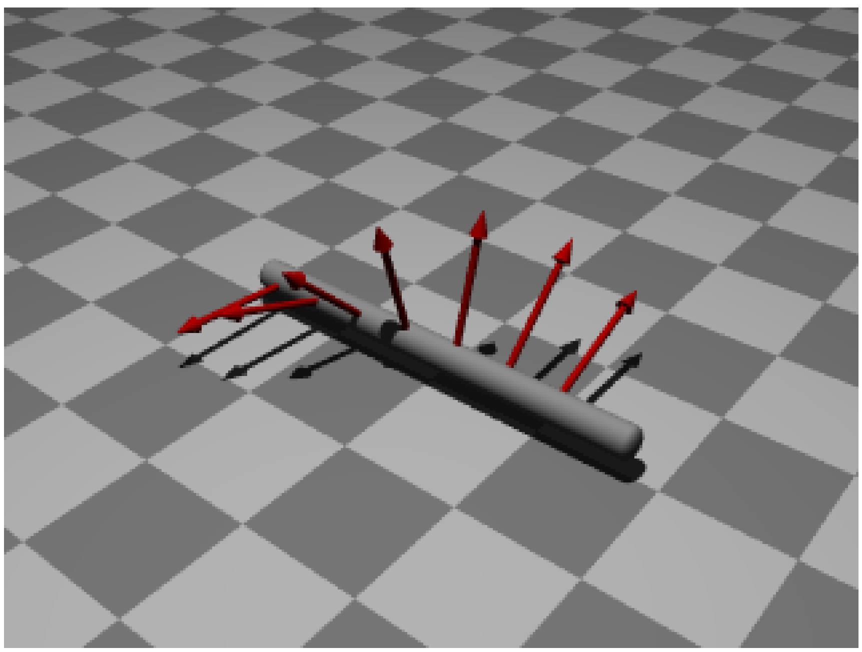
Continuous and discrete aspects of Lagrangian field theories with nonholonomic constraints

Joris Vankerschaver



Proefschrift ingediend aan de Faculteit Wetenschappen
tot het behalen van de graad van
Doctor in de Wetenschappen: Wiskunde

Promotor: Prof. Dr. Frans Cantrijn

Universiteit Gent
Faculteit Wetenschappen
Vakgroep Wiskundige Natuurkunde en Sterrenkunde
Academiejaar 2006-2007

Voorwoord

In dit voorwoord hou ik eraan om een aantal mensen te vermelden wiens inbreng beslissend was voor de uiteindelijke vorm van dit proefschrift. In de eerste plaats wil ik mijn promotor Frans Cantrijn bedanken, die vele jaren geleden mijn enthousiasme voor differentiaalmeetkunde en geometrische mechanica opgewekt heeft, en me sindsdien steeds met de nodige aanmoediging geduldig begeleid heeft. Ik hoop dat deze vruchtbare samenwerking ook in de komende jaren mag voortgezet worden. Ook Willy Sarlet, de grand old man van de vakgroep, verdient hier een woord van oprechte dank.

De collega's van de vakgroep wil ik hier niet vergeten en in het bijzonder wil ik de lichtende voorbeelden Bavo, Fien en Tom, alsook mijn huidige bureaugenoot Wouter, bedanken voor de gezellige sfeer en de stimulerende discussies. Een eervolle vermelding ook voor de niet-zo-lichtende voorbeelden Jos en David, en bij uitbreiding ook de Astridvrienden.

También quisiera agradecer a mis amigos en España: en primer lugar a Manuel de León y David Martín de Diego por su consejo y por su hospitalidad durante mis visitas a Madrid, así como a Eduardo Martínez de Zaragoza.

It is a pleasure to acknowledge many conversations on mathematics, as well as on life, the universe, and everything with David Saunders and Mike Crampin during their visits to Ghent and the workshops abroad. Furthermore, let me also thank Dmitry Zenkov for his interest and encouragement while I was finishing this dissertation. Last and certainly not least, this gratitude is extended to Thomas Bridges and Mark Roberts for their hospitality during my visit to the University of Surrey.

Bedankt ook aan mijn vrienden, voor hun pogingen om mij op tijd en stond van achter mijn computer weg te sleuren en kennis te laten maken met de wereld daarbuiten.

Tot slot wil ik ook mijn ouders Veva en Wilfried en mijn zus Clara bedanken, die me allereerst aangemoedigd hebben om wiskunde te studeren en me daarna ook steeds gesteund hebben tijdens mijn doctoraatsjaren. Ook de diepgaande gesprekken met Bompa en Regina op zondagavond, en de bijhorende Leffe's, hebben in niet te onderschatten mate bijgedragen aan mijn vorming als wetenschapper.

Dit proefschrift kwam tot stand middels een aspirantenmandaat van het Fonds voor Wetenschappelijk Onderzoek – Vlaanderen.

Joris Vankerschaver
januari 2007

Contents

Voorwoord i

Introduction v
  Outline of the dissertation ix

Notations xiii

Chapter 1. Lagrangian field theories 1
  1. Jet bundles 1
  2. Lagrangian field theories 9
  3. The Cauchy formalism 14
  4. Elasticity as a multisymplectic field theory 21

Chapter 2. Variational integrators 27
  1. Geometric integration of mechanical systems 27
  2. Discrete multisymplectic field theory 32

Chapter 3. Discrete Lagrangian field theories on Lie groupoids 41
  1. The discrete jet bundle 41
  2. Discrete fields 45
  3. Lagrangian field theories 51
  4. Examples 60

Chapter 4. Symmetry for discrete Lagrangian field theories 65
  1. Lagrangian reduction 66
  2. The Noether theorem 69

Chapter 5. Euler-Poincaré reduction for discrete field theories 71
  1. Discrete differential geometry 71
  2. Discrete Euler-Poincaré reduction 76
  3. The Noether theorem 80
  4. Extending the Moser-Veselov approach 82
  5. An application: harmonic mappings 84

Chapter 6. Geometric aspects of nonholonomic field theories 87
  1. Nonholonomic Lagrangian field theory 88

2. The nonholonomic projector 93
3. An example from incompressible hydrodynamics 96
4. Cauchy formalism for nonholonomic field theory 99

Chapter 7. The nonholonomic momentum equation 103
1. Further properties of connections on π_1 103
2. Symmetry in the absence of nonholonomic constraints 104
3. The constrained momentum map 105
4. The Cauchy formalism 107

Chapter 8. Holonomic and affine nonholonomic constraints 111
1. Holonomic constraints 112
2. The Skinner-Rusk approach for constrained field theories 114
3. Skinner-Rusk formulation of vakonomic field theories 115
4. Skinner-Rusk formulation of nonholonomic field theories 120
5. Comparison between both approaches 122

Chapter 9. Nonholonomic kinematic constraints in elasticity 125
1. Introduction 125
2. Nonholonomic kinematic constraints 129
3. A Cosserat-type model 136
4. Discrete nonholonomic field theories 143

Chapter 10. Conclusions and future work 151
1. Discrete jets in higher dimensions 151
2. Nonholonomic field theories 152
3. Lie groupoid field theories with nonholonomic constraints 153

Appendix A. Elementary properties of the Frölicher-Nijenhuis bracket 155

Appendix B. Lie groupoids and Lie algebroids 157
1. Lie groupoids 157
2. Lie algebroids 158
3. Prolongations of Lie groupoids and algebroids 160

Nederlandse samenvatting 163

Bibliography 177

Introduction

In a nutshell, this thesis deals with the inclusion of *nonholonomic constraints* into *classical field theories* and their *discretization*. Even though these themes might appear to be quite unrelated at first, they have at least one important property in common: all three are naturally described using the language of differential geometry, which moreover, in each case, leads to new insights, and substantial advances in the development of the theory.

Before outlining the specific accomplishments of this thesis, we will therefore start with a brief summary of the use and advantages of differential geometry in these areas.

Classical field theory

Classical fields are, in most cases, easily modeled as sections of a fibre bundle. A few examples are given below:

- In electromagnetism, the four-potential A_μ can be interpreted as a one-form $A_\mu \mathrm{d}x^\mu$ on space-time.
- In general relativity on a manifold X, the relevant field is the metric g, a section of the bundle of symmetric nondegenerate $(0, 2)$-tensors with appropriate signature.
- A scalar field on a manifold X is just a section of the trivial bundle $\pi : X \times \mathbb{R} \to X$.

More examples may be found throughout the literature, e.g. in [**11**, **22**, **48**, **73**].

Classical field theory can be described very elegantly in terms of the geometry of the underlying fibre bundle and its associated first jet bundle $J^1\pi$. The latter is a natural bundle associated to a given fibre bundle $\pi : Y \to X$. Whereas the tangent bundle of a manifold has fibre coordinates that represent "velocities", a jet bundle is equipped with fibre coordinates that can be interpreted as "generalized velocities", being the derivatives of the coordinates of Y with respect to those of X.

The Poincaré-Cartan form. According to a well-known paradigm in physics, the dynamics of a classical field is fully specified by giving a *Lagrangian*, which is (roughly speaking) just a function on the jet bundle. The specification of a Lagrangian induces a number of interesting geometric objects on the jet bundle, the most important of which is the *Poincaré-Cartan* $(n + 2)$*-form* Ω_L (where $n + 1 = \dim X$). This form, which is the exterior derivative of an $(n+1)$-form Θ_L, called the *Cartan form*, is the generalization to classical field theory of the symplectic form used in the description of classical mechanics.

Using the Poincaré-Cartan form, one can reformulate the Euler-Lagrange equations (as well as various extensions) in an intrinsic way. The Poincaré-Cartan form was discovered in its current inception independently by García [**46**], Goldschmidt and Sternberg [**47**], and Kijowski [**58**], but goes back to the work of Poincaré, Cartan, Weyl, Caratheodory, and many others, at the beginning of the twentieth century.

The Poincaré-Cartan form Ω_L can be derived in a number of different, but equivalent ways. For our purposes, it is interesting to know that Ω_L is intimately related with the variational derivation of the field equations. Let S be the *action functional*, defined as

$$S(\phi) = \int_X L\left(x^\mu, \phi^a(x), \frac{\partial \phi^a}{\partial x^\mu}\right) \mathrm{d}^{n+1}x.$$

The Euler-Lagrange equations are derived by looking for extremals of this action. In [**80**] and subsequent works, Marsden *et al.* point out that by broadening the class of variations, additional terms arise, including the Cartan form. Indeed, let δy^a be an arbitrary "vertical" variation of a field ϕ. By varying S with respect to δy^a, we obtain the following expression:

$$\delta S = \int_U \left(\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\frac{\partial L}{\partial y^a_\mu}\right) \delta y^a \; \mathrm{d}V + \int_{\partial U} \frac{\partial L}{\partial y^a_\mu} \delta y^a \; n_\mu \mathrm{d}A, \tag{1}$$

where U is an open subset of X and n_μ is the normal to the submanifold ∂U in X. Here, the first term on the right-hand side yields the Euler-Lagrange equations, and the second one is just the contraction of the variation written in coordinates as δy^a with the Cartan form Θ_L. The formula (1) was used in [**94**] to characterize the Cartan form, and a similar, but more involved formula holds when vertical variations are replaced by arbitrary variations (eq. 74 in [**70**]). We therefore see that the Poincaré-Cartan form arises naturally in the variational context. Moreover, (1) can be used to derive the Poincaré-Cartan form for situations where it was previously unknown, such as higher-order field theories (see [**61**]) or discrete mechanics and discrete field theory (see [**80**]).

Nonholonomic constraints

Let us now leave classical field theories for a while, and focus on classical mechanics. Consider a convex rigid body, for example a coin, or a homogeneous sphere, rolling without sliding on a fixed horizontal plane. Anyone can readily picture such a motion, and yet even this simple example already exhibits a very striking type of behaviour. For example, cyclic coordinates occur in the Lagrangian of the rolling disc, but the associated momenta are not conserved. More sophisticated examples include the so-called rattleback, whose non-intuitive spinning behaviour has continued to arouse interest for over a century. We refer to [**14**] for an overview of the literature.

For our purposes, it is first and foremost the geometric formulation of systems with nonholonomic constraints which is of interest. In mechanics, most classes of nonholonomic constraints are linear in the velocities and hence can be represented by a distribution D on the configuration space. In order to maintain the constraint, the system is subjected to additional *reaction forces* which are specified by the principle of d'Alembert. In modern terminology, this principle asserts that the reaction forces are co-vectors on TQ which annihilate the vertically lifted distribution D^v (see [**37**]). The equations of motion for a nonholonomic system then become

$$i_\Gamma \omega_L - \mathrm{d}E_L \in (D^v)^\circ \quad \text{and} \quad \operatorname{Im}\Gamma \subset D^c, \tag{2}$$

along D, where D^c is the complete lift of the distribution D. Under suitable regularity conditions, there exists a complement to TD in $T(TQ)$ (defined along D). In [**37**], the authors prove that in that case, any solution Γ of the free problem (*i.e.* where no constraints are present) yields a solution Γ' of (2) by composing Γ with the projection of $T(TQ)$ onto TD.

The nonholonomic equations of motion can also be derived by varying the action functional with respect to admissible variations, where the admissibility of variations is dictated by the principle of d'Alembert. This formulation is especially useful for discrete nonholonomic systems.

Classical field theories with nonholonomic constraints. Mechanical systems with nonholonomic constraints are ubiquitous in nature, and are best described using concepts of differential geometry. Given the fact that classical field theories also have a very geometric description, it is natural to ask oneself whether it makes sense to study nonholonomic constraints for classical field theories as well.

This question can be approached from two different ways. One is the purely geometrical approach, where one tries to use the intrinsic geometry of the jet bundle to find analogues of the constructions outlined in the previous section for mechanics. This is done in chapter 6. Even though this procedure is mathematically rather attractive, it leaves open the question of physical relevance of such theories. Chapter 9 is therefore devoted to a first example of a physically sound nonholonomic field theory.

Discrete mechanics and discrete field theory

Most mechanical systems have equations of motion that cannot be integrated analytically. It is therefore natural to turn to computer methods to help us integrate the Euler-Lagrange equations. In the past, most approaches to numerical analysis consisted of finding general purpose algorithms which could be applied to as large a class of differential equations as possible. However, the Euler-Lagrange or Hamilton equations are characterized by a number of geometric properties, such as conservation of phase space volume (or more generally, *symplecticity*). Over the past decades, it has become clear that numerical preservation (*i.e.* up to rounding errors) of these properties yields qualitatively better algorithms: this is called *geometric integration*.

Discrete mechanics. In this thesis, we do not pretend to make a significant contribution to the construction of geometric integration algorithms. Rather, we will focus on the related area of discrete mechanics, and discrete field theory. While discrete mechanics can be used as a starting point for the development of geometric integrators (as in [**79**]), it is also an interesting area in its own right, with many diverse applications such as integrability (see [**85**]), and links with other areas of geometry such as Lie groupoid theory (see [**75**, **110**]).

In particular the approach of Moser and Veselov [**85**] deserves attention here. They discretized the tangent bundle TQ by replacing it by the product $Q \times Q$, the underlying idea being that a tangent vector on Q can be approximated in some sense by a pair of points (q_0, q_1). Secondly, they discretize the variational principle by considering a discrete Lagrangian and the associated action sum (rather than an action integral as in continuous mechanics) and deriving the discrete equations of motion by varying that discrete action sum. The advantage of this approach is that the resulting discrete system shares many properties with the continuous system. In particular, given a discrete Lagrangian L, there exists a symplectic form Ω_L on $Q \times Q$ and one can show that the discrete flow preserves that form. This is essentially a consequence of the discrete variational principle.

Apart from working with discrete mechanical systems on $Q \times Q$, Moser and Veselov also studied an example where the velocity space is a Lie group $\mathcal{G}$. These two somewhat disconnected examples can be subsumed into one general framework using the theory of *Lie groupoids*; this was done by Weinstein [**110**]. Briefly speaking, a groupoid is a set G resembling a group, but equipped with a "partial" multiplication, in the sense that only specific pairs of elements can be multiplied. Any group is a groupoid; in that case, the multiplication is defined for all pairs of elements. Another example is the pair manifold $Q \times Q$: two elements (x, y) and (u, z) can be multiplied only if $y = u$, and in that case we have

$$(x, y) \cdot (y, z) = (x, z).$$

A Lie groupoid is a groupoid equipped with a smooth structure. Associated to any Lie groupoid G is a certain vector bundle called the *Lie algebroid* AG *of* G: this is similar to the fact that any Lie group gives rise to a Lie algebra. For example: the Lie algebroid of $Q \times Q$ is the tangent bundle TQ. As can be expected from these few examples, the geometry of Lie groupoids is very rich, and in particular, can be used in discrete mechanics in much the same way as was done by Moser and Veselov.

Discrete field theory. Inspired by these powerful and elegant methods, a number of people set out to develop a similar geometric approach to classical field theories. Bridges [**17**] introduced a concept of "multisymplecticity" for Hamiltonian partial differential equations and later Bridges and Reich [**19**] studied numerical integrators that conserve a discretized version of this invariant. Independently, Marsden, Patrick, and Shkoller [**80**] extended the work of Veselov in order to deal with Lagrangian field theories.

Over the years, many people have constructed geometric integrators for field theories which preserve multisymplecticity. An overview of these methods can be found in [**18**]. A theoretical study of discrete field theories and related aspects (such as symmetry reduction, for example) is provided in this thesis.

Outline of this dissertation

The first two chapters of this thesis provide an introduction to Lagrangian field theory and variational integrators, respectively. Most of the material in these chapters is standard and therefore no proofs have been given, except in the case where only partial results were needed, or simpler proofs could be given. After the introductory chapters, the main body of this text contains two different themes, that of *discrete field theories*, and of *nonholonomic field theories*, respectively.

Discrete field theories. In the introductory chapter 2, the current state of affairs in the area of discrete mechanics and field theory is summarized; this treatment is based mainly on the foundational work of Marsden *et al.* [**80**] and Bridges and Reich [**19**]. The work of these authors was developed with a view towards practical applications (*i.e.* the construction of robust geometric integration schemes), but also serves as a source of inspiration for progress on the theoretical front.

As a first step in this direction, we introduce in chapter 3 the concept of discrete fields taking values in a given Lie groupoid G. The use of Lie groupoid techniques for the study of discrete mechanical systems was pioneered by Marrero *et al.* [**75**], and our treatment of discrete field theory is a generalisation of their work. Besides providing a clearer insight into the geometry of discrete mechanics and field theory, Lie groupoids also allow us to treat specific problems that cannot be addressed in the standard framework. An important example is found in discrete reduction theory, where discrete fields taking values in a Lie group naturally arise.

A central element in our treatment is the set $\mathbb{G}^k$ of "k-gons" in G, whose elements are k-tuples of composable elements of G, such that the cyclic multiplication of these elements yields a unit element. Schematically:

$$(g_1, g_2, \ldots, g_k) \in \mathbb{G}^k \quad \text{if} \quad (g_i, g_{i+1}) \in G_2 \quad (\text{for } i = 1, \ldots, k) \quad \text{and} \quad g_1 \cdot g_2 \cdots g_k = e_{\alpha(g_1)},$$

where G_2 is the set of composable elements in G. The set $\mathbb{G}^k$ closely resembles the Lie groupoid G (there is a well-defined way of inverting elements of $\mathbb{G}^k$ and $\mathbb{G}^k$ is equipped with k anchor mappings generalizing the usual source and target map), except for the fact that there is no obvious multiplication of elements of $\mathbb{G}^k$. Nevertheless, the remaining properties of $\mathbb{G}^k$ still allow us to construct a geometric theory of Lie groupoid field theories.

In particular, the Lie algebroid AG of G induces a prolongation bundle $P^k\mathbb{G}$ over $\mathbb{G}^k$, which can be endowed in turn with the structure of a Lie algebroid. In chapter 3, we prove that the Poincaré-Cartan forms are sections of the dual of this bundle, and we introduce a certain kind of Legendre transformations from this bundle to a prolongation of AG, on which a natural symplectic section exists. Pullback of this section along the Legendre maps then yields the original Poincaré-Cartan sections. Furthermore, we derive the discrete field equations and show that the solutions to these equations are "multisymplectic" in the sense of [**80**] and [**19**].

Finally, we construct a reduction procedure for discrete field theories with symmetry. In chapter 4, we first consider the general case of "symmetry with respect to a morphism". We derive a reduction theorem which allows to "factor out" effects of symmetry, and to reformulate the symmetric discrete field theory on a new, reduced groupoid. This reduced field theory shares many features with the original one: it is multisymplectic precisely when the original field theory is multisymplectic, and the Poincaré-Cartan forms of the reduced theory are in a straightforward correspondence with those of the original field theory.

In the second part of that chapter (section 2), as well as in chapter 5, we treat the important special case where a symmetry group acts on the target space of the field theory. As one might expect from the "Noether paradigm" (which says that for each continuous symmetry, there exists a conservation law, and vice versa) the existence of such a symmetry action implies a (discrete) conservation law. In the case where the discrete field theory takes values in a Lie group $\mathcal{G}$, which is at the same time the symmetry group of the theory, we show that these conservation laws are equivalent to the equations of motion, the *discrete Euler-Poincaré equations*. We analyse these equations using some concepts from discrete differential geometry: we show that discrete reduced fields can be considered as *discrete $\mathcal{G}$-connections*, and that the obstruction to reconstruction is precisely the discrete curvature of such a field. We end chapter 5 by showing how the discrete Euler-Poincaré equations may be used to construct a discrete counterpart to the theory of harmonic mappings from $\mathbb{R}^2$ into a semi-simple Lie group.

Nonholonomic field theories. A natural question is whether the geometric methods that proved to be so successful in nonholonomic mechanics, can be extended to the context of classical field theories. A partial answer is provided in chapter 6, where we construct a distribution D along the constraint submanifold $\mathcal{C}$ such that (under some regularity conditions) $T(J^1\pi)$ can be written as the direct sum $T\mathcal{C} \oplus D$. Composition with the projector $\mathcal{P} : T(J^1\pi) \to T\mathcal{C}$ then provides us with a natural way of turning solutions of the free problem into constrained solutions.

Chapter 7 is devoted to the inclusion of symmetry in this framework. One of the cornerstones of mechanics is the Noether theorem. However, this theorem does not always hold if nonholonomic constraints are present, and is replaced by a particular equation which describes the evolution of the associated "conserved currents". This equation was derived in [**13**] for mechanical systems and is extended here to the case of field theories.

In both of these chapters, nonholonomic field theories are also studied from a different point of view, using the so-called *Cauchy formalism* for classical field theories. In particular, we show that nonholonomic field theories give rise to nonholonomic mechanical systems (in the classical sense) on a certain infinite-dimensional manifold, and that the momentum equation derived in chapter 7 induces the corresponding equation for mechanical systems.

The remainder of the thesis is then devoted to the study of certain classes of constraints. In chapter 8 we point out that for field theories with a canonical direction of time, the distinction between holonomic and nonholonomic constraints is not as straightforward as the analogy with mechanics would make us believe. The well-known constraint of inextensibility in fluid dynamics, for example, is traditionally never treated as a nonholonomic constraint, and yet, when we apply the formalism of chapter 6, we still obtain the correct field equations. However, for other nonintegrable constraints, the nonholonomic equations seem to give wrong results.

This enigma is resolved by taking a closer look at the kind of field theories where this behaviour occurs. It turns out that these field theories are characterised by having a base space which can be written as the product of time and space. In that case, one can reformulate the dynamics as an ODE on the (infinite-dimensional) space of fields (a particular application of Cauchy analysis), and it turns out that constraints that only involve spatial derivatives (such as the incompressibility constraint) become integrable in this framework. By contrast, constraints involving time derivatives are genuinely nonholonomic.

In chapter 9, we construct a physical example of such a nonholonomic system. The basic model is that of a Cosserat rod, a special kind of continuum theory. This rod moves in a horizontal plane which is supposed to be sufficiently rough, so that it rolls without sliding. In this way, we obtain what could reasonably be called a continuum version of the vertically rolling disc. This model can be analysed using the techniques developed in chapter 6. Furthermore, we prove the conservation of energy, and we investigate the nonconservation of momentum using the momentum lemma of chapter 7.

Eventually, to obtain a quantitative insight into the dynamics of the nonholonomic Cosserat rod, we have to resort to computer models. At the end of chapter 9, we construct a second-order geometric integration scheme which exactly preserves the nonholonomic constraint. This scheme is based on a suitable discrete version of d'Alembert's principle (as in [**28**]), and is again a testimony to the effectiveness of differential geometry in classical field theory.

References. Some parts of this thesis have already appeared in print, or are submitted for publication. The dynamics of discrete field theories on Lie groupoids was treated in [**108**], while aspects of the theory of symmetry in this context were established in [**105**]. The theoretical framework for nonholonomic field theories was studied in [**102**], while the momentum lemma was derived in [**103**]. The results on linear nonholonomic and holonomic constraints were published in [**104**, **107**]. Finally, the example of the nonholonomic Cosserat rod, as well as its numeric treatment, forms the subject of [**106**].

Notations

In this thesis, we will work in the category of smooth maps and smooth manifolds. All manifolds are finite dimensional, except where indicated otherwise.

The tangent functor is denoted by T: the tangent bundle of a smooth manifold Q is denoted by TQ, and the tangent map of a smooth map f by Tf. The tangent bundle projection will be denoted by $\tau : TQ \to Q$, and the space of sections of τ, or vector fields on Q, by $\mathfrak{X}(Q)$. The cotangent bundle of Q is denoted by T^*Q, and the cotangent bundle projection by $\pi : T^*Q \to Q$. The k-fold exterior product of T^*Q with itself will be denoted by $\bigwedge^k(T^*Q)$. The space of sections of $\bigwedge^k(T^*Q)$ is the module of k-forms, denoted by $\Omega^k(Q)$. The Lie derivative with respect to a vector field X will be denoted by $\mathscr{L}_X$. The contraction of a vector field X with a differential form α will be denoted both by $i_X\alpha$ as well by $X \lrcorner\, \alpha$, and the same convention applies to the contraction of a vector-valued form with a differential form (see appendix A).

A fibre bundle is a triple (Y, π, X), consisting of a base space X, a total space Y and a submersion $\pi : Y \to X$ which is locally trivial. Usually, we will denote a fibre bundle simply by its projection $\pi : Y \to X$. If $\pi : Y \to X$ is a fibre bundle, the vertical bundle is the subbundle of TY denoted by $V\pi$ consisting of those vectors which project onto zero under π. The space of sections of π will be denoted by $\Gamma(\pi)$, or, if no confusion is possible, by $\Gamma(Y)$.

Bundle maps between two fibre bundles π and π' are denoted by pairs $(\underline{\Phi}, \overline{\Phi})$, or simply by Φ, where $\underline{\Phi} : X \to X'$ is the base space map and $\overline{\Phi} : Y \to Y'$ the total space map. When no confusion can arise, we will omit the base space map and refer to $\overline{\Phi}$ as the bundle map.

A special role in this thesis is played by vector fields and forms along a map. Let $f : M \to N$ be a smooth map, then a vector field along f is a map $X : M \to TN$ such that $X(m) \in T_{f(m)}N$. In other words, X is a section of the pullback bundle f^*TN. Forms along f can be defined similarly as maps $\alpha : M \to \bigwedge^k(T^*N)$ such that $\alpha(m) \in \bigwedge^k(T^*_{f(m)}N)$; they are sections of $f^* \bigwedge^k(T^*N)$.

Chapter 1
Lagrangian field theories

As mentioned in the introduction, classical field theory can be quite naturally studied by using the geometry of fibre bundles. In this introductory chapter we intend to give an overview of jet bundle theory in particular, and of the geometric formulation of classical field theory.

The plan of the chapter is as follows: section 1 is devoted to the study of jet bundles and the geometric objects associated with them, including an overview of connection theory in section 1.3. In section 2 we then make the link with Lagrangian field theories on jet bundles. In section 3, a radically different picture of classical field theory is sketched: we break covariance and reformulate the dynamics of the field on an infinite-dimensional space. This is the so-called *Cauchy formalism*. Finally, with section 4 we close this chapter by providing a detailed treatment of nonrelativistic elasticity, a particular example of a Lagrangian field theory.

As most of the material in this chapter is fairly standard, all proofs have been omitted, with the exception of proofs of new results.

1. Jet bundles

The overview of jet bundle theory in this section is based on the book [**94**] by D. J. Saunders.

For the sake of definiteness, from now on $\pi : Y \to X$ will be a fibre bundle of rank m (*i.e.* with m-dimensional fibres), whose base space X is assumed to be an oriented manifold of dimension $n + 1$, equipped with a fixed volume form η.

Throughout, we assume that (x^μ), $\mu = 1, \ldots, n + 1$, is a coordinate system on X compatible with the volume form η, *i.e.* such that

$$\eta \equiv \mathrm{d}^{n+1}x := \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{n+1},$$

and that a system of adapted bundle coordinates (x^μ, y^a) is given on Y, with $a = 1, \ldots, m$.

A *section* of the fibre bundle π is a map $\phi : X \to Y$ such that $\pi \circ \phi = \mathrm{id}_X$, where id_X is the identity map on X. In bundle coordinates, a section ϕ can locally be written as $\phi(x) = (x^i, \phi^a(x))$, for some local functions ϕ^a.

1.1. First-order jets. We define an equivalence relation on the set of sections of π by saying that two sections ϕ and ψ, defined on an open neighbourhood U of a point $x \in X$, are 1-*equivalent* if $\phi(x) = \psi(x)$ and $T_x\phi = T_x\psi$. In other words, ϕ and ψ are 1-equivalent if their Taylor expansions at x agree up to first order. The equivalence class of a local section ϕ at x is denoted by $j^1_x\phi$.

Definition 1.1. *The first jet manifold $J^1\pi$ of π is the set of all such equivalence classes:*

$$J^1\pi := \{j^1_x\phi \quad \text{where}\, x \in M \quad \text{and} \quad \phi \in \Gamma_x(\pi)\},$$

where $\Gamma_x(\pi)$ is the set of sections of π defined in an open neighbourhood of $x \in M$.

It can be shown that $J^1\pi$ can be given the structure of a smooth manifold, but more can be said: the first jet manifold $J^1\pi$ is equipped with two submersions $\pi_1 : J^1\pi \to X$ and $\pi_{1,0} : J^1\pi \to Y$, defined as

$$\pi_1(j^1_x\phi) = x \quad \text{and} \quad \pi_{1,0}(j^1_x\phi) = \phi(x),$$

and $\pi_{1,0} : J^1\pi \to Y$ is an affine bundle modelled over the vector bundle $\pi^*T^*X \otimes V\pi \to Y$, where $V\pi$ is the bundle of π-vertical vectors over Y.

The system of bundle coordinates on the fibre bundle π induces a coordinate system $(x^\mu, y^a; y^a_\mu)$ on $J^1\pi$, where y^a_μ is defined by

$$y^a_\mu(j^1_x\phi) = \frac{\partial \phi^a}{\partial x^\mu}(x).$$

This coordinate expression also shows that $J^1\pi$ has dimension $n + 1 + m + (n + 1)m$.

This system of natural bundle coordinates is also useful in uncovering the affine structure of the bundle $\pi_{1,0} : J^1\pi \to Y$: let $j^1_x\phi$ be an element of $J^1\pi$ and consider an element u of $V\pi \otimes \pi^*T^*X$ in the fibre over $\phi(x)$. In coordinates, u can be written as

$$u = u^a_\mu \mathrm{d}x^\mu \otimes \frac{\partial}{\partial y^a}.$$

The jet $j^1_x\phi + u$ is then the element of $\pi^{-1}_{1,0}(\phi(x))$ with fibre coordinates $y^a_\mu(j^1_x\phi) + u^a_\mu$.

The following alternative interpretation of jets will also be useful: a jet $j^1_x\phi$ can be interpreted as a injective linear map from T_xX to T_yY, where $y = \phi(x)$. Indeed, there is an obvious correspondence $j^1_x\phi \leftrightarrow T_x\phi$, which is easily seen to be well defined.

Remark 1.2. If $\pi : M \times S \to M$ is a trivial bundle, then we denote $J^1\pi$ as $J^1(M, S)$. Note that $J^1(M, S)$ is fibered over M as well as over S: the projection onto M is just π_1, and the projection onto S is $\pi_{1,0}$ composed with the projection $\mathrm{pr}_2 : M \times S \to S$ onto the second factor. ◇

1.1.1. *The vertical endomorphism S on $J^1\pi$.* The tangent bundle TQ of a manifold Q is equipped with a canonical $(1,1)$-tensor field J, called the *vertical endomorphism* (see [**30**]). Its construction relies essentially on the structure of TQ as a vector bundle.[1] In coordinates, J is given by

$$J = \mathrm{d}y^a \otimes \frac{\partial}{\partial \dot{y}^a}, \tag{1.1}$$

in a natural bundle coordinate system $(y^a, \dot{y}^a)$ on TQ. For an intrinsic construction, see the references mentioned above.

There is a similar object in jet theory. Its construction is somewhat more complicated, and depends (among other things) on the existence of a volume form on X. We refer to Saunders [**94**] for a detailed treatment and note only that its construction is intimately tied up with the affine structure of the first jet bundle $\pi_{1,0} : J^1\pi \to Y$. In terms of the fixed volume form η on X, the *vertical endomorphism* is then a vector-valued $(n+1)$-form that has the following expression in the coordinate system described above:

$$S_\eta = (\mathrm{d}y^a - y^a_\nu \mathrm{d}x^\nu) \wedge \mathrm{d}^n x_\mu \otimes \frac{\partial}{\partial y^a_\mu}, \tag{1.2}$$

where

$$\mathrm{d}^n x_\mu := \frac{\partial}{\partial x^\mu} \lrcorner\, \mathrm{d}^{n+1}x.$$

Note in passing that we make no notational distinction between the volume form η on X and its pull-back to Y or to $J^1\pi$ under the respective projections π and π_1.

Remark 1.3. From now on, we will use J to denote the vertical endomorphism on a tangent bundle, and S to denote the vertical endomorphism on a jet bundle. ◇

1.1.2. *Prolongation to $J^1\pi$ of vector fields and bundle maps.* Generally speaking, the prolongation operation j^1 takes certain objects (in this case vector fields and bundle maps) defined on a bundle, and turns them into the corresponding objects on the first jet bundle.

Consider a bundle map $\Phi = (\underline{\Phi}, \overline{\Phi})$ from π to itself and assume that the base space map $\underline{\Phi} : X \to X$ is a diffeomorphism. Recall that Φ is a bundle map if $\pi \circ \overline{\Phi} = \underline{\Phi} \circ \pi$. The *prolongation* of Φ to $J^1\pi$ is the bundle map $j^1\Phi : J^1\pi \to J^1\pi$ defined as

$$j^1\Phi(j^1_x\phi) := j^1_x(\overline{\Phi} \circ \phi \circ \underline{\Phi}^{-1}).$$

Note that the prolongation cannot be defined if $\underline{\Phi}$ is not a diffeomorphism. One can equally well define prolongations of bundle morphisms from one bundle π to another one π', but these will not be needed here.

[1] Indeed, the vertical endomorphism can be introduced on arbitrary vector bundles without conceptual complications (see for instance [**83**]).

If W is a vertical vector field on Y, then its flow $\{\Phi_t\}$ consists of bundle maps over the identity in X. Consequently, the prolongation of W is defined as the vector field $j^1W \in \mathfrak{X}(J^1\pi)$ whose flow is given by the local 1-parameter group $\{j^1\Phi_t\}$. If W is given in coordinates by $W = W^a(x,y)\frac{\partial}{\partial y^a}$, then j^1W is given by

$$j^1W = W^a\frac{\partial}{\partial y^a} + \frac{\mathrm{d}W^a}{\mathrm{d}x^\mu}\frac{\partial}{\partial y^a_\mu}, \quad \text{where} \quad \frac{\mathrm{d}W^a}{\mathrm{d}x^\mu} = \frac{\partial W^a}{\partial x^\mu} + \frac{\partial W^a}{\partial y^b}y^b_\mu.$$

The intrinsic construction of the prolongation of a non-vertical vector field is somewhat harder; we refer to [**94**] for more details. In coordinates, if

$$W = W^\mu(x,y)\frac{\partial}{\partial x^\mu} + W^a(x,y)\frac{\partial}{\partial y^a}\,,$$

then

$$j^1W = W^\mu\frac{\partial}{\partial x^\mu} + W^a\frac{\partial}{\partial y^a} + \left(\frac{\mathrm{d}W^a}{\mathrm{d}x^\mu} - y^a_\nu\frac{\mathrm{d}W^\nu}{\mathrm{d}x^\mu}\right)\frac{\partial}{\partial y^a_\mu}. \tag{1.3}$$

1.1.3. *Semi-basic and contact forms on $J^1\pi$.* On $J^1\pi$, there are a number of classes of distinguished differential forms. These particular types of forms play an important role in the geometric analysis of the calculus of variations (see [**65**]). For the developments in this thesis, however, we only need some introductory definitions.

Definition 1.4. *Consider an arbitrary fibre bundle $\pi : Y \to X$. Let α be a k-form on Y. Then α is* horizontal *or* semi-basic *with respect to the projection π if $i_V\alpha = 0$ for all π-vertical vector fields V on Y.*

We denote the module of semi-basic k-forms with respect to π by $\Omega^k_0(\pi)$. These forms are sections of the pullback bundle $\pi^* \bigwedge^k(X)$, which we denote for the sake of brevity by $\bigwedge^k_0(\pi)$.

Note that the concept of semi-basic k-forms is not limited to the case of jet bundles. The following definition, however, is inextricably tied up with the special structure of the first jet bundle.

Definition 1.5. *Let ω be a k-form on $J^1\pi$. Then ω is a* contact k-form *if $(j^1\phi)^*\omega = 0$ for all sections of π. If ω is a contact k-form, we say that ω is a* 1-contact k-form *if $i_V\omega$ is π_1-horizontal for every π_1-vertical vector field V. We say that ω is an* m-contact k-form *(where $m > 1$ and $m \leq k$) if $i_V\omega$ is an $(m-1)$-contact form for every π_1-vertical vector field V.*

If we denote by θ^a the contact one-form $\mathrm{d}y^a - y^a_\mu\mathrm{d}x^\mu$, the differential ideal of contact forms is algebraically generated by θ^a and $\mathrm{d}\theta^a$, $a = 1,\ldots,m$.

We denote the module of m-contact k-forms on $J^1\pi$ by $\Omega^k_{c,m}(J^1\pi)$ and define the module $\Omega^k_{c,\geq m}(J^1\pi)$ of *at least m-contact* k-forms as

$$\Omega^k_{c,\geq m}(J^1\pi) = \Omega^k_{c,m}(J^1\pi) \oplus \cdots \oplus \Omega^k_{c,k}(J^1\pi).$$

The fundamental result in this area is given below. Its proof can be found in [**64**].

Proposition 1.6. *Every k-form ω has a canonical decomposition of the form $\pi_{1,0}^*\omega = \omega_h + \omega_{c,1} + \cdots + \omega_{c,k}$, where ω_h is a unique horizontal form, and $\omega_{c,m}$ are uniquely defined m-contact k-forms, for $m = 1, \ldots, k$.*

1.2. Higher-order jets. In chapter 9, we will encounter higher-order field theories, in particular with Lagrangians of order 2, *i.e.* depending on the fields and their first- and second-order derivatives. To deal with that kind of field theories, we introduce the manifolds $J^k\pi$ of higher-order jets. Most of the theory of higher-order jet bundles is similar to the first-order case, but there are a number of new aspects, such as iterated jet bundles, that deserve additional attention.

Let x be a point in X. We define an equivalence relation on the set of local sections $\Gamma_x(\pi)$ of π at x and declare two sections to be k-*equivalent* if their Taylor expansions at x in a coordinate chart agree up to the kth order. It can be shown that if these Taylor expansions agree in any one coordinate chart centered at x, then they agree in all coordinate charts at x. By analogy to the first-order case, the equivalence class of a local section ϕ at x is denoted by $j^k_x\phi$.

Definition 1.7. *The kth order jet manifold $J^k\pi$ is the set of all such equivalence classes:*

$$J^k\pi := \{j^k_x\phi \quad \text{where}\, x \in M \quad \text{and} \quad \phi \in \Gamma_x(\pi)\},$$

where $\Gamma_x(\pi)$ has a similar meaning as in definition 1.1.

Note that, according to this definition, $J^0\pi$ is just Y.

The kth order jet bundle is equipped with a number of projections $\pi_{k,l} : J^k\pi \to J^l\pi$ (where $l \leq k$), constructed by "truncating" to order l the Taylor expansion defining an element of $J^k\pi$. Formally, we define $\pi_{k,l}$ as follows: $\pi_{k,l}(j^k_x\phi) = j^l_x\phi$. In addition, there exist projections $\pi_k : J^k\pi \to X$ defined as $\pi_k = \pi \circ \pi_{k,0}$.

A natural coordinate system on $J^k\pi$ is given by $(x^\mu, y^a; y^a_\mu; y^a_{\mu_1\mu_2}, \ldots, y^a_{\mu_1\cdots\mu_k})$, for $a = 1, \ldots, m$ and $\mu_i = 0, \ldots, n$, with the restriction that

$$y^a_{\mu_1\cdots\mu_l} = y^a_{\sigma(\mu_1\cdots\mu_l)} \quad \text{for all}\, 0 < l \leq k, \tag{1.4}$$

and for any permutation $\sigma \in S_l$. These coordinate functions are defined as follows:

$$y^a_{\mu_1\cdots\mu_l}(j^k_x\phi) := \frac{\partial^l\phi}{\partial x^{\mu_1}\cdots\partial x^{\mu_l}}(x).$$

The condition (1.4) hence expresses the commutativity of the partial derivatives.

