For more details about the Nijenhuis-Schouten bracket, see for example [6, 19]

12 Modified Double Poisson Brackets

The main reference for this section is work of S. Arthamonov [20, 32].

12.1 Poisson Brackets for General Associative Algebras

Let $\mathcal{A}$ be an associative algebra. Conventional definition of Poisson bracket becomes too restrictive when $\mathcal{A}$ is essentially non-commutative.

The standard "set" axioms (cf. Definition 2.1) meets with a problem of "non-trivial" example existence. This chapter is devoted to some constructions of "non-commutative" Poisson algebra structures. This subject which has started with the paper of Ping Xu [21], where the author introduces a notion of Poisson structure on noncommutative algebras, and studies some of its properties and applications. Given an associative algebra $\mathcal{A}$ demonstrated that the Hochschild cohomology $HH^*(\mathcal{A}, \mathcal{A})$ can be provided with a graded Lie algebra structure by means of the so-called $G-$bracket. This bracket, which was first introduced by M. Gerstenhaber, is the analogue of the Schouten bracket for multivector fields. A Poisson structure on $\mathcal{A}$ is then defined as an element of $HH^2(\mathcal{A}, \mathcal{A})$ whose G-bracket with itself vanishes. It was shown that such a Poisson structure induces an ordinary Poisson bracket on the center of $\mathcal{A}$.

We shall discuss the drawback of the naive definition of a Poisson structure on a non–commutative algebra $\mathcal{A}$.

First, we describe the following important lemma which appeared in the paper [22] and therefore (by the famous Arnold's statement) is attributed to Victor Ginzburg.

12.1.1 Ginzburg-Voronov Lemma

Lemma 2.7 *If $\mathcal{A}$ is any Poisson algebra, then for all $a, b, c, d \in \mathcal{A}$ the following identity holds*

$$[a, c]\{b, c\} = \{a, c\}[b, d]\ .$$

Proof Take the Poisson bracket { ab,cd}. By the derivation property of the bracket we have

$$\{ab, cd\} = a\{b, cd\} + \{a, cd\}b = ac\{b, d\} + a\{b, c\}d + c\{a, d\}b + \{a, c\}db\ .$$

On the other hand

$$\{ab, cd\} = c\{ab, d\} + \{ab, c\}d = ca\{b, d\} + c\{a, d\}b + a\{b, c\}d + \{a, c\}bd\ .$$

Subtracting the two equations yields the result. □

Definition 2.27 An algebra $\mathcal{A}$ is *prime*, if the product of nonzero ideals in nonzero.

Definition 2.28 An algebra $\mathcal{A}$ is *simple* if it has no non-trivial two-sided ideals and the algebra product is non-trivial.

For example, the algebra given by

$$\{\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \mid a \in \mathbb{R}\}$$

is not simple, as the matrix product is always trivial in this case.

The following theorem is due to D. R. Farkas and G. Letzter [23].

Theorem 2.8 *Let $\mathcal{A}$ be a prime and simple noncommutative Poisson algebra. Then for all $c, d \in \mathcal{A}$*

$$\{c, d\} = \lambda[c, d]$$

for some $\lambda \in \mathrm{Z}_P(\mathcal{A})$.[11]

Definition 2.29 ([24]) A map

$$\{,\}: \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}$$

is an H_0-Poisson bracket if for all $a, b, c \in \mathcal{A}$

1. $\{a, bc\} = b\{a, c\} + \{a, b\}c$ (Right Leibnitz identity),
2. $\{a, \{b, c\}\} - \{b, \{a, c\}\} = \{\{a, b\}, c\}$ (Left Loday-Jacobi identity),
3. $\{a, b\} + \{b, a\} \equiv 0 \bmod [\mathcal{A}, \mathcal{A}]$,
4. $\{ab, c\} = \{ba, c\}$.

Corollary 2.1 ([24]) *An H_0-Poisson bracket induces a Lie Algebra structure $\{\_\}^{Lie}: \mathcal{A}_\natural \otimes \mathcal{A}_\natural \to \mathcal{A}_\natural$ on* abelianization $\mathcal{A}_\natural := \mathcal{A}/[\mathcal{A}, \mathcal{A}]$ *of $\mathcal{A}$.*

12.1.2 Representation Scheme

Following philosophy by M. Kontsevich [25, 34], any algebraic property that makes geometric sense is mapped to its commutative counterpart by *Representation Functor*

$$\mathrm{Rep}_N: \quad \text{fin. gen. Associative algebras} \to \text{Affine schemes}\ ,$$
$$\mathrm{Rep}_N(\mathcal{A}) = Hom(\mathcal{A}, \mathrm{Mat}_N(\mathbb{C}))\ .$$

It assigns to a finitely generated associative algebra $\mathcal{A} = \langle x^{(1)}, \dots, x^{(k)} \rangle / \mathcal{R}$ a scheme of its' $N \times N$ matrix representations. Let

[11] See Definition 2.13 for the definition of Poisson center $\mathrm{Z}_P(\mathcal{A})$.

$$\varphi(x^{(i)}) = \begin{pmatrix} x_{11}^{(i)} & \dots & x_{1N}^{(i)} \\ \vdots & & \vdots \\ x_{N1}^{(i)} & \dots & x_{NN}^{(i)} \end{pmatrix} . \tag{52}$$

Representations of $\mathcal{A}$ then form an affine scheme $\mathcal{V}$ with a coordinate ring $\mathbb{C}[\mathcal{V}] := \mathbb{C}\left[x_{j,k}^{(i)}\right]/\varphi(\mathcal{R})$. Denote as $\mathbb{C}_{\mathcal{V}}$ - the corresponding sheaf of rational functions.

12.1.3 Moduli Space of Representations

Change of basis corresponds to the action $\mathrm{GL}_N(\mathbb{C}) \circlearrowleft \mathrm{Mat}_N(\mathbb{C})$,

$$M \to gMg^{-1} .$$

It induces $\mathrm{GL}_N(\mathbb{C}) \circlearrowleft \mathbb{C}[\mathcal{V}]$. The invariant subalgebra $\mathbb{C}[\mathcal{V}]^{\mathrm{GL}_N(\mathbb{C})} \subset \mathbb{C}[\mathcal{V}]$ is then a coordinate ring of the corresponding moduli space of representations.

$$\varphi_0 : \mathcal{A}_\natural \to \mathbb{C}[\mathcal{V}]^{\mathrm{GL}_N(\mathbb{C})}, \quad \varphi_0(x) = \mathrm{Tr}\,\varphi(x) .$$

Lemma 2.8 (Procesi, 1976) *Subset $\varphi_0(\mathcal{A}_\natural)$ generates $\mathbb{C}[\mathcal{V}]^{\mathrm{GL}_N(\mathbb{C})}$.*

Proposition 2.18 (Crawley-Boevey [24]) *An H_0-Poisson bracket induces a conventional Poisson bracket*

$$\{,\}^{inv} : \mathbb{C}[\mathcal{V}]^{\mathrm{GL}_N(\mathbb{C})} \otimes \mathbb{C}[\mathcal{V}]^{\mathrm{GL}_N(\mathbb{C})} \to \mathbb{C}[\mathcal{V}]^{\mathrm{GL}_N(\mathbb{C})} .$$

12.2 Double Poisson Brackets

Definition 2.30 (M. Van Den Bergh [26]) A map $\{\!\{,\}\!\} : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A} \otimes \mathcal{A}$ is a double Poisson bracket if for all $a, b, c \in \mathcal{A}$:

1. $\{\!\{a, b\}\!\} = -\{\!\{b, a\}\!\}^{op}$,
2. $\{\!\{ab, c\}\!\} = (1 \otimes a)\{\!\{b, c\}\!\} + \{\!\{a, c\}\!\}(b \otimes 1)$,
3. $\{\!\{a, bc\}\!\} = (b \otimes 1)\{\!\{a, c\}\!\} + \{\!\{a, b\}\!\}(1 \otimes c)$,
4. $R_{12}R_{23} + R_{31}R_{12} + R_{23}R_{31} = 0$, where $R_{m,n}(a_1 \otimes \dots \otimes a_k) = a_1 \otimes \dots \otimes a_{m-1} \otimes \{\!\{a_m, a_n\}\!\}' \otimes \dots \otimes \{\!\{a_m, a_n\}\!\}'' \otimes \dots \otimes a_k$.

Proposition 2.19 ([26]) *Double Poisson bracket induces a conventional Poisson bracket*

$$\{,\}^{\mathcal{V}} : \mathbb{C}_{\mathcal{V}} \otimes \mathbb{C}_{\mathcal{V}} \to \mathbb{C}_{\mathcal{V}} \qquad \left\{x_{ij}^{(m)}, x_{kl}^{(n)}\right\}^{\mathcal{V}} = \varphi\left(\{\!\{x^{(m)} \otimes x^{(n)}\}\!\}\right)_{(kj),(il)} .$$

If $\mathcal{A} = \mathbb{C} < x_1, \dots, x_m >$ is the free associative algebra, then $\mathbb{C}[\mathrm{Rep}_n(\mathcal{A})] = \mathbb{C}[x^j_{i,\alpha}]$ where $1 \le \alpha \le m$.

If $\{x_\alpha, x_\beta\}$ is a double Poisson bracket on $\mathcal{A} = \mathbb{C} < x_1, \dots, x_m >$, then, using the Sweedler convention and drop the sign of sum, we obtain the conventional Poisson brackets on $\mathbb{C}[\mathrm{Rep}_n(\mathcal{A})]$:

$$\{x^j_{i,\alpha}, x^l_{k,\beta}\} = \{x_\alpha, x_\beta\}^{'j}_k \{x_\alpha, x_\beta\}^{''l}_i$$

12.3 Quadratic Double Poisson Brackets

Let $\mathcal{A} = \mathbb{C} < x_1, \dots, x_m >$ be the free associative algebra. If double brackets $\{x_i, x_j\}$ between all generators are fixed, then the bracket between two arbitrary elements of $\mathcal{A}$ is uniquely defined by identities (2.30). It follows from (2.30) that constant, linear, and quadratic double brackets are defined by

$$\{x_i, x_j\} = c_{ij} 1 \otimes 1, \qquad c_{i,j} = -c_{j,i}, \tag{53}$$

$$\{x_i, x_j\} = b^k_{ij} x_k \otimes 1 - b^k_{ji} 1 \otimes x_k, \tag{54}$$

and

$$\{x_\alpha, x_\beta\} = r^{uv}_{\alpha\beta}\, x_u \otimes x_v + a^{vu}_{\alpha\beta}\, x_u x_v \otimes 1 - a^{uv}_{\beta\alpha}\, 1 \otimes x_v x_u, \tag{55}$$

where

$$r^{\sigma\epsilon}_{\alpha\beta} = -r^{\epsilon\sigma}_{\beta\alpha}, \tag{56}$$

correspondingly. The summation with respect to repeated indexes is assumed.

It is easy to verify that the bracket (53) satisfies (2.30) for any skew-symmetric tensor c_{ij}. For the bracket (54) the condition (2.30) is equivalent to the identity

$$b^\mu_{\alpha\beta} b^\sigma_{\mu\gamma} = b^\sigma_{\alpha\mu} b^\mu_{\beta\gamma}, \tag{57}$$

which means that $b^\sigma_{\alpha\beta}$ are structure constants of an associative algebra $\mathcal{A}$.

Proposition 2.20 *The bracket (55) satisfies (2.30) iff the following relations hold:*

$$r^{\lambda\sigma}_{\alpha\beta} r^{\mu\nu}_{\sigma\tau} + r^{\mu\sigma}_{\beta\tau} r^{\nu\lambda}_{\sigma\alpha} + r^{\nu\sigma}_{\tau\alpha} r^{\lambda\mu}_{\sigma\beta} = 0, \tag{58}$$

$$a^{\sigma\lambda}_{\alpha\beta} a^{\mu\nu}_{\tau\sigma} = a^{\mu\sigma}_{\tau\alpha} a^{\nu\lambda}_{\sigma\beta}, \tag{59}$$

$$a^{\sigma\lambda}_{\alpha\beta} a^{\mu\nu}_{\sigma\tau} = a^{\mu\sigma}_{\alpha\beta} r^{\lambda\nu}_{\tau\sigma} + a^{\mu\nu}_{\alpha\sigma} r^{\sigma\lambda}_{\beta\tau} \tag{60}$$

and

$$a^{\lambda\sigma}_{\alpha\beta} a^{\mu\nu}_{\tau\sigma} = a^{\sigma\nu}_{\alpha\beta} r^{\lambda\mu}_{\sigma\tau} + a^{\mu\nu}_{\sigma\beta} r^{\sigma\lambda}_{\tau\alpha}. \tag{61}$$

The conventional Poisson bracket corresponding to any double Poisson bracket (55) can be defined on $\mathbb{C}[\mathrm{Rep}_n(\mathcal{A})]$ by the following way [27]:

$$\{x^j_{i,\alpha}, x^{j'}_{i',\beta}\} = r^{\gamma\epsilon}_{\alpha\beta} x^{j'}_{i,\gamma} x^j_{i',\epsilon} + a^{\gamma\epsilon}_{\alpha\beta} x^k_{i,\gamma} x^{j'}_{k,\epsilon} \delta^j_{i'} - a^{\gamma\epsilon}_{\beta\alpha} x^k_{i',\gamma} x^j_{k,\epsilon} \delta^{j'}_i \tag{62}$$

where $x^j_{i,\alpha}$ are entries of the matrix x_α and δ^j_i is the Kronecker delta-symbol. Relations (56), (58)–(61) hold iff (62) is a Poisson bracket.

We may interpret the four index tensors r and a as:

1. operators on $V \otimes V$, where V is an m-dimensional vector space;
2. elements of $\mathrm{Mat}_m(\mathbb{C}) \otimes \mathrm{Mat}_m(\mathbb{C})$;
3. operators on $\mathrm{Mat}_m(\mathbb{C})$.

For the first interpretation let V be a linear space with a basis $e_\alpha,\ \alpha = 1, \ldots, m$. Define linear operators $r,\ a$ on the space $V \otimes V$ by

$$r(e_\alpha \otimes e_\beta) = r^{\sigma\epsilon}_{\alpha\beta} e_\sigma \otimes e_\epsilon, \qquad a(e_\alpha \otimes e_\beta) = a^{\sigma\epsilon}_{\alpha\beta} e_\sigma \otimes e_\epsilon.$$

Then the identities (56), (58)–(61) can be written as

$$\begin{gathered} r^{12} = -r^{21}, \quad r^{23}r^{12} + r^{31}r^{23} + r^{12}r^{31} = 0, \\ a^{12}a^{31} = a^{31}a^{12}, \\ \sigma^{23}a^{13}a^{12} = a^{12}r^{23} - r^{23}a^{12}, \\ a^{32}a^{12} = r^{13}a^{12} - a^{32}r^{13}. \end{gathered} \tag{63}$$

Here all operators act in $V \otimes V \otimes V$, σ^{ij} means the transposition of i-th and j-th components of the tensor product, and $a^{ij},\ r^{ij}$ mean operators $a,\ r$ acting in the product of the i-th and j-th components.

Note that first two relations mean that the tensor r should be skew-symmetric solution of the classical associative Yang-Baxter equation [28].

In the second interpretation we consider the following elements from $\mathrm{Mat}_m(\mathbb{C}) \otimes \mathrm{Mat}_m(\mathbb{C})$: $r = r^{km}_{ij} e^i_k \otimes e^j_m, \quad a = a^{km}_{ij} e^i_k \otimes e^j_m$, where e^i_j are the matrix unities: $e^j_i e^m_k = \delta^j_k e^m_i$. Then (56), (58)–(61) are equivalent to (63), where tensors belong to $\mathrm{Mat}_m(\mathbb{C}) \otimes \mathrm{Mat}_m(\mathbb{C}) \otimes \mathrm{Mat}_m(\mathbb{C})$. Namely, $r^{12} = r^{mk}_{ij} e^i_k \otimes e^j_m \otimes 1$ and so on. The element σ is given by $\sigma = e^j_i \otimes e^i_j$.

For the third interpretation, we shall define operators $r, a, \bar{r}, a^*$: $\mathrm{Mat}_N \to \mathrm{Mat}_N$ by $r(x)^p_q = r^{mp}_{nq} x^n_m$, $a(x)^p_q = a^{mp}_{nq} x^n_m$, $\bar{r}(x)^p_q = r^{pm}_{nq} x^n_m$, $a^*(x)^p_q = a^{pm}_{qn} x^n_m$.
Then (56), (58)–(61) provide the following operator identities:

$$r(x) = -r^*(x), \qquad r(x)r(y) = r(xr(y)) + r(x)y),$$

$$\bar{r}(x) = -\bar{r}^*(x), \qquad \bar{r}(x)\bar{r}(y) = \bar{r}(x\bar{r}(y)) + \bar{r}(x)y),$$

$$a(x)a^*(y) = a^*(y)a(x),$$

$$a^*(ya(x)) = r(xa^*(y)) - r(x)a^*(y),$$

$$a(x)a(y) = -a(r(y)x) - a(yr(x)),$$

$$a^*(a(x)y) = r(a^*(y)x) - a^*(y)r(x),$$

$$a(ya^*(x)) = -\bar{r}(xa(y)) + \bar{r}(x)a(y),$$

$$a^*(x)a^*(y) = a^*(\bar{r}(y)x) + a^*(y\bar{r}(x)),$$

$$a(a^*(x)y) = -\bar{r}(a(y)x) + a(y)\bar{r}(x)$$

for any x, y. First two of these identities mean that operators r and $\bar{r}$ satisfies the Rota-Baxter equation [29] and this fact implies also that the new matrix multiplications $\circ_r$ and $\circ_{\bar{r}}$ defined by

$$x \circ_r y = r(x)y + xr(y), \quad x \circ_{\bar{r}} y = \bar{r}(x)y + x\bar{r}(y)$$

are associative.

12.4 Examples and Classification of Low Dimensional Quadratic Double Poisson Brackets

It is easy to see that for $m = 1$ non-zero quadratic double Poisson brackets does not exist. In the simplest non-trivial case $m = 2$ the system of algebraic equations (56), (58)–(61) can be straightforwardly solved .

Theorem 2.9 *Let $m = 2$. Then the following Cases **1–7** form a complete list of quadratic double Poisson brackets up to equivalence given a linear change of the generators. We present non-zero components of the tensors r and a only.*

Case 1. $r^{21}_{22} = -r^{12}_{22} = 1$. *The corresponding (non-zero) double brackets read*

$$\{v, v\} = v \otimes u - u \otimes v;$$

Case 2. $r^{21}_{22} = -r^{12}_{22} = 1$, $a^{11}_{21} = a^{12}_{22} = 1$. *The corresponding (non-zero) double brackets:*

$$\{v, v\} = v \otimes u - u \otimes v + vu \otimes 1 - 1 \otimes vu,\ \{v, u\} = u^2 \otimes 1,\ \{u, v\} = -1 \otimes u^2;$$

Case 3. $r^{21}_{22} = -r^{12}_{22} = 1$, $a^{11}_{12} = a^{21}_{22} = 1$. *The corresponding (non-zero) double brackets:*

$$\{v, v\} = v \otimes u - u \otimes v + uv \otimes 1 - 1 \otimes uv,\ dbu, v = u^2 \otimes 1,\ \{v, u\} = -1 \otimes u^2;$$

Case 4. $r^{22}_{21} = -r^{22}_{12} = 1$. *The corresponding (non-zero) double brackets:*

$$\{v, u\} = v \otimes v,\ \{u, v\} = -v \otimes v;$$

Case 5. $r^{22}_{21} = -r^{22}_{12} = 1,;$ $a^{21}_{11} = a^{22}_{12} = 1$. *The corresponding (non-zero) double brackets:*

$$\{v, u\} = v \otimes v - 1 \otimes v^2,\ \{u, v\} = -v \otimes v + v^2 \otimes 1,\ \{u, u\} = uv \otimes 1 - 1 \otimes uv;$$

Case 6. $r^{22}_{21} = -r^{22}_{12} = 1,;$ $a^{12}_{11} = a^{22}_{21} = -1$. *The corresponding (non-zero) double brackets:*

$$\{v, u\} = v \otimes v - v^2 \otimes 1,\ \{u, v\} = -v \otimes v + 1 \otimes v^2,\ \{u, u\} = -vu \otimes 1 + 1 \otimes vu;$$

Case 7. $a^{11}_{22} = 1$. *The corresponding (non-zero) double brackets:*

$$\{v, v\} = u^2 \otimes 1 - 1 \otimes u^2.$$

For a proof of (2.9) see [30].

Remark 2.29 Cases **2** and **3** as well as Cases **5** and **6** are linked via the involution.

Remark 2.30 Case **1** is equivalent to the double bracket from Example 1 with $m = 2$.

Remark 2.31 It is easy to verify (see [28]) that there exist only two non-isomorphic anti-Frobenius subalgebras in $\mathrm{Mat}_2(\mathbb{C})$. They are matrices with one zero column and matrices with one zero row. Cases **1** and **4** correspond to them.

Remark 2.32 Notice that the trace Poisson brackets for Cases **2** and **4** are non-degenerate. Corresponding symplectic forms can be found in [31] (Example 5.7 and Lemma 7.1).

Remark 2.33 The corresponding Lie algebra structures on the trace space $\mathcal{A}/[\mathcal{A}, \mathcal{A}]$ are trivial (abelian) in all cases, except the cases **2**, **3** and **4** :

$$[\bar{u}, \bar{v}] = -\bar{u}^2 \quad (\text{Case } 2), \quad [\bar{u}, \bar{v}] = \bar{u}^2 \quad (\text{Case } 3), \quad [\bar{u}, \bar{v}] = -\bar{v}^2 \quad (\text{Case } 4).$$

These cases give the isomorphic Lie algebra structures on $\mathcal{A}/[\mathcal{A},\mathcal{A}]$ with respect to the involutions $u \to v, \quad v \to u$ and $u \to u \quad v \to -v$.
We refer to literature [33, 35, 36] for further details relevant to this section.

Example Consider the trace Poisson bracket (62) corresponding to case **6**. Its Casimir functions are given by

$$\operatorname{tr} v^k, \qquad \operatorname{tr} uv^k, \qquad k = 0, 1, \ldots$$

where $u = x_1, v = x_2$. Functions $\operatorname{tr} u^i$ and $\operatorname{tr} vu^i$, where $i = 2, 3, \ldots$ commute each other with respect to this bracket.

Acknowledgments V. Rubtsov is thankful to the Institut des Hautes Études Scientifiques Université Paris-Saclay for hospitality during the last stages of this work. R. Suchánek is grateful to the University of Angers and to the Masaryk University for their hospitality. The work of R. Suchánek was financially supported under the project GAČR EXPRO GX19-28628X and by the Barrande Fellowship program organized by The French Institute in Prague (IFP) and the Czech Ministry of Education, Youth and Sports (MYES). Authors are also grateful to the organizers of the Winter School and Workshop Wisla 20–21 for a wonderful online event, organized during a difficult time of a covid pandemic.

We are also grateful to the anonymous referees for the careful reading of the manuscript, which permitted us to correct many typos and misprints in the text.

References

1. Kosmann-Schwarzbach, Y.: Poisson algebra. Encyclopedia of Mathematics. http://encyclopediaofmath.org/index.php?title=Poisson_algebra&oldid=51639
2. Camille, L., Anne, P., Pol, V.: Poisson Structures. Springer, Berlin Heidelberg (2013)
3. Vinogradov, A., Krasil'shchik, I.: What is the Hamiltonian formalism. Russian Math. Surv. **30**, 177–202 (1975)
4. Cartiere, P.: Some fundamental techniques in the theory of integrable systems. IHES-M-94-23, Sub. General Theoretical Physics, pp. 42 (1994)
5. Bhaskara, K., Viswanath, K.: Poisson Algebras and Poisson Manifolds. Pitman Research Notes In Mathematics Series, vol. 174 (1988)
6. Krasil'shchik, I.: Schouten bracket and canonical algebras. In: Global Analysis—Studies and Applications III, pp. 79–110 (1988), https://doi.org/10.1007/BFb0080424
7. Brylinski, J.: A differential complex for Poisson manifolds. J. Differential Geom. **28**, 93–114 (1988). https://doi.org/10.4310/jdg/1214442161
8. Koszul, J.: Crochet de Schouten-Nijenhuis et cohomologie. Astérisque **137**, 4–3 (1985)
9. Odesskii, A.V., Rubtsov, V.N.: Polynomial Poisson algebras with regular structure of symplectic leaves. Theoret. Math. Phys. **133**, 1321–1337 (2002)
10. Sklyanin, E.K.: Some algebraic structures connected with the Yang—Baxter equation. Funct. Anal. Appl. **16**, 263–270 (1982)

11. Smith, P.: The 4-dimensional Sklyanin algebras. In: Proc. of Conf. on Algebraic Geometry and Ring Theory in Honor of Michael Artin (Antwerp, 1992). K-Theory, pp. 65–80 (1994)
12. Berger, R., Pichereau, A.: Calabi–Yau algebras viewed as deformations of Poisson algebras. Algebras Represent. Theory **17**, 735–773 (2014)
13. Lichnerowicz, A.: Les variétés de Poisson et leurs algèbres de Lie associées. J. Differential Geom. **12**, 253–300 (1977). https://doi.org/10.4310/jdg/1214433987
14. Kostant, B.: Quantization and Unitary Representations. Lectures in Modern Analysis and Applications III, pp. 87–208 (1970)
15. Kirillov, A.: Method of orbits in the theory of unitary representations of Lie groups. Funct. Anal. Appl. **2**, 90–93 (1968)
16. Kirillov, A.: Lectures on the Orbit Method. Graduate Studies in Mathematics, vol. 64 (2004)
17. Souriau, J.: Quantification géométrique. Applications. Ann. L'I.H.P. Phys. Théor. **6**, 311–341 (1967)
18. Román-Roy, N.: Some properties of multisymplectic manifolds. In: Classical and Quantum Physics, pp. 325–336 (2019)
19. Ibáñez, R., León, M., Marrero, J.: Homology and cohomology on generalized Poisson manifolds. J. Phys. A Math. Gen. **31**, 1253–1266 (1998). https://doi.org/10.1088/0305-4470/31/4/014
20. Arthamonov, S.: Modified double Poisson brackets. J. Algebra **492**, 212–233 (2017). https://www.sciencedirect.com/science/article/pii/S0021869317304805
21. Xu, P.: Noncommutative Poisson algebras. Am. J. Math. **116**, 101–125 (1994). http://www.jstor.org/stable/2374983
22. Voronov, T.: On the Poisson envelope of a Lie algebra. "Noncommutative" moment space. Funct. Anal. Appl. **29**, 196–199 (1995)
23. Farkas, D., Letzter, G.: Ring theory from symplectic geometry. J. Pure Appl. Algebra **125**, 155–190 (1998)
24. Crawley-Boevey, W.: Poisson structures on moduli spaces of representations. J. Algebra **325**, 205–215 (2011). https://www.sciencedirect.com/science/article/pii/S0021869310004710
25. Kontsevich, M.: Formal (non)-commutative symplectic geometry. In: The Gelfand Mathematical Seminars, 1990–1992, pp. 173–187 (1993)
26. Van Den Bergh, M.: Double Poisson algebras. Trans. Am. Math. Soc. **360**, 5711–5769 (2008)
27. Odesskii, A., Rubtsov, V., Sokolov, V.: Bi-Hamiltonian ordinary differential equations with matrix variables. Theoret. Math. Phys. **171**, 442–447 (2012)
28. Aguiar, M.: On the associative analog of Lie bialgebras. J. Algebra. **244**, 492–532 (2001)
29. Rota, G.: Baxter operators, an introduction. In: Kung, J.P.S. (ed.) Gian-Carlo Rota on Combinatorics, Introductory Papers and Commentaries, vol. 1, p. 90. Birkhäuser, Boston (1995)
30. Odesskii, A., Rubtsov, V., Sokolov, V.: Double Poisson brackets on free associative algebras. Noncommutative Birational Geom. Representations Combin. **592**, 225–239 (2013)
31. Bielawski, R.: Quivers and Poisson structures. Manuscripta Math.. **141** (2013)
32. Arthamonov, S.: Noncommutative inverse scattering method for the Kontsevich system. Lett. Math. Phys. **105**, 1223–1251 (2015). https://doi.org/10.1007/s11005-015-0779-5
33. Wolf, T., Efimovskaya, O.: On integrability of the Kontsevich non-Abelian ODE system. Lett. Math. Phys. **100**, 161–170 (2017). https://doi.org/10.1007/s11005-011-0527-4
34. Kontsevich, M.: Noncommutative identities. ArXiv (2011). https://arxiv.org/abs/1109.2469
35. Crawley-Boevey, W., Etingof, P., Ginzburg, V.: Noncommutative geometry and quiver algebras. Adv. Math. **209**, 274–336 (2007). https://www.sciencedirect.com/science/article/pii/S0001870806001587
36. Crawley-Boevey, W.: Preprojective algebras, differential operators and a Conze embedding for deformations of Kleinian singularities. Comment. Math. Helv. **74**, 548–574 (1999)

Some Remarks on Multisymplectic and Variational Nature of Monge-Ampère Equations in Dimension Four

Radek Suchánek

Abstract We describe a necessary condition for the local solvability of the strong inverse variational problem in the context of Monge-Ampère partial differential equations and first-order Lagrangians. This condition is based on comparing effective differential forms on the first jet bundle. To illustrate and apply our approach, we study the linear Klein-Gordon equation, first and second heavenly equations of Plebański, Grant equation, and Husain equation, over a real four-dimensional manifold. Two approaches towards multisymplectic formulation of these equations are described.

1 Introduction

Since the nineteenth and early twentieth century work of mathematicians such as Joseph Liouville, Gaston Darboux, Sophus Lie, Élie Cartan et al., it is well-known that geometry plays an essential role in the study of ordinary and partial differential equations (PDEs).

A special subclass of all non-linear second-order PDEs is Monge-Ampère (M-A) equations. They arise in many examples and have numerous applications throughout mathematics and mathematical physics. One can find them in differential geometry of surfaces, hydrodynamics, acoustics, integrability of various geometric structures, variational calculus, Riemannian, CR, and complex geometry, quantum gravity, and even in theoretical meteorology (semi-geostrophic and quasi-geostrophic theory). Many other instances can be listed. For a detailed exposition of interesting applications of M-A equations, particularly in 2D and 3D, see [1].

R. Suchánek (✉)
Department of Mathematics and Statistics, Masaryk University, Brno, Czech Republic

Angevin Laboratory of Mathematical Research - UMR CNRS 6093, University of Angers, Angers Cedex, France

M. Ulan, S. Hronek (eds.), *Groups, Invariants, Integrals, and Mathematical Physics*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-031-25666-0_3

In this paper, we are mainly interested in the variational structure of M-A equations. In particular, we study whether we can view them as E-L equations for some first-order Lagrangians. Our approach is based on the idea of V. Lychagin to connect the M-A operators with symplectic (on T^*M) and contact (on J^1M) geometries. He also defined a class of variational problems related to M-A equations [1, 2].

Afterwards, we observe the relation between M-A equations and multisymplectic geometry, using the results of two slightly different approaches proposed by F. Hélein [3], and D. Harrivel [4]. We have found some new aspects which could shed light on this connection. We applied our observations in the context of the following 4D PDEs, very famous for their applications in geometry and theoretical physics related to Einstein gravity and relativistic field theories—Plebański heavenly equations and Klein-Gordon equation. We also considered Grant and Husain equations, which are very close to Plebański second equation.

In 1975, J.F. Plebański introduced his first and second heavenly equations [5], which belong to the class of M-A equations in 4D. Their close relatives, Grant and Husain equations, were introduced more recently [6, 7]. These equations appeared firstly in Einstein gravity, and later were studied by numerous authors, both physicists and mathematicians [5–9]. Another significant example of M-A equation is the Klein-Gordon equation, which is a non-homogeneous relativistic wave equation. It was derived in the first quarter of the twentieth century by O. Klein and later reformulated in a more compact form by W. Gordon [10]. The underlying structure of this equation can be found in more general situations than scalar fields, and the knowledge of its solutions is relevant in the relativistic perturbative quantum field theory [11]. The specific form of all the above equations and some further details about them is given below.

In the first section, we define M-A operators and related notions, which will be our main tools in working with M-A equations via differential forms. We also recall the contact and symplectic calculus over J^1M, which we greatly utilize in our computations. The second section describes the construction of the Euler operator on $\Omega^n(J^1M)$ and its relation to variational problems. In the third section, a necessary condition for local solvability of the strong inverse variational problem of a given M-A equation is formulated, together with the corresponding analysis of the aforementioned five M-A equations in four real dimensions. In the fourth section, we present two multisymplectic approaches and provide certain comparison of them, in the context of concrete M-A equations under consideration.

In the sequel, we will be working with smooth real-valued functions $\phi \in C^\infty(M)$ and their first prolongations $j^1\phi\colon M \to J^1M$, where $J^1M \to J^0M = M \times \mathbb{R}$ is the first jet bundle of $\mathrm{pr}_1\colon M \times \mathbb{R} \to M$.

A second-order partial differential equations which are given as a $C^\infty(J^1M)$-linear[1] combination of minors of the Hessian matrix $(\phi_{\mu\nu})_{\mu,\nu}$ are called Monge-

[1] By $C^\infty(J^1M)$-linear we mean that the coefficients can be smooth functions and their first derivatives.

Ampère equations[2] [1, 2, 4, 9]. Consequently, every such equation can be represented by a differential n-form on J^1M via M-A operator $\Delta_\omega\phi := (j^1\phi)^*\omega$. Moreover, one can use *effective* differential forms, which represent M-A equations uniquely (up to a multiple of a non-vanishing function), and without terms corresponding to trivial equations [1, 2, 9]. Effective forms on the first jet space, which produce first-order Lagrangians on the base manifold, have a particularly simple local expression. Their image under the Euler operator represents the Euler-Lagrange (E-L) equations [1]. This feature of the Euler operator, together with the fact that it preserves the effective forms, enables us to study the existence of a first-order Lagrangian for a given M-A equation on the level of differential forms over J^1M. Additionally, some effective forms give rise (in a non-unique way) to multisymplectic forms [4]. This may happen even for an effective form that comes from a M-A equation which does not have a first-order Lagrangian. Since the multisymplectic reformulation usually starts with a Lagrangian [3, 12, 13], this seems to be an interesting property. We will apply the formalism on the following M-A equations: Plebański heavenly, Grant, Husain, and Klein-Gordon equations. We will consider these equations in the real 4D case.

The heavenly equations of Plebański were first derived in [5] in the form

$$\begin{aligned} \phi_{13}\phi_{24} - \phi_{14}\phi_{23} &= 1 \text{ (1st heavenly equation)} \\ \phi_{11}\phi_{22} - (\phi_{12})^2 + \phi_{13} + \phi_{24} &= 0 \text{ (2nd heavenly equation)} \end{aligned}$$

using self-dual 2-forms over a complex 4D Riemannian space. The duality here is given by the Hodge star operator. The Grant equation and the Husain equation are both based on the Ashtekar-Jacobson-Smolin (AJS) equations, which are Einstein self-dual equations. The AJS equations were derived in [14] employing the $3 + 1$ ADS decomposition of spacetime. They characterize 4D complex metrics with self-dual curvature 2-form. Metrics with self-dual curvature form satisfy the vacuum equations of general relativity since they are Ricci flat. In [6], the following equation was introduced

$$\phi_{11} + \phi_{24}\phi_{13} - \phi_{23}\phi_{14} = 0 \text{ (Grant equation)}$$

and subsequently rewritten into a system which enabled the author to construct formal solutions. Notably, *the Grant equation is equivalent with the first heavenly equation of Plebański* [6]. Another reformulation of the AJS equations was provided in [7], in order to identify AJS with a 2D chiral model, and to provide a Hamiltonian formulation. The resulting equation

$$\phi_{13}\phi_{24} - \phi_{14}\phi_{23} + \phi_{11} + \phi_{22} = 0 \text{ (Husain equation)}$$

[2] Note that the minors of rank 1 recover all the second-order semi-linear differential equations, whilst the higher order minors (including the determinant of the whole matrix) add specific non-linear terms.

enabled V. Husain to show the existence of infinitely many non-local conserved currents. Another type of an M-A equation is

$$\phi_{11} - \phi_{22} - \phi_{33} - \phi_{44} + m^2\phi^2 = 0 \text{ (Klein-Gordon equation)}$$

where m is a constant. The Klein-Gordon equations was derived in various ways, for example by W. Gordon [10]. In its real version, it can be interpreted as an equation of motion for a scalar field without charge over a Lorentzian manifold. A key difference between the aforementioned equations is that the Klein-Gordon equation does not arise from self-duality conditions.

2 Preliminary Notions

In this section we fix the notation and introduce basic definitions and statements relevant to our considerations. In particular, we will define the notion of effective forms, Monge-Ampère operators and Monge-Ampère equations. All our considerations are local. We caution the reader about the standard abuse of notation such us denoting a symplectic form by Ω, and by $\Omega(M)$ the exterior algebra of differential forms over M.

We denote by M a smooth n-dimensional manifold, $(q^1, \ldots, q^n)$ are local coordinates over an open subset $U \subset M$, TM and T^*M are the tangent and cotangent bundle, respectively. Let J^1M be the space of 1-jets of smooth functions over M, which is an affine bundle over $M \times \mathbb{R}$

$$\pi : J^1M \to J^0M = M \times \mathbb{R}$$

with typical fiber T^*M. It is also a fiber bundle over M

$$\mathrm{pr}_1 \circ \pi : J^1M \to M ,$$

where $\mathrm{pr}_1 : M \times \mathbb{R} \to M$. We denote by $(q^1, \ldots, q^n, u, p_1, \ldots, p_n)$ the induced local coordinates on J^1M. The first prolongation of $\phi \in C^\infty(M)$[3] is a section $j^1\phi : M \to J^1M$, given by $x \mapsto (j^1\phi)(x) \in J^1M$. Recall that $(j^1\phi)(x)$ is an equivalence class of functions which are equal up to the first order in derivatives at x. In local coordinates,

$$j^1\phi = (q^\mu, \phi, \phi_\mu) ,$$

where $\phi_\mu := \partial_{q^\mu}\phi := \frac{\partial\phi}{\partial_{q^\mu}}$ is the partial derivative in the direction of the coordinate q^μ. The pullbacks of coordinate functions on J^1M are

[3] Each $\phi \in C^\infty(M)$ defines a section $M \to M \times \mathbb{R}$, $x \mapsto (x, \phi(x))$.

$$(j^1\phi)^*q^\mu = q^\mu \qquad (j^1\phi)^*u = \phi \qquad (j^1\phi)^*p_\mu = \phi_\mu \ ,$$

In the local coordinates, we have the identification $J^1U \cong T^*U \times \mathbb{R}$ (which is not canonical). Most relevant for us is that J^1M is naturally equipped with a contact structure [1, 2]. For more details about jet bundles and structures on them, see [15].

2.1 Contact Structure on J^1M

Definition 3.1 Let $\omega \in \Omega^1(M)$ be non-vanishing. Let $\mathcal{D} \subset TM$ be a distribution given by $\mathcal{D} := \ker\omega$. Then ω is called a contact form on M, if $d\omega|_{\mathcal{D}} \colon \mathcal{D} \to \mathcal{D}^*$ is non-degenerate. Manifold with a distribution described by a contact form is called a contact manifold and d is called a contact structure (or contact distribution) on M.

Remark 3.1 Note that the distribution $\mathcal{D} = \ker\omega$ satisfies $\operatorname{codim}\mathcal{D} = 1$. Moreover, the 1-form describing $\mathcal{D}$ is not unique. Consider a class of 1-forms, $[\omega]$, given by $\tilde{\omega} \in [\omega]$ if and only if there is a non-vanishing $f \in C^\infty(M)$ s.t. $\tilde{\omega} = f\omega$. Then every representative of the class $[\omega]$ defines the same distribution d.

The first jet space comes equipped with the *Cartan distribution*, which infinitesimally describes the condition that a section of $J^1M \to M$ is obtained as a prolongation of a function $\phi \in C^\infty(M)$. In the induced coordinates, this requirement can be described by the following *contact form*[4]

$$\mathfrak{c} = \mathrm{d}u - p_\mu \mathrm{d}q^\mu \ . \tag{1}$$

This 1-form satisfies the Definition 3.1 and we can describe the Cartan distribution as $C = \ker\mathfrak{c}$. That is, J^1M is a contact manifold.[5] By the Darboux theorem, every contact form on J^1M is locally given by (1). The contact form defines the Reeb vector field, χ, by the following conditions

$$\chi \lrcorner \mathrm{d}\mathfrak{c} = 0 \text{ and } \mathfrak{c}(\chi) = 1 \ . \tag{2}$$

In the local coordinates s.t. (1) holds, the Reeb field is of the form $\chi = \partial_u$, which immediately follows from (2). Moreover, since $\operatorname{codim} C = 1$, we get the following splitting of TJ^1U

$$TJ^1U \cong C \oplus \operatorname{span}(\chi) \cong \ker\mathfrak{c} \oplus \ker\mathrm{d}\mathfrak{c} \ .$$

[4] We are using the summation convention of summing over the repeated indices.

[5] Cartan distribution exists also on higher jets but the first jets are special due to $\operatorname{codim} C = 1$.

2.2 Symplectic Calculus on the Cartan Distribution

Contact form on J^1M gives rise to a *symplectic form* on C.

Definition 3.2 Let V be a vector space, $\dim V = 2n$. A symplectic form on V is a 2-form $\Omega \in \Lambda^2(V^*)$, which is non-degenerate, i.e. $\Omega^n := \Omega \wedge \ldots \wedge \Omega$ is non-vanishing.

Consider the 2-form $\Omega := \mathrm{d}\mathfrak{c}$ on the contact manifold J^1M. Then Ω is obviously closed. In the chosen coordinates, we have

$$\Omega = \mathrm{d}q^\mu \wedge \mathrm{d}p_\mu \,. \tag{3}$$

Note that Ω is non-degenerate when restricted to C. This means that Ω_x is a symplectic form on C_x at every $x \in M$. Using the symplectic form, we can define various useful operators. This leads to considering the space of differential k-forms which are degenerate along the Reeb field χ. We will denote this C^∞-module by

$$\Omega^k(C) := \{\alpha \in \Omega^k(J^1U) \mid \chi \lrcorner \alpha = 0\} \,. \tag{4}$$

Since the interior product $\lrcorner$ satisfies the graded Leibniz rule with respect to the wedge product, the space

$$\Omega(C) := \bigoplus_{k\leq 0} \Omega^k(C) \subset \Omega(J^1M)$$

has a graded algebra structure. Using suitable projections, $\Omega(C)$ can be turned into a differential graded algebra.

Projection and Projected Derivative Every $\alpha \in \Omega^k(J^1M)$ can be projected on $\Omega^k(C)$ via the projection $p\colon \Omega^k(J^1M) \to \Omega^k(C)$, acting on arbitrary k-form α as

$$p(\alpha) = \alpha - \mathfrak{c} \wedge (\chi \lrcorner \alpha) \,. \tag{5}$$

Let us show that p has the claimed properties. Firstly, $p^2 = p$, since

$$p(p(\alpha)) = \alpha - \mathfrak{c} \wedge (\chi \lrcorner \alpha) - \mathfrak{c} \wedge \big(\chi \lrcorner (\alpha - \mathfrak{c} \wedge (\chi \lrcorner \alpha)\big) = p(\alpha) \,.$$

Secondly, $p(\alpha) \in \Omega^k(C)$, since

$$\chi \lrcorner p(\alpha) = \chi \lrcorner \alpha - \chi \lrcorner \alpha + (\chi \wedge \chi) \lrcorner \alpha \wedge \mathfrak{c} = 0 \,.$$

Note that the property $\alpha \in \Omega(C)$ is not preserved by the exterior derivative $\mathrm{d}\colon \Omega^k(J^1M) \to \Omega^{k+1}(J^1M)$. So with the projection p, we define the degree 1 derivation d_p as the composition

$$\mathrm{d}_p := p \circ \mathrm{d}\colon \Omega^k(J^1M) \to \Omega^{k+1}(C) \,, \tag{6}$$

Bottom Operator Since Ω is non-degenerate on C, the assignment $\xi \mapsto \xi \lrcorner \Omega$ defines an isomorphism $\iota\colon C \to C^*$, which further induces an isomorphism $\Lambda^2\iota^{-1}\colon \Lambda^2 C^* \to \Lambda^2 C$. This enables us to define $X_\Omega := \Lambda^2\iota^{-1}(\Omega)$. In coordinates,

$$X_\Omega = \partial_{q^\mu} \wedge \partial_{p_\mu} \ .$$

Contracting with the 2-vector field X_Ω leads to the *bottom operator* $\perp\colon \Omega^k(J^1U) \to \Omega^{k-2}(J^1U)$. More precisely, for k-form α, $k > 1$,

$$\perp\alpha := X_\Omega \lrcorner \alpha \ . \tag{7}$$

For $k \leq 1$ define $\perp\alpha = 0$. Our convention is such that $\perp\Omega = \partial_{p_\mu} \lrcorner \partial_{q^\mu} \lrcorner (\mathrm{d}q^\mu \wedge \mathrm{d}p_\mu) = n$. The motivation for defining the bottom operator will be more apparent in the next paragraphs.

2.3 Monge-Ampère Operators and Effective Forms

Definition 3.3 Let $\omega \in \Omega^n(J^1M)$ be an arbitrary n-form, $n = \dim M$. The Monge-Ampère operator corresponding to ω, $\Delta_\omega\colon C^\infty(M) \to \Omega^n(M)$, is defined as

$$\Delta_\omega\phi := (j^1\phi)^*\omega \ . \tag{8}$$

The differential equation

$$\Delta_\omega\phi = 0 \tag{9}$$

is called a Monge-Ampère equation.

Notice that the expression $\Delta_\omega\phi = 0$ defines an equation on M only when ω is a $n = \dim M$-form. In this way, the M-A operators enable us to represent M-A equations by differential forms. Note that we have a certain ambiguity in this representation due to

$$(j^1\phi)^*\mathfrak{c} = \mathrm{d}\phi - \phi_\mu \mathrm{d}q^\mu = 0 \ .$$

In full generality, this ambiguity is described by an ideal of the exterior algebra over J^1M, generated by the contact form and its exterior derivative

$$\mathcal{I} =< \mathfrak{c}, \mathrm{d}\mathfrak{c} >\subset \Omega(J^1M) \ . \tag{10}$$

Recall that $\Omega(J^1M)$ is a graded algebra, which implies that $\mathcal{I}$ is a graded ideal

$$\mathcal{I}^k := \mathcal{I} \cap \Omega^k(J^1M) \ .$$

Thus, the redundancy in M-A equations is given by

$$\omega \in \mathcal{I}^n \iff \Delta_\omega \phi = 0 \; \forall \phi \; . \tag{11}$$

This suggest to work with the equivalence classes of $\Omega^n(J^1M)/\mathcal{I}^n$ instead of using arbitrary forms in $\Omega^n(J^1M)$ to describe M-A equations on M. Nevertheless, such an approach is not very convenient for computations in local coordinates. To avoid this problem, we use the following definition of *effective forms*, which captures the above idea of working with forms which do not contain the redundant terms.

Definition 3.4 Let $\omega \in \Omega^k(J^1M)$, $k \leq n$. Then ω is called effective, if

$$\chi \lrcorner \omega = 0 \text{ and } \perp\omega = 0 \; . \tag{12}$$

For further details about effective forms and how the above definition can be linked with the equivalence classes of $\Omega^n(J^1M)/\mathcal{I}^n$, see [1, 2].

Recall that $\chi \lrcorner \omega = 0$ means $\omega \in \Omega^k(\mathcal{C})$ (see (4)). The conditions (12) will be our working definition when dealing with effective forms. Note also that the condition $\perp\omega = 0$ is equivalent to $\Omega \wedge \omega = 0$ if and only if $n = k$.

Example 3.1 Let $\beta = \mathrm{d}q^1 \wedge \mathrm{d}q^2 \wedge \ldots \wedge \mathrm{d}q^n$ and $\beta_\mu := \partial_{q^\mu} \lrcorner \beta$. Then

$$\omega = b_\mu \beta_\mu \wedge \mathrm{d}p_\mu + b\beta$$

is effective for arbitrary choice of $b, b_\mu \in C^\infty(J^1M)$, $\mu = 1, \ldots, n$. Indeed, ω does not contain the $\mathrm{d}u$ term, hence we have $\chi \lrcorner \omega = 0$. Next, we have

$$\perp\omega = b_\mu \perp(\beta_\mu \wedge \mathrm{d}p_\mu) + b\perp\beta$$

due to $C^\infty(J^1M)$-linearity of the interior product $\chi \lrcorner$. Recall that we use the summation convention, so $\beta_\mu \wedge \mathrm{d}p_\mu$ consists of n terms. The first one is $\beta_1 \wedge \mathrm{d}p_1 = \mathrm{d}q^2 \wedge \ldots \wedge \mathrm{d}q^n \wedge \mathrm{d}p_1$. The bottom operator gives

$$\perp(\beta_1 \wedge \mathrm{d}p_1) = (\partial_{q^\mu} \wedge \partial_{p_\mu}) \lrcorner (\beta_1 \wedge \mathrm{d}p_1) = \partial_{p_\mu} \lrcorner \partial_{q^\mu} \lrcorner (\beta_1 \wedge \mathrm{d}p_1) = \partial_{q^1} \lrcorner \beta_1 = 0 \, .$$

Similarly for all the other terms of $\beta_\mu \wedge \mathrm{d}p_\mu$. Obviously $\perp\beta = 0$ since β does not contain any $\mathrm{d}p$ term. We see that ω is effective. Notice that the coefficients of ω might depend on u.

Important result in the theory of effective forms is the Hodge-Lepage decomposition, proved by V. Lychagin in [2] using the representation theory of $\mathfrak{sl}_2(\mathbb{R})$.

Theorem 3.1 *Every $\omega \in \Omega^k(C)$, $k \leq n$, can be written in the form*

$$\omega = \omega_\epsilon + x \wedge \Omega \,, \tag{13}$$

for some $x \in \Omega^{k-2}(C)$ and a uniquely given $\omega_\epsilon \in \Omega^k(C)$ satisfying $\perp\omega_\epsilon = 0$.

Corollary 3.1 *Suppose that $\omega_1, \omega_2 \in \Omega^n(C)$ determine the same Monge-Ampère equation. Then the effective parts satisfy*

$$\omega_{1\epsilon} = k\omega_{2\epsilon} \tag{14}$$

for a non-vanishing function $k \in C^\infty(J^1M)$.

Proof Two forms determine the same equation if and only if for all ϕ

$$\Delta_{\omega_1}\phi = \tilde{k}\Delta_{\omega_2}\phi \,, \tag{15}$$

for some non-vanishing $\tilde{k} \in C^\infty(M)$. Notice that Δ is $C^\infty(J^1M)$-equivariant in the ω argument, i.e. for arbitrary ω and $k \in C^\infty(J^1M)$ we have[6]

$$\Delta_{k\omega}\phi = \big((j^1\phi)^*k\big)\Delta_\omega\phi \,.$$

Moreover, Δ is $\mathbb{R}$-linear in the lower argument, so for arbitrary ω_1, ω_2, and all ϕ

$$\Delta_{\omega_1}\phi - \Delta_{\omega_2}\phi = \Delta_{\omega_1-\omega_2}\phi \,.$$

Hence (15) can be rewritten as

$$\Delta_{\omega_1}\phi - \tilde{k}\Delta_{\omega_2}\phi = \Delta_{\omega_1-k\omega_2}\phi = 0 \,,$$

for appropriate $k \in C^\infty(J^1M)$ s.t. $(j^1\phi)^*k = \tilde{k}$. The above equation holds for all ϕ if and only if

$$\alpha := \omega_1 - k\omega_2 \in \mathcal{I}^n$$

(see (11)). Since every $\alpha \in \mathcal{I}^n$ satisfies $\alpha_\epsilon = 0$ and every $\omega \in \Omega(C)$ satisfies $(k\omega)_\epsilon = k\omega_\epsilon$, we conclude $\omega_{1\epsilon} = k\omega_{2\epsilon}$. □

Using the projection operator (5) together with the Hodge-Lepage decomposition, we know that every k-form ω on J^1M has a unique effective part ω_ϵ (of the same degree). This means that every M-A equation $\Delta_\omega\phi = 0$ can be represented by a unique differential form which does not contain terms generating trivial equation. We will use this observation in order to study the variational nature of the PDEs under consideration.

[6] Note that $(j^1\phi)^*k = k \circ j^1\phi$ since k is a function.

3 Lagrangians, Variational Problems and the Euler Operator

Taking the pullback of a n-form on the jet space results in a n-form on the base manifold M, which can be integrated over M. Let ϕ be compactly supported, $\omega \in \Omega^n(J^1M)$. Define the (action) functional corresponding to $\Delta_\omega\phi$ by

$$\Phi_\omega[\phi] = \int_M \Delta_\omega\phi \, . \tag{16}$$

Definition 3.5 We call an element $\omega \in \Omega^n(J^1M)$ a Lagrangian. A first-order Lagrangian is a n-form ω such that $\Delta_\omega\phi$ depends on ϕ up to the first order.

3.1 First-Order Lagrangians

We are focused on the first-order Lagrangians as defined in Definition 3.5 because they yield all possible first-order Lagrangian functions on M.[7] The following lemma describes the most general form the first-order Lagrangians can have.