An interesting observation is that, for any positive k and l, the $(k+l)$th order jet bundle $J^{k+l}\pi$ is embedded in the iterated jet bundle $J^k\pi_l$. Indeed, there exists an embedding $\iota_{k,l} : J^{k+l}\pi \hookrightarrow J^k\pi_l$ defined by $\iota_{k,l}(j^{k+l}_x\phi) := j^k_x(j^l\phi)$. All this is similar to the case of tangent bundles, where the iterated tangent bundle $T(TQ)$ has a distinguished submanifold T^2Q, whose elements are equivalence classes of "osculating curves".

In order to derive the coordinate form of the embedding $\iota_{k,l}$, we have to set up a number of rather involved coordinate systems. The iterated jet bundle $J^k\pi_l$ is equipped with a coordinate system of the following form:

$$(x^\mu; y^a_{\mu_1\cdots\mu_p;\nu_1\cdots\nu_q}), \quad \text{for all}\, p = 0, \ldots, l, \text{ and } q = 0, \ldots, k.$$

Again, the rough idea is that $y^a_{\mu_1\cdots\mu_p;\nu_1\cdots\nu_q}$ represents the "derivative" of $y^a_{\mu_1\cdots\mu_p}$ with respect to $\{x^{\nu_1}, \ldots, x^{\nu_q}\}$. For the sake of notational simplicity, we have adopted the convention that if $p = 0$, then $y^a_{\mu_1\cdots\mu_p;\nu_1\cdots\nu_q}$ is just $y^a_{;\nu_1\cdots\nu_q}$, and similarly for the case where $q = 0$: $y^a_{\mu_1,\ldots,\mu_p;\nu_1,\ldots,\nu_q}$ is then just $y^a_{\mu_1,\ldots,\mu_p}$

The fibre coordinates satisfy the following symmetry condition (compare with (1.4)):

$$y^a_{\sigma(\mu_1\cdots\mu_p);\tau(\nu_1\cdots\nu_q)} = y^a_{\mu_1\cdots\mu_p;\nu_1\cdots\nu_q} \quad \text{for all}\, \sigma \in S^p, \tau \in S^q,$$

for all $p = 0, \ldots, l$ and $q = 0, \ldots, k$.

By comparison with (1.4), it follows that $J^{k+l}\pi$ is precisely the subset of $J^k\pi_l$ whose elements have coordinates that are symmetric under the full permutation group S^{p+q} for all $p = 0, \ldots, l$ and $q = 0, \ldots, k$.

Finally, $J^{k+l}\pi$ is characterized as a submanifold of $J^{k+l}\pi$ by the condition

$$y^a_{\mu_1\cdots\mu_p;\nu_1\cdots\nu_q} = y^a_{\sigma(\mu_1\cdots\mu_p;\nu_1\cdots\nu_q)} \quad (= y^a_{\mu_1\cdots\mu_p\nu_1\cdots\nu_q}), \tag{1.5}$$

for all $\sigma \in S^{p+q}$, where $p = 0, \ldots, l$ and $q = 0, \ldots, k$. In coordinates, $\iota_{k,l} : J^{k+l}\pi \hookrightarrow J^k\pi_l$ is given by $\iota_{k,l}(y^a_{\mu_1\cdots\mu_p\nu_1\cdots\nu_q}) = y^a_{\mu_1\cdots\mu_p;\nu_1\cdots\nu_q}$, with the same notations as above.

In this thesis, we will only be concerned with $J^2\pi$ and its embedding into $J^1\pi_1$. The latter has a coordinate system $(x^\mu, y^a, y^a_\mu; y^a_{;\nu}, y^a_{\mu;\nu})$, where $a = 1, \ldots, k$ and $\mu, \nu = 1, \ldots, n+1$. Note that $y^a_{\mu;\nu}$ is not symmetric under the exchange of μ and ν. According to (1.5), the image of $J^2\pi$ under $\iota_{1,1}$ is the submanifold of $J^1\pi_1$ determined by the following equation:

$$y^a_{;\mu} = y^a_\mu \quad \text{and} \quad y^a_{\mu;\nu} = y^a_{\nu;\mu} = y^a_{\mu\nu}. \tag{1.6}$$

1.3. Connections on fibre bundles. The concept of a connection on a manifold, or by extension, of a connection on a fibre bundle, is central in differential geometry. In this thesis, we will use the definition of connection given by Charles Ehresmann. In his view, a connection on a fibre bundle $\pi : Y \to X$ is a smooth n-dimensional distribution on Y which is transversal to the vertical bundle $V\pi$. Equivalently, one has the following definition.

Definition 1.8 (see [**94**], def. 3.5.1)**.** *An* Ehresmann connection *on π is a vector-valued one-form $\Upsilon \in \Omega^1_0(\pi) \otimes \mathfrak{X}(Y)$ such that $\Upsilon \lrcorner\, \sigma = \sigma$ for every $\sigma \in \Omega^1_0(\pi)$. (Recall that $\Omega^1_0(Y)$ denotes the module of semi-basic 1-forms with respect to π.)*

In bundle coordinates, such a connection is locally given by

$$\Upsilon = \mathrm{d}x^\mu \otimes \left(\frac{\partial}{\partial x^\mu} + \Gamma^a_\mu(x,y)\frac{\partial}{\partial y^a} \right).$$

Let Υ be an Ehresmann connection. The image H of Υ, regarded as map from π^*TX to TY, is a bundle of n-planes on Y determining a connection in the original sense of Ehresmann, as H is transversal to the vertical bundle $V\pi$. The converse is also true; this is treated in lemma 3.5.3 and lemma 3.5.4 in [**94**]. Summarizing, we have:

Proposition 1.9. *Every Ehresmann connection determines a subbundle H of the tangent bundle TY which is transversal to the vertical bundle: $TY = H \oplus V\pi$. Conversely, each such decomposition determines an Ehresmann connection.*

In coordinates, H is spanned by vector fields of the form

$$\left(\frac{\partial}{\partial x^\mu} \right)^H := \frac{\partial}{\partial x^\mu} + \Gamma^a_\mu(x,y)\frac{\partial}{\partial y^a}.$$

Let $TY = H \oplus V\pi$ be a decomposition of the tangent bundle as in the above proposition. Vectors contained in H will be called *horizontal*; H is the *horizontal distribution.* Associated to such a decomposition is a set of complementary projectors

$$\mathbf{h} : TY \to H \quad \text{and} \quad \mathbf{v} : TY \to V\pi,$$

referred to as the *horizontal* and *vertical* projector, respectively. Based on the concept of projections in the tangent bundle, one can come up with more general definitions of connections (see [**60**]).

A final, particularly fruitful interpretation of a connection is that of a *jet field*, *i.e.* a section of the bundle $\pi_{1,0} : J^1\pi \to Y$, for which we will also use the notation Υ. We may associate a horizontal distribution H to a jet field $\Upsilon : Y \to J^1\pi$ as follows: for each $y \in Y$, $\Upsilon(y)$ is an element of $J^1\pi$ and, hence, can be viewed as a linear map $\Upsilon(y) : T_xX \to T_yY$, where $x = \pi(y)$ (see the alternative characterisation of jets immediately before remark 1.2). Now, define $H(y)$ as the image of $\Upsilon(y)$. The assignment $y \mapsto H(y)$ defines a horizontal distribution on Y. A full proof of the equivalence between jet fields and Ehresmann connections is given in [**94**], proposition 4.6.3. In coordinates, the jet field associated to an Ehresmann connection is given by

$$\Upsilon : (x^\mu, y^a) \mapsto (x^\mu, y^a, \Gamma^a_\mu(x,y)).$$

An *integral section* of a jet field $\Upsilon : Y \to J^1\pi$ is a local section ϕ of π such that $j^1\phi = \Upsilon \circ \phi$. Integral sections do not always exist, not even locally: their existence is related to the curvature of the Ehresmann connection associated to the jet field.

Definition 1.10. *Let $\pi : Y \to X$ be a fibre bundle and Υ an Ehresmann connection on π with horizontal projector $\mathbf{h}$. The* curvature *of Υ is the vector-valued two-form R_Υ on Y defined as*

$$R_\Upsilon(X_1, X_2) := \mathbf{h}([X_1, X_2]) + [\mathbf{h}(X_1), \mathbf{h}(X_2)] - \mathbf{h}([\mathbf{h}(X_1), X_2]) - \mathbf{h}([X_1, \mathbf{h}(X_2)]),$$

for arbitrary vector fields X_1, X_2 on Y.

In coordinates, the curvature is given by $R_\Upsilon = R^a_{\mu\nu} \mathrm{d}x^\mu \otimes \mathrm{d}x^\nu \otimes \frac{\partial}{\partial y^a}$, where

$$R^a_{\mu\nu} = \frac{\partial \Gamma^a_\nu}{\partial x^\mu} - \frac{\partial \Gamma^a_\mu}{\partial x^\nu} + \Gamma^b_\mu \frac{\partial \Gamma^a_\nu}{\partial y^b} - \Gamma^b_\nu \frac{\partial \Gamma^a_\mu}{\partial y^b} \tag{1.7}$$

and Γ^a_μ are the connection coefficients of Υ.

The curvature measures the lack of integrability of the horizontal distribution, and is equal to one half of the Nijenhuis torsion of the horizontal projector $\mathbf{h}$ (proposition 3.5.14 in [**94**]). Similarly, when we interpret the connection as a jet field, the curvature is the obstruction for the existence of integral sections.

Remark 1.11. We now have three equivalent characterizations of Ehresmann connections. As we will not be using other types of connections, we will refer to all three concepts of Ehresmann connections simply as *connections.* ◇

1.3.1. *Connections on $\pi_1 : J^1\pi \to X$.* The definition of connections on the first jet bundle proceeds just as in the case of arbitrary bundles. However, the special nature of the first jet bundle leads to a number of additional interesting properties.

Recall that $J^2\pi$ is a submanifold of the iterated jet bundle $J^1\pi_1$ by (1.6). We define a *second-order jet field* as a jet field on π_1 taking values in $J^2\pi$, *i.e.* a section of $\pi_{2,1} : J^2\pi \to J^1\pi$. A standard (*i.e.* not necessarily second-order) jet field in $(\pi_1)_{1,0} : J^1\pi_1 \to J^1\pi$ has the following coordinate form:

$$\Upsilon : (x^\mu, y^a; y^a_\mu) \mapsto (x^\mu, y^a, y^a_\mu; y^a_{;\mu} = \Gamma^a_\mu(x^\kappa, y^b, y^b_\kappa), y^a_{\mu;\nu} = \Gamma^a_{\mu\nu}(x^\kappa, y^b, y^b_\kappa)). \tag{1.8}$$

From the coordinate expressions (1.6) defining $J^2\pi$ as a submanifold of $J^1\pi_1$, we see that Υ is of second order if and only if the following conditions hold:

$$\Gamma^a_\mu = y^a_\mu \quad \text{and} \quad \Gamma^a_{\mu\nu} = \Gamma^a_{\nu\mu}. \tag{1.9}$$

If a second-order jet field is integrable, then its integral sections are prolongations of sections of $\pi : Y \to X$. As we shall see, second-order jet fields are the geometric counterpart of systems of second-order PDEs. Let us first define a connection with horizontal projector $\mathbf{h}$ to be *semi-holonomic* if

$$i_{\mathbf{h}}\theta = 0, \qquad \text{for each contact 1-form } \theta. \tag{1.10}$$

For the definition of the contraction operator of a $(1, 1)$-tensor field with a k-form, we refer the reader to appendix A.

In coordinates, the horizontal projector $\mathbf{h}$ of a connection on π_1 can be written as

$$\mathbf{h} = \mathrm{d}x^\mu \otimes \left(\frac{\partial}{\partial x^\mu} + \Gamma^a_\mu \frac{\partial}{\partial y^a} + \Gamma^a_{\mu\nu} \frac{\partial}{\partial y^a_\nu} \right), \tag{1.11}$$

and the associated jet field Υ has the form indicated in (1.8). A local section σ of π_1, with $\sigma(x) = (x^\mu, \sigma^a(x), \sigma^a_\mu(x))$, is an integral section of the jet field Υ if

$$\frac{\partial \sigma^a}{\partial x^\mu} = \Gamma^a_\mu(x, \sigma^b, \sigma^b_\kappa) \quad \text{and} \quad \frac{\partial \sigma^a_\mu}{\partial x^\nu} = \Gamma^a_{\mu\nu}(x, \sigma^b, \sigma^b_\kappa). \tag{1.12}$$

From this expression, it follows that σ can be written as $\sigma = j^1\phi$, where ϕ is a section of π, if and only if $\Gamma^a_\mu = y^a_\mu$: in that case, the first condition of (1.12) translates to $\frac{\partial \sigma^a}{\partial x^\mu} = \sigma^a_\mu$, and the connection is at least semi-holonomic.

Similarly, second-order jet fields can be equivalently characterized in terms of certain classes of contact forms (see proposition 5.4.6 in [**94**]). This characterisation will not be needed in this thesis: it suffices to note that second-order jet fields are necessarily semi-holonomic, and if they are integrable, their integral sections satisfy the following system of second-order PDEs:

$$\frac{\partial^2 \phi^a}{\partial x^\mu \partial x^\nu} = \Gamma^a_{\mu\nu} \left(x, \phi^b, \frac{\partial \phi^b}{\partial x^\kappa} \right).$$

A necessary, though not sufficient condition for integrability of this system is that that $\Gamma^a_{\mu\nu} = \Gamma^a_{\nu\mu}$, which is precisely the second condition in (1.9).

2. Lagrangian field theories

The geometry of classical field theories has been studied by many authors and is by now well established. In this section we recall some basic aspect of that theory. We start by giving a brief overview of first-order field theories; for a comprehensive treatment, we refer the interested reader to [**11**,**22**,**36**,**48**,**94**] and the references therein. In section 2.2, we study the effects of a symmetry action, and we recall Noether's theorem. Finally, in subsection 2.3, we turn our attention to field theories of second order (which will play a major role in chapter 9), and we indicate some of the differences with first-order field theories.

2.1. Covariant field theories of first order. Let there be given a fibre bundle $\pi : Y \to X$, whose (local) sections represent fields.

2.1.1. *The Poincaré-Cartan* $(n+2)$*-form.* A fundamental object in Lagrangian field theory is the so-called Poincaré-Cartan form, an $(n+2)$-form associated to a given Lagrangian, which is related to the variational background of the Euler-Lagrange equations. The Poincaré-Cartan form can also be used to recast the Euler-Lagrange equations in an intrinsic form.

Definition 2.1. *A* Lagrangian density *is an* $(n+1)$*-form* $\mathcal{L}$ *along the projection* π_1*. Equivalently, a Lagrangian density is a horizontal* $(n+1)$*-form (with respect to the projection* π_1*) on* $J^1\pi$*.*

Both aspects of this definition can be rephrased as follows: a Lagrangian density is a map $\mathcal{L} : J^1\pi \to \bigwedge^{n+1}(X)$ such that, for all $\gamma \in J^1\pi$, $\mathcal{L}(\gamma) \in \bigwedge_x^{n+1}(X)$, where $x = \pi_1(\gamma)$.

Since X is equipped with a fixed volume form η, any Lagrangian density can be written as $\mathcal{L} = L\eta$, where L is a function on $J^1\pi$, called the *Lagrangian.* From this point of view, the explicit distinction between Lagrangian densities and the associated Lagrangian functions is almost trivial, but note that one can easily think of situations (*i.e.* in general relativity) where it is not desirable to consider a fixed volume form on the base space. Such situations will not occur in this thesis, however.

A Lagrangian L is said to be *regular* if its Hessian matrix is non-degenerate, *i.e.*

$$\det \left(\frac{\partial^2 L}{\partial y^a_\mu \partial y^b_\nu} \right) \neq 0$$

at each point of $J^1\pi$.

Using the vertical endomorphism S_η, we can now construct the following $(n+1)$-form on $J^1\pi$, called the *(first-order) Cartan form*:

$$\Theta_L := S_\eta^* \mathrm{d}L + L\eta$$

and we then define a particular $(n+2)$-form, called the *Poincaré-Cartan form*, as $\Omega_L := -\mathrm{d}\Theta_L$. If L is regular, which we will always assume in the sequel, the Poincaré-Cartan form is a multisymplectic form according to the following definition.

Definition 2.2 (see [**21**, **33**, **35**])**.** *A closed* m*-form* Ω *on a manifold* M *is called* multisymplectic *if the mapping* $v \in T_xM \mapsto i_v\Omega(x) \in \bigwedge_x^{m-1}(M)$ *is injective for all* $x \in M$*.*

Note that symplectic forms ($m = 2$) and volume forms ($m = \dim M$) are particular examples of multisymplectic forms and, moreover, these are the only two cases where the mapping in definition 2.2 is surjective as well as injective (assuming M is finite dimensional). Despite their apparent similarity, multisymplectic geometry is generally quite different from symplectic geometry. One of many important differences is that there is no Darboux theorem for general multisymplectic forms.

Remark 2.3. Contrary to what their definition may suggest, Θ_L and Ω_L are associated to the Lagrangian density rather than to the Lagrangian itself, *i.e.* if $L\eta = L'\eta'$ for some other volume form η' on X, then $\Theta_L = \Theta_{L'}$, provided the volume form η' is used to construct $\Theta_{L'}$. A similar remark can be made about Ω_L. ◇

The following coordinate expressions for Θ_L and Ω_L will often be convenient:

$$\Theta_L = \frac{\partial L}{\partial y^a_\mu}(\mathrm{d}y^a - y^a_\nu \mathrm{d}x^\nu) \wedge \mathrm{d}^n x_\mu + L\mathrm{d}^{n+1}x$$

and

$$\Omega_L = -\frac{\partial L}{\partial y^a}\mathrm{d}y^a \wedge \mathrm{d}^{n+1}x - \mathrm{d}\left(\frac{\partial L}{\partial y^a_\mu}\right) \wedge (\mathrm{d}y^a - y^a_\nu \mathrm{d}x^\nu) \wedge \mathrm{d}^n x_\mu. \tag{1.13}$$

2.1.2. *The Euler-Lagrange equations.* Let U be an open subset of X with compact closure, and define the *action functional* S as

$$S(\phi) = \int_U L(j^1\phi)\eta, \tag{1.14}$$

for each local section ϕ of π whose support is contained in U.

An *infinitesimal variation* of such a section ϕ is a vertical vector field V along ϕ such that $V(x) = 0$ for all $x \in \partial U$. A *finite variation* of ϕ is a local one-parameter group of diffeomorphisms $\{\varphi_\epsilon\}$, (where ϵ takes values in an open interval $(-a, a)$ containing zero) defined on a neighbourhood of $\phi(U)$, and satisfying the following conditions: for each $\epsilon \in (-a, a)$,

(1) φ_ϵ respects the fibered structure of Y, *i.e.* $\pi \circ \varphi_\epsilon = \pi$;
(2) φ_ϵ is the identity on the boundary of $\phi(U)$.

A section ϕ is an *extremal* or *critical point* of (1.14) if

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{S}(\varphi_\epsilon \circ \phi)\Big|_{\epsilon=0} = 0,$$

for any finite variation $\{\varphi_\epsilon\}$. A standard argument then shows that the extremals of (1.14) are characterized by the following set of *Euler-Lagrange equations*:

$$\left[\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^a_\mu}\right)\right](j^2\phi) = 0. \tag{1.15}$$

These partial differential equations can be rewritten in intrinsic form by means of the Poincaré-Cartan form:

Theorem 2.4. *A section ϕ of π is a critical point of the action S, or, equivalently, satisfies the Euler-Lagrange equations (1.15), if and only if*

$$(j^1\phi)^* i_W \Omega_L = 0 \tag{1.16}$$

for all vector fields W on $J^1\pi$.

Proof: See, for instance, [**11**, prop. 7.1.2], [**48**, thm. 3.1] or [**22**]. ◇

In this thesis, we will mostly be concerned with a kind of "linearized version" of the Euler-Lagrange equations obtained by looking for a connection on π_1 whose integral sections will be extremals of the action S. More precisely, we have the following important proposition:

Proposition 2.5. *Let* $\mathbf{h}$ *be the horizontal projector of a holonomic connection on* π_1*. The integral sections of the associated jet field are extremals of (1.14) if and only if*

$$i_{\mathbf{h}}\Omega_L = n\Omega_L. \tag{1.17}$$

Proof: See [**94**, thm. 5.5.5] and [**31**]. ◇

Moreover, a simple coordinate computation shows that, given a regular Lagrangian L on $J^1\pi$, a connection on π_1 satisfying (1.17) will automatically be semi-holonomic. Equation (1.17) is also referred to as the *De Donder-Weyl equation* of Lagrangian field theory.

Remark 2.6. One can prove that the integral sections of a solution $\mathbf{h}$ of the De Donder-Weyl equation are sections σ of π_1 such that $\sigma^*(i_W\Omega_L) = 0$ for all vector fields W on $J^1\pi$, where σ need not be the prolongation of a section of π. These equations are referred to as the *De Donder equations* of Lagrangian field theory. In the case of a regular Lagrangian, one can furthermore show that any solution σ of the De Donder equations is necessarily the prolongation of a solution of the Euler-Lagrange equations (see for instance [**10**]). The De Donder equations will not be used in the remainder of this thesis. ◇

2.2. Symmetries and Noether's theorem. Let $\mathcal{G}$ be a Lie group acting on $\pi : Y \to X$ by bundle automorphisms. By this, we mean that there exist Lie group actions $\underline{\Phi} : \mathcal{G} \times X \to X$ and $\overline{\Phi} : \mathcal{G} \times Y \to Y$, such that, for each $g \in G$, the pair $(\underline{\Phi}_g, \overline{\Phi}_g)$ (collectively denoted by Φ_g) is a bundle automorphism. Here, $\underline{\Phi}_g$ is a shorthand notation for the diffeomorphism $\underline{\Phi}(g, \cdot)$, and similarly for $\overline{\Phi}_g$.

Such an action induces an action on $J^1\pi$ by prolongation, where $\mathcal{G}$ acts on $J^1\pi$ by the action which assigns to each $g \in \mathcal{G}$ the prolongation $j^1\Phi_g$. Consider now an element ξ of the Lie algebra $\mathfrak{g}$ and denote the infinitesimal generator of the prolonged action corresponding to ξ by $\xi_{J^1\pi}$. Note that $\xi_{J^1\pi}$ is just $j^1\xi_Y$, the prolongation of the infinitesimal generator on Y corresponding to ξ.

We say that a Lagrangian density $\mathcal{L}$ is *invariant* under the prolonged action of $\mathcal{G}$ if $(j^1\Phi_g)^*\mathcal{L} = \mathcal{L}$ for all $g \in \mathcal{G}$. Here, we have interpreted $\mathcal{L}$ as a horizontal $(n+1)$-form on $J^1\pi$. If we write the Lagrangian density as $\mathcal{L} = L\eta$, then invariance of $\mathcal{L}$ implies the following equivariance condition for the Lagrangian function:

$$L(j^1\Phi_g(\gamma)) = L(\gamma)\,\mathrm{Jac}(\Phi_g)(x), \quad \text{for all } g \in \mathcal{G}, \gamma \in J^1\pi, \text{ and where } x = \pi_1(\gamma).$$

Here, $\mathrm{Jac}(\Phi_g)$ denotes the Jacobian of the diffeomorphism $\underline{\Phi}_g$. It can be shown that equivariance of the Lagrangian, in the sense defined above, implies invariance of the Cartan $(n+1)$-form, expressed as $(j^1\Phi_g)^*\Theta_L = \Theta_L$ for all $g \in \mathcal{G}$, or, infinitesimally,

$$\mathscr{L}_{\xi_{J^1\pi}}\Theta_L = 0. \tag{1.18}$$

Let L be a $\mathcal{G}$-equivariant Lagrangian. To each $\xi \in \mathfrak{g}$, one can associate an n-form J^L_ξ according to $J^L_\xi := \xi_{J^1\pi} \lrcorner\, \Theta_{\mathcal{L}}$. We now introduce the *momentum map* J^L as the element of $\Omega^n(J^1\pi) \otimes \mathfrak{g}^*$ defined by $\left\langle J^L, \xi \right\rangle = J^L_\xi$ for all $\xi \in \mathfrak{g}$. The importance of the momentum map lies in the following theorem, which we have taken here from [**48**, thm. 4.7]:

Proposition 2.7 (Noether)**.** *Let $\mathcal{L}$ be an invariant Lagrangian density. For all $\xi \in \mathfrak{g}$, the following conservation law holds:*

$$\mathrm{d}[(j^1\phi)^* J^L_\xi] = 0,$$

for all sections ϕ of π that are solutions of the Euler-Lagrange equations (1.15).

Remark 2.8. In this section, we considered only Lie group actions on $J^1\pi$ that are prolongations of actions on Y. This special case will be sufficient for the remainder of this thesis, but one can equally well envisage actions that are not prolongations. In that case, the existence of a momentum map imposes additional conditions on the action. This is also the case in mechanics. ◇

2.3. Second-order field theories. Many of the Lagrangians arising in elasticity are of higher order. In particular, we will encounter a second-order model in chapter 9. We now recall a number of results from the geometric formalism for second-order field theories (see [**61**, **94**] and the references therein). Much of this formalism is similar to the first-order case, but a brief warning is in order here. In the first-order case, both the Lagrangian and the Poincaré-Cartan form reside on $J^1\pi$. For kth order field theories, however, the Lagrangian is a function on $J^k\pi$, as one would expect, but the Poincaré-Cartan form lives on $J^{2k-1}\pi$.

A second-order Lagrangian is a function L on $J^2\pi$. Associated to L is a *second-order Cartan form*, an $(n+1)$-form on $J^3\pi$, whose coordinate expression reads

$$\begin{aligned} \Theta_L \;&=\; \left[\frac{\partial L}{\partial y^a_\nu} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^a_{\nu\mu}}\right)\right] \mathrm{d}y^a \wedge \mathrm{d}^n x_\nu + \frac{\partial L}{\partial y^a_{\nu\mu}} \mathrm{d}y^a_\nu \wedge \mathrm{d}^n x_\mu \\ &+\; \left[L - \frac{\partial L}{\partial y^a_\nu} y^a_\nu + \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^a_{\nu\mu}}\right) y^a_\nu - \frac{\partial L}{\partial y^a_{\nu\mu}} y^a_{\nu\mu}\right] \mathrm{d}^{n+1}x. \end{aligned} \tag{1.19}$$

Let us also define the second-order Poincaré-Cartan form as $\Omega_L := -\mathrm{d}\Theta_L$.

Remark 2.9. In expressions such as (1.19) and (1.20) below, the implied sum in a term of the form

$$\frac{\partial L}{\partial y^a_{\nu\mu}} y^a_{\nu\mu}$$

is over symmetric pairs of indices μ, ν only. ◇

Many results from the previous section on first-order field theories carry over immediately to the higher-order case. Let U be again an open subset of X with compact closure and define the action S as

$$S(\phi) = \int_U L(j^2\phi)\eta,$$

where ϕ is a section of π with support contained in U. A section ϕ is a critical point of this functional (under arbitrary variations defined exactly as in section 2.1.2) if and only if it satisfies the *second-order Euler-Lagrange equations*:

$$\left[\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^a_\mu}\right) + \frac{\mathrm{d}^2}{\mathrm{d}x^\mu \mathrm{d}x^\nu}\left(\frac{\partial L}{\partial y^a_{\mu\nu}}\right)\right](j^4\phi) = 0. \tag{1.20}$$

There also exists an intrinsic formulation of the Euler-Lagrange equations. We quote from [**61**]:

Proposition 2.10. *Let L be a second-order Lagrangian. A section ϕ of π is a solution of the second-order Euler-Lagrange equations if and only if $(j^3\phi)^*(W \lrcorner \Omega_L) = 0$ for all vector fields W on $J^3\pi$.*

Remark 2.11. It should be noted that there always exists a Cartan form for higher-order field theories, but that uniqueness is not always guaranteed (contrary to the first-order case). However, by imposing additional conditions, Saunders [**94**] was able to prove uniqueness for second-order field theories. This unique form, given in (1.19), was derived by Kouranbaeva and Shkoller [**61**] by means of a variational argument. ⋄

Just as in the first-order case, the action of a Lie group $\mathcal{G}$ acting on π by bundle automorphisms gives rise to a prolonged action on $J^2\pi$. If L is a $\mathcal{G}$-equivariant Lagrangian with respect to this action, then the momentum map $J^L \in \Omega^n(J^3\pi) \otimes \mathfrak{g}^*$, defined as $\left\langle J^L, \xi \right\rangle = J^L_\xi$, where $J^L_\xi = \xi_{J^3\pi} \lrcorner \Theta_L$, is conserved: $\mathrm{d}[(j^3\phi)^* J^L_\xi] = 0$ for all sections ϕ of π that are solutions of the Euler-Lagrange equations (1.20).

3. The Cauchy formalism

We now come to a radically different view of classical field theories. Whereas in the previous sections we have stressed the advantages of the covariant, finite-dimensional jet bundle approach, we will now introduce a different, complementary framework based on techniques from infinite-dimensional geometry. This approach is referred to as the Cauchy formalism, because it is a generalisation of the typical Cauchy formulation used for example in general relativity.

In this formalism, it is assumed that the base space X is diffeomorphic to a Cartesian product of the form $\mathbb{R} \times M$. The coordinates on M then play the role of "spatial coordinates", while the coordinate on $\mathbb{R}$ is a "coordinate time". It should be stressed that there can be some amount of arbitrariness in the specification of the diffeomorphism between X and $\mathbb{R} \times M$; as a result, for instance in general relativity the coordinate time need not coincide with physical time of an observer.

Roughly speaking, in the Cauchy formalism the evolution of the system is specified by an ODE on an infinite dimensional space, and the evolution is formally identical to that of a certain mechanical system, whose configuration space consists of embeddings of M into Y, while the velocity space consists of embeddings of M into $J^1\pi$. This setup is sometimes conceptually clearer than the jet bundle approach, but also allows the use of results from the analysis of ODEs, such as the Cauchy-Kowalewska theorem on the existence and uniqueness of solutions, which are not immediately accessible in the covariant framework.

On the other hand, there are many classical field theories for which the base space is itself just $\mathbb{R} \times M$. A typical example is nonrelativistic fluid dynamics, where M is the reference configuration of the fluid. The application of the results from Cauchy theory then leads to the Arnol'd formulation of fluid dynamics as an ODE on the diffeomorphism group of M. These observations were the starting point for the seminal results of Ebin and Marsden [**42**] in 1970 on the short-time existence of well-behaved solutions to the Euler equations in three dimensions.

In this section, we start by giving a brief overview of some formal aspects of Cauchy theory. We completely ignore any questions related to smoothness; our aim is just to show that there is a natural way to transfer the dynamics of a classical field from the first jet bundle to the space of Cauchy data, and that a connection solving the De Donder-Weyl equation (1.17) induces a second-order vector field on the space of Cauchy data which solves the equations of time-dependent mechanics on this space.

Remark 3.1. Of course, the division of the base space X into spatial variables and time is only meaningful for "evolution type" field theories, and excludes, for example, elliptic PDEs. While the covariant results in this thesis hold for arbitrary kinds of field theories, our most important examples will be hyperbolic. See [**25**] for a detailed treatment of this distinction. $\diamond$

3.1. The space of Cauchy data. We first recall some basic aspects of the Cauchy formalism for Lagrangian field theories, following the treatments presented in [**11**, **35**, **91**].

3.1.1. *Definitions.* Consider as usual a fibre bundle $\pi : Y \to X$ whose base space X is an $(n+1)$-dimensional oriented manifold with volume form η. Let M be an n-dimensional compact oriented manifold with volume form η_M. The pair (M, η_M) is called a *Cauchy surface.* Points of M will usually be denoted by u. In the sequel, we will always assume that M has volume one, *i.e.*

$$\int_M \eta_M = 1. \tag{1.21}$$

The volume form η_M on M will not play a significant role until section 3.1.3.

Definition 3.2. *The* space of (parametrized) Cauchy surfaces $\tilde{X}$ *is the space of all embeddings* $\tau : M \hookrightarrow X$.

We now define the infinite-dimensional analogue of, respectively, the “configuration space” and the “velocity space”.

Definition 3.3. *The* space of Dirichlet data *is the manifold* $\tilde{Y}$ *whose elements are embeddings* $\delta : M \hookrightarrow Y$ *having the property that there exists a section* ϕ *of* π *and an element* τ *of* $\tilde{X}$ *such that* $\delta = \phi \circ \tau$.

Definition 3.4. *The* space of Cauchy data *is the manifold* $\tilde{Z}$ *whose elements are embeddings from* M *into* $J^1\pi$, *having the property that for each embedding* $\kappa : M \hookrightarrow J^1\pi$, *there exists a section* ϕ *of* π *and an element* τ *of* $\tilde{X}$ *such that* $\kappa = j^1\phi \circ \tau$.

The respective projections $\pi_{1,0} : J^1\pi \to Y$ and $\pi : Y \to X$ induce by composition natural projections $\tilde{\pi}_{1,0} : \tilde{Z} \to \tilde{Y}$ and $\tilde{\pi} : \tilde{Y} \to \tilde{X}$: $\tilde{\pi}_{1,0}(\kappa)$ is defined as $\pi_{1,0} \circ \kappa$, and $\tilde{\pi}$ is defined similarly. We further introduce the projection $\tilde{\pi}_1$ as $\tilde{\pi}_1 := \tilde{\pi} \circ \tilde{\pi}_{1,0}$.

The spaces $\tilde{X}$, $\tilde{Y}$, and $\tilde{Z}$ can be given smooth manifold structures (see remark 3.5 below). Tangent vectors to each of these manifolds have convenient finite-dimensional interpretations. Take for instance an element $V_\tau \in T_\tau\tilde{X}$. Formally, V_τ is the tangent vector at $\epsilon = 0$ of a curve $\epsilon \mapsto c(\epsilon)$ of embeddings such that $c(0) = \tau$. By taking a fixed element $u \in M$ and applying each embedding $c(\epsilon) : M \hookrightarrow X$ to that element, we obtain a curve in X, denoted by c_u and defined as $c_u(\epsilon) := c(\epsilon)(u)$. Hence, V_τ can be interpreted as the vector field along τ defined by

$$V_\tau : M \to TX, \quad V_\tau : u \mapsto V_\tau(u) = \left.\frac{\mathrm{d}c_u(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \in T_{\tau(u)}X.$$

Conversely, each vector field along τ is an element of $T_\tau\tilde{X}$. As those vector fields are just sections of the pullback bundle τ^*TX, we have the identification

$$T_\tau\tilde{X} \cong \Gamma(\tau^*TX). \tag{1.22}$$

Furthermore, note that since τ is a bijection onto its image, one can identify V_τ with a vector field on X, defined along the submanifold $\tau(M)$. A second point of interest is that any vector field v on X induces a vector field V on $\tilde{X}$ by composition: simply define $V(\tau)$ as $v \circ \tau$.

The tangent vectors to $\tilde{Y}$ and $\tilde{Z}$ have similar interpretations as vector fields along elements of $\tilde{Y}$ and $\tilde{Z}$, respectively. However, because of the additional conditions on the elements of $\tilde{Y}$ and $\tilde{Z}$, not all such vector fields along maps are tangent vectors. Precise characterisations were given in [**91**] but will not be needed here.

Remark 3.5. A few technical remarks are in order here.

(1) The spaces $\tilde{X}$, $\tilde{Y}$, and $\tilde{Z}$ can be made into smooth infinite-dimensional manifolds in a number of ways. For more information, we refer to [**11, 63**]
(2) Usually, X and M are taken to be manifolds with boundary and the embeddings τ belonging to $\tilde{X}$ are then assumed to map the interior and the boundary of M into the interior and the boundary of X, respectively. In this way, boundary conditions can be taken into account. However, in chapter 9 these boundaries are merely zero dimensional and can therefore be included without additional complications. ◇

3.1.2. *Integration of forms.* In the previous section, we saw that there is a close relation between, for instance, vector fields on $J^1\pi$ and tangent vectors to $\tilde{Z}$. Similarly, integration of forms provides a means of turning $(n+k)$-forms on $J^1\pi$ into k-forms on $\tilde{Z}$. Let α be an $(n+k)$-form on $J^1\pi$. We then define $\tilde{\alpha}$ as follows:

$$\tilde{\alpha}(\kappa)(W_1, \ldots, W_k) = \int_M \kappa^*(i_{W_1 \wedge \cdots \wedge W_k}\alpha), \tag{1.23}$$

for $W_1, \ldots, W_k \in T_\kappa \tilde{Z}$. Here, the pull-back is defined by

$$\kappa^*(i_{W_1 \wedge \cdots \wedge W_k}\alpha)(u)(V_1, \ldots, V_n) = \alpha(\kappa(u))(W_1(u), \ldots, W_k(u), T_u\kappa(V_1), \ldots, T_u\kappa(V_n)),$$

for all $V_1, \ldots, V_n \in T_u M$.

Not all k-forms on $\tilde{Z}$ can be obtained by integrating a suitable density. Conversely, the mapping $\alpha \mapsto \tilde{\alpha}$ is not injective either: any $(n+k)$-form α which is at least $(k+1)$-contact induces the zero form on $\tilde{Z}$. This is proved in the next proposition.

Proposition 3.6. *Let κ be an element of $\tilde{Z}$ and consider an $(n+k)$-form α. Then $\tilde{\alpha} = 0$ if and only if α is at least $(k+1)$-contact.*

Proof: Let α be an $(n+k)$-form on $J^1\pi$ and assume that $\tilde{\alpha}(\kappa)(W_1, \ldots, W_k) = 0$ for all $\kappa \in \tilde{Z}$ and all $W_1, \ldots, W_k \in T_\kappa \tilde{Z}$. In particular, consider a set of k arbitrary vector fields $v_1, \ldots, v_k$ on $J^1\pi$ and put $W_i = v_i \circ \kappa$, for $i = 1, \ldots, k$.

We also have that, for every function f on M with compact support $U \subset M$, the following integral is zero:

$$\int_U f(\kappa^* i_{W_1 \wedge \cdots \wedge W_k}\alpha) = 0.$$

Indeed, replace for example W_1 by fW_1 in (1.23). A standard argument then shows that $\kappa^*(i_{W_1 \wedge \cdots \wedge W_k}\alpha) = 0$. Now, it is easy to see that

$$\kappa^*(i_{W_1 \wedge \cdots \wedge W_k}\alpha) = \tau^* j^1\phi^*(i_{v_1 \wedge \cdots \wedge v_k}\alpha),$$

where we have written κ as $j^1\phi \circ \tau$. As $j^1\phi$ and τ are arbitrary, this implies that $i_{v_1\wedge\cdots\wedge v_k}\alpha$ is at least 1-contact, and α itself is then at least $(k+1)$-contact.

Conversely, if α is an at least $(k+1)$-contact form, then it is easily seen that $\tilde{\alpha} = 0$. ⋄

3.1.3. *Splitting the base manifold.* We now introduce an additional element into our discussion: we assume that X is diffeomorphic to the product space $\mathbb{R} \times M$. As mentioned in the introduction to section 3, the philosophy underlying this assumption is that we somehow make a distinction between time and the spatial variables. As a result, we shall see that the sequence of bundles $\tilde{Z} \to \tilde{Y} \to \tilde{X}$ can be reduced to the sequence $J^1\tilde{\pi} \to \tilde{Y} \to \mathbb{R}$, a setup which is well known from time-dependent mechanics (on a infinite dimensional configuration space $\tilde{Y}$).

Let Ψ denote the diffeomorphism between X and $\mathbb{R} \times M$; Ψ is called a *splitting* of X. An element τ of $\tilde{X}$ is then called *admissible* if there exists a (necessarily unique) $t \in \mathbb{R}$ such that $\tau(u) = \Psi(t, u)$ for all $u \in M$.

Instead of considering the whole manifold $\tilde{X}$ as in the previous section, we now restrict $\tilde{X}$ to consist only of admissible embeddings. The spaces $\tilde{Y}$ and $\tilde{Z}$ are restricted accordingly by considering only Dirichlet and Cauchy elements projecting down (under $\tilde{\pi}$ and $\tilde{\pi}_{1,0}$, respectively) onto admissible embeddings.

By restricting our attention to admissible embeddings, we have constructed a bijection between $\mathbb{R}$ and $\tilde{X}$, by mapping each $t \in \mathbb{R}$ to the embedding $\Psi(t, \cdot)$. One can prove that Ψ is a diffeomorphism.

There is a canonically defined vector field $\mathbf{T}$ on $\tilde{X}$, defined as follows: for all $\tau \in \tilde{X}$, $\mathbf{T}(\tau) \in T_\tau\tilde{X}$ is given by

$$\mathbf{T}(\tau)(u) = \frac{\mathrm{d}}{\mathrm{d}s}\Psi(s, u)\bigg|_{s=t} \quad \text{for all } u \in M,$$

where t is such that $\tau(\cdot) = \Psi(t, \cdot)$. This vector field is just the push-forward of the vector field $\frac{\partial}{\partial t}$ under the diffeomorphism between $\mathbb{R}$ and $\tilde{X}$.