Proposition 3.1 *Every effective first-order Lagrangian for one scalar field ϕ is locally of the form*

$$L\beta = L(q^\mu, u, p_\mu)\mathrm{d}q^1 \wedge \ldots \wedge \mathrm{d}q^n \, . \tag{17}$$

for some $L \in C^\infty(J^1M)$.

Proof Let $\omega \in \Omega^n(J^1M)$ be arbitrary. If $\Delta_\omega\phi$ is assumed to depend on the first derivatives of ϕ at most, then ω cannot contain any $\mathrm{d}p_i$ term. Thus

$$\omega = L\beta + L_I \mathrm{d}q^I \wedge \mathrm{d}u \, ,$$

where $\beta = \mathrm{d}q^1 \wedge \ldots \wedge \mathrm{d}q^n$ and $L, L_I \in C^\infty(J^1M)$ with $I = i_1 \ldots i_{k-1}$ running through all possible combinations s.t. $1 \le i_1 \le \ldots \le i_{k-1} \le n$. Now recall that ω can still contain some terms resulting in zero after the pullback. Due to the Hodge-Lepage decomposition (13), every ω has a unique effective part ω_ϵ and the corresponding functionals satisfy

$$\int_M \Delta_\omega\phi = \int_M \Delta_{\omega_\epsilon}\phi \, .$$

So without loss of generality, we may assume that ω is effective. This implies two things: $\chi \lrcorner \omega = 0$ and $\perp\omega = 0$. The first condition rules out the terms containing

[7] After the pullback by $(j^1\phi)^*$ and choice of the volume form on M.

du and we are left with $\omega = L\beta$. It is easy to check that $\perp L\beta = 0$, meaning that $L\beta$ is effective. Thus we conclude that (17) *is the most general first-order Lagrangian for one scalar field ϕ, which does not contain any terms that would vanish after the pullback on M.* □

3.2 Euler-Lagrange Equations and the Euler Operator

Every functional $\Phi_\omega[\phi]$ defines a variational problem $\delta\Phi_\omega[\phi] = 0$ and the corresponding E-L equation. Once we fix a functional, we may compute the E-L equation explicitly. A natural question at this point is whether we can find $\tilde{\omega} \in \Omega^n(J^1M)$ so that the E-L equation $\delta\Phi_\omega[\phi] = 0$ is given by the Monge-Ampère equation $\Delta_{\tilde{\omega}}\phi = 0$. The answer is positive and $\tilde{\omega}$ can be determined using the *Euler operator $\mathcal{E}$.*

Definition 3.6 Euler operator $\mathcal{E}$: $\Omega^n(J^1M) \to \Omega^n(J^1M)$, $n = \dim M$ is defined by

$$\mathcal{E} := \mathrm{d}_p \perp \mathrm{d}_p + \mathcal{L}_\chi \ , \tag{18}$$

where d_p is defined by (6), $\perp$ is defined by (7), and $\mathcal{L}_\chi$ is the Lie derivative along the Reeb field given by (2).

The key motivation for us to work with the Euler operator is the following equivalence

$$\delta\Phi_\omega[\phi] = 0 \iff \Delta_{\mathcal{E}(\omega)}\phi = 0 \, . \tag{19}$$

In other words, the variational problem given by functional of ω is described by $\mathcal{E}(\omega)$. The proof of this statement and many other useful properties, as well as the details about the cohomological origin of the defining equation (18) can be found in [1, 2].

We have the following lemma, which will be used to formulate the necessary conditions for the existence of a first-order Lagrangian of a given PDE (i.e. necessary conditions for the existence of a solution to a given local inverse variational problem).

Lemma 3.1 *Let $L\beta \in \Omega^n(J^1M)$ be a first-order Lagrangian, $\mathcal{E}$ be defined by* (18). *Then*

1. *$\mathcal{E}(L\beta)$ is effective.*
2. *$\Delta_{\mathcal{E}(L\beta)}\phi = 0$ is the E-L equation of $\Phi_{L\beta}[\phi]$.*

Proof Assume the local coordinates satisfying (1) and observe that $\chi \lrcorner L\beta = L\chi \lrcorner \beta = 0$. Direct computation gives

$$\mathrm{d}_p \perp \mathrm{d}_p(L\beta) = \frac{\partial^2 L}{\partial p_\mu \partial p_\nu}\beta_\mu \wedge \mathrm{d}p_\nu - (\frac{\partial^2 L}{\partial q^\mu \partial p_\mu} + p_\mu \frac{\partial^2 L}{\partial u \partial p_\mu})\beta ,$$

where $\beta_\mu := \partial_{q^\mu} \lrcorner \beta = \partial_{q^\mu} \lrcorner (\mathrm{d}q^1 \wedge \ldots \wedge \mathrm{d}q^n)$. Using the Cartan formula $\mathcal{L} = \lrcorner d + d \lrcorner$, we further obtain

$$\mathcal{L}_\chi(L\beta) = \frac{\partial L}{\partial u}\beta + L(\chi \lrcorner \mathrm{d}\beta + \mathrm{d}\chi \lrcorner \beta) = \frac{\partial L}{\partial u}\beta .$$

Thus, following the definition (18), the coordinate expression of $\mathcal{E}(L\beta)$ is

$$\mathcal{E}(L\beta) = \frac{\partial^2 L}{\partial p_\mu \partial p_\nu}\beta_\mu \wedge \mathrm{d}p_\nu - (\frac{\partial^2 L}{\partial q^\mu \partial p_\mu} + p_\mu \frac{\partial^2 L}{\partial u \partial p_\mu} - \frac{\partial L}{\partial u})\beta . \tag{20}$$

Let us denote $B_{\mu\nu} := \frac{\partial^2 L}{\partial p_\mu \partial p_\nu}$ and $\beta_{\mu\nu} := (\partial_{q^\mu} \wedge \partial_{q^\nu}) \lrcorner \beta$. Hence $B_{\nu\mu} = B_{\mu\nu}$, and $\beta_{\nu\mu} = -\beta_{\mu\nu}$. We will check that $\mathcal{E}(L\beta)$ is effective (see Definition 12). Firstly recall that $\chi = \partial_u$ and that (20) does not contain $\mathrm{d}u$, so $\chi \lrcorner \mathcal{E}(L\beta) = 0$. Secondly, since $\perp\beta = 0$,

$$\perp\mathcal{E}(L\beta) = B_{\mu\nu}(\partial_{q^\alpha} \wedge \partial_{p_\alpha}) \lrcorner (\mathrm{d}p_\nu \wedge \beta_\mu) = -B_{\mu\nu}\partial_{q^\nu} \lrcorner \beta_\mu = \begin{cases} -B_{\mu\nu}\beta_{\mu\nu} & \mu = \nu \\ 0 & \mu \neq \nu \end{cases} .$$

Writing the sums over μ, ν explicitly, the term $B_{\mu\nu}\beta_{\mu\nu}$ reads as

$$B_{\mu\nu}\beta_{\mu\nu} = \sum_{\mu<\nu}(B_{\mu\nu}\beta_{\mu\nu} + B_{\nu\mu}\beta_{\nu\mu}) = \sum_{\mu<\nu} B_{\mu\nu}(\beta_{\mu\nu} - \beta_{\mu\nu}) = 0 ,$$

which implies $\perp\mathcal{E}(L\beta) = 0$.

To show the latter statement, we firstly notice that $\mathcal{E}$ is a 0 degree operator, which follows directly from $\deg \mathrm{d}_p = 1, \deg \perp = -2, \deg \mathcal{L} = 0$. Hence starting with $L\beta \in \Omega^n(J^1M)$, the result $\mathcal{E}(L\beta)$ is also a n-form and $\Delta_{\mathcal{E}(\omega)}\phi = 0$ is a well-defined equation on M. The property (19) is then expressed for $\omega = L\beta$ as follows

$$\delta\Phi_{L\beta}[\phi] = \delta \int_M (j^1\phi)^* L\beta = 0 \iff \Delta_{\mathcal{E}(L\beta)}\phi = 0 .$$

Using the coordinate description of $\mathcal{E}(L\beta)$ given by (20), we get

$$\Delta_{\mathcal{E}(L\beta)}\phi = 0 \iff \frac{\partial (j^1\phi)^* L}{\partial \phi} - \frac{\partial}{\partial q^\mu}\frac{\partial (j^1\phi)^* L}{\partial \phi_\mu} = 0 ,$$

which is the standard form of the E-L equation for a first-order Lagrangian function $(j^1\phi)^*L = L(q^\mu, \phi, \phi_\mu)$ on M, corresponding to $\Phi_{L\beta}[\phi] = \int_M \Delta_{L\beta}\phi$. $\square$

4 Effective Forms and the Inverse Variational Problem

In this section, we will see how M-A equations can be described by effective forms, which provide a unique (up to a scalar multiple) representation of the equation by a differential form on the first jet space.[8] This enables us to show that both Plebański heavenly, Husain and Grant equations do not have a first-order Lagrangian which would solve the corresponding (local) inverse variational problem.

The first and easy step is to find a simple representation of the equation (see Definition 3.7). The simple representation might not be effective. Indeed, this is the case in all the aforementioned equations. The Hodge-Lepage decomposition (13) assures that we can always find the effective part of a given form, although it does not give a recipe for doing so. Thus we introduce Lemma 3.3 which provides an efficient algorithmic way to determine the effective form of a M-A equation in the case $\dim M = 4$. The following lemma is an intermediate step.

Lemma 3.2 *Let $\omega \in \Omega^2(C)$ be arbitrary and $\Omega = d\mathfrak{c}$ be the symplectic form on the contact structure $C \subset T(J^1M)$. The following holds*

$$\perp(\omega \wedge \Omega) = (\perp\omega)\Omega + (n-2)\omega\ , \tag{21}$$

where $n = \dim M$.

Proof Recall that, in the local coordinates s.t. (1) holds, we have $\Omega = \mathrm{d}q^\mu \wedge \mathrm{d}p_\mu$ and $\perp\omega = (\partial_{q^\mu} \wedge \partial_{p_\mu}) \lrcorner\, \omega = \partial_{p_\mu} \lrcorner\, \partial_{q^\mu} \lrcorner\, \omega$, which implies $\perp\Omega = n$. Hence

$$\perp(\omega \wedge \Omega) = (\perp\omega)\Omega - \partial_{q^\mu} \lrcorner\, \omega \wedge \partial_{p_\mu} \lrcorner\, \Omega + \partial_{p_\mu} \lrcorner\, \omega \wedge \partial_{q^\mu} \lrcorner\, \Omega + n\omega\ . \tag{22}$$

We will show that the middle two terms add up to -2ω. Note that the basis of $\Omega^2(C)$ consists of pairs $\mathrm{d}q^\mu \wedge \mathrm{d}q^\nu$, $\mathrm{d}q^\mu \wedge \mathrm{d}p_\nu$, $\mathrm{d}p_\mu \wedge \mathrm{d}p_\nu$. Because ∂_q, ∂_p are duals to $\mathrm{d}q, \mathrm{d}p$, the basis of $\Omega^2(C)$ satisfies

$$\begin{aligned}
&\partial_{q^\mu} \lrcorner\, (\mathrm{d}q^\nu \wedge \mathrm{d}q^\xi) = \delta_{\mu\nu}\mathrm{d}q^\xi - \delta_{\mu\xi}\mathrm{d}q^\nu\ , && \partial_{p_\mu} \lrcorner\, (\mathrm{d}q^\nu \wedge \mathrm{d}q^\xi) = 0\ ,\\
&\partial_{q^\mu} \lrcorner\, (\mathrm{d}q^\nu \wedge \mathrm{d}p_\xi) = \delta_{\mu\nu}\mathrm{d}p_\xi\ , && \partial_{p_\mu} \lrcorner\, (\mathrm{d}q^\nu \wedge \mathrm{d}p_\xi) = -\delta_{\mu\xi}\mathrm{d}q^\nu\ ,\\
&\partial_{q^\mu} \lrcorner\, (\mathrm{d}p_\nu \wedge \mathrm{d}p_\xi) = 0\ , && \partial_{p_\mu} \lrcorner\, (\mathrm{d}p_\nu \wedge \mathrm{d}p_\xi) = \delta_{\mu\nu}\mathrm{d}p_\xi - \delta_{\mu\xi}\mathrm{d}p_\nu .
\end{aligned}$$

Since every $\omega \in \Omega^2(C)$ is of the form $\omega = \omega_{IJ}\mathrm{d}q^I \wedge \mathrm{d}p_J$ for some functions $\omega_{IJ} \in C^\infty(J^1M)$, where I, J are ascending multiindices of appropriate length. Due to C^∞-linearity of $\lrcorner$, we can, without loss of generality, assume that all ω_{IJ} are constant functions, say $\omega_{IJ} = 1$, and write

$$\omega = \sum_{\nu<\xi} \mathrm{d}q^\nu \wedge \mathrm{d}q^\xi + \sum_{\nu,\xi} \mathrm{d}q^\nu \wedge \mathrm{d}p_\xi + \sum_{\nu<\xi} \mathrm{d}p_\nu \wedge \mathrm{d}p_\xi\ .$$

[8] The equation can be reconstructed from the differential form via the M-A operator (8).

Using the above relations we obtain

$$\partial_{q^\mu} \lrcorner \omega \wedge \partial_{p_\mu} \lrcorner \Omega = (\delta_{\mu\nu}\mathrm{d}q^\xi - \delta_{\mu\xi}\mathrm{d}q^\nu + \delta_{\mu\nu}\mathrm{d}p_\xi) \wedge (-\mathrm{d}q^\mu) = 2\mathrm{d}q^\nu \wedge \mathrm{d}q^\xi + \mathrm{d}q^\nu \wedge \mathrm{d}p_\xi \ ,$$

and similarly

$$\partial_{p_\mu} \lrcorner \omega \wedge \partial_{q^\mu} \lrcorner \Omega = (-\delta_{\mu\xi}\mathrm{d}q^\nu + \delta_{\mu\nu}\mathrm{d}p_\xi - \delta_{\mu\xi}\mathrm{d}p_\nu)\mathrm{d}p_\mu = -\mathrm{d}q^\nu \wedge \mathrm{d}p_\xi - 2\mathrm{d}p_\nu \wedge \mathrm{d}p_\xi \ .$$

Combining the last two results to fit the terms in (22) yields

$$-\partial_{q^\mu} \lrcorner \omega \wedge \partial_{p_\mu} \lrcorner \Omega + \partial_{p_\mu} \lrcorner \omega \wedge \partial_{q^\mu} \lrcorner \Omega = -2\omega \ ,$$

which proves the formula (21). □

We use the previous lemma to prove the following. A general formula and its proof can be found in [2].

Lemma 3.3 *Let $\omega \in \Omega^4(C)$ be arbitrary, $n = \dim M > 2$. The effective part ω_ϵ is given by*

$$\omega_\epsilon = \omega - \frac{1}{n-2} \perp\omega \wedge \Omega + \frac{\perp^2\omega}{2(n-1)(n-2)} \Omega \wedge \Omega \ . \tag{23}$$

Proof Consider the Hodge-Lepage decomposition

$$\omega = \omega_\epsilon + x \wedge \Omega \ ,$$

where $\omega_\epsilon \in \Omega^k(C)$ is the unique effective part of ω and $x \in \Omega^{k-2}(C)$ is not necessarily effective. Applying $\perp$ twice on the above equation together with the formula (21) gives the following system

$$\begin{aligned} \perp\omega &= (\perp x)\Omega + (n-2)x \ , \\ \perp^2\omega &= 2(n-1)\perp x \ , \end{aligned}$$

which can be solved for x

$$x = \frac{1}{(n-2)} \perp\omega - \frac{\perp^2\omega}{2(n-1)(n-2)} \Omega \ .$$

Substituting this into the Hodge-Lepage decomposition yields the formula for the effective part of a 4-form ω. □

A differential k-form is called *simple* if it contains only one summand, when expressed in the canonical coordinates (1). For example, let $k = 2$. Then $\mathrm{d}q^1 \wedge \mathrm{d}q^2$ is simple while $\mathrm{d}q^1 \wedge \mathrm{d}q^2 + \mathrm{d}q^3 \wedge \mathrm{d}q^4$ is not simple.

Definition 3.7 Consider a M-A equation $\Delta_\omega \phi = 0$. Then ω is called a simple representation of the equation, if it has constant coefficients and contains the minimal number of simple terms.

Remark 3.2 Note that the property of being simple is basis dependent. On the other hand, the effectivity is a basis independent notion.

It seems natural to denote Lagrangian functions and their corresponding counterpart defined on J^1M by the same symbol, i.e. to write $L = L(q^\mu, u, p_\mu)$ as well as $(j^1\phi)^*L = L(q^\mu, \phi, \phi_\mu)$. To avoid any confusion, we distinguish the two in the following proposition as follows. A Lagrangian function that can be integrated over M will be L, its J^1M counterpart will be $\tilde{L}$.

Proposition 3.2 *Let $\Delta_\omega \phi = 0$ be a M-A equation over an open subset of a smooth manifold M, $\dim M = n$. Then a necessary condition for a first-order Lagrangian function $L = L(q^\mu, \phi, \phi_\mu)$ to be a local solution of the inverse variational problem corresponding to $\Delta_\omega \phi = 0$ is*

$$k\omega_\epsilon = \mathcal{E}(\tilde{L}\beta) \ , \tag{24}$$

for some non-vanishing function $k\colon J^1M \to \mathbb{R}$, where ω_ϵ is the effective part of ω, $\mathcal{E}$ is the Euler operator given by (18), $\tilde{L}\colon J^1M \to \mathbb{R}$ is such that $\tilde{L} \circ j^1\phi = L(q^\mu, \phi, \phi_\mu)$, and $\beta = \mathrm{d}q^1 \wedge \ldots \mathrm{d}q^n$.

Proof Let $\alpha \in \Omega^n(J^1M)$ be a first-order Lagrangian in the sense of the Definition 3.5, i.e. $\Delta_\alpha \phi = L\beta$, for some L (possibly defined only locally) which depends smoothly on ϕ up to the first-order in derivatives, $L = L(q^\mu, \phi, \phi_\mu)$. Assume that the E-L equation for L is given by $\Delta_\omega \phi = 0$. Define

$$\Phi_\alpha[\phi] := \int_M \Delta_\alpha \phi = \int_M L\beta$$

(consider only ϕ compactly supported). Without loss of generality, we may restrict α to be effective (see the discussion in the subsection with effective forms) and thus by Proposition 3.1, we (locally) have $\alpha = \tilde{L}\beta$ for appropriate $\tilde{L} \in C^\infty(J^1M)$ satisfying $\tilde{L} \circ j^1\phi = L$. Thus $\Phi_\alpha[\phi] = \Phi_{\tilde{L}\beta}[\phi]$ and, by the second statement of Lemma 3.1, we know that the E-L equation for the functional $\Phi_{\tilde{L}\beta}[\phi]$ is $\Delta_{\mathcal{E}(\tilde{L}\beta)}\phi = 0$. Since we assumed that L locally solves the inverse variational problem given by the equation $\Delta_\omega \phi = 0$, and because ω and ω_ϵ determine the same equation, we have

$$\Delta_{\omega_\epsilon}\phi = 0 \iff \Delta_{\mathcal{E}(\tilde{L}\beta)}\phi = 0 \, .$$

By the first statement of Lemma 3.1, $\mathcal{E}(\tilde{L}\beta)$ is an effective form. Since ω_ϵ and $\mathcal{E}(\tilde{L}\beta)$ are effective forms determining the same equation, the Corollary 3.1 implies that the forms must differ by a multiple of a non-vanishing function. □

Remark 3.3 Although we work locally in a coordinate system, notice that the necessary conditions for the existence of a solution to the inverse variational problem is, in our framework, a tensorial statement and thus independent of the choice of coordinates.

We present the following, simple example in $\dim M = 2$ to show how the Proposition 3.2 can be used.

Example Consider the 1D wave equation (understand one of the two coordinates as time)

$$\phi_{11} - c\phi_{22} = 0\,, \tag{25}$$

where $c > 0$ is a real constant, $\phi\colon M \to \mathbb{R}$, and $\dim M = 2$. We want to find $L(q^\mu, u, p_\mu) \in C^\infty(J^1U)$ s.t. the E-L equation for $(j^1\phi)^*L = L(q^\mu, \phi, \phi_\mu)$ is (25).

The simple representation is

$$\omega = -c\mathrm{d}q^1 \wedge \mathrm{d}p_2 - \mathrm{d}q^2 \wedge \mathrm{d}p_1\ .$$

We can easily see that $\Delta_\omega \phi = 0$ gives the original equation

$$(j^1\phi)^*\omega = -c\mathrm{d}q^1 \wedge \mathrm{d}\phi_2 - \mathrm{d}q^2 \wedge d\phi_1 = (\phi_{11} - c\phi_{22})\mathrm{d}q^1 \wedge \mathrm{d}q^2\ .$$

The simple representation is effective, $\omega = \omega_\epsilon$, since it degenerates along χ

$$\chi \lrcorner\, \omega = \partial_u \lrcorner\, (-c\mathrm{d}q^1 \wedge \mathrm{d}p_2 - \mathrm{d}q^2 \wedge \mathrm{d}p_1) = 0\,,$$

and belongs to the kernel of the bottom operator

$$\perp\omega = \partial_{p_\mu} \lrcorner \partial_{q^\mu} \lrcorner (-c\mathrm{d}q^1 \wedge \mathrm{d}p_2 - \mathrm{d}q^2 \wedge \mathrm{d}p_1) = c\partial_{p_1} \lrcorner (-\mathrm{d}p_2) - \partial_{p_2} \lrcorner \mathrm{d}p_1 = 0\,.$$

The coordinate expression of the Euler operator evaluated on a general first-order Lagrangian n-form is given by (20). For $n = 2$ we have $\beta = \mathrm{d}q^1 \wedge \mathrm{d}q^2$ and $\beta_1 = \partial_{q^1} \lrcorner\, \beta = \mathrm{d}q^2$, $\beta_2 = \partial_{q^2} \lrcorner\, \beta = -\mathrm{d}q^1$, so (20) becomes

$$\begin{aligned}\mathcal{E}(L\beta) = \tfrac{\partial^2 L}{\partial {p_1}^2}\mathrm{d}q^2 \wedge dp_1 + \tfrac{\partial^2 L}{\partial p_1 \partial p_2}\mathrm{d}q^2 \wedge dp_2 - \tfrac{\partial^2 L}{\partial p_2 \partial p_1}\mathrm{d}q^1 \wedge dp_1 - \tfrac{\partial^2 L}{\partial {p_2}^2}\mathrm{d}q^1 \wedge dp_2 \\ -(\tfrac{\partial^2 L}{\partial q^1 \partial p_1} + \tfrac{\partial^2 L}{\partial q^2 \partial p_2} + p_1 \tfrac{\partial^2 L}{\partial u \partial p_1} + p_2 \tfrac{\partial^2 L}{\partial u \partial p_2} + \tfrac{\partial L}{\partial u})\mathrm{d}q^1 \wedge \mathrm{d}q^2\end{aligned}$$

(continued)

We can fix the value of the function in (24) to be constant, say $k = 1$, since two forms which are multiple of each other by a smooth non-vanishing k yields the same M-A equation. Hence we search for $L \in C^\infty(J^1M)$ such that $\omega = \mathcal{E}(L\beta)$, which implies

$$\frac{\partial^2 L}{\partial {p_1}^2} = -1\ , \quad \frac{\partial^2 L}{\partial {p_2}^2} = c\ , \quad \frac{\partial L}{\partial u} = \frac{\partial L}{\partial q^\mu} = 0\ , \ \mu = 1, 2\ .$$

Thus $L = L(p_\mu)$ and we can solve the first two conditions by the choice

$$L = \frac{1}{2}(-{p_1}^2 + c{p_2}^2)\ .$$

because the M-A equation $\Delta_{\mathcal{E}(L\beta)}\phi = 0$ writes

$$\frac{\partial (j^1\phi)^* L}{\partial \phi} - \frac{\partial}{\partial q^\mu}\frac{\partial (j^1\phi)^* L}{\partial \phi_\mu} = \phi_{11} - c\phi_{22} = 0\ .$$

We see that $(j^1\phi)^*L = \frac{1}{2}(-{\phi_1}^2 + c{\phi_2}^2)$ is a solution to the inverse problem for (25). □

4.1 Plebański, Grant, and Husain Equations

Proceeding in a similar fashion as in the previous example, we analysed both Plebański heavenly, Grant, and Husain equations in dim = 4. The following tables summarize simple representations, show their non-effectivity and display effective parts of the simple representations of the aforementioned PDEs, ϕ being a real function. Since the effective forms of M-A equations in four dimensions tend to have lengthy expressions, we introduce the following shorthand notation, which also facilitate the computations. We denote

$$d^\mu := \mathrm{d}q^\mu, \qquad\qquad d_\mu := \mathrm{d}p_\mu\ ,$$

and for the wedge product, we write

$$d^\mu_{\ \nu} := \mathrm{d}q^\mu \wedge \mathrm{d}p_\nu, \qquad\qquad {d_\nu}^\mu := \mathrm{d}p_\nu \wedge \mathrm{d}q^\mu\ .$$

Notice that the position and order of indices matter and there are obvious relations such as $d^\mu_{\ \nu} = -{d_\nu}^\mu$, or for the contractions $\partial_{q^\mu} \lrcorner\, d^\nu = \delta^\mu_\nu$ (the Kronecker delta) and $\partial_{q^\mu} \lrcorner\, d_\nu = \partial_{p_\mu} \lrcorner\, d^\nu = 0$, et cetera. For example, the symplectic form is in the above

Table 1 Simple representations (which are not effective, $\perp\omega \neq 0$) of 1st Plebański (P1), 2nd Plebański (P2), Grant (G), and Husain (H) equations

	Monge-Ampère equation	Simple representation
1st Plebański	$\phi_{13}\phi_{24} - \phi_{14}\phi_{23} = 1$	$\omega_{P1} = d^{12}_{12} - d^{12}_{34}$
2nd Plebański	$\phi_{11}\phi_{22} - (\phi_{12})^2 + \phi_{13} + \phi_{24} = 0$	$\omega_{P2} = d^{123}_{2} - d^{124}_{1} + d^{34}_{12}$
Grant	$\phi_{11} + \phi_{24}\phi_{13} - \phi_{23}\phi_{14} = 0$	$\omega_G = -d^{234}_{1} - d^{12}_{12}$
Husain	$\phi_{13}\phi_{24} - \phi_{14}\phi_{23} + \phi_{11} + \phi_{22} = 0$	$\omega_H = d^{134}_{2} - d^{234}_{1} + d^{12}_{12}$

Table 2 Effective parts of simple representations of P1, P2, G, and H

	Effective form ω_ϵ
1st Plebański	$\omega_{P1\epsilon} = -d^{1234} + \frac{1}{3}(d^{12}_{12} + d^{34}_{34}) - \frac{1}{6}(d^{13}_{13} + d^{14}_{14} + d^{23}_{23} + d^{24}_{24})$
2nd Plebański	$\omega_{P2\epsilon} = \frac{1}{2}(d^{124}_{1} + d^{123}_{2} + d^{234}_{3} + d^{134}_{4}) + d^{34}_{12}$
Grant	$\omega_{G\epsilon} = -d^{234}_{1} + \frac{1}{3}(d^{12}_{12} + d^{34}_{34}) - \frac{1}{6}(d^{13}_{13} + d^{14}_{14} + d^{23}_{23} + d^{24}_{24})$
Husain	$\omega_{H\epsilon} = d^{134}_{2} - d^{234}_{1} + d^{12}_{12} + d^{34}_{34} - \frac{1}{2}(d^{13}_{13} + d^{14}_{14} + d^{23}_{23} + d^{24}_{24})$

notation written as $\Omega = d^1_1 + \ldots + d^n_n$, the volume form on M is $\beta = d^{1234}$, and so on.

Proposition 3.2 yields the following result.

Corollary 3.2 *Monge-Ampère equations from Table 1 do not correspond to a variational problem of a first-order Lagrangian function.*

Proof Table 2 shows the effective forms of Monge-Ampère equations under consideration. In all cases, the effective form contains at least one term of the form $d^{\mu\nu}_{\xi\eta}$. These terms do not occur in the expression (20). Thus the necessary condition for the existence of a first-order Lagrangian, given by the Proposition 3.2, is not satisfied. □

We want to emphasize here that although the Plebański heavenly, Grant, and Husain equations do not have a first-order Lagrangian for which they would be E-L equations, in a different setup a Lagrangian can be found [8, 16]. Let us consider the second heavenly equation

$$\phi_{11}\phi_{22} - (\phi_{12})^2 + \phi_{13} + \phi_{24} = 0\,. \tag{26}$$

If we single-out one coordinate among $q^1, \ldots, q^4$, say q^1, and introduce a new function ψ, then we can write (26) as an evolution system in q^1

$$\psi - \phi_1 = 0\,, \tag{27}$$

$$\psi_1\phi_{22} - {\psi_2}^2 + \psi_3 + \phi_{24} = 0\,, \tag{28}$$

Interestingly, the above system is a variational problem, since it is given by the E-L equations

$$\frac{\partial L}{\partial \phi} - \frac{\partial}{\partial q^\mu}\frac{\partial L}{\partial \phi_\mu} + \frac{\partial^2}{\partial q^\mu \partial q^\nu}\frac{\partial L}{\partial \phi_{\mu\nu}} = 0\,,$$

$$\frac{\partial L}{\partial \psi} - \frac{\partial}{\partial q^\mu}\frac{\partial L}{\partial \psi_\mu} + \frac{\partial^2}{\partial q^\mu \partial q^\nu}\frac{\partial L}{\partial \psi_{\mu\nu}} = 0\,,$$

of the functional

$$L[\phi, \psi] = \psi\phi_1\phi_{22} + \frac{1}{2}\phi_1\phi_3 - \frac{1}{2}\psi^2\phi_{22} + \frac{1}{2}\phi_2\phi_4\,. \tag{29}$$

In [16], a method for treating the general case of Monge-Ampère equations is provided, together with systematic approach of finding Lagrangians for them after the decomposition into an evolution system. For further details regarding the above case, see [8].

The following example shows an equation which has a first-order Lagrangian, the corresponding effective form does not have constant coefficients, and is not a differential form over the cotangent bundle. We will see that the conditions of Proposition 3.2 are satisfied.

4.2 Klein-Gordon Equation

Let M be a four-dimensional Minkowski spacetime with coordinates q^μ and flat metric $\eta_{\mu\nu}$ with signature $(+, -, -, -)$. Consider the (linear) Klein-Gordon equation

$$\phi_{11} - \phi_{22} - \phi_{33} - \phi_{44} + m^2\phi^2 = 0\,, \tag{30}$$

where $m \in \mathbb{R}$ is a constant. We can describe (30) as a M-A equation $\Delta_\omega \phi = 0$ via the form

$$\omega = -\beta_1 \wedge \mathrm{d}p_1 + \sum_{\mu=2}^{4} \beta_\mu \wedge \mathrm{d}p_\mu + m^2 u\beta\,.$$

This 4-form is not a simple representation of (30), due to the non-constant coefficient m^2u, but it is an effective form, see the Example 3.1. Comparing ω with the local form of $\mathcal{E}(L\beta)$ for general L (see (20)), we obtain the following set of conditions

$$\eta_{\mu\nu} = \tfrac{\partial^2 L}{\partial p_\mu \partial p_\nu},\ \mu, \nu = 1, \dots 4\,,$$

$$-m^2 u = \tfrac{\partial^2 L}{\partial q^\mu \partial p_\mu} + p_\mu \tfrac{\partial^2 L}{\partial u \partial p_\mu} - \tfrac{\partial L}{\partial u}\,.$$

One can easily check that the function L

$$L = \frac{1}{2}(-{p_1}^2 + \sum_{\mu=2}^{4} {p_\mu}^2 + m^2u^2) \in C^\infty(J^1M)$$

satisfies all the above conditions. It follows that

$$(j^1\phi)^*L = \frac{1}{2}(-{\phi_1}^2 + \sum_{\mu=2}^{4} {\phi_\mu}^2 + m^2\phi^2)$$

is a first-order Lagrangian for the Klein-Gordon equation.

5 Multisymplectic Formulation

In [3] F. Hélein provided a multisymplectic formulation of the Klein-Gordon equation (30) (in dimension n) over $\mathcal{M} := \Lambda^n T^*(M \times \mathbb{R})$, equipped with the multisymplectic form [18, 19]

$$\mathfrak{m} := \mathrm{d}e \wedge \beta + \mathrm{d}p_\mu \wedge \mathrm{d}\phi \wedge \beta_\mu \ , \tag{31}$$

where e is a fiber coordinate of the trivial line bundle $M \times \mathbb{R} \to M$, p_μ are the cotangent coordinates, $\beta = \mathrm{d}q^1 \wedge \ldots \wedge \mathrm{d}q^n$ and $\beta_\mu = \partial_{q^\mu} \lrcorner \beta$, with q^μ coordinates on a n-dimensional Minkowski spacetime M. Using (31), the following Hamiltonian function on $\mathcal{M}$ is defined in such a way to correspond to solutions of (30)

$$\mathcal{H} := e + \frac{1}{2}\eta_{\mu\nu} p_\mu p_\nu + \frac{1}{2}m^2\phi^2 \ ,$$

where $\eta_{\mu\nu}$ is the Minkowski metric with signature $(+, -, \ldots, -)$. Each solution of (30) is then interpreted as a Hamiltonian n-curve, defined by equations

$$p_\mu = \eta^{\mu\nu}\phi_\nu, \mu = 1, \ldots, n \ ,$$
$$e = -\frac{1}{2}\eta^{\mu\nu}\phi_\mu\phi_\nu - \frac{1}{2}m^2\phi^2 \ ,$$

where $\eta^{\mu\nu}$ is the inverse to $\eta_{\mu\nu}$. In the aforementioned paper, F. Hélein provided a canonical pre-quantization of the Klein-Gordon equation, and defined the notion of observables together with their brackets, which give rise to an infinite dimensional analogue of the Heisenberg algebra. The starting point of the method is the existence of a Lagrangian, which in the context of the Klein-Gordon equation is a first-order one. For more details see [3, 12, 17].

The following theorem is due to D. Harrivel. It enables us to associate to certain effective forms on J^1M their (non-unique) multisymplectic counterpart on the trivial line bundle over J^1M. The proof can be found in [4]. Note that the key difference with respect to the previous multisymplectic formulation of F. Hélein is that them multisymplectic form can be associated with Monge-Ampère equations which are not variational, that is, equations which are not Euler-Lagrange for some first-order Lagrangian. As we have seen in the previous section, this is the case for all the equations in Table 1.

Theorem 3.2 *Let $\omega \in \Omega^n(C)$ be an effective form, $n = \dim M$. Consider a trivial line bundle $\mathcal{T} := J^1M \times \mathbb{R} \to J^1M$ with fiber coordinate e. Define $\mathfrak{m}_\omega \in \Omega^{n+1}(\mathcal{T})$ by*

$$\mathfrak{m}_\omega := \mathrm{d}e \wedge \beta + \mathfrak{c} \wedge \omega \,. \tag{32}$$

Then $\mathfrak{m}_\omega$ is a multisymplectic form if and only if

1. *The set $\mathcal{S}_\omega := \{\partial_{q^1} \lrcorner\, \omega, \dots, \partial_{q^n} \lrcorner\, \omega\}$ is linearly independent over $\Omega^{n-1}(C)$, and,*
2. *$\mathrm{d}_p\omega = 0$.*

Once an equations has a simple representation, the corresponding effective form has constant coefficients, and thus the second assumption of Theorem 3.2 is trivially satisfied since $\mathrm{d}_p = p \circ \mathrm{d}$. The linear independence of the set $\mathcal{S}_\omega$ in the case of 4D equations is decided over $\binom{\dim C}{\dim M - 1} = \binom{8}{3} = 56$-dimensional space of 3-forms on C. In all our cases, this can be determined almost without computation.

5.1 Plebański, Grant, and Husain Equations

For the first heavenly equation we have

$$\mathcal{S}_{P1} = \{-d^{234} + x, d^{134} + y, -d^{124} + z, d^{123} + w\}\,,$$

where x, y, z, w are linear combinations of $d^{\mu}_{\nu\xi}$, for appropriate μ, ν, ξ. We see that $\mathcal{S}_{P1}$ is linearly independent. Similarly for the second heavenly equation

$$\begin{aligned}\mathcal{S}_{P2} = \{\, & \tfrac{1}{2}(d^{24}_{\ 1} + d^{23}_{\ 2} + d^{34}_{\ 4}) + d^{34}_{\ 2}, \tfrac{1}{2}(-d^{14}_{\ 1} - d^{13}_{\ 2} + d^{34}_{\ 3}) - d^{34}_{\ 1}, \\ & \tfrac{1}{2}(d^{12}_{\ 2} - d^{24}_{\ 3} - d^{14}_{\ 4}) + d^{4}_{\ 12}, \tfrac{1}{2}(d^{12}_{\ 1} + d^{23}_{\ 3} + d^{13}_{\ 4}) + d^{3}_{\ 12}\}\,,\end{aligned}$$

which is a linearly independent set as the simple terms are all different. It is not difficult to check that the sets $\mathcal{S}_G$ and $\mathcal{S}_H$ for Grant and Husain equations, respectively, are also linearly independent. Thus the 5-form $\mathfrak{m}_\omega$ is a multisymplectic form on $J^1M \times \mathbb{R}$ in all the four cases described in Table 1.

5.2 Klein-Gordon Equation

Interestingly, and in contrast with the Plebański, Grant, and Husain equations, the 5-form for the Klein-Gordon equation defined by (32) is not a multisymplectic form. To see this, take the differential 4-form

$$\omega = -\beta_1 \wedge \mathrm{d}p_1 + \sum_{\mu=2}^{4} \beta_\mu \wedge \mathrm{d}p_\mu + m^2 u\beta \ ,$$

which, as we already discussed, is effective and represents (30) as a Monge-Ampère equation $\Delta_\omega \phi = 0$. Due to the non-constant $m^2 u$ term, the exterior derivative gives

$$\mathrm{d}\omega = m^2 \mathrm{d}u \wedge \beta \ ,$$

which is not degenerate along the Reeb field. Thus $\mathrm{d}_p \neq \mathrm{d}$ and we have to project the form down to $\Omega(C)$ (see (5) for the definition of p)

$$\mathrm{d}_p\omega = m^2(\mathrm{d}u \wedge \beta - \mathfrak{c} \wedge \chi \lrcorner (\mathrm{d}u \wedge \beta)) = m^2 p_\mu \mathrm{d}q^\mu \wedge \beta \ .$$

We see that the second condition of the Theorem 3.2 is not satisfied and thus $\mathfrak{m}_\omega$ given by (32) is not a multisymplectic form. Notice that the first condition of the theorem is not violated as the set $\mathcal{S}_\omega$ is linearly independent.

6 Conclusion and Discussion

In this work, we mainly focused on the following two questions. Firstly, can we decide whether a first-order Lagrangian for a given Monge-Ampère equation exists? Secondly, motivated by the work of F. Hélein [3] and D. Harrivel [4], can we associate a multisymplectic form to equations which are not variational with respect to a first-order Lagrangian?

Regarding the first question, we provided a partial answer by formulating a necessary condition for the existence of a local solution to this inverse variational problem. This was done by representing a given equation by an effective differential form over the first jet space, and comparing it with an n-form that produces Euler-Lagrange equation for a general, first-order Lagrangian function.

Comparing the effective forms yields a computationally straightforward and simple method for obtaining a non-trivial information about Monge-Ampère equations in the context of strong inverse variational problems. Using the method, we showed that Plebański heavenly equations, Grant equation and Husain equation are not variational in our sense. Recall that the first heavenly equation is equivalent with the Grant equation after appropriate change of coordinates [6]. Using a similar approach, we have shown (as expected) that the Klein-Gordon equation is variational by finding the well-known Lagrangian for it. The hypothesis is that the self-duality conditions imposed to derive the previous four equations creates an obstruction for the existence of the first-order Lagrangian. We want to study this problematics in more detail in our future work.

The presented method is much more suitable for deciding the non-variational nature of a given equation than solving the local inverse problem explicitly. Moreover, it works only when restricted to the case of first-order Lagrangians. Nevertheless, this limitation can be seen as desirable, since the first-order Lagrangians are of great importance throughout the physics.

It is not clear at the moment how to generalize our approach to the case of more functions. The procedure can be naively extended for more scalar fields by introducing multiple Euler operators, the cost being degeneracy issues. This causes further problems, for example in the context of the unique decomposition of differential forms into the effective and non-effective part, which is an essential tool in our approach. In [9], B. Banos used the notion of bi-effective forms to efficiently deal with the complex Monge-Ampère equations, and proved the possibility to always obtain a unique bieffective decomposition. This is not equivalent in an obvious way to the aforementioned naive extension, as the Verbitsky-Bonan relations are not satisfied in our case (see [9], Theorem 1). This is connected with the fact that we do not restrict our forms to have coefficients independent of the u coordinate on J^1M (which allows us to work, for example, with the Klein-Gordon equation). Whether this problems can be resolved will be part of our future investigations.

Regarding the second question focused on the multisymplectic formulation of Monge-Ampère equations. Using the results of [4], we provided multisymplectic 5-forms in the case of real 4-dimensional heavenly Plebański, Grant, and Husain equations, all of which are not variational in our sense. Interestingly, the same approach does not work for the Klein-Gordon equation as the corresponding 5-form is not multisymplectic.

F. Hélein's multisymplectic treatment of the Klein-Gordon equation provided in [3] starts with a first-order Lagrangian function. The other four Monge-Ampère equations we studied cannot be treated in the same way, unless going into higher order Lagrangians. On the other hand, the Theorem 3.2 provides a multisymplectic forms exactly for the four non-variational cases and fails for the Klein-Gordon equation. To provide some explanation of this, it would be interesting to compare the methods of [3, 17] with those in [4] in the situation of a general Monge-Ampère equation.

Acknowledgments This paper was written during my visit in Angers as a part of my PhD research, under the cotutelle agreement between the Masaryk University, Brno, Czech Republic, and the University of Angers, France. I am grateful for the funding provided by the Czech Ministry of Education, and by the Czech Science Foundation under the project GAČR EXPRO GX19-28628X, and I thank the University of Angers for the hospitality during the research period. I also want to express my gratitude to Volodya Rubtsov for his numerous valuable suggestions and detailed comments, and to Jan Slovák for clarification of concepts from the theory of jet bundles. The results were reported at the Winter School and Workshop Wisla 20-21, a European Mathematical Society event organized by the Baltic Institute of Mathematics.

References

1. Kushner, A., Lychagin, V., Rubtsov, V.: Contact Geometry and Nonlinear Differential Equations. Cambridge University Press (2006)
2. Lychagin, V.: Contact geometry and non-linear second-order differential equations. Russian Math. Surv. **34**, 149–180 (1979)
3. Hélein, F.: Multisymplectic Formalism and the Covariant Phase Space. Variational Problems in Differential Geometry, pp. 94–126 (2011)
4. Harrivel, D.: Hamiltonian, Multisymplectic Formalism and Monge-Ampère Equations. Systèmes Intégrables Et Théorie Quantiques Des Champs, pp. 331–354 (2008)
5. Plebanski, J.: Some solutions of complex Einstein equations. J. Math. Phys. **16**, 2395–2402 (1975)
6. Grant, J.: On self-dual gravity. Phys. Rev. D **48**, 2606–2612 (1993). https://link.aps.org/doi/10.1103/PhysRevD.48.2606
7. Husain, V.: Self-dual gravity as a two-dimensional theory and conservation laws. Classical Quantum Gravity **11**, 927–937 (1993)
8. Neyzi, F., Nutku, Y., Sheftel, M.: Multi-Hamiltonian structure of Plebanski's second heavenly equation. J. Phys. A. **38**, 8473 (2005)
9. Banos, B.: Complex solutions of Monge-Ampère equations. J. Geom. Phys. **61**, 2187–2198 (2011). https://www.sciencedirect.com/science/article/pii/S0393044011001641
10. Gordon, W.: Z. Phys. **40**, 117–133 (1926)
11. Radzikowski, M.: Micro-local approach to the Hadamard condition in quantum field theory on curved space-time. Commun. Math. Phys. **179**, 529–553 (1996)
12. Hélein, F., Kouneiher, J.: The notion of observable in the covariant hamiltonian formalism for the calculus of variations with several variables. Adv. Theor. Math. Phys. **8**, 735–777 (2004)
13. Campos, C.M., Guzmán, E., Marrero, J.C.: Classical field theories of first order and Lagrangian submanifolds of premultisymplectic manifolds. J. Geom. Mech. **4** (2012)
14. Ashtekar, A., Jacobson, T., Smolin, L.: A new characterization of half-flat solutions to Einstein's equation. Commun. Math. Phys. **115**, 631–648 (1988)
15. Kolář, I., Michor, P., Slovák, J.: Natural Operations in Differential Geometry. Springer, Berlin Heidelberg (1993). https://www.emis.de/monographs/KSM/
16. Nutku, Y.: Hamiltonian structure of real Monge - Ampère equations. J. Phys. A Math. Gen. **29**, 3257–3280 (1996). https://doi.org/10.1088/0305-4470/29/12/029
17. Hélein, F.: Hamiltonian formalisms for multidimensional calculus of variations and perturbation theory (arXiv, 2002). https://arxiv.org/abs/math-ph/0212036
18. Gaset, J., Román-Roy, N.: Multisymplectic unified formalism for Einstein-Hilbert gravity (arXiv). https://arxiv.org/abs/1705.00569v5
19. Román-Roy, N.: Some Properties of Multisymplectic Manifolds (arXiv). arXiv:1807.11774v2

Generalized Solvable Structures Associated to Symmetry Algebras Isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$

Adrián Ruiz and Concepción Muriel

Abstract Lie symmetry algebras that are isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$ are nonsolvable, hence the standard methods of integration by quadratures cannot be applied to solve ordinary differential equations that are invariant under the action of $\mathrm{GL}(2, \mathbb{R}) \ltimes \mathbb{R}$. In this work it is proved the existence of a generalized solvable structure for the vector field associated with a fifth-order equation admitting a Lie symmetry algebra isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$. As a consequence, the integrability of the given equation splits into two integration processes of second and third-order, respectively. On one hand, two functionally independent first integrals of the equation are computed by quadratures alone. On the other hand, the third-order integration process involves a third-order equation that admits a Lie symmetry algebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$, which is also nonsolvable. Previous results regarding the integrability of $\mathrm{SL}(2, \mathbb{R})$-invariant third-order equations allow us to obtain the general solution to the original fifth-order equation in implicit form and expressed in terms of a fundamental set of solutions to a two-parameter family of Schrödinger-type equations. An example is also included with the aim of showing the effectiveness of the method. Remarkably, the considered example does not have additional Lie point symmetries, apart from the symmetry generators of $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$.

1 Introduction

Many phenomena that appear in physics and engineering can be modelled by the use of both ordinary and partial differential equations. Therefore it is important to develop new techniques for the search of exact solutions for differential equations with the goal of advancing in the knowledge of the events under study. In this regard, the Norwegian mathematician Sophus Lie developed one of the most powerful tools

A. Ruiz (✉) · C. Muriel
Department of Mathematics, University of Cádiz, Campus Río San Pedro, Cádiz, Spain
e-mail: adrian.ruiz@uca.es; concepcion.muriel@uca.es

M. Ulan, S. Hronek (eds.), *Groups, Invariants, Integrals, and Mathematical Physics*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-031-25666-0_4

for solving differential equations. Motivated by the notion of symmetry group of the roots of a polynomial equation established by Galois, Lie introduced the concept of symmetry group of a differential equation. In general terms, and following a parallel analogy with Galois' theory, a symmetry group of a differential equation is a local group of local transformations that maps solutions of the equation into solutions of the equation [1–8]. In the context of ordinary differential equations, one of the main results states that the existence of a solvable n-dimensional symmetry algebra for an nth-order equation guarantees its integrability by quadrature [2–8].

However, there exist some cases in which the classical Lie method fails to obtain exact solutions of the equation because either the symmetry algebra is nonsolvable or the dimension of the symmetry algebra is strictly less than the order of the equation [38]. This fact has motivated the appearance in the recent years of different generalizations of the classical concept of Lie symmetry such as hidden symmetries [9], nonlocal symmetries [10–12], λ-symmetries [13–15], μ-symmetries [16], σ-symmetries [17] and solvable structures [18–23].

Among the above mentioned extensions, we focus on the concept of solvable structure because constitutes a natural generalization of the notion of solvable symmetry algebra [24] that characterizes the integrability by quadratures of an involutive distribution of vector fields. The application of solvable structures to address the problem of the integrability of equations admitting nonsolvable symmetry algebras has been particularly useful in the recent literature. The case of third-order equations admitting a Lie symmetry algebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$ has been widely studied in [25–30]. In [31, 39] it was proved the existence of a solvable structure for $\mathrm{SL}(2, \mathbb{R})$-invariant third-order equations by using the symmetry generators of $\mathfrak{sl}(2, \mathbb{R})$. This theoretical result was used in [32] to provide the general solution for any $\mathrm{SL}(2, \mathbb{R})$-invariant third-order equation in parametric form in terms of a fundamental set of solutions to a related second-order linear equation. This complemented the results previously reported in the recent literature [25–30].

The notion of solvable structure has also been extended by means of the concept of generalized solvable structure, which is defined as a usual solvable structure not just for the vector field associated to the equation, but for an involutive distribution of vector fields that includes the vector associated to the equation [33, 34]. This is particularly useful for the case of equations of arbitrary order n that admits a Lie symmetry algebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$, because the determination of a generalized solvable structure for that type of equations permits to obtain the general solution in parametric form and expressed in terms of a $(n-3)$-parameter family of second-order linear equations [34]. This approach has been successfully applied to fourth-order ordinary differential equations admitting a four-dimensional nonsolvable symmetry algebra [35], extending the results previously reported for $\mathrm{SL}(2, \mathbb{R})$-invariant third-order equations.

The goal of this paper is to address the problem of the integrability of fifth-order equations admitting a Lie symmetry algebra isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$, which is five-dimensional and nonsolvable. The paper is organized as follows. Firstly, in Sect. 2, we provide an introduction regarding the foundations of solvable structures. In Sect. 3, we explicitly determine a generalized solvable structure for the vector field associated to the equation by using the symmetry generators of $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$. This permits to split the problem into two different integration processes. One of them consists of calculating two functionally independent first integrals by quadrature, whereas the second one involves a third-order equation admitting a Lie symmetry algebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$. As a consequence, it is proved that the general solution of the original fifth-order equation can be expressed in implicit form in terms of a fundamental set of solutions to a two-parameter family of Schrödinger-type equations. We would like to emphasize that the obtainment of closed-form solutions for this type equations may be a challenging task due the nonsolvability nature of the underlying symmetry algebra, as well as for its high dimension. Finally, we include in Sect. 4 an example of a fifth-order equation whose Lie symmetry algebra is five-dimensional and isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$.

2 Preliminaries: Solvable Structures

In this section, with the aim of being self-contained, we recall the basics regarding solvable structures and its application to integrate ordinary differential equations. For a more extensive study, the reader can consult [18, 19, 22]. From this point on, functions, vector fields and differential 1-forms are assumed to be smooth on a simply-connected open set U of an n-dimensional manifold $\mathcal{M}_n$.

The notion of solvable structure is established for systems of vector fields of the form $\mathcal{A} = \{\mathbf{A}_1, \ldots, \mathbf{A}_r\}$, $1 \leq r \leq n$, that are in involution, i.e.:

$$[\mathbf{A}_i, \mathbf{A}_j] = \sum_{k=1}^{r} c_{i,j}^k \mathbf{A}_k, \quad 1 \leq i, j \leq r \quad \text{and } c_{i,j}^k \in C^\infty(U).$$

The classical notion of symmetry can be extended for involutive systems of vector fields as follows [31]:

Definition 4.1 Let $\mathcal{A} = \{\mathbf{A}_1, \ldots, \mathbf{A}_r\}$, with $1 \leq r < n$, be an involutive system of pointwise linearly independent vector fields on U. A vector field $\mathbf{V}$ is called a symmetry of $\mathcal{A}$ if and only if

1. $\mathbf{A}_1, \ldots, \mathbf{A}_r$, and $\mathbf{V}$ are pointwise linearly independent on U;
2. $[\mathbf{V}, \mathbf{A}_i] \in \operatorname{span}(\mathcal{A})$, for $i = 1, \ldots, r$.

Previous definition permits to establish the concept of solvable structure [18, 19, 22]:

Definition 4.2 Let $\mathcal{S} = \langle \mathbf{X}_1, \ldots, \mathbf{X}_{n-r} \rangle$ be an ordered set of pointwise linearly independent vector fields on U. The ordered system

$$\mathcal{A} \cup \mathcal{S} = \langle \mathbf{A}_1, \ldots, \mathbf{A}_r, \mathbf{X}_1, \ldots, \mathbf{X}_{n-r} \rangle$$

is a **solvable structure** with respect to $\mathcal{A}$ if

1. $\mathcal{S}_j = \{\mathbf{A}_1, \ldots, \mathbf{A}_r, \mathbf{X}_1, \ldots, \mathbf{X}_j\}$ is involutive, for $j = 1, \ldots, n-r$;
2. $\mathbf{X}_1$ is a symmetry of $\mathcal{A}$;
3. $\mathbf{X}_{j+1}$ is a symmetry of $\mathcal{S}_j$, for $j = 1, \ldots, n-r-1$.

The existence of a solvable structure characterizes the (local) integrability by quadrature of the system [19, Proposition 6]:

Proposition 4.1 *An involutive system $\mathcal{A}$ is locally integrable by quadrature if and only if there exists a solvable structure with respect to $\mathcal{A}$.*

These results can be applied to the trivially involutive system associated to an nth-order ordinary differential equation written in explicit form

$$u_n = \varphi(x, u, u_1 \ldots, u_{n-1}), \tag{1}$$

where x is the independent variable, u is the dependent variable and $u_i = \frac{d^i u}{dx^i}$, for $i = 1, \ldots, n$. Equation (1) is assumed to be defined on a suitable domain $U \subset J^n(\mathbb{R}, \mathbb{R})$, where, for $k \in \mathbb{N}$, $J^k(\mathbb{R}, \mathbb{R})$ stands for the kth-order jet space [36]. We denote by M the projection of this domain to the zero-order jet space and $M^{(k)}$ the corresponding jet space of order $k \geq 1$. The vector field associated to Eq. (3) is defined on $M^{(n-1)}$ and is given by

$$\mathbf{A} = \partial_x + u_1 \partial_u + u_2 \partial_{u_1} + \cdots + \varphi(x, u, u_1, \ldots, u_{n-1}) \partial_{u_{n-1}}.$$

If a solvable structure $\langle \mathbf{A}, \mathbf{X}_1, \cdots, \mathbf{X}_n \rangle$ with respect to $\{\mathbf{A}\}$ is known, then one can find, at least locally, a complete set of first integrals of $\mathbf{A}$ by quadrature alone. In order to do that, let $\boldsymbol{\Omega} = dx \wedge \cdots \wedge du_{n-1}$ denote the volume form on $M^{(n-1)}$ and define the differential 1-forms given by

$$\boldsymbol{\omega}_i = \frac{\mathbf{X}_n \lrcorner \cdots \lrcorner \widehat{\mathbf{X}_i} \lrcorner \cdots \lrcorner \mathbf{X}_1 \lrcorner \mathbf{A} \lrcorner \boldsymbol{\Omega}}{\mathbf{X}_n \lrcorner \cdots \lrcorner \mathbf{X}_1 \lrcorner \mathbf{A} \lrcorner \boldsymbol{\Omega}}, \quad \text{for} \quad i = 1, \cdots, n, \tag{2}$$

where $\lrcorner$ is the interior product and $\widehat{\mathbf{X}_i}$ indicates omission of $\mathbf{X}_i$. The 1-forms (2) have distinguishing closure properties [22]:

- $\boldsymbol{\omega}_n$ is locally exact and a primitive I_n is a first integral of $\mathbf{A}$.
- The restriction of $\boldsymbol{\omega}_{n-1}$ to each submanifold defined by $I_n = c_n, c_n \in \mathbb{R}$, is closed and then locally exact (i.e., $\boldsymbol{\omega}_{n-1}$ is locally exact module $\boldsymbol{\omega}_n$). In consequence, a primitive I_{n-1} of $\boldsymbol{\omega}_{n-1}$ (module $\boldsymbol{\omega}_n$) can be found by quadrature.
- We can continue this process until all the 1-forms (2) have been integrated. This permits to compute a complete system of functionally independent first integrals of $\mathbf{A}$.

These results provide the following consequence [31, Theorem 2.3]:

Theorem 4.1 *Let* $\mathbf{A}$ *be the vector field associated to an nth-order ordinary differential equation. If* $\langle \mathbf{A}, \mathbf{X}_1, \dots, \mathbf{X}_n \rangle$ *is a solvable structure with respect to* $\{\mathbf{A}\}$*, then the corresponding equation can be (at least locally) solved by quadrature alone.*

The notion of solvable structure for an nth-order ordinary differential equation can be adapted to include involutive systems of vector fields containing the vector field associated to the equation [33, 34]:

Definition 4.3 Let $\mathbf{A}$ be the vector field associated to an nth-order ordinary differential equation. An ordered set of vector fields $\langle \mathbf{A}, \mathbf{X}_1, \dots, \mathbf{X}_n \rangle$ is a generalized solvable structure with respect to $\{\mathbf{A}\}$ if

- there exists $k \in \mathbb{N}$, with $1 \leq k \leq n-1$, such that $\langle \mathbf{A}, \mathbf{X}_1, \dots, \mathbf{X}_k \rangle$ is involutive,
- the ordered set $\langle \mathbf{A}, \mathbf{X}_1, \dots, X_k, \dots, \mathbf{X}_n \rangle$ is a solvable structure with respect to $\langle \mathbf{A}, \mathbf{X}_1, \cdots, \mathbf{X}_k \rangle$.

Generalized solvable structures are specially useful for solving ordinary differential equations of arbitrary order n admitting nonsolvable symmetry algebras [34, 35], or for equations whose Lie symmetry algebra is insufficient to integrate them by quadrature [33].

3 Generalized Solvable Structures for $\mathrm{GL}(2, \mathbb{R}) \ltimes \mathbb{R}$-Invariant Fifth-Order Equations

Once the foundations of solvable structures and generalized solvable structures have been presented, we aim to apply such concepts to integrate $\mathrm{GL}(2, \mathbb{R}) \ltimes \mathbb{R}$-invariant fifth-order equations. Thus, let us consider a fifth-order ordinary differential equation written in explicit form

$$u_5 = \varphi(x, u, u_1, u_2, u_3, u_4), \tag{3}$$

and assume that it admits a Lie symmetry algebra isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$. There exists only one inequivalent action of the Lie group $\mathrm{GL}(2, \mathbb{R}) \ltimes \mathbb{R}$ on a real two-dimensional manifold, which can be modelled by the Lie algebra of vector fields spanned by [37, 40]

$$\mathbf{v}_1 = \partial_x, \quad \mathbf{v}_2 = x^2 \partial_x, \quad \mathbf{v}_3 = x\partial_x, \quad \mathbf{v}_4 = \partial_u, \quad \mathbf{v}_5 = u\partial_u, \tag{4}$$

which satisfy the following commutation relations:

$$\begin{array}{lll} [\mathbf{v}_1, \mathbf{v}_3] = \mathbf{v}_1, & [\mathbf{v}_1, \mathbf{v}_2] = 2\mathbf{v}_3, & [\mathbf{v}_3, \mathbf{v}_2] = \mathbf{v}_2, \\ [\mathbf{v}_1, \mathbf{v}_4] = 0, & [\mathbf{v}_2, \mathbf{v}_4] = 0, & [\mathbf{v}_3, \mathbf{v}_4] = 0, \\ [\mathbf{v}_1, \mathbf{v}_5] = 0, & [\mathbf{v}_2, \mathbf{v}_5] = 0, & [\mathbf{v}_3, \mathbf{v}_5] = 0, \\ [\mathbf{v}_4, \mathbf{v}_5] = \mathbf{v}_4. & & \end{array} \tag{5}$$

Any other basis of generators of the Lie algebra $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$ can be therefore mapped to the basis elements given in (4) by means of a local change of variables. If we denote by $\mathbf{V}_i = \mathbf{v}_i^{(4)}$, for $i = 1, 2, 3, 4, 5$, then the symmetry condition implies that

$$[\mathbf{V}_i, \mathbf{A}] = -\mathbf{A}(\mathbf{V}_i(x))\mathbf{A}, \quad i = 1, 2, 3, 4, 5. \tag{6}$$

3.1 Construction of a Generalized Solvable Structure

As a result of the commutation relations (5) and (6), the following theorem holds:

Theorem 4.2 *Let* $\mathbf{A}$ *be the vector field associated to a fifth-order equation admitting a Lie symmetry algebra isomorphic to* $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$ *spanned by* $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5\}$ *and denote* $\mathbf{V}_i = \mathbf{v}_i^{(4)}$*, for* $i = 1, 2, 3, 4, 5$*. Then the ordered set of vector fields*

$$\langle \mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{V}_5 \rangle$$

is a solvable structure with respect to the integrable distribution $\langle \mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3 \rangle$*. As a consequence,* $\langle \mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{V}_5 \rangle$ *is a generalized solvable structure for* $\{\mathbf{A}\}$*, for* $k = 3$ *(see Definition 4.3).*

The explicit determination of a solvable structure with respect to the involutive distribution of vector fields $\langle \mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3 \rangle$ can be used to compute two functionally independent first integrals of Eq. (3) by quadratures alone. We consider the differential 1-forms

$$\boldsymbol{\omega}_5 = \frac{\mathbf{V}_4 \lrcorner \mathbf{V}_3 \lrcorner \mathbf{V}_2 \lrcorner \mathbf{V}_1 \lrcorner \mathbf{A} \lrcorner \boldsymbol{\Omega}}{\mathbf{V}_5 \lrcorner \mathbf{V}_4 \lrcorner \mathbf{V}_3 \lrcorner \mathbf{V}_2 \lrcorner \mathbf{V}_1 \lrcorner \mathbf{A} \lrcorner \boldsymbol{\Omega}} \tag{7}$$

and

$$\boldsymbol{\omega}_4 = \frac{\mathbf{V}_5 \lrcorner \mathbf{V}_3 \lrcorner \mathbf{V}_2 \lrcorner \mathbf{V}_1 \lrcorner \mathbf{A} \lrcorner \boldsymbol{\Omega}}{\mathbf{V}_5 \lrcorner \mathbf{V}_4 \lrcorner \mathbf{V}_3 \lrcorner \mathbf{V}_2 \lrcorner \mathbf{V}_1 \lrcorner \mathbf{A} \lrcorner \boldsymbol{\Omega}}. \tag{8}$$

The 1-form $\boldsymbol{\omega}_5$ is closed, and then locally exact. A corresponding primitive

$$I_5 = I_5(x, u, u_1, u_2, u_3, u_4)$$

is a first integral common to the distribution of vector fields $\{\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4\}$. Next, we have that $\boldsymbol{\omega}_4$ is closed, and then locally exact, on each leaf defined by $I_5(x, u, u_1, u_2, u_3, u_4) = c_5$, with $c_5 \in \mathbb{R}$, i.e.,

$$d\boldsymbol{\omega}_4 = 0 \text{ mod. } \boldsymbol{\omega}_5.$$

If we determine a function $J_4 = J_4(x, u, u_1, u_2, u_3; c_5)$ such that

$$dJ_4 = \boldsymbol{\omega}_4 \text{ mod. } \boldsymbol{\omega}_5$$

then we have that

$$I_4 = J_4(x, u, u_1, u_2, u_3, I_5(x, u, u_1, u_2, u_3, u_4)) \tag{9}$$

is a first integral common to the set of vector fields $\{\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\}$ and functionally independent to I_5. As a result of the previous discussion, the next theorem has been proved:

Theorem 4.3 *If a fifth-order ordinary differential equation admits a Lie symmetry algebra isomorphic to* $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$ *then two functionally independent first integrals* I_5 *and* I_4 *of the given equation can be calculated by quadratures alone.*

Once two functionally independent first integrals I_5 and I_4 to Eq. (3) have been explicitly computed, we can consider the submanifold $N \subset M^{(4)}$ given by the level set of such first integrals, i.e.:

$$I_5(x, u, u_1, u_2, u_3, u_4) = c_5, \quad I_4(x, u, u_1, u_2, u_3, u_4) = c_4, \quad c_5, c_4 \in \mathbb{R}. \tag{10}$$

From (9) and (10) we obtain the third-order ordinary differential equation:

$$J_4(x, u, u_1, u_2, u_3; c_5) = c_4. \tag{11}$$

Besides, both I_5 and I_4 are first integrals common to the distribution of vector fields $\{\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\}$, which means that (11) turns out to be a two-parameter family of third-order equations admitting a Lie symmetry algebra spanned by

$$\mathbf{v}_1 = \partial_x, \quad \mathbf{v}_2 = x^2 \partial_x, \quad \mathbf{v}_3 = x \partial_x. \tag{12}$$

According to the commutation relations (5), the vector fields (12) generate the nonsolvable symmetry algebra $\mathfrak{sl}(2, \mathbb{R})$. At this stage we observe that the original nonsolvability problem of order five has been reduced to a nonsolvability problem of order three. In order to obtain the general solution to the family of Eqs. (11) in closed-form, we consider the differential invariants s and m common to the third-order prolongations of the vector fields given in (12), which are [32]:

$$s = u, \quad m = \frac{3u_2^2 - 2u_1 u_3}{4u_1^4}. \tag{13}$$

Thus, the submanifold N can be locally expressed in terms of the invariants s and m in the form

$$m = C(s; c_4, c_5), \quad c_4, c_5 \in \mathbb{R}, \tag{14}$$

or, in terms of the original variables, by (13):

$$u_3 = \frac{3u_2^2}{2u_1} - 2u_1^3 C(u; c_4, c_5). \tag{15}$$

Equation (15) is one of the canonical $\mathrm{SL}(2,\mathbb{R})$-invariant third-order equations appearing in [32]. The general solution to Eq. (15), and then to the original equation (3), is implicitly given by [32, Eq. (38)]:

$$x = c_3(c_1 - c_2)\frac{c_1\psi_2(u; c_4, c_5) - \psi_1(u; c_4, c_5)}{c_2\psi_2(u; c_4, c_5) - \psi_1(u; c_4, c_5)},$$

where $c_i \in \mathbb{R}$, $i = 1, 2, 3, 4, 5$, $c_3 \neq 0$, $c_1 \neq c_2$ and $\psi_1 = \psi_1(u; c_4, c_5)$, $\psi_2(u; c_4, c_5)$ form a fundamental set of solutions to the Schrödinger-type equation

$$\psi''(u) + C(u; c_4, c_5)\psi(u) = 0.$$

Thus, the following theorem has been demonstrated:

Theorem 4.4 *The general solution to a fifth-order ordinary differential equation admitting a Lie symmetry algebra isomorphic to $\mathfrak{gl}(2,\mathbb{R}) \ltimes \mathbb{R}$ can be obtained in implicit form as follows*

$$x = c_3(c_1 - c_2)\frac{c_1\psi_2(u; c_4, c_5) - \psi_1(u; c_4, c_5)}{c_2\psi_2(u; c_4, c_5) - \psi_1(u; c_4, c_5)},$$

where $c_i \in \mathbb{R}$, $i = 1, 2, 3, 4, 5$, $c_3 \neq 0$, $c_1 \neq c_2$ and $\psi_1 = \psi_1(u; c_4, c_5)$, $\psi_2(u; c_4, c_5)$ form a fundamental set of solutions to the Schrödinger-type equation

$$\psi''(u) + C(u; c_4, c_5)\psi(u) = 0.$$

The function $C = C(u; c_4, c_5)$ can be determined by expressing the submanifold defined in (11) in terms of the invariants s and m given in (13) and isolating m.

In Fig. 1 it is sketched the procedure previously explained for finding the general solution in closed-form of a fifth-order equation admitting a Lie symmetry algebra isomorphic to $\mathfrak{gl}(2,\mathbb{R}) \ltimes \mathbb{R}$.

4 Example

Consider the following fifth-order ordinary differential equation:

$$u_5 = \frac{50u_3^2u_1^2 + 30u_1^2u_2u_4 - 240u_2^2u_3u_1 + 180u_2^4}{3u_1^3}, \tag{16}$$

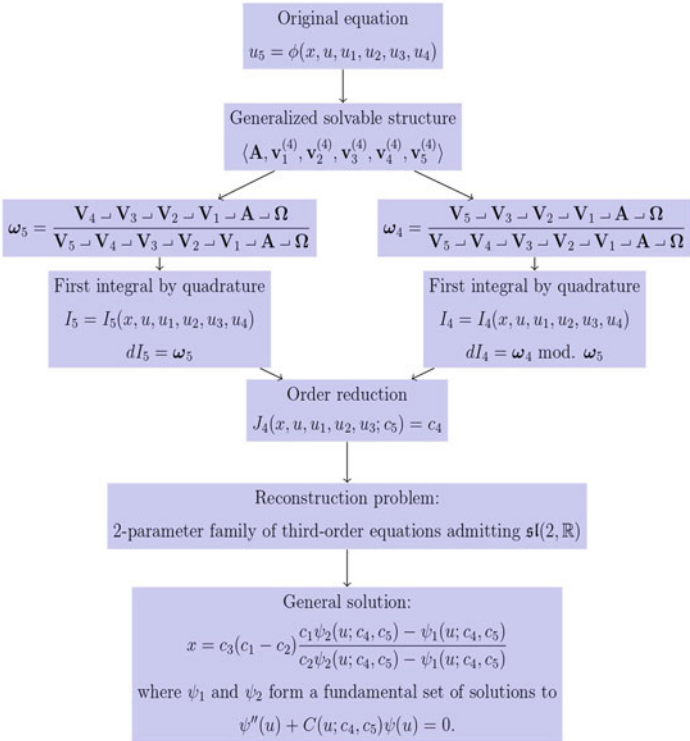

Fig. 1 Integration of fifth-order ODEs admitting $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$ via generalized solvable structures

defined on a suitable domain $U \subset J^5(\mathbb{R}, \mathbb{R})$ such that $u_1 \neq 0$. It can be checked that its Lie symmetry is five-dimensional, isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$ and spanned by the vector fields

$$\mathbf{v}_1 = \partial_x, \ \mathbf{v}_2 = x^2\partial_x, \ \mathbf{v}_3 = x\partial_x, \ \mathbf{v}_4 = \partial_u, \ \mathbf{v}_5 = u\partial_u.$$

For easier handling, we consider the local coordinates (x, u, u_1, u_2, u_3, z), where

$$z = \frac{u_4u_1^2 - 6u_2u_3u_1 + 6u_2^3}{(3u_2^2 - 2u_3u_1)^{3/2}} \tag{17}$$

is a fourth-order differential invariant common to $\{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4, \mathbf{V}_5\}$. In terms of these new coordinates the 1-form (7) becomes:

$$\boldsymbol{\omega}_5 = \frac{3(u_3u_1 - 2u_2^2)}{u_1(2u_3u_1 - 3u_2^2)}du_1 + \frac{3u_2}{2u_3u_1 - 3u_2^2}du_2 - \frac{u_1}{2u_3u_1 - 3u_2^2}du_3 - \frac{3z}{9z^2 + 8}dz,$$

and a corresponding primitive is locally defined by

$$\widetilde{I}_5 = 2\ln(u_1) - \frac{1}{2}\ln(2u_3u_1 - 3u_2^2) - \frac{1}{6}\ln(9z^2 + 8),$$

which turns out to be a first integral common to the distribution of vector fields $\{\mathbf{A}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3, \mathbf{V}_4\}$. With the aim of simplifying further computations, we choose the following functionally dependent first integral:

$$I_5 = e^{6\widetilde{I}_5} = \frac{u_1^{12}}{8(3u_2^2 - 2u_3u_1)^3 + 9(u_4u_1^2 - 6u_2u_3u_1 + 6u_2^3)^2}. \tag{18}$$

Once we have found a first integral of the equation, we can restrict the 1-form $\boldsymbol{\omega}_4$ defined in (8) to a generic leaf given by $I_5 = c_5$, with $c_5 \in \mathbb{R}$. Since on the open set U we have that $u_1 \neq 0$, then $c_5 \neq 0$ and we can isolate the coordinate u_4 from the expression

$$I_5(x, u, u_1, u_2, u_3, u_4) = c_5,$$

which produces:

$$u_4 = T(u_1, u_2, u_3; c_5) = \frac{18c_5u_1u_2u_3 - 18c_5u_3^2 \pm H(u_1, u_2, u_3; c_5)}{3c_5u_1^2}, \tag{19}$$

where the function $H = H(u_1, u_2, u_3; c_5)$ is locally given by

$$H = \sqrt{64c_5^2(u_3u_1)^3 + c_5u_1^{12} - 216(c_5^2u_2^6 - 2c_5^2u_2^4u_3u_1) - 288(c_5u_2u_3u_1)^2}. \tag{20}$$

By considering the local transformation

$$\varphi(x, u_1, u_2, u_3, u_4) = (x, u, u_1, u_2, u_2, T(u_1, u_2, u_3; c_5))$$

we have that $\varphi^*\boldsymbol{\omega}_4$ takes the form:

$$-du \mp \frac{9c_5u_1(u_3u_1 - 2u_2^2)}{H}du_1 \mp \frac{9c_5u_2u_1^2}{H}du_2 \pm \frac{3c_5u_1^3}{H}du_3, \tag{21}$$

where the expression of $H = H(u_1, u_2, u_3; c_5)$ is defined by (20). It can be checked that (21) is closed and then locally exact. In order to calculate a corresponding primitive, we introduce new local coordinates (x, u, u_1, u_2, s), where

$$s = H(u_1, u_2, u_3; c_5).$$

In terms of these new local coordinates the 1-form (21) becomes:

$$\varphi^*\boldsymbol{\omega}_4 = -du \mp \frac{3c_5^{1/3}u_1 s}{(c_5 u_1^{12} - s^2)^{2/3}}du_1 \pm \frac{c_5^{1/3}u_1^2}{2(c_5 u_1^{12} - s^2)^{2/3}}ds, \tag{22}$$

and a primitive, once expressed in the original coordinates, takes the form

$$-u \pm G\left(\frac{3u_2^2 - 2u_3u_1}{4u_1^4}; c_5\right), \tag{23}$$

where the function $G = G(m; c_5)$ is given by

$$G(m; c_5) = 4m\sqrt{c_5 - 512c_5^2m^3}\,{}_2F_1\left(\frac{5}{6}, 1, \frac{3}{2}, 1 - 512c_5m\right), \tag{24}$$

being ${}_2F_1$ the generalized hypergeometric function [41, 42]. Therefore, a first integral for the original equation, functionally independent with I_5, is given by

$$I_4 = -u \pm G\left(\frac{3u_2^2 - 2u_3u_1}{4u_1^4}; I_5\right),$$

where the expression of I_5 appears in (18). In consequence, by restricting to a generic leaf given by

$$I_5(x, u, u_1, u_2, u_3, u_4) = c_5, \qquad I_4(x, u, u_1, u_2, u_3, u_4) = c_4,$$

the following two-parameter family of third-order equations arises:

$$-u \pm G\left(\frac{3u_2^2 - 2u_3u_1}{4u_1^4}; c_5\right) = c_4, \qquad c_5, c_4 \in \mathbb{R},\ c_5 \neq 0. \tag{25}$$

Equations in the family (25) turn out to be SL(2, $\mathbb{R}$)-invariant and they can be locally expressed as follows

$$u_3 = \frac{3u_2^2}{2u_1} - 2u_1^3G^{-1}(c_4 \pm u; c_5), \tag{26}$$

where G^{-1} denotes the inverse function of G. As a consequence, we can conclude that the general solution to Eq. (26), and therefore the general solution to the original equation (16), is implicitly given by

$$x = c_3(c_1 - c_2)\frac{c_1\psi_2(u; c_4, c_5) - \psi_1(u; c_4, c_5)}{c_2\psi_2(u; c_4, c_5) - \psi_1(u; c_4, c_5)}, \tag{27}$$

where $c_i \in \mathbb{R}$, $i = 1, 2, 3, 4, 5$, $c_5, c_3 \neq 0$, $c_1 \neq c_2$ and $\psi_1 = \psi_1(u; c_4, c_5)$, $\psi_2(u; c_4, c_5)$ form a fundamental set of solutions to

$$\psi''(u) + G^{-1}(c_4 \pm u; c_5)\psi(u) = 0,$$

where G^{-1} denotes the inverse function of (24).

5 Concluding Remarks

In this paper we have addressed the problem of the integrability by quadrature of ordinary differential equations admitting a nonsolvable symmetry algebra of dimension higher than three. Specifically, we have focused on fifth-order ordinary differential equations admitting a Lie symmetry algebra isomorphic to $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$. The classical Lie reduction method cannot be applied to solve this kind of equations by quadrature because $\mathfrak{gl}(2, \mathbb{R}) \ltimes \mathbb{R}$ is nonsolvable.

The presented method is based on the determination of a generalized solvable structure for the equation, which led to the expression of the general solution of the equation in closed-form. Such solution can always be given in implicit form and expressed in terms of a fundamental set of solutions to a two-parameter family of Schrödinger-type equations. This extends the results previously obtained for lower dimensional nonsolvable Lie symmetry algebras.

Acknowledgments The authors acknowledge the financial support from *FEDER–Ministerio de Ciencia, Innovación y Universidades–Agencia Estatal de Investigación* by means of the project PGC2018-101514-B-I00 and from *Junta de Andalucía* to the research group FQM–377.

References

1. Lie, S.: Klassifikation und Integration von gewöhnlichen Differentialgleichungen zwischen x, y, die eine Gruppe von Transformationen gestatten I, II. Mat. Ann. **32**, 213–281 (1888)
2. Bluman, G., Anco, S.: Symmetry and Integration Methods for Differential Equations. Springer, New York (2002)
3. Hydon, P.E.: Symmetry Methods for Differential Equations: A Beginner's Guide. Cambridge University Press, Cambridge (2000)
4. Ibragimov, N.H.: A Practical Course in Differential Equations and Mathematical Modelling: Classical and New Methods, Nonlinear Mathematical Models, Symmetry and Invariance Principles. ALGA Publications, Karlskrona (2004)
5. Olver, P.J.: Applications of Lie Groups to Differential Equations. Springer, New York (2000)
6. Ovsiannikov, L.V.: Group Analysis of Differential Equations. Academic Press, New York (1982)
7. Schwarz, F.: Algorithmic Lie Theory for Solving Ordinary Differential Equations. Chapman & Hall/CRC Pure and Applied Mathematics. CRC Press, Boca Raton (2007)
8. Stephani, H.: Differential Equations: Their Solution Using Symmetries. Cambridge University Press, Cambridge (1989)

9. Abraham-Shrauner, B.: Hidden symmetries and nonlocal group generators for ordinary differential equations. IMA J. Appl. Math. **56**(3), 235–252 (1996)
10. Adam, A., Mahomed, F.M.: Integration of ordinary differential equations via nonlocal symmetries. Nonlinear Dyn. **30**(2), 267–275 (2002)
11. Govinder, K.S., Leach, P.G.L.: On the determination of non-local symmetries. J. Phys. A Math. Gen. **28**(18), 5349–5359 (1995)
12. Govinder, K.S., Leach, P.G.L.: A group-theoretic approach to a class of second-order ordinary differential equations not possessing Lie point symmetries. J. Phys. A Math. Gen. **30**(6), 2055–2068 (1997)
13. Muriel, C., Romero, J.L.: New methods of reduction for ordinary differential equations. IMA J. Appl. Math. **66**(2), 111–125 (2001)
14. Muriel, C., Romero, J.L.: First integrals, integrating factors and λ-symmetries of second-order differential equations. J. Phys. A Math. Theor. **42**, 365207–365224 (2009)
15. Muriel, C., Romero, J.L.: The λ-symmetry reduction method and Jacobi last multipliers. Commun. Nonlinear Sci. Numer. Simul. **19**(4), 807–820 (2014)
16. Cicogna, G., Gaeta, G., Morando, P.: On the relation between standard and μ-symmetries for PDEs. J. Phys. A Math. Gen. **37**, 9467–9486 (2004)
17. Cicogna, G., Gaeta, G., Walcher, S.: A generalization of λ-symmetry reduction for systems of ODEs: σ-symmetries. J. Phys. A Math. Theor. **45**(35), 355205–355234 (2012)
18. Barco, M.A., Prince, G.E.: Solvable symmetry structures in differential form applications. Acta Appl. Math. **66**(1), 89–121 (2001)
19. Basarab-Horwath, P.: Integrability by quadratures for systems of involutive vector fields. Ukr. Math. J. **43**(10), 1236–1242 (1991)
20. Catalano-Ferraioli, D., Morando, P.: Applications of solvable structures to the nonlocal symmetry-reduction of ODEs. Nonlinear Math. Phys. **16**(1), 27–42 (2009)
21. Catalano-Ferraioli, D., Morando, P.: Local and nonlocal solvable structures in the reduction of ODEs. J. Phys. A Math. Theor. **42**(3), 035210–035225 (2009)
22. Hartl, T., Athorne, C.: Solvable structures and hidden symmetries. J. Phys. A Math. Gen. **27**, 3463–3474 (1994)
23. Prince, G., Sherring, J.: Geometric aspects of reduction of order. Trans. Amer. Math. Soc. **334**, 433–453 (1992)
24. Duzhin, S.V., Lychagin, V.V.: Symmetries of distributions and quadrature of ordinary differential equations. Acta Appl. Math. **24**, 29–57 (1991)
25. Clarkson, P., Olver, P.J.: Symmetry and the Chazy equation. J. Differ. Equations **124**(1), 225–246 (1996)
26. Gat, O.: Symmetry algebras of third-order ordinary differential equations. J. Math. Phys. **33**(9), 2966–2971 (1992)
27. Ibragimov, N.H., Nucci, M.C.: Integration of third order ordinary differential equations by Lie's method: equations admitting three-dimensional Lie algebras. Lie Groups Appl. **1**(2), 49–64 (1994)
28. Muriel, C., Romero, J.L.: $\mathcal{C}^{\infty}$-symmetries and non-solvable symmetry algebras. IMA J. Appl. Math. **66**, 477–498 (2001)
29. Schmucker, A., Czichowski, G.: Symmetry algebras and normal forms of third order ordinary differential equation. J. Lie Theory **8**, 129–137 (1998)
30. Mgaga, T.C., Govinder, K.S.: On the linearization of some second-order ODEs via contact transformations. J. Phys. A Math. Theor. **44**, 015203–015210 (2011)
31. Ruiz, A., Muriel, C.: Solvable structures associated to the non-solvable symmetry algebra $\mathfrak{sl}(2,\mathbb{R})$. SIGMA **077**, 18 pp. (2016)
32. Ruiz, A., Muriel, C.: First integrals and parametric solutions of third-order ODEs with Lie symmetry algebra isomorphic to $\mathfrak{sl}(2,\mathbb{R})$. J. Phys. A Math. Theor. **50**, 205401–205222 (2017)
33. Muriel, C., Romero, J.L., Ruiz, A.: Integration methods for equations without enough Lie point symmetries. AIP Conf. Proc. **2153**, 020013–020021 (2018)
34. Morando, P., Muriel, C., Ruiz, A.: Generalized solvable structures and first integrals for ODEs admitting an $\mathfrak{sl}(2,\mathbb{R})$ symmetry algebra. J. Nonlinear Math. Phys. **26**(2), 188–201 (2018)

35. Ruiz, A., Muriel, C.: On the integrability of GL(2, $\mathbb{R}$)-invariant fourth-order ordinary differential equations. Math. Methods Appl. Sci. **2021**, 1–12 (2021)
36. Sardanashvily, G.: Advanced Differential Geometry for Theoreticians. Lap Lambert Academic Publishing, Saarbrucken (2013)
37. Olver, P.J.: Equivalence, Invariants, and Symmetry. Cambridge University Press, Cambridge (1995)
38. González-López, A.: Symmetry and integrability by quadratures of ordinary differential equations. Phys. Lett. A. **133**(4–5), 190–194 (1988)
39. Ruiz, A., Muriel, C.: Applications of $\mathcal{C}^{\infty}-$symmetries in the construction of solvable structures. In: Ortegón-Gallego, F., Redondo-Neble, M.V., Rodríguez-Galván, R. (eds.) Trends in Differential Equations and Applications. SEMA SIMAI Springer Series, Switzerland, pp. 387–403 (2016)
40. Kamran, N., Olver, P.J., González-López, A.: Lie algebras of the vector fields in the real plane. Proc. Lond. Math. Soc. **64**, 339–368 (1992)
41. Olver, F.W.J., Lozier, D.W., Boisvert, R.F., Clark, C.W.: NIST Handbook of Mathematical Functions. Cambridge University Press, Cambridge (2010)
42. Whittaker, E.T., Watson, G.N.: A Course of Modern Analysis, 4th edn. Cambridge University Press, Cambridge (1935)

Fundamental Groupoids and Homotopy Types of Non-compact Surfaces

Sergiy Maksymenko and Oleksii Nikitchenko

Abstract The paper contains an application of van Kampen theorem for groupoids to computation of homotopy types of certain class of non-compact foliated surfaces obtained by at most countably many strips $\mathbb{R} \times (0, 1)$ with boundary intervals in $\mathbb{R} \times \{\pm 1\}$ along some of those intervals.

1 Introduction

The present paper is devoted to applications of van Kampen theorem for groupoids to computation of homotopy types of a certain class of non-compact foliated surfaces called *striped surfaces*.

It was mentioned by M. Morse that a smooth function f with non-degenerate critical points (a Morse function) on a compact manifold Z "contains" a lot of homological information about the manifold itself (the famous Morse inequalities). In particular, if $\dim M = 2$, so Z is a "surface", one can even determine the topological type of Z by the numbers of critical points of distinct indices via any Morse function $f : M \to \mathbb{R}$. Motivated by study functions of complex variable, in particular, harmonic functions being real and imaginary parts of holomorphic functions, Morse extended his observations in the book [1] to pseudoharmonic functions f defined on compact domains Z in the complex plane $\mathbb{C}$.

By definition, a pseudoharmonic function $f : \mathbb{C} \supset Z \to \mathbb{R}$ is locally a composition $g \circ h$, where g is a harmonic function, and h is a homeomorphism of $\mathbb{C}$. Such a function is continuous, all its critical (in a proper sense) points belonging to

S. Maksymenko (✉)
Institute of Mathematics of NAS of Ukraine, Kyiv, Ukraine
e-mail: maks@imath.kiev.ua

O. Nikitchenko
Kyiv Academic University, Kyiv, Ukraine

M. Ulan, S. Hronek (eds.), *Groups, Invariants, Integrals, and Mathematical Physics*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-031-25666-0_5

the interior of Z are isolated and are not local extremes (due to maximum principle for holomorphic functions). Moreover, it is assumed that the restriction of f to ∂M has only finitely many local minimums and maximums. Absence of local extremes in the interior of Z implies (by Jordan curve theorem) that the foliation of Z into connected components of level sets of f has no closed curves.

W. Kaplan [2, 3] characterized such foliation for pseudoharmonic function on $\mathbb{R}^2$. Namely, he shown that every foliation $\mathcal{F}$ on the plane $\mathbb{R}^2$ has the following properties:

(1) every leaf ω of $\mathcal{F}$ is an image of a proper embedding $\omega : \mathbb{R} \to \mathbb{R}^2$, so $\lim\limits_{t\to\pm\infty} \omega(t) = \infty$;
(2) there exists at most countably many leaves $\{\omega_i\}_{i\in\Lambda}$ of $\mathcal{F}$, such that the complement $\mathbb{R}^2 \setminus \bigcup\limits_{i\in\Lambda} \omega_i$ is a disjoint union of "open strips" $\mathbb{R} \times (0, 1)$ foliated into lines $\mathbb{R} \times t$, $t \in (0, 1)$;
(3) there exists a pseudoharmonic function $f : \mathbb{R}^2 \to \mathbb{R}$ without critical points such that $\mathcal{F}$ is a partition of $\mathbb{R}^2$ into connected components of level-sets of f.

Property (3) is in a spirit of de Rham theory: if $\mathcal{F}$ is smooth, then it is defined by some closed differential 1-form, and since $\mathbb{R}^2$ is contractible, $\omega = df$ for some function f satisfying (3).

It also gives a certain connection between foliations on surfaces and pseudoharmonic functions on the plane. Let $\mathcal{F}'$ be a foliation on a connected surface Z' without boundary distinct from 2-sphere and projective plane, and $p : Z \to Z'$ be the universal covering map. Then we get a well-defined foliation $\mathcal{F}$ on Z whose leaves are connected components of the inverses under p of leaves of $\mathcal{F}'$, and the group $\pi_1 Z'$ of covering transformations interchanges the leaves of $\mathcal{F}$. By Epstein [4, Corollary 1.8] Z is homeomorphic to $\mathbb{R}^2$, whence $\mathcal{F}$ satisfies (1)–(3), while $\mathcal{F}'$ may loose all of those properties. One may say that $\mathcal{F}'$ is obtained from a foliation on $\mathbb{R}^2$ into connected components of level sets of some pseudoharmonic function by some free and properly discontinuous action of $\pi_1 Z$. In general, such a function is not invariant with respect to the action of $\pi_1 Z$.

Notice also that property (2) proposes a certain classification of foliations on $\mathbb{R}^2$ by studying the way in which such a foliation is glued from open strips. This "combinatorics of gluing" would give certain topological invariants of (pseudo)harmonic functions and potentially information about symmetries of differential equations whose coefficients are harmonic functions. Moreover, one could try to extend Kaplan's technique to arbitrary pseudoharmonic functions $f : \mathbb{R}^2 \to \mathbb{R}$ just by removing the set of singular points Σ_f of f and considering the remained foliation on $\mathbb{R}^2 \setminus \Sigma_f$. Such an approach was used in the papers by W. Boothby [5, 6], M. Morse and J. Jenkins [7], M. Morse [8].

Kaplan tried to minimize the total number of leaves, and for that reason his construction was not "canonical". Namely, the choice of leaves $\{\omega_i\}_{i\in\Lambda}$ (and thus a "cutting of the foliation $\mathcal{F}$" into open strips) was not unique. Moreover, if S is a connected component of $\mathbb{R}^2 \setminus \bigcup_{i\in\Lambda} \omega_i$, then the homeomorphism $S \to \mathbb{R}\times(0,1)$ does not necessarily extend to an embedding $\overline{S} \to \mathbb{R}\times[0,1]$.

In a series of papers by S. Maksymenko, Ye. Polulyakh [9–12] and Yu. Soroka [13, 14] it was studied a class of foliations on non-compact surfaces Z (called striped) glued from open strips in a certain "canonical" way. In a joint paper [15] of the above three authors it was also described an analogue of mapping class group for foliated homeomorphisms (sending leaves into leaves) of such foliations $\mathcal{F}$ and proved that it is isomorphic with an automorphism group of a certain graph (one-dimensional CW-complex) G encoding an information about gluing a surface from strips. This graph is in a certain sense *dual* to the space of leaves of $\mathcal{F}$.

The aim of the present paper is to prove that the connection between a striped surface Z and its graph G is more deep: namely they are homotopy equivalent, see Theorems 5.2 and 5.3. One of the difficulties of proving such a result is that there is no canonical map $Z \to G$. We construct a continuous injection $\varphi : G \to Z$, and then prove (using van Kampen theorem for groupoids established by R. Brown and A. Salleh [16]) that φ induces an isomorphism of fundamental groupoids of G and Z. This will imply that φ is a homotopy equivalence, since G and Z are aspherical. For instance the ranks of their homology groups $H_1(Z,\mathbb{Z})$ and $H_1(G,\mathbb{Z})$ are the same.

In fact, the result is rather simple when Z is glued of finitely many strips: in this case the above map $\varphi : G \to Z$ is an embedding and its image $\varphi(G)$ is a strong deformation retract of Z. On the other hand, if the number of strips is infinite, G can be not a locally finite CW-complex, having thus no countable local bases at some vertices. Therefore there will be no embeddings of G into a manifold Z, thus the image $\varphi(G)$ will not be a strong deformation retract of Z. Nevertheless, van Kampen theorem allows to accomplish the result.

Actually, the obtained result has no deal with a foliation itself but only with a way in which a surface is glued from strips. Nevertheless, suppose we are given a foliation $\mathcal{F}$ on a non-compact surface Z whose leaves are non-compact closed subsets of Z. Now, if $\mathcal{F}$ has "not so much singular leaves", see Theorem 5.1 below and Fig. 4, then it is a striped surface. Therefore we get a partition into strips and our result shows that the foliation "contains" an information about the homotopy type of the underlying surface. Thus our statement could be viewed in the frame of Morse theory in which the gradient lines connecting critical points of a Morse function $f : Z \to \mathbb{R}$ (or equivalently decomposition of Z into handles in the sense of S. Smale corresponding to those critical points) determine a CW-partition of Z and this relates f with homological and even topological structure of Z.

The exposition of the paper is intended to be elementary in order to make it accessible to a large audience of readers, and thus to propagate and popularize usage of homotopy methods (like van Kampen theorem for groupoids) to more applied problems.

2 Striped Surface and Its Graph

Definition 5.1 ([9]) A subset $S \subset \mathbb{R} \times [-1, 1]$ will be called a *model strip* if:

(1) $\mathbb{R} \times (-1, 1) \subset S$;
(2) the intersection $S \cap (\mathbb{R} \times \{\pm 1\})$ is a union of at most countably many mutually disjoint open intervals.

For instance, $\mathbb{R} \times (-1, 1)$, $\mathbb{R} \times [-1, 1]$, $\big(\mathbb{R} \times (-1, 1)\big) \cup \big((-2, 3) \times \{1\}\big)$ are model strips. Of course one can replace $[-1, 1]$ with any other closed segment $[a, b] \subset \mathbb{R}$.

Notice that condition (2) is equivalent to the assumption that S is open as a subset of $\mathbb{R} \times [-1, 1]$. Define the following subsets of S:

$$\partial_- S := S \cap (\mathbb{R} \times \{-1\}), \qquad \partial_+ S := S \cap (\mathbb{R} \times \{1\}), \qquad \partial S := \partial_- S \cup \partial_+ S.$$

Definition 5.2 ([9]) Let Z be a two-dimensional manifold. A *striped atlas* on Z is a map $q : Z_0 \to Z$ having the following properties:

(1) $Z_0 = \bigsqcup_{\alpha \in A} S_\alpha$ is at most countable disjoint union of model strips;
(2) q is a quotient map, so a subset $U \subset Z$ is open iff $q^{-1}(U)$ is open in Z_0;
(3) there are two disjoint families $X = \{X_\beta\}_{\beta \in B}$, $Y = \{Y_\beta\}_{\beta \in B} \subset \bigcup_{\alpha \in A} \partial S_\alpha$ of boundary intervals such that:
    (a) q is an injective on $Z_0 \setminus (X \cup Y)$;
    (b) $q(X_\beta) = q(Y_\beta)$ for each $\beta \in B$ and the restrictions $q|_{X_\beta} : X_\beta \to q(X_\beta)$ and $q|_{Y_\beta} : Y_\beta \to q(Y_\beta)$ are embeddings;
    (c) $q(X_\beta) \cap q(X_{\beta'}) = \varnothing$ for each $\beta \neq \beta' \in B$;
    (d) $q(X_\beta) \cap q(Z_0 \setminus (X \cup Y)) = \varnothing$ for each $\beta \in B$.

Figures 1, 4 and 2 below contain examples of striped atlases.

The pair (Z, q) will also be called a *striped structure* on Z. A *striped surface* is a surface admitting a striped atlas. When talking about a striped surface Z we will also assume that some striped atlas (Z, q) is fixed. Notice that a striped surface is a non-compact two-dimensional manifold which can be non-orientable and disconnected. Moreover, each connected component of its boundary is an open interval.

2.1 Seams

Let $\beta \in B$. Then we have a homeomorphism $\gamma_\beta : Y_\beta \to X_\beta$ given by

$$\gamma_\beta = (q|_{X_\beta})^{-1} \circ q|_{Y_\beta}.$$

Therefore, a striped surface Z can also be regarded as a quotient space obtained by gluing some pairs of boundary intervals of model strips via homeomorphisms

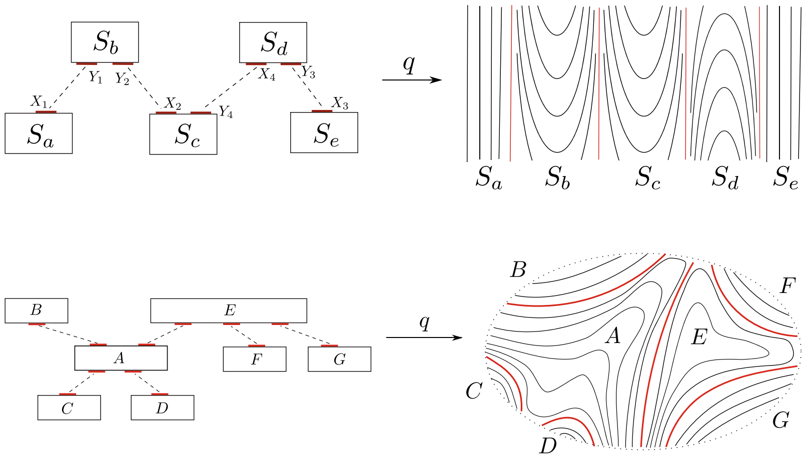

Fig. 1 Striped atlases

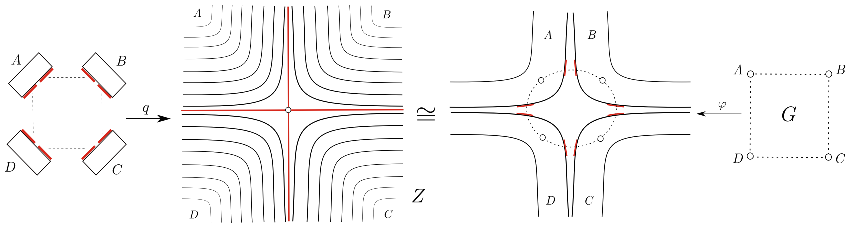

Fig. 2 Striped structure and the map φ on $\mathbb{R}^2 \setminus (0, 0)$ for canonical foliation by level sets of the function $f(x, y) = xy$

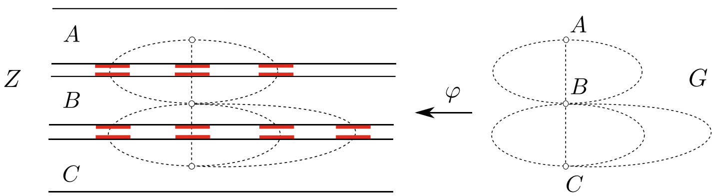

Fig. 3 Striped surface Z, its graph G, and a map $\varphi : G \to Z$

$\{\gamma_\beta\}_{\beta \in B}$. The image

$$\omega_\beta := q(X_\beta) = q(Y_\beta)$$

will be called a *seam* of Z (as well as of q). In Fig. 1 seams are colored in red color.

Since $q^{-1}(\omega_\beta) = X_\beta \cup Y_\beta$ is closed in Z_0, and q is a quotient map, it follows that each seam is a closed subset of Z.

2.2 Foliated Characterization of Striped Surfaces

Though the notion of a striped surface looks rather restrictive, it nevertheless covers a large class of surfaces equipped with foliations which agree in a certain sense with a decomposition into strips. We recall here the principal statement from [12]. It will not be used for the proof of our main result, however we present it to describe the general picture.

By a foliated surface $(Z, \mathcal{F})$ we will mean a two-dimensional manifold Z equipped with a foliation $\mathcal{F}$. A *saturation* $\mathrm{Sat}(U)$ of a subset $U \subset Z$ is the union of all leaves of $\mathcal{F}$ intersecting U, that is

$$\mathrm{Sat}(U) := \bigcup_{\gamma \cap U \neq \varnothing} \gamma.$$

A subset $U \subset Z$ is *saturated* whenever $U = \mathrm{Sat}(U)$ i.e. U is a union of leaves of $\mathcal{F}$.

Given an open subset $U \subset Z$, denote by $\mathcal{F}|_U$ the foliation on U consisting of connected components of the intersections of leaves of $\mathcal{F}$ with U. We will call $\mathcal{F}|_U$ the *restriction of* $\mathcal{F}$ to U.

A homeomorphism $h : Z \to Z'$ between foliated surfaces $(Z, \mathcal{F})$ and $(Z, \mathcal{F}')$ is *foliated* whenever for each leaf γ of $\mathcal{F}$ its image $h(\gamma)$ is a leaf of $\mathcal{F}'$.

Notice that each model strip S admits a natural foliation into boundary intervals and lines $\mathbb{R} \times \{t\}$, $t \in (-1, 1)$. We will call this foliation *canonical*. Now if (Z, q) is a striped surface, then canonical foliations on the corresponding model strips induce a foliation on all of Z which we will also call the *canonical foliation (of the striped atlas q)*.

Let $(Z, \mathcal{F})$ be a foliated surface. Say that a leaf $\gamma \subset \mathrm{Int}(Z)$, resp. $\gamma \subset \partial Z$, of $\mathcal{F}$ is *regular* if there exists an open saturated neighborhood U of γ such that the pair $(\overline{U}, U)$ is foliated homeomorphic to the model strip

$$\big(\mathbb{R} \times [-1, 1],\ \mathbb{R} \times (-1, 1)\big),$$

resp. $\big(\mathbb{R} \times [0, 1],\ \mathbb{R} \times [0, 1)\big)$, via a homeomorphism sending γ to $\mathbb{R} \times 0$.

A leaf which is not regular will be called *singular*. For example, in the above figures, the seams (red leaves) are precisely singular leaves.

For the case of striped surfaces and its canonical foliation, it is easy to see that a leaf γ is singular if and only if it satisfies one of the following conditions:

(i) $\gamma = q(\delta)$, for some boundary interval $\delta = (a, b) \times \varepsilon 1 \subset \partial_\varepsilon S_\alpha$, where $\varepsilon \in \{\pm\}$ and $\alpha \in A$;
(ii) $\partial_\varepsilon S_\alpha$ contains some other boundary intervals distinct from δ.

Theorem 5.1 ([12], Theorem 4.4) *Let $(Z, \mathcal{F})$ be a foliated surface such that every leaf γ of $\mathcal{F}$ is a non-compact closed subset of Z. Then the following conditions are equivalent:*

(1) Z admits a striped atlas q such that $\mathcal{F}$ coincides with the canonical foliation of q;
(2) the collection of all singular leaves of $\mathcal{F}$ is locally finite.

In fact, not every foliation on the plane admits a striped atlas, see [15, Examples 7.6, 7.7]. One of the reasons is that seams can converge to other seams. We will briefly recall [15, Example 7.6]. Let $\mathcal{F}$ be the foliation on $\mathbb{R}^2$ into parallel lines $\mathbb{R}\times t$, $t \in \mathbb{R}$. Let also $\{a_n\}_{n\in\mathbb{N}}\mathbb{R}$ be a strictly decreasing sequence of reals converging to some $a \in \mathbb{R}$, and $K = \{a\} \cup \{a_n\}_{n\in\mathbb{N}}$. Consider the open subset $Z = \mathbb{R}^2 \setminus (0 \times K)$ and let $\mathcal{F}|_Z$ be the restriction of $\mathcal{F}$ to Z. Notice that every point $b \in K$ splits $\mathbb{R} \times b$ into two arcs $\gamma_b^- = (-\infty, 0) \times b$ and $\gamma_b^+ = (0, +\infty) \times b$ being the leaves of $\mathcal{F}|_Z$. By property (ii) the latter leaves are singular for $\mathcal{F}|_Z$ and they converse to the leaves γ_a^- and γ_a^+. Hence the family of all singular leaves of $\mathcal{F}|_Z$ is not locally finite, and by Theorem 5.1 $(Z, \mathcal{F}|_Z)$ does not admit a striped atlas.

2.3 Graph of a Striped Surface

The above figures propose to consider a graph which encodes the gluing of strips. Such a graph was introduced in [15, Section 6] and also takes to account boundary components of striped surface which are not seams. We will consider here a more simplified version of it.

Let $q : \underset{\alpha\in A}{\sqcup} S_\alpha \to Z$ be a striped atlas on a surface Z. Then one can associate to q a one-dimensional CW-complex ("topological graph") G whose vertices are strips of the atlas and the edges correspond to gluing strips along boundary intervals. More precisely,

(0) 0-skeleton of G is $G^{(0)} = A$;
(1) Let $\beta \in B$, so we have a pair of boundary intervals $\{X_\beta, Y_\beta\}$ such that $X_\beta \subset \partial S_\alpha$ and $Y_\beta \subset \partial S_{\alpha'}$ for some *vertices* $\alpha, \alpha' \in A$. Let also $I_\beta = [-1, 1]$. Glue I_β to A via the following map:

$$\chi_\beta : \partial I_\beta = \{\pm 1\} \to A, \qquad \chi_\beta(-1) = \alpha, \qquad \chi_\beta(1) = \alpha'.$$

Then the resulting CW-complex:

$$G = \Big(\underset{\beta\in B}{\sqcup} I_\beta\Big) \bigcup_{\chi_\beta,\ \beta\in B} A.$$

will be called the *graph of the striped atlas q*. We will also denote by the same letter $\chi_\beta : I_\beta \to G$ the characteristic map of the 1-cell I_β. Thus

- $\chi_\beta|_{(-1,1)} : (-1, 1) \to G$ is an embedding and we will denote by $e_\beta = \chi_\beta\big((-1, 1)\big)$ its image being an open 1-cell;
- $\chi_\beta|_{\partial I_\beta} : \partial I_\beta \to A$ is the corresponding gluing map.

2.4 Canonical Injection $\varphi : G \to Z$

We will construct here a continuous injective map $\varphi : G \to Z$.

For each $\alpha \in A$ let $s_\alpha := (0, 0) \in \mathbb{R} \times (-1, 1) \subset S_\alpha$ be the origin. Define the map $\varphi : G^{(0)} \to Z$ by $h(\alpha) = s_\alpha$.

Furthermore, for each $\beta \in B$ fix a point $z_\beta \in \omega_\beta = q(X_\beta) = q(Y_\beta)$, and let

$$(x_\beta, \varepsilon) = q^{-1}(z_\beta) \cap X_\beta, \qquad (y_\beta, \varepsilon') = q^{-1}(z_\beta) \cap Y_\beta, \tag{1}$$

where $\varepsilon, \varepsilon' \in \{\pm 1\}$. Assume that $X_\beta \subset S_\alpha$ and $Y_\beta \subset S_{\alpha'}$ for some $\alpha, \alpha' \in A$.