The elements of $\tilde{Y}$ and $\tilde{Z}$ are assumed to project down onto admissible embeddings. One can easily prove that, similarly, tangent vectors to these spaces have to project down onto $\mathbf{T}$ (or a multiple thereof).

Proposition 3.7. *Let δ be an element of $\tilde{Y}$ and consider $W_\delta \in T_\delta\tilde{Y}$. Then, for any $u \in M$, there exists a constant $k \in \mathbb{R}$ such that*

$$T_{\delta(u)}\pi(W_\delta(u)) = k\mathbf{T}(t, u), \quad \textit{or} \quad T_\delta\tilde{\pi}(W_\delta) = k\mathbf{T},$$

where $t = \tilde{\pi}(\delta)$. Similarly, if X_κ is any element of $T_\kappa\tilde{Z}$, then there exists for any $u \in M$ a constant k such that

$$T_{\kappa(u)}\pi_1(X_\kappa(u)) = k\mathbf{T}(t, u), \quad \textit{or} \quad T_\kappa\tilde{\pi}_1(X_\kappa) = k\mathbf{T}.$$

Proof: This is proposition 5.1.1 in [**91**]. ⋄

Using the diffeomorphism Ψ we now consider a special volume form η on X, defined as $\eta = \mathrm{d}t \wedge \eta_M$, where η_M is the volume form on M (see the definition of a Cauchy surface in section 3.1.1). In the sequel we will always assume that X is oriented in terms of this volume form. Note that, if $\tilde{\eta}$ is the one-form induced by η using (1.23), then this particular choice for η implies that

$$\tilde{\eta} = \tilde{\pi}_1^* \mathrm{d}t. \tag{1.24}$$

The main consequence of these restrictions is that there exists a diffeomorphism between $\tilde{Z}$ and the jet bundle $J^1\tilde{\pi}$. This was not the case for the space of Cauchy data as defined in the previous section.

Theorem 3.8. *There exists a bijection between* $\Gamma(\pi)$ *and* $\Gamma(\tilde{\pi})$*. As a result, there also exists a bijective correspondence between* $\tilde{Z}$ *and* $J^1\tilde{\pi}$*.*

Proof: See [**91**, par. 5.2.]. The one-to-one correspondence between sections ϕ of π and sections φ of $\tilde{\pi} : \tilde{Y} \to \tilde{X}$ is determined by

$$\phi(x) = \varphi(\tau)(u), \quad \text{where } x = \Psi(t, u) \text{ and } \tau(\cdot) = \Psi(t, \cdot).$$

The bijection between $\tilde{Z}$ and $J^1\tilde{\pi}$ is then defined by mapping $\kappa \in \tilde{Z}$ to the one-jet $j^1_\tau \varphi \in J^1\tilde{\pi}$. Here, we have written κ as $j^1\phi \circ \tau$, and φ is the section of $\tilde{\pi}$ associated to ϕ. The proof that this is a well-defined bijection can be found in [**91**, par. 5.2.]. ⋄

We will use the bijection between $\Gamma(\pi)$ and $\Gamma(\tilde{\pi})$ implicitly, but we will always stick to the notation "ϕ" for a section of π and "φ" for the corresponding section of $\tilde{\pi}$.

3.2. The dynamics on the space of Cauchy data. In the previous section, we have reduced the spaces of Cauchy and Dirichlet data to a form which resembles the bundle setup in time-dependent mechanics, with a configuration space fibered over the reals. We now show that the specification of a Lagrangian on $J^1\pi$, and the associated geometric objects, induce a system of ODEs on the space of Cauchy data, which are formally identical to the Euler-Lagrange equations in time-dependent mechanics.

Let L be a regular Lagrangian on $J^1\pi$ and consider the Poincaré-Cartan form Ω_L. This $(n+2)$-form induces a induces a 2-form $\tilde{\Omega}_L$ on the space $\tilde{Z}$ of Cauchy data according to (1.23). Likewise, the volume form η induces a one-form $\tilde{\eta}$ on $\tilde{Z}$. One can prove that both $\tilde{\Omega}_L$ and $\tilde{\eta}$ are closed forms and, in particular, it turns out that $\tilde{\Omega}_L = -\mathrm{d}\tilde{\Theta}_L$, where $\tilde{\Theta}_L$ is the one-form on $\tilde{Z}$ induced by the Cartan $(n+1)$-form Θ_L following the prescription (1.23). See [**91**] for more details.

In his thesis [**91**, section 5.2.3], Santamaría showed furthermore that the space of Cauchy data $\tilde{Z}$ can be equipped with a vertical endomorphism $\tilde{S}_{\tilde{\eta}}$ induced by the corresponding object S_η on $J^1\pi$. In the case under consideration, the base space $\tilde{X}$ is 1-dimensional, and $\tilde{S}_{\tilde{\eta}}$ therefore is a vector valued one-form that takes a similar form as in time-dependent mechanics. It is constructed as follows: take any $\kappa \in \tilde{Z}$, with $\kappa = j^1\phi \circ \tau$, where τ is an element of $\tilde{X}$ and ϕ a section of π. In view of theorem 3.8, we write κ as $j^1_\tau\varphi$. For arbitrary $W_\kappa \in T_\kappa\tilde{Z}$, we then put

$$\tilde{S}_{\tilde{\eta}}(W_\kappa) = \left(T_{j^1_\tau\varphi}\tilde{\pi}_{1,0}(W_\kappa) - T_\tau\varphi \circ T_{j^1_\tau\varphi}\tilde{\pi}_1(W_\kappa)\right)^v , \tag{1.25}$$

where the superscript 'v' denotes the natural vertical lift operation from $T\tilde{Y}$ to $V\tilde{\pi}_{1,0}$. We deliberately ignore the precise definition of this operation: the only fact that will be needed below (especially in the proof of proposition 3.9) is that it is a linear bundle map.

In accordance with the established terminology in time-dependent mechanics, we say that a vector field Γ on $\tilde{Z}$ is *a second-order vector field* (or a *SODE* for short) if

$$\tilde{S}_{\tilde{\eta}}(\Gamma) = 0 \quad \text{and} \quad i_\Gamma\tilde{\eta} = 1\,. \tag{1.26}$$

Consider now a connection Υ on $\pi_1 : J^1\pi \to X$, with horizontal projector $\mathbf{h}$. One can then construct a vector field Γ on $\tilde{Z}$ as follows. For $\kappa \in \tilde{Z}$, with $\kappa = j^1\phi \circ \tau$, define the vector $\Gamma(\kappa) \in T_\kappa\tilde{Z}$ by

$$\Gamma(\kappa)(u) = \mathbf{h}\left(T_{\tau(u)}j^1\phi(\mathbf{T}(\tau)(u))\right), \tag{1.27}$$

i.e. $\Gamma(\kappa)(u) \in T_{\kappa(u)}J^1\pi$ is the horizontal lift of $\mathbf{T}(\tau)(u) \in T_{\tau(u)}X$ under the given connection Υ. We then have the following interesting property.

Proposition 3.9. *If Υ is a semi-holonomic connection on π_1, then the vector field Γ on $\tilde{Z}$, defined by (1.27), is a second-order vector field.*

Proof: For the contraction of Γ with $\tilde{\eta}$ we find that

$$(i_\Gamma\tilde{\eta})(\kappa) = \int_M \kappa^*(i_{\Gamma(\kappa)}\eta) = \int_M \tau^*(i_{\mathbf{T}(\tau)}\eta) = 1\,,$$

where the last equality follows from the normalization assumption (1.21) and for the second equality we have used the fact that (with previous conventions) $i_{\Gamma(\kappa)}\eta = \pi_1^*\left(i_{\mathbf{T}(\tau)}\eta\right)$ and $\pi_1 \circ \kappa = \tau$. Hence, we have shown that Γ verifies the second condition of (1.26).

Next, we investigate the first condition of (1.26). Since the given connection Υ is semi-holonomic, it is easily checked in coordinates that $\mathbf{h}$ satisfies

$$T_\gamma\pi_{1,0}(\mathbf{h}(v_\gamma)) = T_\gamma(\phi \circ \pi_1)(v_\gamma), \tag{1.28}$$

where $\gamma = j^1_x\phi$ and $v_\gamma \in T_\gamma J^1\pi$. We now compute $\tilde{S}_{\tilde{\eta}}(\Gamma(\kappa))$. With $W_\kappa = \Gamma(\kappa)$, the first term on the right-hand side of (1.25) becomes

$$\begin{aligned} T_\kappa\tilde{\pi}_{1,0}(\Gamma(\kappa))(u) &= T_{\kappa(u)}\pi_{1,0}(\Gamma(\kappa)(u)) \\ &= T_{\kappa(u)}\pi_{1,0}(\mathbf{h}(K_{\kappa(u)})), \end{aligned}$$

where $\kappa = j^1_\tau\varphi$ and where, for notational convenience, we have written $Tj^1\phi(\mathbf{T}(\tau)(u))$ as $K_{\kappa(u)}$. Using property (1.28), we further obtain

$$\begin{aligned} T_\kappa\tilde{\pi}_{1,0}(\Gamma(\kappa))(u) &= T_{\kappa(u)}(\phi \circ \pi_1)(K_{\kappa(u)}) \\ &= T_{\kappa(u)}\phi(\mathbf{T}(\tau)(u)), \end{aligned}$$

so that

$$T_\kappa\tilde{\pi}_{1,0}(\Gamma(\kappa)) = T_\tau\varphi(\mathbf{T}(\tau)),$$

from which it follows that $\tilde{S}_{\tilde{\eta}}(\Gamma(\kappa)) = 0$, which completes the proof that Γ defines a second-order ODE. $\diamond$

We are now in a position to state the main theorem: the vector field Γ, induced by a solution $\mathbf{h}$ of the De Donder-Weyl equation, satisfies the equations from time-dependent mechanics on $J^1\tilde{\pi}$.

Theorem 3.10. *If* $\mathbf{h}$ *satisfies the De Donder-Weyl equation (1.17), then the vector field* Γ *on* $\tilde{Z}$*, defined by (1.27), satisfies the equations*

$$i_\Gamma\tilde{\Omega}_L = 0 \quad \text{and} \quad i_\Gamma\tilde{\eta} = 1\,.$$

Proof: See [**91**, chapter 5]. $\diamond$

Remark 3.11. Using the integration of forms (1.23), one can find a more manageable form for the vertical endomorphism $\tilde{S}$ in (1.25):

$$\left\langle \alpha, \tilde{S}(X_\kappa) \right\rangle \equiv \left\langle \tilde{S}^*(\alpha), X_\kappa \right\rangle = \int_M \kappa^*(X_\kappa \lrcorner\, S^*_\eta(\alpha)). \tag{1.29}$$

This is easily verified by a coordinate calculation, using the coordinate expressions for $\tilde{S}$ in [**91**]. $\diamond$

4. Elasticity as a multisymplectic field theory

Following Marsden *et al.* [**81**], we will show in this section that the classical theory of elastodynamics can be interpreted as a multisymplectic field theory. We derive the equations of motion using the jet-bundle approach, and we make the link between the Cauchy formulation and the conventional geometric formulation of elasticity.

Let (M, G) be an n-dimensional orientable Riemannian manifold with metric G. We refer to M as the *reference configuration*; the points of M label the points of the abstract continuum. Secondly, let (S, g) be a Riemannian manifold of dimension m with metric g: S is the physical space in which the body moves. In most cases, S is just the Euclidean space $\mathbb{R}^3$, or a subset thereof, but this is by no means necessary. Coordinates on M will be denoted by x^i $(i = 1, \ldots, n)$ and those on S by y^a $(a = 1, \ldots, m)$. We denote the space of embeddings of M into S by $C^\infty(M, S)$. Just as the spaces of Cauchy data in section 3, $C^\infty(M,S)$ can be given the structure of a smooth infinite-dimensional manifold.

Note that the dimensions of M and S do not have to agree. For instance, in the important case of Cosserat rods, M will be one-dimensional, and S will be three-dimensional.

A *configuration* of a continuum is an embedding of M into S. Physically, such a configuration assigns to each point of M (*i.e.* each point of the abstract continuum) its location in S. A *motion* of a continuum is an assignment of a configuration to each time t in an interval $]a,b[\subset \mathbb{R}$, *i.e.* a curve in the space of embeddings $C^\infty(M,S)$. Elastodynamics thus reduces to the study of mechanical systems on the infinite-dimensional configuration space $C^\infty(M,S)$. This point of view was taken by many authors; see, among others, [**44**, **76**, **96**].

Remark 4.1. In elasticity, coordinates on M are usually denoted as X^I, $I = 1,\ldots,n$, and coordinates on S as x^i, $i = 1,\ldots,m$. It is also common to denote the coordinates on $J^1\pi$ by $(t, X^I, x^i, v^i, F^i_I)$. We will not use this convention here. $\diamond$

4.1. Covariant field theory. Consider the manifolds $X := \mathbb{R}\times M$ and $Y := X \times S$ and let $\pi : Y \to X$ be the natural projection given by $\pi(t,x,y) := (t,x)$. A motion of a continuum induces a section of π: let $\{\varphi_t\}$ be a family of configurations; the map $\phi : (t,x) \mapsto (t,x,\varphi_t(x))$ is then a section of π. The converse is not immediately true: a section of π induces a family of mappings $f_t : M \to S$, but these mappings need not be embeddings. We will say that f_t is *regular* if it is an embedding.

Due to its special structure (the triviality of π together with the fact that X is a product manifold), the jet bundle $J^1\pi$ can be written as a product of more elementary constituents. Roughly speaking, we separate each jet $\gamma \in J^1\pi$ into a part involving spatial derivatives, and a part involving the time derivative. This is a special case of the space + time decomposition of the jet bundle in Cauchy analysis.[2]

Lemma 4.2. *The first jet bundle $J^1\pi$ is isomorphic, as an affine bundle over $Y = \mathbb{R}\times M \times S$, to $\mathbb{R}\times[J^1(M,S)\times_S TS]$. Here, the bundle $\mathbb{R}\times[J^1(M,S)\times_S TS]$ consists of triples (t,κ,v) such that $\pi_S(\kappa) = \tau_S(v)$, where $\tau_S : TS \to S$ is the tangent bundle projection, and $\pi_S : J^1(M,S) \to S$ was defined in remark 1.2.*

Proof: Take any point (t,m,s) in $\mathbb{R}\times M\times S$ and consider a 1-jet $\gamma \in J^1\pi$ such that $\pi_{1,0}(\gamma) = (t,m,s)$. An alternative interpretation of γ is that of a linear map $\gamma : T_{(t,m)}(\mathbb{R}\times M) \to T_sS$. Consider now the map $\Psi_{(t,m,s)}$, mapping γ to the element of

[2] This aspect of Cauchy analysis was not treated in section 3, but can be found, for example, in [**49**, paragraph 6B].

$\mathbb{R} \times [J^1(M,S) \times_S TS]$ given by

$$\Psi_{(t,m,s)}(\gamma) = \left(t, \gamma(0_t, \cdot), \gamma\left(\left.\frac{\partial}{\partial t}\right|_t, 0_m\right)\right), \tag{1.30}$$

where 0_m and 0_t are the zero vectors in T_mM and in $T_t\mathbb{R}$, respectively. Here, we have identified $T_{(t,m)}(\mathbb{R} \times M)$ with $T_t\mathbb{R} \times T_mM$, and $\gamma(0_t, \cdot) : T_mM \to T_sS$ is the linear map obtained by restricting γ to $\{0_t\} \times T_mM$. Note that $\gamma(0_t, \cdot) \in J^1(M,S)$.

Let us now construct the map $\Phi : J^1\pi \to \mathbb{R} \times [J^1(M,S) \times_S TS]$ as follows:

$$\Phi : \gamma \mapsto \Phi(\gamma) := \Psi_{\pi_{1,0}(\gamma)}(\gamma).$$

It is easy to check that Φ is an isomorphism of affine bundles. ◇

In natural bundle coordinates $(y^a, \dot{y}^a)$ on TS and $(x^i, y^a; y^a_{;i})$ on $J^1(M,S)$, the isomorphism of lemma 4.2 is given by $(x^\mu, y^a; y^a_\mu) \mapsto (t; x^i, y^a, y^a_{;i} = y^a_i; \dot{y}^a = y^a_0)$.

Assume that a mass density $\rho : M \to \mathbb{R}$ is given. A suitable Lagrangian for continuum mechanics is then given by

$$L(\gamma) = \frac{1}{2}\sqrt{\det[G]}\rho(x)g(v,v) - \sqrt{\det[G]}\rho(x)W(x, G(x), g(y), F), \tag{1.31}$$

(compare with [**81**, expr. 2.3]), where $v \in TS$ and $F \in J^1(M,S)$ are determined by $(t, v, F) = \Phi(\gamma)$, where Φ is the isomorphism introduced in lemma 4.2. The function W in (1.31) is the *stored energy density*, which depends only on the spatial derivatives of the field, represented by F.

The field equations associated to (1.31) are the following: (see [**81**, eq. 2.13])

$$\rho g_{ab}\left(\frac{D\dot{\phi}}{Dt}\right)^b - \frac{1}{\sqrt{\det[G]}}\frac{\partial}{\partial x^i}\left(\rho\frac{\partial W}{\partial y^a_i}(j^1\phi)\sqrt{\det[G]}\right) = -\rho\frac{\partial W}{\partial g_{bc}}\frac{\partial g_{bc}}{\partial y^a}(j^1\phi), \tag{1.32}$$

where D denotes covariant differentiation with respect to g. In the case where M and S are Euclidian, these equations reduce to the well-known equations from continuum mechanics: $\ddot{y}^a - \partial_{x^i}\sigma^i_a = 0$, where σ^i_a is the Piola-Kirchhoff stress tensor (see [**76**]):

$$\sigma^i_a = \frac{\partial W}{\partial y^a_i}.$$

These equations should be supplemented by the balance law of angular momentum, which is a consequence of the fact that the Lagrangian (1.31) is invariant under orthogonal transformations (see [**69**, **70**]).

Remark 4.3. Fluid dynamics can be treated as a special sub-case of the general theory outlined above. In the case of a fluid moving in a fixed container M, the base space of π is given by $\mathbb{R} \times M$ and the fibre is M. The Lagrangian is still given by (1.31) but now the metric g coincides with G. ◇

4.2. The infinite-dimensional setting. Having described above a covariant geometric framework for elasticity, we now turn to Cauchy analysis. Our aim in this section is to show that $\tilde{Y}$ is in this case just $\mathbb{R}\times C^\infty(M,S)$, and that $\tilde{Z}$ is $\mathbb{R}\times TC^\infty(M,S)$. Hence, one can expect a close relationship between the Cauchy formalism and the traditional geometric approach, where the configuration space is also $C^\infty(M,S)$.

Note that $X=\mathbb{R}\times M$ is naturally equipped with a foliation of hypersurfaces of constant time (equivalently, the natural splitting of X is just the identity). Therefore, we expect the Cauchy analysis to have less of an arbitrary character than in section 3.

Remark 4.4. Throughout the remainder of this chapter, we denote the space $C^\infty(M,S)$ of embeddings of M into S by $\mathcal{Q}$. ◇

It is easy to see that the space of Dirichlet data $\tilde{Y}$ is diffeomorphic to $\mathbb{R}\times\mathcal{Q}$. On the other hand, we also have the following proposition, where $\tilde{Z}$, the space of Cauchy data, again consists of embeddings $\kappa: M\hookrightarrow J^1\pi$ such that there exists a fixed $t\in\mathbb{R}$ and a section ϕ of π for which $\kappa(m)=j^1\phi(t,m)$, for all $m\in M$.

Proposition 4.5. *The space of Cauchy data $\tilde{Z}$ is diffeomorphic to $\mathbb{R}\times T\mathcal{Q}$.*

Proof: Consider an element κ of $\tilde{Z}$ and for a fixed $t\in\mathbb{R}$, let ϕ be a section of π such that $\kappa(m)=j^1\phi(t,m)$ for all $m\in M$. The section ϕ induces an element $\delta:=\phi(t,\cdot)$ of $\mathcal{Q}$, and, because of lemma 4.2, κ induces a map $X_\delta: M\to TS$ along δ. By assigning to each κ the corresponding pair (t,X_δ), we obtain a map Θ from $\tilde{Z}$ to $\mathbb{R}\times T\mathcal{Q}$.

Conversely, let (t,δ) be an element of $\mathbb{R}\times\mathcal{Q}$, and consider an element X_δ of $T_\delta\mathcal{Q}$. Consider a curve $\epsilon\mapsto\delta_\epsilon\in\mathcal{Q}$, defined in a neighbourhood of 0, and such that $\delta_0=\delta$ and $\dot{\delta}_0=X_\delta$. Let ϕ be the local section of π, defined on an open neighbourhood of $\{t\}\times M$ in $\mathbb{R}\times M$ by the following prescription:

$$\phi:(t',m)\mapsto(t',m,\delta_{t'-t}(m)).$$

Now, let $\kappa: M\hookrightarrow J^1\pi$ be the Cauchy map given by $\kappa(m)=j^1\phi(t,m)$. Note that κ does not depend on the actual curve δ_ϵ, but only on the tangent vector X_δ. Indeed, the components of $j^1\phi(t,m)$ are $(t,m,\delta(m),\phi^a_\mu(m,t))$, where

$$\phi^a_0(t,m)=\frac{\mathrm{d}}{\mathrm{d}t'}\delta_{t'-t}(m)\Big|_{t'=t}=X_\delta(m),\quad\text{and}\quad\phi^a_i(t,m)=\frac{\partial\delta^a}{\partial x^i}(m).$$

Finally, we consider the map $\Xi:\mathbb{R}\times T\mathcal{Q}\to\tilde{Z}$ taking (t,X_δ) to κ. It is easy to check that Θ and Ξ are each other's inverse. ◇

For future reference, we also mention that $T\mathcal{Q}$ is equipped with the usual geometric objects known from tangent bundle geometry, such as a vertical lift operation, a vertical endomorphism, and a Liouville vector field. Because of the special nature of $\mathcal{Q}$ as a manifold of embeddings, these objects are induced by their finite-dimensional counterparts on TS. This provides us with a convenient way of avoiding any functional-theoretic aspects that would arise in a direct definition.

Definition 4.6. *The Liouville vector field* Δ *on* TS *induces a vector field* $\tilde{\Delta}$ *on* $T\mathcal{Q}$ *by composition:* $\tilde{\Delta}(X) = \Delta \circ X$ *for all* X *in* $T\mathcal{Q}$*. Similarly, the vertical endomorphism* S *on* TS *induces a* $(1,1)$*-tensor* $\tilde{S}$ *on* $T\mathcal{Q}$ *by composition:* $\tilde{S}(X) = S \circ X$*. We will refer to* $\tilde{\Delta}$ *as the Liouville vector field, and to* $\tilde{S}$ *as the vertical endomorphism on* $T\mathcal{Q}$*.*

By putting $\hat{S} = \tilde{S} - \tilde{\Delta} \otimes \mathrm{d}t$, we obtain a vertical endomorphism on $\mathbb{R} \times T\mathcal{Q} = \tilde{Z}$.

The model introduced in this section will be used as an illustrative example throughout our discussion on nonholonomic field theories (starting from chapter 6). The Cauchy analysis of this kind of field theory (where the base space can be written canonically as $\mathbb{R} \times M$) will allow us to distinguish a special class of constraints in chapter 8, and will also form the basis of the theory in chapter 9.

Chapter 2
Variational integrators

As we mentioned in the introduction, it is often useful and desirable to use numerical methods to gain an insight into the dynamics of a mechanical system. However, traditional numerical integration schemes often do not respect the geometric background of the equations of motion and therefore lead to relatively inaccurate results, at least in the long run.

In this chapter, we report on a different class of integration schemes, designed to remedy this defect. In particular, they are derived by use of a discrete variational principle on a suitable space. Consequently, they have a number of interesting properties, the most important of which is that the discrete flow is symplectic with respect to a natural symplectic form. The purpose of this chapter is twofold: to serve as an introduction for the theory of discrete field theories in chapter 3, and to give a quick overview of geometric integration methods, which will be used in chapter 9.

The construction of discrete mechanical systems is the subject of section 1. In section 2, we then turn our attention to geometric integration of field theories. While the last few years have seen substantial advances in this particular area, an extensive background theory for this kind of discrete fields is lacking. Most of our investigations will therefore proceed by analogy to the case of mechanics.

The presentation in section 1 is inspired by the survey paper [**79**] of Marsden and West, and also by [**52**, chapter VI.6]. For more information on symplectic integrators, see also [**68**]. A good overview of multisymplectic methods can be found in [**18**].

1. Geometric integration of mechanical systems

In this section, we recall some basic elements from discrete mechanics, and we give an overview of how they may be used in the construction of geometric integrators.

1.1. Discrete mechanics. We start with a quick overview of discrete mechanics. The tangent bundle TQ of a configuration space Q is discretized by considering the Cartesian product $Q \times Q$. The idea is that a suitable approximation to a tangent vector v_q is obtained by considering a pair (q, q'), with q' close to q such that q and q' belong to the same chart (U, ψ), such that $(\psi(q') - \psi(q))/h$ (where $h \sim d(q, q')$) provides a "good" approximation to $T_q\psi(v_q)$. Of course, the idea of what constitutes a good approximation is model dependent, but it is clear that the Cartesian product $Q \times Q$ plays a fundamental role in discretizing TQ.

A discrete dynamical system is a diffeomorphism $\Upsilon : Q \times Q \to Q \times Q$ of the form $\Upsilon(q_0, q_1) = (q_1, q_2)$. A discrete Lagrangian is a function L_d on $Q \times Q$. From now on, we will assume that an appropriate discrete Lagrangian L_d is given; its construction will be dealt with later. One can then consider the following discrete action:

$$S(q_0, q_1, \ldots, q_N) = \sum_{i=0}^{N-1} L_d(q_i, q_{i+1}). \tag{2.1}$$

As in the case of continuous mechanical systems, we are interested in discrete trajectories, *i.e.* sequences $\{q_0, q_1, \ldots, q_N\}$, that extremize this action with respect to appropriate variations. These extremal sequences are characterised in theorem 1.2.

Definition 1.1. *Let $Q^{\times k}$ denote the k-fold Cartesian product of Q with itself. If f is a function on $Q^{\times k}$, we define the i*th differential of f, *denoted by $D_i f : Q^{\times k} \to T^*Q$, as*

$$D_i f(q_1, q_2, \ldots, q_k) = \mathrm{d}[f(q_1, \ldots, q_{i-1}, \cdot, q_{i+1}, \ldots, q_k)]_{q_i} \in T^*_{q_i}Q.$$

Note that $D_i f$ is a one-form along the projection pr_i onto the ith factor.

Theorem 1.2. *A sequence $\{q_0, q_1, \ldots, q_N\}$ is an extremum of (2.1) under arbitrary variations that keep the end points fixed if and only if it satisfies the following set of* discrete Euler-Lagrange equations*:*

$$D_1 L_d(q_k, q_{k+1}) + D_2 L_d(q_{k-1}, q_k) = 0, \quad \text{for } k = 1, \ldots, N-1. \tag{2.2}$$

Let $\mathcal{H}$ be the matrix with entries $\mathcal{H}_{ij}$ given by

$$\mathcal{H}_{ij} = \frac{\partial^2 L_d}{\partial q_0^i \partial q_1^j}.$$

If the Lagrangian is regular, in the sense that $\mathcal{H}$ is invertible, the implicit equations (2.2) can be reformulated as an explicit map $\Upsilon : (q_{k-1}, q_k) \mapsto (q_k, q_{k+1})$, where the triple (q_{k-1}, q_k, q_{k+1}) satisfies (2.2).

Associated to a discrete Lagrangian are two *discrete Legendre transformations*, denoted by $\mathbb{F}^+ L_d, \mathbb{F}^- L_d : Q \times Q \to T^*Q$, and defined as $\mathbb{F}^+ L_d = D_1 L_d$ and $\mathbb{F}^- L_d = -D_2 L_d$. These maps may be used to pull back the canonical symplectic potential Θ on T^*Q to $Q \times Q$, which gives $\Theta^+_{L_d} = (\mathbb{F}^+ L_d)^*\Theta$ and $\Theta^-_{L_d} = (\mathbb{F}^- L_d)^*\Theta$. Moreover, these one-forms satisfy $\Theta^+_{L_d} - \Theta^-_{L_d} = \mathrm{d}L_d$, and a symplectic form Ω_{L_d} can hence be defined by $\Omega_{L_d} = \mathrm{d}\Theta^+_{L_d} = \mathrm{d}\Theta^-_{L_d}$. In analogy with the continuous case, $\Theta^+_{L_d}$ and $\Theta^-_{L_d}$ are called the *(discrete) Cartan forms*, and Ω_{L_d} is called the *(discrete) Poincaré-Cartan form.*

Remark 1.3. A well-known property of the canonical symplectic form on T^*Q is that $\alpha^*\Theta = \alpha$ for all one-forms α on Q. By using a slight extension of this result to forms along the projection, we see that $\Theta^+_{L_d}$, resp. $\Theta^-_{L_d}$ can be identified in a natural way with $\mathbb{F}^+L_d$ and $-\mathbb{F}^-L_d$. This is the discrete counterpart of the well-known result from the continuum case that the Poincaré-Cartan form and the Legendre transformation are two aspects of the same object (see for instance [**92**]). ⋄

Theorem 1.4. *Let $\Upsilon : Q \times Q \to Q \times Q$ be a solution of the discrete Euler-Lagrange equations (2.2). Then Υ is a symplectic mapping with respect to the symplectic form Ω_{L_d}, i.e. $\Upsilon^*\Omega_{L_d} = \Omega_{L_d}$.*

Until now, the discrete Lagrangian was simply assumed to be given. However, if we know a Lagrangian $L : TQ \to \mathbb{R}$ for the corresponding continuous problem, a natural discrete Lagrangian suggests itself.

Let there be given a regular Lagrangian L on TQ. Associated to L is the so-called exact discrete Lagrangian L_E, defined on a neighbourhood of the diagonal in $Q \times Q$, and given explicitly by

$$L_E(q_0, q_1; h) = \int_0^h L(q(t), \dot{q}(t))\mathrm{d}t, \tag{2.3}$$

where $t \mapsto q(t)$ is the unique solution of the Euler-Lagrange equations such that $q(0) = q_0$ and $q(h) = q_1$. It can be proved that such a solution exists for q_0, q_1 nearby points and h sufficiently small (see [**82**] for a proof). The importance of the exact discrete Lagrangian lies in the following theorem. In a nutshell, this theorem states that the solution trajectories of the discrete Euler-Lagrange equations follow exactly the continuous solutions.

Theorem 1.5. *Consider a pair (q_0, q_1) in $Q \times Q$ and let $q(t)$ be a solution of the Euler-Lagrange equations (defined for $t \in [0, 2h]$) such that $q(0) = q_0$ and $q(h) = q_1$. Let q_2 be the point determined in terms of q_0 and q_1 by the discrete Euler-Lagrange equations (2.2) associated to the exact discrete Lagrangian L_E. Then $q(2h) = q_2$.*

Proof: This is a weaker version of theorem 1.6.4 in [**79**]. ⋄

Example 1.6. Let Q be equipped with a Riemannian metric g and consider the kinetic energy Lagrangian $L(v) = 1/2g(v, v)$. Between any two points q_0 and q_1 that are sufficiently close, there exists a unique length minimizing geodesic; the exact discrete Lagrangian is then the Riemannian distance: $L_E(q_0, q_1; h) = d(q_0, q_1)$. Note that unless (Q, g) is geodesically complete, L_E is only defined on a neighbourhood of the diagonal in $Q \times Q$, in contrast to the Riemannian distance d, which is defined everywhere. ⋄

In practice, the computation of the exact discrete Lagrangian (2.3) requires the solution of the Euler-Lagrange equations. One therefore resorts to using an approximation of

L_E, which leads to discrete trajectories that approximate the exact flow. To make these claims more rigorous, a number of definitions are needed.

Definition 1.7. *A discrete Lagrangian* $L_d : Q \times Q \to \mathbb{R}$ *is a* kth-order discrete approximation of L_E *if there exists an open neighbourhood* U *of the diagonal in* $Q \times Q$*, and a constant* C *such that* $\|L_d(q_0, q_1) - L_E(q_0, q_1; h)\| \leq Ch^{k+1}$ *for all* $(q_0, q_1) \in U$*.*

Definition 1.8. *A discrete flow* $\Upsilon : Q \times Q \to Q \times Q$ *is a* kth-order discrete flow *if there exists a constant* C *and a neighbourhood* U *of the diagonal in* $Q \times Q$ *such that* $\|q_2 - q(2h)\| \leq Ch^{k+1}$ *for all* $(q_0, q_1) \in U$*. Here,* $q(t)$ *is the exact flow such that* $q(0) = q_0$ *and* $q(h) = q_1$*, and* q_2 *is determined by* $\Upsilon(q_0, q_1) = (q_1, q_2)$*.*

One can then prove that a kth-order approximation of the exact discrete Lagrangian induces a kth-order discrete flow. In practice, one usually employs second-order approximations. As an illustration, we show how the midpoint rule for the numerical evaluation of definite integrals can be used to construct a class of second-order discrete Lagrangians.

Let us assume that the configuration space Q is a vector space, or a convex subset thereof. Applying the midpoint rule to the integral in the definition of L_E gives

$$L_E(q_0, q_1; h) = hL(q(h/2), \dot{q}(h/2)) + \mathcal{O}(h^3) = hL\left(\frac{q_0 + q_1}{2}, \frac{q_1 - q_0}{h}\right) + \mathcal{O}(h^3),$$

where, in the second step, we have expanded $q(h/2)$ and $\dot{q}(h/2)$ in a Taylor series. The approximation L_d, defined as

$$L_d(q_0, q_1) = hL\left(\frac{q_0 + q_1}{2}, \frac{q_1 - q_0}{h}\right), \tag{2.4}$$

is therefore a second-order approximation, and the resulting discrete flow will also be of second order. Of course, other discretizations can be obtained by using different quadrature formulas, such as the trapezium rule, for the exact discrete Lagrangian.

One can think of the discrete Lagrangian L_d in (2.4) as being obtained by pulling back the continuous Lagrangian L by the *discretization mapping* $\Phi : Q \times Q \to TQ$, defined as

$$\Phi(q_0, q_1) = \left(\frac{q_0 + q_1}{2}, \frac{q_1 - q_0}{h}\right).$$

Example 1.9. Consider the following Lagrangian $L = v^2/2 - V(q)$, describing the motion of a particle of unit mass under the influence of a potential V. An associated second-order discrete Lagrangian is given by

$$L_d = \frac{h}{2}\left(\frac{q_1 - q_0}{h}\right)^2 - hV\left(\frac{q_0 + q_1}{2}\right),$$

and the resulting discrete Euler-Lagrange equations are then

$$\frac{q_{k+1} - 2q_k + q_{k-1}}{h^2} = -\frac{1}{2}\nabla V\left(\frac{q_{k-1}+q_k}{2}\right) - \frac{1}{2}\nabla V\left(\frac{q_k+q_{k+1}}{2}\right),$$

for all k. It is straightforward to check that these equations are indeed of second order. This discrete method is equivalent to the well-known *midpoint method* (see [**79**]). ◇

Remark 1.10. The Cartesian product $Q \times Q$ is a particular example of a *Lie groupoid* and it turns out that many constructions from the preceding paragraph can be extended to the case where the configuration space is a Lie groupoid. This particular insight was formulated by Weinstein [**110**], and extended by Marrero *et al.* [**75**]. ◇

1.2. Nonholonomic integrators. The framework of the previous section can be extended in order to deal with mechanical systems with linear nonholonomic constraints. Here, we follow [**28**], where a discrete d'Alembert principle was introduced.

If D is a constraint distribution, whose annihilator is spanned by the k one-forms $A^\alpha \in \Omega^1(Q)$, and if $\varphi^\alpha(q, v) = \langle v, A^\alpha(q)\rangle$ are k functions whose vanishing defines D, we can define k discrete constraint functions $\varphi^\alpha_d : Q \times Q \to \mathbb{R}$ by a similar construction as in (2.4): we put

$$\varphi^\alpha_d(q_0, q_1) = \varphi^\alpha\left(\frac{q_0+q_1}{2}, \frac{q_1-q_0}{h}\right).$$

Theorem 1.11 (see [**28**])**.** *Let L be a Lagrangian and D a constraint distribution defined as above. The algorithm defined by*

$$D_1L_d(q_n, q_{n+1}) + D_2L_d(q_{n-1}, q_n) = \lambda_\alpha A^\alpha(q_n), \tag{2.5}$$

where the multipliers λ_α are determined by the requirement that $\varphi^\alpha_d(q_n, q_{n+1}) = 0$, is second-order, symmetric, and satisfies the discrete constraints exactly.

Nonholonomic integrators were first introduced in [**28**]. Despite some theoretical investigations (see e.g. [**34**]), it is safe to say that nonholonomic integrators are not well understood yet, and there are clues that nonholonomic integrators behave fundamentally different from ordinary symplectic integrators. Even though the flow of a nonholonomic system is explicitly nonsymplectic, this result is nevertheless quite surprising. Given the previous successes in using geometric methods in nonholonomic mechanics, one would expect similar conservation results as obtained by backward error analysis in the symplectic case (see [**52**]). It seems that the contrary is true: the energy is conserved for mechanical systems with linear nonholonomic constraints, but in [**84**], a simple geometric integrator was presented that clearly shows energy diffusion.

2. Discrete multisymplectic field theory

In this section, we give a very brief overview of discrete Lagrangian field theories. The purpose of this overview is twofold: to pave the way for subsequent generalization to Lie groupoid field theories in chapter 3, as well as allowing us to construct numerical methods for nonholonomic field theories in chapter 9.

Our treatment is mainly based on [**80**], and, to a lesser extent, on [**19**]. For a recent overview of multisymplectic integrators and their properties, see [**3**, **18**].

For the remainder of this chapter, we will only consider bundles π of the form $\pi : \mathbb{R}^2 \times Q \to \mathbb{R}^2$, *i.e.* π is trivial, and the base space is $\mathbb{R}^2$. In this case, a field is just a map $\phi : \mathbb{R}^2 \to Q$. In addition, one can easily prove (see for instance [**86**]) that the jet bundle is in this case isomorphic to $\mathbb{R}^2 \times [TQ \oplus TQ]$. For reasons of simplicity, we shall also consider only Lagrangians that do not depend on the coordinates of the base space, *i.e.* Lagrangians of the form $L : TQ \oplus TQ \to \mathbb{R}$.

2.1. Discrete fields. In this section, we begin by discretizing the base space, and by defining discrete fields. Using these definitions, we then recall the definition of the discrete jet bundle, as proposed in [**80**].

Definition 2.1. *A* mesh *in* $\mathbb{R}^2$ *is a discrete subset*[1] V *of* $\mathbb{R}^2$*. A* discrete field *is a map* $\phi : V \to Q$*.*

Of course, the idea behind this definition is that a discrete field is determined by its values at certain discrete points in space-time. Note that the subset $V \subset \mathbb{R}^2$ should be specified at the outset: in the cases that we consider here, V will be just a regular lattice in $\mathbb{R}^k$, but less regular subsets are equally possible. For example, one might want to increase the density of mesh points in regions where the field is expected to vary wildly, or where other "extreme conditions" apply; see for example [**69**, **70**].

Throughout this thesis, we will mostly illustrate our results on the following quadrangular mesh:

$$V = \{x_{i,j} = (hi, kj) \in h\mathbb{Z} \times k\mathbb{Z}\}.$$

Here, the grid size is determined by the parameters h and k. A possible extension to more general meshes will be treated in the last chapter of this thesis.. Finally, we denote the value of a field $\phi : \mathbb{R}^2 \to Q$ at a point $x_{i,j}$ simply by $\phi_{i,j} := \phi(x_{i,j})$.

Remark 2.2. In remark 3.1 in chapter 1, we already hinted that most of the examples in this thesis would be of "evolution type". A similar remark can be made for this chapter, and the reader is encouraged, when confronted for example with a variable

[1]Recall that a subspace S of a topological space X is called *discrete* if, for every $x \in S$, there exists an open neighbourhood U of x in X such that $S \cap U = \{x\}$.

$x_{i,j}$, to think of i and j as a discrete time and a discrete spatial variable, respectively. ◇

For future reference, we introduce the sets of *triangles* $\mathbb{X}^3$ and *squares* $\mathbb{X}^4$ in V as follows:

$$\mathbb{X}^3 := \{(x_{i,j}, x_{i+1,j}, x_{i,j+1}) \in V^{\times 3}\} \quad \text{and} \quad \mathbb{X}^4 := \{(x_{i,j}, x_{i+1,j}, x_{i+1,j+1}, x_{i,j+1}) \in V^{\times 4}\}.$$

Note that the elements of $\mathbb{X}^3$ and $\mathbb{X}^4$ are ordered sets. We denote the elements of $\mathbb{X}^3$ by $[x]$, and for $[x] = (x_{i,j}, x_{i+1,j}, x_{i,j+1})$, we put $[x]_1 := x_{i,j}$, $[x]_2 := x_{i+1,j}$, and $[x]_3 := x_{i,j+1}$. Elements of $\mathbb{X}^4$ are also denoted as $[x]$ (it will be clear from the context to which set a generic element $[x]$ belongs), and the subscript notation $[x]_i$, $i = 1, \ldots, 4$ is defined similarly.