Now define the following path $\varphi_\beta : I_\beta \equiv [-1, 1] \to Z$ by the following formula, see Fig. 4:

$$\varphi_\beta(t) = \begin{cases} q\big(2(1+t)x_\beta, (1+t)\varepsilon\big), & t \in [-1, -\frac{1}{2}], \\ q\big(x_\beta, (1+t)\varepsilon\big), & t \in [-\frac{1}{2}, 0], \\ q\big(y_\beta, (1-t)\varepsilon'\big), & t \in [0, \frac{1}{2}], \\ q\big(2(1-t)y_\beta, (1-t)\varepsilon'\big), & t \in [\frac{1}{2}, 1]. \end{cases}$$

It easily follows that

$$\varphi_\beta(-1) = h(\alpha) = s_\alpha, \qquad \varphi_\beta(0) = z_\beta, \qquad \varphi_\beta(1) = h(\alpha') = s_{\alpha'}.$$

Moreover, φ_β is a simple path if $\alpha \neq \alpha'$, and a simple closed path if $\alpha = \alpha'$.

Since the paths φ_β agree with φ at end-points, they induce the required map $\varphi : G \to Z$, see Fig. 3.

The aim of the present paper is to prove the following

Theorem 5.2 *Let (Z, q) be a striped surface and G be its graph. Then for each $x \in G$ the above map $\varphi : G \hookrightarrow Z$ induces an isomorphism of the fundamental groups*

$$\varphi_* : \pi_1(G, x) \to \pi_1\big(Z, \varphi(x)\big).$$

In particular, φ is a homotopy equivalence.

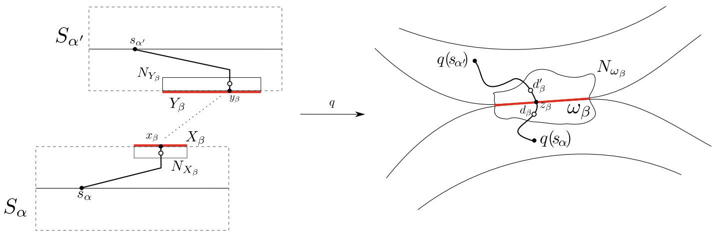

Fig. 4 The map φ_β

In fact we will establish a more precise statement

Theorem 5.3 *Let (Z, q) be a striped surface and G be its graph. Then there exists a subset $Z' \subset Z$ such that the map $\varphi : G \hookrightarrow Z$ induces an isomorphism of the corresponding fundamental groupoids*

$$\varphi^* : \Pi_1(G, \varphi^{-1}(Z')) \to \Pi_1(Z, Z').$$

The proof of these theorems will be given in Sect. 4. They are based on application of van Kampen theorem for groupoids, see Lemma 5.3.

Notice that when Z_0 has only finitely many seams, the result is rather simple and in this case $\varphi(G)$ is a strong deformation retract of Z.

On the other hand, if the number of seams is infinite, the graph G might be not a locally finite CW-complex and Z can not be deformed onto the image $\varphi(G)$.

Example 5.1 Let $Z = \mathbb{R}^2 \setminus (\mathbb{Z} \times 0)$. Then Z has an atlas consisting of two model strips:

$$S_0 = \mathbb{R} \times (-1, 1) \bigsqcup \underset{n\in\mathbb{Z}}{\cup} (n, n+1) \times \{1\},$$

$$S_1 = \mathbb{R} \times (-1, 1) \bigsqcup \underset{n\in\mathbb{Z}}{\cup} (n, n+1) \times \{-1\}.$$

and its graph has two vertices connected with countably many edges, see Fig. 5. In this case G is not locally finite at its vertices and therefore it has no countable local base, whence $\varphi : G \to Z$ is not an embedding.

Let us also mention the following statement describing certain topological properties of striped surfaces.

Lemma 5.1 *Let Z be a connected non-compact surface. Then Z has the homotopy type of an aspherical CW-complex, and $\pi_1 Z$ is free.*

Proof For the fact that every separable manifold has homotopy type of a CW-complex see [17, Corollary 5.7.].

Further, let $p : \tilde{Z} \to Z$ be the universal covering of Z. Then by Epstein [4, Corollary 1.8], the interior $\mathrm{Int}(\tilde{Z})$ is homeomorphic to $\mathbb{R}^2$. Moreover, the inclusion

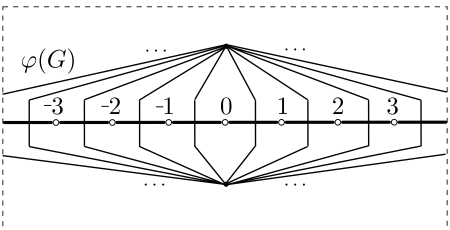

Fig. 5 Non locally finite graph G

$\mathrm{Int}(\tilde{Z}) \subset \tilde{Z}$ is a homotopy equivalence due to existence of collars of the boundary of metrizable manifolds, [18, Theorem 2]. Hence p induces isomorphisms

$$0 = \pi_k \mathrm{Int}(\tilde{Z}) \cong \pi_k \tilde{Z} \cong \pi_k Z,$$

$k \geq 2$. This, by definition, means asphericity of Z.

Finally, the statement that $\pi_1 Z$ is free, is proved in distinct sources, e.g. [19], [20, Theorem 44A], or [21, Section 4.4.2]. Another proof is that since Z is non-compact, it's cohomological dimension $\mathrm{cd}\, Z \leq 1$, whence by Swan [22, Corollary to Theorem A], $\pi_1 Z$ is free. See also [23] for discussions and [24]. □

3 Fundamental Groupoids

In this section we will briefly recall the notion of fundamental groupoid, list some of its properties and formulate van Kampen theorem for groupoids.

3.1 Small Categories

A *small category* $\mathfrak{C}$ is given by the following data:

- a set $\mathrm{Ob}(\mathfrak{C})$, called *set of objects*;
- a set $\mathrm{Hom}_{\mathfrak{C}}(X, Y)$ for each pair of objects $X, Y \in \mathrm{Ob}(\mathfrak{C})$, called *set of morphisms between X and Y*;
- for each pair of morphisms $f \in \mathrm{Hom}_{\mathfrak{C}}(X, Y)$, $g \in \mathrm{Hom}_{\mathfrak{C}}(Y, Z)$ there defined a unique morphism $g \circ f \in \mathrm{Hom}_{\mathfrak{C}}(X, Z)$, called *composition of f and g* such that the following axioms are satisfied:

– **associativity:** for any three morphisms $f \in \mathrm{Hom}_{\mathfrak{C}}(X, Y)$, $g \in \mathrm{Hom}_{\mathfrak{C}}(Y, Z)$ and $h \in \mathrm{Hom}_{\mathfrak{C}}(Z, W)$ we have

$$h \circ (g \circ f) = (h \circ g) \circ f$$

– **identity:** for each object $X \in \mathrm{Ob}(\mathfrak{C})$ there exists a morphism $\mathrm{id}_A \in \mathrm{Hom}_{\mathfrak{C}}(X, X)$, called the *identity*, such that for each $f \in \mathrm{Hom}_{\mathfrak{C}}(X, Y)$ we have

$$f \circ \mathrm{id}_X = \mathrm{id}_Y \circ f = f.$$

3.2 Functors

A *functor* $F\colon \mathfrak{C} \to \mathfrak{D}$ from category $\mathfrak{C}$ to category $\mathfrak{D}$ is the following collection of maps (usually denoted by the same letter F):

(1) a map $F : \mathrm{Ob}(\mathfrak{C}) \to \mathrm{Ob}(\mathfrak{D})$, associating to each object $X \in \mathfrak{C}$ an object $F(X) \in \mathfrak{D}$,
(2) a map $F : \mathrm{Hom}_{\mathfrak{C}}(X, Y) \to \mathrm{Hom}_{\mathfrak{D}}(F(X), F(Y))$ for every pair of objects $X, Y \in \mathfrak{C}$,

such that the following conditions are satisfied:

(a) $F(\mathrm{id}_X) = \mathrm{id}_{F(X)}$ for each $X \in \mathrm{Ob}(\mathfrak{C})$;
(b) $F(g \circ f) = F(g) \circ F(f)$ for any morphisms f, g in category $\mathfrak{C}$.

3.3 Coequalizers

Let $f, g : X \to Y$ be two morphisms. Then a *coequalizer* of f, g is a morphism $h : Y \to Z$ such that

(1) $h \circ f = h \circ g$;
(2) for any other morphism $h' : Y \to Z'$ with $h' \circ f = h' \circ g$ there exists a unique morphism $q : Z \to Z'$ such that $h' = q \circ h$.

In other words, we have the following diagram:

$$\begin{tikzcd} X \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] & Y \arrow[r, "h"] \arrow[dr, "h'"'] & Z \arrow[d, "q"] \\ & & Z' \end{tikzcd}$$

which is commutative except for paths f, g connecting X and Y.

3.4 Groupoids

A morphism $f \in \mathrm{Hom}_{\mathfrak{C}}(X, Y)$ is called an *isomorphism* whenever there exists $g \in \mathrm{Hom}_{\mathfrak{C}}(Y, X)$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$.

A *groupoid* is a category in which all morphisms are isomorphisms. Moreover, a morphism of groupoids is just a functor between the corresponding categories.

Example 5.2 For any set X the Cartesian product $X \times X$ has a natural groupoid structure.

Evidently, if $f : X \to Y$ is a bijection of sets, then the induced mapping

$$f \times f : X \times X \to Y \times Y, \qquad (f \times f)(a, b) = \big(f(a), f(b)\big)$$

is an isomorphism of groupoids.

3.5 Fundamental Groupoid

Let X be a topological space and $X' \subset X$ be a subset. Let also $I = [0, 1]$, and $C\big((I, \partial I), (X, X')\big)$ be the set of continuous paths with ends at X'. Say that two paths $\alpha, \beta \in C\big((I, \partial I), (X, X')\big)$ are equivalent ($\alpha \sim \beta$) if

1. $\alpha(0) = \beta(0), \alpha(1) = \beta(1)$;
2. α is homotopic to β relatively $\{0, 1\}$.

Then the set $\Pi_1(X, X') = C\big((I, \partial I), (X, X')\big)/\sim$ of the corresponding equivalence classes is called the *fundamental groupoid of the pair* (X, X') and the set X' is *the set of its base points*.

The equivalence class of a path $\alpha \in C\big((I, \partial I), (X, X')\big)$ will be denoted by $[\alpha]$. For a class $q \in \Pi_1(X, X')$ we will denote by $q(0)$ and $q(1)$ the common start and end-points of all representatives of q.

Then $\Pi_1(X, X')$ admits a natural "partial" operation of composition of paths turning it into groupoid. Evidently, if X' consists of a unique point, then the multiplication is defined for all elements of $\Pi_1(X, X')$, and in this case $\Pi_1(X, X')$ is the same as the fundamental group of X at the point X'.

Notice also that we have a natural map

$$r_X : \Pi_1(X, X') \to X' \times X', \qquad r_X(q) = \big(q(0), q(1)\big), \tag{2}$$

associating to each homotopy class of paths its start and end-points.

Moreover, every continuous map of pairs $f : (X, X') \to (Y, Y')$ induces the following commutative diagram:

$$\begin{array}{ccc} \Pi_1(X, X') & \xrightarrow{f^*} & \Pi_1(Y, Y') \\ {\scriptstyle r_X}\downarrow & & \downarrow{\scriptstyle r_Y} \\ X' \times X' & \xrightarrow{f \times f} & Y' \times Y'. \end{array} \tag{3}$$

The following simple lemma is left for the reader:

Lemma 5.2

(a) The map (2) *is a morphism of groupoids.*
(b) If X is path connected and simply connected, then r_X is an isomorphism.
(c) Suppose $f : (X, X') \to (Y, Y')$ is a continuous map, X and Y are path connected and simply connected, and the restriction map $f|_{X'}\colon X' \to Y'$ is a bijection. Then all morphisms in (3) *are isomorphisms.*

3.6 Coproducts

Let $\{X_\lambda\}_{\lambda\in\Lambda}$ be a collection of objects of a category $\mathfrak{C}$. Then their *coproduct* is an object $Y = \bigsqcup\limits_{\lambda\in\Lambda} X_\lambda$ together with a collection of morphisms $i_\lambda : X_\lambda \to Y$ such that for any other collection of morphisms $\{f_\lambda : X_\lambda \to Z\}_{\lambda\in\Lambda}$, to the same object Z there exists a unique morphism $f\colon Y \to Z$ such that $f_\lambda = f \circ i_\lambda$. In other words, for each $\lambda \in \Lambda$ the following diagram is commutative:

$$\begin{tikzcd} X_\lambda \arrow[r, "i_\lambda"] \arrow[dr, "f_\lambda"'] & \bigsqcup\limits_{\lambda\in\Lambda} X_\lambda \arrow[d, "f=\bigsqcup\limits_{\lambda\in\Lambda} f_\lambda"] \\ & Z \end{tikzcd}$$

Consider examples:

(a) In the category of sets a coproduct is just a disjoint union of sets.
(b) Let $\{G_\lambda\}_{\lambda\in\Lambda}$ be a collection of groupoids with multiplications

$$\mu_\lambda : G_\lambda \times G_\lambda \supset Q_\lambda \to G_\lambda,$$

and sets of objects E_λ. Then their disjoint union

$$G = \bigsqcup_{\lambda\in\Lambda} G_\lambda$$

has a natural groupoid structure in which the partial multiplication

$$\mu : G \times G \supset \bigsqcup_{\lambda\in\Lambda} Q_\lambda \to G,$$

is defined by $\mu(a, b) = \mu_\lambda(a, b)$, when $(a, b) \in Q_\lambda$. This groupoid is the coproduct in the category of groupoids.

Notice also that if $\{f_\lambda : G_\lambda \to H\}_{\lambda\in\Lambda}$ is a collection of morphisms of groupoids, then the induced morphism $f = \bigsqcup\limits_{\lambda\in\Lambda} f_\lambda : G \to H$ is given by $f(a) = f_\lambda(a)$ if $a \in G_\lambda$.

(c) Let $\{(X_\lambda, X'_\lambda)\}_{\lambda\in\Lambda}$ be a collection of pairs of topological spaces,

$$X = \bigsqcup_{\lambda\in\Lambda} X_\lambda, \qquad\qquad X' = \bigsqcup_{\lambda\in\Lambda} X'_\lambda$$

be the corresponding coproducts of sets. For each $\lambda \in \Lambda$ the natural inclusion of pairs $j_\lambda : (X_\lambda, X'_\lambda) \subset (X, X')$ induces the corresponding morphism of groupoids

$$\Pi_1(j_\lambda) : \Pi_1(X_\lambda, X'_\lambda) \to \Pi_1(X, X').$$

Hence, by coproduct property, we have a unique morphism ζ making the following diagram commutative:

$$\begin{tikzcd} \Pi_1(X_\lambda, X'_\lambda) \arrow[r, hook, "i_\lambda"] \arrow[dr, "\Pi_1(j_\lambda)"'] & \bigsqcup_{\lambda\in\Lambda} \Pi_1(X_\lambda, X'_\lambda) \arrow[d, "\zeta=\bigsqcup_{\lambda\in\Lambda}\Pi_1(j_\lambda)"] \\ & \Pi_1(X, X') \end{tikzcd} \tag{4}$$

where i_λ is a natural morphism into coproduct. We claim that ζ is an isomorphism, i.e. it is bijective.

Let us show that ζ is *surjective*. Indeed, let $\alpha : (I, \partial I) \to (X, X')$ be a continuous map. We should show that $[\alpha] = \zeta(q)$ for some $q \in \bigsqcup_{\lambda\in\Lambda} \Pi_1(X_\lambda, X'_\lambda)$. Since I is path-connected and $X_\lambda \cap X_\nu = \varnothing$ for $\lambda \neq \nu$, it follows that $\alpha(I)$ is contained in some X_λ. Hence $\alpha = j_\lambda \circ \beta$ for some path $\beta : (I, \partial I) \to (X_\lambda, X'_\lambda)$. Therefore, $[\alpha] = \Pi_1(j_\lambda)\big([\beta]\big) = \zeta \circ i_\lambda\big([\beta]\big) = \zeta\big([i_\lambda \circ \beta]\big)$.

Let us prove that ζ is *injective*. Let $p, q \in \bigsqcup_{\lambda\in\Lambda} \Pi_1(X_\lambda, X'_\lambda)$ be two classes such that $\zeta(p) = \zeta(q)$. By definition, $p = i_\lambda([\alpha])$ and $q = i_\nu([\beta])$ for some paths

$$\alpha : (I, \partial I) \to (X_\lambda, X'_\lambda), \qquad \beta : (I, \partial I) \to (X_\nu, X'_\nu).$$

Then

$$\Pi_1(j_\lambda)\big([\alpha]\big) = \zeta\big(i_\lambda([\alpha])\big) = \zeta(p) = \zeta(q) = \zeta\big(i_\nu([\beta])\big) = \Pi_1(j_\nu)\big([\beta]\big)$$

which means that the paths $j_\lambda \circ \alpha, j_\lambda \circ \beta : (I, \partial I) \to (X, X')$ are homotopic relatively their ends. Hence $\lambda = \nu$. Moreover, let $H : [0, 1] \times [0, 1] \to X$ be the corresponding homotopy between $H_0 = \alpha$ and $H_1 = \beta$. Since $[0, 1] \times [0, 1]$ is path-connected, it follows that $H([0, 1] \times [0, 1]) \subset X_\lambda$, and thus H is a homotopy between α and β in X_λ relatively their ends. This means that $\lambda = \nu$, $[\alpha] = [\beta] \in \Pi_1(X_\lambda, X'_\lambda)$, whence $p = i_\lambda([\alpha]) = i_\nu([\beta]) = q$. Thus ζ is injective.

3.7 van Kampen Theorem for Groupoids

Let Λ be a set. It will be convenient to consider a category $\mathfrak{C}$ whose objects are triples $\big(X, X', \mathcal{U}\big)$ where

1. X is a topological space;
2. $X' \subset X$ a subset;
3. $\mathcal{U} = \{U_\lambda\}_{\lambda\in\Lambda}$ is an open cover of X enumerated by the same set of indices Λ;

and morphisms between triples $\big(X, X', \mathcal{U} = \{U_\lambda\}_{\lambda\in\Lambda}\big)$ and $\big(Y, Y', \mathcal{V} = \{V_\lambda\}_{\lambda\in\Lambda}\big) \in \mathfrak{C}$ are continuous maps of pairs $f\colon (X, X') \to (Y, Y')$ such that $f\big(V_\lambda\big) \subset U_\lambda$ for all $\lambda \in \Lambda$.

Given a triple $\big(X, X', \mathcal{U}\big) \in \mathfrak{C}$, for each n-tuple $\nu = (\nu_1, \dots, \nu_n) \in \Lambda^n$ put

$$U_\nu := U_{\nu_1} \cap \cdots \cap U^{\nu_n}, \qquad\qquad U'_\nu := U_\nu \cap X'.$$

In particular, for $n = 2$ and $\nu = (j, k) \in A^2$ we have the following two morphisms of fundamental groupoids:

$$a_{jk} : \Pi_1(U_j \cap U_k, U'_{(j,k)}) \to \Pi_1(U_j, U'_j),$$
$$b_{jk} : \Pi_1(U_j \cap U_k, U'_{(j,k)}) \to \Pi_1(U_k, U'_k)$$

induced by natural inclusions of $U_j \cap U^k$ into U^j and U^k respectively.

These morphisms yield morphisms of the corresponding coproducts:

$$a, b : \bigsqcup_{(j,k)\in\Lambda^2} \Pi_1(U_{(j,k)}, U'_{(j,k)}) \to \bigsqcup_{\lambda\in\Lambda} \Pi_1(U_\lambda, U'_\lambda)$$

Similarly, the inclusion $(U_j, U'_j) \subset (X, X')$ induces a morphism

$$c_j : \Pi_1(U_j, U'_j) \to \Pi_1(X, X').$$

Then the Π_1-*diagram of the cover* $\mathcal{U}$ is the following diagram:

$$\bigsqcup_{(j,k)\in\Lambda^2} \Pi_1(U_{(j,k)}, U'_{(j,k)}) \underset{b}{\overset{a}{\rightrightarrows}} \bigsqcup_{\lambda\in\Lambda} \Pi_1(U_\lambda, U'_\lambda) \xrightarrow{\ c\ } \Pi_1(X, X'). \tag{5}$$

Theorem 5.4 (van Kampen Theorem for Fundamental Groupoids, [16]) *Suppose that an open cover $\mathcal{U} = \{U_\lambda\}_{\lambda\in\Lambda}$ of X has the following property:*

(∗) *a subset $X' \subset X$ meets each path-component of each non-empty twofold and threefold intersection of distinct sets of $\mathcal{U}$.*

Then in the Π_1-diagram (5) of the cover $\mathcal{U}$ the morphism c is a coequalizer for morphisms a, b in the category of fundamental groupoids.

3.8 Π_1-Diagram for Covers by Simply Connected Sets

Let

$$f : \big(X, X', \{U_\lambda\}_{\lambda\in\Lambda}\big) \to \big(Y, Y', \{V_\lambda\}_{\lambda\in\Lambda}\big)$$

be a morphism in the category $\mathfrak{C}$. Then it is evident that f induces the following diagram:

$$\begin{tikzcd}
\bigsqcup\limits_{(j,k)\in\Lambda^2} (U_{(j,k)}, U'_{(j,k)}) \arrow[r, shift left, "a"] \arrow[r, shift right, "b"'] \arrow[d, "f_2"] & \bigsqcup\limits_{\lambda\in\Lambda} (U_\lambda, U'_\lambda) \arrow[r, "c"] \arrow[d, "f_1"] & (X, X') \arrow[d, "f"] \\
\bigsqcup\limits_{(j,k)\in\Lambda^2} (V_{(j,k)}, V'_{(j,k)}) \arrow[r, shift left, "a'"] \arrow[r, shift right, "b'"'] & \bigsqcup\limits_{\lambda\in\Lambda} (V_\lambda, V'_\lambda) \arrow[r, "c'"] & (Y, Y')
\end{tikzcd} \tag{6}$$

where $f_1 = \bigsqcup\limits_{\lambda\in\Lambda} f|_{V_\lambda}$ and $f_2 = \bigsqcup\limits_{(j,k)\in\Lambda^2} f|_{V_{(j,k)}}$. This diagram is commutative except for the left square in which we have only the following identities:

$$f_1 \circ a = a' \circ f_2, \qquad\qquad f_1 \circ b = b' \circ f_2. \tag{7}$$

Since Π_1 is a functor from the category of sets to the category of groupoids, (6) implies the following diagram combined from two Π_1-diagrams:

$$\begin{tikzcd}
\bigsqcup\limits_{(j,k)\in\Lambda^2} \Pi_1(U_{(j,k)}, U'_{(j,k)}) \arrow[r, shift left, "a"] \arrow[r, shift right, "b"'] \arrow[d, "f_2"] & \bigsqcup\limits_{\lambda\in\Lambda} \Pi_1(U_\lambda, U'_\lambda) \arrow[r, "c"] \arrow[d, shift left, "f_1"] & \Pi_1(X, X') \arrow[d, shift left, "f"] \\
\bigsqcup\limits_{(j,k)\in\Lambda^2} \Pi_1(V_{(j,k)}, V'_{(j,k)}) \arrow[r, shift left, "a'"] \arrow[r, shift right, "b'"'] & \bigsqcup\limits_{\lambda\in\Lambda} \Pi_1(V_\lambda, V'_\lambda) \arrow[r, "c'"] \arrow[u, shift left, "g_1"] \arrow[ur, "k"] & \Pi_1(Y, Y') \arrow[u, shift left, "g"]
\end{tikzcd} \tag{8}$$

Here we used the isomorphism of groupoids (4). Notice that diagram (8) is also commutative except for left square in which only the identities (7) hold.

The following lemma plays a key role in proving the basic theorem.

Lemma 5.3 *Let $f : \big(X, X', \{U_\lambda\}_{\lambda\in\Lambda}\big) \to \big(Y, Y', \{V_\lambda\}_{\lambda\in\Lambda}\big)$ be a morphism in the category $\mathfrak{C}$. Suppose that*

(1) for each $\lambda \in \Lambda$ spaces U_λ and V_λ are simply connected;
(2) for each $\lambda \in \Lambda$ the restriction $f|_{U'_\lambda} : U'_\lambda \equiv U_\lambda \cap X' \to V'_\lambda \equiv V_\lambda \cap Y'$ is a bijection;

(3) the open covers $\{U_\lambda\}_{\lambda\in\Lambda}$ and $\{V_\lambda\}_{\lambda\in\Lambda}$ satisfy condition (∗) of van Kampen Theorem 5.4, that is X' (resp. Y') meets each path component of each non-empty twofold and threefold intersections of elements of $\mathcal{U}$ (resp. $\mathcal{V}$).

Then the induced morphism $f^* : \Pi_1(X, X') \to \Pi_1(Y, Y')$ of fundamental groupoids from diagram (8) *is an isomorphism.*

Proof Conditions (1) and (2) mean that for each $\lambda \in \Lambda$ the restriction map

$$f|_{U_\lambda} : (U_\lambda, U'_\lambda) \to (V_\lambda, V'_\lambda)$$

satisfies condition (c) of Lemma 5.2. Therefore the induced morphism of groupoids $f_\lambda : \Pi_1(U_\lambda, U'_\lambda) \to \Pi_1(V_\lambda, V'_\lambda)$ is an isomorphism. Hence so is the middle vertical morphism f_1 of diagram (8).

Let $g_1 : \bigsqcup_{\lambda\in\Lambda} \Pi_1(V_\lambda, V'_\lambda) \to \bigsqcup_{\lambda\in\Lambda} \Pi_1(U_\lambda, U'_\lambda)$ be the isomorphism inverse to f_1. Denote $k = c \circ g_1$, see (9):

$$\begin{tikzcd}
\bigsqcup\limits_{(j,k)\in\Lambda^2} \Pi_1(U_{(j,k)}, U'_{(j,k)}) \arrow[r, shift left, "a"] \arrow[r, shift right, "b"'] \arrow[d, "f_2"] & \bigsqcup\limits_{\lambda\in\Lambda} \Pi_1(U_\lambda, U'_\lambda) \arrow[r, "c"] \arrow[d, shift left, "f_1"] & \Pi_1(X, X') \arrow[d, shift left, "f"] \\
\bigsqcup\limits_{(j,k)\in\Lambda^2} \Pi_1(V_{(j,k)}, V'_{(j,k)}) \arrow[r, shift left, "a'"] \arrow[r, shift right, "b'"'] & \bigsqcup\limits_{\lambda\in\Lambda} \Pi_1(V_\lambda, V'_\lambda) \arrow[r, "c'"] \arrow[u, shift left, "g_1"] \arrow[ur, "k"] & \Pi_1(Y, Y') \arrow[u, shift left, "g"]
\end{tikzcd} \tag{9}$$

Then $k \circ a' = k \circ b'$, whence by van Kampen Theorem 5.4 there exists a *unique* morphism $g : \Pi_1(Y, Y') \to \Pi_1(X, X')$ such that $k = g \circ c'$. Since f_1 and g_1 are inverse each to other, it follows that f and g must also be inverse each to other and therefore f is an isomorphism. □

4 Proof of Theorem 5.3

Let (Z, q) be a striped surface, G its graph, and $\varphi : G \to Z$ the continuous injective map defined in Sect. 2.4. We should construct a subset $Z' \subset Z$ such that the map $\varphi : G \hookrightarrow Z$ induces an isomorphism of the corresponding fundamental groupoids $\varphi_* : \Pi_1(G, \varphi^{-1}(Z')) \to \Pi_1(Z, Z')$.

In fact we will also define a special open cover $\mathcal{V}$ of Z, and then consider a cover $\mathcal{U} = \varphi^{-1}(\mathcal{V})$ of G consisting of inverse of elements of $\mathcal{V}$. Then φ will evidently induce a morphism of triples $\varphi : (G, G', \mathcal{U}) \to (Z, Z', \mathcal{V})$ in the category $\mathfrak{C}$, and we will show that conditions (1)–(3) of Lemma 5.3 are satisfied. This will imply that φ_* is an isomorphism.

An Open Cover $\mathcal{V}$ of Z

Let S be a model strip and $X = (a, b) \times \varepsilon \subset \partial S$, $\varepsilon \in \{\pm 1\}$, be a boundary interval. Then the following *open* neighborhood of X in S:

$$N_X = \begin{cases} (a, b) \times (0.8, 1], & \varepsilon = 1, \\ (a, b) \times [-1, 0.8), & \varepsilon = -1. \end{cases}$$

will be called the *standard* neighborhood of X.

It is evident, that standard neighborhoods of boundary intervals are mutually disjoint. Hence the family of boundary intervals of S is *discrete*[1] and in particular locally finite. Therefore, the union of any number of boundary intervals is a closed set.

By a *standard* neighborhood N_{ω_β} of a seam ω_β we will mean the union of images of standard neighborhoods of X_β and Y_β (see Fig. 4):

$$N_{\omega_\beta} = q(N_{X_\beta}) \cup q(N_{Y_\beta}).$$

Then N_{ω_β} is open in Z. Moreover, $N_{\omega_\beta} \cap N_{\omega_{\beta'}} = \varnothing$ for $\beta \neq \beta'$, whence the family of seams is discrete and therefore locally finite. In particular, the union of any collections of seams is also closed.

Furthermore, for $\alpha \in A$ put

$$N_{S_\alpha} := q\big(S_\alpha \setminus (X \cup Y)\big).$$

Thus N_{S_α} is image of a model strip S_α without boundary intervals corresponding to seams. It follows that N_{S_α} is open in Z. Hence we get the following open cover of Z.

$$\mathcal{V} := \{N_{S_\alpha}\}_{\alpha \in A} \ \cup \ \{N_{\omega_\beta}\}_{\beta \in B}.$$

The Set Z'

Notice that $N_{\omega_\beta} \setminus \omega_\beta$ consists of two connected components each homeomorphic to an open rectangle and the image $\varphi(G)$ intersects each of those components. Choose any two points

$$d_\beta \in \varphi(G) \cap \big(q(N_{X_\beta}) \setminus \omega_\beta\big), \qquad d'_\beta \in \varphi(G) \cap \big(q(N_{Y_\beta}) \setminus \omega_\beta\big).$$

belonging to distinct components of $N_{\omega_\beta} \setminus \omega_\beta$. In Fig. 4 such points are denoted by circles.

[1] Recall that a collections of subsets $\{Q_i\}_{i \in \Lambda}$ of a topological space X is called *discrete*, if for each $i \in \Lambda$ there exists an open neighborhood U_i of Q_i such that $U_i \cap U_j = \varnothing$ for $i \neq j \in \Lambda$.

Let J' be the set of isolated vertices of G. Then every strip S_α with $\alpha \in J$ is not glued to any other strips, and no boundary intervals of S are glued together. Let also $K' = \{s_\alpha = (0,0) = \varphi(\alpha) \in S_\alpha \mid \alpha \in J'\} \subset Z$ be the set of origins of such strips in Z. Put

$$Z' = K' \cup \{d_\beta, d'_\beta \mid \beta \in B\},$$

and

$$G' = \varphi^{-1}(Z'), \qquad L_{s_\alpha} := \varphi^{-1}(N_{S_\alpha}), \qquad L_\beta := \varphi^{-1}(N_{\omega_\beta}).$$

Thus L_β is an open arc in some 1-cell of G containing two points $\varphi^{-1}(d_\beta)$ and $\varphi^{-1}(d'_\beta)$, while L_{s_α} is a "star"-neighborhood of the vertex $\alpha \in G^0 = A$ such that each edge of L_{s_α} contains a unique point $\varphi^{-1}(d_\beta)$ or $\varphi^{-1}(d'_\beta)$ for some $\beta \in B$.

It follows that

$$\mathcal{U} := \{L_{s_\alpha}\}_{\alpha \in A} \cup \{L_\beta\}_{\beta \in B}$$

is an open cover of G, and $\varphi : (G, G', \mathcal{U}) \to (Z, Z', \mathcal{V})$ induces a morphism in the category $\mathfrak{C}$.

Verification of Conditions of Lemma 5.3

(1) Evidently, the elements of $\mathcal{U}$ and $\mathcal{V}$ are even contractible and therefore they are simply connected.

(2) Since φ is injective, it follows that for any subset $Q \subset \varphi(G) \subset Z$, we have that $\varphi|_{\varphi^{-1}(Q)} : \varphi^{-1}(Q) \to Q$ is a bijection. In particular, so are the restrictions

$$\varphi : G' \cap L_{s_\alpha} \to Z' \cap N_{S_\alpha}, \qquad \varphi : G' \cap L_\beta \to Z' \cap N_{\omega_\beta}$$

for each $\alpha \in A$ and $\beta \in B$.

(3) First notice that Z' (resp. G') meets every path component of Z (resp. G). Furthermore, $N_{S_\alpha} \cap N_{S_{\alpha'}} = N_{\omega_\beta} \cap N_{\omega_{\beta'}} = L_{s_\alpha} \cap L_{s_{\alpha'}} = L_\beta \cap L_{\beta'} = \varnothing$ for $\alpha \neq \alpha' \in A$ and $\beta \neq \beta' \in B$, which implies that all threefold intersections of elements of $\mathcal{V}$ and $\mathcal{U}$ are empty.

Also, $N_{S_\alpha} \cap N_{\omega_\beta} \neq \varnothing$ iff either X_β or Y_β or both of them are contained in ∂S_α. In this case each connected component of $N_{S_\alpha} \cap N_{\omega_\beta}$ contains either d_β or d'_β. It follows that each connected component of $L_{s_\alpha} \cap L_\beta$ contains either $\varphi^{-1}(d_\beta)$ or $\varphi^{-1}(d'_\beta)$. Thus Z' (resp. G') meets all path components of all twofold intersections of elements of $\mathcal{V}$ (resp. $\mathcal{U}$). Hence all conditions of Lemma 5.3 holds, whence φ_* is an isomorphism of groupoids. □

5 Proof of Theorem 5.2

By Theorem 5.3, φ yields an isomorphism of fundamental groupoids. In particular, if $x \in G'$, then φ also induces an isomorphism of the fundamental groups $\pi_1(G, x) \to \pi_1(Z, \varphi(x))$. Since Z' (resp. G') meets every path component of Z (resp. G), it follows that φ induces isomorphism of fundamental groups are each point $x \in G$.

Notice that every connected component of G is covered by at most countable tree, and therefore G is aspherical. Moreover, Z is also aspherical by Lemma 5.1. Hence φ induces isomorphisms between all the corresponding homotopy groups of G and Z. Moreover, since Z has the homotopy type of an infinite CW-complex, see Lemma 5.1, we have from the Whitehead theorem that φ is a homotopy equivalence between all corresponding connected components of G and Z. □

References

1. Morse, M.: Topological Methods in the Theory of Functions of a Complex Variable. Princeton University Press (1947)
2. Kaplan, W.: Regular curve-families filling the plane, I. Duke Math. J. **7**, 154–185 (1940)
3. Kaplan, W.: Regular curve-families filling the plane, II. Duke Math. J. **8**, 11–46 (1941). http://projecteuclid.org/euclid.dmj/1077492492
4. Epstein, D.: Curves on 2-manifolds and isotopies. Acta Math. **115**, 83–107 (1966). https://doi.org/10.1007/BF02392203
5. Boothby, W.: The topology of regular curve families with multiple saddle points. Amer. J. Math. **73**, 405–438 (1951)
6. Boothby, W.: The topology of the level curves of harmonic functions with critical points. Amer. J. Math. **73**, 512–538 (1951)
7. Jenkins, J., Morse, M.: Contour equivalent pseudoharmonic functions and pseudoconjugates. Amer. J. Math.. **74**, 23–51 (1952)
8. Morse, M.: The existence of pseudoconjugates on Riemann surfaces. Fund. Math. **39**, 269–287 (1952)
9. Maksymenko, S., Polulyakh, E.: Foliations with non-compact leaves on surfaces. Proc. Geom. Center **8**, 17–30 (2015)
10. Maksymenko, S., Polulyakh, E.: One-dimensional foliations on topological manifolds. Proc. Int. Geom. Cent. **9**, 1–23 (2016)
11. Maksymenko, S., Polulyakh, E.: Foliations with all non-closed leaves on non-compact surfaces. Methods Funct. Anal. Topol. **22**, 266–282 (2016). https://doi.org/10.1080/03155986.1984.11731927
12. Maksymenko, S., Polulyakh, E.: Characterization of striped surfaces. Proc. Int. Geom. Cent. **10**, 24–38 (2017)
13. Soroka, Y.: Homeotopy groups of rooted tree like non-singular foliations on the plane. Methods Funct. Anal. Topology **22**, 283–294 (2016)
14. Soroka, Y.: Homeotopy groups of nonsingular foliations of the plane. Ukr. Math. Journ.. **69**, 1000–1008 (2017). https://doi.org/10.1007/s11253-017-1423-6
15. Maksymenko, S., Polulyakh, E., Soroka, Y.: Homeotopy groups of one-dimensional foliations on surfaces. Proc. Int. Geom. Cent. **10**, 22–46 (2017)
16. Brown, R., Salleh, A.: A van Kampen theorem for unions on nonconnected spaces. Arch. Math. (Basel). **42**, 85–88 (1984). https://doi.org/10.1007/BF01198133

17. Lundell, A., Weingram, S.: The topology of CW complexes. Van Nostrand Reinhold Co. (1969). https://doi.org/10.1007/978-1-4684-6254-8
18. Brown, M.: Locally flat imbeddings of topological manifolds. Ann. Math. (2) **75**, 331–341 (1962). https://doi.org/10.2307/1970177
19. Johansson, I.: Topologische Untersuchungen über unverzweigte Überlagerungsflächen. Skr. Norske Vid.-Akad., Oslo, Math.-Naturv. Kl. , pp. 1–69 (1931)
20. Ahlfors, L., Sario, L.: Riemann Surfaces. Princeton University Press (1960)
21. Stillwell, J.: Classical Topology and Combinatorial Group Theory. Springer, New York (1993). https://doi.org/10.1007/978-1-4612-4372-4
22. Swan, R.: Groups of cohomological dimension one. J. Algebra **12**, 585–610 (1969). https://doi.org/10.1016/0021-8693(69)90030-1
23. Putman, A.: Fundamental groups of noncompact surfaces (MathOverflow). https://mathoverflow.net/q/18454 (version: 2015-10-27)
24. Putman, A.: Spines of manifolds and the freeness of fundamental groups of noncompact surfaces. https://www3.nd.edu/~andyp/notes/NoncompactSurfaceFree.pdf

A Geometric Framework to Compare PDEs and Classical Field Theories

Lukas Silvester Barth

Abstract In this contribution, a mathematical framework is constructed to relate and compare non-linear partial differential equations (PDEs) in the category of smooth manifolds. In particular, it can be used to compare those aspects of field theories (e.g. of classical (Newtonian) mechanics, hydrodynamics, electrodynamics, relativity theory, classical Yang-Mills theory and so on) that are described by such equations.

Employing a geometric (jet space) approach, a suitable notion of shared structure of two systems of PDEs is identified. It is proven that this shared structure can serve to transfer solutions from one theory to another and a generalization of so-called Bäcklund transformations is derived that can be used to generate non-trivial solutions of some non-linear PDEs.

A procedure (based on formal integrability) is introduced with which one can explicitly compute the minimal consistency conditions that two systems of PDEs need to fulfill in order to share structure under a given correspondence. Furthermore, it is shown how symmetry groups can be used to identify useful correspondences and structure that is shared up to symmetries. Thereby, the role that Bäcklund transformations play in the theory of quotient equations is clarified.

Explicit examples illustrate the general ideas throughout the text and in the last chapter, the framework is applied to systems related to electrodynamics and hydrodynamics.

1 Introduction

Studying relationships of different theories can serve to identify their underlying central features. Once shared structure of two theories is known, methods for solving

---

L. S. Barth (✉)
Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany
e-mail: Lukas.Barth@mis.mpg.de

M. Ulan, S. Hronek (eds.), *Groups, Invariants, Integrals, and Mathematical Physics*, Tutorials, Schools, and Workshops in the Mathematical Sciences,
https://doi.org/10.1007/978-3-031-25666-0_6

a problem in one domain can be transferred to another. In the long run, a structured overview could set free innovation for the development of these theories.

1.1 Previous Attempts to Compare Theories

In the physics literature, comparisons were usually restricted to analogies of two specific theories established by juxtaposition of the corresponding equations of motion. For instance, [1] introduced new effective quantities to rewrite the Navier-Stokes equations in a form very similar to Maxwell's equations. [2] established an analogy between general relativity and electrodynamics by showing that a certain linear combination of derivatives of the Faraday tensor has an irreducible representation with 16 components, 10 of which can be associated with the 10 components of the Weyl tensor of general relativity. Visser [3] explained the analogy of mathematical aspects of the description of black holes and supersonic flows which resulted in research of so-called analogue experiments (cf. [4]). All those analogies are however rather specific and a general framework for comparisons is missing.

In the philosophy of science literature, some more abstract, category theoretical approaches are outlined. Weatherall uses groupoids (categories in which all morphisms are isomorphisms) to compare theories that differ in their formulation but describe the same physics (cf. [5, 6]). More specifically, the objects in those groupoids are the formal solutions of systems of PDEs and the morphisms are symmetries of the underlying spacetime that preserve those solutions. Weatherall then defines an equivalence of two such theories as a categorical equivalence between their corresponding groupoids that preserves the empirical content of the physical theories. This idea was subsequently used by others to compare formulations of other theories, e.g. [7] compare the geometric and algebraic formulation of general relativity and [8] compares the Lagrangian and Hamiltonian formulation of classical mechanics.

The problem of this approach is however that categorical equivalence can only serve to render equivalent formulations that differ up to invertible (symmetry) transformations but is not capable of providing a framework to compare entirely different theories, to identify their intersection or subtheories. And it does not provide any means for understanding which solutions can be transferred from one theory to another.

Comparisons between the Hamiltonian and the Lagrangian view of mechanics are also discussed in the mathematics literature, see e.g. [9] or [10]. But again, such discussions are not aimed at the formalization of a general framework for the comparison of theories. The most general discussion of relationships between systems of differential equations, known to me, involves the powerful concept of so-called coverings in the category of diffieties (cf. [11–13]). However, coverings were constructed to investigate generalized, nonlocal symmetries of PDEs and are not designed for the comparison of arbitrary systems of equations. Furthermore, since they are defined over infinitely prolonged differential equations, they can not serve to find integrability conditions (which requires the inclusion of methods of formal integrability at the level of finitely prolonged equations) that arise upon the comparison of different theories.

Apart from these mathematical approaches, there is also literature that discusses the differences and transitions of physical theories heuristically. For instance, the ideas regarding the structure of scientific progress developed by Kuhn [14] are well-known. Kuhn describes progress in a recurring loop of eras with three stages which might roughly be described as follows: Confusion about how to describe a process in nature, determination of a unifying model and finally application of this model—until new ideas and experiments lead to another stage of confusion.

Another example of a heuristic discussion of the conceptual structure of physical theories is provided by Stamatescu et al. [15]. He takes into account the role of the symbols that we use for the description of physics and emphasises as a guideline the so-called Hertzian principle (cf. [16]). According this principle, the concordance of reality and symbolic description must be such that any consequences of an initial experimental setup due to the laws of nature must correspond to thought consequences of the symbols that describe this initial setup due to the laws of the mathematical formalism. Stamatescu also discusses the transition of theories and the development of their concepts. The problem with more heuristic discussions is that they are very hard to formalize. Indeed, the geometric framework presented here can not account for transitions of physical theories. However, it might be extended in the future to do so as suggested in the outlook in Sect. 10.2.

Finally, note that the present work builds on research of the author's Master's thesis but contains several generalizations. For example, within the thesis, the notion of a correspondence between theories was defined as a differential operator, which, in contrast to the correspondence on the natural product bundle introduced below, did not allow for implicit comparisons of systems of equations. Moreover, the present approach is conceptually cleaner because the two compared theories determine the natural space in which the intersection takes place before the correspondence is imposed and both, the compared theories as well as the correspondence are all treated as geometric spaces. Most importantly, the approach in the thesis did not allow to generalize Bäcklund transformations whose inclusion allows for a much more powerful transfer of solutions.

1.2 Requirements for the Framework

A classical field theory is here understood as a system of partial differential equations (PDEs) on some manifold (possibly called spacetime), together with a physical interpretation. This physical interpretation specifies

- how the mathematical quantities are related to experimental measurements,
- which initial/boundary conditions are physically plausible

- and strictly speaking should also include validity bounds for the mathematical formalism.[1]

In this article however, only the PDEs themselves are compared without considering their interpretation because of the aspiration to identify common causes. It is desired to understand which models are structurally similar even if they can be associated with different experimental setups because exactly this abstraction facilitates to obtain a new intuition for the phenomena described by the equations and to transfer methods. If desired, it is always possible to impose an interpretation later to discriminate theories further.[2]

The previous subsection illustrates that there are many different aspects of classical field theories that can be compared. Some approaches focus on symmetries, orbit spaces and conservation laws, some on the dynamics, others on structural similarities or on the solution spaces. However, if the underlying systems of PDEs of the field theories are equivalent, then all of those aspects are equivalent as well. At the same time, each single aspect can also be studied at the level of the PDEs. As a conclusion, a very wholesome approach to the comparison of the mathematical structures of field theories consists in the formulation of a framework that compares PDEs.

Such a framework then should be able to answer the following questions in a mathematically precise way.

(Q.1) Are two systems of PDEs equivalent?
(Q.2) Do two PDEs share any subsystem?
(Q.3) When are two systems equivalent up to a symmetry?
(Q.4) How to transfer solutions from one system to another?

It is important for the framework to provide an answer to the last question because it requires a degree of formalization that exceeds a purely heuristic comparison and because the determination of the space of common solutions is arguably one of the best measures for the similarity of two theories.

1.3 Methods

To summarise the above, the aim of this article is to compare field theories by comparing their PDEs, preferably in a well-defined category. In order to do that, one needs to define what two systems of PDEs have in common but there is usually no canonical way to define this common part. However, if one could comprehend a

[1] For example, classical mechanics is only valid on certain scales, only produces predictions within acceptable errors up to certain velocities and so on.

[2] To take into account the validity bounds that go along with an interpretation would require a lot of work, both because those bounds are not always clearly defined and because one might have to add inequalities that restrict the range of the variables.

PDE as a geometric object, then the common part could be naturally identified as the intersection of those objects in a suitable space. Fortunately, the language of *jet spaces*, in which PDEs are understood as submanifolds, allows for such an approach which is the main reason that the present framework is formulated in this language. Another important reason is that it also allows for the implementation of methods from the area of *formal integrability* that serve to calculate the minimal consistency conditions that arise when comparing two systems of PDEs.

Jet spaces arose with Cartan's concept of a prolongation and were defined by Ehresmann in 1953. Their theory steadily evolved, giving rise to the theories of formal integrability (cf. [17, 18]), involution (cf. [19]), differential Galois theory (cf. [20]), to the invention of so-called diffieties which generalise algebraic varieties (cf. [21]), and a whole new calculus called secondary calculus (cf. [22, 23]). Furthermore, they were used for the study of (variational) boundary value problems (cf. [24, 25], [26]), control theory (here the algebraic reformulation is particularly useful, cf. [27, 28]), the application of (co)homological methods and moving frames to PDEs (cf. [29–31], [32]), and especially to investigate local and nonlocal symmetries (cf. [33], [13], [34]), their invariants and quotients (cf. [35], [36]).

The present framework is restricted to jet spaces in the category of smooth manifolds, i.e. PDEs are assumed to be smooth submanifolds. However, this does not imply that their solutions are necessarily smooth or that the framework can only compare the spaces of smooth solutions of two systems of PDEs. Instead, this smooth category is a convenient setting to study certain singular solutions as well, like e.g. shock waves, whose singularity vanishes on higher order jet spaces (cf. [37], [26]). However, distributional solutions are indeed excluded in the present framework.

1.4 Outline

Given two systems of PDEs, each represented as a submanifold of a jet space, it becomes possible to define another submanifold (subject to some conditions), called *correspondence*, in the fibered product of those jet spaces, that connects the two systems and gives rise to a meaningful notion of an *intersection*. However, this intersection only has solutions if certain *integrability conditions* are fulfilled. Those conditions can in turn be calculated with methods of formal integrability. Once the two systems are shown to be compatible, their intersection is called *shared structure*. Solutions of this shared structure can be shown to be solutions of both intersected theories. Thus, (Q.2) can be answered because the shared structure corresponds to a subsystem with shared solutions.[3] It also allows the transfer of solutions and is subsequently shown to naturally include so-called Bäcklund-transformations that

[3] Moreover, since formal integrability serves to calculate the *minimal integrability conditions*, it is the largest possible subsystem given a chosen correspondence.

can serve to generate non-trivial solutions of some non-linear PDEs. Thus, (Q.4) can be answered. Also (Q.1) is answered by defining two subsystems to be equivalent if their shared structure possesses all solutions of both theories. Finally, (Q.3) can be answered by defining two systems to be equivalent up to symmetry if one of the systems possesses a symmetry group such that its corresponding quotient system is equivalent to the other system.

2 Notation and Preliminaries

The aim of this section is to fix the notation and to introduce subsequently necessary notions. A detailed introduction would require more space than available. Therefore, the reader unfamiliar with manifolds is referred to [38] and the reader unfamiliar with fibered manifolds and jet bundles is referred to [39]. A shorter introduction to jet bundles can be found in section 2 of chapter 3 of [13]. An advanced introduction that also includes the preliminaries for the theory of formal integrability is given in the article [17]. Further introductory material can be found in the [40].

1. M denotes a smooth manifold with dimension m. A point of M is denoted by x.
2. $\pi : E \to M$ denotes a smooth fibered manifold over M with dimension $d := m + e$, i.e. e is the dimension of the fiber.[4] p denotes a point of E. The local coordinates of E may be expressed by $C_E = (x^i, u^j)$. The convention is used that tuples like (x^i, u^j) stand for tuples like $(x^1, \cdots, x^m, u^1, \cdots, u^e)$. $\xi : F \to M$ is also a fibered manifold with local coordinates $C_F = (x^i, w^h)$ and dimension $m + f$.
3. Let $\alpha = \alpha_1 \cdots \alpha_n$ be a multi-index. It is a tuple of $n \in \mathbb{N}_0$ numbers $\alpha_i \in \{0, 1, \cdots, m = \dim(M)\}$ for which one defines the length $|\alpha| = n$. The tuple is commutative, i.e. $\alpha_i \alpha_j = \alpha_j \alpha_i$. One can multiply multi-indices as follows:

$$\alpha\sigma := \alpha_1 \cdots \alpha_n \sigma_1 \cdots \sigma_l \qquad \Rightarrow \qquad |\alpha\sigma| = n + l. \tag{1}$$

If $s : U \subset M \to E$ is a section of our fibered manifold $\pi : E \to M$, and $i \in \{0, \cdots, m\}$ an index and $\alpha = \alpha_1 \cdots \alpha_n$ a multi-index, then define

$$s_i^j := \frac{\partial s^j}{\partial x^i}, \qquad s_\alpha^j := \frac{\partial^n s^j}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_n}} \qquad \Rightarrow \qquad s_{\alpha i}^j = \frac{\partial^{n+1} s^j}{\partial x^{\alpha_1} \cdots \partial x^{\alpha_n} \partial x^i}. \tag{2}$$

[4] A fibered manifold $\pi : E \to M$ is a differentiable manifold E together with a differentiable surjective submersion π called projection.

A surjective submersion is a differentiable surjective map such that its pushforward π_* is also surjective at each point.

A fiber bundle is a fibered manifold with a local trivialization.

A vector bundle is a fiber bundle in which the fibers are vector spaces and whose transition maps are linear.

4. $J^k(E)$ denotes the k-th order jet bundle of E. It can be endowed with the structure of a smooth manifold. Locally, $J^k(E)$ may be described by the coordinates (x^i, u^j, u^j_σ) where σ is a multi-index and $1 \leq |\sigma| \leq k$. Please note that in contrast to (2), u^j_σ is not the derivative of u^j. Here the multi-index only serves as a label. $\pi^n_m : J^n(E) \to J^m(E)$ denotes the projection for all $0 \leq m \leq n$. $J^0(E) := E$ and $\pi^n : J^n(E) \to M$. $J^n(E)_x := (\pi^n)^{-1}(x)$ denotes the fiber of $J^n(E)$ over $x \in M$. Counting local coordinates, one obtains

$$\begin{aligned} \dim(J^k(E)) &= m + e\binom{m+k}{k}, \\ \dim(J^k(E)) - \dim(J^{k-1}(E)) &= e\binom{m-1+k}{k}, \end{aligned} \tag{3}$$

where $\dim J^k(E) - \dim J^{k-1}(E)$ is the dimension of the fiber of $\pi^k_{k-1} : J^k(E) \to J^{k-1}(E)$.
5. Let $s_E : M \to E$ and $s := s_F : M \to F$ be sections. $j^l(s)$ denotes the l-th prolongation of s. If $s(x) = (x^i, s^h(x))$ are the local coordinates of the section, then one can use the multi-index notation to give an explicit formulation of the prolongation[5]

$$j^l(s)(x) = (x^i, s^h(x), s^h_\alpha(x)),\ 1 \leq |\alpha| \leq l \tag{4}$$

6. A *differential equation* $\mathcal{E}$ is defined to be a fibered submanifold of $J^k(E)$. One can show that this is a geometric generalization of the usual notion of a (possibly non-linear) partial differential equation.
7. An essential notion in the algebro-geometric theory of PDEs, that is also heavily used in the present article, is the *differential consequence* or *prolongation* of a differential equation. To prolong a differential equation $\mathcal{E} \subset J^k(E)$ to a submanifold in $J^{k+l}(E)$, one needs the concept of repeated Jets: Since $\mathcal{E}$ is a fibered submanifold of $J^k(E)$, one can consider the fibered manifold $\pi^k|_\mathcal{E} : \mathcal{E} \to M$ and one can consider the space of jets of sections of $\pi^k|_\mathcal{E}$, called $J^l(\mathcal{E})$. Since $\mathcal{E}$ is a submanifold of $J^k(E)$, $J^l(\mathcal{E})$ is naturally a submanifold of the jet bundle $J^l(J^k(E))$.

   If $J^k(E)$ is locally described by the coordinates (x^i, u^j_σ), then the coordinates of $J^l(J^k(E))$ are $(x^i, (u^j_\sigma)_\alpha)$ where $|\sigma| \leq k$ and $|\alpha| \leq l$.[6] The subset of *repeated jets* in $J^l(J^k(E))$ consists of the image of the embedding

$$i_{k,l} : J^{k+l}(E) \to J^l(J^k(E)),\ j^{k+l}(s)(x) \mapsto j^l(j^k(s))(x) \tag{5}$$

[5] By Borel's lemma, given any point $\theta \in J^k(E)$, one can always find a section s_E such that $j^k(s_E)(x) = \theta$. However, given a submanifold O of $J^k(E)$, it is not always possible to find a section $s : \pi(O) \to E$ whose prolongation lies in O.

[6] Note that this is not the same as $u^j_{\sigma\alpha}$ because one "double-counts" those coordinates that arise from jets of sections whose derivatives would usually commute.

In local coordinates, this embedding reads $(x^i, u^j_{\sigma\alpha} = s^j_{\sigma\alpha}(x)) \mapsto (x^i, (u^j_\sigma)_\alpha = s^j_{\sigma\alpha}(x))$. One can show that it is well defined (see [39]).

Now one can prolong a fibered submanifold $\mathcal{E} \subset J^k(E)$ to a submanifold in $J^{k+l}(E)$ as follows. First take the intersection $J^l(\mathcal{E}) \cap i_{k,l}(J^{k+l}(E))$ within $J^l(J^k(E))$. In this intersection are only points of the form $j^l(j^k(s))(x)$ and therefore the projection $p : J^l(\mathcal{E}) \cap i_{k,l}(J^{k+l}(E)) \to J^{k+l}(E)$, $j^l(j^k(s))(x) \mapsto j^{k+l}(s)(x)$ is well-defined. Thus, define the l-th *prolongation of a PDE* $\mathcal{E} \subset J^k(E)$ *(into* $J^{k+l}(E)$*)* by

$$P^l(\mathcal{E}) := p(J^l(\mathcal{E}) \cap i_{k,l}(J^{k+l}(E))) \tag{6}$$

An intersection must not necessarily be a smooth manifold and therefore, a prolongation does not always exist in the category of smooth manifolds. In particular, the intersection might be empty.

8. Define the total differential operators $D^k_i,\ i \in \{1, \cdots, m = \dim(M)\}$ as vector fields on $J^k(E)$ locally by

$$D^k_i := \frac{\partial}{\partial x^i} + \sum_{j=1}^{e} \sum_{|\sigma|<k} u^j_{\sigma i} \frac{\partial}{\partial u^j_\sigma}, \qquad D_i := D^\infty_i. \tag{7}$$

If $\alpha = \alpha_1 \cdots \alpha_n$ is a multi-index, define $D_\alpha := D_{\alpha_1} \circ \cdots \circ D_{\alpha_n}$.

9. If $\pi : E \to M$ and $\pi' : E' \to M'$ are fibered smooth manifolds, then a smooth map $\Phi : E \to E'$ is called a *morphism of fibered (smooth) manifolds* if there exists a map $\phi : M \to M'$ such that $\pi' \circ \Phi = \phi \circ \pi$. A special case is $M = M'$, $\phi = \mathrm{id}$. Then the map Φ is a morphism of fibered manifolds if $\pi' \circ \Phi = \pi$. In the following, a morphism of fibered manifolds shall always refer to this special case if nothing else is mentioned.
10. A differential operator $\varphi : J \subset J^k(E) \to F$ is defined as a morphism of fibered manifolds. Its l-th prolongation is defined by

$$p^l(\varphi) : P^l(J) \to J^l(F), \qquad j^{k+l}(x) \mapsto j^l(\varphi(j^k(s)(x))) \tag{8}$$

In local coordinates, it is given by

$$p^l(\varphi)(x^i, u^j_{\sigma\alpha}) = (x^i, D_\alpha \varphi^h(x^i, u^j_\sigma)),\ 0 \le |\sigma| \le k,\ 0 \le |\alpha| \le l. \tag{9}$$

(Most often, one considers $J = J^k(E)$ and then $P^l(J) = J^{k+l}(E)$.)

11. Let $s : M \to F$ be a section. Define the kernel of a differential operator by

$$\ker_s(\varphi) := \left\{\theta \in J \mid \varphi(\theta) = s(\pi^k(\theta))\right\}. \tag{10}$$

12. Note that [17, Proposition 2.1] includes the statement that for any morphism $\varphi : A \to B$ of fibered manifolds (over the same base space) and any section s

of B, $\ker_s(\varphi)$ is a fibered submanifold of A if

$$s(M) \subset \varphi(J) \text{ and } \operatorname{rank}(\varphi) \text{ is locally constant.} \tag{11}$$

This holds in particular for a differential operator $\varphi : J \subset J^k(E) \to F$ (which, by definition, is a morphism of fibered manifolds) and therefore $\ker_s(\varphi)$ is a fibered submanifold of $J \subset J^k(E)$ and hence a differential equation whenever (11) holds for any differential operator φ.

13. If $\mathcal{E} = \ker_s(\varphi)$ is a differential equation, then the following equality holds,

$$\begin{aligned} P^l(\mathcal{E}) &= \ker_{j^l(s)}(p^l(\varphi)) \\ &= \left\{\theta \in P^l(J) \;\middle|\; D_\alpha \varphi^h(\theta) = D_\alpha s^h(\pi^{k+l}(\theta)),\, |\alpha| \le l\right\} \end{aligned} \tag{12}$$

14. For a section $s : U \subset M \to E$, denote by Γ_s^k the image of $j^k(s) : U \subset M \to J^k(E)$ and, for any section s and for any point $\theta \in \Gamma_s^k$, call $T_\theta \Gamma_s^k \subset T_\theta J^k(E)$ an R-plane. The span of all R-planes at a point $\theta \in J^k(E)$ is denoted by C_θ and is called Cartan-plane. The map $C : J^k(E) \to TJ^k(E),\ \theta \mapsto C_\theta$ is called Cartan distribution (sometimes also Vessiot distribution).
15. An integral submanifold of the Cartan distribution is defined to be a submanifold $W \subset J^k(E)$ such that $T_\theta W \subset C_\theta$ for all $\theta \in W$. An integral submanifold W is called locally maximal if no open subset of W can be embedded into an integral submanifold of greater dimension.
16. A solution S of a differential equation $\mathcal{E}$ is a locally maximal, $\dim(M)$-dimensional integral submanifold of C with $S \subset \mathcal{E}$. As emphasized before, this definition includes certain singular solutions (cf. [26]).

3 Correspondence and Intersection

This section develops the framework for the comparison of systems of differential equations. To this end, the most important concepts are those of a correspondence and an intersection which are described below.

3.1 Motivating Example

Consider the equations of magneto-statics and of the viscous Navier-Stokes equation (in a dimensionless form):

1. Magneto-statics:

$$\nabla \times \mathbf{B} = \mathbf{j}, \qquad \nabla \cdot \mathbf{B} = 0. \tag{13}$$

Here $\mathbf{B} = (B^1, B^2, B^3)^T$ denotes the magnetic field vector and $\mathbf{j} = (j^1, j^2, j^3)^T$ the charge current density.

2. Viscous, incompressible Navier-Stokes equations (without external forcing):

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right) \mathbf{u} = -\nabla \left(\frac{p}{\rho}\right) + \nu\Delta\mathbf{u}, \qquad \nabla \cdot \mathbf{u} = 0. \tag{14}$$

Here $\mathbf{u}$ is the velocity vector, p is the pressure, ρ is the density and ν is the viscosity coefficient.

Now let us make the following additional assumptions that might occur in some physical settings:

$$\begin{aligned} &(1) \text{ The current density } \mathbf{j} \text{ is the gradient of a function } \psi, \text{ i.e. } \mathbf{j} = -\nabla\psi, \\ &(2) \text{ The velocity flow is static, i.e. } 0 = d\mathbf{u}/dt = \partial u/\partial t + (\mathbf{u} \cdot \nabla)\mathbf{u}. \end{aligned} \tag{15}$$

If we apply those assumptions to the equations above and use the vector identity $\Delta\mathbf{u} = \nabla(\nabla \cdot \mathbf{u}) - \nabla \times (\nabla \times \mathbf{u})$ as well as $\nabla \cdot \mathbf{u} = 0$ and $\nabla \cdot (\nabla \times \mathbf{u}) = 0$ (because of grad $\circ$ rot $= 0$), the systems of equations above become:

$$\begin{aligned} &\nabla \times \mathbf{B} = -\nabla\psi, \qquad \nabla \cdot \mathbf{B} = 0 \\ \text{and} \qquad &\nabla \times (\nabla \times \mathbf{u}) = -\nabla\phi, \qquad \nabla \cdot (\nabla \times \mathbf{u}) = 0, \qquad \nabla \cdot \mathbf{u} = 0. \end{aligned} \tag{16}$$

where $\phi := p/(\rho\nu)$. It is apparent that those equations aquire a similar form under the "correspondence"

$$\mathbf{B} = \nabla \times \mathbf{u}. \tag{17}$$

Or, put differently, if one replaced $\mathbf{B}$ by $\nabla \times \mathbf{u}$, then the system of all equations together would be consistent.

And in fact, because $\nabla \cdot \mathbf{B} = 0$, we can use the Poincaré Lemma (in any star-shaped region) to conclude that there exists a vector potential $\mathbf{A}$ such that $\nabla \times \mathbf{A} = \mathbf{B}$ and because gauge transformations do not change the physics of classical electrodynamics (and in particular of magneto-statics), we can use them to gauge $\mathbf{A}$ in such a way that $\nabla \cdot \mathbf{A} = 0$. Therefore, under the above assumptions, there is a direct correspondence between $\mathbf{A}$ (in some gauge) and $\mathbf{u}$. The physical interpretation is that a static fluid velocity field behaves like the vector potential of magneto-statics with certain charge current densities.[7] This can give a new intuition about the corresponding physical phenomena.

As this example illustrates in an intuitive way, (16) and (17) describe "shared structure" of the equations (13) and (14) under the conditions (15). But what is the appropriate space in which the correspondence (17) holds and in which the

[7] Of course the initial and boundary conditions additionally influence the solutions.

shared structure can be obtained? Is there a way to compute the assumptions (15) instead of guessing them, given the correspondence (17)? And how to generalize the procedure? To answer those and other questions already motivated in the introduction, a general framework is constructed in the next subsections.

3.2 Formal Definitions

Suppose that $\pi : E \to M$ and $\xi : F \to M$ are fibered manifolds with the same base space (a generalization to different base spaces is work in progress). Suppose further that we are given two PDEs $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$. We want to relate the PDEs in a space in which we can embed both of them. A natural choice is their pullback in the category of smooth manifolds, i.e. their so-called fibered product

$$J := J^k(E) \times_M J^l(F) := \left\{ J^k(E)_x \times J^l(F)_x \mid x \in M \right\} . \tag{18}$$

Now the canonical projections $\pi_E : J \to J^k(E)$ and $\pi_F : J \to J^l(F)$ allow to pull $\mathcal{E}$ and $\mathcal{F}$ back to J:

$$\mathcal{E}_J := \pi_E^{-1}(\mathcal{E}) \qquad \text{and} \qquad \mathcal{F}_J := \pi_F^{-1}(\mathcal{F}). \tag{19}$$

π_E and π_F can also be used to pull back the Cartan distribution defined on $J^k(E)$ and $J^l(F)$: If Σ_E is the module of Cartan forms (the differential forms that annihilate the Cartan distribution) on $J^k(E)$ and Σ_F is the module of Cartan forms on $J^l(F)$, then the module Σ on $J^k(E) \times_M J^l(F)$ is generated by $\pi_E^* \Sigma_E$ and $\pi_F^* \Sigma_F$.

Though the two equations are now pulled back into a natural common space, they are not yet related. Directly intersecting $\mathcal{E}_J$ and $\mathcal{F}_J$ would result in a space

$$\mathcal{E}\mathcal{F} := \mathcal{E}_J \cap \mathcal{F}_J \tag{20}$$

that is big enough to accomodate all solutions of both $\mathcal{E}$ and $\mathcal{F}$, even if $\mathcal{E}$ and $\mathcal{F}$ are completely unrelated. Therefore, one additionally needs to intersect $\mathcal{E}_J$ and $\mathcal{F}_J$ with a third submanifold $\Phi \subset J$ in order to relate them.

But what kind of submanifold is Φ supposed to be? One would not like Φ to be of the form $\pi_E^{-1}(\phi)$ or $\pi_F^{-1}(\psi)$ for $\phi \subset J^k(E)$ and $\psi \subset J^l(F)$ because this would only impose additional relations on one of the pulled back equations. Instead Φ is supposed to relate the fibers of $J^k(E)$ with those of $J^l(F)$ without imposing such additional conditions. To ensure that, one might require that Φ is large enough to fulfill $\mathcal{E} \subset \pi_E(\Phi)$ and $\mathcal{F} \subset \pi_F(\Phi)$. This would in particular imply that $M \subset \Phi$, i.e. Φ would not impose any relations on M. However, the condition $\mathcal{E} \subset \pi_E(\Phi)$ might be considered too weak because it does not necessarily ensure that Φ does not impose any relations on $\mathcal{E}_J$ locally. At the same time, the condition $\mathcal{E} \subset \pi_E(\Phi)$ in a different sense might also be considered too strong because it does not allow to restrict the comparison of the PDEs to a particular open neighbourhood (for example, by adding some inequalities to the local definition of Φ). Both issues can be resolved, however, by requiring instead that for all open $W \subset \mathcal{E}_J \cap \Phi$, one has $\pi_E(W)$ open in $\mathcal{E}$.

The previous condition ensures that the dimension of Φ is locally sufficiently large. At the same time, it should not be arbitrarily large because this would again not impose any relations and thus render the intersection meaningless. Since every submanifold can locally be described by a set of equations where the number of independent equations is equal to the codimension of the submanifold (see also Sect. 4.1), the codimension of Φ quantifies the number of (global) relations it imposes. To ensure that the dependent variables of at least either $\mathcal{E}_J$ or $\mathcal{F}_J$ are determined in terms of the other, this codimension should at least equal $n(x) := \min(e(x), f(x))$ where $e(x) := \dim(E_x)$ and $f(x) := \dim(F_x)$ are the dimensions of the fibers of E and F over x. (Often they are constant and do not depend on x. They are always locally constant because we work in the category of smooth manifolds.) The above thoughts can be summarized in the following definitions.

Definition 1 Let $p : Y \to X$ and $q : Z \to X$ be fibered manifolds, let S_Y be a submanifold of Y and S_Z be a submanifold of Z, let $Y \times_X Z$ be the fibered product of Y and Z over X and let $\pi_Y : Y \times_X Z \to Y$ and $\pi_Z : Y \times_X Z \to Z$ be the canonical projections.

A submanifold $S \subset Y \times_X Z$ is called *almost diagonal* iff for all open subsets $U \subset S$, the set $\pi_Y(U)$ is an open subset of Y and the set $\pi_Z(U)$ is an open subset of Z.

A submanifold $S \subset Y \times_X Z$ is called *almost diagonal to S_Y and S_Z* iff for all open subsets $U \subset S$, the set $\pi_Y(U) \cap S_Y$ is an open subset of S_Y and the set $\pi_Z(U) \cap S_Z$ is an open subset of S_Z.

As said above, intuitively, the definition is supposed to ensure that the submanifold S is defined by equations, that either only relate fiber coordinates of Y with fiber coordinates of Z within the fibered product $Y \times_X Z$, or, if it imposes additional relations on the coordinates of Y or Z alone within the fibered product, then those relations must already be imposed by $\pi_Y^{-1}(S_Y)$ and $\pi_Z^{-1}(S_Z)$. The previous definition now allows to define a correspondence.

Definition 2 A *correspondence* between $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$, is a fibered submanifold Φ of $J^k(E) \times_M J^l(F)$ with $\mathrm{cod}(\Phi)(x) \geq \min(\dim(E_x), \dim(F_x))$ which is almost diagonal to $\mathcal{E}$ and $\mathcal{F}$.

Given a natural space to relate two PDEs, one can define their common part as a set-theoretic intersection.

Definition 3 Given a correspondence Φ between $\mathcal{E}$ and $\mathcal{F}$, their *intersection* $\mathcal{I}$ is defined by

$$\mathcal{I} := \mathcal{E}_J \cap \mathcal{F}_J \cap \Phi. \tag{21}$$

Those definitions allow to define shared structure in Sect. 6.

3.3 Local Description

Given a smooth manifold X, we denote its local coordinates (in some suitably adapted chart) by C_X. If m denotes the dimensions of M and if e and f denote the dimensions of the fibers of E and F, then the local coordinates of the manifolds described above are given by

$$\begin{aligned}
&C_M = (x^i), \quad C_E = (x^i, u^j), \quad C_{J^k(E)} = (x^i, u^j_\alpha), \quad i \le m,\ j \le e,\ |\alpha| \le k \\
&C_F = (x^i, v^g), \quad C_{J^l(F)} = (x^i, v^g_\beta), \quad i \le m,\ g \le f,\ |\beta| \le l \\
&C_{J^k(E)\times_M J^l(F)} = (x^i, u^j_\alpha, v^g_\beta), \quad i \le m,\ j \le e,\ g \le f,\ |\alpha| \le k,\ |\beta| \le l
\end{aligned} \tag{22}$$

If $\mathcal{E} \subset J^k(E)$ is a submanifold, then (by Proposition 2) it can locally always be described as the kernel of independent functions, i.e. by equations $F_a(x^i, u^j_\alpha) = 0$, where $1 \le a \le r$. In other words, the submanifold $\mathcal{E}$ is locally, in some neighbourhood $U \subset J^k(E)$ defined by those points (x^i, u^j_α) contained in U that are subject to the conditions $F_a(x^i, u^j_\alpha) = 0$. Instead of $\mathcal{E} = \{\ (x^i, u^j_\alpha) \in U \subset J^k(E) \mid F_a(x^i, u^j_\alpha) = 0\ \}$, the following shorthand notation is used.

$$\mathcal{E} : \left\{ F_a(x^i, u^j_\alpha) = 0 \right\} \tag{23}$$

If we now pull back $\mathcal{E}$ to $\mathcal{E}_J = \pi_E^{-1}(\mathcal{E}) \subset J^k(E) \times_M J^l(F)$, then $\mathcal{E}_J$ is locally described by those equations that define the points in the inverse image $\pi^{-1}(U \cap \mathcal{E})$. This inverse image consists of all points $(x^i, u^j_\alpha, v^g_\beta)$ such that $\pi(x^i, u^j_\alpha, v^g_\beta) = (x^i, u^j_\alpha) \in U \cap \mathcal{E}$. But $(x^i, u^j_\alpha) \in U \cap \mathcal{E}$, precisely iff $F_a(x^i, u^j_\alpha) = 0$. Thus, the points in $\pi^{-1}(U \cap \mathcal{E})$ are described by the same equations as the points in $U \cap \mathcal{E}$. As a consequence, $\mathcal{E}_J$ is locally described by the conditions $F_a(x^i, u^j_\alpha) = 0$ but now imposed on an open neighbourhood of $J^k(E) \times_M J^l(F)$.

Furthermore, if $\mathcal{E}_J$ and $\mathcal{F}_J$ are locally defined by points fulfilling the equations $F^{\mathcal{E}}_a(x^i, u^j_\alpha) = 0$ and $F^{\mathcal{F}}_b(x^i, v^g_\beta) = 0$, then their intersection is necessarily locally defined by those points that simultaneously fulfill both equations. In other words, the *intersection* of $\mathcal{E}_J$ and $\mathcal{F}_J$ is locally described by the *union* of their equations.[8] As a consequence, all local descriptions can be summarized as follows.

$$\begin{aligned}
J^k(E) \supset \mathcal{E} &: \{F^{\mathcal{E}}_a(x^i, u^j_\alpha) = 0, \qquad 0 \le a \le r_{\mathcal{E}}\ \}, \\
J^l(F) \supset \mathcal{F} &: \{F^{\mathcal{F}}_b(x^i, v^g_\beta) = 0, \qquad 0 \le b \le r_{\mathcal{F}}\}, \\
J^k(E) \times_M J^l(F) \supset \Phi &: \{\phi_c(x^i, u^j_\alpha, v^g_\beta) = 0, \qquad 0 \le c \le r_\Phi\ \}, \\
J^k(E) \times_M J^l(F) \supset \mathcal{I} &: \{F^{\mathcal{E}}_a(x^i, u^j_\alpha) = F^{\mathcal{F}}_b(x^i, v^g_\beta) = \phi_c(x^i, u^j_\alpha, v^g_\beta) = 0\}.
\end{aligned} \tag{24}$$

[8] The intuitive reason is that each equation represents a constraint on the space of solutions and therefore the intersection, which is smaller than both original solution spaces, must be described by the union of those constraints.

Example Here, a simple version of the motivating example of the previous Sect. 3.1 is rephrased in the present terminology. The equations are modeled on a flat, Euclidean spacetime $\mathbb{R}^3 \times \mathbb{R}$ with local coordinates (x^i, t), $i \in \{1, 2, 3\}$.

For Hydrodynamics, let the fibered manifold (which is now a trivial vector bundle) be $\pi : E := M \times \mathbb{R}^3 \to M$ with local coordinates (x^i, t, u^i) and dimension $\dim(E) = m + e = 4 + 3$. Let $u^{i,j}$ denote the coordinates corresponding to $\partial u^i / \partial x^j$ and recall that the sum convention is used. Let $p : M \to \mathbb{R}$ be a given function called pressure and $\rho, \nu \in \mathbb{R}$ be the constant density and viscocity. Denote by $p^{,i}$ the components of the gradient of p. Describe $J^2(E)$ with local coordinates $(x^i, t, u^i, u^{i,j}, u^i_t, u^{i,jk}, u^{i,j}_t)$ where $u^{i,jk} = u^{i,kj}$. The Navier-Stokes equations described in (14) in this setting are then given by

$$\mathcal{E} : \left\{ u^i_t + u^j u^{i,j} = -\frac{1}{\rho} p^{,i} + \nu u^{i,jj}, \qquad u^{i,i} = 0 \right\} \tag{25}$$

Magneto-statics is also modeled on M, even though the equations do not involve any time-component. (Since a realistic experiment always takes place in space and time, even though the fields might not change over time, this is not a bad assumption.) Thus, for magneto-statics, the vector bundle $\xi : F := M \times \mathbb{R}^3 \to M$ with local coordinates (x^i, t, B^i) is defined and the magneto-static equations (corresponding to (13)) are described on $J^1(F)$ with local coordinates $(x^i, t, B^i, B^{i,j})$ via

$$\mathcal{F} : \left\{ \varepsilon_{ijk} B^{k,j} = I^i, \qquad B^{i,i} = 0 \right\} \tag{26}$$

(the letter I is used for the current density instead of j to avoid confusion with other j's). The natural product bundle $J := J^2(E) \times_M J^1(F)$ has local coordinates $(x^i, t, u^i, B^i, u^{i,j}, u^i_t, u^{i,jk}, u^{i,j}_t, B^{i,j}, B^i_t)$ one can define $\Phi \subset J$ as the submanifold locally given by

$$\Phi : \left\{ B^i = \varepsilon_{ijk} u^{k,j} \right\} \tag{27}$$

Since Φ does not contain any equations that relate the coordinates of $J^2(E)$ or $J^1(F)$ among themselves, the projection $\pi_E(U)$, of all of its open subsets $U \subset \Phi$, is open in $J^2(E)$ and $\pi_F(U)$ is open in $J^1(F)$. Hence Φ is almost diagonal. As a consequence, $\pi_E(U) \cap \mathcal{E}$ is also always open in $\mathcal{E}$ and $\pi_F(U) \cap \mathcal{F}$ is always open in $\mathcal{F}$. Hence, Φ is almost diagonal to $\mathcal{E}$ and $\mathcal{F}$ and is therefore a correspondence in the sense of Definition 2. We can thus define a valid intersection $\mathcal{I}$ by the following equations.

(continued)

$$\mathcal{I} = \pi_E^{-1}(\mathcal{E}) \cap \pi_F^{-1}(\mathcal{F}) \cap \Phi : \quad \left\{ \begin{array}{l} u_t^i + u^j u^{i,j} = -\frac{1}{\rho} p^{,i} + \nu u^{i,jj}, \; u^{i,i} = 0 \\ \varepsilon_{ijk} B^{k,j} = I^i, \quad B^{i,i} = 0 \\ B^i = \varepsilon_{ijk} u^{k,j} \end{array} \right\} \tag{28}$$

The shared structure contained in this intersection can be computed with the methods that are going to be introduced in the following sections. In Sect. 9.4, it is shown that this shared structure indeed corresponds to the one described in (16), and the assumption of a static fluid flow, guessed in (15), is the result of the computation of the minimal consistency conditions for shared structure to arise.

4 Consistency Conditions

The conditions that need to be satisfied in order to be able to speak of a meaningful intersection of two PDEs are related to at least two areas, namely transversality theory in differential topology and the theory of formal integrability. The next subsections, as well as Sect. 5, provide all corresponding background information in those areas that are needed to understand the rest of the article.

4.1 Smoothness Conditions

In this subsection is investigated under which circumstances the intersection $\mathcal{I}$ of two differential equations is actually again a differential equation, that means a smooth submanifold of a jet space.

The intersection theory of differential topology can answer this question. The remainder of this subsection largely follows [41] and those theorems that are needed in the present context are cited. The starting point is the preimage theorem which is a quite straightforward consequence of the inverse function theorem and the local submersion theorem.

Definition 4 For a smooth map $f : X \to Y$, a point $y \in Y$ is called a *regular value* if the pushforward (or differential) $df_x : T_x X \to T_{f(x)} Y$ is surjective for all $x \in f^{-1}(y)$.

Proposition 1 (Preimage Theorem) *If y is a regular value of $f : X \to Y$, then $f^{-1}(y)$ is a smooth submanifold with dimension $dim(X) - dim(Y)$.*

Note that it is often not hard to check if the pushforward of a smooth map is surjective. It amounts to checking the rank of the Jacobian matrix.

There is also a partial converse to the theorem, namely

Proposition 2 *If $Z \subset X$ is a smooth submanifold, then it can locally be defined as the kernel of independent smooth functions.*

The following proposition is also useful.

Proposition 3 *Let $f : X \to Y$ be smooth and y a regular value of f. The tangent space of $Z := f^{-1}(y)$ is given by $T_x Z = \ker(df_x)$ for any $x \in Z$.*

The next step is to consider what happens if one does not only look at the preimage of a single regular value but at the preimage of a submanifold. Then one can use the definition of transversality to prove the following theorem.

Definition 5 The map $f : X \to Y$ is said to be transversal to the submanifold $Z \subset Y$, abbreviated $f \pitchfork Z$, if the equation

$$\mathrm{im}(df_x) + T_{f(x)} Z = T_{f(x)} Y \tag{29}$$

holds true at each point x in $f^{-1}(Z)$.

Proposition 4 *If the smooth map $f : X \to Y$ is transversal to a submanifold $Z \subset Y$, then the preimage $f^{-1}(Z)$ is a submanifold of X. Moreover,*

$$cod(f^{-1}(Z) \subset X) = cod(Z \subset Y) \tag{30}$$

Given the manifold Y and two submanifolds $X \subset Y$ and $Z \subset Y$, one can apply the above theorem to their intersection $X \cap Z$ as follows: If $i : X \to Y$ is the canonical inclusion that embeds X into Y then $X \cap Z = i^{-1}(Z)$. Since $\mathrm{im}(di_x) = T_x X$, and $T_{i(x)} Z = T_x Z$, one obtains

Proposition 5 *If X and Z are smooth submanifolds of Y, then $X \cap Z$ is a smooth submanifold of Y iff $X \pitchfork Z$, that means*

$$T_x X + T_x Z = T_x Y \tag{31}$$

for all $x \in X \cap Z$. In this case $cod(X \cap Z) = cod(X) + cod(Z)$.

Condition (31) can be checked locally. Indeed, we obtain the following proposition as a consequence.

Proposition 6 *If X and Z are submanifolds of Y, locally described by equations of the form $F^X_a(y^i) = 0$ with $1 \le a \le r$ and $F^Z_b(y^i) = 0$ with $1 \le b \le q$, then $X \pitchfork Z$ iff dF as defined in Eq. (32) has full rank.*

Proof If Y has local coordinates (y^i), then, since X is a smooth submanifold of Y, by Proposition 2, every local chart $U \subset X$ is described as the kernel of independent functions $F^X : Y \to \mathbb{R}^r$, i.e. $U = (F^X)^{-1}(0)$, or, equivalently, we write as before

$X : \{F_a^X(y^i) = 0\}$ with $1 \leq a \leq r$. Using Proposition 3, we can then compute $T_x X$ as the kernel of dF^X. Similarly, if Z is locally described by $F_b^Z(y^i) = 0$ with $1 \leq b \leq q$, then $T_x Z = \ker(dF^Z)$ and $X \cap Z$ is locally described by the joint system of those equations, i.e. by

$$0 = F_c(y^i) = \begin{cases} F_c^X(y^i), & 1 \leq c \leq r \\ F_{c-r}^Z(y^i), & r+1 \leq c \leq r+q \end{cases} \tag{32}$$

By Proposition 5, $X \pitchfork Z$ iff $X \cap Z = F^{-1}(0)$ is a smooth submanifold, which, by the preimage theorem, 1, is true if dF is surjective, i.e. has full (row) rank. □

Furthermore, using Sard's theorem, one can prove the transversality theorem which guarantees that almost all maps of a family of smooth maps are transversal to some submanifold in the codomain.

Proposition 7 (Sard) *The set of values of a smooth map $f : X \to Y$ which are not regular has Lebesgue measure zero.*

This means "almost all" points of a smooth map are regular. However, sets of measure zero can be quite large, for example the subset $\mathbb{R}^n$ has measure zero in $\mathbb{R}^{n+1}$.

Proposition 8 (Transversality Theorem) *Suppose that $F : X \times S \to Y$ is a smooth map between smooth manifolds, where only X has boundary, and let Z be any boundaryless submanifold of Y. One can use F to define a smooth family of homotopic maps by $f_s(x) := F(x, s)$. If both F and ∂F are transversal to Z, then for almost every $s \in S$, both f_s and ∂f_s, are transversal to Z.*

For a map $f : X \to \mathbb{R}^m$ this immediately implies that transversality is a generic feature because one can simply define S as an open subset of $\mathbb{R}^m$ and define $F(x, s) := f(x) + s$. As S is open in $\mathbb{R}^m$, this means that F is surjective everywhere and therefore Definition 5 is always fulfilled. Following this thought further, one can prove the so-called **transversality homotopy theorem**.

Proposition 9 *For any smooth map $f : X \to Y$ and any boundaryless submanifold Z of the boundaryless manifold Y, there exists a smooth map $g : X \to Y$ homotopic to f such that $g \pitchfork Z$ and $g \pitchfork \partial Z$.*

Now reconsider two differential equations $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$. Using the above, one can show the following.

Proposition 10 *$\mathcal{EF} = \mathcal{E}_J \cap \mathcal{F}_J$ as defined in Eq. (20) is a PDE, i.e. a smooth submanifold.*

Proof To check that $\mathcal{EF}$ is a smooth submanifold, it suffices, by Proposition 5, to check that $\mathcal{E}_J \pitchfork \mathcal{F}_J$, which in turn, can be checked locally using Proposition 6. So if $\mathcal{E}$ and $\mathcal{F}$ are locally described by the independent smooth functions $F_a^{\mathcal{E}}(x^i, u_\alpha^j)$ and $F_b^{\mathcal{F}}(x^i, v_\beta^g)$ as in Eq. (24), then we must check if the differential of the joint system of equations $F_c = 0$ as defined in Eq. (32) has full rank. Since $\mathcal{E}$ and $\mathcal{F}$ are assumed to be fibered submanifolds of $J^k(E)$ and $J^l(F)$, they do not impose any conditions on M. Furthermore, $F^{\mathcal{E}}$ does not depend on v^g_β and $F^{\mathcal{F}}$ does not depend on u^j_α. Hence, $dF^{\mathcal{E}}$ and $dF^{\mathcal{F}}$ are linearly independent. Since they are both assumed to have full rank and they are independent, the joint system dF must also have full rank. □

This theorem implies the following.

Corollary 1 *$\mathcal{I} = \mathcal{EF} \cap \Phi$ is a PDE iff $\mathcal{EF} \pitchfork \Phi$.*

Whether $\mathcal{EF} \pitchfork \Phi$ or not depends on the definition of Φ and can not be proven in general. To check it explicitly in practice for a given $\mathcal{EF}$ and a given Φ, one can calculate the rank of the joint system as described in Propositions 5 and 6. If this rank is locally maximal, then transversality is guaranteed. If the rank is not locally maximal but locally constant, then we can restrict the codomain such that the smooth system becomes locally maximal. Therefore, the intersection is also a well-defined smooth submanifold at those points around which the system is locally constant.

This is also the reason why the preimage of a differential operator which has locally constant rank is a smooth submanifold, i.e. a differential equation (if it is not empty), see condition (11).

However, even if $\mathcal{EF}$ is not transversal to Φ, then the transversality theorem 9 implies that it suffices to deform Φ (locally this means to perturb the smooth functions describing Φ) just ever so slightly in order to obtain an intersection that is a well-defined object in the category of smooth manifolds. Furthermore, if some smooth functions ϕ_c locally describes our manifold Φ, one way to make it transversal to $\mathcal{EF}$ is to use $F(x, s) := \phi(x) + s$ for some very small s. Thus, the theorem assures us that taking intersections of $\mathcal{EF}$ and Φ is not a hopeless endeavor but to the contrary can always lead to a smooth manifold at least after slight deformations.

4.2 Differential Consistency

In the last subsection was clarified when the intersection of two differential equations is actually again a differential equation. As a next step, it is assumed that the intersection *is* a differential equation, i.e. a smooth submanifold, and it is asked if the PDE has solutions.

Ultimately, one is interested in the existence of smooth (or even more general) solutions but since there is not yet any general theory that allows to compute whether a solution (in any non-analytic category) of a PDE exists or not, it is necessary at this point to ask for something weaker. The next best thing after a general condition that allows to compute the existence of solutions is to ask for the existence of so-called formal solutions. Formal solutions are formal power series that formally solve the PDE (i.e. the series satisfies all algebraic equations describing the smooth solution spaces that characterize the PDE and its prolongations) but is not guaranteed to converge or might converge to something that is not a solution.

Formal solutions are tractable because their existence is encoded in the *differential consequences* of a PDE.

In particular, if one can prolong an equation infinitely many times in a certain smooth way without obtaining any contradiction, then one "point" of the infinite prolongation $P^{\infty}(\mathcal{E})$ can be seen as the sequence of coefficients for a (not necessarily converging) taylor expansion that solves the equation locally around the projection of that point. However, as remarked below Eq. (6), the prolongations of $\mathcal{E}$ do not necessarily exist.

This means that in order to check if formal solutions exist, one needs a general formalism to determine if a PDE is differentially consistent in the sense that all of its prolongations exist.

Furthermore, recall that in the motivating example in Sect. 3.1, we had to make certain physical assumptions (15). It would be beautiful if those assumptions could be obtained in a systematic way. In general, if one could obtain the minimal amount of assumptions that must be made to make a system differentially consistent (if such assumptions exist), then this would be optimal. Fortunately, one can use the theory of *formal integrability* for this purpose. In particular, the "physical assumptions" come out of the formalism as "integrability conditions" that are needed for consistency.

Since the theory is somewhat involved, the next section provides an introduction to the theory of formal integrability. In the section after the next, those notions of formal integrability are combined with the notions of correspondence and intersection defined above to define what it means for two theories to share structure.

5 Formal Integrability

Subsequently, the introduction follows [42], [17], and [18] (chapter IX) to introduce the notion of formal integrability. The first subsection contains the necessary definitions and the derivation of explicit coordinate expressions which are missing in Goldschmidt's publications, as well as the derivation of Proposition 11 that can simplify some computations.

The second subsection describes the main theorems of the formal theory. The third subsection discusses integrability conditions which are especially important for subsequent constructions. The reader already familiar with formal integrability can directly proceed with Sect. 5.3. The reader who prefers to learn with examples is referred to Sect. 5.4.

5.1 Definitions and Preliminaries

1. Recall that if X, Y and N are manifolds and $f : Y \to X$ is a smooth map and $\pi : N \to X$ is a fiber bundle with fibers denoted by $N_x,\ x \in X$, then f^*N denotes the pullback bundle over Y and it is defined as follows:

$$f^*N := \left\{ N_{f(y)} \mid y \in Y \right\} . \tag{33}$$

   To each point $y \in Y$, we attach the fiber $N_{f(y)}$ that would usually be attached to the point $x = f(y) \in X$.

   Suppose we are given the following configuration of smooth maps between smooth manifolds:

$$\begin{tikzcd}
 & X_1 & N^1 \arrow[l, "\pi_1"'] \\
Y \arrow[ur, "f"] \arrow[dr, "g"'] & & \\
 & X_2 & N^2 \arrow[l, "\pi_2"]
\end{tikzcd}$$

   where $\pi_i : N^i \to X_i$ are vector bundles (not just fiber bundles). Then we define

$$N^1 \otimes_Y N^2 := f^*N^1 \otimes g^*N^2 = \left\{ N^1_{f(y)} \otimes N^2_{g(y)} \mid y \in Y \right\} . \tag{34}$$

   which is a vector bundle over Y.
2. Now, for any $k \geq 0$, let $V(J^k(E)) \to J^k(E)$ denote the vertical subbundle of the tangent bundle $TJ^k(E)$ of $J^k(E)$ containing those vectors which are tangent to the fibers of $\pi : J^k(E) \to M$. It is a bundle over $J^k(E)$. In a local neighbourhood $U \subset J^k(E)$ with coordinates (x^i, u^j, u^j_σ), $V(J^k(E))$ is the span of the vector fields

$$V(J^k(E)) = \text{span} \left(\frac{\partial}{\partial u^j}, \frac{\partial}{\partial u^j_\sigma} \right) \tag{35}$$

   and we have

$$\pi_{*,\theta} \left(\left. \frac{\partial}{\partial u^j} \right|_\theta \right) = 0 = \pi_{*,\theta} \left(\left. \frac{\partial}{\partial u^j_\sigma} \right|_\theta \right) \in T_{\pi(\theta)} M \tag{36}$$

   at every point $\theta \in U \subset J^k(E)$.
3. If M denotes our base manifold as before, we denote by T^* its tangent bundle, by $S^k T^*$ the k-th symmetric power of the tangent bundle and by $\Lambda^k T^*$ the k-th anti-symmetric power.

In local coordinates, general elements of those spaces are written

$$
\begin{aligned}
T^* \ni v &= v_i dx^i, \ i \in \{1, \cdots, m\} \\
S^k T^* \ni a &= a_{i_1 \dots i_k} dx^{i_1} \vee \cdots \vee dx^{i_k}, \ i_j \in \{1, \cdots, m\} \\
\Lambda^k T^* \ni w &= w_{i_1 \dots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}, \ i_j \in \{1, \cdots, m\}
\end{aligned}
\tag{37}
$$

where the sum convention is always used. $S^k T^*$ and $\Lambda^k T^*$ are different in that $dx^{i_j} \vee dx^{i_k} = dx^{i_k} \vee dx^{i_j}$ but $dx^{i_j} \wedge dx^{i_k} = -dx^{i_k} \wedge dx^{i_j}$. As a consequence, $\dim \Lambda^k T^* \leq \dim(S^k T^*)$. To calculate the dimension, note that there are as many symmetric basis elements as there are ways to put k balls between $m-1$ sticks. Thus,

$$
\dim(S^k T^*) = \binom{m-1+k}{k}, \qquad \dim(\Lambda^k T^*) = \binom{m}{k} \tag{38}
$$

If one has a multi-index α with $|\alpha| = k$, one can define $dx^\alpha_\vee := dx^{\alpha_1} \vee \cdots \vee dx^{\alpha_k}$ and $dx^\alpha_\wedge := dx^{\alpha_1} \wedge \cdots \wedge dx^{\alpha_k}$ to write more concisely

$$
S^k T^* \ni a = a_\alpha dx^\alpha_\vee, \ |\alpha| = k, \qquad \Lambda^k T^* \ni w = w_\alpha dx^\alpha_\wedge, \ |\alpha| = k. \tag{39}
$$

4. Define the map $\Delta_{l,k} : S^{l+k} T^* \rightarrow S^l T^* \otimes S^k T^*$ as the composition

$$
S^{l+k} T^* \xrightarrow{\quad i \quad} \otimes^{l+k} T^* \xrightarrow{\quad s_{l,k} \quad} S^l T^* \otimes S^k T^*
$$

where i is the injection given by

$$
i(dx^{i_1} \vee \ldots \vee dx^{i_{k+l}}) := \sum_{\sigma \in \mathfrak{S}_{k+l}} dx^{\sigma(i_1)} \otimes \ldots \otimes dx^{\sigma(i_{l+k})} \tag{40}
$$

where the sum goes over all entries σ of the permutation group $\mathfrak{S}$. And $s_{l,k}$ is the projection given by

$$
s_{l,k}(dx^{i_1} \otimes \cdots \otimes dx^{i_{l+k}}) := dx^{i_1} \vee \cdots \vee dx^{i_l} \otimes dx^{i_{l+1}} \vee \cdots \vee dx^{i_{l+k}} \tag{41}
$$

Thus, all in all, we obtain

$$
\begin{aligned}
\Delta_{l,k}(dx^{i_1} \vee \ldots \vee dx^{i_{k+l}}) = \sum_{\sigma \in \mathfrak{S}_{k+l}} & dx^{\sigma(i_1)} \vee \ldots \vee dx^{\sigma(i_l)} \\
& \otimes dx^{\sigma(i_{l+1})} \vee \cdots \vee dx^{\sigma(i_{l+k})}.
\end{aligned}
\tag{42}
$$

5. Given some smooth manifold Y and maps $\pi : Y \rightarrow M$ and $\pi_0 : Y \rightarrow E$, define

$$
F^k_Y := S^k T^* \otimes_Y V(E) \tag{43}
$$

Call it k-fiber (over Y). The k-fiber is the vector bundle whose fibers have as many dimensions (and hence local coordinates) as there are local coordinates of order k on $J^k(E)$. This can be seen by observing that

$$\dim((F^k_Y)_{p\in Y}) \overset{(38)}{=} \dim(S^kT^*)\cdot e \overset{(3)}{=} \dim(J^k(E)) - \dim(J^{k-1}(E)). \tag{44}$$

In local coordinates, an element $p \in F^k_Y$ can be written $p = (\theta, a)$ where $\theta \in Y$ and

$$a = a^j_{i_1\dots i_k} dx^{i_1}\vee\dots\vee dx^{i_k}|_{\pi(\theta)} \otimes \frac{\partial}{\partial u^j}\bigg|_{\pi_0(\theta)} = a_\alpha dx^\alpha_\vee|_{\pi(\theta)} \otimes \frac{\partial}{\partial u^j}\bigg|_{\pi_0(\theta)}. \tag{45}$$

6. One can show (see [17, Proposition 5.1]) that for $k \geq 1$, the jet bundle $J^k(E)$ is an affine bundle over $J^{k-1}(E)$, modeled on the vector bundle $S^kT^* \otimes_{J^{k-1}(E)} V(E)$ over $J^{k-1}(E)$ ((44) shows that the dimensions match).
   As described in chapter IX.§3 of [18], if $\theta \in J^{k-1}(E)$, the vector space $S^kT^*_{\pi(\theta)} \otimes V_{\pi_0(\theta)}(E)$ considered as an additive group acts freely and transitively on the fiber of $J^k(E)$ over θ. As a consequence, for $a \in S^kT^*_{\pi(\theta)} \otimes V_{\pi_0(\theta)}(E)$, we can denote by $q+a$ the image of the element q of the fiber $J^k(E)_\theta$ under the action of a. If (x, u^j, u^j_σ) are the local coordinates of q, the local coordinates of $q + a$ are (x, u^j, z^j_σ) where

$$\begin{pmatrix} z^j_\sigma = u^j_\sigma, & \text{if } |\sigma| < k \\ z^j_\sigma = u^j_\sigma + a^j_{\alpha=\sigma}, & \text{if } |\sigma| = k \end{pmatrix} \tag{46}$$

   (Goldschmidt also provides an intrinsic definition of this map in §5.)
7. The above described action on the fibers of $\pi^k_{k-1} : J^k(E) \to J^{k-1}(E)$ induces a map

$$\mu : F^k_{J^k(E)} \to V(J^k(E)),\ (\theta, a) \mapsto \frac{d}{dt}(\theta + ta)|_{t=0} \overset{(46)\text{ and }(36)}{=} a^j_\alpha \frac{\partial}{\partial u^j_\alpha}\bigg|_\theta \tag{47}$$

   where $|\alpha| = |i_1\cdots i_k| = k$.
8. Because of (36), $(\pi^k_{k-1})_{*,\theta} V_\theta(J^k(E)) = V_{\pi^k_{k-1}(\theta)}(J^{k-1}(E))$, i.e. the pushforward of π^k_{k-1} restricted to $V_\theta(J^k(E))$ is a surjective map whose kernel consists of the vectors tangent to the fibers of $\pi^k_{k-1} : J^k(E) \to J^{k-1}(E)$. Those vectors are precisely those contained in $\mu(F^k_{\theta\in J^k(E)})$. Therefore, we have the exact sequence of vector spaces

$$0 \longrightarrow F^k_{\theta\in J^k(E)} \xrightarrow{\mu} V_\theta(J^k(E)) \xrightarrow{(\pi^k_{k-1})_{*,\theta}} V_{\pi^k_{k-1}(\theta)}(J^{k-1}(E)) \longrightarrow 0$$

   which we can pull back to a sequence of vector bundles using (33) and (34):

$$0 \longrightarrow F^k_{J^k(E)} \xrightarrow{\mu} V(J^k(E)) \xrightarrow{(\pi^k_{k-1})_*} (\pi^k_{k-1})^* V(J^{k-1}(E)) \longrightarrow 0$$

This is an exact sequence (see also [18] or [17]) of vector bundles over $J^k(E)$.

9. Given a differential operator $\varphi : J^k(E) \supset J \to F$, we can restrict its pushforward φ_* to the vertical subbundle $V(J)$ of TJ. By definition, a differential operator is a morphism of fibered manifolds. That means, we have $\xi \circ \varphi = \pi$ (where $\pi : J^k(E) \to M$ and $\xi : F \to M$ are projections). This implies that vertical vectors of J are mapped to vertical vectors of F. Thus, we obtain a map $\varphi_* : V(J) \to V(F)$. Now define the *symbol* $\sigma(\varphi)$ (of φ) as the composition

$$\begin{aligned} \sigma(\varphi) &:= \varphi_* \circ \mu : F^k_J \to V(F) \\ p = (\theta, a) &\mapsto \varphi_{*,\theta} \left(a^j_{\sigma=i_1 \dots i_k} \frac{\partial}{\partial u^j_\sigma}\bigg|_\theta \right) = a^j_\sigma \frac{\partial \varphi^h}{\partial u^j_\sigma} \frac{\partial}{\partial w^h}\bigg|_{\varphi(\theta)} \end{aligned} \tag{48}$$

Here $h \in \{1, \cdots, \dim(F_{\pi(\theta)})\}$ and $|\sigma| = k$ because of (44), (46) and (47).

10. The l-th prolongation $\sigma^l(\varphi)$ of the symbol $\sigma(\varphi)$ of φ is defined as the composition (see [18], end of chapter IX)

$$F^{l+k}_J = S^{l+k}T^* \otimes_J V(E) \xrightarrow{\Delta_{l,k} \otimes \mathrm{id}} S^l T^* \otimes F^k_J \xrightarrow{\mathrm{id} \otimes \sigma(\varphi)} S^l T^* \otimes_F V(F).$$

In local coordinates, we can express a point $p \in F^{l+k}_J$ as a tuple $p = (\theta \in J \subset J^k(E), a \in S^{l+k}T^* \otimes V(E))$ such that

$$\begin{aligned} \sigma^l(\varphi)(p) &= \sum_{\sigma \in \mathfrak{S}_{k+l}} a^j_{\sigma(i_1) \dots \sigma(i_{l+k})} dx^{\sigma(i_1)} \vee \dots \vee dx^{\sigma(i_l)} \\ &\qquad \otimes \sigma(\varphi) \left(dx^{\sigma(i_{l+1})} \vee \cdots \vee dx^{\sigma(I_{l+k})} \otimes \frac{\partial}{\partial u^j}\bigg|_\theta \right) \\ &\overset{(a \text{ is symmetric})}{=} \sum_{\sigma \in \mathfrak{S}_{k+l}} a^j_{i_1 \dots i_{l+k}} dx^{\sigma(i_1)} \vee \dots \vee dx^{\sigma(i_l)} \\ &\qquad \otimes \frac{\partial \varphi^h}{\partial u^j_{\sigma(i_{l+1}) \dots \sigma(i_{l+k})}} \frac{\partial}{\partial w^h}\bigg|_{\varphi(\theta)} \end{aligned} \tag{49}$$

$\sigma^1(\varphi)$ is especially important, and is explicitly rewritten as follows.

$$\sigma^1(\varphi)(p) = \sum_{\sigma \in \mathfrak{S}_{1+k}} a^j_{i_1 \dots i_{1+k}} \frac{\partial \varphi^h}{\partial u^j_{\sigma(i_1) \dots \sigma(i_k)}} dx^{\sigma(i_{1+k})} \otimes \frac{\partial}{\partial w^h}\bigg|_{\varphi(\theta)} \tag{50}$$

11. The following proposition is useful for practical calculations.

Proposition 11 *The following diagram commutes.*

$$\begin{array}{ccc} S^{l+k}T^* \otimes_J V(E) & \xrightarrow{\sigma^l(\varphi)} & S^l T^* \otimes_F V(F) \\ \downarrow \mu & & \downarrow \mu \\ V(P^l(J)) & \xrightarrow{p^l(\varphi)_*} & V(J^l(F)) \end{array}$$

Locally, one can thus use (52) with $|\alpha| = l$ instead of (49).

Proof If one applies μ to Eq. (49), one obtains

$$\mu(\sigma^l(\varphi)(p)) = \sum_{\sigma \in \mathfrak{S}_{k+l}} a^j_{\sigma(i_1)\ldots\sigma(i_{l+k})} \frac{\partial \varphi^h}{\partial u^j_{\sigma(i_{l+1})\ldots\sigma(i_{l+k})}} \frac{\partial}{\partial w^h_{\sigma(i_1)\ldots\sigma(i_l)}}\bigg|_{p^l(\varphi)(\theta)} \tag{51}$$

One can furthermore define the composition[9]

$$F^{l+k}_J = S^{l+k}T^* \otimes_J V(E) \xrightarrow{\quad\mu\quad} V(P^l(J)) \xrightarrow{\quad p^l(\varphi)_*\quad} V(J^l(F)).$$

In local coordinates, this means, for $|\sigma| = l + k,\ 0 \leq |\alpha| \leq l$, that

$$\begin{aligned} p^l(\varphi)_* \circ \mu : F^{l+k}_J &\to V(J^l(F)) \\ p = (\theta, a) \mapsto p^l(\varphi)_{*,\theta} \left(a^j_{\sigma=i_1\ldots i_{l+k}} \frac{\partial}{\partial u^j_\sigma}\bigg|_\theta \right) &= a^j_\sigma \frac{\partial(D_\alpha \varphi^h)}{\partial u^j_\sigma} \frac{\partial}{\partial w^h_\alpha}\bigg|_{p^l(\varphi)(\theta)} \end{aligned} \tag{52}$$

Note that, if $\varphi : J^k(E) \supset J \to F$ is a differential operator of order k, then it involves at most coordinates u^j_β with $0 \leq |\beta| \leq k$. As a consequence $\partial D_\alpha \varphi^h / \partial u^j_\sigma$ for $|\sigma| = l + k$ must be zero for all $0 \leq |\alpha| < l$. Hence, to obtain the non-zero components of $p^l(\varphi)_* \circ \mu$, it suffices to calculate (52) for $|\alpha| = l$.