The main idea behind these definitions is that the values of a discrete field at the vertices of a triangle or a square can be used to define a *discrete jet*, which is an approximation to a continuous jet in the same sense that a pair (q_0, q_1) is an approximation to a tangent vector. The idea behind this definition will be explained more fully in the next section.

Definition 2.3. *A* discrete jet *is a pair* $([x], [q])$*, where* $[x]$ *is a triangle in* V *(*i.e. *an element of* $\mathbb{X}^3$*), and* $[q]$ *is an element of* $Q^{\times 3}$*. The set of all discrete jets is denoted by* $J^1_d\pi \cong \mathbb{X}^3 \times Q^{\times 3}$*.*

A similar definition exists for the case of quadrangles. In that case, a discrete jet is also a pair $([x], [q])$, but now $[x]$ is a quadrangle and $[q]$ is an element of $Q^{\times 4}$.

2.2. Discretizing the Lagrangian: the set of triangles. Recall that $J^1\pi$ is isomorphic to $\mathbb{R}^2 \times [TQ \oplus TQ]$. Inspired by the discretization mappings for TQ in section 1.1, we now propose the discretization map $\Phi_{h,k} : J^1_d\pi \to J^1\pi$ defined as

$$\Phi_{h,k}([x], [q]) = \left(\frac{x_0 + x_1 + x_2}{3}; \frac{q_0 + q_1 + q_2}{3}; \frac{q_1 - q_0}{h}, \frac{q_2 - q_0}{k} \right). \tag{2.6}$$

One should think of the last three factors in the expression on the right-hand side as being a discretization of an element in $TQ \oplus TQ$, where the second factor represents the base point, while the two last factors are discretizations of elements of TQ.

Again, such a map obviously depends on the choice of a coordinate system on Q. Moreover, we are interested only in Lagrangians that do not depend on the base space coordinates, and therefore, we will simply omit the base space coordinates in (2.6) and view $\Phi_{h,k}$ as a map from $Q^{\times 3}$ to $TQ \oplus TQ$.

Let $L : TQ \oplus TQ \to \mathbb{R}$ be a continuous Lagrangian; we then define its discrete counterpart $L_d : Q^{\times 3} \to \mathbb{R}$ by $L_d := hk\Phi^*_{h,k}L$.

Example 2.4. The Lagrangian for the wave equation is just $L = \frac{1}{2}((y_0)^2 - (y_1)^2)$. Its discretization is therefore given by

$$L_d(q_0, q_1, q_2) = \frac{hk}{2}\left(\frac{q_1 - q_0}{h}\right)^2 - \frac{hk}{2}\left(\frac{q_2 - q_0}{k}\right)^2. \tag{2.7}$$

The discrete field equations will be derived in the next section. ◇

2.3. Discretizing the Lagrangian: the set of quadrangles. The discretization (2.6) is the most straightforward possible. However, as we shall see, there are certain cases where other discretizations are more appropriate. For example, in [**80**], the authors proposed the map $\Psi_{h,k} : Q^{\times 4} \to J^1\pi$ defined as follows:

$$\Psi_{h,k}(q_0, q_1, q_2, q_3) = \left(\frac{q_0 + \cdots + q_3}{4}; \frac{q_1 - q_0}{h}, \frac{1}{2}\left(\frac{q_2 - q_1}{k} + \frac{q_3 - q_0}{k}\right)\right), \tag{2.8}$$

If $L : TQ \oplus TQ \to \mathbb{R}$ is a Lagrangian, we again define the associated discrete Lagrangian as $L_d := hk\Psi_{h,k}^* L$.

Example 2.5. Let L be the Lagrangian for the wave equation as in example 2.4. The corresponding discrete Lagrangian is then given by

$$L_d(q_0, q_1, q_2, q_3) = \frac{hk}{2}\left(\frac{q_1 - q_0}{h}\right)^2 - \frac{hk}{8}\left(\frac{q_2 - q_1}{k} + \frac{q_3 - q_0}{k}\right)^2.$$

As we shall see, the effect of using the average in (2.8) is that the resulting field equations are implicit. The advantage is that these equations will be unconditionally stable (*i.e.* for all values of h and k), in contrast with the equations derived from (2.7). ◇

2.4. The field equations. Once the concept of discrete Lagrangian is defined, the derivation of the Euler-Lagrange equations follows quite easily. For the sake of definiteness, we use here the definition of the discrete jet bundle based on the set of triangles. Hence, the discrete Lagrangian is a function on $Q^{\times 3}$. Note that all definitions can easily be extended to the case of quadrangles.

Definition 2.6 (see [**80**])**.** *A triangle* $[x]$ touches *a vertex* $y \in V$ *if* y *is a vertex of* $[x]$.

The following concepts have been taken from [**80**]. Let U_F be a finite subset of $\mathbb{X}^3$. The set U_F induces a finite subset U_V of V, where a vertex x is an element of U_V if and only there exists a triangle $[y] \in \mathbb{X}^3$ such that $[y]$ touches x. We define the *interior* of U_V, denoted by $\operatorname{int} U_V$, as the set of vertices x such that all three triangles touching x are elements of U_F. The *closure* of U_V, denoted by $\operatorname{cl} U_V$, is then defined as the set of vertices of the triangles that touch vertices in U_V. Finally, we define the *boundary* of U_V, denoted by ∂U_V, as the set of vertices $x \in U_V$ which belong also to $\operatorname{cl} U_V$ but not to $\operatorname{int} U_V$.

From now on, we restrict our attention to *regular* sets of vertices U_V, which are such that U_V is exactly the union of its interior and its boundary.

Definition 2.7. *Let $\phi : U_V \to Q$ be a discrete field restricted to U_V. An* infinitesimal variation *of ϕ is a vector field V along ϕ such that $V(x) = 0$ for all $x \in \partial U_V$. A* finite variation *of ϕ is a local one-parameter group of diffeomorphisms $\{\varphi_\epsilon\}$, defined on a neighbourhood of $\phi(U_V)$, such that φ_ϵ is the identity on $\phi(\partial U_V)$.*

The action functional S is defined in the usual way:

$$S(\phi) = \sum_{[x] \in U_F} L(\psi([x])),$$

where we have introduced the map $\psi : U_F \to Q^{\times 3}$ associated to the field ϕ as follows:

$$\psi([x]) = (\phi([x]_1), \phi([x]_2), \phi([x]_3)). \tag{2.9}$$

We now look for discrete fields ϕ such that

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} S(\varphi_\epsilon \circ \phi)\Big|_{\epsilon=0} = 0$$

for all variations φ_ϵ of ϕ. By the same reasoning as that leading to theorem 1.2, we then obtain the following set of discrete Euler-Lagrange equations:

$$D_1 L(\phi_{i,j}, \phi_{i+1,j}, \phi_{i,j+1}) + D_2 L(\phi_{i-1,j}, \phi_{i,j}, \phi_{i-1,j+1}) + D_3 L(\phi_{i,j-1}, \phi_{i+1,j-1}, \phi_{i,j}) = 0, \tag{2.10}$$

for all $(i, j) \in U_V$.

Example 2.8. For the discrete wave Lagrangian introduced in example 2.4, the discrete Euler-Lagrange equations become

$$\frac{\phi_{i+1,j} - 2\phi_{i,j} + \phi_{i-1,j}}{h^2} = \frac{\phi_{i,j+1} - 2\phi_{i,j} + \phi_{i,j-1}}{k^2},$$

the standard second-order finite difference approximation to the wave equation. We know from numerical analysis that such an explicit scheme becomes unstable if $h/k > 1$: this is the famous CFL bound (see [**1**]). In order to avoid the CFL bound, implicit schemes are needed. To this end, the authors of [**80**] introduced the four-point discretization (2.8). The Euler-Lagrange equations associated to the discrete Lagrangian in example 2.5 can be derived in a similar way as above, and become

$$\frac{\phi_{i+1,j} - 2\phi_{i,j} + \phi_{i-1,j}}{h^2} =$$
$$\frac{1}{4}\left(\frac{\phi_{i-1,j+1} - 2\phi_{i-1,j} + \phi_{i-1,j-1}}{k^2} + 2\frac{\phi_{i,j+1} - 2\phi_{i,j} + \phi_{i,j-1}}{k^2} + \frac{\phi_{i+1,j+1} - 2\phi_{i+1,j} + \phi_{i+1,j-1}}{k^2}\right).$$

These equations are implicit, and hence present a new level of complication. The advantage is that they are unconditionally stable. ◇

2.5. The Poincaré-Cartan forms. For the purpose of this section, we assume that a discrete Lagrangian is a function L_d on $Q^{\times k}$. So far, we have more or less distinguished between $k = 3$ and $k = 4$, but only to emphasize the different rationales behind both approaches. If we compare the final results, we may conclude that the particular value of k does not have any profound consequences. Therefore, we will no longer make the explicit distinction between both cases. This also underlines the fact that almost everything in this chapter carries over without change to the case of an irregular mesh.

Associated to L_d are k *Poincaré-Cartan* 1*-forms* $\theta^{(i)}_{L_d} : Q^{\times k} \to T^*Q$ defined by

$$\theta^{(i)}_{L_d}(q_1, \ldots, q_k) := D_i L_d(q_1, \ldots, q_k) \in T^*_{q_i}Q. \tag{2.11}$$

Compare with the Poincaré-Cartan forms defined in section 1.1. It is straightforward to see that $\theta^{(1)}_{L_d} + \cdots + \theta^{(k)}_{L_d} = \mathrm{d}L_d$. Furthermore, we define $\Omega^{(i)}_{L_d} := \mathrm{d}\theta^{(i)}_{L_d}$, for $i = 1, \ldots, k$.

The k Poincare-Cartan forms share many properties with their counterparts $\theta^{+,-}_L$ from mechanics; we will return to this once we treat their generalization to the context of Lie groupoids in chapter 3. The immediate purpose of defining the Poincaré-Cartan forms here is that they are needed for the multisymplectic form formula in the next section.

Remark 2.9. In some ways, the discretization of Lagrangian field theory using k Poincaré-Cartan forms is not satisfactory. Indeed, our approach started by taking a discretization of the base space, while leaving the fibres continuous. By enforcing this distinction between the base space and the fibres, we end up with a discrete theory which resembles more the so-called k-symplectic approach to Lagrangian field theory (see [**50**, **86**]), which uses only the geometry of the standard fibre, while neglecting the base space. In the k-symplectic approach, the dynamics is formulated using k presymplectic 2-forms rather than one multisymplectic form, and the same holds for the dynamics of the discrete field theories in this chapter (see (2.11)). However, the similarities do not end here. In the k-symplectic approach, the fields take value in the following manifold:

$$T^{(k)}Q := TQ \oplus \cdots \oplus TQ,$$

instead of the jet bundle. By replacing each factor TQ by its Moser-Veselov discretization $Q \times Q$, one can see that $T^{(k)}Q$ is naturally discretized by the k-fold product $Q^{\times k}$. For $k = 3, 4$, this is the space of triangles and quadrangles, respectively.

Currently, there appears to be no discretization of multisymplectic Lagrangian field theory which does not rely implicitly or explicitly on k-symplectic theory. We shall have more to say on this in the last chapter.

Note that the k-symplectic theory is also related to Bridges's multisymplectic theory (see [**17**]), which was subsequently used for the construction of geometric integrators as well (see [**19**, **68**]). The relation between Bridges's theory and conventional multisymplectic field theory was studied by Marsden and Shkoller [**78**]. ◇

2.6. Multisymplecticity. In mechanics, it is well known that the flow of a Lagrangian or a Hamiltonian vector field is symplectic with respect to an appropriate symplectic form. This fact lies at the basis of many important results in mechanics. Moreover, preserving symplecticity in a discrete mechanical system generally leads to qualitatively better results.

There is a similar, but less known result for classical field theory, known as the *multisymplectic form formula*, which embodies the conservation of multisymplecticity. In this section, following [**80**], we first recall the multisymplectic form for continuous field theories, and then we show that solutions of the discrete Euler-Lagrange equations satisfy a similar property.

Finally, we recall a different interpretation of conservation of multisymplecticity due to Bridges & Reich [**19**], and we show that their conservation law can be derived from the discrete multisymplectic form formula.

2.6.1. *Continuous field theories.* Let $\pi : Y \to X$ be a fibre bundle. We denote by $\mathcal{M}$ the set of sections $\Gamma(\pi)$ of π. The set $\mathcal{M}$ can be given the structure of a smooth manifold; see also remark 3.5 in chapter 1.

Let $L : J^1\pi \to \mathbb{R}$ be a Lagrangian and define the set $\mathcal{S}_L$ of solutions of the Euler-Lagrange equations as follows:

$$\mathcal{S}_L = \{\phi \in \mathcal{M} : (j^1\phi)^*(W \lrcorner\, \Omega_L) = 0 \quad \text{for all } W \in \mathfrak{X}(J^1\pi)\}.$$

It is well known that $\mathcal{S}_L$ is not always a true submanifold of $\mathcal{M}$. However, for the purpose of this exposition, we will assume that this is nevertheless the case. Similarly, we define a distribution $\mathcal{F}$ on $\mathcal{S}_L$ as follows:

$$\mathcal{F}(\phi) = \{V \in T_\phi\mathcal{M} : (j^1\phi)^*\mathscr{L}_{j^1V}(W \lrcorner\, \Omega_L) = 0 \quad \text{for all } W \in \mathfrak{X}(J^1\pi)\},$$

for all $\phi \in \mathcal{S}_L$. The elements of $\mathcal{F}(\phi)$ are referred to as *first variations*, as they solve the linearised form of the Euler-Lagrange equations. We now have the necessary tools to state the multisymplectic form formula:

Theorem 2.10. *If $\phi \in \mathcal{M}$ is a solution of the Euler-Lagrange equations, and $V, W \in \mathcal{F}(\phi)$ are first variations, then the following multisymplectic form formula holds:*

$$\int_U (j^1\phi)^*(j^1V \lrcorner\, j^1W \lrcorner\, \Omega_L) = 0. \tag{2.12}$$

Proof: See [**80**, thm. 4.1]. ◇

Remark 2.11. In [**80**], a more general situation is considered, namely where $\mathcal{M}$ is not just a space of sections, but a space of embeddings $\phi : U \subset X \to Y$. The advantage of this generalization is that it also allows for horizontal variations and horizontal symmetries, rather than just for vertical variations and symmetries. ◇

2.6.2. *Discrete field theories.* For the case of discrete field theories, there is a multisymplectic form formula which is similar to (2.12).

Let U_F be again a finite subset of $\mathbb{X}^k$ ($k = 3, 4$), and let $\mathcal{M}_d$ be the set of all maps $\phi : U_V \to Q$. The set $\mathcal{S}_{d,L}$ is then defined to be the subset of $\mathcal{M}_d$ whose elements are solutions of the Euler-Lagrange equations (2.10).

Inspired by the developments in the previous section, we define finite and infinitesimal first variations as follows:

Definition 2.12. *A* finite first variation *of an element ϕ of $\mathcal{S}_{d,L}$ is a local one-parameter group of diffeomorphisms $\{\varphi_\epsilon\}$, defined on a neighbourhood of $\phi(U_V)$, such that $\varphi_\epsilon \circ \phi$ is again an element of $\mathcal{S}_{d,L}$, for all ϵ.*

In other words, a finite first variation transforms solutions of the discrete Euler-Lagrange equations into new solutions. We then define a *infinitesimal first variation* of $\phi \in \mathcal{S}_{d,L}$ somewhat circularly as a vector field V along ϕ such that there exists a finite first variation φ_ϵ with the property that

$$V(x) = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\varphi_\epsilon(\phi(x))\Big|_{\epsilon=0} \quad \text{for all } x \in U_V.$$

Let φ_ϵ now be a first variation with infinitesimal counterpart V. Note that φ_ϵ is not necessarily zero on $\phi(\partial U_V)$. Varying the action with respect to such a variation gives

$$\mathrm{d}S(\phi) \cdot V = \sum_{x \in \operatorname{int} U_V} \mathcal{E}_L(x) \cdot V(x) + \sum_{[x]:[x]\cap\partial U_V \neq \varnothing} \left(\sum_{i=1}^k D_i L(\psi([x])) \cdot V(x_i) \right)$$

where x_i is the ith vertex of $[x]$, and $\mathcal{E}_L(x)$ is just a shorthand form for the left hand side of (2.10), and $\psi : \mathbb{X}^k \to Q^{\times k}$ is map associated to ϕ as in (2.9). Hence, since ϕ is a solution of the discrete Euler-Lagrange equations, $\mathrm{d}S$ can be written as

$$\mathrm{d}S(\phi) \cdot V = \sum_{[x]:[x]\cap\partial U_V \neq \varnothing} \left(\sum_{i=1}^k \theta_L^{(i)}(\psi([x])) \cdot V(x_i) \right), \tag{2.13}$$

where x_i is again the ith vertex of $[x]$.

This equation shows that the Poincaré-Cartan 1-forms arise in the same variational way as their continuous counterparts: by submitting the action to variations that do not vanish on the boundary, and inspecting the remaining terms. Expression (2.13) can also be used to prove the discrete version of the multisymplectic form formula: consider two infinitesimal first variations V_1 and V_2 of $\phi \in \mathcal{S}_{d,L}$, and use (2.13) to expand the right hand side of the trivial identity $\mathrm{d}^2 S(\phi)(V_1, V_2) \equiv 0$. The result is given in the following theorem.

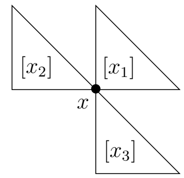


Figure 2.1. The triangles $[x_1], [x_2]$ and $[x_3]$ touching a given vertex x.

Theorem 2.13. *If ϕ is a solution of (2.10), and V_1, V_2 are first variations of ϕ, then the following discrete multisymplectic form formula holds:*

$$\sum_{[x]:[x]\cap\partial U_V \neq \varnothing} \left(\sum_{i=1}^{k} (V_1(x_i) \lrcorner V_2(x_i) \lrcorner \Omega_L^{(i)})(\psi([x])) \right) = 0. \tag{2.14}$$

Proof: The proof of this theorem can also be found in [**80**]. ◇

2.6.3. *A special case.* A related formula was derived by Bridges & Reich in [**19**]. As we shall see, a version of their result can be derived from (2.14) in the case that $k = 3$.

Note first of all that (2.13) and hence (2.14) still hold if (with some abuse of notation) ∂U_V consists of a single grid point x. The sum in (2.14) is then over all triangles that touch x:

$$\sum_{l=1}^{3} \Omega_L^{(l)}(\psi([x_l]))(V_1(x), V_2(x)) = 0, \tag{2.15}$$

where $[x_l]$, $l = 1, 2, 3$, are the triangles that touch $x_{i,j}$, with the convention that $[x_l]_l = x$ (see figure 2.1). We now introduce the following notation:

$$\eta_{i,j}^{(x)} = \Omega_L^{(2)}(\psi([x_2]))(V_1(x_{i,j}), V_2(x_{i,j})) \quad \text{and} \quad \eta_{i,j}^{(y)} = \Omega_L^{(3)}(\psi([x_3]))(V_1(x_{i,j}), V_2(x_{i,j})).$$

By using the fact that $\Omega_L^{(1)} + \Omega_L^{(2)} + \Omega_L^{(3)} = 0$ in (2.15), we then obtain the following discrete conservation law:

$$\eta_{i,j}^{(x)} - \eta_{i-1,j}^{(x)} + \eta_{i,j}^{(y)} - \eta_{i,j-1}^{(y)} = 0,$$

to be compared with [**19**, Proposition 1].

Chapter 3

Discrete Lagrangian field theories on Lie groupoids

In the previous chapter, we have recalled a number of basic aspects from discrete field theories. Although geometric objects such as the Poincaré-Cartan forms already made a brief appearance there, it might appear to the untrained eye that the role of geometry in discrete field theory is rather limited.

The present chapter is partly designed to dispel that impression. We examine a new class of field theories, taking values in a *Lie groupoid*. As we shall see shortly, the description of such field theories is strongly influenced by the geometry of the target Lie groupoid. Since the developments of the previous chapter are encompassed by the current framework, this provides a sounder foundation for some of the constructions in the previous chapter.

Moreover, the use of Lie groupoid field theories is not limited to acting as a new framework for old results: in chapters 4 and 5, we will show that these field theories arise naturally in the context of symmetry and reduction.

1. The discrete jet bundle

As in the previous chapter, we consider only field theories where the bundle is of the form $\pi : \mathbb{R}^2 \times Q \to \mathbb{R}^2$ (see the discussion at the beginning of section 2 in chapter 2). In the last chapter, an extension of this formalism to arbitrary bundles will be sketched.

In addition we assume the existence of a Lie groupoid G over Q. A Lie groupoid is a generalization of the usual concept of a Lie group, with the important distinction that the multiplication of two arbitrary elements is not always defined. Briefly, it is a small category in which all the arrows are invertible, and which is equipped with a suitable smooth structure. In appendix B, we have collected some basic definitions and examples regarding Lie groupoids.

In section 1.1, we start by reviewing the mesh concept from chapter 2. It isn't until section 1.2, when we introduce an appropriate discretization of the jet bundle, that the Lie groupoid G is needed.

1.1. Discretizing the base space. This section is essentially a recapitulation of the discussion following definition 2.1 in chapter 2. The material is not really new, but a number of points are emphasized that were glossed over in chapter 2.

1.1.1. *The mesh.* In chapter 2, a mesh is defined as a discrete subset V of $\mathbb{R}^2$. This definition has to be amended somewhat for Lie groupoid field theories: in addition to knowing the vertices of the mesh, we also need some way of telling which vertices "belong together". This can be made more rigorous by means of some elementary concepts from graph theory, which we now review.

A *graph* is a pair of sets (V, E), where the elements of V are called *vertices*, while the elements of E are pairs $\{x_0, x_1\}$ of vertices called *edges*. In contrast to what is usually assumed in graph theory, we will allow V and E to be (countably) infinite. Note that the edges in E are *undirected*, and that our class of graphs is automatically *simple* since there is by definition at most one edge connecting each pair of distinct vertices.

A *path* between two vertices x and y is a sequence of edges $\{x, p_1\}, \{p_1, p_2\}, \ldots, \{p_l, y\}$. A graph is said to be *connected* if there exists a path between any two vertices. In the sequel, we will only consider connected graphs.

A *planar graph* is a graph (V, E) where V is a subset of $\mathbb{R}^2$ and the edges are curves in $\mathbb{R}^2$ connecting pairs of vertices such that if any two edges intersect, they do so in a common vertex. For a planar graph, there is a notion of *face*, defined as follows. Consider the *geometric realisation* $|E|$ of (V, E), which is just the union of all edges. The complement $\mathbb{R}^2 \backslash |E|$ of $|E|$ is a disconnected set, whose connected components are the *faces* of the planar graph (V, E). A face is therefore a region in the plane, bounded by a number of edges.

The *degree of a face* is defined as the number of edges that make up the boundary of that face. Dually, the *degree of a vertex* is defined as the number of edges arriving in that vertex.

Armed with these definitions, we now come to the following definition of the concept of a "mesh". In some sense, this definition was already implicit in chapter 2: the sets of triangles or quadrangles from section 2.1 can be viewed as the set of faces of a certain graph, whose set of edges is determined implicitly.

Definition 1.1. *A* mesh *in* $X = \mathbb{R}^2$ *is a connected planar graph* (V, E) *in* X *such that the following conditions are satisfied:*

1. *the edges are realised as segments of straight lines in* $\mathbb{R}^2$*;*
2. *the degree of the faces is constant and equal to some natural number* $k > 2$*;*
3. *the degree of the vertices is always larger than two.*

Again, the precise characteristics of such a mesh have to be dictated by the problem under scrutiny. For most of this chapter, we will use the triangular and quadrangular

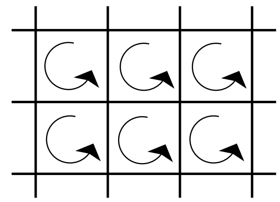


Figure 3.1. Square mesh in $\mathbb{R}^2$, with counterclockwise orientation.

meshes from chapter 2, but note that definition 1.1 also allows for example hexagonal meshes. Furthermore, the developments in this chapter can be easily extended to the case of irregular meshes as well. Other generalizations, such as meshes where the edges are arcs, are also conceivable.

Remark 1.2. A few trivial remarks concerning notation are in order:

(1) A typical element $\mathfrak{e}$ of E is a straight line segment, and is hence determined by specifying its begin and end vertices x_0 and x_1. This will be reflected in our notation: we denote $\mathfrak{e}$ simply as $\{x_0, x_1\} \in V \times V$.
(2) Each face $\mathfrak{f}$ is a k-sided polygon, and will hence be denoted by specifying its corner vertices:
$$\mathfrak{f} \rightsquigarrow \{x_1, \ldots, x_k\} \in V^{\times k}.$$
Furthermore, the set of all faces will be denoted as F. ◇

1.1.2. *The local groupoid E.* In order to bring to the fore the algebraic character of the set of edges E of a given mesh (V, E), we construct a new set E', whose elements are *ordered* pairs $(x, y) \in V \times V$ satisfying the following axioms:

(1) $(x, x) \in E'$ for all $x \in V$;
(2) if $\{x, y\}$ is an element of E, then $(x, y) \in E'$ and $(y, x) \in E'$.

The important difference between E and E' is that the elements of E are *undirected* edges, whereas the elements of E' are *directed*. As we will no longer have a use for E, no confusion can arise if we, henceforth, denote E' simply by E.

If we define the source and target mappings $\alpha_X, \beta_X : E \to V$ in the usual way as $\alpha_X(x, y) = x$ and $\beta_X(x, y) = y$, then E is a subset of the pair groupoid $V \times V$, satisfying all but one of the axioms of a discrete groupoid: if $\mathfrak{e}_1 = (x, y)$ and $\mathfrak{e}_2 = (y, z)$ are elements of E such that $\beta_X(\mathfrak{e}_1) = \alpha_X(\mathfrak{e}_2)$, then the multiplication $\mathfrak{e}_1 \cdot \mathfrak{e}_2$, defined as $\mathfrak{e}_1 \cdot \mathfrak{e}_2 = (x, z)$, is an element of $V \times V$ but not necessarily of E.

This is strongly reminiscent of the concept of *local groupoid* introduced by Van Est in [**101**] in the context of Lie groupoids as, roughly speaking, differentiable groupoids in which the condition $\beta(\mathfrak{e}_1) = \alpha(\mathfrak{e}_2)$ is necessary but not sufficient for the product $\mathfrak{e}_1 \cdot \mathfrak{e}_2$

to exist. Even though in its original definition this concept makes no sense for discrete spaces, the name is nevertheless quite appropriate and so we will continue to refer to E as a local groupoid.

1.1.3. *The set of k-gons $\mathbb{X}^k$.* We now introduce the set of k-gons $\mathbb{X}^k$. The elements of this set are the faces of the mesh, but with a consistent orientation. Indeed, the natural orientation of $X = \mathbb{R}^2$ allows us to write down the edges of each face $\mathfrak{f}$ in (say) counterclockwise direction:

$$\mathfrak{f} = \big((x_k, x_1), (x_1, x_2), \ldots, (x_{k-1}, x_k)\big).$$

We now introduce $\mathbb{X}^k$ as the set of all faces, considered as k-tuples of edges written down in the counterclockwise direction:

$$\mathbb{X}^k = \left\{ \big((x_k, x_1), (x_1, x_2), \ldots, (x_{k-1}, x_k)\big) \quad \text{where } \{x_1, \ldots, x_k\} \in F \right\}.$$

We will also refer to the elements of $\mathbb{X}^k$ as *k-gons* and denote them as

$$[x] := \big((x_k, x_1), (x_1, x_2), \ldots, (x_{k-1}, x_k)\big).$$

To refer to the ith component of a k-gon $[x]$, we will use the subscript notation: $[x]_1 = (x_k, x_1)$ and $[x]_i = (x_{i-1}, x_i)$ for $i = 2, \ldots, k$. In the following, we will assume that the indices are defined "modulo k, plus one", which allows us to write $[x]_i = (x_{i-1}, x_i)$, for all $i = 1, \ldots, k$.

It is useful to note that a k-gon is not changed by a cyclic permutation of its elements and that the common edge of two adjacent k-gons is traversed in opposite directions.

Example 1.3. In the example given in figure 3.1, the degree of the faces is four. The elements of $\mathbb{X}^4$ are the faces with the counterclockwise orientation indicated on the figure. ◇

1.2. The discrete jet space $\mathbb{G}^k$. We now complete our programme of discretizing the jet bundle of π. Recall that there exists a Lie groupoid G over Q with source map α and target map β, and where Q is regarded as a submanifold of G. We will now use G to construct a manifold $\mathbb{G}^k$, which is similar to $\mathbb{X}^k$, playing the role of discrete jet bundle. Roughly speaking, the elements of $\mathbb{G}^k$ are sequences of k elements in G, which are composable, and such that, when multiplied together, they yield a unit element.

Definition 1.4. *The* discrete jet bundle *is the manifold $\mathbb{G}^k$ consisting of all ordered k-tuples $(g_1, \ldots, g_k) \in G \times \cdots \times G$ such that*

$$(g_1, g_2), (g_2, g_3), \ldots, (g_k, g_1) \in G_2 \quad \text{and} \quad g_1 \cdot g_2 \cdots g_k = \alpha(g_1) (= \beta(g_k)).$$

(Recall that G_2 is the set of composable pairs; see appendix B.)

Elements of $\mathbb{G}^k$ will be referred to as "k-gons" in G, and will be denoted as $[g] = (g_1, \ldots, g_k)$. A subscript refers to the individual components of a k-gon: $[g]_i = g_i$. Note that, whereas $\mathbb{X}^k$ is a discrete set due to its compatibility with the mesh, $\mathbb{G}^k$ is a smooth manifold and $\dim \mathbb{G}^k \geq \dim G$.

Example 1.5. At this point, it is perhaps useful to show how definition 1.4 ties in with the developments in chapter 2. Consider the pair groupoid $Q \times Q$ over Q; it is then easy to see that the discrete jet bundle associated to this particular choice of groupoid is just the k-fold product $Q^{\times k}$. This is precisely the fibre part of the discrete jet bundle defined in chapter 2 (see definition 2.3). ◇

The discrete jet bundle $\mathbb{G}^k$ can be equipped with the following two operations:

(1) the inverse of a given k-gon $[g]$, denoted as $[g]^{-1}$ and defined as
$$[g]^{-1} = (g_k^{-1}, g_{k-1}^{-1}, \ldots, g_1^{-1});$$
(2) a collection of k mappings $\alpha^{(i)} : \mathbb{G}^k \to Q$, called *generalized source maps* and defined as $\alpha^{(i)}([g]) = \alpha(g_i)$.

2. Discrete fields

Formerly, a discrete field was a map from the set of vertices V to the manifold Q. In the present context, a discrete field will be a certain kind of map from the set of edges E to the Lie groupoid G, with the important property that composable pairs of edges (*i.e.* edges having a vertex in common) are mapped to composable pairs in G. The definition from chapter 2 can be recovered by considering discrete fields taking values in the pair groupoid $Q \times Q$.

Definition 2.1. *A* discrete field *is a pair $\phi = (\phi_{(0)}, \phi_{(1)})$, where $\phi_{(0)}$ is a map from V to Q and $\phi_{(1)}$ is a map from E to G such that*

(1) *$\alpha(\phi_{(1)}(x, y)) = \phi_{(0)}(x)$ and $\beta(\phi_{(1)}(x, y)) = \phi_{(0)}(y)$;*
(2) *for each $(x, y) \in E$, $\phi_{(1)}(y, x) = [\phi_{(1)}(x, y)]^{-1}$.*
(3) *for all $x \in V$, $\phi_{(1)}(x, x) = \phi_{(0)}(x)$.*

The definition we have given here is strongly reminiscent of that of a groupoid morphism (see section 1 in appendix B). Of course E is not a proper groupoid but just a subset of $V \times V$. However, a discrete field can be uniquely extended to a groupoid morphism from $V \times V$ into G, as we now show.

Proposition 2.2. *Let $\phi = (\phi_{(0)}, \phi_{(1)})$ be a discrete field. Then there exists a unique groupoid morphism $\varphi : V \times V \to G$ extending ϕ.*

Proof: First of all, we define a map $f : V \to Q$ by putting $f(x) := \phi_{(0)}(x) \in Q$ for all $x \in V$. This will be the base map of the morphism φ.

Now, let (x, y) be any element of $V \times V$. If $(x, y) \in E$, then we put $\varphi(x, y) := \phi_{(1)}(x, y)$. If $(x, y) \notin E$, then, because of the connectivity of the mesh (see definition 1.1), there exists a sequence $(x, u_1), (u_1, u_2), \ldots, (u_l, y)$ in E such that in the pair groupoid $V \times V$,

$$(x, y) = (x, u_1) \cdot (u_1, u_2) \cdots (u_l, y). \tag{3.1}$$

We now put $\varphi(x, y) = \phi(x, u_1) \cdot \phi(u_1, u_2) \cdots \phi(u_l, y)$. As each factor on the right-hand side is composable with the next (see property (1) in def. 2.1), this multiplication is well defined. We only have to prove that $\varphi(x, y)$ does not depend on the sequence used in (3.1). Therefore, consider any other decomposition of (x, y) as a product in $V \times V$ of elements of E, *i.e.*

$$(x, y) = (x, u'_1) \cdot (u'_1, u'_2) \cdots (u'_m, y). \tag{3.2}$$

and form the product

$$(x, x) = (x, u_1) \cdot (u_1, u_2) \cdots (u_l, y) \cdot (y, u'_m) \cdot (u'_m, u'_{m-1}) \cdots (u'_1, x).$$

By acting on both sides with φ, we obtain

$$f(x) = \varphi(x, u_1) \cdots \varphi(u_l, y) \cdot [\varphi(u'_m, y)]^{-1} \cdots [\varphi(x, u'_1)]^{-1}$$

and therefore

$$f(x)\varphi(x, u'_1) \cdots \varphi(u'_m, y) = \varphi(x, u_1) \cdots \varphi(u_l, y).$$

By noting that $f(x) = \alpha(\varphi(x, u'_1))$, a left-sided unit, we obtain the desired path independence.

To prove that φ is unique, we consider a second groupoid morphism φ', with base map f', extending ϕ, *i.e.* such that

$$\varphi'(x, y) = \varphi(x, y) = \phi_{(1)}(x, y) \quad \text{for } (x, y) \in E.$$

Then, let (x, y) be an arbitrary element of $V \times V$. By writing (x, y) as a sequence of elements in E as in (3.2), and applying φ' to this product, we may conclude that φ' coincides with φ on the whole of $V \times V$. ◇

Remark 2.3. [1] The preceding proposition makes clear why property 3 of Definition 2.1 cannot be omitted. Indeed, consider the Lie group $G = GL(2, \mathbb{R})$, and let $(\phi_{(0)}, \phi_{(1)})$ be the pair of constant maps defined as

$$\phi_{(0)}(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \phi_{(1)}(x, y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The pair $(\phi_{(0)}, \phi_{(1)})$ satisfies the requirements of definition 2.1 except for property 3, but cannot be extended to a groupoid morphism. ◇

[1] We are grateful to R. Benito and D. Martín de Diego for pointing out to us this example, as well as the absence of property 3 from Definition 2.1 in an earlier version of [**108**].

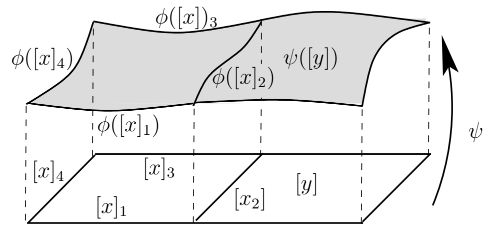


Figure 3.2. A discrete field ϕ and its associated mapping $\psi : \mathbb{X}^k \to \mathbb{G}^k$

Henceforth, we will use the notation 'ϕ' for both the discrete field $(\phi_{(0)}, \phi_{(1)})$ and the groupoid morphism extending it. The notation Φ (used in appendix B) will be reserved for morphisms from G to itself.

Let us briefly recapitulate these developments. We started from the mesh (V, E) in X, and defined a discrete field essentially as a mapping attaching a groupoid element to each element of E. We then showed that such discrete fields are equivalent to morphisms of groupoids.

It now remains to make the link between morphisms from $V \times V$ to G on the one hand, and mappings from $\mathbb{X}^k$ to $\mathbb{G}^k$ on the other hand. It is straightforward to see that a morphism $\phi : V \times V \to G$ induces a map $\psi : \mathbb{X}^k \to \mathbb{G}^k$ by putting

$$\psi([x]) = (\phi([x]_1), \ldots, \phi([x]_k)). \tag{3.3}$$

(see also figure 3.2). The map ψ has some properties reminiscent of those of groupoid morphisms. Of particular importance is the following:

Morphism property: if $[x]$ and $[y]$ are elements of $\mathbb{X}^k$ having an edge in common, then the images of $[x]$ and $[y]$ under ψ have the corresponding edge in $\mathbb{G}^k$ in common. Explicitly:

$$[x]_l = ([y]_m)^{-1} \quad \text{implies that} \quad \psi([x])_l = (\psi([y])_m)^{-1}. \tag{3.4}$$

Proposition 2.4. *There is a one-to-one correspondence between groupoid morphisms $\phi : V \times V \to G$ and mappings $\psi : \mathbb{X}^k \to \mathbb{G}^k$ satisfying the morphism property.*

Proof: We have already associated with a groupoid morphism ϕ a map ψ satisfying the morphism property. To prove the converse, let $\psi : \mathbb{X}^k \to \mathbb{G}^k$ be a map satisfying the morphism property. Define first $\phi : E \to G$ as follows.

(1) For $(u, u) \in E$, we take a k-gon $[x]$ having u as its lth vertex: $u = \alpha_X([x]_l)$ and we put

$$\phi(u, u) = \alpha^{(l)}(\psi([x])).$$

It is straightforward but rather tedious to show that this expression does not depend on the choice of $[x]$. Let $[y]$ be another k-gon, with u as its mth vertex.

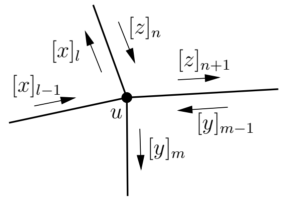


Figure 3.3. A vertex of degree four.

Let us assume for the sake of simplicity that u has degree four (the general case can be dealt with by repeated application of this special case). Then the edges that emerge from u are $[x]_l$ and $[y]_m$, as well as $([x]_{l-1})^{-1}$ and $([y]_{m-1})^{-1}$ (see figure 3.3) and there exists exactly one k-gon $[z]$ such that

$$[z]_n = ([x]_l)^{-1} \quad \text{and} \quad [z]_{n+1} = ([y]_{m-1})^{-1}.$$

By definition, we have

$$\beta(\psi([z])_n) = \alpha(\psi([z])_{n+1}) \quad \text{and} \quad \beta(\psi([y])_{m-1}) = \alpha(\psi([y])_m).$$

On the other hand, the morphism property ensures that

$$\psi([x])_l = (\psi([z])_n)^{-1} \quad \text{and} \quad \psi([y])_{m-1} = (\psi([z])_{n+1})^{-1}.$$

By applying α to the left equality and β to the right equality, we finally obtain

$$\alpha^{(l)}(\psi([x])) = \alpha^{(m)}(\psi([y])),$$

which shows that $\phi(u,u)$ does not depend on $[x]$.

(2) For $(u,v) \in E$, $u \neq v$, we take $[x]$ in $\mathbb{X}^k$ such that $(u,v) = [x]_l$ and we put

$$\phi(u,v) = \psi([x])_l.$$

This is well defined because of the morphism property and, moreover, ϕ satisfies $\phi(y,x) = (\phi(x,y))^{-1}$.

By applying proposition 2.2 we obtain the desired morphism $\phi : V \times V \to G$. $\diamond$

In a way, the mapping ψ associated to a morphism ϕ plays the role of “first jet prolongation” of ϕ. Similarly, the morphism property is in some sense a discrete analogue of the distinction between holonomic and non-holonomic sections of $J^1\pi$.