As all terms of $D_\alpha \varphi^h$ in (52) vanish if they are not highest order, let us calculate what is left of $D_\alpha \varphi^h$ if we only look at its highest order terms. Suppose that φ is a differential operator of order k, then

$$\begin{aligned} D_\alpha \varphi^h = D_{\alpha_1\ldots\alpha_{l-1}} D_{\alpha_l} \varphi^h &\overset{(7)}{=} D_{\alpha_1\ldots\alpha_{l-1}} \left(\frac{\partial \varphi^h}{\partial x^{\alpha_l}} + \cdots + \frac{\partial \varphi^h}{\partial u^j_\theta} u^j_{\theta\alpha_l} \right) \quad \text{with } |\theta| = k \\ &\xrightarrow{\text{(highest order)}} \frac{\partial \varphi^h}{\partial u^j_\theta} D_{\alpha_1\ldots\alpha_{l-1}} u^j_{\theta\alpha_l} = \frac{\partial \varphi^h}{\partial u^j_\theta} u^j_{\theta\alpha_1\ldots\alpha_l} = \frac{\partial \varphi^h}{\partial u^j_\theta} u^j_{\theta\alpha} \end{aligned} \tag{53}$$

This means the calculation of terms of order $k + l$ of $D_\alpha \varphi^h$ only involves derivatives of φ^h of order k. Thus,

$$\frac{\partial(D_\alpha \varphi^h)}{\partial u^j_\sigma} = \frac{\partial \varphi^h}{\partial u^k_\theta} \frac{\partial u^k_{\theta\alpha}}{\partial u^j_\sigma} \tag{54}$$

[9] To recall the definition of $P^l(J)$, see Eq. (6).

Now $\partial u^k_{\theta\alpha}/\partial u^j_\sigma = 1$ only if $k = j$ and $\theta_1 \cdots \theta_k \alpha_1 \cdots \alpha_l = \sigma_1 \cdots \sigma_{l+k}$ *or any permutation thereof.*

Therefore, when summing over everything, one obtains

$$
\begin{aligned}
& a^j_\sigma \frac{\partial (D_\alpha \varphi^h)}{\partial u^j_\sigma} \frac{\partial}{\partial w^h_\alpha}\bigg|_{p^l(\varphi)(\theta)} \\
&\quad = \sum_{\sigma \in \mathfrak{S}_{k+l}} a_{\sigma(i_1)\ldots\sigma(i_{k+l})} \frac{\partial \varphi^h}{\partial u^j_{\sigma(i_1)\ldots\sigma(i_k)}} \frac{\partial}{\partial w^h_{\sigma(i_{k+1})\ldots\sigma(i_{k+l})}}\bigg|_{p^l(\varphi)(\theta)}
\end{aligned}
\tag{55}
$$

As the sum goes through all permutations, this is equivalent to Eq. (51). Thus, we obtain $\mu \circ \sigma^l(\varphi) = p^l(\varphi)_* \circ \mu$. $\square$

12. Given a differential equation $\mathcal{E}$, define

$$
g^k := V(\mathcal{E}) \cap \mu(F^k_{\mathcal{E}}) \tag{56}
$$

and also call it the *symbol* (of $\mathcal{E}$). It's l-th prolongation is defined as

$$
g^{k+l} := (S^l T^* \otimes_{\mathcal{E}} V(\mathcal{E})) \cap F^{l+k}_{\mathcal{E}} \tag{57}
$$

If a differential operator $\varphi : J \to F$ is given such that (11) holds, [17] shows that the symbol of $\mathcal{E} := \ker_s(\varphi)$ and its l-th prolongation are given by

$$
g^k = \ker(\sigma(\varphi))|_{\mathcal{E}}, \qquad g^{k+l} = \ker(\sigma^l(\varphi))|_{\mathcal{E}}. \tag{58}
$$

Set $g^{k+l} = F^{k+l}_{\mathcal{E}}$ for $l < 0$ and $F^{-1}_{\mathcal{E}} = 0$.

13. Define a map $\delta : S^{1+k}T^* \to T^* \otimes S^k T^*$ by setting $\delta = \Delta_{1,k}$ (see (42)). Then extend this map by letting the same letter δ denote the map

$$
\begin{aligned}
& \delta : T^* \otimes S^k T^* \to \Lambda^2 \otimes S^{k-1} T^* \\
& dx^{h_1} \otimes dx^{i_1} \vee \cdots \vee dx^{i_k} \mapsto (-1) dx^{i_1} \wedge \Delta_{1,k-1}(dx^{i_1} \vee \cdots \vee dx^{i_k})
\end{aligned}
\tag{59}
$$

Now let n be any natural number and $w \in \Lambda^j$ and extend the map again as follows:

$$
\begin{aligned}
& \delta : \Lambda^j \otimes F^n_Y \to \Lambda^{j+1} \otimes F^{n-1}_Y \\
& w \otimes dx^{i_1} \vee \cdots \vee dx^{i_n} \otimes \frac{\partial}{\partial u^l} \mapsto (-1)^j w \wedge \Delta_{1,n-1}\left(dx^{i_1} \vee \cdots \vee dx^{i_n}\right) \otimes \frac{\partial}{\partial u^l}.
\end{aligned}
\tag{60}
$$

If we set $S^l T^* = 0$ for $l < 0$, one can now use this map δ to obtain the sequence

$$
\begin{array}{ccccccc}
0 & \longrightarrow & S^k T^* & \xrightarrow{\quad\delta\quad} & & & T^* \otimes S^{k-1} T^* \\
 & & & & & & \big\downarrow \delta \\
0 & \longleftarrow & \Lambda^m \otimes S^{k-m} T^* & \xleftarrow{\delta} & \ldots & \xleftarrow{\delta} & \Lambda^2 \otimes S^{k-2} T^*
\end{array}
$$

(where $m = \dim(M)$.) This sequence is exact (see [17, Lemma 6.1]).

As $g^{k+l} \subset F_{\mathcal{E}}^{k+l}$ and $\delta(g^n) \subset T^* \otimes_{\mathcal{E}} g^{n-1}$, the above map (60) also gives rise to the sequence

$$\begin{array}{ccccccc} 0 & \longrightarrow & g^n & \xrightarrow{\delta} & & & T^* \otimes_{\mathcal{E}} g^{n-1} \\ & & & & & & \downarrow \delta \\ \Lambda^{n-k+1} \otimes_{\mathcal{E}} F_{\mathcal{E}}^{k-1}. & \longleftarrow & \Lambda^{n-k} \otimes_{\mathcal{E}} g^k & \xleftarrow{\delta} & \dots & \xleftarrow{\delta} & \Lambda^2 \otimes_{\mathcal{E}} g^{n-2} \end{array}$$

The cohomology groups of this sequence are denoted by

$$H^{n,j} := \frac{\ker(\delta : \Lambda^j \otimes_{\mathcal{E}} g^n \ \to \ \Lambda^{j+1} \otimes_{\mathcal{E}} g^{n-1})}{\mathrm{Im}(\delta : \Lambda^{j-1} \otimes_{\mathcal{E}} g^{n+1} \ \to \ \Lambda^j \otimes_{\mathcal{E}} g^n)} \tag{61}$$

and are called *Spencer cohomology* groups.

One says that

$$g^k \text{ is } r-\textit{acyclic} \text{ if } H^{n,j} = 0 \text{ for all } n \geq k \text{ and } 0 \leq j \leq r \tag{62}$$

and that g^k is *involutive* if

$$g^k \text{ is } \infty\text{-acyclic, i.e. if } H^{n,j} = 0 \,\forall\, n \geq k,\ \ j \geq 0. \tag{63}$$

14. Finally, one needs the notion of a *quasi-regular basis*. To this end, define the space

$$S^{k,j}T^* := \left\{ \mathrm{span}(dx^{i_1} \vee \cdots \vee dx^{i_k}) | j+1 \leq i_1 \leq \cdots \leq i_k \leq m = \dim(M) \right\} \tag{64}$$

Its dimension can be calculated as before by counting the number of possibilities of putting k balls between $m-(j+1)$ sticks. The result is

$$\dim(S^{k,j}T^*) = \binom{m-j-1+k}{k} \tag{65}$$

Using this definition, define the k, j-fiber

$$F_Y^{k,j} := S^{k,j}T^* \otimes_{\mathcal{E}} V(E) \tag{66}$$

and use this to define the k, j-symbol and its prolongation

$$g^{k,j} := g^k \cap F_{\mathcal{E}}^{k,j}, \qquad g^{k+l,j} := g^{k+l} \cap F_{\mathcal{E}}^{k+l,j}. \tag{67}$$

If $\varphi : J \ \to \ F$ is a differential operator such that (11) holds, then the k, j-symbol of φ and its prolongation are defined as the restrictions of $\sigma(\varphi)$ and

$\sigma^l(\varphi)$ to $F_J^{k,j}$ and $F_J^{l+k,j}$. Explicitly, we have

$$\begin{aligned}
\sigma(\varphi)^j &:= \sigma(\varphi)|_{F_J^{k,j}} : F_J^{k,j} \rightarrow V(F), \\
\sigma^l(\varphi)^j &:= \sigma^l(\varphi)|_{F_J^{l+k,j}} : F_J^{l+k,j} \rightarrow V(F), \\
\Rightarrow \quad g^{k,j} &= \ker(\sigma(\varphi)^j)|_{R_k}, \qquad g^{l+k,j} = \ker(\sigma^l(\varphi)^j)|_{\mathcal{E}}.
\end{aligned} \tag{68}$$

Now say that a basis $\{\partial_1, \cdots, \partial_m\}\,|_{x\in M}$ of $T_x M$ is quasi-regular for g^k at $p \in \mathcal{E}$ if[10]

$$\dim(g_p^{k+1}) = \dim(g_p^k) + \sum_{j=1}^{m-1} \dim(g_p^{k,j}). \tag{69}$$

And say that there is a quasi-regular basis for g^k if there is a quasi-regular basis for g^k at every $p \in \mathcal{E}$.

5.2 *Formal Theory*

Now with all definitions at hand, we can proceed with a motivation for the definition of formal integrability. Given a differential equation $\mathcal{E}$, one would like to find its solutions. In general, solutions around a point are difficult to find. Recall that a horizontal solution can be described by a section $s : M \supset U \rightarrow E$ such that $j^k(s)(U) \subset \mathcal{E}$. If a section fulfills this property and it is smooth, then its prolongations also fulfill the prolonged equations, i.e. $j^{k+l}(s)(U) \subset P^l(\mathcal{E})$. In particular, this means, if s is a solution and one chooses a fixed $x \in U$, then it holds true that

$$j^{k+l}(s)(x) \in P^l(\mathcal{E}) \text{ for all } l \geq 0. \tag{70}$$

Thus, (70) is a necessary condition for the existence of a smooth solution $s : U \rightarrow E$.

A point $\theta \in \mathcal{E}$ is called a solution of order k at $x = \pi(\theta)$. It is called a solution of order k at x because by Borel's lemma, one can always find a section s that fulfills $j^k(s)(x) = \theta$. However, this section does not necessarily fulfill the condition (70).

Therefore, given a solution $\theta \in \mathcal{E}$ of order k, one wishes to check if there exists a section such that (70) holds. If this condition holds at $x = \pi(\theta)$, then one says that $\mathcal{E}$ has a formal solution at θ. If one can find formal solutions at all points of $\theta \in \mathcal{E}$, then one says that $\mathcal{E}$ is *formally integrable*.

---

[10] The condition on g^k locally imposes a condition on the dual basis and thus also on the basis.

As higher derivatives are promoted to coordinates in the jet bundle approach, $\mathcal{E}$ is usually the kernel of an algebraic (most often polynomial) equation. Therefore, *to find solutions of order k is comparatively easy* because it does not involve any analysis but algebraic operations are sufficient.

Finally, suppose that a formal solution consisting of a section s that fulfills (70) at the point $x = \pi(\theta)$ is given. Then the section s we have found is precisely the section whose taylor expansion is equal to the expansion whose coefficients are $j^l(s)(x)$. This taylor expansion does not necessarily converge. It may also happen that it does only converge at x and in no neighbourhood of x. Therefore, it is not necessarily a solution of $\mathcal{E}$ in the usual sense.

However, suppose that it does converge in a neighbourhood of x, then it is a smooth solution of $\mathcal{E}$. In general, it is possible to show that a formal solution always converges if one works in the analytic category where all functions are locally given by a converging taylor expansion. Therefore, in this category, formal integrability is also a sufficient condition for the existence of (local) solutions.

To motivate the precise definition of formal integrability, note that the requirement that any solution of order k can be extended to a solution of infinite order can only be fulfilled if the prolongation of any order of the equation does not impose new constraints on the coordinates of the solution up to order k ("new constraints" means new equations involving coordinates up to order k which are not equivalent to the equations one started with). For suppose we started with a solution of order k that did not fulfill those constraints, then this solution could not be extended to a solution of the order which imposes those constraints.

If no new constraints are imposed on the coordinates of order k by the prolongation, this means geometrically that $P^l(\mathcal{E})$ is a surface which can be given local coordinates that agree with those of $\mathcal{E}$ up to order k. Then,

$$\pi^{k+l+1}_{k+l} : P^{l+1}(\mathcal{E}) \;\to\; P^l(\mathcal{E}) \qquad \text{is surjective for all } l \geq 0. \tag{71}$$

One might define formal integrability using just this condition. However, in most cases one would like to work in the smooth category in order to find out if smooth solutions exist for some equation. This requires us to impose an additional smoothness condition. To ask if a smooth solution exists given some k-th order solution is equivalent to asking whether the prolongation is smooth to all orders. As a solution of order $k+l$ is a section such that $j^{k+l}(s)(U) \subset P^l(\mathcal{E})$ (with $U \subset M$), smoothness of the section can only be guaranteed if $P^l(\mathcal{E})$ is a smooth submanifold of $J^{k+l}(E)$. Goldschmidt shows in [17, proposition 7.1] that $\pi^{1+k}_k : P^1(\mathcal{E}) \to \mathcal{E}$ is a smooth fibered submanifold of $\pi^{1+k}_k : J^{1+k}(E)|\mathcal{E} \to \mathcal{E}$ if and only if g^{1+k} (defined in (57)) is a vector bundle over $\mathcal{E}$ and $\pi^{1+k}_k : P^1(\mathcal{E}) \to \mathcal{E}$ is surjective. Those considerations motivate the following definition:

Definition 6 A differential equation $\mathcal{E}$ is said to be *formally integrable* if

1. $\pi^{k+l+1}_{k+l} : P^{l+1}(\mathcal{E}) \to P^l(\mathcal{E})$ is surjective,
2. g^{k+l+1} is a vector bundle

for all $l \in \{0, 1, 2, \cdots\}$.

The above definition requires to check an infinite amount of conditions. Goldschmidt proved a theorem that facilitates to determine formal integrability in a finite amount of steps. It is based on theorem 8.1 of [17] which we cite here:

Proposition 12 *If $\mathcal{E}$ is a differential equation, then it is formally integrable if and only if*

1. $\pi_k^{k+1} : P^1(\mathcal{E}) \rightarrow \mathcal{E}$ *is surjective,*
2. g^{k+1} *is a vector bundle over* $\mathcal{E}$,
3. g^k *is 2-acyclic.*

Recall that g^k is 2-acyclic (cf. (62)) if the Spencer cohomology groups $H^{n,j}$ (see (61)) vanish for all $n \geq k$ and $0 \leq j \leq 2$. However, [17, Lemma 6.2] also proves that g^k is always 1-acyclic, i.e. $H^{n,j} = 0$ for all $n \geq k$ and $0 \leq j \leq 1$. Therefore, one can replace the last condition by the requirement that

$$H^{n,2} = 0 \text{ for all } n \geq k. \tag{72}$$

This still seems to require an infinite number of calculations. However, [17, Lemma 6.4] shows the following.

Proposition 13 *If the dimension of $V_{\pi(\theta)}(E)$ does not depend on $\theta \in \mathcal{E} \subset J^k(E)$, then there exists an integer $k_0 > k$ depending only on $dim(M)$ and k and $dim V_{\pi(\theta)}(E)$ such that g^{k_0} is involutive, i.e. that g^{k_0} is ∞-acyclic, i.e. $H^{n,j} = 0\ \forall\, n \geq k_0,\ j \geq 0$.*

Similarly, [17, Lemma 6.4 and proposition 7.2] is used to prove [17, theorem 8.2] which reads

Proposition 14 *If the dimensions of all components of E are the same and $\mathcal{E} \subset J^k(E)$ is a differential equation, then there exists an integer $k_0 > k$ depending only on $dim(M)$ and k and $dim(E)$ such that $\mathcal{E}$ is formally integrable if and only if*

1. $\pi_{k+l}^{k+1+l} : P^{1+l}(\mathcal{E}) \rightarrow P^l(\mathcal{E})$ *is surjective,*
2. g^{k+1+l} *is a vector bundle over* $\mathcal{E}$,

for all $0 \leq l \leq k_0 - k$.

The last two propositions in particular imply that whenever the dimension of E is constant (which is the case in most applications, where one often chooses $E = \mathbb{R}^m \times \mathbb{R}^e$ or some other manifold with constant dimension), then it must be possible to determine whether $\mathcal{E}$ is formally integrable in **finitely many steps**. This means that one actually does not have to compute the infinitely many cohomology groups appearing in (72). Nevertheless, Proposition 14 does not tell us how large this finite k_0 might be. In general, there does not seem to be a simple way to estimate this, which can be problematic. However, it turns out that one can prove stronger statements about the stronger condition of ∞-acyclicity/involutivity (defined in (63)). In [18, theorem 2.14 in Chapter IX] (according to them, going back to Serre), they state:

Proposition 15 *The following conditions are equivalent:*

1. *There exists a quasi-regular basis (cf. (69)) of g^k at $\theta \in \mathcal{E}$,*
2. *g^k is involutive at θ, i.e. $H_\theta^{n,j} = 0\ \forall n \geq k,\ j \geq 0$.*

This means, if there is a quasi-regular basis, then g^k is ∞-acyclic and therefore also 2-acyclic. Hence, combining Proposition 12 with the last proposition, we obtain

Proposition 16 *If $\mathcal{E}$ is a differential equation, then it is formally integrable if*

1. *$\pi_k^{k+1} : P^1(\mathcal{E}) \rightarrow \mathcal{E}$ is surjective,*
2. *g^{k+1} is a vector bundle over $\mathcal{E}$,*
3. *There exists a quasi-regular basis for g^k.*

Definition 7 A PDE $\mathcal{E}$ is called *involutive* if and only if it satisfies the conditions of Proposition 16.

As a corollary, an involutive equation is also formally integrable but the converse is not true (because 2-acyclicity does not imply ∞-acyclicity). Indeed there are examples of equations that are formally integrable but not involutive. Thus, though the above Proposition 16 is more readily used in practice than Propositions 12 or 14, it only provides a sufficient but not a necessary condition for formal integrability.

Remark An extensive treatment including possible subtleties of involution and formal integrability can be found in [19].

The above Propositions 14 and 16 are the central propositions of this subsection. In practice, one can use them to determine formal integrability and involutivity in finitely many steps. In actual calculations of the rank of g^k (which is necessary for validating condition 3 of Proposition 16), it may happen that one must determine the rank of a larger matrix. As written above, the code for a small program computing it can be found in [43] but much more sophisticated algorithms are provided in [19].

Given formal integrability of an equation $\mathcal{E}$, it becomes possible to show the existence of local solutions in the analytic category as mentioned at the beginning of the subsection. The precise definition of analyticity is

Definition 8 A map is called analytic if, around any point, it can locally be defined by a convergent power series. (Note that this definition can also be applied to real functions. If the condition holds, they are called real-analytic).

A manifold is called analytic if all of its transition functions are analytic. The analytic category is defined as the category in which the objects are analytic manifolds and the morphisms are analytic maps between them.

The existence of local solutions in the analytic category is guaranteed by theorem 9.1 of [17] which is here rephrased as follows:

Proposition 17 *Suppose that $\mathcal{E}$ is a formally integrable differential equation which is analytic. Then given a point $\theta \in P^l(\mathcal{E})$ (for any $l \in \{0, 1, 2, \cdots\}$), it is possible to find an analytic section $s : U \to E$ where U is a neighbourhood of $x = \pi(\theta)$ such that $j^{k+l}(s)(x) = \theta$ and s is a local solution of $\mathcal{E}$.*

One might wonder if it is possible to prove something stronger, for example that smoothness guarantees existence of local solutions. This is not possible because of "Lewy's example", a well-known counter-example.

5.3 Integrability Conditions

When checking for formal integrability or involutivity of a system of differential equations, it may happen that the first prolongation $P^1(\mathcal{E})$ does not project surjectively to $\mathcal{E}$ via π_k^{k+1}, or that g^{k+1} is not a smooth vector bundle or that there exists no quasi-regular basis for g^k but that the PDE can become formally integrable if certain integrability conditions $\mathcal{B}(\mathcal{E})$ are *added to* $\mathcal{E}$, i.e. by defining $\mathcal{B} := \mathcal{B}(\mathcal{E}) \cap \mathcal{E}$, $\mathcal{B}$ can become formally integrable. This subsection gives a definition for $\mathcal{B}(\mathcal{E})$ that is useful for identifying minimal consistency conditions when comparing systems of differential equations and field theories.

The definition is motivated by the following example in which surjectivity fails to hold.

Example Let $\pi : E := \mathbb{R}^2 \times \mathbb{R}$ be a fibered manifold with local coordinates (x, t, u). Define $F := \mathbb{R}^2 \times \mathbb{R}^2$ and consider the differential operator

$$\Phi : J^2(E) \to F, \qquad (x, t, u, u_x, u_t, u_{xx}, u_{tt}, u_{xt}) \mapsto (x, t, u_x, u_{tt}) \tag{73}$$

Now the differential equation $\mathcal{E} = \ker_0(\Phi)$ is given by

$$\mathcal{E} = \{(x, t, u, 0, u_t, u_{xx}, 0, u_{xt})\} \tag{74}$$

The first prolongation is

$$\begin{aligned} P^1(\mathcal{E}) = \ker(\Phi^1) &= \ker\left(u_x,\ u_{tt},\ u_{xx},\ u_{xt},\ u_{ttx},\ u_{ttt}\right) \\ &= \{(x, t, u, 0, u_t, 0, 0, 0, u_{xxx}, u_{xxt}, 0, 0)\} \end{aligned} \tag{75}$$

so that $\pi_2^3(P^1(\mathcal{E})) = \ker\ \{(x, t, u, 0, u_t, 0, 0, 0)\}$ which is much smaller than $\mathcal{E}$. The reason is that due to the prolongation, there arise additional constraints on coordinates **of the order of** $\mathcal{E}$, here of second order, namely on u_{xx} and

(continued)

u_{tx}. Concretely, they are given by

$$\mathcal{B}(\mathcal{E}) : \{ u_{xx} = 0, \ u_{tx} = 0 \}. \tag{76}$$

$\mathcal{B}(\mathcal{E})$ are the *integrability or consistency conditions* of this system $\mathcal{E}$. Therefore, to restore surjectivity, one can try to include those additional constraints right from the start and define

$$\mathcal{B} := \mathcal{E} \cap \mathcal{B}(\mathcal{E}) = \ker\left(u_x, \, u_{tt}, \, u_{xt}, \, u_{xx}\right) = \{(x, t, u, 0, u_t, 0, 0, 0)\} \tag{77}$$

Now if we prolong *this* system, then the result is similar to $P^1(\mathcal{E})$ except for the fact that the additional constraints $u_{xtx} = 0$, $u_{xxt} = 0$ are additionally imposed. However, those are only new constraints on the coordinates of order 3. Therefore, the projection $\pi^3_2|_{P^1(\mathcal{B})}$ is now indeed surjective.

Furthermore, we can read off a solution from $\mathcal{B}$, namely

$$u(x, t) = At + B \tag{78}$$

which is a meaningful solution because it *also* is a solution of $\mathcal{E}$. Indeed, for reasons of consistency just shown above, those are the only solutions of $\mathcal{E}$. Therefore, the procedure to define a new system for which surjectivity is guaranteed is meaningful as long as $\mathcal{B}$ is again a PDE (in particular, it must be non-empty).

Motivated by the observations above, we define consistency/integrability conditions as follows.

Definition 9 The *integrability condition* of a given PDE $\mathcal{E} \subset J^k(E)$ is defined to be the biggest smooth submanifold $\mathcal{B}(\mathcal{E})$ of the lowest order jet space $J^l(E)$ (with $l \geq k$) such that $\mathcal{B} := \mathcal{B}(\mathcal{E}) \cap (\pi^l_k)^{-1}(\mathcal{E})$ is formally integrable (where $\pi^l_k : J^l(E) \to J^k(E)$ is the canonical projection).

If $\mathcal{B}$ is non-empty, smooth, and has a component with dimension bigger zero, it is called the *formal closure* of $\mathcal{E}$. Otherwise the formal closure of $\mathcal{E}$ does not exist and $\mathcal{E}$ is said to be *(formally) non-integrable*.

Note that the formal closure (or its non-existence) can always be computed in finitely many steps because formal integrability can be checked in finitely many steps using Proposition 14 which is of practical importance.

Furthermore, it can often be useful to attempt to compute the *involutive closure/completion* of the PDE $\mathcal{E}$ instead because the conditions of Proposition 16 are easier to check. If the PDE in question admits such a completion, one does not need to check formal integrability anymore. If it does not admit such a completion, one can still resort to checking the conditions of Proposition 14.

For the intersections of our physical theories, it might occur quite often that the intersections are formally integrable only after redefining them as systems that take the consistency conditions into account. Those consistency conditions that are automatically found when checking formal integrability are precisely *the minimal amount of assumptions that must be made in order to make the system consistent.* Therefore, they are actually really useful for us because *they can be interpreted as the minimal physical assumptions under which a correspondence becomes meaningful.*

This means that without knowing exactly what assumptions are reasonable to relate two systems, we can just define a correspondence and then find it out. This happens later in the example where magneto-statics and hydrodynamics are shown to share an intersection whose consistency conditions had to be guessed in Eq. (15) in the motivating example in Sect. 3.1.

The correspondences themselves still have to be guessed. However, symmetries can provide clues about which correspondences might be especially meaningful as explained in Sect. 8.

5.4 Explicit Example of the Application of Proposition 16

In this subsection, involutivity and thus also formal integrability for one simple example is proved using Proposition 16. Despite the simplicity of the equation, the example is very detailed to illustrate the formalism. The reader not interested in this illustration can directly continue with the next section. The reader interested in more explicit examples is referred to the author's thesis, [43].

Below, $\mathcal{E}$ and $P^l(\mathcal{E})$ are defined as kernel of a differential operator φ and its prolongation, using (10) and (12). Thus, it is possible to obtain g^k and g^{k+l} as the kernel of $\sigma(\varphi)$ and its prolongation using (58). Furthermore, $g^{k,j}$ and $g^{k+l,j}$ can be obtained using (68). First define $\pi : E \to M$ as follows:

$$M := \mathbb{R},\ E := \mathbb{R}\times\mathbb{R},\ \pi = \mathrm{pr}_1 : E \to M \text{ is the projection onto the first factor.} \tag{79}$$

Let $J := J^1(E) \simeq \mathbb{R}^3$ with local coordinates (x, u, u_x). Then define a differential operator $\varphi : J \to F := E$ by

$$\varphi(x, u, u_x) = (x, \varphi^1(x, u, u_x)) := (x, u_x - u) \tag{80}$$

which is a first order linear operator. Its kernel

$$\begin{aligned}\mathcal{E} := \ker\varphi &\overset{(10))}{=} \{\theta \in J \mid \varphi(\theta) = 0(\pi(\theta)) = (x, 0)\ \} \\ &= \{\theta \in J \mid u_x = u\} = \{(\rho, \lambda, \lambda) \mid \rho, \lambda \in \mathbb{R}\}\end{aligned} \tag{81}$$

is a first order linear differential equation corresponding to a two-dimensional subspace of J. We know it to have the general solution

$$u(x) := u(s_E(x)) = A\exp(x),\ \ A \in \mathbb{R} \tag{82}$$

but want to show formal integrability of $\mathcal{E}$ to illustrate the general methods introduced above.

To show all 3 conditions of Proposition 16, we first need to calculate $P^1(\mathcal{E})$ and $g^{k+1} = g^2$. To this end, note that the prolongation of $J = J^1(E)$ is $J^2(E) \simeq \mathbb{R}^4$ with local coordinates (x, u, u_x, u_{xx}). Thus, we can use (9) to prolong φ to obtain

$$p^1(\varphi)(\theta \in J^2(E)) = (\varphi, D_x\varphi)(\theta) = (x, u_x - u, u_{xx} - u_x) \in J^1(E). \tag{83}$$

such that

$$\begin{aligned} P^1(\mathcal{E}) = \ker p^1(\varphi) &\overset{(12)}{=} \Big\{\theta \in J^2(E) \mid (\varphi(\theta), D_x\varphi(\theta)) = 0(\pi(\theta))\Big\} \\ &= \Big\{\theta \in J^2(E) \mid u_x - u = 0,\ u_{xx} - u_x = 0\Big\} = \{(\rho, \lambda, \lambda, \lambda) \mid \rho, \lambda \in \mathbb{R}\}. \end{aligned} \tag{84}$$

Now that $P^1(\mathcal{E})$ and $\mathcal{E}$ are explicitly given, one can see that the restriction of π^2_1 to $P^1(\mathcal{E})$ **surjectively** projects down to $\mathcal{E}$. Explicitly,

$$\pi^2_1 P^1(\mathcal{E}) = \Big\{\pi^2_1(\rho, \lambda, \lambda, \lambda)\Big\} = \{(\rho, \lambda, \lambda)\} = \mathcal{E}. \tag{85}$$

This means condition 1. of Proposition 16 is fulfilled. In fact, there even is an inverse map sending (ρ, λ, λ) back to $(\rho, \lambda, \lambda, \lambda)$, so $P^1(\mathcal{E}) \simeq \mathcal{E}$. This continues for higher orders. We have

$$p^l(\varphi) = (\varphi, D_x\varphi, \cdots, D^l_x\varphi) \ \ \Rightarrow \ \ P^l(\mathcal{E}) = \{\ (\rho, \lambda, \cdots, \lambda) \mid \rho, \lambda \in \mathbb{R}\ \} \ \simeq \ \mathcal{E}. \tag{86}$$

Now let us calculate g^1 and g^2. To do so, we must first calculate the symbol of φ. To do this, we must first clarify how an element $p \in F^1_{\mathcal{E}}$ looks like. This can be done using (45). Note that our manifold $M = \mathbb{R}$ is one dimensional and therefore T^*M has basis dx while $V(E)$ has basis $\partial/\partial u$. Thus,

$$F^1_{\mathcal{E}} \ni p = (\theta, a) = \left((x, u, u),\ \ \left(a^1_1 dx \otimes \frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)}\right)\right). \tag{87}$$

As a consequence $\dim(F^1_{\mathcal{E}})_{\theta\in\mathcal{E}} = 1$ and we obtain

$$\sigma(\varphi)(p) \overset{(48)}{=} a^1_1 \frac{\partial\varphi^1}{\partial u_x}\frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)} \overset{(80)}{=} a^1_1 \frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)} \tag{88}$$

$$
\begin{aligned}
g^1 &\overset{(58)}{=} \ker(\sigma(\varphi))|_{\mathcal{E}} \overset{(88)}{=} \left\{ p \in F^1_{\mathcal{E}} \,\middle|\, \theta \in \mathcal{E} \text{ and } a^1_1 \frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)} = 0 \right\} \\
&= \{(\theta, 0) \mid \theta \in \mathcal{E}\} \;\simeq\; \mathcal{E}
\end{aligned}
\tag{89}
$$

This shows that g^1 is the trivial vector bundle over $\mathcal{E}$ whose fibers consist of the zero-point only. Similarly,

$$
F^2_{\mathcal{E}} \ni p = (\theta, a) = \left((x, u, u), \; \left(a^1_{11} dx \vee dx \otimes \frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)} \right) \right). \tag{90}
$$

whose fibers are also one-dimensional and therefore

$$
\sigma^1(\varphi)(p) \overset{(49)}{=} 2a^1_{11} \frac{\partial \varphi^1}{\partial u_x} \frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)} \overset{(80)}{=} 2a^1_{11} \frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)} \tag{91}
$$

such that

$$
\begin{aligned}
g^2 &\overset{(58)}{=} \ker(\sigma^1(\varphi))|_{\mathcal{E}} \overset{(91)}{=} \left\{ p \in F^2_{\mathcal{E}} \mid \theta \in \mathcal{E} \text{ and } a^1_{11} \frac{\partial}{\partial u}\bigg|_{\pi_0(\theta)} = 0 \right\} \\
&= \{(\theta, 0) \mid \theta \in \mathcal{E}\} \simeq g^1 \;\simeq\; \mathcal{E}
\end{aligned}
\tag{92}
$$

As a consequence, g^2 is also a trivial *vector bundle over* $\mathcal{E}$. This proves that condition 2. of Proposition 16 is fulfilled. In fact, one can see that $g^{1+l} \simeq g^1$ for all l. As a consequence, we do not even need to test condition 3 of proposition 5 because this together with $P^l(\mathcal{E}) \simeq \mathcal{E}$ directly shows that the Definition 6 of formal integrability is fulfilled.

Nevertheless, let us test condition 3 of Proposition 16 explicitly. To this end, we must check condition (69) for all $p \in \mathcal{E}$. To do so, we must calculate $g^{1,j}$. However, the definition of $S^{k,j}T^*$ (see (64)) requires that $j + 1 \leq i_1 \leq \cdots \leq i_k \leq m = 1$ which is only possible for $j = 0$, i.e. $S^{k,j>0}T^* = 0$. But the sum in (69) only goes from $j = 1$ to $j = m - 1 = 0$. As a consequence, we only have to verify that

$$
\dim(g^2) = \dim(g^1). \tag{93}
$$

This does hold because $g^2 \simeq g^1$ as shown above. This shows that all conditions of Proposition 16 are satisfied and our Eq. (81) is involutive and thus formally integrable.

6 Shared Structure

6.1 Definition

Now that the notions of intersection and correspondence are developed and that the theory of formal integrability has been reviewed, everything can be combined to define what it means for two theories to share structure.

So suppose we are given two fibered manifolds $\pi : E \to M$ and $\xi : F \to M$ and would like to compare the differential equations $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$. Consider $\mathcal{EF}$ as defined in Eq. (32) which is a PDE by Proposition 10. However, given a correspondence Φ, Corollary 1 shows that $\mathcal{EF} \cap \Phi$ is only a PDE if $\mathcal{EF} \pitchfork \Phi$. Thus, the following definition is useful.

Definition 10 $\mathcal{E}$ and $\mathcal{F}$ *share an intersection* $\mathcal{I} := \mathcal{EF} \cap \Phi$ *(under the correspondence* Φ*)* if $\mathcal{EF} \pitchfork \Phi$.

Now let us suppose that $\mathcal{E}$ and $\mathcal{F}$ do share an intersection under Φ. From the discussion in Sect. 4.2, it is clear that sharing an intersection is not enough for saying that two theories share structure in a meaningful way. Instead, one should require that the system is differentially consistent/formally integrable as well.

Definition 11 (Shared Structure) Two differential equations $\mathcal{E}$ and $\mathcal{F}$ *share structure* if they share an intersection $\mathcal{I}$ that has a formal closure $\mathcal{B}$ (in the sense of Definition 9).

Note that if only an open subset of $\mathcal{I}$ has a formal closure, then one can always restrict Φ such that $\mathcal{I}'$ has a formal closure.

This definition is meaningful because formal integrability guarantees that all N-th order solutions on an open subset of $\mathcal{B}$ can be prolonged to formal solutions. As explained in Sect. 5.2, those N-th order solutions can be constructed very easily by defining a Taylor expansion using as coefficients the entries of any point $\theta = (x^i, u^j_\alpha, v^g_\beta)$ in this open subset of $\mathcal{B}$. So if two differential equations share structure, then this usually means that the formal closure of their intersection has a lot of formal solutions and in this case the corresponding theories have quite a lot in common.

Given the geometric theory of shared structure, one can also obtain a natural notion of equivalence of PDEs. Most canonically, equivalence is perhaps defined as follows.

Definition 12 Two systems of PDEs $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$ are said to be *equivalent* if there exists a diffeomorphism $L : \mathcal{E} \to \mathcal{F}$ that preserves the Cartan distribution, i.e.

$$dL_\theta(C_\theta \cap T_\theta\mathcal{E}) = C_{L(\theta)} \cap T_{L(\theta)}\mathcal{F} \quad \forall\theta \in \mathcal{E}. \tag{94}$$

This diffeomorphism is very similar to a Lie transformation that is used to define a symmetry, below in Sect. 8. However, to integrate the above definition into the product bundle setting defined above, one could define equivalence also as follows.

Definition 13 Two systems of PDEs $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$ are said to be *equivalent* if there exists a correspondence Φ and intersection $\mathcal{I} = \Phi \cap \mathcal{EF}$ s.t. $\pi_E|_{\mathcal{I}}$ and $\pi_F|_{\mathcal{I}}$ map $\mathcal{I}$ diffeomorphically onto $\mathcal{E}$ and $\mathcal{F}$ and preserve the Cartan distribution.

This second definition is equivalent to the first. It might look somewhat more convoluted but the product space is then a more suitable setting for investigating relationships that are weaker than equivalence, such as shared subsystems in the form of shared structure that still allow for the transfer of some shared solutions, as demonstrated by the numerous constructions of the subsequent subsections.

6.2 Solution Transfer

In this subsection, we assume that the intersection $\mathcal{I}$ of two differential equations $\mathcal{E}$ and $\mathcal{F}$ is itself a differential equations with solutions and investigate the relationship between those solutions and the solutions of $\mathcal{E}$ and $\mathcal{F}$. Recall that a solution S of $\mathcal{I}$ is a locally maximal $\dim(M)$-dimensional integral submanifold of C with $S \subset \mathcal{I}$ as described in item 16. Let $J := J^k(E) \times_M J^l(F)$ and $\Pi := \pi^k \times_M \xi^l : J \to M$, the natural projection to the base space. A submanifold $S \subset J$ is called *horizontal* if $d\Pi|_\theta : T_\theta S \to T_{\Pi(\theta)} M$ is injective for all $\theta \in S$.

Proposition 18 *If $\mathcal{E}$ and $\mathcal{F}$ share the intersection $\mathcal{I} = \mathcal{EF} \cap \Phi$ (where $\mathcal{EF} = \pi_E^{-1}(\mathcal{E}) \cap \pi_F^{-1}(\mathcal{F})$) and $\mathcal{I}$ has a horizontal solution S, then $\pi_E(S)$ is a solution of $\mathcal{E}$ and $\pi_F(S)$ is a solution of $\mathcal{F}$.*

Proof In order to verify that $S_E := \pi_E(S)$ is a solution of $\mathcal{E}$, we must verify that

1. $S_E \subset \mathcal{E}$,
2. S_E is a smooth submanifold with dimension $m = \dim(M)$,
3. S_E is an integral submanifold of C, i.e. $T_\theta S_E \subset C_\theta$ for all $\theta \in S_E$,
4. S_E is a locally maximal integral submanifold.

Since $S \subset \mathcal{I} = \mathcal{EF} \cap \Phi$, we have in particular $S \subset \mathcal{EF} = \mathcal{E}_J \cap \mathcal{F}_J$ and $S \subset \mathcal{E}_J = \pi_E^{-1}(\mathcal{E}) = \{\theta \in J^k(E) \times_M J^l(F) \mid \pi_E(\theta) \in \mathcal{E}\}$. Thus, $\pi_E(S) \subset \mathcal{E}$ and the first item is verified.

As S is horizontal, it can locally be described as the image of the prolongation, $U_S := \mathrm{im}(j^k(s_E) \times_M j^l(s_F))$ of a local section $s = s_E \times_M s_F : U \subset M \to E \times_M F,\ x^i \mapsto (x^i, (s_E)^j(x), (s_F)^g(x))$. The prolongation and hence also S can locally be described by the tuple $(x^i, (s_E)^j_\alpha(x), (s_F)^g_\beta(x))$. As a consequence, $\pi_E(S_U)$ has the local description $(x^i, (s_E)^j_\alpha(x))$ on $J^k(E)$. At each point of $\pi_E(S_U)$, one can thus define $m = \dim(M)$ tangent vectors, the n-th of which is given by

$$v_n := \frac{\partial x^i}{\partial x^n}\frac{\partial}{\partial x^i} + \sum_{j=1}^{e}\sum_{|\alpha|<k}\frac{\partial (s_E)^j_\alpha(x)}{\partial x^n}\frac{\partial}{\partial u^j_\alpha} \tag{95}$$

They are all non-zero and linearly independent because $\partial x^i/\partial x^n = \delta^{in}$. Thus, $\dim(\pi_E(S)) \geq \dim(M)$. However, since $\dim(\pi_E(S)) \leq \dim(S) = \dim(M)$, we obtain $\dim(\pi_E(S_U)) = \dim(M)$. Hence, $d\pi_E|_S$ is locally a bijection and the vectors defined in (95) span the tangent space around a generic point of S_E. Furthermore, since π_E is smooth, $\pi_E|_S$ is also smooth. Therefore, $\pi_E|_S$ is a local diffeomorphism.

Now suppose we have two local neighbourhoods $O, O' \subset S_E = \pi_E(S)$ which are such that $O \cap O' \neq \varnothing$. Then, since π_E is a local diffeomorphism, we obtain corresponding open subsets $U = \pi_E^{-1}(O)$ and $U' = \pi_E^{-1}(O')$ in S. Furthermore, $\pi_E^{-1}(O \cap O') = \pi_E^{-1}(O) \cap \pi_E^{-1}(O') = U \cap U'$ because inverse images always preserve intersections. Then, since S is a smooth manifold, we also have a smooth transition map $\varphi : U \to U'$. As a consequence, since composition of smooth maps are smooth, $\pi_E \circ \varphi \circ \pi_E^{-1}|_{O\cap O'} : O \cap O' \to O'$ is a smooth transition map on S_E. Therefore, all local pieces $\pi_E(U)$ coming from the local pieces $U \subset S$ of the solution S piece together to form a global smooth, $\dim(M)$-dimensional submanifold $\pi_E(S)$ of $\mathcal{E}$. This verifies the second item.

To show that S_E is an integral submanifold of C, it suffices to show that the tangent vectors (95) that locally span the tangent space of S_E are annihilated by the Cartan forms $w^j_\alpha = du^j_\alpha - u^j_{\alpha i}dx^i$. Indeed we immediately obtain $w^j_\alpha|_{(x^i,u^j_\alpha)=(x^i,s^j_\alpha(x))}(v_n) = 0$ which verifies the third item. Since S_E is already $\dim(M)$ dimensional, no open subset of it can be embedded into a solution of higher dimension which implies the fourth item. Thus, S_E is a solution of $\mathcal{E}$.

Since the above did not make any assumptions about $\mathcal{E}$ which are not shared by $\mathcal{F}$, the same conclusion also holds for $\mathcal{F}$ and $\pi_F(S)$ is a solution of $\mathcal{F}$. □

In the case of non-horizontal / singular solutions, one has to be a bit more careful. In that case, not all solutions are projected to smooth submanifolds via π_E and π_F.

Example Consider $\pi : E \to M$ with $M := \mathbb{R}$, $E := M \times M$ and π the projection to the first factor and consider another, identical bundle $\xi : F \to M$. Assume we are given the differential equations $\mathcal{E} \subset J^1(E)$ and $\mathcal{F} \subset J^1(F)$ described by

$$\mathcal{E} : \{x^2 + u_x^2 = 1\}, \qquad \mathcal{F} : \{v_x^2 = 2v + \frac{1}{2}\}. \tag{96}$$

(continued)

Note that the solution of $\mathcal{E}$ is singular because the smooth integral submanifold described by

$$\mathbf{x} = \begin{pmatrix} x = \sin(2t) \\ u = t + \frac{1}{4}\sin(4t) \\ u_x = \cos(2t) \end{pmatrix}, \qquad \mathbf{v} := \frac{\partial \mathbf{x}}{\partial t} = \begin{pmatrix} 2\cos(2t) \\ 1 + \cos(4t) \\ -2\sin(2t) \end{pmatrix} \tag{97}$$

gives rise to a section $s : \mathbb{R} \to E,\ t \mapsto (x(t), s(x(t)))$ with singular points at $x = x(t = \pi/4 + n\pi),\ n \in \mathbb{Z}$ because $d\pi_0^1|_{t=(\pi/4+n\pi)}(\mathbf{v}) = \mathbf{0}$ (where $\pi_0^1 : J^1(E) \to E$). The Eq. $\mathcal{E}$ and its singular solution were described in a talk by Luca Vitagliano, in relation to the publication [26].

In the present example, the aim is to illustrate how such singular solutions relate to the notion of a correspondence. To this end, define such a correspondence between $\mathcal{E}$ and $\mathcal{F}$ on $J := J^1(E) \times_M J^1(F)$ by

$$\Phi : \{u_x^2 + v_x^2 = 1\}. \tag{98}$$

Then one solution S of the submanifold $\mathcal{I} = \mathcal{E} \cap \mathcal{F} \cap \Phi$ is described by

$$\mathbf{x} = \begin{pmatrix} x = \sin(2t) \\ u = t + \frac{1}{4}\sin(4t) \\ u_x = \cos(2t) \\ w = -\frac{1}{4}\cos(4t) \\ w_x = \sin(2t) \end{pmatrix}, \ \mathbf{v} := \frac{\partial \mathbf{x}}{\partial t} = \begin{pmatrix} 2\cos(2t) \\ 1 + \cos(4t) \\ -2\sin(2t) \\ \sin(4t) \\ 2\cos(2t) \end{pmatrix},$$

$$\mathbf{v}(t = \pi/4 + n\pi) = \begin{pmatrix} 0 \\ 0 \\ -2 \\ 0 \\ 0 \end{pmatrix}$$

In the present situation, we obtain $d\pi_F|_{t=(\pi/4+n\pi)}(\mathbf{v}) = \mathbf{0}$. This means that $\pi_F(S)$ is not a smooth manifold because it contains singular points. However, after removing those, $\pi_F(S)$ becomes smooth (but disconnected).

Note that $\pi_E(S)$ is, however, a smooth submanifold even though it is, by definition, a singular solution. This means that singular solutions might lead to singular points of $\pi_E(S)$ or $\pi_F(S)$ but not in all cases. The next proposition answers under which conditions it does not.

For non-horizontal solutions, the following proposition still holds.

Proposition 19 *If $\mathcal{E}$ and $\mathcal{F}$ share the intersection $\mathcal{I} = \mathcal{EF} \cap \Phi$ (where $\mathcal{EF} = \pi_E^{-1}(\mathcal{E}) \cap \pi_F^{-1}(\mathcal{F})$) and $\mathcal{I}$ has a (possibly singular) solution S, then $\pi_E(S)$ is a solution of $\mathcal{E}$ if $d\pi_E|_S$ is injective and $\pi_F(S)$ is a solution of $\mathcal{F}$ if $d\pi_F|_S$ is injective.*

Proof If $d\pi_E|_S$ is injective, then, since π_E is smooth, $\pi_E|_S$ is a local diffeomorphism onto its image. As shown in the proof of Proposition 18, this implies that $\pi_E(S)$ is a smooth submanifold of $\mathcal{E} \subset J^k(E)$. To show that it is a solution, it only remains to show that π_E preserves the Cartan distribution, i.e. $d\pi_E(v \in C_\theta) \subset C_{\pi_E(\theta)}\ \forall\theta$. Since $v \in C_\theta$ locally lies in the span of the vector fields

$$\begin{aligned} D_q &= \frac{\partial}{\partial x^q} + \sum_{j=1}^{e}\sum_{|\alpha|<k} u^j_{\alpha q}\frac{\partial}{\partial u^j_\alpha} + \sum_{g=1}^{f}\sum_{|\beta|<l} v^j_{\beta q}\frac{\partial}{\partial v^j_\beta} \\ \text{and } D^j_\delta &:= \frac{\partial}{\partial u^j_\delta},\ |\delta| = k, \quad \text{as well as } D^g_\kappa := \frac{\partial}{\partial v^g_\kappa},\ |\kappa| = l, \end{aligned} \tag{99}$$

and $\pi_E(x^i, u^j_\alpha, v^g_\beta) = (x^i, u^j_\alpha)$, one obtains

$$d\pi_E = \sum_{i=1}^{m} \frac{\partial}{\partial x^i} \otimes dx^i + \sum_{j=1}^{e}\sum_{|\alpha|\le k} \frac{\partial}{\partial u^j_\alpha} \otimes du^j_\alpha \tag{100}$$

and consequently

$$\begin{aligned} d\pi_E(D_q) &= \frac{\partial}{\partial x^q} + \sum_{j=1}^{e}\sum_{|\alpha|<k} u^j_{\alpha q}\frac{\partial}{\partial u^j_\alpha} \\ d\pi_E(D^j_\delta) &= D^j_\delta, \qquad d\pi_E(D^g_\kappa) = 0. \end{aligned} \tag{101}$$

Thus, $d\pi_E(C_\theta) = C_{\pi_E(\theta)}$. As a consequence, since S was an integral submanifold of C, i.e. $v \in C_\theta\ \forall v \in T_\theta S$, and those vectors are mapped to the Cartan distribution of $J^k(E)$ by $d\pi_E$, it follows that $\pi_E(S)$ must also be an integral submanifold of the Cartan distribution. Since it is a smooth submanifold of dimension $\dim(M)$, this implies that it is a (possibly singular) solution of $\mathcal{E}$. □

In particular, the above proposition yields the following corollary.

Corollary 2 *If $\mathcal{E}$ and $\mathcal{F}$ share the intersection $\mathcal{I} = \mathcal{EF} \cap \Phi$ (where $\mathcal{EF} = \pi_E^{-1}(\mathcal{E}) \cap \pi_F^{-1}(\mathcal{F})$) and $\mathcal{I}$ has a (possibly singular) solution S, then if*

$$S_{inj(E)} := \{\theta \in S \mid d\pi_E|_S \text{ is injective}\} \tag{102}$$

has dimension $\dim(M)$, $\pi_E(S_{inj(E)})$ is a (possibly singular) solution of $\mathcal{E}$. The same holds for $\pi_F(S_{inj(F)})$.

7 Bäcklund Correspondences

In this section, it is shown how the present framework naturally generalizes Bäcklund transformations which can sometimes serve to generate non-trivial solutions of non-linear PDEs.

Another definition of Bäcklund transformations within the beautiful theory of coverings can be found in subsection 3.8 of [12] and also in subsection 1.11 of chapter 6 of [13]. However, the theory of coverings takes place in the category of infinitely prolonged differential equations which is not convenient in the present situation for two reasons: First, the present setting was developed to compare two differential equations that might not share enough structure to be formally integrable which forces us to stay on the level of finite jets. Second, singular solutions are more difficult to deal with on infinite jet spaces because the Cartan distribution becomes purely horizontal. Therefore, a generalization of Bäcklund transformations on the level of finite jets is useful for the present purposes.

As a first step, the definition of a Bäcklund transformation described on p. 134–140 in [44] is rewritten and somewhat simplified using the present notation. As before, let $\pi : E \to M$ be a fibered manifold, $J^k(E)$ the k-th order jet space over E and $\xi : F \to M$ another fibered manifold with the same base space M. If $\psi : J^k(E) \times_M J^0(F) \to J^1(F)$ is a morphism of fibered manifolds, then $p^1(\psi) : J^{k+1}(E) \times_M J^1(F) \to J^1(J^1(F))$ denotes the prolongation of ψ. As already explained around Eq. (5), there is a well-defined inclusion $i_{1,1} : J^2(F) \to J^1(J^1(F))$ that embeds $J^2(F)$ into $J^1(J^1(F))$.

Definition 14 A *Bäcklund map* is a morphism of fibered manifolds, $\psi : J^k(E) \times_M J^0(F) \to J^1(F)$, such that

$$\xi_0^1 \circ \psi = \pi_2, \tag{103}$$

where $\xi_0^1 : J^1(F) \to F$ and $\pi_2 : J^k(E) \times J^0(F) \to J^0(F) = F$. The *Bäcklund compatibility condition*

$$\mathrm{im}(p^1(\psi)) \subset i_{1,1}(J^2(F)) \tag{104}$$

gives rise to a subset $\mathcal{P}(\psi) \subset J^{k+1}(E) \times_M J^1(F)$ given by

$$\mathcal{P}(\psi) := \{\theta \in J^{k+1}(E) \times_M J^1(F) \mid p^1(\psi)(\theta) \in i_{1,1}(J^2(F))\} \tag{105}$$

Definition 15 If $\mathcal{P}(\psi)$ contains a system that only depends on the coordinates of $J^{k+1}(E)$, i.e. if one has

$$\mathcal{E} = \pi_E'(\mathcal{P}(\psi)) \text{ for some PDE } \mathcal{E} \subset J^{k+1}(E), \tag{106}$$

where $\pi'_E : J^{k+1}(E) \times_M J^1(F) \to J^{k+1}(E)$, then ψ is called an *ordinary Bäcklund map* for $\mathcal{E}$.

To provide a better understanding of this definition, a brief description of all conditions in local coordinates is given. Let the coordinates of $J^k(E) \times_M J^0(F)$ be (x^i, u^j_α, v^g), $|\alpha| \leq k$ and those of $J^1(F)$ be (x^i, w^g, w^g_b), $i, b \in \{1, \cdots, m\}$. The condition that ψ is a morphism of fibered manifolds locally translates into the description

$$(x^i, u^j_\alpha, v^g) \mapsto (x^i, w^g = \psi^g(x, u, v), w^g_b = \psi^g_b(x, u, v)) \tag{107}$$

The condition (103) then locally implies

$$w^g = \psi^g(x, u, v) = v^g, \tag{108}$$

and the compatibility condition (104) can locally be understood as follows. Let $J^{k+1}(E) \times_M J^1(F)$ have local coordinates $(x^i, u^j_\alpha, v^g_\beta)$, this time with $|\alpha| \leq k+1$ and $|\beta| \leq 1$ and the local coordinates of $J^1(J^1(F))$ be $(x^i, w^g, w^g_b, (w^g)_b, (w^g_b)_c)$, $b, c \in \{1, \cdots, m\}$. Then, for $p^1(\psi) : J^{k+1} \times_M J^1(F) \to J^1(J^1(F))$, one obtains

$$p^1(\psi) \begin{pmatrix} x^i \\ u^j_\alpha \\ v^g_\beta \end{pmatrix} = \begin{pmatrix} x^i \\ w^g = \psi^g(x, u, v) \overset{(108)}{=} v^g \\ w^g_b = \psi^g_b(x, u, v)) \\ (w^g)_b = D_b \psi^g(x, u, v) \overset{(108)}{=} D_b v^g = v^g_b \\ (w^g_b)_c = D_c \psi^g_b(x, u, v) \end{pmatrix}, \tag{109}$$

where D_b, as before, is the total differential operator.

$$D_b = \frac{\partial}{\partial x^b} + \sum_{j=1}^{e} \sum_{|\alpha|<k+1} u^j_{\alpha b} \frac{\partial}{\partial u^j_\alpha} + \sum_{g=1}^{f} \sum_{|\beta|<1} v^g_{\beta b} \frac{\partial}{\partial v^g_\beta} \tag{110}$$

Since the subset $i_{1,1}(J^2(F))$ in $J^1(J^1(F))$ has local coordinates $(x^i, w^g, w^g_b, (w^g)_b = w^g_b, (w^g_b)_c = w^g_{bc} = w^g_{cb} = (w^g_c)_b)$, the local equations describing $\mathcal{B}$ defined in (105) by the compatibility condition (104) are finally given by

$$\mathcal{P}(\psi) : \{ v^g_b = \psi^g_b, \quad D_c \psi^g_b = D_b \psi^g_c \}. \tag{111}$$

This concludes the descriptions of the local coordinates involved in the definition of a Bäcklund map.

The next step is to use a Bäcklund map to define a Bäcklund transformation. To this end, note first that, since the restriction of $p^1(\psi)$ to $\mathcal{P}(\psi)$ by construction has an image that lies in $J^2(F)$, one can define a map $\psi^1 : \mathcal{P}(\psi) \to J^2(F)$, simply given by $\psi^1(\theta) := p^1(\psi)|_{\mathcal{P}(\psi)}(\theta)$. This procedure can be iterated to obtain a map $\psi^r : P^{r-1}(\mathcal{P}(\psi)) \to J^{r+1}(F)$ where $P^{r-1}(\mathcal{P}(\psi))$ is the $r-1$-th prolongation of $\mathcal{P}(\psi)$.

Definition 16 If $\psi : J^k(E) \times_M J^0(F) \to J^1(F)$ is an ordinary Bäcklund map for $\mathcal{E}$ and if, for some r, a system of differential equations $\mathcal{F} \subset J^{r+1}(F)$ contains the image of $\psi^r : \mathcal{P}(\psi)^{r-1} \to J^{r+1}(F)$, then ψ is called a *Bäcklund transformation* between $\mathcal{E}$ and $\mathcal{F}$.

The idea behind those definitions is to reduce the equations locally describing $\mathcal{F}$ to first order equations with the help of $\mathcal{E}$ and ψ. One usually obtains the following proposition that is reproven in the present terminology, for convenience.

Proposition 20 *Suppose that ψ is a Bäcklund transformation between $\mathcal{E} \subset J^{k+1}(E)$ and $\mathcal{F} \subset J^{r+1}(F)$. If $s_E : U \subset M \to E$ is a horizontal solution of $\mathcal{E}$ (i.e. $im(j^{k+1}(s_E)) \subset \mathcal{E}$), then a solution $s_F : U \subset M \to F$ of $\mathcal{F}$ (with $im(j^{r+1}(s_F)) \subset \mathcal{F}$) can be obtained by solving the following system of PDEs*

$$j^1(s_F) = \psi(j^k(s_E) \times_M s_F) \tag{112}$$

which is first-order in s_F (recall that s_E is already given) and locally described by

$$\frac{\partial s_F^g(x)}{\partial x^b} = \psi_b^g \left(x^i, \partial_\alpha s_E^j(x), s_F^h(x)\right), \tag{113}$$

where $i, q \in \{1, \cdots, m\}$, $j \in \{1, \cdots, e\}$, $g, h \in \{1, \cdots, f\}$ and $0 \leq |\alpha| \leq k+1$.

Proof A horizontal solution of $\mathcal{P}(\psi)$ is described by a section $s = s_E \times_M s_F : U \subset M \to E \times_M F$ such that

$$\mathrm{im}(j^{k+1}(s_E) \times_M j^1(s_F)) \subset \mathcal{P}(\psi). \tag{114}$$

Since s is assumed to be smooth, (114) holds if

$$\mathrm{im}(j^{k+r}(s_E) \times_M j^r(s_F)) \subset P^{r-1}(\mathcal{P}(\psi)). \tag{115}$$

Since by assumption $\psi^r(P^{r-1}(\mathcal{P}(\psi))) \subset \mathcal{F}$, (115) in turn implies $\mathrm{im}(\psi^r(j^{k+r}(s_E) \times_M j^r(s_F))) \subset \mathcal{F}$. At the same time,

$$\begin{aligned}\psi^r(j^{k+r}(s_E) \times_M j^r(s_F)) &= p^r(\psi)(j^{k+r}(s_E) \times_M j^r(s_F)) \\ &\overset{(8)}{=} j^r(\psi(j^k(s_E) \times_M s_F))\end{aligned} \tag{116}$$

Thus, if $s = s_E \times_M s_F$ is a solution of $\mathcal{P}(\psi)$ and one can find a section $s_{F'} : U \subset M \to F$ such that

$$j^1(s_{F'}) = \psi(j^k(s_E) \times_M s_F), \tag{117}$$

then $s_{F'}$ is a solution of $\mathcal{F}$. Since $\pi_2 = \xi_0^1 \circ \psi$ by (103), we also have

$$s_F = \pi_2(j^k(s_E) \times_M s_F) = \xi_0^1(\psi(j^k(s_E) \times_M s_F)), \tag{118}$$

and since s_F is holonomic, this implies

$$j^1(s_F) = \psi(j^k(s_E) \times_M s_F) \tag{119}$$

In other words, $s_{F'} = s_F$ always solves (117). As a conclusion, whenever $s = s_E \times_M s_F$ solves $\mathcal{P}(\psi)$, then s_F itself is such that it solves $\mathcal{F}$.

Hence, if a solution s_E of $\mathcal{E}$ is given, a solution $s = s_E \times_M s_F$ of $\mathcal{P}(\psi)$ can be found by finding s_F s.t. (114) holds. As we also assume that ψ is ordinary for $\mathcal{E}$, Eq. (106) holds, which implies that $\pi_E^{-1}(\mathrm{im}(s_E))$ contains the image of a section $s = s_E \times_M s_F$ which is contained in $\mathcal{P}(\psi)$. Therefore given a solution s_E, we get s_F by solving the remaining equation describing $\mathcal{P}(\psi)$, $v_b^g = \psi_b^g$ (cf. (111)) that is Eq. (119), which in local coordinates is described by the system (113). □

As a next step, Bäcklund transformations are identified as a special case of the present framework.

Proposition 21 *Every Bäcklund transformation $\psi : J_B := J^k(E) \times_M J^0(F) \to J^1(F)$ between $\mathcal{E} \subset J^{k+1}(E)$ and $\mathcal{F} \subset J^{r+1}(F)$ gives rise to an intersection $\mathcal{I} = \mathcal{EF} \cap \Phi$ where $\mathcal{EF}$ is constructed as in Eq. (20) on the natural product bundle $J := J^{k+1}(E) \times_M J^{r+1}(F)$ of $\mathcal{E}$ and $\mathcal{F}$ and Φ is completely determined by ψ.*

Φ fulfills a condition equivalent to (103) and the projection of the prolongation $\pi_{k+1,\,1}^{k+2,\,r+2}(P^1(\Phi))$ corresponds to the compatibility condition $\mathcal{P}(\psi)$ defined in (105).

Proof As before, given $\mathcal{E} \subset J^{k+1}(E)$ and $\mathcal{F} \subset J^{r+1}(F)$, one can form the natural product bundle $J := J^{k+1}(E) \times_M J^{r+1}(F)$ and pull $\mathcal{E}$ and $\mathcal{F}$ back to $\mathcal{E}_J$ and $\mathcal{F}_J$ via $\pi_E : J \to J^{k+1}(E)$ and $\pi_F : J \to J^{r+1}(F)$, i.e. $\mathcal{EF} := \pi_E^{-1}(\mathcal{E}) \cap \pi_F^{-1}(\mathcal{F})$ as in Eq. (20). Next, one can define a correspondence Φ as follows

$$\Phi : \{\xi_1^{r+1} \circ \pi_F = \psi \circ \pi_B\} \tag{120}$$

where $\xi_1^{r+1} : J^{r+1}(F) \to J^1(F)$ and $\pi_B : J \to J_B$.

Recall that $\pi_2 : J_B \to J^0(F)$. Since $\pi_2 \circ \pi_B = \xi_0^{r+1} \circ \pi_F$ and $\xi_0^1 \circ \xi_1^{r+1} = \xi_0^{r+1}$, applying ξ_0^1 to both sides of the Eq. (120) defining Φ results in

$$\pi_2 \circ \pi_B = \xi_0^{r+1} \circ \pi_F = \xi_0^1 \circ \xi_1^{r+1} \circ \pi_F \overset{(120)}{=} \xi_0^1 \circ \psi \circ \pi_B, \tag{121}$$

which is equivalent to condition (103) but this time imposed on Φ on J instead of on ψ on J_B. Note that the condition here is trivially fulfilled because we are only considering a submanifold Φ on one product bundle with one set of coordinates (x, u, v) instead of a morphism ψ between two different fibered manifolds with two different sets of coordinates (x, u, v) and (x, w) that required the additional condition $v = w$. This is an indication that the present approach is more natural.

If the coordinates of J are $(x^i, u^j_\alpha, v^g_\beta)$ and of $J^1(F)$ are (x^i, w^g, w^g_b), then $\psi \circ \pi_B$ and $\xi_1^{r+1} \circ \pi_F$ are locally given by

$$\begin{pmatrix} x^i \\ v^g \\ v^g_b \end{pmatrix} = \xi_1^{r+1}(\pi_F(x^i, u^j_\sigma, v^g_\lambda)) \overset{(120)}{=} \psi(\pi_B(x^i, u^j_\sigma, v^g_\lambda)) = \begin{pmatrix} x^i \\ \psi^g(x^i, u^j_\delta, v^g) \\ \psi^g_b(x^i, u^j_\delta, v^g) \end{pmatrix} \tag{122}$$

which correspond to the equations described in Eq. (108) and the left equation in (111). (Note that, in the eq. above, $|\sigma| \leq k+1$, $|\lambda| \leq r+1$ but $|\delta| \leq k$.)

Condition (104) is a projected version of the compatibility condition that is enforced by the intersection in the definition of a prolongation, cf. Eq. (6),

$$P^1(\Phi) = p(\, J^1(\Phi) \,\cap\, (i_{k+1,1} \times_M i_{r+1,1}(P^1(J)))\,), \tag{123}$$

where

$$P^n(J) = J^{k+1+n}(E) \times_M J^{r+1+n}(F). \tag{124}$$

Indeed, by Eqs. (12) and (122),

$$P^1(\Phi) : \left\{ \theta \in P^1(J) \;\middle|\; D_b\psi^g(\kappa) = v^g_b = \psi^g_b(\kappa),\ D_a\psi^g_b(\kappa) = v^g_{ab} = D_b\psi^g_a(\kappa) \right\} \tag{125}$$

where $\kappa = \pi^{k+2,\ r+2}_{k+1,\ 1}(\theta)$ and $\pi^{a,\ b}_{c,\ d} : J^a(E) \times_M J^b(F) \to J^c(E) \times_M J^d(F)$ is the canonical projection. Since $\kappa \in J^{k+1}(E) \times_M J^1(F)$, those equations (apart from the condition $v^g_{ab} = D_b\psi^g_a$) are preserved under projection, and one obtains

$$\begin{aligned} \pi^{k+2,\ r+2}_{k+1,\ 1}(P^1(\Phi)) &= \left\{ \theta \in J^{k+1}(E) \times_M J^1(F) \mid v^g_b = \psi^g_b,\ D_a\psi^g_b = D_b\psi^g_a \right\} \\ &\overset{(111)}{=} \mathcal{P}(\psi) \end{aligned} \tag{126}$$

As a result, the compatibility conditions of a Bäcklund map can be understood as the equations arising upon prolongation of the correspondence Φ. □

Prolonging (126), one obtains

$$\pi^{k+1+n,\ r+1+n}_{k+n,\ n}(P^n(\Phi)) = P^{n-1}(\mathcal{P}(\psi)) \tag{127}$$

Proposition 22 *$\mathcal{I}$ as defined in Proposition 21 allows to transfer solutions from $\mathcal{E}$ to $\mathcal{F}$ in the sense of Proposition 20.*

Proof To show that Φ facilitates to transfer solutions from $\mathcal{E}$ to $\mathcal{F}$ by solving a first-order system, one can proceed as follows. By Proposition 18, we know that any solution S of $\mathcal{I} = \mathcal{E}\mathcal{F} \cap \Phi$ can be projected to solutions $\pi_E(S)$ and $\pi_F(S)$ of $\mathcal{E}$ and $\mathcal{F}$ respectively. What's special about Bäcklund transformations, is that solving Φ alone is actually sufficient. The reason is that the differential consequences of Φ contain the equations describing $\mathcal{E}$ and $\mathcal{F}$. To show that, we will show that $\pi^r_J(P^r(\Phi)) \subset \mathcal{E}\mathcal{F}$ where $\pi^r_J : P^r(J) \to J$ (and $P^r(J)$ is given by Eq. (124)).

Since ψ is a Bäcklund transformation between $\mathcal{E}$ and $\mathcal{F}$, Definition 15 holds, i.e. $\pi'_E(\mathcal{P}(\psi)) = \mathcal{E}$. Recall that $\pi_E = \pi'_E \circ \pi^{k+1,\ r+1}_{k+1,\ 1}$. Then $\pi'_E(\mathcal{P}(\psi)) = \mathcal{E}$ implies that

$$\begin{aligned}
&\pi_E^{-1}(\pi'_E(\mathcal{P}(\psi))) \subset \pi_E^{-1}(\mathcal{E}) = \mathcal{E}_J \quad \text{where} \\
&\pi_E^{-1}(\pi'_E(\mathcal{P}(\psi))) = (\pi^{k+1,\ r+1}_{k+1,\ 1})^{-1} \circ (\pi'_E)^{-1} \circ \pi'_E(\mathcal{P}(\psi)) \\
&\qquad = (\pi^{k+1,\ r+1}_{k+1,\ 1})^{-1}(\mathcal{P}(\psi)) \\
&\qquad \overset{(126)}{=} (\pi^{k+1,\ r+1}_{k+1,\ 1})^{-1}\big(\pi^{k+2,\ r+2}_{k+1,\ 1}(P^1(\Phi))\big)
\end{aligned} \tag{128}$$

Note also that apart from $v^g_{ab} = D_a\psi^g_b$, the eqs describing $P^1(\Phi)$ are first order in v (cf. Eqs. (125) and (126)). Therefore,

$$\begin{aligned}
(\pi^{k+1,\ r+1}_{k+1,\ 1})^{-1}\big(\pi^{k+2,\ r+2}_{k+1,\ 1}(P^1(\Phi))\big) \cap \ \{v^g_{ab} = D_a\psi^g_b\} &= \pi^{k+2,\ r+2}_{k+1,\ r+1}(P^1(\Phi)) \\
&= \pi^1_J(P^1(\Phi))
\end{aligned} \tag{129}$$

and thus

$$\pi^1_J(P^1(\Phi)) \subset \big(\mathcal{E}_J \ \cap \ \{v^g_{ab} = D_a\psi^g_b\}\ \big) \subset \mathcal{E}_J \tag{130}$$

Since projections of further prolongations can only increase the number of constraints/equations, we can conclude

$$\pi^n_J(P^n(\Phi)) \subset \pi^1_J(P^1(\Phi)) \subset \mathcal{E}_J, \qquad \forall\, n \geq 1. \tag{131}$$

Next, we want to show that we also have $\pi^r_J(P^r(\Phi)) \subset \mathcal{F}_J$. To do so, we use Definition 16 that guarantees that a Bäcklund transformation satisfies $\mathrm{im}(\psi^r) \subset \mathcal{F}$ which implies

$$\pi_F^{-1}(\mathrm{im}(\psi^r)) \subset \pi_F^{-1}(\mathcal{F}) = \mathcal{F}_J \tag{132}$$

where

$$\begin{aligned}\mathrm{im}(\psi^r) &= \mathrm{im}\left(p^r(\psi)|_{P^{r-1}(\mathcal{P}(\psi))}\right) = p^r(\psi)(P^{r-1}(\mathcal{P}(\psi))\\ &\overset{(127)}{=} p^r(\psi)(\pi^{k+1+r,\ r+1+r}_{k+r,\ r}(P^r(\Phi)))\\ &= p^r(\psi\circ\pi^{k+1,\ r+1}_{k,\ 0})(P^r(\Phi))\\ &\overset{(120)}{=} p^r(\xi^{r+1}_1\circ\pi_F)(P^r(\Phi)) = \xi^{r+1+r}_{1+r}\circ p^r(\pi_F)(P^r(\Phi))\end{aligned} \tag{133}$$

Since

$$J^{k+1+r}(E)\times_M J^{r+1+r}(F) \xrightarrow{p^r(\pi_F)} J^{r+1+r}(F) \xrightarrow{\xi^{r+1+r}_{1+r}} J^{r+1}(F) \tag{134}$$

commutes with

$$J^{k+1+r}(E)\times_M J^{r+1+r}(F) \xrightarrow{\pi^r_J} J^{k+1}\times_M J^{r+1}(F) \xrightarrow{\pi_F} J^{r+1}(F)\ , \tag{135}$$

we obtain

$$\begin{aligned}\mathrm{im}(\psi^r) &\overset{(133)}{=} \pi_F\circ\pi^r_J(P^r(\Phi))\\ \Rightarrow \pi^r_J(P^r(\Phi)) &\overset{(132)}{\subset} F_J\end{aligned} \tag{136}$$

Together with (131), we thus finally obtain

$$\pi^r_J(P^r(\Phi)) \subset \mathcal{EF} := \mathcal{E}_J\cap\mathcal{F}_J \tag{137}$$

which expresses the essential property of a Bäcklund transformation: The equations $\mathcal{EF}$ are differential consequences of Φ. (Since $\pi^r_J(P^r(\Phi))$, which includes the differential consequence of Φ up to order $k+1$ in u and $r+1$ in v, is contained in $\mathcal{EF}$, the equations that locally describe $\mathcal{EF}$ are in turn a subset of the equations of the (smaller) space $\pi^r_J(P^r(\Phi))$).

Since we assume that solutions are smooth, every solution of Φ must also be a solution of any prolongation $P^n(\Phi)$. Since the prolongation $P^r(\Phi)$ of Φ contains both, the equations describing $\mathcal{E}$ and those describing $\mathcal{F}$, the solution of Φ must also solve $\mathcal{E}$ and $\mathcal{F}$. By Proposition 18, solutions of $\mathcal{I}$ can be projected to solutions of $\mathcal{E}$ and $\mathcal{F}$ via π_E and π_F (even singular ones if the conditions in Proposition 19 are fulfilled).

As a final step, let us show that solving a first-order system is sufficient if a general solution to $\mathcal{E}$ is given. Suppose that $\mathcal{E}$ has a general family of solutions S^α_E, parameterized by α, that is locally described by sections s^α_E. Since $\pi^r_J(P^r(\Phi))\subset\mathcal{E}_J$, the pullback of the family of solutions $\pi_E^{-1}(\mathrm{im}(s^\alpha_E))\subset\mathcal{E}_J$ should intersect solutions of $P^r(\Phi)$ that can be found by looking for a section s_F such that the prolongation of $s^\alpha_E\times_M s_F$ is contained in Φ for some α. The resulting system of equations is then first order in s_F, namely locally described by (122). This solution can then be mapped to $\mathcal{F}$ as explained above, by Proposition 18. Hence, given s_E, it suffices to solve the first order PDE (113) to obtain a solution of $\mathcal{F}$ which concludes an alternative proof of Proposition 20 in a more general setting. □

The proof above makes it clear that the exact form of Φ is not really essential for transferring solutions and reducing the order of equations as long as $\mathcal{E}_J$ and $\mathcal{F}_J$ are differential consequences of Φ, i.e. as long as (137) is satisfied. In particular, staying in the natural product bundle makes it unnecessary to impose conditions like (103) or to require that the codomain of ψ is a first order jet space. Thus, the following generalization seems appropriate.

Definition 17 A correspondence $\Phi \subset J := J^k(E) \times_M J^l(F)$ between two differential equations $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$ is said to be a *Bäcklund correspondence* or to *have the Bäcklund property* if, for some $r \geq 1$,

$$\mathcal{P}(\Phi) := \pi^r_J(P^r(\Phi)) \subset \mathcal{E}\mathcal{F}. \tag{138}$$

where $\mathcal{P}(\Phi)$ is the projection of the prolongation of Φ. It is said to be a *strict Bäcklund correspondence* if

$$\pi_E(\mathcal{P}(\Phi)) = \mathcal{E} \text{ and } \pi_F(\mathcal{P}(\Phi)) = \mathcal{F}. \tag{139}$$

Note that (139) implies (138) because

$$\begin{aligned} (139) \;\Rightarrow\; \mathcal{P}(\Phi) &\subset \begin{cases} \pi_E^{-1}(\pi_E(\mathcal{P}(\Phi))) = \pi_E^{-1}(\mathcal{E}) = \mathcal{E}_J \\ \pi_F^{-1}(\pi_F(\mathcal{P}(\Phi))) = \pi_F^{-1}(\mathcal{E}) = \mathcal{F}_J \end{cases} \\ \Rightarrow\; \mathcal{P}(\Phi) &\subset \mathcal{E}_J \cap \mathcal{F}_J = \mathcal{E}\mathcal{F}. \end{aligned} \tag{140}$$

With this definition, the following proposition holds.

Proposition 23 *Whenever a correspondence Φ between two PDEs $\mathcal{E}$ and $\mathcal{F}$ is Bäcklund, every solution of Φ is a solution of both, $\mathcal{E}$ and $\mathcal{F}$.*

Proof Since $\mathcal{P}(\Phi) \subset \mathcal{E}\mathcal{F}$, and, since $\mathcal{P}(\Phi) := \pi^r_J(P^r(\Phi))$, also $\mathcal{P}(\Phi) \subset \Phi$, we obtain $\mathcal{P}(\Phi) \subset \mathcal{I} = \mathcal{E}\mathcal{F} \cap \Phi$. Hence, what solves $\mathcal{P}(\Phi)$ also solves $\mathcal{I}$. But $\mathcal{P}(\Phi)$ is solved when Φ is solved because $\mathcal{P}(\Phi)$ is the projection of differential consequences of Φ. Hence, a solution of Φ solves $\mathcal{I}$ and then this solution can be mapped to $\mathcal{E}$ and $\mathcal{F}$ by Proposition 18. □

Again, given the solution of one of the equation might allow to reduce the order of the other:

Proposition 24 *If Φ is a strict Bäcklund correspondence between $\mathcal{E}$ and $\mathcal{F}$, and a solution S of $\mathcal{E}$ is given, then a solution of $\mathcal{F}$ can be found by finding a solution of $\pi_E^{-1}(S) \cap \Phi$.*

Proof Since Φ is a strict Bäcklund correspondence, one has $\pi_E(\mathcal{P}(\Phi)) = \mathcal{E}$. This means that, apart from $\mathcal{E}_J$, the prolongation of Φ does not impose additional equations, purely in terms of coordinates of $J^k(E)$, on J. Hence $\pi_E^{-1}(S) = S \times_M J^l(F)$ intersects the solution space of Φ. If a solution in this intersection can be found, it also solves $\mathcal{F}$ by Proposition 23. □

The present approach generalizes the usual definition of a Bäcklund transformation because one can now define a correspondence of any order and the dependence on the coordinates of $J^l(F)$ can be arbitrary apart from the requirement that Φ should be an almost diagonal fibered submanifold of $J^k(E) \times_M J^l(F)$. Despite the increased generality, solutions can still be transferred in a similar way to the simpler case.

Example A very classical example that illustrates Bäcklund transformations is the one involving the Liouville equation $u_{12} = \exp(u)$. It is briefly rephrased in the present terminology to illustrate the general ideas above. Consider $\pi : E := \mathbb{R} \times \mathbb{R} \to \mathbb{R} =: M$ with local coordinates (x, y, u) and $\xi : F \simeq E \to M$ with local coordinates (x, y, v) and the equations $\mathcal{E} : \{ u_{12} = e^u \} \subset J^2(E)$ and $\mathcal{F} : \{ v_{12} = 0 \} \subset J^2(F)$. We relate them on $J^2(E) \times_M J^2(F)$ by a correspondence Φ determined by the equations

$$\Phi : \left\{ v_1 = u_1 + \beta \exp\left(\frac{u+v}{2}\right) , \quad v_2 = -u_2 - \frac{2}{\beta} \exp\left(\frac{u-v}{2}\right) \right\} \tag{141}$$

First, we check that this Φ is indeed a correspondence. Since it is defined by two independent equations, it has codimension $2 \geq 1 = \min(\dim(E_x, F_x))$. Its projection to both, $J^2(E)$ and $J^2(F)$ does not impose any conditions and thus, it is almost diagonal to $\mathcal{E}$ and $\mathcal{F}$. The prolongation $P^1(\Phi)$ of Φ is described by the equations describing Φ and additionally by the following ones.

$$\left\{ \begin{aligned} v_{11} &= u_{11} + \beta \exp\left(\frac{u+v}{2}\right) \frac{u_1 + v_1}{2} \\ v_{12} &= u_{12} + \beta \exp\left(\frac{u+v}{2}\right) \frac{u_2 + v_2}{2} \\ v_{21} &= -u_{21} - \frac{2}{\beta} \exp\left(\frac{u-v}{2}\right) \frac{u_1 - v_1}{2} \\ v_{22} &= -u_{22} - \frac{2}{\beta} \exp\left(\frac{u-v}{2}\right) \frac{u_2 - v_2}{2} \end{aligned} \right\} \tag{142}$$

(continued)

The compatibility conditions $v_{12} = v_{21}$ and $u_{12} = u_{21}$, that must be imposed when taking the prolongation, result in the following two equations, which, together with Eqs. (141) and (142) describe $\mathcal{P}(\Phi)$:

$$\left\{u_{12} = \exp(u), \qquad v_{12} = 0\right\} \tag{143}$$

As can be seen, the correspondence was designed such that its differential consequences are included in both of the intersected equations, i.e.

$$\mathcal{P}(\Phi) \subset \mathcal{EF} \cap \Phi =: \mathcal{I} \subset \mathcal{EF}, \tag{144}$$

Thus, Φ is a Bäcklund correspondence. Furthermore, since Φ is almost diagonal, we obtain $\pi_E(\mathcal{P}(\Phi)) = \mathcal{E}$ and $\pi_F(\mathcal{P}(\Phi)) = \mathcal{F}$. Hence, Φ is a strict Bäcklund correspondence.

Thus, by Proposition 24, solutions an be transferred between the PDEs. The general solution of $v_{12} = 0$ is given by $v(x, y) = A(x) + B(y)$ and plugging this into (141) results in a PDE for $u(x, y)$ that can be integrated (though it is not completely trivial), and one obtains the solution

$$u(x, y) = 2\ln\left(\frac{\exp\left(\frac{A(x)-B(y)}{2}\right)}{\frac{\beta}{2}\int_{x_0}^{x}\exp\left(A(x')\right)\,dx' + \frac{1}{\beta}\int_{y_0}^{y}\exp\left(-B(y')\right)\,dy'}\right) \tag{145}$$

As mentioned by [44], this encouraging result was an important motivation for the search of Bäcklund transformations.

8 Equivalence Up to Symmetry and Quotient Equations

When comparing two theories in mathematically different formulations that only differ up to a symmetry which is physically not relevant, then one would like to find a way to compare the two theories after removing this symmetry. For example, classical electrodynamics can be formulated in terms of gauge potentials and in terms of Faraday tensors. At least classically, those two theories are physically equivalent because only the fields are measurable quantities. To formalise this physical equivalence mathematically, Weatherall invented the solution-Category approach described in [5] and [6] which was already mentioned in the introduction in Sect. 1.1. The idea behind this formalism was, among other things, to show that those mathematical structures in which the morphisms between the objects of the solution categories are induced (via the pushforward or pullback) by the diffeomorphisms of the underlying manifold are more natural than those in which those symmetries have to be "added by hand" in order to achieve an equivalence to other physically equivalent formulations.

The aim of the present section is to show how one can approach those ideas in the category of smooth manifolds.

The section describes the general idea how to "quotient out" a symmetry of an equation and how to obtain the corresponding invariant equation. Basically, the invariant equation is realised by replacing the variables in the equation by the invariants of the symmetry. So the real work consists in finding all functionally independent invariants. Though the present approach was developed somewhat independently, quotient equations are a well-known concept (cf. [13] (chapter 3.6), [35, 36, 45], also [32] is related).

We start with the geometric definition of a symmetry of a PDE (taken from [13])

Definition 18 A *Lie transformation* is a diffeomorphism $L : J^k(E) \to J^k(E)$ such that $dL_\theta(C_\theta) = C_{L(\theta)}\ \forall\theta \in J^k(E)$ (where C is the Cartan distribution on $J^k(E)$). A vector field X on the manifold $J^k(E)$ is called a *Lie field*, if shifts along its flow are Lie transformations.

Definition 19 A Lie transformation S which is such that $S(\mathcal{E}) = \mathcal{E}$ is called a *symmetry* of the differential equation $\mathcal{E} \subset J^k(E)$. A Lie field X is called an *infinitesimal symmetry* of the equation $\mathcal{E} \subset J^k(E)$, if it is tangent to $\mathcal{E}$.

Having defined Symmetries, we can proceed to define the concept of an invariant of a symmetry (taken from [46]).

Definition 20 Given a Lie transformation S on $J^k(E)$, an *invariant* of this transformation is a map $I : J^k(E) \to \mathbb{R}$ such that $S^*I = I$, i.e. $I(\theta) = I(S(\theta))\ \forall\theta \in J^k(E)$.

Now suppose that S is a symmetry of the equation $\mathcal{E}$, i.e. $S(\mathcal{E}) = \mathcal{E}$. If the equation is given as the kernel of a differential operator $\Phi : J \subset J^k(E) \to F$, where $\pi' : F \to M$ is another fibered manifold, i.e. $\mathcal{E} = \ker_s(\Phi)$, where $s : M \to F$ is a suitable section, then this implies that $\Phi(\theta) = s(\pi^k(\theta))$ iff $S^*\Phi(\theta) = \Phi(S(\theta)) = s(\pi^k(S(\theta)))$.

Observe that Φ itself does not have to be invariant but the condition $\Phi(\theta) = s(\pi^k(\theta))$ only holds for $\theta \in \mathcal{E}$ which is invariant. But this means that it should be possible to perform algebraic operations on the equation $\ker_s(\Phi)$ which facilitate to reformulate the equation in terms of invariants of the symmetry, at least at all those points where those algebraic operations are well-defined. In other words, it should be possible to find a $\Phi' : J' \subset J^k(E) \to F$ such that $\mathcal{E} = \ker_{s'}(\Phi')$ and $S^*\Phi' = \Phi'$, at least at all those points making up J' where the algebraic operations on $\ker_s(\Phi)$ do not lead to a division by zero.

To find out how to find this Φ', let us suppose that we have a Lie group G that acts on $J^k(E)$. We write this action as $g \cdot \theta := S_g(\theta)$ where $S_g : J^k(E) \to J^k(E)$ is the Symmetry on our bundle corresponding to the action of $g \in G$. Given such a symmetry group, we can try to find the generating functions of all S_g-Invariants on $J^k(E)$. They can be found in a systematic way using the following proposition (also taken from [46]):

Proposition 25 *If G is a group of symmetries acting on $J^k(E)$, then all invariants I of this symmetry group fulfill the equations*

$$\boxed{X(I) = 0} \tag{146}$$

where X are the infinitesimal symmetries corresponding to the action of the Lie algebra of G.

Proof For an invariant I of a group it is true by definition that $S_g^* I = I$, $\forall g \in G$. As we assume a Lie group, we can write $S_g = \exp(aX_g)$ where X_g is the infinitesimal generator corresponding to the action of g. Thus,

$$\begin{aligned} 0 &= \frac{d}{da} I(\theta)\bigg|_{a=0} = \frac{d}{da} I(S_g(\theta))\bigg|_{a=0} \\ &= \frac{d}{da} I(\exp(aX_g)\theta)\bigg|_{a=0} = I'(\theta) X_g|_\theta = X_g(I). \end{aligned} \tag{147}$$

This is true for all g and thus for all X in the Lie algebra. □

This means that if we have a finite number of generators for our symmetry group, then it becomes possible to find all functionally independent invariants by finding the most general solution of a finite number of equations of the form (146).

Now suppose we have found out that any invariant of a given group action on a given bundle must be a function of the functionally independent invariants $(I_1, \cdots, I_r)$. Furthermore, suppose that the equation $\mathcal{E}$ on $J^k(E)$ is also invariant under the group action. Then, as explained before, it must be possible to express Φ', whose kernel is $\mathcal{E}$, almost everywhere as a function of $I_1, \cdots, I_r$. To formalize this idea, one can create a new fibered manifold using those invariants on which this quotient equation emerges. To do so, one must choose $\dim(M)$ functionally independent invariants that act as coordinates of the base space N of this new fibered manifold. The remaining invariants can then serve to indicate how many dimensions the fibers F_θ of the new manifold $\xi : F \to N$ should have. In general, the base coordinates do not agree with those of M and then one needs to invoke Tresse derivatives to construct a jet space over F or modify the Cartan distribution. However, in the following, the simpler special case, in which the coordinates of M are invariant under the symmetry, is assumed because the main purpose is to illustrate how quotient equations naturally fit into the present setting involving correspondence and intersection. There are quite a number of symmetries like translations and dilations of the dependent coordinates that are included in this special assumption. The more general case is also compatible with the present approach and might be described more explicitly in future work.

Thus, for now we assume $(I_1, \cdots, I_m) = (x^1, \cdots, x^m)$ and therefore set $N = M$ and create a new fibered manifold $\xi : F \to M$ where the fibers F_x are chosen as the spaces where the invariants live and consist of $l = r - m$ dimensions (i.e. locally they are isomorphic to $\mathbb{R}^l$) where $r > m$ is the number of the functionally independent invariants found in the previous step and $m = \dim(M)$. Then denote the corresponding local coordinates of the fibers by $(v^g) = (v^1, \cdots, v^l)$. Now the invariants $(I_1, \cdots, I_r)$ naturally determine a *correspondence* $\Phi(I)$ on the product bundle

$$J(I) := J^k(E) \times_M J^0(F), \tag{148}$$

namely

$$\Phi(I) : \{ v^1 = I_{m+1}(x^i, u^j_\alpha), \cdots, v^l = I_{r=m+l}(x^i, u^j_\alpha) \} \tag{149}$$

If one computes the prolongations $P^l(Q(I))$ of the intersection

$$Q(I) := (\pi'_E)^{-1}(\mathcal{E}) \cap \Phi(I), \tag{150}$$

where $\pi'_E : J^k(E) \times_M J^0(F) \to J^k(E)$, then, since $\mathcal{E}$ is invariant with respect to the symmetry used to construct the invariances expressed by the correspondence $\Phi(I)$ which relates the equation to the coordinates $(v^1, \cdots, v^l)$, $\mathcal{E}$ must necessarily give rise to an equation $(\mathcal{F}_{P^l(J(I))} \supset P^l(Q(I))) \subset P^l(J(I)) = J^{k+l}(E) \times_M J^l(F)$, for some l, whose local description solely involves (x^i, v^g_β), $|\beta| \leq l$. (The exact number l is determined by the minimal amount of prolongations needed to arrive at such an expression for $\mathcal{F}_{P^l(J(I))}$.)

Since the expression describing $\mathcal{F}_{P^l(J(I))}$ only depends on coordinates of $J^l(F)$, this local description is preserved under the projection $\pi^{k+l,\, l}_{k,\, l}(\mathcal{F}_{P^l(J(I))}) =: \mathcal{F}_J \subset J^k(E) \times J^l(F) =: J$. Finally, $\mathcal{F} := \pi_F(\mathcal{F}_J)$ is then called the *quotient equation*.

Note that one can take the pullback of $\Phi(I)$ to arrive at the usual notion of a correspondence

$$\Phi := (\pi^{k,\, l}_{k,\, 0})^{-1}(\Phi(I)) \subset J \tag{151}$$

on J, between the two equations $\mathcal{E}$ and $\mathcal{F}$. Furthermore, defining

$$Q := (\pi^{k,\, l}_{k,\, 0})^{-1}(Q(I)) = \mathcal{E}_J \cap \Phi, \qquad \mathcal{E}_J := \pi_E^{-1}(\mathcal{E}), \quad \pi_E : J \to J^k(E), \tag{152}$$

one can also express $\mathcal{F}$ as the projection of $\mathcal{P}(Q) := \pi^l_J(P^l(Q))$, i.e.

$$\mathcal{F} := \pi_F(\mathcal{P}(Q)), \qquad \mathcal{F}_J := \pi_F^{-1}(\mathcal{F}), \tag{153}$$

(where, as usual, $\pi^l_J : P^l(J) = J^{k+l}(E) \times_M J^{l+l}(F) \to J$.)

The quotient equation can be understood as the system which one obtains after quotienting out the action of the Group G because locally it represents $\mathcal{E}$ in terms of coordinates that were constructed from the invariants of this group. Those ideas are summarized in the following definition.

Definition 21 If there is a symmetry group G acting on $J^k(E)$ such that the action S_g is a symmetry of the PDE $\mathcal{E} \subset J^k(E)$ for all $g \in G$, then the correspondence $\Phi(I)$ defined in (149) (on the product bundle $J(I)$ defined in (148)), determined by the functionally independent invariants $I = (I_1, \ldots, I_r)$ (which can be computed by solving (146)), is called a *quotient correspondence* for $\mathcal{E}$.

At this point, it is important to notice that the symmetry completely determines the correspondence. This means that **symmetries can help to find meaningful correspondences**.

The explanations above then show that the following corollary holds.

Corollary 3 *Given a quotient correspondence $\Phi(I) \subset J^k(E) \times_M F$ for $\mathcal{E}$, the prolongations $P^l(Q(I))$ of the intersection $Q(I)$, defined in (150), for sufficiently high l, give rise to an equation on $J^l(F)$, called quotient equation, defined as in* (153)*, and, defining Φ as in* (151)*, a quotient intersection*

$$\mathcal{I} := \mathcal{E}_J \cap \mathcal{F}_J \cap \Phi \subset J. \tag{154}$$

Thus, the definition of $\mathcal{I}$ is in harmony with the usual notion of an intersection, cf. Definition 3.

The present framework allows to show that a quotient correspondence gives always rise to a special kind of Bäcklund correspondence.

Proposition 26 *A quotient correspondence $\Phi(I)$ for some equation $\mathcal{E}$ determines a strict Bäcklund correspondence Q where Q is defined as in* (152)*.*

Proof By construction, we already have $\pi_F(\mathcal{P}(Q)) = \mathcal{F}$, cf. Eq. (153). What remains to be shown is that $\pi_E(\mathcal{P}(Q)) = \mathcal{E}$.

Since $\Phi(I)$ is locally explicitly defined by (149), always relating v-coordinates to u-coordinates, it is almost diagonal and since there are no other equations involving v-coordinates, all additional conditions that arise upon prolongation of $Q = \mathcal{E}_J \cap \Phi$, apart from the differential consequences of $\mathcal{E}_J$ (which we assume here not to impose conditions of lower order on u-coordinates, i.e. $\mathcal{E}$ is assumed to be in involutive form), can always be written as expressions also involving v-coordinates and thus do not impose additional equations involving only u-coordinates. Hence $\mathcal{P}(Q) = \pi^l_J(P^l(Q))$ is almost diagonal to $\mathcal{E}$ and $\mathcal{F}$ and such that $\pi_E(\mathcal{P}(Q)) = \mathcal{E}$ and $\pi_F(\mathcal{P}(Q)) = \mathcal{F}$. Thus, Q is a strict Bäcklund correspondence. □

If $\Phi(I)$ is understood to contain the information about the symmetry group G, then this last proposition demonstrates that *Bäcklund correspondences are generalized symmetries*.

As usual, a Bäcklund transformation allows to transfer solutions between $\mathcal{E}$ and $\mathcal{F}$. However, because of the specific nature of $\Phi(I)$, one can even give an explicit description of the transferred solution, as described by the following proposition.

Proposition 27 *If S_E is a solution of $\mathcal{E}$ and $\mathcal{F}$ is a quotient equation of $\mathcal{E}$, then $S_F = \pi_F(\mathcal{P}(\pi_E^{-1}(S_E) \cap \Phi))$ is a solution of $\mathcal{F}$.*

Proof Since Q is a strict Bäcklund correspondence, $\pi_E^{-1}(S_E) = S_E \times_M J^l(F)$ intersects the solution space of $Q = \mathcal{E}_J \cap \Phi$. Since the constraints imposed by $\mathcal{E}_J$ are described by the same equations as those describing $\mathcal{E}$, which are already solved by S_E, one only needs to find a solution of $\pi_E^{-1}(S_E) \cap \Phi$.

At the same time, Φ, described by equations of the form (149), explicitly and uniquely defines the values of v^g as functions of (x^i, u^j_α). Thus, the prolongation $P^l(\Phi)$ of Φ determines, without solving any equations, the values of $v^g_\sigma,\ |\sigma| \leq l$ in terms of (x^i, u^j_δ) with $|\delta| \leq k + l$. However, when considering $P^l(\pi_E^{-1}(S_E) \cap \Phi)$, all coordinates u^j_δ are locally expressible as functions of x^i because S_E is an m-dimensional integral submanifold. Hence, one can solve v^g_σ for those x^i and project $P^l(\pi_E^{-1}(S_E) \cap \Phi)$ back to J, and then to $J^l(F)$, i.e. taking $\pi_F(\pi^l_J(P^l(\pi_E^{-1}(S_E) \cap \Phi))) = \pi_F(\mathcal{P}(\pi_E^{-1}(S_E) \cap \Phi))$, preserving those solutions. □

As a result, the following definition becomes meaningful.

Definition 22 Two differential equations $\mathcal{E} \subset J^k(E)$ and $\mathcal{F} \subset J^l(F)$ are said to be *equivalent up to the action of the symmetry Group G on $J^k(E)$* if $\mathcal{F}$ is the quotient equation of $\mathcal{E}$ with respect to a quotient correspondence determined by G.

An extended example is given in Sect. 9.3 where Maxwell's equations formulated in terms of Faraday tensors are shown to be a quotient equation of Maxwell's equations formulated in terms of gauge potentials.

A brief example that is supposed to illustrate the general formalism is given below:

Example On the bundle $\pi : E := \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2 =: M$ with coordinates (x, y, u), consider, on $J^2(E)$, the heat equation $\mathcal{E} : \ \{ u_2 = \beta u_{11} \}$. It is invariant under prolongations of dilations $X = u\frac{\partial}{\partial u}$ along u. The prolongation of X is given by

$$X^{(2)} = \sum_{|\beta| \leq 2} u_\beta \frac{\partial}{\partial u_\beta} \tag{155}$$

The generators of the differential algebra of all invariants of $X^{(2)}$ are given by the solution of (146).

$$I_1 = x, \quad I_2 = y, \quad I_3 = \frac{u_1}{u}, \quad I_4 = \frac{u_2}{u} \tag{156}$$

For later convenience, we renorm I_3 and write $I_3 = -2\beta\frac{u_1}{u}$. According to the general procedure above, we now construct a new bundle, $\xi : F := \mathbb{R}^2 \times M \to$

(continued)

M with coordinates (x, y, v, w). On the product bundle $J^2(E) \times_M F$, we can define the correspondence $\Phi(I)$ by

$$\Phi(I) : \left\{ v = -2\beta \frac{u_1}{u}, \quad w = \frac{u_2}{u} \right\} \tag{157}$$

We now want to find the compatibility conditions $\mathcal{P}(Q)$ where $Q(I) := \Phi(I) \cap (\pi'_E)^{-1}(\mathcal{E})$ (with $\pi'_E : J^2(E) \times F \to J^2(E)$). $Q(I)$ is locally given by

$$Q(I) : \left\{ u_2 = \beta u_{11}, \quad v = -2\beta \frac{u_1}{u}, \quad w = \frac{u_2}{u} \right\} \tag{158}$$

Note that the equations imply $w = \beta u_{11}/u$. The prolongation of $Q(I)$ imposes the following additional conditions

$$\left\{ \begin{array}{l} u_{12} = \beta u_{111}, \quad u_{22} = \beta u_{112}, \\ v_1 = -2\beta \dfrac{u_{11}}{u} + 2\beta \left(\dfrac{u_1}{u}\right)^2, \quad v_2 = -2\beta \dfrac{u_{12}}{u} + 2\beta \dfrac{u_1 u_2}{u^2}, \\ w_1 = \dfrac{u_{12}}{u} - \dfrac{u_1 u_2}{u^2}, \quad w_2 = \dfrac{u_{22}}{u} - \left(\dfrac{u_2}{u}\right)^2 \end{array} \right\} \tag{159}$$

The equations imply

$$v_1 = -2w + \frac{v^2}{2\beta} \Rightarrow \frac{v^2}{4\beta} - \frac{v_1}{2} = w = \frac{u_2}{u} \tag{160}$$

Thus, w is a function of v which implies that, on $(\pi'_E)^{-1}(\mathcal{E})$, the second generator in (157) depends on the first one, i.e. on $(\pi'_E)^{-1}(\mathcal{E})$ there is only one independent generator of the symmetry. We can thus expect to find one quotient equation of $\mathcal{E}$ purely in terms of v. Indeed, the differential consequences of (160) reveal the following relations.

$$w_1 = \frac{v v_1}{2\beta} - \frac{v_{11}}{2}, \quad w_2 = \frac{v v_2}{2\beta} - \frac{v_{12}}{2} \tag{161}$$

Furthermore, we can rewrite the eq for v_2 in (159) to obtain a 2nd condition on w_1:

$$v_2 = -2\beta \left(\frac{u_{12}}{u} - \frac{u_1 u_2}{u^2} \right) = -2\beta w_1 \tag{162}$$

(continued)

Combining the last two expressions for w_1, we obtain the following quotient equation on $J := J^2(E) \times_M J^2(F)$, purely in terms of v and its derivatives:

$$\mathcal{F}_J : \left\{v_2 = \beta v_{11} - v v_1\right\} \tag{163}$$

This is Burger's equation, i.e. we computed the well-known Hopf-Cole reduction.[11]

By Proposition 27, solutions of $\mathcal{E}$ can be transferred to the quotient equation $\mathcal{F} := \pi_F(\mathcal{F}_J)$ (where $\pi_F : J \to J^2(F)$). Note that the coordinate w is not involved and we could thus also consider $\mathcal{F}$ as an equation on $J^2(G)$ where $\rho : G \to M$ has local coordinates (x, y, v). One can e.g. solve the following boundary value problem. On the $((x, y) = (x, t))$ plane:

$$\begin{cases} v = A(x), & t = 0, \\ \mathcal{F}: \{v_2 + v v_1 - \beta v_{11} = 0\}, & t > 0. \end{cases} \tag{164}$$

The correspondence $v = -2\beta \frac{u_1}{u} \;\; \Rightarrow \;\; u(x,t) = \exp(-1/(2\beta) \int dx \; v(x,t))$, transforms this into an initial value problem for $\mathcal{E}$:

$$\begin{cases} u = \exp\left(-\frac{1}{2\beta} \int_{x_0}^{x} d\sigma \; A(\sigma)\right), & t = 0, \\ \mathcal{E}: \{u_2 - \beta u_{11} = 0\}, & t > 0 \end{cases} \tag{165}$$

The general solution of the heat equation given the initial condition $u(x, 0) = g(x)$ is the convolution

$$\begin{aligned} u(x,t) &= \int_{-\infty}^{\infty} dz \; f(x-z,t) g(z) \quad \text{where} \\ f(x,t) &= \frac{1}{\sqrt{4\pi\beta t}} \exp\left(-\frac{x^2}{4\beta t}\right) \quad \text{is the fundamental solution.} \end{aligned} \tag{166}$$

In the present case where $g(x)$ is given by (165), this leads to

$$u(x,t) := \frac{1}{\sqrt{4\pi\beta t}} \int_{-\infty}^{\infty} dz \; \exp\left(-\frac{1}{2\beta}\left[\frac{(x-z)^2}{2t} + \int_{z_0}^{z} d\sigma \; A(\sigma)\right]\right) \tag{167}$$

(continued)

[11] Note that (160) can be seen as a derivation of a correspondence $\Phi' : \{ v = -2\beta u_1/u, \;\; v^2/(4\beta) - v_1/2 = u_2/u \}$. which is a Bäcklund correspondence with diff. consequences $\mathcal{E}_J$ and $\mathcal{F}_J$, that, in contrast to Q, does not require the coordinate w anymore.

and using the correspondence $v = -2\beta u_1/u$ again, we obtain, without solving any further equations (as described in Proposition 27), the quite general solution of Burgers' equation:

$$v(x,t) = -2\beta \frac{u_x(x,t)}{u(x,t)} = \frac{\int_{-\infty}^{\infty} dz\, \frac{x-z}{t} \exp\left(-\frac{1}{2\beta}\left[\frac{(x-z)^2}{2t} + \int_{z_0}^{z} d\sigma\ A(\sigma)\right]\right)}{\int_{-\infty}^{\infty} dz\, \exp\left(-\frac{1}{2\beta}\left[\frac{(x-z)^2}{2t} + \int_{z_0}^{z} d\sigma\ A(\sigma)\right]\right)} \tag{168}$$

This well-known result also appears as a Bäcklund transformation in [44]. The example is supposed to show how it arises in the present framework as a special case of a solution transfer, relating symmetries/quotient equations to correspondences which in turn can give rise to generalized notions of Bäcklund transformations.

9 Application to Electrodynamics and Hydrodynamics

In this section, the framework is applied to study some aspects of electrodynamics and hydrodynamics in order to illustrate the general aspects outlined in the last sections.

1. In the first subsection, formal integrability of Maxwell's equations is shown. This is a well-known result but provided for completeness.
2. In the second subsection, the shared structure of Maxwell's equations in vacuum and the wave equations is computed and Maxwell's equations in vacuum are identified as an auto-Bäcklund correspondence of the wave equation.
3. In the third subsection, it is shown that electrodynamics, formulated in terms of gauge potentials, is equivalent up to gauge symmetries to electrodynamics, formulated in terms of Faraday tensors, in the precise sense of Definition 22.
4. The fourth subsection picks up the motivating example of Sect. 3.1 and the shared structure of magneto-statics and the incompressible, viscous Navier-Stokes equation. It is shown that the integrability conditions coming out of the formalism are exactly those physical assumptions that had to be guessed in the motivating example.