Remark 2.5. It is perhaps useful to illustrate the theory developed so far by applying it to groupoid mechanics. In this case, the base space X is $\mathbb{R}$, but all of the constructions for $X = \mathbb{R}^2$ carry through to this case. As a discretization of $\mathbb{R}$, we choose the canonical injection $\iota : \mathbb{Z} \hookrightarrow \mathbb{R}$. A discrete field can then be identified with a bi-infinite sequence of pairwise composable groupoid elements $\ldots, g_{-2}, g_{-1}, g_0, g_1, \ldots$, which is precisely the definition of an *admissible sequence* in [**75, 110**]. $\diamond$

2.1. The prolongation $P^k\mathbb{G}$. Let AG denote the Lie algebroid associated to G (see appendix B).

We recall that the discrete jet bundle $\mathbb{G}^k$ is equipped with k generalized source maps $\alpha^{(i)} : \mathbb{G}^k \to Q$, defined as $\alpha^{(i)}([g]) = \alpha([g]_i)$. By use of these maps, we define the *prolongation* $P^k\mathbb{G}$ of $\mathbb{G}^k$ through the following commutative diagram:

$$\begin{array}{ccc} P^k\mathbb{G} & \longrightarrow & AG \times \cdots \times AG \\ \downarrow & & \downarrow \\ \mathbb{G}^k & \longrightarrow & Q \times \cdots \times Q \end{array}$$

It follows that $P^k\mathbb{G}$ consists of elements $([g]; v_1, \ldots, v_k)$, where $v_i \in A_{\alpha(g_i)}G$ for each $i = 1, \ldots, k$. We denote by $\pi^{(k)} : P^k\mathbb{G} \to \mathbb{G}^k$ the projection which maps $([g]; v_1, \ldots, v_k)$ onto $[g]$. Furthermore, there exist k bundle morphisms $(P^{(i)}, p^{(i)}) : P^k\mathbb{G} \to PG$, where PG is the prolongation of the Lie groupoid G, defined in section 3.1.1 of appendix B. These morphisms are defined as follows. The base space map $p^{(i)} : \mathbb{G}^k \to G$ is the projection onto the ith factor, $p^{(i)}([g]) = [g]_i$, and the total space map $P^{(i)}$ is defined as

$$P^{(i)}([g]; v_1, \ldots, v_k) = (g_i; v_i, v_{i+1}). \tag{3.5}$$

The definition of $P^k\mathbb{G}$ is strongly reminiscent of that of the prolongation of a Lie groupoid over a fibration (see section 3.1), although in general $\mathbb{G}^k$ is not a groupoid. The exact nature of $P^k\mathbb{G}$ is unclear at this stage, but we will show in section 2.1.2 that the algebroid structure of PG can be used to equip $P^k\mathbb{G}$ with a Lie algebroid structure by requiring that the maps $(P^{(i)}, p^{(i)})$ are Lie-algebroid morphisms.

Remark 2.6. For $k = 2$, the manifold $\mathbb{G}^2$ is diffeomorphic to G, with the diffeomorphism φ mapping each pair (g, g^{-1}) onto g. Note that $p^{(1)} = \varphi$. In addition, we have that

$$\alpha^{(1)} = \alpha \circ \varphi \quad \text{and} \quad \alpha^{(2)} = \beta \circ \varphi,$$

confirming our intuition that the maps $\alpha^{(i)}$ are some sort of "generalized source maps". Furthermore, the projection $P^{(1)}$ is given by

$$P^{(1)}(g, g^{-1}; u_{\alpha(g)}, v_{\beta(g)}) = (g; u_{\alpha(g)}, v_{\beta(g)}),$$

and so in fact it is just the natural identification of $P^2\mathbb{G}$ with PG. On the other hand, $P^{(2)}$ is given by

$$P^{(2)}(g, g^{-1}; u_{\alpha(g)}, v_{\beta(g)}) = (g^{-1}; v_{\beta(g)}, u_{\alpha(g)}).$$

We recalled in section 3.1 that PG is a groupoid over AG in a natural way. A brief comparison shows that $P^{(2)}$ is just the inversion mapping of PG, once we use $P^{(1)}$ to identify PG and $P^2\mathbb{G}$. ◇

2.1.1. *The injection $\mathcal{I} : P^k\mathbb{G} \hookrightarrow T\mathbb{G}^k$.* Of central importance for the following developments is the fact that there exists a bundle injection $\mathcal{I}$ of $P^k\mathbb{G}$ into $T\mathbb{G}^k$. In order to define $\mathcal{I}$, we recall that a section v of the Lie algebroid AG defines on G a left-invariant vector field v^L and a right-invariant vector field v^R (see expression (B.1)). We also recall that we use the same notation for the point-wise operation (see remark 2.3).

Now, let $([g]; v_1, \ldots, v_k)$ be any element of $P^k\mathbb{G}$, and define $\mathcal{I}([g]; v_1, \ldots, v_k) \in T_{[g]}\mathbb{G}^k$ as

$$\mathcal{I}([g]; v_1, \ldots, v_k) = (v_1^R(g_1) + v_2^L(g_1), v_2^R(g_2) + v_3^L(g_2), \ldots, v_k^R(g_k) + v_1^L(g_k)). \tag{3.6}$$

To prove that the right-hand side is a tangent vector to $\mathbb{G}^k$ at $[g]$, we take for each $i = 1, \ldots, k$ a curve $\epsilon \mapsto h_i(\epsilon) \in \mathcal{F}^\alpha(g_i)$ in the α-fibre through g_i such that $h_i(0) = \alpha(g_i)$ and $\dot{h}_i(0) = v_i$. Then the vector on the right-hand side is the tangent vector at 0 to the following curve in $\mathbb{G}^k$:

$$\epsilon \mapsto \big(h_1^{-1}(\epsilon)g_1h_2(\epsilon), h_2^{-1}(\epsilon)g_2h_3(\epsilon), \ldots, h_k^{-1}(\epsilon)g_kh_1(\epsilon)\big).$$

Definition 2.7. *Let $[g]$ be an element of $\mathbb{G}^k$. The* ith tangent lift *is the map $L^{(i)}_{[g]} : A_{\alpha(g_i)}G \to T_{[g]}\mathbb{G}^k$ defined as*

$$L^{(i)}_{[g]}(v) = \mathcal{I}([g]; 0, \ldots, 0, v, 0, \ldots, 0) \quad \textit{for } v \in A_{\alpha(g_i)}G,$$

where v occupies the ith position among the arguments of $\mathcal{I}([g]; \ldots)$. We will frequently use the notation $v^{(i)}_{[g]}$ for the element $L^{(i)}_{[g]}(v)$.

Remark 2.8. We pointed out that $P^2\mathbb{G}$ is isomorphic to PG. In this case, the injection $\mathcal{I}$ is given by

$$\mathcal{I} : (g; u_{\alpha(g)}, v_{\beta(g)}) \mapsto T(r_g \circ i)(u_{\alpha(g)}) + Tl_g(v_{\beta(g)}) \in V_g\beta \oplus V_g\alpha,$$

and coincides with the isomorphism $\Theta : PG \to V\beta \oplus V\alpha$ (see section 3.1 of appendix A). In this case, the map $\mathcal{I}$ can also be seen as the anchor of the Lie algebroid PG. This theme will return in the next section, when we endow $P^k\mathbb{G}$ with the structure of a Lie algebroid, with $\mathcal{I}$ as its anchor map. ◇

2.1.2. *The Lie algebroid structure on $\pi^{(k)} : P^k\mathbb{G} \to \mathbb{G}^k$.* In order to endow $P^k\mathbb{G}$ with the structure of a Lie algebroid, we introduce the concept of the *lift of a section* of AG to $P^k\mathbb{G}$. Let v be a section of AG; we then define $v_{(i)}$ as the section of $P^k\mathbb{G}$ given by

$$v_{(i)}([g]) := ([g]; 0, \ldots, 0, v(\alpha(g_i)), 0, \ldots, 0),$$

where $v(\alpha(g_i))$ occupies the ith place. This lift operation should not be confused with the tangent lift of definition 2.7 (although they are related), and will only be used in this section.

We now define a Lie algebroid structure on $P^k\mathbb{G}$. The anchor $\rho^{(k)} : P^k\mathbb{G} \to T\mathbb{G}^k$ is the injection $\mathcal{I}$ defined in (3.6), and the bracket is defined component-wise as follows. Let v and w be sections of AG, and define

$$[v_{(i)}, w_{(j)}]_{P^k\mathbb{G}} = \begin{cases} 0 & \text{if } i \neq j, \\ [v, w]_{(i)} & \text{if } i = j. \end{cases} \tag{3.7}$$

It is straightforward to prove that the bracket and the anchor satisfy the requirements of definition 2.1 in appendix B. We denote the associated exterior differential on $\bigwedge(P^k\mathbb{G})^*$ by $\mathrm{d}^{(k)}$.

Corollary 2.9. *The projection mappings $P^{(i)} : P^k\mathbb{G} \to PG$ defined in (3.5) are Lie algebroid morphisms.*

Proof: The projection map $P^{(i)}$ is fibrewise surjective; hence we will use proposition 2.4 in appendix B. The first part, $\hat{\rho} \circ P^{(i)} = Tp_{(i)} \circ \rho^{(k)}$, follows easily from the definitions.

Now, let v and w be sections of AG. Note that $v_{(i)}$ and $v_{(i-1)}$ are $P^{(i)}$-related to respectively $v^{(1,0)}$ and $v^{(0,1)}$, where we have used the notations of section 3.1.1 in appendix B. If $j \neq i-1, i$, then $v_{(j)}$ is $P^{(i)}$-related to the zero section of AG.

To complete the proof, we now need to show that $[v_{(i)}, w_{(i)}]$ is $P^{(i)}$-related to $[v, w]^{(1,0)}$, $P^{(i-1)}$-related to $[v, w]^{(0,1)}$, and $P^{(j)}$-related ($j \neq i-1, i$) to the zero section. Explicitly:

$$P^{(j)} \circ [v_{(i)}, w_{(i)}]_{P^k\mathbb{G}} = \begin{cases} [v, w]^{(1,0)} \circ p_i & \text{if } j = i, \\ [v, w]^{(0,1)} \circ p_{i-1} & \text{if } j = i-1, \\ 0 & \text{all other cases.} \end{cases} \tag{3.8}$$

This follows immediately from (3.7). ◇

3. Lagrangian field theories

Having thus prepared the geometric stage, we now turn to the analysis of discrete field theories. In essence, this is not very different from the procedure followed in chapter 2: the discrete Euler-Lagrange equations characterise the extremals of a certain discrete action sum, etc. The difference lies of course in the details: a discrete Lagrangian is now a function on $\mathbb{G}^k$, the Poincaré-Cartan forms are sections of the dual of $P^k\mathbb{G}$, and the discrete Legendre transformations are defined accordingly.

3.1. The Poincaré-Cartan forms. Let $L : \mathbb{G}^k \to \mathbb{R}$ be a discrete Lagrangian. To L one can associate k sections $\theta_L^{(i)}$ of $(\pi^{(k)})^* : (P^k\mathbb{G})^* \to \mathbb{G}^k$, called *Poincaré-Cartan forms*, which are defined as follows:

$$\left\langle \theta_L^{(i)}([g]), ([g]; v_1, \ldots, v_k) \right\rangle = (v_i^{(i)})_{[g]}(L),$$

where $v_i \in A_{\alpha(g_i)}G$ and $v_i^{(i)}$ is the ith tangent lift of v_i to $\mathbb{G}^k$ (cf. definition 2.7). As $\sum v_i^{(i)} = \mathcal{I}([g]; v_1, \ldots, v_k)$, we may conclude that

$$\mathrm{d}^{(k)} L = \sum_{i=1}^{k} \theta_L^{(i)}.$$

Remark 3.1. In the case $k = 2$, it follows from remark 2.8 that $\theta_L^{(1)}$, resp. $\theta_L^{(2)}$, can be identified with the Poincaré-Cartan forms θ_L^-, resp. θ_L^+, defined in [**75**] as

$$\theta_L^-(g; u_{\alpha(g)}, v_{\beta(g)}) = \mathrm{d}L(g)(u^R(g)) \quad \text{and} \quad \theta_L^+(g; u_{\alpha(g)}, v_{\beta(g)}) = \mathrm{d}L(g)(v^L(g)).$$

Indeed, let us consider the function L_{mech} on G given by $L_{\mathrm{mech}} = \varphi_* L$, where $\varphi : \mathbb{G}^2 \to G$ is the diffeomorphism introduced in remark 2.6, or, explicitly, $L_{\mathrm{mech}}(g) = L(g, g^{-1})$. Then, by definition,

$$\theta_L^{(1)}(g, g^{-1}; u_{\alpha(g)}, v_{\beta(g)}) = \frac{\mathrm{d}}{\mathrm{d}t} L(h^{-1}(t)g, g^{-1}h(t))\Big|_0,$$

where $h(t) \in \mathcal{F}^\alpha(g)$ is such that $h(0) = \alpha(g)$ and $\dot{h}(0) = u_{\alpha(g)}$. The right-hand side can now be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t} L_{\mathrm{mech}}(h^{-1}(t)g)\Big|_0 = \big\langle \mathrm{d}L_{\mathrm{mech}}, T(r_g \circ i)(u_{\alpha(g)})\big\rangle = \theta_L^-(g; u_{\alpha(g)}, v_{\beta(g)}).$$

There is a similar identification of $\theta_L^{(2)}$ with $\theta_{L_{\mathrm{mech}}}^+$. ◇

3.2. The field equations. We derive the discrete field equations for a Lie groupoid morphism $\phi : V \times V \to G$ by varying a discrete action sum. Let $L : \mathbb{G}^k \to \mathbb{R}$ be a discrete Lagrangian and consider a finite subset U_F of $\mathbb{X}^k$. Define the *action sum* S as follows:

$$S(\phi) = \sum_{[x] \in U_F} L(\psi([x])), \tag{3.9}$$

where ψ is the map from $\mathbb{X}^k$ to $\mathbb{G}^k$ associated to the morphism ϕ (see proposition 2.4). In order to derive the discrete Euler-Lagrange equations, we first need a suitable definition of finite and infinitesimal variations.

3.2.1. *Variations.* Let us start with the definition of a finite variation. A key property is of course that the variation of a groupoid morphism should yield a new groupoid morphism.

There are two ways in which this property can be implemented. First of all, if we consider discrete fields as morphisms $\phi : V \times V \to G$, a finite variation Φ_ϵ has to be a groupoid morphism of G to itself. Only then is the composition of Φ_ϵ with ϕ again a morphism. However, if we view discrete fields as mappings $\psi : \mathbb{X}^k \to \mathbb{G}^k$ satisfying the morphism property, then it is not so obvious how to define a finite variation. We want the composition $\Psi_\epsilon \circ \psi$ to satisfy again the morphism property: this is achieved by imposing a number of additional conditions, which can be thought of as "morphism properties" for $\mathbb{G}^k$.

Of course, in view of proposition 2.4, one expects that maps Ψ satisfying these morphism properties (to be defined below) are just groupoid morphisms from G to itself in disguise. This is proved in proposition 3.3 below.

Let us first introduce a slight modification of the source mappings $\alpha^{(i)}$:

$$\hat{\alpha}^{(i)} : \mathbb{G}^k \to \mathbb{G}^k, \quad \hat{\alpha}^{(i)}([g]) = (\alpha([g]_i), \ldots, \alpha([g]_i)).$$

It is obvious that for any $l \leq k$, $\big(\hat{\alpha}^{(i)}([g])\big)_l = \alpha^{(i)}([g])$.

Definition 3.2. *A map $\Psi : \mathbb{G}^k \to \mathbb{G}^k$ is said to satisfy the* morphism properties *if, for all $[g], [h] \in \mathbb{G}^k$,*

(1) $\Psi \circ \hat{\alpha}^{(i)} = \hat{\alpha}^{(i)} \circ \Psi$ *for $i = 1, \ldots, k$;*
(2) *if $[g]_l = [h]_m$, then $\Psi([g])_l = \Psi([h])_m$.*

Proposition 3.3. *There is a one-to-one correspondence between groupoid morphisms $\Phi : G \to G$ and mappings $\Psi : \mathbb{G}^k \to \mathbb{G}^k$ satisfying the morphism properties.*

Proof: Let Φ be a morphism from G to itself. As in (3.3), Φ induces a mapping $\Psi : \mathbb{G}^k \to \mathbb{G}^k$ satisfying the morphism properties, namely:

$$\Psi([g]) = (\Phi([g]_1), \ldots, \Phi([g]_k))\,.$$

Conversely, let $\Psi : \mathbb{G}^k \to \mathbb{G}^k$ be a mapping satisfying the morphism properties and let g be any element of G. In order to define $\Phi(g)$, we take any $[\eta] \in \mathbb{G}^k$ such that there exists a natural number $l \leq k$ for which $g = [\eta]_l$. We then put

$$\Phi(g) := \Psi([\eta])_l.$$

Morphism property 2 ensures that $\Phi(g)$ depends only on g and not on the other components of $[\eta]$. We now have to check that Φ is a morphism of G to itself.

(1) In order to prove that $\alpha \circ \Phi = \Phi \circ \alpha$, we take any $g \in G$ and consider $[\eta] \in \mathbb{G}^k$ such that $[\eta]_l = g$. Then $\alpha(\Phi(g)) = \alpha(\Psi([\eta])_l) = \alpha^{(l)}(\Psi([\eta]))$.
However, because of morphism property 1 we have

$$\hat{\alpha}^{(l)}\big(\Psi([\eta])\big) = \Psi\big(\hat{\alpha}^{(l)}([\eta])\big) = \Psi\big((\alpha(g), \ldots, \alpha(g))\big). \tag{3.10}$$

For any arbitrary $m \leq k$, we have that $\Phi(\alpha(g)) = \Psi\big((\alpha(g), \ldots, \alpha(g))\big)_m$, and so, by considering the mth component of (3.10),

$$\begin{aligned}\Phi(\alpha(g)) &= \big(\hat{\alpha}^{(l)}\big(\Psi([\eta])\big)\big)_m \\ &= \alpha^{(l)}\big(\Psi([\eta])\big),\end{aligned}$$

from which we conclude that $\alpha(\Phi(g)) = \Phi(\alpha(g))$ for all $g \in G$. A similar argument can be used to show that Φ commutes with β.

(2) We now show that $\Phi(g^{-1}) = \Phi(g)^{-1}$ for any $g \in G$. Let

$$[\xi] = (g, g^{-1}, \alpha(g), \ldots, \alpha(g)),$$

then $\Psi([\xi])_1 = \Phi(g)$, $\Psi([\xi])_2 = \Phi(g^{-1})$ and $\Psi([\xi])_j = \Phi(\alpha(g))$ for $j = 3, \ldots, k$. Moreover, since $\Psi([\xi]) \in \mathbb{G}^k$, we have, by definition of $\mathbb{G}^k$, that $\Psi([\xi])_1 \cdots \Psi([\xi])_k = \alpha(\Psi([\xi])_1)$, or

$$\Phi(g)\Phi(g^{-1})\Phi(\alpha(g)) \cdots \Phi(\alpha(g)) = \alpha(\Phi(g)),$$

which, after simplification, leads to $\Phi(g^{-1}) = \Phi(g)^{-1}$.

(3) Finally, we have to show that if (g, h) is a composable pair, *i.e.* $\beta(g) = \alpha(h)$, then $(\Phi(g), \Phi(h))$ is also composable, and moreover, $\Phi(gh) = \Phi(g)\Phi(h)$. The proof of this property is similar to the proof of the previous property.
Consider the following k-gon:

$$[\eta] = (g, h, (gh)^{-1}, \alpha(g), \ldots, \alpha(g)).$$

Then, as $\Psi([\eta]) \in \mathbb{G}^k$, we conclude that, first of all, $\beta(\Phi(g)) = \alpha(\Phi(h))$, and secondly

$$\Phi(g)\Phi(h)\Phi((gh)^{-1}) = \alpha\big(\Phi(g)\big).$$

By using the previous properties, as well as some of the standard properties of the groupoid G, we find that $\Phi(gh) = \Phi(g)\Phi(h)$.

We conclude that $\Phi : G \to G$ is a groupoid morphism. ◇

Corollary 3.4. *Let $\Psi : \mathbb{G}^k \to \mathbb{G}^k$ be a map satisfying the morphism properties. Then for each $[g] \in \mathbb{G}^k$,*

$$\Psi([g]^{-1}) = \Psi([g])^{-1}.$$

Proof: This can be proved directly, or by noting that Ψ induces a groupoid morphism Φ such that

$$\Psi([g]) = \big(\Phi([g]_1), \ldots, \Phi([g]_k)\big),$$

and writing out the definition of $[g]^{-1}$ and $\Psi([g])^{-1}$. ◇

After these introductory lemmas, we now turn to the concepts of finite and infinitesimal variations of a morphism $\phi : V \times V \to G$. Let U_F be the finite subset of $\mathbb{X}^k$ used in defining the discrete action S in (3.9), and recall from definition 2.6 in the previous chapter that U_F induces a finite subset U_V of V. In a similar fashion, U_F induces a finite subset U_E of E, where an edge (x_0, x_1) is an element of U_E if and only if there exists a k-gon $[x]$ in U_F and an index i such that either $(x_0, x_1) = [x]_i$ or $(x_1, x_0) = [x]_i$. We also define the boundary ∂U_E of U_E to be the following set:

$$\partial U_E := \big\{(u, v) \in V \times V : \exists [x], [y] \in \mathbb{X}^k \text{ such that } [x]_l = (u, v), [y]_m = (v, u) \\ \text{and } [x] \in U_F, [y] \notin U_F\big\}.$$

In other words, the boundary ∂U_E consists of edges that, when traversed in opposite directions, are part of two k-gons $[x]$ and $[y]$, one of which is contained in U_F, while the other one is not.

Definition 3.5. *A* finite variation *of a discrete field ϕ is a one-parameter family of maps $h_\epsilon : U_V \to G$ such that $\alpha \circ h_\epsilon = \phi_{(0)}$ for all ϵ. In addition, $h_\epsilon(x) \equiv \phi_{(0)}(x)$ for all $x \in \partial U_E$.*

Recall that $\phi_{(0)} : V \to Q$ was introduced in definition 2.1 as the "base map" of the discrete field ϕ.

In other words, a finite variation is just an assignment of a curve in the α-fibre through $\phi_{(0)}(x)$ to each vertex in U_F. Such a finite variation induces a one-parameter family of maps $\Phi_\epsilon : \phi(U_E) \to G$, defined as

$$\Phi_\epsilon(g) = h_\epsilon(x_0)^{-1} g h_\epsilon(x_1) \quad \text{where } g = \phi(x_0, x_1). \tag{3.11}$$

Even though these maps are not defined on the whole of G, they satisfy the defining properties of a groupoid morphism and hence give rise to a one-parameter family $\Psi_\epsilon : \psi(U_F) \to \mathbb{G}^k$ through proposition 3.3, where $\psi : \mathbb{X}^k \to \mathbb{G}^k$ is the map associated to ϕ as in proposition 2.4. These maps have the following form:

$$\Psi_\epsilon([g]) = (k_1(\epsilon)^{-1} g_1 k_2(\epsilon), k_2(\epsilon)^{-1} g_2 k_3(\epsilon), \ldots, k_k(\epsilon)^{-1} g_k k_1(\epsilon)), \tag{3.12}$$

where $k_i(\epsilon) = h_\epsilon(x_i)$ if $[g] = \psi([x])$.

Based on these constructions, infinitesimal variations can now be defined in a natural way, by taking the derivative of a finite variation:

Definition 3.6. *A* infinitesimal variation *of a discrete field ϕ is a map $V : U_V \to AG$ along $\phi_{(0)}$ (*i.e. *such that $V(x) \in A_q G$, where $q := \phi_{(0)}(x)$). In addition, $V(x) = 0$ if $x \in \partial U_V$.*

An infinitesimal variation V of a field ϕ gives rise to a section Γ of $\pi^{(k)}$ along ψ in the obvious way by putting $\Gamma([x]) = (V(x_1), \ldots, V(x_k))$. This point of view will hardly ever be needed, except in our derivation of the multisymplectic form formula.

3.2.2. *The field equations.* We now have at our disposal all the tools required to derive the discrete field equations. For the sake of notational clarity, this will be done for the quadrangular mesh only; the generalization to non-regular meshes is straightforward but notationally quite intricate.

Let U_F be again the finite subset of $\mathbb{X}^k$ used in (3.9) to define the discrete action, and let U_V and U_E be the associated subsets of V and $V \times V$ respectively. Consider a discrete field $\phi : U_E \to G$ and let $h_\epsilon : U_V \to G$ be a finite variation of ϕ as in definition 3.5. Recall from (3.11) that h_ϵ defines a one-parameter family of maps $\Phi_\epsilon : \phi(U_E) \to G$; we will denote by ϕ_ϵ the composition $\Phi_\epsilon \circ \phi$.

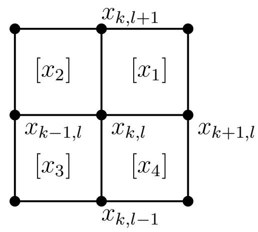


Figure 3.4. The vertex $x_{k,l}$ and the four surrounding faces

The morphism $\phi : U_E \to G$ is an extremum of (3.9) under arbitrary variations if and only if

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} S(\phi_\epsilon)\Big|_{\epsilon=0} = 0. \tag{3.13}$$

From this condition the discrete Euler-Lagrange equations easily follow. The remainder of this section is devoted to this derivation.

Let $x_{k,l}$ be an element of V; naturally, $x_{k,l}$ is a common vertex of four quadrangles, denoted by $[x_i]$, for $i = 1, \dots, 4$. Here, we have employed the convention that $\alpha([x_i]_i) = x_{k,l}$; *i.e.* $x_{k,l}$ is the ith vertex of $[x_i]$ (see figure 3.4). We denote the image of $[x_i]$ under ψ by $[g_i]$.

The effect of the variation h_ϵ on (for example) the quadrangle $[g_1]$ is to map it to a new quadrangle $[g_1(\epsilon)]$, given by

$$[g_1(\epsilon)] = (h_\epsilon(x_{k,l})^{-1} g_1 h_\epsilon(x_{k+1,l}), \dots, h_\epsilon(x_{k,l+1})^{-1} g_4 h_\epsilon(x_{k,l})).$$

Here, $[g_1]$ is written as (g_1, g_2, g_3, g_4).

Similar expressions can be written down for $[g_2]$, $[g_3]$, and $[g_4]$. By substituting these formulae into (3.13), we eventually obtain

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} S(\phi_\epsilon)\Big|_{\epsilon=0} = \sum_{x_{k,l} \in U_V} \left((v_{k,l})^{(1)}_{[g_1]}(L) + (v_{k,l})^{(2)}_{[g_2]}(L) + (v_{k,l})^{(3)}_{[g_3]}(L) + (v_{k,l})^{(4)}_{[g_4]}(L) \right), \tag{3.14}$$

where the superscript i denotes the ith tangent lift of an element of AG to $T\mathbb{G}^k$ (see definition 2.7), and $v_{k,l} := \frac{\mathrm{d}}{\mathrm{d}\epsilon} h_\epsilon(x_{k,l})|_{\epsilon=0}$ is the infinitesimal variation associated to h_ϵ, evaluated at $x_{k,l}$. As the elements $v_{k,l}$ are independent, we obtain the following characterization of the extremals of S.

Theorem 3.7. *Let $\phi : U_E \to G$ be a discrete field defined on a finite set $U_E \subset V \times V$. Then ϕ is an extremum of the action sum (3.9) if and only if, for all $v \in A_q G$, where $q = \phi_{(0)}(x_{i,j})$, the following* discrete Euler-Lagrange equations *hold:*

$$v^{(1)}_{[g_1]}(L) + v^{(2)}_{[g_2]}(L) + v^{(3)}_{[g_3]}(L) + v^{(4)}_{[g_4]}(L) = 0. \tag{3.15}$$

Here, the quadrangles $[g_i]$, $i = 1, \ldots, 4$ are defined as in the discussion preceding this theorem.

In case where G is the pair groupoid $Q \times Q$, the equations (3.15) reduce to the discrete Euler-Lagrange equations derived in chapter 2 following [**80**].

3.3. The Legendre transformation. In this section, we introduce a notion of Legendre transformation and use it to show that the pullback of the canonical section of a suitable dual bundle yields the Poincaré-Cartan forms constructed in section 3.1. More precisely, the Legendre transformation will be a collection of k bundle maps from $P^k\mathbb{G}$ to the bundle $P^{\tau^*}(AG) \to A^*G$. As sketched in section B-3.2.1, the dual of the latter is equipped with a canonical section θ and the pullback of this section by each of the bundle maps corresponding to the Legendre transformation, will provide the full set of Poincaré-Cartan forms.

We first introduce the pullback bundles $P^{(i)}(AG)$, $i = 1, \ldots, k$, constructed by means of the following commutative diagram:

$$\begin{array}{ccc} P^{(i)}(AG) & \longrightarrow & T\mathbb{G}^k \\ \downarrow & & \downarrow{\scriptstyle T\alpha^{(i)}} \\ AG & \underset{\rho}{\longrightarrow} & TQ \end{array}$$

The bundles $P^{(i)}(AG)$ bear the same relation to $\mathbb{G}^k$ as $P^\alpha(AG)$ and $P^\beta(AG)$ to G.

3.3.1. *The mappings $\mathfrak{P}^{(i)} : P^k\mathbb{G} \to P^{(i)}(AG)$.* For each $i = 1, \ldots, k$, there is a natural injection $\varphi^{(i)} : G \to \mathbb{G}^k$ defined as

$$\varphi^{(i)}(g) = (\alpha(g), \ldots, \alpha(g), g, g^{-1}, \alpha(g), \ldots, \alpha(g))$$

where g and g^{-1} occupy the ith and the $(i+1)$th position, respectively. For $i = k$, $\varphi^{(k)}$ is defined as

$$\varphi^{(k)}(g) = (g^{-1}, \alpha(g), \ldots, \alpha(g), g).$$

The projections $P^{(i)} : P^k\mathbb{G} \to PG$, as defined in section 2.1, can be used to define projection mappings $\mathfrak{P}^{(i)} : P^k\mathbb{G} \to P^{(i)}(AG)$ by means of the composition

$$\mathfrak{P}^{(i)} : P^k\mathbb{G} \xrightarrow{P^{(i)}} PG \xrightarrow{A(\Phi^\alpha)} P^\alpha(AG) \xrightarrow{\mathrm{id}\times T\varphi^{(i)}} P^{(i)}(AG),$$

where $A(\Phi^\alpha) : PG \to P^\alpha(AG)$ was defined in section B-3.2.2.

Remark 3.8. For $k = 2$, we now show that $\mathfrak{P}^{(1)}$ and $\mathfrak{P}^{(2)}$ can be identified with $A(\Phi^\alpha)$ and $A(\Phi^\beta)$, respectively. We recall that $P^2\mathbb{G}$ is isomorphic to PG and that there is a diffeomorphism $\varphi : \mathbb{G}^2 \to G$ sending each (g, g^{-1}) to g (see remark 2.6). Hence, $\varphi^{(1)}$ is just φ^{-1} and $\varphi^{(2)}$ equals $\varphi^{-1} \circ i$. There is a natural identification of $P^{(1)}(AG)$ with $P^\alpha(AG)$, and of $P^{(2)}(AG)$ with $P^\beta(AG)$. Using these identifications, it is straightforward to see that $\mathfrak{P}^{(1)}$ can be identified with $A(\Phi^\alpha)$. The identification of $\mathfrak{P}^{(2)}$ with $A(\Phi^\beta)$ takes some more work. Consider first the composition

$$P^\alpha(AG) \stackrel{\mathrm{id}\times T\varphi^{(2)}}{\longrightarrow} P^{(2)}(AG) \cong P^\beta(AG),$$

which is easily seen to be equal to $\mathrm{id} \times Ti$. We then obtain the following for the map $\mathfrak{P}^{(2)}$, considered as a map into $P^\beta(AG)$:

$$\begin{aligned}
\big((\mathrm{id} \times Ti) \circ A(\Phi^\alpha) \circ P^{(2)}\big)&(g, g^{-1}; u_{\alpha(g)}, v_{\beta(g)}) \\
&= (\mathrm{id} \times Ti \circ A(\Phi^\alpha))(g^{-1}; v_{\beta(g)}, u_{\alpha(g)}) \\
&= (\mathrm{id} \times Ti)(g^{-1}; v_{\beta(g)}, T(r_{g^{-1}} \circ i)(v_{\beta(g)}) + Tl_{g^{-1}}(u_{\alpha(g)})) \\
&= (g : v_{\beta(g)}, T(r_g \circ i)(u_{\alpha(g)}) + Tl_g(v_{\beta(g)})) \\
&= A(\Phi^\beta)(g; u_{\alpha(g)}, v_{\beta(g)}),
\end{aligned}$$

where we again refer to section 3.2.2 for the definition of $A(\Phi^\beta)$. ⋄

3.3.2. *Definition of the Legendre transformations.* Given a Lagrangian $L : \mathbb{G}^k \to \mathbb{R}$, there are k distinguished bundle maps $(P\mathbb{F}L^{(i)}, \mathbb{F}L^{(i)})$ from $P^k\mathbb{G}$ to the bundle $P^{\tau^*}(AG)$ over A^*G, which we call *Legendre transformations*.

For each $i = 1, \ldots, k$, the base map $\mathbb{F}L^{(i)} : \mathbb{G}^k \to A^*G$ is defined as follows. For each $[g] \in \mathbb{G}^k$, $\mathbb{F}L^{(i)}([g])$ is the element of $A^*_{\alpha(g_i)}G$ defined by

$$\mathbb{F}L^{(i)}([g])(v_{\alpha(g_i)}) = v^{(i)}_{\alpha(g_i)}(L) \quad \text{for all } v_{\alpha(g_i)} \in A_{\alpha(g_i)}G.$$

Recall that $v^{(i)}_{\alpha(g_i)}$ is the ith tangent lift of $v_{\alpha(g_i)}$ to $T_{[g]}\mathbb{G}^k$. The total space map $P\mathbb{F}L^{(i)} : P^k\mathbb{G} \to P^{\tau^*}(AG)$ is defined as the composition $(\mathrm{id} \times T\mathbb{F}L^{(i)}) \circ \mathfrak{P}^{(i)}$.

Proposition 3.9. *Let θ be the canonical section of $[P^{\tau^*}(AG)]^* \to A^*G$ defined in section B-3.2.1. Then, for $i = 1, \ldots, k$,*

$$(P\mathbb{F}L^{(i)}, \mathbb{F}L^{(i)})^\star\theta = \theta^{(i)}_L.$$

Proof: Let $([g]; v_1, \ldots, v_k)$ be an element of $P^k\mathbb{G}$ and consider

$$[(P\mathbb{F}L^{(i)}, \mathbb{F}L^{(i)})^\star\theta]_{[g]}([g]; v_1, \ldots, v_k) = \theta_{\mathbb{F}L^{(i)}([g])}(P\mathbb{F}L^{(i)}([g]; v_1, \ldots, v_k)). \tag{3.16}$$

Now, the canonical section θ is defined by the following rule: for $\alpha \in A^*G$ and (v, X_α) in $(P^{\tau^*}(AG))_\alpha$, we have that $\theta_\alpha(v, X_\alpha) = \alpha(v)$. Noting that

$$P\mathbb{F}L^{(i)}([g]; v_1, \ldots, v_k) = (v_i, \cdot)$$

(the precise form of the second argument doesn't matter), the right-hand side of (3.16) then becomes

$$\mathbb{F}L^{(i)}([g])(v_i) = \theta^{(i)}_L([g]; v_1, \ldots, v_k),$$

where the last equality follows by comparing the definition of the ith Poincaré-Cartan form with the ith Legendre transformation. ⋄

3.4. Variational interpretation of the Poincaré-Cartan forms. In this section, we derive a multisymplectic form formula for discrete field theories on Lie groupoids. The existence of such a formula is a consequence of the variational background of the discrete Euler-Lagrange equations and should therefore come as no surprise.

In fact, the analysis in this section is very similar to the developments in chapter 2, and to those in [**80**]. Conversely, when the target groupoid is $Q \times Q$, then our results reduce to those given before.

3.4.1. *Arbitrary variations.* Consider again a finite subset U_F of $\mathbb{X}^k$ and let ∂U_E be the boundary of the associated subset $U_E \subset V \times V$. Let $\phi : U_E \to G$ be a morphism defined on U_E.

We now define a finite *first variation* of ϕ as a one-parameter family of maps $h_\epsilon : U_V \to G$ such that $\alpha \circ h_\epsilon = \phi_{(0)}$ for all ϵ. In addition, we require that the composition of Φ_ϵ with ϕ is a solution of the discrete field equations for all ϵ. Here, $\Phi_\epsilon : \phi(U_E) \to G$ is defined as in (3.11). First variations hence resemble ordinary variations (see definition 3.5) except for the fact that they are not necessarily trivial on the boundary of the image of ϕ.

Associated to a finite first variation h_ϵ of ϕ is again an infinitesimal first variation V, which is a section of AG along $\phi_{(0)}$ defined by

$$V(x_{i,j}) = \frac{\mathrm{d}}{\mathrm{d}\epsilon} h_\epsilon(x_{i,j})\Big|_{\epsilon=0}.$$

When extremizing the action sum (3.9) under first variations, there are now two different kinds of contributions. The first comes from the interior of U_V, and yields the discrete Euler-Lagrange equations, as we saw before. The second is the contribution from the boundary ∂U_V, which takes the following form (with the notations of section 3.2):

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} S(\phi_\epsilon)\Big|_{\epsilon=0} = \sum_{[x]\cap \partial U_V \neq \varnothing} \left(\sum_{l;\alpha([x]_l)\in \partial U_V} \left(\theta_L^{(l)}(\psi([x])) \cdot V_{\psi([x])} \right) \right), \tag{3.17}$$

where V is the infinitesimal variation associated to φ_ϵ. Once again, we see how the Poincaré-Cartan forms arise naturally in the context of discrete Lagrangian field theories.

3.4.2. *Multisymplecticity.* By exactly the same reasoning as in chapter 2, we now derive a criterion for multisymplecticity. We will not repeat the entire proof, but we only highlight some of the key points. For more information, the reader is referred to [**80**].

Define $\mathcal{M}_d$ to be the manifold of morphisms $\phi : U_V \to G$ and consider the subset $\mathcal{S}_{d,L}$ of morphisms that solve the discrete field equations. We now erect a certain Lie algebroid over $\mathcal{S}_{d,L}$, denoted by $\mathcal{AG}$, and defined as follows:

$$\mathcal{AG}(\phi) = \Gamma(\phi^*_{(0)} AG) \quad \text{for all } \phi \in \mathcal{S}_{d,L}.$$

In other words, the fibre of $\mathcal{AG}$ over ϕ is the module of sections of the pullback bundle $\phi^*_{(0)}AG$. Formally, $\mathcal{AG}$ inherits the structure of a Lie algebroid from AG. Indeed, sections v and w of AG induce sections $\tilde{v}$ and $\tilde{w}$ of $\mathcal{AG}$ by composition: $\tilde{v}(\phi) := v \circ \phi_{(0)}$, and similarly for $\tilde{w}$. For this class of sections, the bracket and the anchor are determined by

$$[\tilde{v}, \tilde{w}]_{\mathcal{AG}}(\phi) := [v, w] \circ \phi_{(0)} \quad \text{and} \quad \rho_{\mathcal{AG}}(\tilde{v}) := \rho \circ \tilde{v}.$$

In the case of standard discrete field theories as in chapter 2, AG is the tangent bundle TQ, and $\mathcal{AG}$ is just $T\mathcal{M}_d$. By definition, an infinitesimal first variation V of a discrete field ϕ is an element of $\mathcal{AG}(\phi)$. On the other hand, not all such elements are first variations.

The action sum S can be interpreted as a function on $\mathcal{M}_d$, and by restriction also on $\mathcal{S}_{d,L}$. As $\mathcal{AG}$ is a Lie algebroid, its space of sections is equipped with an exterior derivative 'd'; after some thought, it can be seen that $\mathrm{d}S \cdot V$ can be expressed as

$$\mathrm{d}S \cdot V = \sum_{[x] \cap \partial U_V \neq \varnothing} \left(\sum_{l; \alpha([x]_l) \in \partial U_V} \left(\theta_L^{(l)}(\psi([x])) \cdot V_{\psi([x])} \right) \right).$$

By an argument similar to [**80**, thm. 4.1], it can then be shown that, for any $\phi \in \mathcal{S}_{d,L}$ and V_1, V_2 first variations of ϕ, the identity $\mathrm{d}^2 S(\phi)(V_1, V_2) \equiv 0$ can be written as

$$0 = \sum_{[x] \cap \partial U \neq \varnothing} \left(\sum_{l; \alpha^{(l)}([x]) \in \partial U} \left(\Omega_L^{(l)}(\psi([x]))(V_1, V_2) \right) \right), \tag{3.18}$$

which is the desired multisymplectic form formula (compare with 2.13).

4. Examples

In this section, we treat some special examples of Lie groupoid field theories. By taking for G the pair groupoid $Q \times Q$, we show that the discrete multisymplectic field theory of section 2 in chapter 2 is a special case of Lie groupoid field theory. Secondly, in subsection 4.2 we study the case of discrete fields taking values in a Lie group $\mathcal{G}$. (Recall from appendix B that Lie groups are particular examples of groupoids). These field theories go by the name of *Euler-Poincaré* theories and play an important role in the theory of reduction, which will be the topic of the next chapter.

4.1. Relation with the standard formalism. Throughout this section, G will be the pair groupoid $Q \times Q$. As we pointed out before, a morphism $\phi : V \times V \to Q \times Q$ can be seen as an assignment of an element of Q to each vertex in V. This is formalized in the next lemma.

Lemma 4.1. *Consider a morphism ϕ from $V \times V$ to $Q \times Q$ and let $f : V \to Q$ be the associated map between the sets of units. Then $\phi(x, y) = (f(x), f(y))$ for all $(x, y) \in V \times V$.*

Proof: If $\phi(x, y) = (q_0, q_1)$, then

$$q_0 = \alpha_Q(\phi(x, y)) = f(\alpha_V(x, y)) = f(x).$$

Similarly, $q_1 = f(y)$ and therefore $\phi(x, y) = (f(x), f(y))$ for all $(x, y) \in V \times V$. We conclude that ϕ is completely determined by specifying f: in the future, we will therefore identify ϕ and f. ◇

It is easy to see that $\mathbb{G}^k$ is just $Q^{\times k}$: the identification is given by

$$\big((q_k, q_1), (q_1, q_2), \ldots, (q_{k-1}, q_k)\big) \mapsto (q_1, q_2, \ldots, q_k).$$

Furthermore, the prolongation algebroid $P^k\mathbb{G}$ can be identified with the k-fold Cartesian product of TQ with itself. For a vector field v on Q, the ith tangent lift of v is the following section of $(TQ)^{\times k}$:

$$v^{(i)} : (q_1, q_2, \ldots, q_k) \mapsto (0, \ldots, 0, v(q_i), 0, \ldots, 0),$$

where $v(q_i)$ occupies the ith place.

Now, let $L : \mathbb{G}^k \to \mathbb{R}$ be a Lagrangian and denote by $\hat{L}$ the induced map on $Q^{\times k}$. The ith Poincaré-Cartan form then becomes

$$\theta_L^{(i)}(q_1, \ldots, q_k; v_1, \ldots, v_k) = \mathrm{d}\hat{L}(q_1, \ldots, q_{i-1}, \cdot, q_{i+1}, \ldots, q_k) \cdot v_i,$$

where $v_i \in T_{q_i}Q$ for $i = 1, \ldots, k$. This agrees with our original definition of the Poincaré-Cartan forms in (2.11).

It is instructive to see what becomes of the concepts of finite and infinitesimal variations in this case: an infinitesimal variation is just a vector field on Q defined along a discrete field $\phi : V \to Q$. A finite variation is then a one-parameter family of diffeomorphisms defined in a neighbourhood of $\operatorname{Im} \phi$.

For the case of the square mesh of figure 3.1, we may describe the field by assigning a value $\phi_{i,j} \in Q$ to each vertex (i, j). Let $\hat{L}(q_1, q_2, q_3, q_4)$ be a Lagrangian density; then $\{\phi_{i,j}\}$ is a solution of the field equations (3.15) associated to L if and only if, for all $(i, j) \in V$,

$$\begin{aligned}
&D_1L(\phi_{i,j}, \phi_{i+1,j}, \phi_{i+1,j+1}, \phi_{i,j+1}) + D_2L(\phi_{i-1,j}, \phi_{i,j}, \phi_{i,j+1}, \phi_{i-1,j+1}) + \\
&D_3L(\phi_{i-1,j-1}, \phi_{i,j-1}, \phi_{i,j}, \phi_{i-1,j}) + D_4L(\phi_{i,j-1}, \phi_{i+1,j-1}, \phi_{i+1,j}, \phi_{i,j}) = 0,
\end{aligned}$$

and we obtain the same discrete Euler-Lagrange equations as in chapter 2.

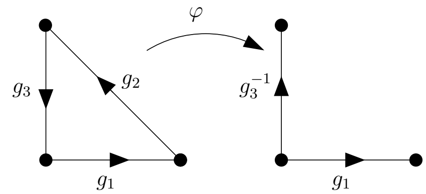


Figure 3.5. The map $\varphi : \mathbb{G}^3 \to \mathcal{G} \times \mathcal{G}$.

4.2. The Euler-Poincaré equations. Let $\mathcal{G}$ be a Lie group. As shown in example 1.1 of appendix B, $\mathcal{G}$ can be interpreted as a Lie groupoid over a singleton, and therefore one can study discrete fields taking values in $\mathcal{G}$ using the framework for Lie groupoid field theories developed in this chapter. The resulting equations are called the *Euler-Poincaré equations*. As we shall see in the next chapters, the Euler-Poincaré equations can also be derived using a symmetry reduction procedure.

Let $\mathbb{G}^3$ be the manifold of triangles corresponding to the Lie group $\mathcal{G}$. The elements of $\mathbb{G}^3$ are triples (g_1, g_2, g_3) such that $g_1 \cdot g_2 \cdot g_3 = e$ and $\mathbb{G}^3$ can therefore be identified with $\mathcal{G} \times \mathcal{G}$ using the following diffeomorphism (see also figure 3.5):

$$\varphi : \mathbb{G}^3 \to \mathcal{G} \times \mathcal{G}, \quad \text{with } \varphi(g_1, g_2, g_3) = (g_1, g_3^{-1}). \tag{3.19}$$

A few notational conventions are in order here. Let L be a Lagrangian on $\mathbb{G}^3$: we denote the induced function on $\mathcal{G} \times \mathcal{G}$ by ℓ; *i.e.* $\ell(g, h) = L(g, g^{-1}h, h^{-1})$. Secondly, if $\phi : V \times V \to \mathcal{G}$ is a morphism, then we use the following notation for the image of ϕ:

$$u_{i,j} := \phi((i,j), (i+1, j)) \quad \text{and} \quad v_{i,j} := \phi((i,j), (i, j+1)). \tag{3.20}$$

Let us now derive the discrete field equations for a field theory taking values in the Lie group $\mathcal{G}$. This derivation is similar to the one in the proof of theorem 3.7, but as the resulting field equations (the so-called *Euler-Poincaré* equations) will play an important role in the next two chapters, it is useful to repeat it for this special case.

Proposition 4.2. *A morphism $\phi : V \times V \to \mathcal{G}$ is an extremum for the action defined by L if and only if it satisfies the following set of* discrete Euler-Poincaré equations*:*

$$\begin{aligned} &\left[\left(R^*_{u_{i,j}} \mathrm{d}\ell(\cdot, v_{i,j})\right)_e - \left(L^*_{u_{i-1,j}} \mathrm{d}\ell(\cdot, v_{i-1,j})\right)_e\right] + \\ &\left[\left(R^*_{v_{i,j}} \mathrm{d}\ell(u_{i,j}, \cdot)\right)_e - \left(L^*_{v_{i,j-1}} \mathrm{d}\ell(u_{i,j-1}, \cdot)\right)_e\right] = 0. \end{aligned} \tag{3.21}$$

Proof: Let $h_\epsilon : U_V \to \mathcal{G}$ be a variation of a field ϕ as in definition 3.5. The varied action is then given by

$$S(\phi_\epsilon) = \sum_{(i,j) \in U} \ell(h_\epsilon(i,j)^{-1} u_{i,j} h_\epsilon(i+1, j), h_\epsilon(i,j)^{-1} v_{i,j} h_\epsilon(i, j+1)),$$

and therefore ϕ is an extremum if $\frac{\mathrm{d}}{\mathrm{d}\epsilon} S(\phi_\epsilon)\Big|_{\epsilon=0} = 0$, or explicitly,

$$\sum_{(i,j)\in U_V} \Big(\; D_1\ell(u_{i-1,j}, v_{i-1,j}) \cdot TL_{u_{i-1,j}}(V(i,j)) - D_1\ell(u_{i,j}, v_{i,j}) \cdot TR_{u_{i,j}}(V(i,j))$$
$$+ D_2\ell(u_{i,j-1}, v_{i,j-1}) \cdot TL_{v_{i,j-1}}(V(i,j)) - D_2\ell(u_{i,j}, v_{i,j}) \cdot TR_{v_{i,j}}(V(i,j))\Big) = 0,$$

where $V : U_V \to \mathfrak{g}$ is the infinitesimal variation associated to h_ϵ. The equations (3.21) follow immediately. ◇

In [**15**, **16**, **77**], and later also in [**75**], the authors derived equations for discrete mechanics on a Lie group $\mathcal{G}$, where $\mathcal{G}$ is also the symmetry group. These equations were also referred to as the *discrete Euler-Poincaré equations*. Roughly speaking one can recognize in (3.21) two copies of these equations, one for the "spatial" direction and one for the "time" direction.

Chapter 4

Symmetry for discrete Lagrangian field theories

In the preceding chapter, we established a Lie groupoid framework for discrete Lagrangian field theories. In this chapter and the next one, we will study the role of symmetry in this framework. The purpose of the present chapter is to point out a number of general aspects of symmetry; in the next chapter, we will then focus on a special case, namely the Euler-Poincaré equations of section 4.2 in the last chapter.

In section 1, we start by proving a reduction theorem for Lie groupoid field theories with symmetry. As we will show in section 1.3, this reduction theorem is "multisymplectic" in the sense that it maps multisymplectic field theories into new multisymplectic field theories. Finally, in section 2 we introduce the concept of a Noether symmetry, and we prove that every such symmetry gives rise to a conservation law.

A word of explanation is in order here concerning our definition of a symmetry action. In this chapter, we will assume that there exists a surjective Lie groupoid morphism Φ of the original Lie groupoid G to a new Lie groupoid G', which we call the *reduced groupoid*, and that there exists a *reduced Lagrangian* L' on $\mathbb{G}'^k$ such that the pullback of L' by Φ yields the original Lagrangian. The reduction theorem then relates the solutions of the original discrete Euler-Lagrange equations with those on the reduced Lie groupoid.

There is no better way to justify this abstract approach than by considering an example: let $\mathcal{G}$ be a Lie group and consider the pair groupoid $\mathcal{G} \times \mathcal{G}$. If L is a left $\mathcal{G}$-invariant Lagrangian, then the reduced Lie groupoid is $(\mathcal{G} \times \mathcal{G})/\mathcal{G}$, where $\mathcal{G}$ acts on $\mathcal{G} \times \mathcal{G}$ by the left diagonal action. Note that $(\mathcal{G} \times \mathcal{G})/\mathcal{G}$ is isomorphic to $\mathcal{G}$ by the isomorphism mapping $[(g, h)]$ to $g^{-1}h$. The Lie groupoid morphism Φ mentioned above is then just the quotient morphism:

$$\Phi : \mathcal{G} \times \mathcal{G} \to (\mathcal{G} \times \mathcal{G})/\mathcal{G} \cong \mathcal{G} \quad \text{where } \Phi(g, h) = g^{-1}h.$$

As we shall see later on, if L is left $\mathcal{G}$-invariant, then there exists a reduced Lagrangian on this reduced Lie groupoid, and the eventual effect of the reduction procedure is that we have eliminated the $\mathcal{G}$-symmetry. It should be noted that the equations thus obtained are the Euler-Poincaré equations: this will be the subject of section 1.1 below. However, a detailed study of these equations is postponed to the next chapter.

In this chapter, as well as in the next, we will work mostly with the triangular mesh described in chapter 2.

1. Lagrangian reduction

In this section, we turn to Lagrangian reduction in the context of discrete field theories with values in Lie groupoids. Although we have insisted on using a triangular mesh, it will be apparent that the reduction theorem 1.3, as well as the accompanying propositions 1.4 and 1.5, can be proved for general meshes as well.

Definition 1.1. *Let $\Phi : (G, Q) \to (G', Q')$ be a morphism of Lie groupoids. Associated to Φ there is a bundle map $\Psi : P^k\mathbb{G} \to P^k\mathbb{G}'$, whose base map $\underline{\Psi} : \mathbb{G}^k \to \mathbb{G}'^k$ and total space map $\overline{\Psi} : P^k\mathbb{G} \to P^k\mathbb{G}'$ are defined by*

$$\underline{\Psi}([g]) = (\Phi(g_1), \dots, \Phi(g_k)) \quad \text{and} \quad \overline{\Psi}([g]; v_1, \dots, v_k) = (\underline{\Psi}([g]); A\Phi(v_1), \dots, A\Phi(v_k)),$$

where $A\Phi$ is the Lie algebroid morphism induced by Φ (see appendix B).

It will often happen in this chapter that Φ is a submersion. In that case, $A\Phi$ is fibrewise surjective and we may find a local basis of sections $\{e_A\} = \{e_\alpha, e_a\}$ of AG and a local basis of sections $\{\bar{e}_\alpha\}$ of AG' adapted to $A\Phi$, i.e. such that

$$A\Phi \circ e_\alpha = \bar{e}_\alpha \circ \underline{\Phi} \quad \text{and} \quad A\Phi \circ e_a = 0. \tag{4.1}$$

In other words, the sections e_α are $A\Phi$-related to $\bar{e}_\alpha$ while the sections e_a are $A\Phi$-related to the zero section.

Lemma 1.2. *Let $\Phi : (G, Q) \to (G', Q')$ be a Lie groupoid morphism, which is a surjective submersion. The bundle map Ψ associated to Φ is a morphism of Lie algebroids.*

Proof: The proof is again an application of theorem 2.4 in appendix B.

Consider the basis of sections defined in the paragraph before. It follows from (4.1) that $(e_\alpha)_{(i)}$ is Ψ-related with $(\bar{e}_\alpha)_{(i)}$, and $(e_a)_{(i)}$ with the zero section. We now have to prove a similar property for the commutators $[(e_\alpha)_{(i)}, (e_\beta)_{(j)}]$, $[(e_\alpha)_{(i)}, (e_b)_{(j)}]$, and $[(e_a)_{(i)}, (e_b)_{(j)}]$.

For the first commutator, we have

$$\overline{\Psi}([(e_\alpha)_{(i)}, (e_\beta)_{(i)}]([g])) = \overline{\Psi}([e_\alpha, e_\beta]_{(i)}([g])) = [A\Phi([e_\alpha, e_\beta](\alpha(g_i))]_{(i)}([g]).$$

But $A\Phi$ is a Lie algebroid morphism and hence $A\Phi \circ [e_\alpha, e_\beta] = [\bar{e}_\alpha, \bar{e}_\beta] \circ \underline{\Phi}$. By plugging this into the last equation, we obtain

$$\overline{\Psi}([(e_\alpha)_{(i)}, (e_\beta)_{(i)}]([g])) = [\bar{e}_\alpha, \bar{e}_\beta]_{(i)}(\underline{\Psi}([g])).$$

The other commutators are easily seen to be zero. We now show that Ψ intertwines the anchor mappings of $P^k\mathbb{G}$ and $P^k\mathbb{G}'$. Let $([g]; v_1, \dots, v_k)$ be an element of $P^k\mathbb{G}$ and notice that the ith factor of

$$(\rho^{(k)} \circ \overline{\Psi})([g]; v_1, \dots, v_k) = \\ \left(\underline{\Psi}([g]); [A\Phi(v_1)]^R(g'_1) + [A\Phi(v_2)]^L(g'_2), \dots, [A\Phi(v_k)]^R(g'_k) + [A\Phi(v_1)]^L(g'_1)\right),$$

where $g_i' = \Phi(g_i)$, is just $(\hat{\rho} \circ P\Phi)(g_i; v_i, v_{i+1})$. Here, $P\Phi : PG \to PG'$ is the Lie algebroid morphism defined at the end of section 3.1.1 in appendix B. Because of the morphism properties of $P\Phi$, this is

$$(T\underline{\Phi} \circ \hat{\rho})(g_i; v_i, v_{i+1}) = T\underline{\Phi}(v_i^R(g_i) + v_{i+1}^L(g_i)).$$

The other components can be treated in a similar way and finally we end up with $\rho^{(k)\prime} \circ \overline{\Psi} = T\underline{\Psi} \circ \rho^{(k)}$. ◇

Theorem 1.3 (Reduction)**.** *Consider a morphism $\Phi : (G, Q) \to (G', Q')$, which is a surjective submersion. Furthermore, let $L : \mathbb{G}^3 \to \mathbb{R}$ be a Lagrangian on $\mathbb{G}^3$ and assume that there exists a reduced Lagrangian L' on $\mathbb{G}'^3$ such that $L = \Psi^\star L'$ on $\mathbb{G}'^3$, where Ψ is the map associated to Φ as in definition 1.1.*

A morphism $\phi : V \times V \to G$ is a solution of the discrete field equations for L if and only if the induced morphism $\Phi \circ \phi : V \times V \to G'$ satisfies the field equations for L'.

Proof: The proof relies on the following equality: for $i \leq k$, $[g] \in \mathbb{G}^3$, and $v \in A_qG$, where $q := \alpha(g_i)$,

$$v_{[g]}^{(i)}(L) = [A\Phi(v)]_{\Psi([g])}^{(i)}(L'), \tag{4.2}$$

which is relatively straightforward to prove. With the same notations as above, this implies that

$$\mathcal{E}_L([g_1], [g_2], [g_3]) \cdot v = \mathcal{E}_{L'}(\Psi([g_1]), \Psi([g_2]), \Psi([g_3])) \cdot (A\Phi(v)). \tag{4.3}$$

Here, $\mathcal{E}_L([g_1], [g_2], [g_3])$ is the left-hand side of the Euler-Lagrange equations (3.15) in chapter 3 (for the triangular mesh), and $[g_1]$, $[g_2]$ and $[g_3]$ are such that there exists an element $q \in Q$ such that $\alpha_{(i)}([g_i]) = q$ for $i = 1, 2, 3$ (as in the formulation of theorem 3.7 in the previous chapter). The subscript L indicates that we consider the Euler-Lagrange equations on $\mathbb{G}^k$ with respect to L. The expression $\mathcal{E}_{L'}$ is defined in a similar way, but now one considers the Euler-Lagrange equations on $\mathbb{G}'^3$ associated to L'.

Therefore, if ϕ is such that $\Phi \circ \phi$ is a solution of the Euler-Lagrange equations for L', then ϕ itself is a solution of the Euler-Lagrange equations for L.

Conversely, if Φ is a submersion, $A\Phi$ is surjective, and it follows from (4.3) that if ϕ is a solution of the field equations for L, then $\Phi \circ \phi$ is a solution of the field equations for L'. ◇

1.1. The Euler-Poincaré equations. Let $\mathcal{G}$ be a Lie group, and consider the pair groupoid $\mathcal{G} \times \mathcal{G}$. Recall from the last chapter (lemma 4.1) that discrete fields taking values in $\mathcal{G} \times \mathcal{G}$ can be identified with mappings from V to $\mathcal{G}$, and that the manifold of triangles $\mathbb{G}^3$ for the pair groupoid $\mathcal{G} \times \mathcal{G}$ is just the triple product $\mathcal{G}^{\times 3}$.

Assume that $L : \mathcal{G}^{\times 3} \to \mathbb{R}$ is left $\mathcal{G}$-invariant in the following sense:

$$L(g_1, g_2, g_3) = L(hg_1, hg_2, hg_3) \quad \text{for all } (g_1, g_2, g_3) \in \mathcal{G}^{\times 3} \quad \text{and} \quad h \in \mathcal{G}. \tag{4.4}$$

Such a Lagrangian induces a Lagrangian ℓ on $\mathcal{G} \times \mathcal{G}$ as follows: $\ell(g_1^{-1}g_2, g_1^{-1}g_3) = L(g_1, g_2, g_3)$. Similarly, a discrete field $\phi : V \to \mathcal{G}$ gives rise to a reduced field $\varphi : V \times V \to \mathcal{G}$ defined as $\varphi(x_0, x_1) := \phi(x_0)^{-1}\phi(x_1)$.

Let us now show how the discrete Euler-Poincaré equations fit into the framework of Lagrangian reduction. Let $\mathcal{G}$ be a Lie group and consider the pair groupoid $\mathcal{G} \times \mathcal{G}$ over $\mathcal{G}$. Let $\Phi : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ be the morphism defined as $\Phi(g, h) = g^{-1}h$. Let L be a $\mathcal{G}$-invariant Lagrangian on $\mathcal{G}^{\times 3}$ as in (4.4) and consider the induced Lagrangian ℓ on $\mathcal{G} \times \mathcal{G}$.

Let $\phi : V \times V \to \mathcal{G} \times \mathcal{G}$ be a discrete field, and $\varphi := \Phi \circ \phi : V \times V \to \mathcal{G}$ the reduced field. According to theorem 1.3, ϕ is a solution to the Euler-Lagrange equations associated to L if and only if φ is a solution to the field equations associated to ℓ. We will come back to this situation once we have proved theorem 2.1 in the next chapter.

Reduction in discrete field theories is thus very similar to the corresponding theory in mechanics. We glossed over some subtle differences, however, mainly related to the reconstruction problem. This will be treated in more detail in the next chapter. Briefly speaking, not every solution of the reduced field equations is necessarily of the form $\varphi = \Phi \circ \phi$.

1.2. The reduced Poincaré-Cartan forms. Let $L : \mathbb{G}^3 \to \mathbb{R}$ be a Lagrangian, $\Phi : (G, Q) \to (G', Q')$ a morphism of Lie groupoids, and assume that there exists a reduced Lagrangian L' on $\mathbb{G}'^k$ such that $L = \Psi^\star L'$ (see also the formulation of theorem 1.3). The Poincaré-Cartan forms associated to L are then related to those associated to L', as shown in the following proposition.

Proposition 1.4. *Let Ψ be the bundle map associated to the morphism Φ as in definition 1.1. Then $\Psi^\star \theta_{L'}^{(i)} = \theta_L^{(i)}$, as well as $\Psi^\star \Omega_{L'}^{(i)} = \Omega_L^{(i)}$.*

Proof: For all $([g]; v_1, v_2, v_3) \in P^3\mathbb{G}$, we have that

$$\begin{aligned} \Psi^\star \theta_{L'}^{(i)}([g]; v_1, v_2, v_3) &= \theta_{L'}^{(i)}\big(\underline{\Psi}([g]); A\Phi(v_1), \ldots, A\Phi(v_3)\big) \\ &= [A\Phi(v_i)]^{(i)}_{\underline{\Psi}([g])}(L') = v^{(i)}_{[g]}(L) \\ &= \theta_L^{(i)}([g]; v_1, v_2, v_3), \end{aligned}$$

where, in the third step, we have used (4.2). The corresponding statement for $\Omega_L^{(i)}$ and $\Omega_{L'}^{(i)}$ then follows immediately, using the fact that Ψ is a morphism of Lie algebroids (lemma 1.2). ◇

1.3. Multisymplecticity of the reduced equations. We now show that the reduced field equations also conserve multisymplecticity. Of course, this is obvious if we think of the variational derivation of the multisymplectic form formula: since the reduced equations are also variational, it follows that a similar formula can be derived for these equations as well. Here, we follow a different route, and show that the multisymplectic form formula for the reduced equations follows from the original formula for the unreduced equations.

Proposition 1.5. *Assume that Φ is a submersion, and that the solutions of the discrete field equations derived from L are multisymplectic. Then the same holds for the reduced solutions of the field equations associated to L'.*

Proof: The solutions of the discrete field equations are multisymplectic if (3.18) from chapter 3 holds. Using proposition 1.4, we may rewrite the summand as

$$\Omega_L^{(l)}(\psi([x]))(V_1, V_2) = \Omega_{L'}^{(l)}(\psi'([x]))(\overline{\Psi}(V_1), \overline{\Psi}(V_2)),$$

where $\psi' = \underline{\Psi} \circ \psi$ is a reduced solution. Plugging this into (3.18) then shows that ψ' is multisymplectic with respect to first variations of the form $\overline{\Psi}(V)$. However, if Φ is a submersion, then all reduced first variations are of this form. ◇

Remark 1.6. As in continuum field theory (see [**23**]), it is possible that not all solutions of the reduced equations are of the form $\psi' = \underline{\Psi} \circ \psi$, where ψ is a solution of the unreduced problem. For this class of solutions, proposition 1.5 need not be true. ◇

2. The Noether theorem

Let $L : \mathbb{G}^3 \to \mathbb{R}$ be a discrete Lagrangian. We now turn to the *Noether symmetries* of L. In this context a Noether symmetry is a section v of AG with a number of properties listed below in definition 2.1. Briefly speaking, these amount to asking that L be invariant under the flow of v, up to a term which does not contribute to the discrete Euler-Lagrange equations.

Definition 2.1. *A section v of AG is a Noether symmetry of the Lagrangian L if there exist functions f_2, f_3 on Q such that*

$$v^{(1)}(L) + v^{(2)}(L) + v^{(3)}(L) = (\alpha^{(2)*} f_2 - \alpha^{(1)*} f_2) + (\alpha^{(3)*} f_3 - \alpha^{(1)*} f_3).$$

Just as in the continuous case, each such Noether symmetry gives rise to a conservation law. This is the content of the celebrated Noether theorem, proved in the context of discrete Lagrangian field theories in theorem 2.2. For an overview of symmetries in classical field theory, see [**35**].

Theorem 2.2 (Discrete Noether theorem)**.** *Let v be a Noether symmetry of L and consider the functions*

$$\eta^{(x)} = \theta_L^{(2)}(\omega) - \alpha^{(2)*} f_2 \quad \text{and} \quad \eta^{(y)} = \theta_L^{(3)}(\omega) - \alpha^{(3)*} f_3,$$

where ω is the section of $P^3\mathbb{G}$ defined by

$$\omega([g]) := \left([g]; v(\alpha^{(1)}([g])), v(\alpha^{(2)}([g])), v(\alpha^{(3)}([g]))\right).$$

Then for any solution ϕ of the discrete Euler-Lagrange equations, the following conservation law holds:

$$\eta^{(x)}([g_1]) - \eta^{(x)}([g_2]) + \eta^{(y)}([g_1]) - \eta^{(y)}([g_3]) = 0, \tag{4.5}$$

where $[g_1], [g_2], [g_3]$ are triangles in the image of ϕ such that $\alpha^{(1)}([g_1]) = \alpha^{(2)}([g_2]) = \alpha^{(3)}([g_3])$.

Proof: We have that

$$\begin{aligned}\eta^{(x)}([g_1]) &- \eta^{(x)}([g_2]) + \eta^{(y)}([g_1]) - \eta^{(y)}([g_3]) \\ &= \theta_L^{(2)}([g_1])(\omega) - f_2(\alpha^{(2)}([g_1])) - \theta_L^{(2)}([g_2])(\omega) + f_2(\alpha^{(2)}([g_2])) \\ &+ \theta_L^{(3)}([g_1])(\omega) - f_3(\alpha^{(3)}([g_1])) - \theta_L^{(3)}([g_3])(\omega) + f_3(\alpha^{(3)}([g_3])).\end{aligned} \tag{4.6}$$

Note that $\theta_L^{(i)}([g])(\omega) = v_{[g]}^{(i)}(L)$. The section v is a Noether symmetry, and therefore

$$\begin{aligned}&\theta_L^{(2)}([g_1])(\omega) + \theta_L^{(3)}([g_1])(\omega) = \\ &\quad - \theta_L^{(1)}([g_1])(\omega) + f_2(\alpha^{(2)}([g_1])) - f_2(\alpha^{(1)}([g_1])) + f_3(\alpha^{(3)}([g_1])) - f_3(\alpha^{(1)}([g_1])).\end{aligned}$$

Substituting this in (4.6), we finally obtain

$$\eta^{(x)}([g_1]) - \eta^{(x)}([g_2]) + \eta^{(y)}([g_1]) - \eta^{(y)}([g_3]) = \mathcal{E}_L([g_1], [g_2], [g_3]) \cdot v(q), \tag{4.7}$$

where $q = \alpha^{(1)}([g_1]) = \alpha^{(2)}([g_2]) = \alpha^{(3)}([g_3])$. ◇

In the next chapter, we will show that if the Lagrangian L is $\mathcal{G}$-invariant as in (4.4), then the conservation law given by Noether's theorem and the Euler-Poincaré equations are intimately related.

Chapter 5

Euler-Poincaré reduction for discrete field theories

We now focus on a special class of discrete Lagrangian field theories with symmetry, namely the Euler-Poincaré theories. These field theories were first derived in chapter 3; in chapter 4, we showed that they also arise through a reduction process. In this chapter, we take a closer look at the reduction theorem and the field equations for field theories taking values in a *Lie group* $\mathcal{G}$. In this case, alternative interpretations of the concept 'field' are possible. If $\mathcal{G}$ is Abelian, then a field can be viewed as a $\mathcal{G}$-valued cochain on the mesh. On the other hand, if $\mathcal{G}$ is nonabelian, then a new interpretation suggests itself: that of a discrete $\mathcal{G}$-connection (definition 1.5 below).

The latter is particularly fruitful: in this way, we stay close to continuous framework (see [**23**, **24**]) where the reduced fields are also (continuous) connections. Moreover, it is known that such a reduced field can be "reconstructed" to a solution of the original, unreduced problem if and only if the reduced field has curvature zero. In section 2.3, we show that a similar obstruction in terms of discrete curvature arises in the reconstruction of the discrete Euler-Poincaré equations.

In section 3, we then take an alternative route to the Euler-Poincaré equations: inspired by a similar treatment in [**23**], we show that the Noether theorem of the last chapter yields a conservation law which is equivalent to the Euler-Poincaré equations. The key to this identification is the set of discrete Legendre transformations from chapter 3.

As a modest final application, we consider (from section 4 onwards) the Lagrangian of harmonic mappings into a Lie group: we propose an extension of the well-known Moser-Veselov algorithm, demonstrate its equivalence to the Euler-Poincaré equations, and establish a special form of the field equations of discrete harmonic mappings, using the concepts of discrete geometry established earlier.

1. Discrete differential geometry

Some elementary concepts of discrete differential geometry come up quite naturally in the study of discrete fields taking values in a Lie group $\mathcal{G}$. Recall that a discrete field taking values in the pair groupoid $\mathcal{G} \times \mathcal{G}$ (an "unreduced field" as in section 1.1 in the previous chapter) can be identified with a map from V to $\mathcal{G}$ (lemma 4.1 in chapter 3). On the other hand, reduced fields are maps from the set of edges E to $\mathcal{G}$. As we shall see in this section, if $\mathcal{G}$ is Abelian, unreduced and reduced fields correspond to discrete

0-forms and discrete 1-forms, respectively. This is the main reason for the introduction of discrete geometry in the study of Euler-Poincaré reduction. Moreover, as we shall see later on, if $\mathcal{G}$ is nonabelian, then reduced fields have a natural interpretation as discrete $\mathcal{G}$-connections.

1.1. Discrete differential forms. This section is dedicated to a review of some elementary concepts of algebraic topology, but for the sake of clarity we only introduce what is strictly necessary for the developments in subsequent sections. Almost everything in this section can be extended to much more general settings; the reader wishing to do so is referred to [**39**, **56**, **112**], or to the text book [**53**].

Consider again a mesh (V, E) in $\mathbb{R}^2$ as in definition 1.1 in chapter 3. The collection of sets $\{V, E, F\}$ together with its various incidence relations determines a CW-complex (see [**53**, p. 5]) and this leads us naturally to the concepts of homology and cohomology. It is therefore not unreasonable to expect that some of these concepts will resurface in our study of discrete field theories later on.

A fundamental concept in topology is that of an *n-chain*. This is a formal linear combination (with coefficients in $\mathbb{R}$) of "n-dimensional elements". More precisely, the vector space of 0-chains consists of finite linear combinations of elements of V:

$$C_0 = \{\alpha_1 x_1 + \cdots + \alpha_m x_m : \alpha_1, \ldots, \alpha_m \in \mathbb{R}, x_1, \ldots, x_m \in V\}, \tag{5.1}$$

where it should be stressed that the elements of C_0 are *formal* linear combinations of elements of V. Similarly, the vector space C_1 is generated by elements of E, and C_2 by elements of F.

Remark 1.1. We recall that each edge in E is realized as a segment of a straight line in $\mathbb{R}^2$, and hence is determined by its begin and end vertex. This will be reflected in our notation by writing an edge $\mathfrak{e}$ simply as an ordered pair (x_0, x_1). In particular, the fact that $\mathfrak{e} - \mathfrak{e} = 0$ in C_1 for any edge $\mathfrak{e} = (x_0, x_1)$ implies that $-\mathfrak{e}$ can be identified with (x_1, x_0). ◇

It is customary to define *discrete n-forms* as n-dimensional cochains, *i.e.* elements of the dual vector space C_n^*. From this definition, it follows immediately that a discrete zero-form induces a function $\phi : V \to \mathbb{R}$. Conversely, such a function gives rises to a zero-form through linear extension.

Similarly, in view of remark 1.1, discrete one-forms can be identified with functions $\varphi(x_0, x_1)$ on the set of edges E. In contrast to what this notation may suggest, it should be borne in mind that these functions are not necessarily defined on the whole of $V \times V$, but only on the subset of edges E. We note that $\varphi(x_0, x_1) = -\varphi(x_1, x_0)$ for any edge (x_0, x_1).

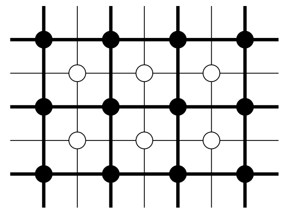


Figure 5.1. Square mesh (black) and its dual (light).

We continue in a similar vein by defining discrete two-forms as functions $\psi(q_1, \ldots, q_k)$ on the set of faces. Again, these functions can be extended unambiguously by linearity to a proper cochain. To summarize, we have the following definition:

Definition 1.2. *For $n = 0, 1, 2$, a* discrete n-form *is a linear map $f : C_n \to \mathbb{R}$. For $n > 2$, all discrete n-forms are zero. The set of all discrete n-forms is denoted by C_n^*.*

For the sake of self-containedness, we recall the explicit form of the coboundary operator $\mathrm{d} : C_n^* \to C_{n+1}^*$. For a zero-form ϕ, $\mathrm{d}\phi(q_0, q_1) = \phi(q_1) - \phi(q_0)$. For a one-form φ,

$$\mathrm{d}\varphi(q_1, \ldots, q_k) = \varphi(q_1, q_2) + \varphi(q_2, q_3) + \cdots + \varphi(q_{k-1}, q_k). \tag{5.2}$$

The coboundary of a discrete two-form is defined to be zero. We will sometimes refer to d as the "discrete differential".

1.2. The discrete Hodge star. As could be expected from the continuous theory, the discrete Hodge star $\star$, to be introduced below, maps discrete n-forms into $(2-n)$-forms. However, there is an additional complication in the discrete case: the forms $\star f$ are not defined on the mesh itself, but rather on a *dual mesh*, which we now define.

The *dual mesh* (V^*, E^*) is constructed as follows. For every face in F, there is a vertex in V^* (for which one usually takes the barycentric dual; see [**39**]). There is an edge in E^* between two vertices $q_0, q_1 \in V^*$ if and only if the faces in F corresponding to q_0 and q_1 have an edge in common. This determines the sets V^* and E^*; the set F^* consists of the faces of this dual graph. It is easy to see that to each face in F^*, there corresponds a vertex in V. See figure 5.1 for an illustration.

Implicit in this definition is the existence of a duality operator $*$ between n-chains on the mesh and $(2-n)$-chains on the dual mesh. This duality is well defined, but only up to orientation. We now use the orientation of $\mathbb{R}^2$ to settle this point. The definition used here agrees with the algorithm for the orientation of dual cells proposed by Hirani (see [**56**, remark 2.5.1]).

We define $*$ first on the elements of V, E, and F. By linearity, it will then be determined on the whole of C_n. There is no ambiguity in determining the orientation of the dual vertex $*\mathfrak{f}$ of a face $\mathfrak{f}$. We define the dual $*x$ of a vertex x to be the corresponding face in F^*: $*x = (r_1, r_2, r_3, r_4)$, with the natural orientation, *i.e.* the vertices are listed in anticlockwise order. Finally, the definition of $*$ on E is slightly more intricate; here, we follow [**56**]. Let (x_0, x_1) be an edge in E and let $\{r_0, r_1\}$ be the corresponding dual edge, considered as an unordered set. The line segments $[x_0, x_1]$ and $[r_0, r_1]$ determine a basis of $\mathbb{R}^2$: if this basis is positively oriented, then $*(x_0, x_1) = (r_0, r_1)$, otherwise, $*(x_0, x_1) = (r_1, r_0)$. In the case of the square mesh of figure 5.1, the action of $*$ on E corresponds to an anticlockwise rotation over $\pi/2$.

On the dual mesh, one can again introduce discrete forms. We will denote the vector space of discrete n-forms on the dual mesh by D_n^*.

In the case of the square mesh of figure 5.1, the dual mesh is again square and hence there is a natural way to extend $*$ to an operator from (V^*, E^*) to (V, E). It is then easy to check that $**v = (-1)^{n(2-n)}v$ for any $v \in C_n$.

Definition 1.3. *The* discrete Hodge star $\star : C_n^* \to D_{2-n}^*$ *is defined by*

$$(\star\alpha)(*v) = \alpha(v).$$

The definition given here is (up to a constant) a special case of the one proposed in [**39**]. Note that it follows immediately that $\star\star\alpha = (-1)^{n(2-n)}\alpha$, where we have defined the Hodge star on D_n^* as in definition 1.3, but using the duality operator $*$ defined on (V^*, E^*).

With the discrete Hodge star and the coboundary operator of the previous paragraph, we now arrive at the definition of the discrete codifferential.

Definition 1.4. *Let α be a discrete n-form. Then the* discrete codifferential $\delta\alpha$ *is the discrete $(n-1)$-form $\delta\alpha$ defined as $\delta\alpha = \star \mathrm{d} \star \alpha$.*

It is useful to write out this definition for a few explicit cases. If φ is a discrete one-form, then $\delta\varphi$ is given by

$$(\delta\varphi)(x) = \varphi(x_1, x) + \varphi(x_2, x) + \varphi(x_3, x) + \varphi(x_4, x),$$

where x_1, x_2, x_3, x_4 are the end points of the edges that emanate from x. In other words, $\delta\varphi$ assigns to each vertex x the sum of contributions from φ on the edges that have x as a vertex. Finally, for a discrete two-form ψ we note that $\delta\psi$ is the discrete one-form given by

$$\delta\psi(x_0, x_1) = \psi(\mathfrak{f}_0) - \psi(\mathfrak{f}_1),$$

where $\mathfrak{f}_0$ and $\mathfrak{f}_1$ are the faces that have (x_0, x_1) as a common edge, and where $\mathfrak{f}_0$ is the face where the orientation of the boundary edges agrees with the ordering of (x_0, x_1), whereas $\mathfrak{f}_1$ is the face with the opposite ordering.

1.3. Discrete connections. In the preceding sections, we introduced discrete one-forms as assignments of a real number to each edge $\mathfrak{e} \in E$. This theory can be extended in a straightforward way to discrete forms taking values in an arbitrary *Abelian* Lie group $\mathcal{G}$, the only significant difference being that we have to redefine the spaces of n-chains C_n as consisting of formal linear combinations with coefficients in $\mathbb{Z}$.

For instance, if $\varphi : C_1 \to \mathcal{G}$ is a discrete one-form, then $\mathrm{d}\varphi$ is determined by its action on the set F by (5.2) and can be extended by linearity to yield a map from C_2 to $\mathcal{G}$, where it should be borne in mind that the elements of C_2 are still formal linear combinations of elements in F, but now with coefficients in $\mathbb{Z}$.

The theory of discrete forms with values in an Abelian Lie group will be used in section 4, but in the general case, we will be confronted with mappings from E to a *non-Abelian* Lie group $\mathcal{G}$. Such maps can no longer be interpreted as discrete one-forms. Luckily, it turns out that these maps have a natural interpretation as *discrete $\mathcal{G}$-connections*, which we now define.

Definition 1.5. *A* discrete $\mathcal{G}$-connection *is a map $\omega : E \to \mathcal{G}$, such that, for all edges $\mathfrak{e} \in E$, $\omega(\mathfrak{e}^{-1}) = \omega(\mathfrak{e})^{-1}$. The* curvature *of such a connection is the map $\Omega : F \to \mathcal{G}$ defined as $\Omega(\mathfrak{f}) = \omega(\mathfrak{e}_1) \cdots \omega(\mathfrak{e}_k)$, where $\mathfrak{e}_1, \ldots, \mathfrak{e}_k$ are the boundary edges of the face $\mathfrak{f}$. A discrete $\mathcal{G}$-connection is said to be* flat *if $\Omega(\mathfrak{f}) = e$ for all $\mathfrak{f} \in F$.*

Note that in the case of a non-flat connection, $\Omega(\mathfrak{f})$ depends not only on $\mathfrak{f}$, but also on the exact representation of $\mathfrak{f}$ as a set of edges $\mathfrak{e}_1, \ldots, \mathfrak{e}_4$ (any cyclic permutation of this set represents the same face). However, this indeterminacy does not occur for flat connections, the only case that we will consider later on.

The theory of discrete $\mathcal{G}$-connections closely mimics the usual theory of connections. As an example, we mention the following proposition, from which a number of interesting properties may be deduced.

Proposition 1.6. *Consider a discrete $\mathcal{G}$-connection $\omega : E \to \mathcal{G}$. If ω is flat, then there exists a unique mapping $\phi : V \times V \to \mathcal{G}$ such that $\phi_{|E} = \omega$.*

Proof: See [**108**, prop. 7]. Note that this theorem remains trivially unchanged if we replace $\mathbb{R}^2$ by any other simply connected manifold. ◇

There are some immediate consequences of this proposition that are worth mentioning. Let $\omega : E \to \mathcal{G}$ be any discrete $\mathcal{G}$-connection. In particular, ω need not be flat. Then, ω induces a morphism $\hat{\omega}$ from the groupoid of paths $\mathcal{P}$ to $\mathcal{G}$ as follows:

$$\hat{\omega}(\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_m) = \omega(\mathfrak{e}_1)\omega(\mathfrak{e}_2) \cdots \omega(\mathfrak{e}_m).$$

Here, the *groupoid of paths* $\mathcal{P}$ is the set of paths in E, *i.e.* sequences of composable elements $\mathfrak{e}_1, \mathfrak{e}_2, \ldots, \mathfrak{e}_m$, equipped with the natural source and target mappings.