9.1 Formal Integrability of Maxwell's Equations

Let M be our spacetime with local coordiantes $(x^0, \cdots, x^3)$ and $E = TM$ an 8-dimensional bundle, $\pi : E \to M$, which locally has the form $U \times \mathbb{R}^4$, $U \subset M$ with local coordinates $(x^0, \cdots, x^3, A^0, \cdots, A^3)$. We abbreviate those local coordinates with (x^μ, A^μ). A^μ are the local coordinates of the *gauge potential* of electrodynamics. In the present context, they are coordinate functions $A^\mu : J^0(\pi) \to \mathbb{R}$ and they should not be confused with sections $A^\mu : M \to J^0(\pi)$, $\mathbf{x} \mapsto A^\mu(\mathbf{x})$. One can prolong $J^0(\pi)$ to $J^2(\pi)$ to obtain the local coordinates

$$(x^\mu, A^\mu, A^{\mu,\nu}, A^{\mu,\nu\lambda}). \tag{169}$$

As second derivatives commute, the relation $A^{\mu,\nu\lambda} = A^{\mu,\lambda\nu}$ holds for the corresponding coordinate functions of the prolongation. Thus, $J^2(\pi) = J^2(4,4)$ is a space with $4 + 4 + 4^2 + 4 \cdot 4 \cdot (4+1)/2 = 24 + 40 = 64$ dimensions. Furthermore, we let $g : TM \otimes TM \to C^\infty(M)$ be the Lorentzian metric of our spacetime M. It is an element of $T^*M \otimes T^*M$. In local coordinates, it can be written $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$. If one assumes that the metric is given (e.g. as solution of the Einstein equations) and that the sources $J^\nu : M \to \mathbb{R}$ are also given, one can locally describe Maxwell's equations as the kernel of the differential operator[12]

$$\varphi : J^2(E) \to E, \qquad (x^\mu, A^\mu, A^{\mu,\nu}, A^{\mu,\nu\lambda}) \mapsto (x^\mu, g_{\nu\lambda} A^{[\nu,\mu]\lambda} - J^\mu). \tag{170}$$

Proposition 28 $\mathcal{E} = \ker(\varphi)$ *is involutive and thus formally integrable.*

Proof The prolongation $P^1(\mathcal{E})$ only involves new constraints on 3rd order coordinates. As a result, $\pi^3_2 : P^1(\mathcal{E}) \to \mathcal{E}$ is surjective. Let us check if the other two conditions of Proposition 16 are fulfilled.

$$\begin{aligned}\sigma(\varphi) &= a^{\rho,\kappa\theta} \frac{\partial \varphi^h}{\partial A^{\rho,\kappa\theta}} \frac{\partial}{\partial w^h} \\ &= a^{\rho,\kappa\theta} \frac{g_{\nu\lambda} \partial A^{\nu,\mu\lambda}}{\partial A^{\rho,\kappa\theta}} \frac{\partial}{\partial w^\mu} - a^{\rho,\kappa\theta} \frac{g_{\nu\lambda} \partial A^{\mu,\nu\lambda}}{\partial A^{\rho,\kappa\theta}} \frac{\partial}{\partial w^\mu} \\ &= \left(a^{\nu,\mu\lambda} g_{\nu\lambda} - a^{\mu,\nu\lambda} g_{\nu\lambda} \right) \frac{\partial}{\partial w^\mu} = g_{\nu\lambda} a^{[\nu,\mu]\lambda} \frac{\partial}{\partial w^\mu}\end{aligned} \tag{171}$$

Over pairs of indices is summed and thus, those are in total 4 equations. When calculating the rank of the symbol, those 4 equations impose 4 constraints. This means (recall that $\dim(E) = m + e = 4 + 4$)

---

[12] The notation $a_{[\mu_1 \dots \mu_n]}$ means antisymmetrisation of the indices, e.g. $F_{[\nu\lambda,\mu]} = F_{\nu\lambda,\mu} - F_{\mu\lambda,\nu} + F_{\mu\nu,\lambda} - F_{\lambda\nu,\mu} + F_{\lambda\mu,\nu} - F_{\nu\mu,\lambda}$ or $g_{\mu\lambda} A^{[\nu,\mu]\lambda} = g_{\mu\lambda} A^{\nu,\mu\lambda} - g_{\mu\lambda} A^{\mu,\nu\lambda}$. The Einstein sum convention is used.

$$\dim(g^2) \overset{(38)}{=} e\binom{m-1+2}{2} - e \overset{(m=4=e)}{=} 4\binom{5}{2} - 4 = 4 \cdot 10 - 4 = 36. \tag{172}$$

Next, calculate the prolongation:

$$\mu \circ \sigma^1(\phi) \overset{(52)}{=} a^{\nu,\mu\lambda\theta} \frac{\partial(D_o\phi^h)}{\partial A^{\nu,\mu\lambda\theta}} \frac{\partial}{\partial w_o^h} = g_{\nu\lambda} a^{[\nu,\mu]\lambda\theta} \frac{\partial}{\partial w_\theta^\mu} \tag{173}$$

Those are in total 16 equations. However, the rank of the system might be lower if some of them are functionally dependent. A small program was implemented that generates the corresponding matrix and calculates the rank. The code of this program is given in [43]. The program delivers the rank 15 for the system above for 4 dimensions This means, one of the functions depends on the others. Therefore, we obtain

$$\dim(g^3) = 4\binom{6}{3} - 15 = 80 - 15 = 65. \tag{174}$$

The dimension is constant for every local neighbourhood and thus g^3 is a smooth vector bundle over $\mathcal{J}^2$.

We can use (65) to obtain the dimensions of $F^{2,j}_{\mathcal{J}^2}$.

$$\dim(F^{2,j}_{\mathcal{J}^2}) = e \cdot \binom{m-1-j+2}{2} \tag{175}$$

Let us give an explicit basis for them

$$F^{2,j}_{\mathcal{J}^2} = \left\{ \text{span}\left(dx^l \vee dx^n \otimes \partial_u^k\right) \mid j+1 \leq l \leq n \leq m \right\} \tag{176}$$

We can obtain the intersection by restricting $\sigma(\phi)$ to $F^{2,j}_{\mathcal{J}^2}$:

$$\begin{aligned} g^{2,j} &= g^2 \cap F^{2,j}_{\mathcal{J}^2} = \ker(\sigma(\phi)|_{F^{2,j}_{\mathcal{J}^2}}) \\ &= \left\{\left(0 = g_{\nu\lambda}(a^{\mu,\nu\lambda} - a^{\nu,\mu\lambda}) \mid \text{ last two indices } \in \{j+1, \cdots, m\}\right)\right\} \end{aligned} \tag{177}$$

For $j < m-1$, the above equation always gives rise to $e = 4$ different conditions on the components $a^{\mu,\nu\lambda}$ because the last two indices can be chosen differently. However, for $j = m-1$ one obtains the equation

$$g_{mm}a^{\mu,mm} - g_{\nu m}a^{\nu,\mu m}\delta^{\mu,m} \tag{178}$$

And this means that for $\mu = m$, the last term of the matrix of derivatives of the equation above with respect to $a^{\mu,\nu\lambda}$ (whose rank corresponds to the rank of the system) vanishes. Then they impose one condition less.

In accordance with this, the computer program delivers:

$$\mathrm{rank}\sigma(\phi)|_{F^{2,1}} = 4, \qquad \mathrm{rank}\sigma(\phi)|_{F^{2,2}} = 4, \qquad \mathrm{rank}\sigma(\phi)|_{F^{2,3}} = 3 \tag{179}$$

All in all, we obtain

$$\begin{aligned}\dim(g^2) + \sum_{j=1}^{3} \dim(g^{2,j}) &= 36 + 4\binom{4}{2} - 4 + 4\binom{3}{2} - 4 + 4\binom{2}{2} - 3 \\ &= \dim(g^3).\end{aligned} \tag{180}$$

Thus, the system is formally integrable. □

Note that formal integrability of the Yang-Mills-Higgs equations was shown for arbitrary dimensions in 1996 by Giachetta and Mangiarotti [47].

9.2 Embedding of Vacuum Electrodynamics in Wave Equations

As is well-known, when considering Maxwell's equations in flat spacetime in vacuum (without sources and in Gaussian units)

$$\nabla \cdot \mathbf{E} = 0, \qquad \nabla \times \mathbf{E} = -\frac{1}{c}\partial_t \mathbf{B}, \qquad \nabla \times \mathbf{B} = \frac{1}{c}\partial_t \mathbf{E}, \qquad \nabla \cdot \mathbf{B} = 0, \tag{181}$$

one can derive wave equations as follows

$$\begin{aligned}\partial_t^2 \mathbf{B} &= -c\,\partial_t(\nabla \times \mathbf{E}) = -c^2\,\nabla \times (\nabla \times \mathbf{B}) = -c^2\,\nabla \cdot (\nabla \cdot \mathbf{B}) + c^2 \Delta \mathbf{B} = c^2 \Delta \mathbf{B} \\ \partial_t^2 \mathbf{E} &= -c\,\partial_t(\nabla \times \mathbf{B}) = -c^2\,\nabla \times (\nabla \times \mathbf{E}) = -c^2\,\nabla \cdot (\nabla \cdot \mathbf{E}) + c^2 \Delta \mathbf{E} = c^2 \Delta \mathbf{E}\end{aligned} \tag{182}$$

In the following is shown how Maxwell's equations in vacuum can be understood as a Bäcklund correspondence for the wave equations.

As can be seen, the wave equations are differential consequences of Maxwell's equations. Furthermore, the consequences separate into constraints imposed solely on **B** and **E**. This suggests to understand (181) as a correspondence between the two wave equations.

Indeed, if one defines the bundle $\pi : E := M \times \mathbb{R}^3 \to M$ where $M = \mathbb{R}^4$ in this case, with local coordinates (t, x^i, E^j), $i, j \in \{1, \cdots, 3\}$ and the bundle $\xi : F \simeq E \to M$ with coordinates (t, x^i, B^j), then one can define Maxwell's equations in vacuum as a correspondence Φ on the product space $J := J^2(E) \times_M J^2(F)$ by

$$\Phi : \begin{cases} E^{i,i} = 0, & \varepsilon_{ijk} E^{k,j} = -\frac{1}{c} B^i_t \\ B^{i,i} = 0, & \varepsilon_{ijk} B^{k,j} = \frac{1}{c} E^i_t \end{cases} \tag{183}$$

The compatibility conditions $\mathcal{P}(\Phi) = \pi^1_J(P^1(\Phi))$ for Φ are given by

$$\mathcal{P}(\Phi) = \ker \begin{pmatrix} E^{i,i}, & E^{i,ij}, & E^{i,i}_t \\ c\ \varepsilon_{ijk} E^{k,j} + B^i_t, & c\ \varepsilon_{ijk} E^{k,jl} + B^{i,l}_t, & c\ \varepsilon_{ijk} E^{k,j}_t + B^i_{tt} \\ c\ \varepsilon_{ijk} B^{k,j} - E^i_t, & c\ \varepsilon_{ijk} B^{k,jl} - E^{i,l}_t, & c\ \varepsilon_{ijk} B^{k,j}_t - E^i_{tt} \\ B^{i,i}, & B^{i,il}, & B^{i,i}_t \end{pmatrix}, \tag{184}$$

Now, as already shown above, the entries [2,2] and [3,2] of the matrix can be inserted into the entries [3,2] and [3,3] to obtain the wave equations via the ε_{ijk}-identities.

$$\mathcal{E}_J : \left\{ E^i_{tt} = c^2\ E^{i,jj} \right\}, \quad \mathcal{F}_J : \left\{ B^i_{tt} = c^2\ B^{i,jj} \right\}. \tag{185}$$

Together, $\mathcal{E}_J \cap \mathcal{F}_J = \mathcal{E}\mathcal{F} \supset \pi^1_J(P^1(\Phi)) = \mathcal{P}(\Phi)$. Furthermore, since no other equations purely in terms of E or B coordinates are imposed, we have $\pi_E(\mathcal{P}(\Phi)) = \mathcal{E}$ and $\pi_F(\mathcal{P}(\Phi)) = \mathcal{F}$. Hence Φ is a strict Bäcklund correspondence.

Furthermore, there is a diffeomorphism $\mathcal{E} \simeq \mathcal{F}$ and therefore this Bäcklund correspondence is actually an auto-Bäcklund correspondence. This is a useful fact because (it is well-known that) auto-Bäcklund correspondences allow to generate an infinite amount of solutions. Indeed, by Proposition 24, solutions can be transferred from $\mathcal{E}$ to $\mathcal{F}$ by solving Φ. Since the process involves solving Φ, the solution obtained for $\mathcal{F}$ is in general different to the solution coming from $\mathcal{E}$. However, once such a solution of $\mathcal{F}$ is obtained, one can repeat the process because $\mathcal{E} \simeq \mathcal{F}$ and obtain a new solution of $\mathcal{F}$ and so on.

Another aspect that is shown quite clearly in this geometric product bundle setting, is that the space of all differential solutions of $\mathcal{E}_J$ and $\mathcal{F}_J$ contain the space of all differential solutions of Φ (because they are a differential consequence of Φ, i.e. $\mathcal{P}(\Phi) \subset \mathcal{E}\mathcal{F}$). Thus, one could say that the solution space of electrodynamics in vacuum is *embedded* into the solution spaces of the wave equations. Hence, once the most general solution of the wave equations is found (possibly by utilizing the auto-Bäcklund correspondence), one can restrict this general solution to the subspace of solutions of Maxwell's equations (that can be obtained simply by inserting the solutions into those equations) to obtain the general solution of Maxwell's equations in vacuum.

9.3 Equivalence Up to Gauge Symmetry

In this subsection, the aim is to derive Maxwell's equations in terms of Faraday tensors[13]

$$\mathcal{F}: \; \{ \, g_{\nu\lambda} F^{\nu\mu,\lambda} = J^{\mu}, \quad F^{[\nu\lambda,\mu]} = 0 \, \} \tag{186}$$

as a quotient equation by quotienting out gauge symmetries from the equations in terms of vector potentials

$$\mathcal{E}: \; \{ \, g_{\nu\lambda} A^{[\nu,\mu]\lambda} = J^{\mu} \, \} \tag{187}$$

using the methods introduced in Sect. 8. Among other things, this shall illustrate that the framework is versatile enough to answer the questions that the solution-Category approach described in [5] answers—though the way the answer is obtained is quite different.

The first equation above can be modeled on the jet bundle $J^1(F)$ where F is the total space of the bundle $\xi : F := TM \otimes TM \;\; \rightarrow \;\; M$ with local coordinates $(x^{\mu}, F^{\mu\nu})$. M is a Lorentzian spacetime, equipped with a Lorentzian metric $g \in T^*M \otimes T^*M$. Its local description reads $g = g_{\mu\nu} dx^{\mu} \otimes dx^{\nu}$.

The second equation can be modeled as submanifold $\mathcal{E} \subset J^2(E)$ over the bundle $\pi : E := TM \;\rightarrow\; M$ with local coordinates (x^{μ}, A^{μ}). $J^2(E)$ has local coordinates $(x^{\mu}, A^{\mu}, A^{\mu,\nu}, A^{\mu,\nu\lambda})$. As second derivatives commute, the corresponding relation $A^{\mu,\nu\lambda} = A^{\mu,\lambda\nu}$ also holds for the jet bundle coordinates.

The differential equation $\mathcal{E}$ is invariant under so called *gauge transformations*

$$x^{\mu} \;\rightarrow\; x^{\mu}, \qquad A^{\mu} \;\rightarrow\; A^{\mu} + \chi^{\,,\mu} \tag{188}$$

which prolonged to $J^2(E)$ take the form

$$A := \begin{pmatrix} x^{\mu} \\ A^{\mu} \\ A^{\mu,\nu} \\ A^{\mu,\nu\lambda} \end{pmatrix} \qquad \rightarrow \qquad A' := \begin{pmatrix} x^{\mu} \\ A^{\mu} \\ A^{\mu,\nu} \\ A^{\mu,\nu\lambda} \end{pmatrix} + \begin{pmatrix} 0 \\ \chi^{\,,\mu} \\ \chi^{\,,\mu\nu} \\ \chi^{\,,\mu\nu\lambda} \end{pmatrix}. \tag{189}$$

Note that because of the prolongation, we have $\chi^{\,,\mu\nu} = \chi^{\,,\nu\mu}$ and therefore, if we contract it with some tensor $T_{\mu\nu}$, we obtain

$$\chi^{\,,\mu\nu} T_{\mu\nu} = \chi^{\,,\mu\nu} \left(\frac{T_{\mu\nu} + T_{\nu\mu}}{2} + \frac{T_{\mu\nu} - T_{\nu\mu}}{2} \right) = \chi^{\,,\mu\nu} \frac{T_{\mu\nu} + T_{\nu\mu}}{2} \tag{190}$$

[13] As already mentioned in the footnote above Eq. (170), the notation $a_{[\mu_1 \ldots \mu_n]}$ means antisymmetrisation of the indices, e.g. $F_{[\nu\lambda,\mu]} = F_{\nu\lambda,\mu} - F_{\mu\lambda,\nu} + F_{\mu\nu,\lambda} - F_{\lambda\nu,\mu} + F_{\lambda\mu,\nu} - F_{\nu\mu,\lambda}$ or $g_{\mu\lambda} A^{[\nu,\mu]\lambda} = g_{\mu\lambda} A^{\nu,\mu\lambda} - g_{\mu\lambda} A^{\mu,\nu\lambda}$. The Einstein sum convention is used.

because the anti-symmetric part vanishes upon contraction. Similarly,

$$\begin{aligned}\chi^{,\mu\nu\lambda}T_{\mu\ ,\nu\lambda} &= \chi^{,\mu\nu\lambda}\left(\frac{T_{\mu\ ,\nu\lambda}+T_{\lambda\ ,\mu\nu}+T_{\nu,\ \lambda\mu}+T_{\mu\ ,\lambda\nu}+T_{\lambda\ ,\nu\mu}+T_{\nu,\ \mu\lambda}}{3!}\right)\\ &= \chi^{,\mu\nu\lambda}\frac{T_{\mu\ ,\nu\lambda}+T_{\lambda\ ,\mu\nu}+T_{\nu,\ \lambda\mu}}{3}\end{aligned} \tag{191}$$

The gauge transformation can be rewritten as the action of group elements on A to extract the generators X of this transformation.

$$\begin{aligned}A' &= \begin{pmatrix} x^{\mu} \\ A^{\mu} \\ A^{\mu,\nu} \\ A^{\mu,\nu\lambda}\end{pmatrix} + \begin{pmatrix} 0 \\ \chi^{,\mu} \\ \chi^{,\mu\nu} \\ \chi^{,\mu\nu\lambda}\end{pmatrix}\\ &= \exp\left[\chi^{,\mu}\frac{\partial}{\partial A^{\mu}} + \chi^{,\mu\nu}\frac{\partial}{\partial A^{\mu,\nu}} + \chi^{,\mu\nu\lambda}\frac{\partial}{\partial A^{\mu,\nu\lambda}}\right]\begin{pmatrix} x^{\mu} \\ A^{\mu} \\ A^{\mu,\nu} \\ A^{\mu,\nu\lambda}\end{pmatrix}\\ &\overset{(191)}{=} \exp\left[\chi^{,\mu}\underbrace{\frac{\partial}{\partial A^{\mu}}}_{=:X^{\mu}} + \frac{1}{2}\chi^{,\mu\nu}\underbrace{\left(\frac{\partial}{\partial A^{\mu,\nu}} + \frac{\partial}{\partial A^{\nu,\mu}}\right)}_{=:X^{\mu\nu}}\right.\\ &\qquad\left. + \frac{1}{3}\chi^{,\mu\nu\lambda}\underbrace{\left(\frac{\partial}{\partial A^{\mu,\nu\lambda}} + \frac{\partial}{\partial A^{\nu,\mu\lambda}} + \frac{\partial}{\partial A^{\lambda,\mu\nu}}\right)}_{=:X^{\mu\nu\lambda}}\right]\begin{pmatrix} x^{\mu} \\ A^{\mu} \\ A^{\mu,\nu} \\ A^{\mu,\nu\lambda}\end{pmatrix}\end{aligned} \tag{192}$$

If $\chi^{,\mu\nu}$ were different from $\chi^{,\nu\mu}$, then $\partial/\partial A^{\mu,\nu}$ and $\partial/\partial A^{\nu,\mu}$ would be two different generators but because $\chi^{,\mu\nu} = \chi^{,\nu\mu}$, we obtain the generator $X^{\mu\nu}$. Similarly for $\chi^{,\mu\nu\lambda}$ and $X^{\mu\nu\lambda}$.

As a consequence, to obtain a functionally independent set of Invariants of gauge transformations, we use Eq. (146) and obtain

$$X^{\mu}(I) = 0, \qquad X^{\mu\nu}(I) = 0, \qquad X^{\mu\nu\lambda}(I) = 0. \tag{193}$$

Proposition 29 *This system can only be solved if I is a function of x^{μ} and*

$$I^{\mu\nu} := A^{\mu,\nu} - A^{\nu,\mu} = A^{[\mu,\nu]} \tag{194}$$

and its prolongations $I^{\mu\nu,\lambda} = A^{[\mu,\nu]\lambda}, \cdots, I^{\mu\nu,\lambda_1\ldots\lambda_k} = A^{[\mu,\nu]\lambda_1\ldots\lambda_k}$ and so on.

Proof That the generators annihilate x^μ is trivial because they only contain derivatives w.r.t. the dependent variables. For deriving (194), let us consider the equations order by order:

1. $0 = X^\mu(I) = \partial I/\partial A^\mu$ implies that I does not depend on A^μ.
2. Now we have $0 = X^{\mu\nu}(I) = \partial I/\partial A^{\mu,\nu} + \partial I/\partial A^{\nu,\mu}$. The general dependence of I can be found by a coordinate transformation. First, let us fix some indices μ, ν, λ and then define $x_1 := A^{\mu,\nu}$ and $x_2 := A^{\nu,\mu}$ such that the above equation takes the form $0 = \partial I/\partial x_1 + \partial I/\partial x_2$. Now, we introduce the transformation

$$\begin{pmatrix} x_1' := x_1 \\ x_2' := x_2 - x_1 \end{pmatrix} \quad \Rightarrow \quad \frac{\partial}{\partial x_i} = \frac{\partial x_1'}{\partial x_i}\frac{\partial}{\partial x_1'} + \frac{\partial x_2'}{\partial x_i}\frac{\partial}{\partial x_2'} \tag{195}$$

   Thus, $\partial/\partial x_1 = \partial/\partial x_1' - \partial/\partial x_2'$ and $\partial/\partial x_2 = \partial/\partial x_2'$. Therefore

$$0 = \left(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right) I = \frac{\partial I}{\partial x_1'}. \tag{196}$$

   This implies that I can only be any function of $I^{\mu\nu} := x_2' = x_2 - x_1 = A^{\mu,\nu} - A^{\nu,\mu}$. This goes through for any choice of μ, ν, λ.
3. $0 = X^{\mu\nu\lambda}(I)$ implies

$$0 = \left(\frac{\partial}{\partial A^{\mu\ ,\nu\lambda}} + \frac{\partial}{\partial A^{\nu\ ,\mu\lambda}} + \frac{\partial}{\partial A^{\lambda\ ,\mu\nu}}\right) I$$

   We employ the same method as above. We define $x_1 = A^{\mu,\nu\lambda}, \cdots, x_3 = A^{\lambda,\mu\nu}$ and the transformation

$$\begin{pmatrix} x_1' = x_1, \\ x_2' = x_2 - x_1, \\ x_3' = x_3 - x_1 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} \frac{\partial}{\partial x_1} = \frac{\partial}{\partial x_1'} - \frac{\partial}{\partial x_2'} - \frac{\partial}{\partial x_3'}, \\ \frac{\partial}{\partial x_2} = \frac{\partial}{\partial x_2'}, \\ \frac{\partial}{\partial x_3} = \frac{\partial}{\partial x_3'}. \end{pmatrix} \tag{197}$$

   Thus,

$$\sum_i \frac{\partial I}{\partial x_i} = \frac{\partial I}{\partial x_1'} \tag{198}$$

   implying that I is a function of $x_2' = x_2 - x_1$ and $x_3' = x_3 - x_1$. Observe that $x_2' - x_3' = x_2 - x_3$ which means that this system is linearly equivalent to the system $x_i - x_j, i, j \in \{1, 2, 3\}$.

   Thus, we can say I to this order only depends on

$$I^{\mu\nu\lambda} := A^{\mu,\nu\lambda} - A^{\nu,\lambda\mu} = A^{[\mu,\nu]\lambda} \tag{199}$$

   or any permutation thereof in μ, ν, λ.

If $A^{\mu}(\mathbf{x})$ is a section, then $A^{[\mu,\nu]\lambda}(\mathbf{x}) = \nabla^{\lambda} A^{[\mu,\nu]}(\mathbf{x})$ and therefore $I^{\mu\nu\lambda} = I^{\mu\nu,\lambda}$ as desired.

If we prolong the bundle further, this idea continuous for higher orders. For order n, the equation $X(I) = 0$ gives

$$0 = \left(\frac{\partial}{\partial A^{\lambda_1,\lambda_2\ldots\lambda_n}} + \frac{\partial}{\partial A^{\lambda_2,\lambda_3\ldots\lambda_1}} + \cdots + \frac{\partial}{\partial A^{\lambda_n,\lambda_1\ldots\lambda_{n-1}}}\right) I =: \sum_{i=1}^{n} \frac{\partial I}{\partial x_i} \tag{200}$$

Thus, with the transformation

$$\begin{pmatrix} x'_1 = x_1, \\ x'_2 = x_2 - x_1, \\ \cdots \\ x'_n = x_n - x_1 \end{pmatrix} \Rightarrow \begin{pmatrix} \frac{\partial}{\partial x_1} = \frac{\partial}{\partial x'_1} - \frac{\partial}{\partial x'_2} - \cdots - \frac{\partial}{\partial x'_n}, \\ \frac{\partial}{\partial x_2} = \frac{\partial}{\partial x'_2}, \\ \cdots \\ \frac{\partial}{\partial x_n} = \frac{\partial}{\partial x'_n}. \end{pmatrix} \tag{201}$$

we obtain $\partial I/\partial x'_1 = 0$ and therefore I only depends on $x_i - x_1$ or, equivalently, on

$$x_i - x_j = A^{[\lambda_i,\lambda_j]\lambda_1\ldots\lambda_{i-1}\lambda_{i+1}\ldots\lambda_{j-1}\lambda_{j+1}\ldots\lambda_n} = I^{\mu\nu,\lambda_1\ldots\lambda_k} \tag{202}$$

proving the claim. □

Thus, apart from x^{μ}, the $I^{\mu\nu}$ are our only functionally and differentially independent Invariants. Their degree is $n = 1$ because they only involve functions from $J^1(E)$. Define $N = \max(k, n) = k = 2$, $L = N - n = 1$. As described in Sect. 8, one can now create a new bundle $\xi : Q \rightarrow M$ with the same base space M and where Q is the bundle on which the $I^{\mu,\nu}$ live, i.e. $TM \otimes TM$. It is given the local coordinates $(x^{\mu}, F^{\mu\nu})$ whose number coincides with the number of the $I^{\mu\nu}$. Next, the quotient correspondence $\Phi(I)$ is defined on $J^2(E) \times_M J^0(Q)$,

$$\Phi(I) : \{ F^{\mu\nu} = I^{\mu\nu} = A^{[\mu,\nu]} \} \tag{203}$$

By our general theory, prolonging the equation $Q(I) := (\pi'_E)^{-1}(\mathcal{E}) \cap \Phi(I)$ should give rise to compatibility conditions only involving the $F^{\mu\nu}$-coordinates. Indeed, a prolongation of $\Phi(I)$ results in

$$P^1(\Phi(I)) : \{ F^{\mu\nu} = A^{[\mu,\nu]},\ F^{\mu\nu,\lambda} = A^{[\mu,\nu]\lambda} \} \tag{204}$$

Thus, intersection with the prolongation of $(\pi'_E)^{-1}(\mathcal{E})$ (cf. Eq. (187)) results in the compatibility condition

$$\begin{aligned} P^1(Q(I)) \ : \ \{F^{\mu\nu} = A^{[\mu,\nu]},\ F^{\mu\nu,\lambda} = A^{[\mu,\nu]\lambda} \\ g_{\nu\lambda}A^{[\nu,\mu]\lambda} = J^{\mu},\ g_{\nu\lambda}A^{[\nu,\mu]\lambda\theta} = J^{\mu,\theta}, \\ \Rightarrow g_{\nu\lambda}F^{\nu\mu,\lambda} = g_{\nu\lambda}A^{[\nu,\mu]\lambda} = J^{\mu}, \\ F^{[\nu\lambda,\mu]} = A^{[\nu,\lambda\mu]} = 0\,\} \ \subset J^3(E) \times J^1(Q), \end{aligned} \tag{205}$$

where $A^{[\nu,\lambda\mu]} = 0$ always holds because $A^{\mu,\nu\lambda} = A^{\mu,\lambda\nu}$.

Equations purely in terms of coordinates of Q thus arise already after one prolongation. The natural product bundle is thus $J := J^2(E) \times_M J^1(Q)$, and defining $\mathcal{Q} := (\pi^{2,\,1}_{2,\,0})^{-1}(Q(I))$ as in Eq. (152), one obtains the following equation $\mathcal{F} \subset J^1(Q)$ from the compatibility conditions $\mathcal{P}(\mathcal{Q}) = \pi^1_J(P^1(\mathcal{Q}))$:

$$\mathcal{F} \overset{(153)}{=} \pi_F(\mathcal{P}(\mathcal{Q})) : \{\, g_{\nu\lambda}F^{\nu\mu,\lambda} = J^{\mu},\ \ F^{[\nu\lambda,\mu]} = 0\,\} \subset J^1(Q). \tag{206}$$

Therefore, one indeed obtains Maxwell's equations in terms of Faraday tensors. Hence, as defined in Definition 22, $\mathcal{E}$ and $\mathcal{F}$ are equivalent up to symmetry and $\mathcal{F} = \pi_F(\mathcal{P}(\mathcal{Q}))$ is the quotient equation of $\mathcal{E}$.

Thus, "adding Morphisms of some group" in the solution-Category can be compared with "finding the invariant equation with respect to some group" in the category of smooth manifolds where differential equations are submanifolds of jet spaces. The procedure in the category of smooth manifolds might be computationally more involved but in contrast to the solution-Category approach, it delivers all invariants of the symmetry and it produces the corresponding quotient equation without the need to know it before. Furthermore, it enables to see connections and find solutions of many systems of PDEs that result from solution transfer to the quotient as detailed in Proposition 27 and also from the quotient back to the original equation (here, for example, the quotient equation is a system of lower order).

9.4 Shared Structure of Magneto-Statics and Hydrodynamics

In this subsection, the motivating example in Sect. 3.1 is picked up. In particular, the assumption of a static fluid flow, guessed in (15), arise as the result of the computation of the minimal integrability conditions for shared structure under the given correspondence.

The notation that is used in the following computations is the one introduced in Example 3.3. In particular, $\mathcal{E}$ is given by (25), $\mathcal{F}$ by (26), the correspondence by (27) and the intersection by (28), copied here for convenience:

$$\mathcal{I} = \pi_E^{-1}(\mathcal{E}) \cap \pi_F^{-1}(\mathcal{F}) \cap \Phi : \left\{ \begin{aligned} & u_t^i + u^j u^{i,j} = -\frac{1}{\rho} p^{,i} + \nu u^{i,jj}, \, u^{i,i} = 0 \\ & \varepsilon_{ijk} B^{k,j} = I^i, \;\; B^{i,i} = 0 \\ & B^i = \varepsilon_{ijk} u^{k,j} \end{aligned} \right\} \tag{207}$$

The first prolongation of Φ leads to

$$P^1(\Phi) : \left\{ B^i = \varepsilon_{ijk} u^{k,j} \quad \middle| \quad \begin{aligned} B^{i,l} &= \varepsilon_{ijk} u^{k,jl} \\ B_t^i &= \varepsilon_{ijk} u_t^k \end{aligned} \right\} \tag{208}$$

With the additional relations, all equations in (207) can be expressed in terms of the coordinates of $J^2(E)$. In particular, we obtain for the middle row of (207),

$$\begin{aligned} \varepsilon_{ijk} B^{k,j} = \varepsilon_{ijk} \varepsilon_{klm} u^{m,lj} = (\delta_l^i \delta_m^j - \delta_m^i \delta_l^j) u^{m,lj} = u^{j,ij} - u^{i,jj} - I^i \\ B^{i,i} = \varepsilon_{ijk} u^{k,ji} = -\varepsilon_{ijk} u^{k,ji} = 0 \end{aligned} \tag{209}$$

where we used that ε_{ijk} is antisymmetric and thus annihilates $u^{k,ji}$ because it is symmetric in ji.

Since all relations in (207) are now expressed in terms of coordinates of $J^2(E)$ (and $B^{i,i} = 0$ is trivially fulfilled), formal integrability of the whole system amounts to formal integrability of the following system on $J^2(E)$.

$$\mathcal{I}^2 : \left\{ \begin{aligned} & u_t^i + u^j u^{i,j} + \frac{1}{\rho} p^{,i} = \nu u^{i,jj}, \qquad u^{i,i} = 0 \\ & u^{j,ij} - u^{i,jj} = I^i \end{aligned} \right\} \tag{210}$$

On $\mathcal{I}^2$, $u^{i,jj} = u^{j,ij} - I^i$. Thus, we can rewrite the first line as $u_t^i + u^j u^{i,j} + \frac{1}{\rho} p^{,i} + \nu I^i - \nu u^{j,ij}$. To simplify the problem, let us assume that

$$\nu I^i = -p^{,i}/\rho, \tag{211}$$

corresponding to the first of the two assumptions in (15). Then the above system is equivalent to the system

$$\mathcal{I}^2 = \ker \begin{pmatrix} u_t^i + u^j u^{i,j} - \nu u^{j,ij} \\ u^{i,i} \\ u^{j,ij} - u^{i,jj} - I^i \end{pmatrix} \tag{212}$$

Proposition 30 *The system (212) is not formally integrable without adding the integrability condition*

$$\mathcal{B}(\mathcal{I}^2) : \left\{ u^{i,ik} = u^{j,kj} = 0 \right\} \quad \Rightarrow \quad \frac{d\boldsymbol{u}}{dt} = 0. \tag{213}$$

and thereafter, for $u^i \neq 0$, becomes involutive and thus formally integrable.

Proof Consider the first prolongation

$$\mathcal{I}^3 = \ker \left(\begin{array}{c|c} u^i_t + u^j u^{i,j} - \nu u^{j,ij} & \begin{array}{c} u^{i,k}_t + u^{j,k} u^{i,j} + u^j u^{i,jk} - \nu u^{j,ijk} \\ u^i_{tt} + u^j_t u^{i,j} + u^j u^{i,j}_t - \nu u^{j,ij}_t \end{array} \\ u^{i,i} & \begin{array}{c} u^{i,ik} \\ u^{i,i}_t \end{array} \\ u^{j,ij} - u^{i,jj} - I^i & \begin{array}{c} u^{j,ijk} - u^{i,jjk} - I^{i,k} \\ u^{j,ij}_t - u^{i,jj}_t - I^i_t \end{array} \end{array} \right) \tag{214}$$

Due to the term $u^{i,ik}$ on the right side, which is set to 0 when considering the kernel, constraints on coordinates of order 2 are imposed. Furthermore, the third equation simplifies to $u^{i,jj} = I^i$. As a consequence, $\pi^3_2 : \mathcal{I}^3 \rightarrow \mathcal{I}^2$ is not surjective, violating the first condition of Proposition 14. Thus, the system is not formally integrable without adding those integrability conditions to $\mathcal{I}^2$.

As explained in detail in Sect. 5.3, those new constraints can be understood as the *minimal conditions under which the intersection is differentially consistent*. The conditions are

$$u^{i,ik} = u^{j,kj} = 0 \tag{215}$$

and thus $du^i/dt \mathrel{\widehat{=}} u^i_t + u^j u^{i,j} = 0$. This means the consistency conditions induce the constraint of a static fluid flow.

Let us therefore define a new system (as explained in Sect. 5.3) which takes those consistency conditions up to order two into account:

$$\Rightarrow \mathcal{J}^2 = \ker \left(\begin{array}{c} u^{j,j},\ u^{j,j}_t,\ u^{j,ji} \\ u^i_t + u^j u^{i,j},\ (u^i_t + u^j u^{i,j})^{,k},\ (u^i_t + u^j u^{i,j})_t \\ u^{i,jj} + I^{i,k} \end{array} \right) \tag{216}$$

The prolongation $\mathcal{J}^3$ now by construction either does not lead to equations not contained in $\mathcal{J}^2$ or the prolonged terms always involve at least one 3rd order coordinate. For example, the term $(u^i_t + u^j u^{i,j})^{,kl}$ can be solved for $u^{i,kl}_t$ and is thus only turned into a constraint on a coordinate of order 3.

As a result, $\pi^3_2 : \mathcal{J}^3 \rightarrow \mathcal{J}^2$ is surjective. To verify involutivity, let us check if the other two conditions of Proposition 16 are fulfilled.

$$\begin{aligned} \sigma(\phi) &= a^{j,kl} \frac{\partial \phi^h}{\partial u^{j,kl}} \frac{\partial}{\partial w^h} + a^{j,k}_t \frac{\partial \phi^h}{\partial u^{j,k}_t} \frac{\partial}{\partial w^h} + a^j_{tt} \frac{\partial \phi^h}{\partial u^j_{tt}} \frac{\partial}{\partial w^h} \\ &= a^{j,j}_t \partial^2_w + a^{j,ji} \partial^3_w + \left(a^{i,k}_t + u^j a^{i,jk} \right) \partial^5_w \\ &\qquad + \left(a^i_{tt} + u^j a^{i,j}_t \right) \partial^6_w + a^{i,jj} \partial^7_w \end{aligned} \tag{217}$$

This and the system (218) below are quite high dimensional systems. Thus, a small computer program was implemented to determine their rank. The code is given in the appendix of the Master's thesis (cf. [43]). It facilitates to generate the matrix corresponding to the tensor equations automatically.

When counting all components of the above equations, one obtains 19 but calculating the rank with the program gives us 18 constraints (i.e. there is one linear dependence). Note that even though $a_t^{i,k}$ and a_{tt}^{i} depend on u^j due to the non-linearity, they depend on it in a smooth way and thus g^2 has the same dimension everywhere and is a smooth vector bundle over $\mathcal{J}^2$.

Next, we have to calculate the prolongation:

$$\begin{aligned}\ker\sigma^1(\phi) &\overset{(52)}{=} \ker\Bigg(a^{j,klm}\frac{\partial(D_n\phi^h)}{\partial u^{j,klm}}\frac{\partial}{\partial w_n^h} + a_{tt}^{j,k}\frac{\partial(D_t\phi^h)}{\partial u_{tt}^{j,k}}\frac{\partial}{\partial w_t^h}\\ &\qquad + a_t^{j,kl}\frac{\partial(D_t\phi^h)}{\partial u_t^{j,kl}}\frac{\partial}{\partial w_t^h} + a_t^{j,kl}\frac{\partial(D_n\phi^h)}{\partial u_t^{j,kl}}\frac{\partial}{\partial w_n^h}\Bigg) \\ &= \ker\begin{pmatrix} a_t^{j,jn},\ a_{tt}^{j,j},\ a^{j,jin} \\ a_t^{i,kn} + u^j a^{i,jkn},\ a_{tt}^{i,k} + u^j a_t^{i,jk},\ a_{ttt}^{i} + u^j a_{tt}^{i,j} \\ a^{i,jjn},\ a_t^{i,jj}\end{pmatrix}\end{aligned} \tag{218}$$

If all equations of this system are taken to be independent, then this imposes $3 + 1 + 3! + 3\cdot 3! + 3\cdot 3 + 3 + 3\cdot 3 + 3 = 52$ constraints. However, the program computes the rank to be 44 (i.e. there are 8 linear dependencies). If one sets $u^i = 0$, the program still returns 8 in accordance to what was said before (in particular this constancy means that g^{k+1} is a smooth vector bundle everywhere). Thus, so far we obtain

$$\dim(g^2) \overset{(38)}{=} 3\binom{4-1+2}{2} - 18 = 12,\ \dim(g^3) = 3\binom{4-1+3}{3} - 44 = 16. \tag{219}$$

If we want to show that the system is formally integrable, then it remains to show that $\dim(g^{2,1}) + \dim(g^{2,2}) + \dim(g^{2,3}) = 4$. To calculate this, we consider the kernel of $\sigma(\phi)$ restricted to $F_{\mathcal{J}^2}^{2,j}$. For $g^{2,1}$, this means that all a_t's fall away. Thus, we obtain

$$\sigma(\phi)|_{F_{\mathcal{J}^2}^{2,1}} = a^{j,ji}\partial_w^3 + u^j a^{i,jk}\partial_w^5 + a^{i,jj}\partial_w^7 \tag{220}$$

Using the program again, one obtains the rank 14. Thus,

$$\dim(g^{2,j}) = \dim(F_{\mathcal{J}^2}^{2,j}) - e = e\cdot\binom{m-j+1}{2} - e \tag{221}$$

$$\dim(g^{2,1}) = \dim(F^{2,1}_{\mathcal{J}^2}) - 14 = e \cdot \binom{m-j+1}{2} - 14 = 3 \cdot \binom{4}{2} - 14 = 18 - 14 = 4. \tag{222}$$

Note, however, that the rank changes to 6 if one sets $u^i = 0$ above in Eq. (220). This means that the system is *not* involutive for $u^i = 0$.

Now, for $u^i \neq 0$, it remains to show that $\dim(g^{2,2}) = 0 = \dim(g^{2,3})$. For them, we obtain the same system as above but the range of the derivatives now only covers the coordinates 2 and 3. For $\ker(\varphi)|_{F^{2,2}_{\mathcal{J}^2}}$, the program gives us the rank 9 (and the rank 5 for $u^i = 0$). For $\ker(\varphi)|_{F^{2,3}_{\mathcal{J}^2}}$, it delivers rank 3 (and also rank 3 for $u^i = 0$). Thus,

$$\dim(g^{2,2}) = \dim(F^{2,2}_{\mathcal{J}^2}) - 9 = e \cdot \binom{m-j+1}{2} - 9 = 3 \cdot \binom{3}{2} - 9 = 9 - 9 = 0.$$
$$\dim(g^{2,3}) = \dim(F^{2,3}_{\mathcal{J}^2}) - 3 = e \cdot \binom{m-j+1}{2} - 3 = 3 \cdot \binom{2}{2} - 3 = 3 - 3 = 0. \tag{223}$$

Therefore, for $u^i \neq 0$, the system is involutive and thus formally integrable. □

Remark As explained in Sect. 5.3, the integrability conditions (213), can be interpreted as the minimal amount of physical assumptions that have to be made in order to reach consistency. Observe how, at this point, the consistency conditions emerge from the formalism without the need to guess them as in the motivating Example 3.1 (second assumption of (15)).

Now using the definitions introduced in Sect. 6 about shared structure, one can make the following conclusions. Hydrodynamics of an incompressible fluid and magneto-statics *share structure* under a *linear correspondence of first order* Φ in case that the fluid flow strength $\mathbf{u}$ is not zero and condition (211) holds. (The formal closure is then $\mathcal{B} := \mathcal{I} \cap \mathcal{B}(\mathcal{I})$.) As was explained already in the motivating example, $\mathbf{u}$ in that case takes the role of $\mathbf{A}$ in a fixed gauge in magneto-statics.

All solutions of $\mathcal{I}$ are solutions of both the Navier-Stokes equation and, via the correspondence Φ, of magneto-statics by proposition 18. Finally, note that this correspondence might not be the only one under which those two theories share structure.

10 Discussion

In this section, some conclusions are presented that are supposed to show that the aims, that were described in the introduction (Sect. 1), were reached, and an outlook to possible future research directions is given.

10.1 Conclusion

A geometric framework was developed to compare classical field theories, or more generally, any two systems of PDEs in the category of smooth manifolds, in a mathematically precise sense. For every two theories there might be multiple correspondences relating them, enabling a very versatile comparison, both of subtheories of a single theory with themselves and with subtheories of other theories.

The methods developed in this contribution allow to give an answer to all requirements (Q.1)–(Q.4) described in Sect. 1.2 in the following way.

1. A geometric answer to (Q.1) ("Are two systems of PDEs equivalent?") is given by Definition 13 in Sect. 6.
2. (Q.2) ("Do two PDEs share any subsystem?") can be answered by computing the shared structure described in Definition 11, using the methods from differential topology and formal integrability introduced in Sects. 4 and 5.
3. (Q.3) ("When are two systems equivalent up to a symmetry?") was answered by Definition 22 via the introduction of quotient equations (Definition 3) which can be computed using (146) in combination with differential consequences of $Q(I)$ as defined in (150).
4. Finally, (Q.4) ("How to transfer solutions from one system to another?") was answered by Propositions 18, 19 and Corollary 2, with the generalization of Bäcklund transformations in Proposition 22, Definition 17 and Propositions 23, 24 and the proposition about the transfer of solutions to quotient equations 27.

Hence, theoretical analogies of similar systems can now be analysed, new analogies can be found using symmetries, and methods to solve systems can be transferred with a generalization of Bäcklund transformations, that in particular can help to solve some otherwise barely tractable non-linear PDEs.

10.2 Outlook

It would be interesting to apply the framework to the comparison of more complex theories, for example to understand the relations between general relativity, hydrodynamics and electrodynamics.

Perhaps the description of analogue experiments can be made more transparent with the present approach.

Something that is still missing in the present framework is a way to find the best possible correspondence (e.g. the ones that maximizes the solution space of the intersection) between two given theories. A starting point for making progress in this direction might be the relationship between correspondences and symmetries as outlined in Sect. 8.

An interesting endeavor might be to study how Bäcklund transformations from eq. $\mathcal{E}$ to $\mathcal{F}$ and from $\mathcal{F}$ to $\mathcal{G}$ could give rise to Bäcklund transformations between $\mathcal{E}$ and $\mathcal{G}$ and if those could be used to build up chains of generalized relations between multiple equations that facilitate to map solutions of rather simple equations to ever more complex ones.

Another future aim would be to describe transitions between theories and approximations of theories in a mathematically precise way. They are important both for conceptual reasons—namely, to identify how one theory prepares the rise of another—and for practical purposes—namely, in order to be able to understand how one should approximate a complicated equation by a simpler one.

In the geometric framework, an equation is a submanifold of a jet bundle which locally is the kernel of some system of equations. Therefore, a slight approximation to this system would correspond to a slight deformation of the submanifold. Thus, deformation and homotopy theory might serve to describe such transitions.

A natural question is whether it would be possible to extend the framework to compare quantum theories. To a certain extend, it can be applied to quantum mechanics because the Schrödinger equation is also a PDE. However, in quantum field theory it would perhaps be necessary to consider functional equations because the Dyson-Schwinger equations, whose solution is the path integral, is a functional differential equation. At some points, [22] points out that cohomology theory could be used to study problems usually approached by functional analysis. The advantage would be that cohomology theory directly connects with all areas of geometry, topology, homological algebra, abstract algebra and would provide many tools to study quantum field theoretical problems in new ways. However, it is not yet clear how to set up such a theory.

Acknowledgments Since this work builds on the research of my Master's thesis, I want to express my gratitude towards my two supervisors, James Owen Weatherall (University of California, Irvine) and Ion Stamatescu (University of Heidelberg). Furthermore, I obtained valuable comments and support from Luca Vitagliano (University of Salerno) and Igor Khavkine (Czech Academy of Sciences). Finally, I'd like to thank my dear friend Thomas Mikhail for his continuous feedback.

References

1. Marmanis, H.: Analogy between the Navier–Stokes equations and Maxwell's equations: Application to turbulence. Phys. Fluids **10**, 1428–1437 (1998)
2. Goulart, E., Falciano, F.: Formal analogies between gravitation and electrodynamics (2008). http://arxiv.org/abs/0807.2777

3. Visser, M.: Acoustic black holes: horizons, ergospheres, and Hawking radiation (1997). https://arxiv.org/abs/gr-qc/9712010v2
4. Steinhauer, J., de Nova, J.R.M.: Self-amplifying Hawking radiation and its background: A numerical study. Phys. Rev. A **95**, 033604 (2017)
5. Weatherall, J.O.: Are Newtonian gravitation and geometrized Newtonian gravitation theoretically equivalent? (2014). https://arxiv.org/abs/1411.5757v3
6. Weatherall, J.O.: Understanding Gauge (2015). https://arxiv.org/abs/1505.02229v2
7. Rosenstock, S., Barrett, T.W., Weatherall, J.O.: On einstein algebras and relativistic spacetimes (2015). https://arxiv.org/abs/1506.00124 .
8. Barrett, T.W.: Equivalent and inequivalent formulations of classical mechanics (2017). http://philsci-archive.pitt.edu/13092/1/eaifocm.pdf
9. Abraham, R., Marsden, J.: Foundations of Mechanics. AMS Chelsea Pub./American Mathematical Society (2008)
10. Román-Roy, N.: Multisymplectic lagrangian and hamiltonian formalisms of classical field theories. Symmetry Integrability Geom. Methods Appl. **5**, 100 (2009)
11. Vinogradov, A.M., Krasilshchik, I.S.: Nonlocal symmetries and the theory of coverings: an addendum to A. M. Vinogradov's "local symmetries and conservation laws". Acta Appl. Math. **2**, 79–96 (1984)
12. Vinogradov, A.M., Krasil'shchik, I.S.: Nonlocal Trends in the Geometry of Differential Equations: Symmetries, Conservation Laws, and Bäcklund Transformations, pp. 161–209. Springer, Dordrecht (1989)
13. Krasil'shchik, I.S., Vinogradov, A.M., Bocharov, A.V., Chetverikov, V.N., Duzhin, S.V., Khor'kova, N.G., Samokhin, A.V., Torkhov, Y.N, Verbovetsky, A.M.: Symmetries and Conservation Laws for Differential Equations of Mathematical Physics. American Mathematical Society, Providence (1999) (Translations of Mathematical Monographs)
14. Kuhn, T.S.: The Structure of Scientific Revolutions. University of Chicago Press, Chicago (1996)
15. Stamatescu, I.-O., et al.: Symbol and Physical Knowledge - On the Conceptual Structure of Physics. Springer, Berlin (2013)
16. Hertz, H.: Gesammelte Werke von Heinrich Hertz - Band III - Die Prinzipien der Mechanik. Johann Ambrosius Barth (Arthur Meiner), Druck von Metzger und Wittich, Leipzig (1894)
17. Goldschmidt, H.: Integrability criteria for systems of nonlinear partial differential equations. J. Differ. Geom. **1**(3–4), 269–307 (1967)
18. Bryant, R.L., Gardner, R.B., Chern, S.S., Goldschmidt, H.L., Griffiths, P.A.: Exterior Differential Systems. Mathematical Sciences Research Institute Publications. Springer, New York (1991)
19. Seiler, W.M.: Involution: The Formal Theory of Differential Equations and its Applications in Computer Algebra. Algorithms and Computation in Mathematics. Springer, Berlin (2009)
20. Pommaret, J.F.: Partial Differential Equations and Group Theory: New Perspectives for Applications. Mathematics and Its Applications. Springer, Dordrecht (1994)
21. Vinogradov, A.M.: Local Symmetries and Conservation Laws. Springer, Berlin (1984)
22. Vinogradov, A.M.: Cohomological Analysis of Partial Differential Equations and Secondary Calculus. American Mathematical Society, Providence (2001)
23. Vitagliano, L.: Secondary calculus and the covariant phase space (2010). https://arxiv.org/abs/0809.4164v5
24. Vinogradov, A.M., Moreno, J.: Domains in infinite jet spaces: -spectral sequences. Dokl. Math. **75**(2), 204–207 (2007)
25. Moreno, G.: The geometry of the space of cauchy data of nonlinear PDEs (2012). https://arxiv.org/abs/1207.6290
26. Vitagliano, L.: Characteristics, bicharacteristics and geometric singularities of solutions of PDEs. Int. J. Geom. Methods Modern Phys. **11**(09), 1460039 (2014)
27. Pommaret, J.F.: Partial Differential Control Theory and Causality, pp. 599–605. Birkhäuser, Boston (1991)

28. Sorokina, M.: Poisson structures on manifolds with singularities (2013). https://arxiv.org/abs/1312.6262
29. Krasil'shchik, I.S., Verbovetsky, A.M.: Homological methods in equations of mathematical physics (1998). https://arxiv.org/pdf/math/9808130.pdf
30. Kogan, I.A., Olver, P.J.: Invariant euler–lagrange equations and the invariant variational bicomplex. Acta Appl. Math. **76**(2), 137–193 (2003)
31. Thompson, R., Valiquette, F.: On the cohomology of the invariant euler-lagrange complex. Acta Appl. Math. **116**(2), 199 (2011)
32. Valiquette, F.: Group foliation of differential equations using moving frames. Forum Math. Sigma **3** (2015)
33. Olver, P.J.: Equivalence, Invariants and Symmetry. Cambridge University Press, Cambridge (1995)
34. Kruglikov, B.: Symmetry approaches for reductions of PDEs, differential constraints and Lagrange-Charpit method. Acta Appl. Math. **101**(1), 145–161 (2008)
35. Kruglikov, B., Lychagin, V.: Global lie–tresse theorem. Sel. Math. **22**(3), 1357–1411 (2016)
36. Schneider, E.: Solutions of second-order PDEs with first-order quotients (2020). https://arxiv.org/abs/2005.06794
37. Kant, U., Seiler, W.M.: Singularities in the geometric theory of differential equations. Conf. Publ. **2011**, 784–793 (2011)
38. Tu, L.W.: An Introduction to Manifolds. Universitext. Springer, New York (2010)
39. Saunders, D.J.: The Geometry of Jet Bundles. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge (1989)
40. Uniffiety: Webpage created to collect research about the geometric approach to PDEs, Wiki. https://uniffiety.com
41. Guillemin, V., Pollack, A.: Differential Topology. AMS Chelsea Publishing Series. Prentice-Hall, Hoboken (2010)
42. Goldschmidt, H.: Existence theorems for analytic linear partial differential equations. Ann. Math. **86**(2), 246–270 (1967)
43. Barth, L.: A mathematical framework to compare classical field theories (2019). https://arxiv.org/abs/1910.08614
44. Rogers, C., Shadwick, W.F.: Bäcklund Transformations and Their Applications. Conference Series/Institute of Mathematics and Its Applica. Academic Press, Cambridge (1982)
45. Svinolupov, S.I, Sokolov, V.V.: Factorization of evolution equations. Russian Math. Surv. **47**(3), 127 (1992)
46. Reincke-Collon, C.: Entwurf invarianter Folgeregler für Systeme mit Lie-Symmetrien. Logos-Verlag, Berlin (2012)
47. Giachetta, G., Mangiarotti, L.: Gauge invariance and formal integrability of the yang-mills-higgs equations. Int. J. Theor. Phys. **35**(7), 1405–1422 (1996)