We call $\hat{\omega}$ the "discrete holonomy mapping". Note that if ω is a flat connection, $\hat{\omega}$ maps closed paths to the unit in $\mathcal{G}$. Furthermore, in the case of a flat connection, it can be easily established that simplicially homotopic paths have the same image under $\hat{\omega}$. This is the discrete counterpart of a well-known theorem of continuous connections (see [**59**, p. 93]): if ω is a flat connection, then any two closed homotopic loops have the same holonomy.

In the case of a flat connection, $\hat{\omega}$ descends to a map from Π, the *path groupoid*, consisting of paths in E modulo simplicial homotopy (keeping the end points fixed), to $\mathcal{G}$. For a simply connected manifold, Π is isomorphic to $V \times V$ and we have established the existence of a map $\hat{\omega} : V \times V \to \mathcal{G}$. This is basically the proof of proposition 1.6.

Remark 1.7. The concept of discrete $\mathcal{G}$-connections used here is common in lattice gauge theories (see [**4**, **113**]). A related concept was put forward by Novikov in [**88**]. $\diamond$

2. Discrete Euler-Poincaré reduction

In this section, we pick up the thread from section 1.1 in chapter 4 and continue our investigation of Euler-Poincaré reduction.

We consider again fields ϕ taking values in the pair groupoid $\mathcal{G} \times \mathcal{G}$ (where $\mathcal{G}$ is a Lie group). Recall from lemma 4.1 in chapter 3 that such fields can be identified with maps that associate an element of $\mathcal{G}$ to each vertex. After reduction by the natural left action of $\mathcal{G}$ on $\mathcal{G} \times \mathcal{G}$, they induce mappings φ that associate a group element to each edge. Explicitly, if $\phi : V \to \mathcal{G}$ is an unreduced field, then the reduced field $\varphi : E \to \mathcal{G}$ is given by $\varphi(\mathfrak{e}) = \phi(x_0)^{-1}\phi(x_1)$, where $\mathfrak{e} = (x_0, x_1)$. From the last section, we know that such maps have a natural interpretation as discrete $\mathcal{G}$-connections.

In the forthcoming theorem 2.1, it is shown how the field equations for the unreduced fields ϕ are equivalent to the discrete Euler-Poincaré equations for the reduced fields φ. Both sets of equations arise by extremizing a certain action functional. In theorem 2.4, we deal with the reconstruction problem. Starting from a reduced field $\varphi : E \to \mathcal{G}$, it is shown that φ gives rise to a solution $\phi : V \to \mathcal{G}$ of the original field equations if and only if the curvature of φ vanishes. This treatment was inspired by the work of Castrillón *et al.* [**24**], who developed Lagrangian reduction for field theories in the continuous case.

2.1. Review: discrete fields. Let $\mathcal{G}$ be an arbitrary Lie group. Recall from lemma 4.1 in chapter 4 that morphisms from $V \times V$ into $\mathcal{G} \times \mathcal{G}$ can be identified with maps $\phi : V \to \mathcal{G}$ assigning a value $\phi_{i,j} = \phi(x_{i,j})$ in $\mathcal{G}$ to each vertex $x_{i,j}$. Both interpretations will be used interchangeably throughout the remainder of this chapter.

The Lie group $\mathcal{G}$ has a natural diagonal action by left translations on the groupoid $\mathcal{G} \times \mathcal{G}$: $g \cdot (g_1, g_2) = (gg_1, gg_2)$. As the quotient groupoid $(\mathcal{G} \times \mathcal{G})/\mathcal{G}$ is naturally isomorphic to the Lie group $\mathcal{G}$ itself, reduced fields are defined to be morphisms $\varphi : V \times V \to \mathcal{G}$ attaching a group element to each edge in E.

Let $\Phi : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ be the morphism defined as $\Phi(g, g') = g^{-1}g'$. A morphism $\phi : V \to \mathcal{G}$ induces a reduced field $\varphi : V \times V \to \mathcal{G}$, defined as $\varphi = \Phi \circ (\phi \times \phi)$, or explicitly, by $\varphi(\mathfrak{e}) = \phi(x_0)^{-1}\phi(x_1)$, where $\mathfrak{e} = (x_0, x_1)$. This reduced field can be described as an assignment of a group element $u_{i,j}$ to each "vertical" edge $((i, j+1), (i, j))$, and of a group element $v_{i,j}$ to each "horizontal" edge $((i, j), (i+1, j))$, where

$$u_{i,j} = \phi_{i,j}^{-1}\phi_{i+1,j} \quad \text{and} \quad v_{i,j} = \phi_{i,j}^{-1}\phi_{i,j+1}. \tag{5.3}$$

See also (3.20). We will use these notations throughout this chapter.

For the pair groupoid $\mathcal{G} \times \mathcal{G}$, the manifold of triangles $\mathbb{G}^3$ is naturally isomorphic to the triple product $\mathcal{G}^{\times 3}$. In this case, a discrete Lagrangian is therefore just a function $L : \mathcal{G}^{\times 3} \to \mathbb{R}$. On the other hand, the set of triangles $\mathbb{G}'^3$, associated to the Lie group $\mathcal{G}$ (viewed as a Lie groupoid), is easily seen to be diffeomorphic to $\mathcal{G} \times \mathcal{G}$ (this is the isomorphism (3.19), as depicted on figure 3.5). For the remainder of this chapter, the identification of $\mathbb{G}^3$ with $\mathcal{G}^{\times 3}$, and of $\mathbb{G}'^3$ with $\mathcal{G} \times \mathcal{G}$, will be understood.

Let us now turn to the prolongation algebroid $P^k\mathbb{G}$ over $\mathbb{G}^k$. In the case of the pair groupoid, this algebroid is just the Cartesian product $(T\mathcal{G})^{\times 3}$. In the case of a Lie group $\mathcal{G}$, the prolongation is $(\mathcal{G} \times \mathcal{G}) \times (\mathfrak{g} \oplus \mathfrak{g} \oplus \mathfrak{g})$. The bundle map Ψ associated to the morphism Φ (as in definition 1.1 in chapter 4) then has a particularly simple interpretation: $\underline{\Psi} : \mathcal{G}^{\times 3} \to \mathcal{G} \times \mathcal{G}$ is given by $\underline{\Psi}(g_1, g_2, g_3) = (g_1^{-1}g_2, g_1^{-1}g_3)$, while the map $\overline{\Psi}$ between the total spaces is just left translation:

$$\overline{\Psi}(g_1, g_2, g_3; v_1, v_2, v_3) = (g_1^{-1}g_2, g_1^{-1}g_3; TL_{g_1^{-1}}(v_1), TL_{g_2^{-1}}(v_2), TL_{g_3^{-1}}(v_3)). \tag{5.4}$$

In the remainder of this chapter, we wish to study the situation where a discrete Lagrangian $L : \mathcal{G}^{\times 3} \to \mathbb{R}$ is given, which is invariant under the natural diagonal left action of $\mathcal{G}$ on $\mathcal{G}^{\times 3}$. In that case, L gives rise to a reduced Lagrangian $\ell : \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ defined by

$$\ell(g_1^{-1}g_2, g_1^{-1}g_3) = L(g_1, g_2, g_3).$$

Note that, in line with the general theory of chapter 4, ℓ and L are related as follows: $\ell \circ \underline{\Psi} = L$.

2.2. The reduction problem. Given a discrete Lagrangian L, the discrete action sum S is given by

$$S(\phi) = \sum_{(i,j)} L(\phi_{i,j}, \phi_{i+1,j}, \phi_{i,j+1}),$$

where ϕ is a map from V to $\mathcal{G}$. Recall that ϕ is an extremum of this action if and only ϕ satisfies the discrete Euler-Lagrange equations (3.15). In the case of a triangular mesh, these equations are given by

$$D_1L(\phi_{i,j},\phi_{i+1,j},\phi_{i,j+1})+D_2L(\phi_{i-1,j},\phi_{i,j},\phi_{i-1,j+1})+D_3L(\phi_{i,j-1},\phi_{i+1,j-1},\phi_{i,j})=0. \quad (5.5)$$

Similarly, we may define the reduced action sum s as

$$s(\varphi)=\sum_{(i,j)}\ell(u_{i,j},v_{i,j}).$$

A morphism $\varphi : V\times V\to\mathcal{G}$ is an extremum of s if and only if it satisfies the discrete Euler-Poincaré equations derived in proposition 4.2 of chapter 4. The central aspects of discrete Euler-Poincaré reduction are summarized in the following theorem. This theorem, as well as its proof, are very similar to the discrete reduction process in mechanics (see [**77**]). Moreover, this theorem is in fact a special case of theorem 1.3 in chapter 4.

Theorem 2.1 (Reduction)**.** *Let L be a $\mathcal{G}$-invariant Lagrangian on $\mathcal{G}^{\times 3}$ and consider the reduced Lagrangian ℓ on $\mathcal{G}\times\mathcal{G}$. Consider a discrete field $\phi : V\to\mathcal{G}$ and let $\varphi : V\times V\to\mathcal{G}$ be the associated reduced field defined as $\varphi=\Phi\circ(\phi\times\phi)$. Then the following are equivalent:*

(a) *ϕ is a solution of the discrete Euler-Lagrange equations for L;*
(b) *ϕ is an extremum of the action sum S for arbitrary variations;*
(c) *the reduced morphism φ is a solution of the discrete Euler-Poincaré equations:*

$$\begin{aligned}&\left[\left(R^*_{u_{i,j}}\mathrm{d}\ell(\cdot,v_{i,j})\right)_e-\left(L^*_{u_{i-1,j}}\mathrm{d}\ell(\cdot,v_{i-1,j})\right)_e\right]+\\&\left[\left(R^*_{v_{i,j}}\mathrm{d}\ell(u_{i,j},\cdot)\right)_e-\left(L^*_{v_{i,j-1}}\mathrm{d}\ell(u_{i,j-1},\cdot)\right)_e\right]=0;\end{aligned}\quad (5.6)$$

(d) *the reduced morphism φ is an extremum of the reduced action sum s for variations of the form*

$$\delta u_{i,j}=TR_{u_{i,j}}(\theta_{i,j+1})-TL_{u_{i,j}}(\theta_{i,j})\in T_{u_{i,j}}\mathcal{G} \quad (5.7)$$

and

$$\delta v_{i,j}=TR_{v_{i,j}}(\theta_{i,j+1})-TL_{v_{i,j}}(\theta_{i,j})\in T_{v_{i,j}}\mathcal{G}, \quad (5.8)$$

where $\theta_{i,j}=TL_{\phi^{-1}_{i,j}}(\delta\phi_{i,j})\in\mathfrak{g}$.

Proof: The equivalence of (a) and (b) follows from a standard argument in discrete Lagrangian field theories, and was proved in [**108**, sec. 5.1]. The equivalence of (c) and (d) follows from theorem 1.3 in chapter 4.

In order to prove the equivalence of (b) and (d), we note that $L=\Psi^\star\ell$, from which we conclude that if $\varphi=\Phi\circ\phi$, then $S(\phi)=s(\varphi)$. Now, consider the components $\{u_{i,j}\}$ and $\{v_{i,j}\}$ of the reduced field, as in (5.3). It is easy to check that an arbitrary variation $\epsilon\mapsto\phi_{i,j}(\epsilon)$ of ϕ induces corresponding variations $\delta u_{i,j}$ and $\delta v_{i,j}$ of $u_{i,j}$ and $v_{i,j}$, given by (5.7) and (5.8).

Finally, we show the equivalence of (d) with the Euler-Poincaré equations (5.6). The reduced action sum $s(\varphi)$ is given by:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}s(\varphi(\epsilon))\Big|_{\epsilon=0} &= \sum_{i,j}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\ell(u_{i,j}(\epsilon),v_{i,j}(\epsilon))\Big|_{\epsilon=0} \\ &= \sum_{i,j}\left(\mathrm{d}\ell(\cdot,v_{i,j})\cdot\delta u_{i,j}+\mathrm{d}\ell(u_{i,j},\cdot)\cdot\delta v_{i,j}\right).\end{aligned}$$

Substitution of (5.7) and (5.8) into this expression then yields (after relabelling some of the summation indices) the discrete Euler-Poincaré equations (5.6). ◇

Remark 2.2. The variations $\delta u_{i,j}$ and $\delta v_{i,j}$ are of a more general kind than those in chapter 3 (definition 3.6). In chapter 3, a variation is thought of as a map $V : U_V \to \mathfrak{g}$, whereas $\delta u_{i,j}$ and $\delta v_{i,j}$ are in fact maps from U_E, a subset of the set of edges, to TG (along the discrete field). ◇

2.3. The reconstruction problem. Let there be given, as in theorem 2.1, a $\mathcal{G}$-invariant Lagrangian L, a solution $\phi : V \to \mathcal{G}$ of the discrete Euler-Lagrange equations, and a reduced morphism $\varphi = \Phi \circ \phi$. The reduced morphism φ is a map from E to $\mathcal{G}$ and has a natural interpretation as a discrete connection in the sense of definition 1.5. This connection is easily seen to be flat.

To tackle the converse problem, we use the following consequence of proposition 1.6.

Proposition 2.3. *Let ω be a flat discrete $\mathcal{G}$-connection with associated discrete holonomy $\hat{\omega} : V \times V \to \mathcal{G}$. Then there exists a map $\phi : V \to \mathcal{G}$ such that $\omega(x_0, x_1) = \phi(x_0)^{-1}\phi(x_1)$. The map ϕ is unique up to the choice of an element of $\mathcal{G}$.*

Proof: Choose an arbitrary vertex x_0 and a group element g_0, and define $\phi(x_0) = g_0$. Let x_1 be any other vertex and put $\phi(x_1) = g_0\hat{\omega}(x_0, x_1)$. This map is well defined. ◇

Let $\varphi : E \to \mathcal{G}$ be a solution of the discrete Euler-Poincaré equations (5.6). We now wish to construct a solution ϕ of the original problem, such that $\varphi = \Phi \circ \phi$. The map ϕ is provided by proposition 2.3, on the condition that φ is a flat connection. As soon as φ is not flat, the holonomy $\hat{\omega}$ is path dependent, and no such ϕ can exist. Therefore, we have the following theorem.

Theorem 2.4 (reconstruction)**.** *Let $\varphi : E \to \mathcal{G}$ be a solution of the Euler-Poincaré equations (5.6). There exists a solution $\phi : V \to \mathcal{G}$ of the unreduced Euler-Lagrange equations (5.5) if and only if φ is flat. In that case, ϕ is unique up to the choice of an element of $\mathcal{G}$.*

Remark 2.5. In some cases, the element $g_0 \in \mathcal{G}$ determining the map ϕ is fixed by considering boundary conditions, or initial values. ◇

3. The Noether theorem

In section 2 of chapter 4, we introduced the concept of Noether symmetries and the associated conservation laws. From (4.7) in the proof of the Noether theorem (theorem 2.2 in the previous chapter), one can see that the fact that ϕ satisfies the Euler-Lagrange equations (the right hand side) implies that a certain conservation law holds. However, in general, there are not enough Noether symmetries to conclude the converse: if ϕ is an arbitrary discrete field satisfying the conservation law (4.5), then ϕ is not necessarily a solution of the Euler-Lagrange equations.

The case of a left-invariant Lagrangian on a configuration space of the form $\mathcal{G} \times \mathcal{G}$ is special in this regard, because in this case the converse does hold and the Euler-Lagrange equations can be rewritten as, and are equivalent to, a certain conservation law. This is the discrete counterpart of a similar construction in [**23**] for continuum field theories.

The Poincaré-Cartan forms for an unreduced Lagrangian $L : \mathcal{G}^{\times 3} \to \mathbb{R}$ are given by

$$\theta^L_{(1)} : \mathcal{G}^{\times 3} \to (T^*\mathcal{G})^{\times 3}, \quad \left\langle \theta^L_{(1)}(g_1, g_2, g_3), (v_1, v_2, v_3) \right\rangle = \left\langle \mathrm{d}L(\cdot, g_2, g_3)_{g_1}, v_1 \right\rangle,$$

and similarly for $\theta^L_{(2)}$ and $\theta^L_{(3)}$. Note that $\theta^L_{(1)} + \theta^L_{(2)} + \theta^L_{(3)} = \mathrm{d}L$. The Poincaré-Cartan forms associated to the reduced Lagrangian $\ell = \Psi^\star L : \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ are given by

$$\theta^\ell_{(1)} : \mathcal{G}^2 \to \mathfrak{g}^* \oplus \mathfrak{g}^* \oplus \mathfrak{g}^*, \tag{5.9}$$
$$\left\langle \theta^\ell_{(1)}(u_1, u_2), (\xi_1, \xi_2, \xi_3) \right\rangle = - \left\langle \mathrm{d}\ell(\cdot, u_2), TR_{u_1}(\xi_1) \right\rangle - \left\langle \mathrm{d}\ell(u_1, \cdot), TR_{u_2}(\xi_1) \right\rangle,$$

where $\xi_1, \xi_2, \xi_3 \in \mathfrak{g}$. The other Poincaré-Cartan forms $\theta^\ell_{(2)}$ and $\theta^\ell_{(3)}$ are then

$$\left\langle \theta^\ell_{(2)}(u, v), (\xi_1, \xi_2, \xi_3) \right\rangle = \left\langle \mathrm{d}\ell(\cdot, v), TL_u(\xi_2) \right\rangle \tag{5.10}$$

and

$$\left\langle \theta^\ell_{(3)}(u, v), (\xi_1, \xi_2, \xi_3) \right\rangle = \left\langle \mathrm{d}\ell(u, \cdot), TL_v(\xi_3) \right\rangle. \tag{5.11}$$

Note that $\theta^\ell_{(i)}$ and $\theta^L_{(i)}$, for $i = 1, 2, 3$, are related by proposition 1.4 in chapter 4.

3.1. The unreduced Lagrangian. Assume that $L : \mathcal{G}^{\times 3} \to \mathbb{R}$ is a left $\mathcal{G}$-invariant Lagrangian, in the sense that $L(hg_1, hg_2, hg_3) = L(g_1, g_2, g_3)$ for all h in $\mathcal{G}$. According to Noether's theorem, associated to this symmetry there is a conservation law.

Let ξ be an element of $\mathfrak{g}$. Infinitesimal invariance of the Lagrangian under the flow generated by ξ is expressed as

$$\left\langle \mathrm{d}L(g_1, g_2, g_3), (\xi_\mathcal{G}(g_1), \xi_\mathcal{G}(g_2), \xi_\mathcal{G}(g_3)) \right\rangle = 0, \tag{5.12}$$

where $\xi_\mathcal{G}$, defined by $\xi_\mathcal{G}(g) = TR_g(\xi)$, is the fundamental vector field associated to ξ. This shows that ξ is a Noether symmetry of L according to definition 2.1 in chapter 4.

From the Noether theorem (theorem 2.2 in chapter 4), we then conclude that the following conservation law holds:

$$\eta^{(x)}(\psi(\mathfrak{f}_1)) - \eta^{(x)}(\psi(\mathfrak{f}_2)) + \eta^{(y)}(\psi(\mathfrak{f}_1)) - \eta^{(y)}(\psi(\mathfrak{f}_3)) = 0, \tag{5.13}$$

where $\eta^{(x)} = \left\langle \theta_L^{(2)}, (\xi_{\mathcal{G}}, \xi_{\mathcal{G}}, \xi_{\mathcal{G}}) \right\rangle$, $\eta^{(y)} = \left\langle \theta_L^{(3)}, (\xi_{\mathcal{G}}, \xi_{\mathcal{G}}, \xi_{\mathcal{G}}) \right\rangle$, and $\mathfrak{f}_1$, $\mathfrak{f}_2$, $\mathfrak{f}_3$ are three triangles that touch a common vertex (see figure 2.1). As in the proof of theorem 2.2, the conservation law (5.13) can be written as

$$\langle D_1 L(\psi(\mathfrak{f}_1)) + D_2 L(\psi(\mathfrak{f}_2)) + D_3 L(\psi(\mathfrak{f}_3)), \xi_G \rangle = 0. \tag{5.14}$$

As ξ ranges over the whole of $\mathfrak{g}$, we conclude that the conservation law (5.13) is *equivalent* with the discrete Euler-Lagrange equations.

Remark 3.1. The above discussion can rephrased in terms of *discrete momentum maps.* For $i = 1, 2, 3$, we define the functions J_ξ^i on $\mathcal{G}^{\times 3}$ as

$$J_\xi^i(g_1, g_2, g_3) = \left\langle \theta_{(i)}^L(g_1, g_2, g_3), (\xi_{\mathcal{G}}(g_1), \xi_{\mathcal{G}}(g_2), \xi_{\mathcal{G}}(g_3)) \right\rangle.$$

Because of (5.12), we have $J_\xi^1 + J_\xi^2 + J_\xi^3 = 0$ for all $\xi \in \mathfrak{g}$. For more information on discrete momentum maps, see [**80**]. ◇

3.2. The reduced Lagrangian. Not only is the Noether theorem equivalent to the unreduced discrete field equations, it turns out that it contains the Euler-Poincaré equations as well. To show this, we start from the discrete conservation law as expressed in (4.6). We use the same notational conventions as in the preceding section and write

$$\Psi(\psi(\mathfrak{f}_i)) = (u_i, v_i), \quad i = 1, 2, 3.$$

Furthermore, we note that, by definition, $(\psi(\mathfrak{f}_1))_1 = (\psi(\mathfrak{f}_2))_2 = (\psi(\mathfrak{f}_3))_3$; this unique element of $\mathcal{G}$ is denoted by g. By rewriting each of the three expressions in (4.6) in terms of the reduced Lagrangian ℓ only, we obtain

$$\left\langle \theta_{(1)}^L, (\xi_{\mathcal{G}}, \xi_{\mathcal{G}}, \xi_{\mathcal{G}}) \right\rangle (\psi(\mathfrak{f}_1)) = -\left\langle R_{u_1}^* \mathrm{d}\ell(\cdot, v_1), \eta \right\rangle - \left\langle R_{v_1}^* \mathrm{d}\ell(u_1, \cdot), \eta \right\rangle,$$

where $\eta = \mathrm{Ad}_{g^{-1}}\xi$, as well as

$$\left\langle \theta_{(2)}^L, (\xi_{\mathcal{G}}, \xi_{\mathcal{G}}, \xi_{\mathcal{G}}) \right\rangle (\psi(\mathfrak{f}_2)) = \left\langle L_{u_2}^* \mathrm{d}\ell(\cdot, v_2), \eta \right\rangle$$

and

$$\left\langle \theta_{(3)}^L, (\xi_{\mathcal{G}}, \xi_{\mathcal{G}}, \xi_{\mathcal{G}}) \right\rangle (\psi(\mathfrak{f}_3)) = \left\langle L_{v_3}^* \mathrm{d}\ell(u_3, \cdot), \eta \right\rangle.$$

Putting all of these expressions together gives the following new form of the conservation law (5.13):

$$\left\langle \left[R_{u_1}^* \mathrm{d}\ell(\cdot, v_1) - L_{u_2}^* \mathrm{d}\ell(\cdot, v_2) \right] + \left[R_{v_1}^* \mathrm{d}\ell(u_1, \cdot) - L_{v_3}^* \mathrm{d}\ell(u_3, \cdot) \right], \eta \right\rangle = 0.$$

As η ranges over the whole of $\mathfrak{g}$, we conclude that the conservation law (5.13) implies the discrete Euler-Poincaré equation (5.6).

4. Extending the Moser-Veselov approach

In their seminal paper [**85**], Moser and Veselov approached the problem of finding an integrable discretization of the rigid-body equations by embedding the rotation group $SO(n)$ into a linear space, namely $\mathfrak{gl}(n)$. Somewhat later, Marsden, Pekarsky, and Shkoller [**77**] then developed a general procedure of Lagrangian reduction for discrete mechanical systems, and showed that the Moser-Veselov equations are equivalent to the discrete Lie-Poisson equations.

Here, we intend to do the same thing for a fundamental model in field theory: that of harmonic mappings from $\mathbb{R}^2$ into a Lie group $\mathcal{G}$. We will show that it is possible to develop a Moser-Veselov type discretization of these field equations, provided that $\mathcal{G}$ is embedded in a linear space. As could be expected, these discrete field equations are equivalent to the Euler-Poincaré equations.

In the continuous case, the harmonic mapping Lagrangian is given by

$$L = \frac{1}{2}\left\langle \phi^{-1}\phi_x, \phi^{-1}\phi_x \right\rangle + \frac{1}{2}\left\langle \phi^{-1}\phi_y, \phi^{-1}\phi_y \right\rangle, \tag{5.15}$$

where $\langle \cdot, \cdot \rangle$ is the Killing form on $\mathfrak{g}$. For the sake of clarity, we will only treat the case of harmonic maps that take values in $SO(n)$, embedded in $\mathfrak{gl}(n)$, in which case the Killing form is just the trace. We stress that the entire theory can be generalized to the case of an arbitrary semi-simple group $\mathcal{G}$, embedded in a linear space.

Consider the quadrangular mesh from section 2.1 in chapter 2 and denote the mesh spacing by h. As usual, we denote the values of the field ϕ on the vertices by $\phi_{i,j}$. We discretize the reduced partial derivatives $\phi^{-1}\phi_x$ and $\phi^{-1}\phi_y$ by writing them as follows:

$$\phi^{-1}\phi_x \approx \frac{1}{h}\phi^T_{i+1,j}(\phi_{i+1,j} - \phi_{i,j}) \quad \text{and} \quad \phi^{-1}\phi_y \approx \frac{1}{h}\phi^T_{i,j+1}(\phi_{i,j+1} - \phi_{i,j}),$$

where $\phi^T_{i+1,j}$ denotes the transpose of $\phi_{i+1,j} \in SO(n)$. Substituting this into (5.15) yields the following discrete Lagrangian (up to an unimportant constant):

$$L_d = -\frac{1}{h^2}\mathrm{tr}(\phi^T_{i,j}\phi_{i+1,j}) - \frac{1}{h^2}\mathrm{tr}(\phi^T_{i,j}\phi_{i,j+1}).$$

In order to ensure that $\phi_{i,j} \in SO(n)$, we need to impose the constraint that $\phi^T_{i,j}\phi_{i,j} = I$. We are thus led to consider the following constrained action involving Lagrange multipliers:

$$S(\phi) = \sum_{i,j}\left(\mathrm{tr}(\phi^T_{i,j}\phi_{i+1,j}) + \mathrm{tr}(\phi^T_{i,j}\phi_{i,j+1}) - \frac{1}{2}\mathrm{tr}\left(\Lambda_{i,j}(\phi^T_{i,j}\phi_{i,j} - I)\right)\right), \tag{5.16}$$

where we have redefined the Lagrange multipliers $\Lambda_{i,j}$ to get rid of the factor $-1/h^2$ (see [**77**]). Note that $\Lambda_{i,j}$ is a symmetric matrix of Lagrange multipliers.

The field equations are obtained by requiring that S be stationary under arbitrary variations; they are given by

$$\phi_{i+1,j} + \phi_{i-1,j} + \phi_{i,j+1} + \phi_{i,j-1} = \phi_{i,j}\Lambda_{i,j}. \tag{5.17}$$

We multiply these equations by $\phi_{i,j}^T$ from the right, and remark that $\phi\Lambda\phi^T$ is symmetric in order to get rid of the Lagrange multipliers:

$$\begin{aligned}\phi_{i+1,j}\phi_{i,j}^T + \phi_{i,j+1}\phi_{i,j}^T + \phi_{i-1,j}\phi_{i,j}^T + \phi_{i,j-1}\phi_{i,j}^T = \\ \phi_{i,j}\phi_{i+1,j}^T + \phi_{i,j}\phi_{i,j+1}^T + \phi_{i,j}\phi_{i-1,j}^T + \phi_{i,j}\phi_{i,j-1}^T.\end{aligned}$$

By introducing the following quantities,

$$m_{i+1,j} = \phi_{i+1,j}\phi_{i,j}^T - \phi_{i,j}\phi_{i+1,j}^T \quad \text{and} \quad n_{i,j+1} = \phi_{i,j+1}\phi_{i,j}^T - \phi_{i,j}\phi_{i,j+1}^T,$$

the field equations can be rephrased as the following set of conservation laws:

$$m_{i+1,j} + n_{i,j+1} = m_{i,j} + n_{i,j}. \tag{5.18}$$

Finally, let us introduce the *discrete momenta* $M_{i,j}$ and $N_{i,j}$, defined as

$$M_{i,j} = \phi_{i-1,j}^T m_{i,j}\phi_{i-1,j} \quad \text{and} \quad N_{i,j} = \phi_{i,j-1}^T n_{i,j}\phi_{i,j-1}.$$

The field equations governing the behaviour of these quantities are then easily determined to be, on the one hand

$$\begin{cases} M_{i,j} = \alpha_{i,j} - \alpha_{i,j}^T & \text{where } \alpha_{i,j} = u_{i-1,j}; \\ N_{i,j} = \beta_{i,j} - \beta_{i,j}^T & \text{where } \beta_{i,j} = v_{i,j-1}, \end{cases} \tag{5.19}$$

as well as, on the other hand, the counterpart of (5.18):

$$M_{i+1,j} + N_{i,j+1} = \mathrm{Ad}_{\alpha_{i,j}^T} M_{i,j} + \mathrm{Ad}_{\beta_{i,j}^T} N_{i,j}. \tag{5.20}$$

The similarities with the Moser-Veselov equations for the discrete rigid body are obvious (see [**85**, eq. 4]).

Remark 4.1. It is now straightforward to see the equivalence between the Moser-Veselov and the Euler-Poincaré equations. Indeed, starting from the reduced Lagrangian ℓ, put

$$M_{i+1,j} = R_{u_{i,j}}^* \mathrm{d}\ell(\cdot, v_{i,j}) \quad \text{and} \quad N_{i,j+1} = R_{v_{i,j}}^* \mathrm{d}\ell(u_{i,j}, \cdot),$$

which can be interpreted as a discrete Legendre transformation. Furthermore, put $\alpha_{i,j} = u_{i-1,j}$ and $\beta_{i,j} = v_{i,j-1}$. The Euler-Poincaré equations (5.6) then reduce to the Moser-Veselov equations derived above. ◇

5. An application: harmonic mappings

In our extension of the Moser-Veselov algorithm, we already made a brief start on the study of discrete harmonic mappings. Here, we would like to summarize some further points of interest. However, before doing so, it is perhaps useful to start with a brief review of harmonic mappings in the continuous case. For more information, see [**23**, **114**].

5.1. The continuum formulation. In this section, we consider harmonic mappings $\phi : \mathbb{R}^2 \to \mathcal{G}$ with values in an arbitrary Lie group $\mathcal{G}$ with *bi-invariant* metric $\langle \cdot, \cdot \rangle$. The continuum Lagrangian is given by (5.15); the associated field equations are $\tau(\phi) = 0$, where $\tau(\phi)$ is the tension of ϕ, defined as

$$\tau(\phi)^a = h^{ij} \left(\frac{\partial^2 \phi^a}{\partial x^i \partial x^j} - \Gamma^k_{ij} \frac{\partial \phi^a}{\partial x^k} + C^a_{bc} \frac{\partial \phi^b}{\partial x^i} \frac{\partial \phi^c}{\partial x^j} \right),$$

where h_{ij} are the components of the metric on $\mathbb{R}^2$, with associated Christoffel symbols Γ^k_{ij}, and C^a_{bc} are the Christoffel symbols of the bi-invariant metric on $\mathcal{G}$. In our case, h_{ij} is of course just the flat Euclidian metric.

The Euler-Poincaré equations for harmonic mappings are partial differential equations for a $\mathfrak{g}$-valued one-form α on $\mathbb{R}^2$. These equations were derived from a reduced variational principle in [**23**]; here, we present a more direct derivation (see [**114**]).

Let θ be the (left) Maurer-Cartan form on $\mathcal{G}$, defined as $\theta_g(v_g) = TL_{g^{-1}}(v_g)$, which satisfies the Maurer-Cartan equation $\mathrm{d}\theta + \frac{1}{2}[\theta, \theta] = 0$. For any mapping $\phi : \mathbb{R}^2 \to \mathcal{G}$, we now consider the pull-back form $\alpha = \phi^* \theta$. Naturally, α satisfies a Maurer-Cartan type equation:

$$\mathrm{d}\alpha + \frac{1}{2}[\alpha, \alpha] = 0. \tag{5.21}$$

Conversely, if a $\mathfrak{g}$-valued one-form α on $\mathbb{R}^2$ satisfies this equation, then it is possible to find a map $\phi : \mathbb{R}^2 \to \mathcal{G}$ such that $\alpha = \phi^* \theta$.

In addition, one can prove that ϕ is harmonic if and only if α is co-closed:

$$\delta\alpha = 0. \tag{5.22}$$

At this point, one usually introduces a spectral parameter allowing one to write both equation (5.21) and (5.22) as a single equation (see [**114**]). We will not go this far; rather, we will prove in the next section that there exists a natural discrete counterpart of these two equations.

5.2. Discrete harmonic mappings. In this section, we derive the discrete field equations for harmonic maps by use of the Euler-Poincaré reduction procedure. We recall the expression (5.16) for the discrete action. By starting from the unreduced action sum S, we may easily derive the field equations (5.17). We now multiply these equations from the *left* (rather than from the right as in the derivation of the Moser-Veselov equations) to obtain the following set of Euler-Poincaré equations:

$$u_{i,j} + v_{i,j} + u^T_{i-1,j} + v^T_{i,j-1} = \Lambda_{i,j}, \tag{5.23}$$

together with the integrability condition

$$u_{i,j} v_{i+1,j} u^{-1}_{i,j+1} v^{-1}_{i,j} = e. \tag{5.24}$$

Using then the symmetry of $\Lambda_{i,j}$, we eliminate the multipliers $\Lambda_{i,j}$ to arrive at the following expression:

$$u_{i,j} + v_{i,j} - u_{i-1,j} - v_{i,j-1} = u^T_{i,j} + v^T_{i,j} - u^T_{i-1,j} - v^T_{i,j-1}.$$

If we view φ as a $\mathfrak{gl}(n)$-valued discrete one-form in the sense of section 1.1, then the Euler-Poincaré equations can be conveniently expressed using the discrete codifferential:

$$\delta\varphi = (\delta\varphi)^T,$$

or $[\![\delta\varphi]\!] = 0$, where $[\![\cdot]\!]$ denotes the antisymmetric part of a matrix: $[\![A]\!] = \frac{1}{2}(A - A^T)$.

Chapter 6

Geometric aspects of nonholonomic field theories

In this chapter, we give a geometric interpretation of classical field theories with nonholonomic constraints. In a nutshell, a constrained problem consists of looking for sections ϕ of a bundle π that extremize a certain action density, but with the added condition that the prolongations $j^1\phi$ take values in a certain submanifold $\mathcal{C}$ of $J^1\pi$. Of course, an arbitrary solution of the Euler-Lagrange equations will not satisfy the latter condition automatically; rather, one has to introduce certain *reaction forces*, that keep the solution constrained to $\mathcal{C}$.

In mechanics, this problem setting is well known, and two different approaches exist for the determination of the dynamics.

- One may restrict the action functional S to curves which satisfy the constraint. This is the so-called constrained variational or *vakonomic* approach (which stands for dynamics of the "Variational Axiomatic Kind").[1]
- On the other hand, one might leave the domain of S completely free, and impose restrictions on the variations instead. In particular, one can obtain the reaction forces from the d'Alembert principle, which specifies that variations satisfying the constraints should perform no work. The resulting approach is referred to as the *nonholonomic method*.

Somewhat surprisingly at first sight, these two approaches are not equivalent, unless the constraints are integrable (see [**27**]). Nowadays, it is generally accepted that the nonholonomic approach is the correct one for the dynamics of mechanical systems. This was already known to pioneers such as Korteweg, and was verified experimentally by Murray & Lewis [**71**]. The vakonomic approach, on the other hand, is used in subriemannian geometry and control theory.

Both the nonholonomic and the vakonomic method have their counterparts in classical field theory. Here, we propose an extension of the methods from nonholonomic mechanics. The vakonomic method was treated by García-Perez *et al.* [**45**] and, in the context of elasticity theory, by Marsden *et al.* [**81**]. We also mention the work of Bibbona *et al.* [**8**,**9**], who showed that the nonholonomic framework in this chapter is inappropriate for certain kinds of constraints. These observations will be dealt with in chapter 9. The

[1] We prefer the etymology of Kozlov [**62**], according to whom the roots of the word "vakonomic" go back to the Italian verb "vacare" (to be free, to have nothing to do).

distinction between the vakonomic and the nonholonomic treatment in general will be made clearer in chapter 8.

In this chapter, we will mainly focus on the geometry of nonholonomic constraints. In particular, we construct in section 2 a special projection operator which maps solutions of the free De Donder-Weyl equation (1.17) to solutions of the nonholonomic De Donder-Weyl equation (introduced below). Finally, in section 4, we invoke the Cauchy framework to show that a nonholonomic field theory induces a nonholonomic mechanical system (in the traditional sense) on the space of Cauchy data, thus extending the analogy of section 3 in chapter 1.

1. Nonholonomic Lagrangian field theory

As in chapter 1, we consider a fibre bundle $\pi : Y \to X$ with oriented $(n+1)$-dimensional base space X, with η a volume form on X. The fibre bundle π is assumed to have rank m and is equipped with bundle coordinates (x^μ, y^a) as in the first chapter.

Let a first-order Lagrangian $L : J^1\pi \to \mathbb{R}$ be given. Assume that $\mathcal{C} \hookrightarrow J^1\pi$ is a submanifold of $J^1\pi$ of codimension k, representing some external constraints imposed on the system. For the sake of clarity, two assumptions will be made regarding the nature of $\mathcal{C}$: first, that $\mathcal{C}$ projects onto the whole of Y, *i.e.* $\pi_{1,0}(\mathcal{C}) = Y$, and secondly, that the restriction $(\pi_{1,0})_{|\mathcal{C}} : \mathcal{C} \to Y$ of $\pi_{1,0}$ to $\mathcal{C}$ is a fibre bundle, which, however, need not be an affine subbundle of $\pi_{1,0}$.

Of these two assumptions, the latter in particular is quite restrictive. With proper caution, one can probably carry out the further analysis under some weaker condition. However, from the point of view of practical examples, not much would be gained by such an extension.

Since $\mathcal{C}$ is a submanifold of $J^1\pi$, one can always find a covering of $\mathcal{C}$ consisting of open subsets U of $J^1\pi$, with $U \cap \mathcal{C} \neq \varnothing$, such that on each $U \in \mathcal{U}$ there exist k functionally independent smooth functions φ^α that locally determine $\mathcal{C}$, *i.e.*

$$\mathcal{C} \cap U = \{\gamma \in J^1\pi : \varphi^\alpha(\gamma) = 0 \text{ for } 1 \leq \alpha \leq k\}. \tag{6.1}$$

Additionally, the assumption that $(\pi_{1,0})_{|\mathcal{C}}$ be a fibre bundle implies that the matrix with entries $\frac{\partial \varphi^\alpha}{\partial y^a_\mu}$ has maximal rank k at each point $\gamma \in \mathcal{C} \cap U$.

1.1. The bundle of constraint forms. In nonholonomic mechanics, the derivation of the equations of motion of a mechanical system with nonholonomic constraints is based on the so-called *d'Alembert principle* and involves, among others, the specification of a suitable bundle of admissible "reaction forces" (and a corresponding bundle of admissible virtual velocities), defined along the constraint submanifold. This choice relies on an additional rule or principle. In nonholonomic mechanics it is quite common to use the so-called Chetaev principle, whereby the bundle of reaction forces is constructed directly in terms of the given constraints. In principle, however, the specification of the appropriate bundle of reaction forces (or virtual displacements), compatible with the given constraints, is problem dependent and need not necessarily be based on Chetaev's rule. For a critical discussion of this matter we refer to [**74**]; see also [**99**].

In general, we shall model these "reaction forces" as certain $(n+1)$-forms. The reasoning behind this particular model is as follows. Recall that in geometric mechanics a force is modelled as a one-form α on TQ, and the work done by α along a variation X (a vector field on TQ) is then simply given by the pairing $W := \langle \alpha, X \rangle$. In the Cauchy framework, an $(n+1)$-form induces a one-form on the space of Cauchy data, which is then a force in the sense alluded to above.

Returning to the case of first-order field theory with external constraints, we follow a similar procedure and introduce a special subbundle F of rank k of the bundle of exterior $(n+1)$-forms on $J^1\pi$ defined along the constraint submanifold $\mathcal{C}$, where we recall that k is the codimension of $\mathcal{C}$. This bundle, called the bundle of *constraint forms*, will play a role similar to that of the bundle of reaction forces in nonholonomic mechanics.

We stress that the module of constraint forms is to some extent an independent unknown of the model. We begin by introducing a submodule $\mathcal{F}$ of the module of $(n+1)$-forms on $J^1\pi$; the bundle F will then be determined by the fact that $\mathcal{F}$ is the module of sections of F. The only conditions that we impose on $\mathcal{F}$ are of a technical nature: an element of $\mathcal{F}$ is an $(n+1)$-form Φ defined along $\mathcal{C}$, satisfying the following two assumptions:

(1) Φ is n-horizontal, *i.e.* Φ vanishes when contracted with any two π_1-vertical vector fields;
(2) Φ is 1-contact, *i.e.* $(j^1\phi)^*\Phi = 0$ for any section ϕ of π.

Apart from these assumptions, the nature of $\mathcal{F}$ is left entirely free.

Regarding the local expression of the constraint forms, one can find an open cover $\mathcal{U}$ of $\mathcal{C}$ such that on each open set $U \in \mathcal{U}$, the module $\mathcal{F}$ is generated by k independent $(n+1)$-forms Φ^α that locally read

$$\Phi^\alpha = (C^\alpha)^\mu_a(\mathrm{d}y^a - y^a_\nu \mathrm{d}x^\nu) \wedge \mathrm{d}^n x_\mu = (C^\alpha)^\mu_a \theta^a \wedge \mathrm{d}^n x_\mu, \tag{6.2}$$

for some smooth functions $(C^\alpha)^\mu_a$ on U. Independence of the forms Φ^α clearly implies that the $(k \times (n+1)m)$-matrix whose elements are the functions $(C^\alpha)^\mu_a$, has constant maximal rank k. The *bundle of constraint forms* (or *constraint forces*) is then defined by

$$F = \cup_{\gamma\in\mathcal{C}} F_\gamma \quad \text{with} \quad F_\gamma = \{\Phi(\gamma) \,|\, \Phi \in \mathcal{F}\}\,.$$

At this point, the reason for selecting a constraint bundle of the type described above is primarily based on the analogy with nonholonomic mechanics.

Remark 1.1. In [**10**] the authors have constructed the bundle of constraint forms by considering a natural extension of the *Chetaev principle* commonly used in mechanics when dealing with nonlinear nonholonomic constraints. More precisely, they define the local generators Φ^α of the bundle of constraint forms by putting

$$\Phi^\alpha := S_\eta^*(\mathrm{d}\varphi^\alpha)\,, \tag{6.3}$$

where the φ^α are the local constraint functions defined in (6.1). One easily verifies that these Φ^α are indeed of the form (6.2), with

$$(C^\alpha)_a^\mu = \frac{\partial \varphi^\alpha}{\partial y_\mu^a}$$

In the case we are considering, the linear independence of these Φ^α is guaranteed by our initial assumption that $\mathcal{C}$ has the structure of a fibre bundle over Y. ◇

1.2. The constraint distribution. As we will now show, the constraint bundle F gives rise to a distribution D along $\mathcal{C}$, called the *constraint distribution*. As above, consider an open cover $\mathcal{U}$ of $\mathcal{C}$ such that on each $U \in \mathcal{U}$, the module $\mathcal{F}$ is generated by k independent $(n+1)$-forms Φ^α of the form (6.2).

Proposition 1.2. *For each α, there exists a unique vector field $Z^\alpha \in \mathfrak{X}(U)$ such that*

$$i_{Z^\alpha}\Omega_L = -\Phi^\alpha. \tag{6.4}$$

Proof: Take Z^α to be a $\pi_{1,0}$-vertical vector field on U, *i.e.*

$$Z^\alpha = (Z^\alpha)_\mu^a \frac{\partial}{\partial y_\mu^a}.$$

With this choice, equation (6.4) reduces to

$$(Z^\alpha)_\mu^a \frac{\partial^2 L}{\partial y_\mu^a \partial y_\nu^b} = (C^\alpha)_b^\nu, \tag{6.5}$$

which determines the $(Z^\alpha)_\mu^a$ uniquely, as L is supposed to be regular. This already proves the existence of a solution of (6.4). Uniqueness then follows from the fact that Ω_L is multisymplectic. ◇

The vector fields Z^α span a k-dimensional distribution D_U on U. It is not difficult to check that for any two open sets $U, V \in \mathcal{U}$ with nonempty intersection, and for each $\gamma \in U \cap V$, we have that $D_U(\gamma) = D_V(\gamma)$. Indeed, assume that $\mathcal{F}$ is generated on U by k independent forms Φ^α and on V by k independent forms $\bar{\Phi}^\alpha$. Then, there exists a nonsingular matrix of functions A_β^α on $U \cap V$ such that

$$\Phi^\alpha = A_\beta^\alpha \bar{\Phi}^\beta\,.$$

If we denote the corresponding generators of D_U by Z^α and those of D_V by $\bar{Z}^\alpha$, it readily follows from the previous proposition that

$$Z^\alpha_{|U\cap V} = A^\alpha_\beta \bar{Z}^\beta_{|U\cap V},$$

which proves that $D_U = D_V$ on $U \cap V$. Consequently, the local distributions described in the previous proposition induce a well-defined (global) distribution D along the constraint submanifold $\mathcal{C}$, whose sections are $\pi_{1,0}$-vertical vector fields. Moreover, using a similar argument as above, one easily verifies that this distribution does not depend on the initial choice we made for an open cover $\mathcal{U}$ of $\mathcal{C}$.

1.3. The nonholonomic field equations. Summarizing the above, we are looking for a field theory built on the following data:

(1) a Lagrangian density $L\eta$ with regular Lagrangian $L \in C^\infty(J^1\pi)$;
(2) a constraint submanifold $\mathcal{C} \hookrightarrow J^1\pi$ which can be locally represented by equations of the form $\varphi^\alpha(x^\mu, y^a, y^a_\mu) = 0$, for $\alpha = 1, \ldots, k$ and where the matrix $(\partial\varphi^\alpha/\partial y^a_\mu)$ has maximal rank k;
(3) a bundle F of constraint forms and an induced constraint distribution D, both defined along $\mathcal{C}$, whereby F is locally generated by k independent $(n+1)$-forms (6.2), and D is defined according to the construction described in proposition 1.2.

To complete our model for nonholonomic field theory, we now have to specify the field equations. Proceeding along the same lines as in [**10**] we introduce the following definition, using a generalization of d'Alembert's principle.

Definition 1.3. *A local section ϕ of $\pi : Y \to X$, defined on an open set $U \subset X$ with compact closure, is a solution of the nonholonomic constrained problem described above if $j^1\phi(U) \subset \mathcal{C}$ and*

$$\int_U (j^1\phi)^* \mathscr{L}_{j^1V} L\eta = 0\,,$$

for all π-vertical vector fields V on Y that vanish on the boundary of $\phi(U)$ and such that

$$j^1V \lrcorner\, \Phi = 0 \quad \textit{along} \ \mathrm{Im}\, j^1\phi \tag{6.6}$$

for all sections Φ of the bundle F of constraint forms.

Putting $V = V^a(x,y)\partial/\partial y^a$ and taking into account the expression (1.3) for the prolonged vector field j^1V, it is easily seen that the condition (6.6) translates into

$$(C^\alpha)^\mu_a V^a = 0 \quad \text{along } \mathrm{Im}\, j^1\phi,$$

where the $(C^\alpha)^\mu_a$ are the coefficients of the constraint forms introduced in (6.2). One can then verify that if $\phi(x) = (x^\mu, \phi^a(x))$ is a solution of the constrained problem, then the functions $\phi^a(x)$ satisfy the following system of partial differential equations:

$$\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^a_\mu}\right) = \lambda_{\alpha\mu}(C^\alpha)^\mu_a \quad (a = 1, \ldots, m)\,, \tag{6.7}$$

$$\varphi^\alpha\left(x^\mu, \phi^a(x), \frac{\partial \phi^a}{\partial x^\mu}(x)\right) = 0 \quad (\alpha = 1, \ldots, k)\,. \tag{6.8}$$

As usual, the (a priori) unknown functions $\lambda_{\alpha\mu}$ play the role of "Lagrange multipliers". The equations (6.7) are called the *nonholonomic field equations* for the constrained problem. Note that if the bundle F of constraint forms is defined according to a Chetaev-type prescription (see remark 1.1), then we recover the nonholonomic field equations derived in [**10**].

Let $\mathcal{I}(F)$ be the ideal of differential forms, defined along $\mathcal{C}$, generated by the constraint forms: i.e any element of $\mathcal{I}(F)$ is of the form $\sum_i \lambda_i \wedge \Phi^i$, for some $\Phi^i \in \mathcal{F}$ and arbitrary differential forms λ_i. Again proceeding along the same lines as in [**10**] we can formulate the following modification of the De Donder-Weyl problem for nonholonomic Lagrangian field theory: find a connection on $\pi_1 : J^1\pi \to X$ with horizontal projector $\mathbf{h}$ such that along the constraint submanifold $\mathcal{C}$

$$i_{\mathbf{h}}\Omega_L - n\Omega_L \in \mathcal{I}(F) \quad \text{and} \quad \mathrm{Im}\,\mathbf{h} \subset T\mathcal{C}\,. \tag{6.9}$$

For simplicity we will refer to (6.9) as the *nonholonomic De Donder-Weyl equation.* In coordinates, if we represent $\mathbf{h}$ by (1.11), one can easily check that the relation on the left of (6.9) leads to the following set of equations for the connection coefficients of the connection we are looking for:

$$(\Gamma^b_\nu - y^b_\nu)\left(\frac{\partial^2 L}{\partial y^a_\mu \partial y^b_\nu}\right) = 0\,,$$

$$\frac{\partial L}{\partial y^a} - \frac{\partial^2 L}{\partial x^\tau \partial y^a_\tau} - \Gamma^b_\tau \frac{\partial^2 L}{\partial y^b \partial y^a_\tau} - \Gamma^b_{\tau\nu}\frac{\partial^2 L}{\partial y^b_\tau \partial y^a_\nu} + (\Gamma^b_\nu - y^b_\nu)\frac{\partial^2 L}{\partial y^a \partial y^b_\nu} = \lambda_{\alpha\tau}(C^\alpha)^\tau_a\,,$$

for $a = 1, \ldots, m$ and $\mu = 1, \ldots, n+1$ and some Lagrange multipliers $\lambda_{\alpha\tau}$. These expressions should still be supplemented by the requirement that for any $\gamma \in \mathcal{C}$ and any $v \in T_\gamma J^1\pi$, $\mathbf{h}(v) \in T_\gamma\mathcal{C}$. This is equivalent to requiring that $\mathbf{h}(v)(\varphi^\alpha) = 0$ for all $v \in T_{\mathcal{C}}J^1\pi$, where φ^α $(\alpha = 1, \ldots, k)$ are the (local) constraint functions. If, locally, $\mathbf{h}$ is written in the form (1.11), then the previous condition translates into the following additional equations for the connection coefficients in points of $\mathcal{C}$:

$$\frac{\partial \varphi^\alpha}{\partial x^\mu} + \Gamma^b_\mu \frac{\partial \varphi^\alpha}{\partial y^b} + \Gamma^b_{\mu\nu}\frac{\partial \varphi^\alpha}{\partial y^b_\nu} = 0 \quad \text{for all} \quad \mu = 1, \ldots, n+1;\ \alpha = 1, \ldots, k.$$

One can prove that in case of a regular Lagrangian, integral sections of a connection satisfying (6.9) will be 1-jet prolongations of solutions of the nonholonomic field equations (see [**10**] for details).

2. The nonholonomic projector

The purpose of the present section is to show that for a nonholonomic first-order field theory in the sense described above, one can construct, under an appropriate additional condition, a projection operator which maps solutions of the De Donder-Weyl equation (1.17) for the free (*i.e.* unconstrained) Lagrangian problem into solutions of the nonholonomic De Donder-Weyl equation (6.9).

Given a constrained problem as described in the previous section, with regular Lagrangian L, constraint manifold $\mathcal{C} \hookrightarrow J^1\pi$ and constraint distribution D, we now impose the following *compatibility condition*: for each $\gamma \in \mathcal{C}$, we require that

$$D(\gamma) \cap T_\gamma \mathcal{C} = \{0\}. \tag{6.10}$$

If $\mathcal{C}$ is locally defined by k equations $\varphi^\alpha(x^\mu, y^a, y^a_\mu) = 0$ and if D is locally generated by the vector fields Z^α (see subsection 1.2), a straightforward computation shows that the compatibility condition is satisfied if and only if

$$\det\left(Z^\alpha(\varphi^\beta)(\gamma)\right) \neq 0\,,$$

at each point $\gamma \in \mathcal{C}$. Indeed, take $v \in T_\gamma\mathcal{C} \cap D(\gamma)$, then $v = v_\alpha Z^\alpha(\gamma)$ for some coefficients v_α. On the other hand, $0 = v(\varphi^\beta) = v_\alpha Z^\alpha(\varphi^\beta)(\gamma)$. Hence, if the matrix $\left(Z^\alpha(\varphi^\beta)(\gamma)\right)$ is invertible, we may conclude that $v = 0$ and the compatibility condition holds. The proof of the converse is similar.

We now have the following result.

Proposition 2.1. *If the compatibility condition (6.10) holds, then at each point $\gamma \in \mathcal{C}$ we have the decomposition*

$$T_\gamma J^1\pi = T_\gamma\mathcal{C} \oplus D(\gamma).$$

Proof: The proof immediately follows from (6.10) and a simple counting of dimensions: $\dim T_\gamma\mathcal{C} = \dim T_\gamma J^1\pi - k$ and $\dim D(\gamma) = k$. ◇

The direct sum decomposition of $T_\mathcal{C} J^1\pi$ determines two complementary projection operators $\mathcal{P}$ and $\mathcal{Q}$:

$$\mathcal{P} : T_\mathcal{C} J^1\pi \to T\mathcal{C} \quad \text{and} \quad \mathcal{Q} = I - \mathcal{P} : T_\mathcal{C} J^1\pi \to D\,,$$

where I is the identity on $T_\mathcal{C} J^1\pi$. We will call $\mathcal{P}$ *the nonholonomic projector* associated to the given constrained problem.

Given a connection on π_1 such that the associated horizontal projector $\mathbf{h}$ is a solution of the free De Donder-Weyl equation (1.17), we will prove that the operator $\mathcal{P} \circ \mathbf{h}_{|T_\mathcal{C} J^1\pi}$ satisfies the constrained De Donder-Weyl equation (6.9). Note that this operator is only defined along $\mathcal{C}$ and, therefore, strictly speaking, is not the horizontal projector of a connection on π_1. However, one can show (lemma 2.3) that its restriction to $T\mathcal{C}$ induces a genuine connection on the restricted bundle $(\pi_1)_{|\mathcal{C}} : \mathcal{C} \to X$, and so the constrained De Donder-Weyl equation still makes sense for this kind of map.

Remark 2.2. In the following, we will introduce the concept of "connection on a subbundle of π_1". A similar concept was introduced, from a slightly different point of view, in [**32**, appendix C]. $\diamond$

Lemma 2.3. *The map* $\mathcal{P} \circ \mathbf{h}_{|T_{\mathcal{C}}J^1\pi} : T_{\mathcal{C}}J^1\pi \to T\mathcal{C}\ (\subset T_{\mathcal{C}}J^1\pi), v \mapsto \mathcal{P}(\mathbf{h}(v))$ *is a projector whose restriction* $\mathbf{h}_{\mathcal{P}}$ *to* $T\mathcal{C}$ *induces a connection on* $(\pi_1)_{|\mathcal{C}} : \mathcal{C} \to X$.

Proof: First of all, we check that for each $\gamma \in \mathcal{C}$ the map $\mathcal{P}_\gamma \circ \mathbf{h}_\gamma$ is a projector. Indeed, taking into account that $\operatorname{Im} \mathcal{Q} = D$ is $\pi_{1,0}$-vertical, it follows that for all $v \in T_\gamma J^1\pi$

$$(\mathbf{h}_\gamma \circ \mathcal{P}_\gamma)(v) = \mathbf{h}_\gamma(v) - (\mathbf{h}_\gamma \circ \mathcal{Q}_\gamma)(v) = \mathbf{h}_\gamma(v).$$

and therefore

$$(\mathcal{P}_\gamma \circ \mathbf{h}_\gamma)^2 = \mathcal{P}_\gamma \circ \mathbf{h}_\gamma.$$

The restriction $\mathbf{h}_{\mathcal{P}}$ of $\mathcal{P} \circ \mathbf{h}_{|T_{\mathcal{C}}J^1\pi}$ to $T\mathcal{C}$ obviously is still a projector. The key point we now have to prove is that $\operatorname{Im}(\mathbf{h}_{\mathcal{P}})$ is a complementary bundle to $V(\pi_1)_{|\mathcal{C}}$ in $T\mathcal{C}$, *i.e.*

$$\operatorname{Im}(\mathbf{h}_{\mathcal{P}}) \oplus V(\pi_1)_{|\mathcal{C}} = T\mathcal{C}\,. \tag{6.11}$$

For that purpose we start by observing that along $\mathcal{C}$ we have $T\mathcal{C} \cap V\pi_1 = V(\pi_1)_{|\mathcal{C}}$. In view of Proposition 2.1 one can then easily derive the following direct sum decomposition:

$$V(\pi_1)_{|\mathcal{C}} \oplus D = V\pi_1 \quad (\text{along } \mathcal{C}). \tag{6.12}$$

Next, by taking into account the fact that the constraint distribution D is vertical, and therefore that $\mathbf{h}_{\mathcal{P}}(T_\gamma\mathcal{C}) = (\mathcal{P} \circ \mathbf{h})(T_\gamma J^1\pi)$ for every $\gamma \in \mathcal{C}$, it is a routine exercise to verify that

$$\dim(\mathcal{P} \circ \mathbf{h})(T_\gamma J^1\pi) = \dim \mathbf{h}(T_\gamma J^1\pi)\,. \tag{6.13}$$

We now prove the direct sum decomposition (6.11). Take any $v \in T\mathcal{C}$ with $v \in \operatorname{Im}(\mathbf{h}_{\mathcal{P}}) \cap V(\pi_1)_{|\mathcal{C}}$, then there exists a vector $w \in T\mathcal{C}$ such that $v = \mathcal{P}(\mathbf{h}(w)) = \mathbf{h}(w) - \mathcal{Q}(\mathbf{h}(w))$. Since v is π_1-vertical, we conclude that $\mathbf{h}(w) = 0$ and, hence, $v = 0$. This already implies that $\operatorname{Im}(\mathbf{h}_{\mathcal{P}}) \cap V(\pi_1)_{|\mathcal{C}} = 0$. The equality (6.11) now follows from a simple dimensional argument. Indeed, relying on Proposition 2.1 as well as on (6.12) and (6.13), we have at each point $\gamma \in \mathcal{C}$:

$$\begin{aligned}
\dim(\mathbf{h}_{\mathcal{P}}(T_\gamma\mathcal{C})) + \dim V_\gamma(\pi_1)_{|\mathcal{C}} &= \dim(\mathbf{h}(T_\gamma J^1\pi)) + \dim V_\gamma\pi_1 - \dim D(\gamma) \\
&= \dim(T_\gamma J^1\pi) - \dim D(\gamma) \\
&= \dim T_\gamma\mathcal{C}\,.
\end{aligned}$$

This concludes the proof that $\mathbf{h}_{\mathcal{P}} = \mathcal{P} \circ \mathbf{h}_{|T\mathcal{C}}$ is the horizontal projector of a connection on $(\pi_1)_{|\mathcal{C}}$. $\diamond$

Although $(\pi_1)_{|\mathcal{C}} : \mathcal{C} \to X$ is not a first-order jet bundle, we will say that a connection on $(\pi_1)_{|\mathcal{C}}$, with associated horizontal projector $\mathbf{h}_\mathcal{C}$, is *semi-holonomic* if for each contact 1-form θ on $J^1\pi$

$$i_{\mathbf{h}_\mathcal{C}} \iota^* \theta = 0, \tag{6.14}$$

where $\iota : \mathcal{C} \hookrightarrow J^1\pi$ is the canonical injection. Suppose $\tau : X \to \mathcal{C}$ is an integral section of a connection on $(\pi_1)_{|\mathcal{C}}$, in the sense that $T\tau(T_xX) \subset \mathbf{h}_\mathcal{C}(T_{\tau(x)}\mathcal{C})$ for all x in the domain of τ. Then, if the given connection is semi-holonomic one can verify that, locally, τ can be written as the first jet prolongation of a (local) section of π.

As mentioned at the end of subsection 2.1, the regularity of L together with the fact that $\mathbf{h}$ satisfies the free De Donder-Weyl equation, imply that $\mathbf{h}$ is a semi-holonomic connection on $J^1\pi$. Herewith one can prove the following result.

Lemma 2.4. *The connection on $(\pi_1)_{|\mathcal{C}}$ defined in Lemma 2.3, with horizontal projector $\mathbf{h}_\mathcal{P}$, is semi-holonomic.*

Proof: We will use the fact that $\mathbf{h}$ is semi-holonomic and therefore satisfies (1.10). Let $v \in T_\gamma J^1\pi$ be a $\pi_{1,0}$-vertical vector, then for any contact 1-form θ on $J^1\pi$ we have that $i_v\theta(\gamma) = 0$. Now, for each $v \in T_\mathcal{C}J^1\pi$ we have that $(\mathcal{P}\circ\mathbf{h}-\mathbf{h})(v) = -\mathcal{Q}(\mathbf{h}(v)) \in D$ and, hence, $(\mathcal{P}\circ\mathbf{h}-\mathbf{h})(v)$ is $\pi_{1,0}$-vertical. Therefore $i_{\mathcal{P}\circ\mathbf{h}}\theta(v) = i_\mathbf{h}\theta(v) = 0$ for any contact 1-form θ and any $v \in T_\mathcal{C}J^1\pi$. From this one can readily deduce that $\mathbf{h}_\mathcal{P}$ satisfies (6.14) and so we may conclude that the induced connection on $(\pi_1)_{|\mathcal{C}}$ is indeed semi-holonomic. $\diamond$

We now arrive at the main result of this section. From now on, for ease of notation, we will use the projector $\mathcal{P}\circ\mathbf{h}$ without further indication of its domain. The latter should be clear from the context.

Theorem 2.5. *Consider a constrained problem of the type described above, with regular Lagrangian L, constraint submanifold $\mathcal{C} \hookrightarrow J^1\pi$ and bundle of constraint forms F, and assume the compatibility condition (6.10) holds. Let $\mathbf{h}$ be the horizontal projector of a connection on π_1, satisfying the free De Donder-Weyl equation (1.17) and let $\mathcal{P}$ be the nonholonomic projector associated to the constrained problem. Then the projector $\mathcal{P}\circ\mathbf{h}$ determines a solution of the constrained De Donder-Weyl problem (6.9) and restricts to the horizontal projector of a semi-holonomic connection on $(\pi_1)_{|\mathcal{C}} : \mathcal{C} \to X$.*

Proof: Along $\mathcal{C}$ we can rewrite the free De Donder-Weyl equation as

$$i_{\mathcal{P}\circ\mathbf{h}}\Omega_L - n\Omega_L = -i_{\mathcal{Q}\circ\mathbf{h}}\Omega_L.$$

Therefore, in order to prove that $\mathcal{P}\circ\mathbf{h}$ satisfies the constrained De Donder-Weyl equation, we only need to verify that the right-hand side is an element of $\mathcal{I}(F)$.

We can write the projector $\mathbf{h}$ as $\mathbf{h} = \mathrm{d}x^\mu \otimes X_\mu^H$, with $X_\mu^H = (\partial/\partial x^\mu) + \Gamma_\mu^a(\partial/\partial y^a) + \Gamma_{\mu\nu}^a(\partial/\partial y_\nu^a)$ (see (1.11)). Along $\mathcal{C}$ we can then put $\mathcal{Q}(X_\mu^H) = \lambda_{\alpha\mu}Z^\alpha$ for some functions

$\lambda_{\alpha\mu}$ and with the vector fields Z^α as defined in Proposition 1.2. Then, at each point $\gamma \in \mathcal{C}$ and for any $v_1, \ldots, v_{n+2} \in T_\gamma J^1\pi$ we obtain

$$\begin{aligned}(i_{\mathcal{Q}\circ\mathbf{h}}\Omega_L)(v_1,\ldots,v_{n+2}) &= \sum_{i=1}^{n+2}(-1)^{i+1}\Omega_L((\mathcal{Q}\circ\mathbf{h})(v_i), v_1,\ldots,\hat{v}_i,\ldots,v_{n+2}) \\ &= \sum_{i=1}^{n+2}(-1)^{i+1}\lambda_{\alpha\mu}\mathrm{d}x^\mu(v_i)(i_{Z^\alpha}\Omega_L)(v_1,\ldots,\hat{v}_i,\ldots,v_{n+2}) \\ &= -\lambda_{\alpha\mu}(\mathrm{d}x^\mu \wedge \Phi^\alpha)(v_1,\ldots,v_{n+2})\,.\end{aligned}$$

This shows that, along $\mathcal{C}$,

$$i_{\mathcal{Q}\circ\mathbf{h}}\Omega_L = -\lambda_{\alpha\mu}\mathrm{d}x^\mu \wedge \Phi^\alpha \in \mathcal{I}(F),$$

which completes the proof of the first part of the theorem.

The proof that $\mathcal{P}\circ\mathbf{h}$ induces a semi-holonomic connection on $(\pi_1)_{|\mathcal{C}}$ follows from the previous lemmas 2.3 and 2.4. ◇

Note that even in case a connection on π_1, with horizontal projector $\mathbf{h}$ satisfying the free De Donder-Weyl equation, is holonomic (or integrable), the 'projected' semi-holonomic connection $\mathbf{h}_\mathcal{P} = \mathcal{P}\circ\mathbf{h}$ on $(\pi_1)_{|\mathcal{C}}$ need not admit integral sections in general.

3. An example from incompressible hydrodynamics

As an example of a field theory with an external constraint, we consider the case of an incompressible fluid flow. This problem is traditionally treated using the constrained variational approach (see for instance [**81**] for a geometric treatment). From the point of view of nonholonomic field theory it is therefore an a-typical example since (contrary to prior expectation) the constrained field equations resulting from the nonholonomic approach agree with those derived using the vakonomic approach.

The reason for this unexpected agreement stems from the fact that the incompressibility constraint can be written as a divergence — recall that for a mechanical system with a nonholonomic constraint that arises from a total time derivative of a function on the configuration space, the nonholonomic and the vakonomic equations of motion are equivalent.

3.1. The constrained problem. We recall the geometric formalism for fluid dynamics outlined in section 4.1 of chapter 1. Here, we shall assume that both the material manifold M and the spatial manifold S are open subsets of $\mathbb{R}^3$. Consider now the function $\mathcal{J} : J^1\pi \to \mathbb{R}$ given by

$$\mathcal{J}(\gamma) := \det(y^a_i(\gamma)) \quad \text{where } y = \pi_{1,0}(\gamma) \quad \text{and} \quad x = \pi_1(\gamma).$$

Here, y^a_i (for $i = 1, \dots, n$) represents the "spatial part" of y^a_μ. Note that (y^a_i) is a square matrix, which is invertible when restricted to regular sections. For any section ϕ of π, $\mathcal{J} \circ j^1\phi$ measures the volume change of a small fluid element under the fluid motion represented by ϕ. The incompressibility constraint is expressed by imposing the condition $\mathcal{J}(j^1\phi) = 1$, *i.e.* we have the constraint function

$$\varphi(\gamma) := \mathcal{J}(\gamma) - 1, \tag{6.15}$$

defining a constraint submanifold $\mathcal{C}$ in $J^1\pi$. This constraint was examined in [**81**] using the vakonomic formalism (as is customary); here, we will use the nonholonomic framework and see what equations result. For the bundle F of constraint forms, we adopt the generalized Chetaev principle (see remark 1.1); F is the line bundle along $\mathcal{C}$ generated by the 4-form

$$\begin{aligned}\Phi &:= S^*_\eta(\mathrm{d}\varphi)\\ &= \mathcal{J}(y^{-1})^i_a(\mathrm{d}y^a - y^a_\nu \mathrm{d}x^\nu) \wedge \mathrm{d}^3x_i\,,\end{aligned}$$

3.2. The nonholonomic field equations. Before proceeding towards the field equations, we make the additional assumption that we are dealing with a *barotropic fluid* which, in particular, implies that W depends on the y^a_i through $\mathcal{J}$, *i.e.* $W = W(\mathcal{J})$. The nonholonomic field equations (6.7) for a barotropic fluid with Lagrangian (1.31), subject to the incompressibility constraint (6.15) and with constraint form Φ, then become

$$\rho\frac{\mathrm{d}y^a_0}{\mathrm{d}t} - \frac{\mathrm{d}}{\mathrm{d}x^j}\left(\rho W'\mathcal{J}(y^{-1})^j_a\right) = \lambda_i \mathcal{J}(y^{-1})^i_a \quad (a = 1, 2, 3)\,, \tag{6.16}$$

which should be considered together with the constraint equation $\mathcal{J}(\gamma) - 1 = 0$. This should be compared with equation (4.8) in [**81**]. In that paper, the field equations for an incompressible barotropic fluid were derived by means of a constrained variational approach. Since there is only one constraint equation, this approach gives rise to only one Lagrangian multiplier P, which is commonly interpreted as a pressure.

Our approach, however, runs into trouble here. In mechanics, one usually determines the Lagrange multipliers by taking the derivative with respect to time of the constraints, and uses the equations of motion to eliminate the accelerations. Under some modest assumptions, this procedure fully determines the multipliers in terms of the positions and the velocities.

No such approach is possible in the context of field theories. In order to do so, one would need to be able to write the field equations as

$$y^a_{,\mu\nu} = f^a_{\mu\nu}(x^\kappa, y^b_\lambda),$$

where the "comma notation" is used to denote partial derivatives. However, it is clear from (6.16) that the field equations for hydrodynamics cannot be written in this form.

Hence, it follows that there may be several sets of multipliers λ_i, all of whom satisfy the constraint equations and the field equations.

We therefore make an additional assumption, that there exists a function P such that $\lambda_i = \frac{\partial P}{\partial x^i}$. This assumption is rather ad hoc, but leads to a set of equations which coincides with the field equations derived using the vakonomic approach. That these two sets of equations should agree is at first sight rather remarkable. The reason for this is to be found in the fact that the incompressibility constraint is determined by a divergence. More precisely, we have the following property.

Proposition 3.1. *The constraint function φ can be written (locally) as a total divergence,* i.e. *there exist functions ψ^μ such that $\varphi = \frac{\mathrm{d}\psi^\mu}{\mathrm{d}x^\mu}$.*

Proof: One can easily verify that

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial\varphi}{\partial y^a_\mu}\right) - \frac{\partial\varphi}{\partial y^a} \equiv 0,$$

i.e. φ is a "null-Lagrangian", which is equivalent to φ being a divergence (see e.g. [**89**, thm. 4.7]). More directly, if we consider the functions

$$\psi^0 = 0 \quad \text{and} \quad \psi^i = \frac{1}{3}\mathcal{J} y^a (y^{-1})^i_a - x^i\,,$$

with y^{-1} the inverse of the matrix (y^a_i), which are well defined on a neighborhood of $\mathcal{C}$, a rather tedious but straightforward computation shows that $\varphi = \mathrm{d}\psi^\mu/\mathrm{d}x^\mu$. ◇

Note that the nonholonomic approach would have given the wrong kind of field theories if proposition 3.1 did not hold. This can be observed in other, more sophisticated field theories, and was claimed in [**8**, **9**]. Our response to this criticism can be found in chapter 8, where we will also make a detailed study of the comparison between the constrained variational approach and the nonholonomic approach (in the case of affine constraints).

3.3. The nonholonomic projector. To illustrate some further concepts defined in the preceding sections, we now turn to the explicit form of the nonholonomic projector $\mathcal{P}$ for the example of incompressible fluid (not necessarily barotropic). As there is only one constraint, the constraint distribution D is spanned by a single vector field $Z = Z^a_\mu \partial/\partial y^a_\mu$. The coefficients of this vector field can be derived from (6.5) where, in the present case, $C^\mu_a = \partial\varphi/\partial y^a_\mu$:

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{\partial^2 W}{\partial y^a_i \partial y^b_j} \end{pmatrix} \begin{pmatrix} Z^b_0 \\ Z^b_j \end{pmatrix} = \begin{pmatrix} 0 \\ \mathcal{J}(\gamma)\,(y^{-1})^i_a \end{pmatrix}.$$

If, for brevity, we denote the Hessian matrix of W with respect to the y^a_i by $\mathcal{H}$, then Z is the vector field along $\mathcal{C}$ given by

$$Z = (\mathcal{H}^{-1})^{ab}_{ij} \mathcal{J} (y^{-1})^j_b \frac{\partial}{\partial y^a_i}.$$

Let us consider the function $C := Z(\varphi)$, or explicitly

$$C = (\mathcal{H}^{-1})^{ab}_{ij} \mathcal{J}^2 (y^{-1})^i_a (y^{-1})^j_b.$$

For each $\gamma \in \mathcal{C}$, $C(\gamma) \neq 0$ from which it follows that the compatibility condition (6.10) holds. The nonholonomic projector $\mathcal{P}$ is then found to be

$$\mathcal{P} = I - \frac{1}{C} \mathrm{d}\varphi \otimes Z.$$

4. Cauchy formalism for nonholonomic field theory

We will now describe the transition from the multisymplectic covariant treatment of nonholonomic field theory, discussed in the previous sections, to the formulation of the problem on the space of Cauchy data. In chapter 1, it was shown that a covariant field theory on $J^1\pi$ formally induces a mechanical system on the space of Cauchy data. Here, we will extend that analogy by proving that a nonholonomic field theory gives rise in a similar fashion to a nonholonomic mechanical system on the space of Cauchy data.

4.1. Introductory definitions. For the remainder of this chapter, let us assume that the compatibility condition (6.10) holds. In order to adapt the Cauchy formalism to the nonholonomic case, we first define a subset $\tilde{\mathcal{C}}$ of $\tilde{Z}$ as follows:

$$\tilde{\mathcal{C}} := \left\{ \kappa \in \tilde{Z} : \mathrm{Im}\, \kappa \subset \mathcal{C} \right\}. \tag{6.17}$$

This set can be equipped with a smooth manifold structure such that $\tilde{\mathcal{C}}$ becomes a (infinite-dimensional) submanifold of $\tilde{Z}$, and should be thought of as the infinite-dimensional (global) analogue to the (local) constraint submanifold. There exists a particularly convenient expression for tangent vectors to $\tilde{\mathcal{C}}$:

Lemma 4.1. *For every $\kappa \in \tilde{\mathcal{C}}$, there exists a natural bijection between $T_\kappa \tilde{\mathcal{C}}$ and the space of sections of the pullback bundle $\kappa^* T\mathcal{C}$.*

Proof: The proof is similar to the proof of (1.22). ◇

For each $\kappa \in \tilde{\mathcal{C}}$, let

$$\tilde{D}_\kappa := \left\{ W_\kappa \in T_\kappa \tilde{Z} : \mathrm{Im}\, W_\kappa \subset D \right\},$$

where D is the constraint distribution along $\mathcal{C}$. Equivalently, $\tilde{D}_\kappa$ can be identified with the set of sections of κ^*D. Putting $\tilde{D} = \cup_{\kappa\in\tilde{\mathcal{C}}} D_\kappa$, one may verify that $\tilde{D}$ determines a smooth distribution on $\tilde{Z}$ along $\tilde{\mathcal{C}}$.

Next, for $\kappa \in \tilde{\mathcal{C}}$ and for each section α of the bundle F of constraint forms along $\mathcal{C}$, we define an element $\tilde{\alpha}_\kappa$ of $T^*_\kappa \tilde{Z}$ by

$$\tilde{\alpha}_\kappa(W_\kappa) = \int_M \kappa^*(i_{W_\kappa}\alpha), \quad \text{for all } W_\kappa \in T_\kappa\tilde{Z}. \tag{6.18}$$

The set of all such covectors $\tilde{\alpha}_\kappa$ determines a subspace $\tilde{F}_\kappa$ of $T^*_\kappa\tilde{Z}$ and $\tilde{F} = \bigcup_{\kappa\in\tilde{\mathcal{C}}} \tilde{F}_\kappa$ is a codistribution on $\tilde{Z}$ along $\tilde{\mathcal{C}}$.

The Chetaev principle. In the remainder of this section, we assume that the Chetaev principle (see remark 1.1) holds, and we show that the bundle of reaction forces $\tilde{F}$, determined by (6.18), can also be constructed starting from the submanifold $\tilde{\mathcal{C}}$ and using the vertical endomorphism $\tilde{S}_{\tilde{\eta}}$. We prove that

$$\tilde{F} = \tilde{S}^*_{\tilde{\eta}}(T^\circ\tilde{\mathcal{C}}),$$

which is precisely the geometric form of the Chetaev principle in mechanics (see [**26**]). This result is a first example of the fact that, using Cauchy analysis, nonholonomic field theories become genuine nonholonomic mechanical systems.

Using the volume form η_M on M, we may establish a correspondence between one-forms on $J^1\pi$ and one-forms on $\tilde{Z}$ by putting, for $\alpha \in \Omega^1(J^1\pi)$,

$$\tilde{\alpha}_\kappa(X_\kappa) = \int_M \kappa^*(\alpha(X_\kappa))\eta_M.$$

In other words, $\tilde{\alpha}$ is the one-form on $\tilde{Z}$ associated to the $(n+1)$-form $\alpha \wedge \eta_M$ according to the prescription (1.23).

Lemma 4.2. *For every $\kappa \in \tilde{\mathcal{C}}$, there exists a natural bijection between the annihilator subspace $T^\circ_\kappa\tilde{\mathcal{C}}$ and the space of sections of $\kappa^*T^\circ\mathcal{C}$.*

Proof: Only the inclusion of $T^\circ_\kappa\tilde{\mathcal{C}}$ in the space of sections of $\kappa^*T^\circ\mathcal{C}$ is not entirely obvious. Consider $\alpha \in T^\circ_\kappa\tilde{\mathcal{C}}$, then for all $X_\kappa \in T_\kappa\tilde{\mathcal{C}}$, we have

$$\alpha(X_\kappa) = \int_M \kappa^*(X_\kappa \lrcorner\, \alpha)\eta_M = 0.$$

By a standard argument, it follows that $X_\kappa(u) \lrcorner\, \alpha(u)$ is a contact form for any $u \in M$. But as $\alpha(u)$ is a one-form, we have $X_\kappa(u) \lrcorner\, \alpha(u) = 0$, or $\alpha(u) \in T^\circ_{\kappa(u)}\mathcal{C}$. ◇

Proposition 4.3. *The bundle of reaction forces $\tilde{F}$, defined in (6.18), satisfies*

$$\tilde{F} = \tilde{S}^*_{\tilde{\eta}}(T^\circ\tilde{\mathcal{C}}).$$

Proof: Let κ be an element of $\tilde{\mathcal{C}}$ and consider $\alpha \in \tilde{F}(\kappa)$. Then, by definition, there exist k functions λ_α on M such that

$$\alpha(X_\kappa) = \int_M \kappa^*(X_\kappa \lrcorner\, S^*_\eta(\lambda_\alpha \mathrm{d}\varphi^\alpha)) = \left\langle \tilde{S}^*_{\tilde{\eta}}(\lambda_\alpha \mathrm{d}\varphi^\alpha), X_\kappa \right\rangle,$$

where we have used the form (1.29) of $\tilde{S}_{\tilde{\eta}}$. We conclude that $\alpha = \tilde{S}^*_{\tilde{\eta}}(\lambda_\alpha \mathrm{d}\varphi^\alpha)$; the converse is similar. $\diamond$

4.2. The nonholonomic equations of motion. Since we assume that the given constrained problem satisfies the compatibility condition, we can use the nonholonomic projector $\mathcal{P}$ and the complementary projector $\mathcal{Q} = I - \mathcal{P}$ (cf. section 2) to define two operators $\tilde{\mathcal{P}}$, $\tilde{\mathcal{Q}} : T_{\tilde{\mathcal{C}}}\tilde{Z} \to T_{\tilde{\mathcal{C}}}\tilde{Z}$ by composition. For each $\kappa \in \tilde{\mathcal{C}}$ and $W_\kappa \in T_\kappa\tilde{Z}$, put

$$\tilde{\mathcal{P}}_\kappa(W_\kappa) = \mathcal{P} \circ W_\kappa \ (\in T_\kappa\tilde{Z}), \qquad \tilde{\mathcal{Q}}_\kappa(W_\kappa) = \mathcal{Q} \circ W_\kappa \ (\in T_\kappa\tilde{Z}).$$

Using the properties of $\mathcal{P}$ and $\mathcal{Q}$, it is not hard to check that, for each $\kappa \in \tilde{\mathcal{C}}$, $\tilde{\mathcal{P}}_\kappa$ and $\tilde{\mathcal{Q}}_\kappa$ define complementary projectors in $T_\kappa\tilde{Z}$, *i.e.*

$$(\tilde{\mathcal{P}}_\kappa)^2 = \tilde{\mathcal{P}}_\kappa,\ (\tilde{\mathcal{Q}}_\kappa)^2 = \tilde{\mathcal{Q}}_\kappa \quad \text{and} \quad \tilde{\mathcal{P}}_\kappa + \tilde{\mathcal{Q}}_\kappa = I_\kappa\,,$$

with I_κ the identity on $T_\kappa\tilde{Z}$. This implies that $T_\kappa\tilde{Z} = \mathrm{Im}\,\tilde{\mathcal{P}}_\kappa \oplus \mathrm{Im}\,\tilde{\mathcal{Q}}_\kappa$. Again relying on the definitions of $\tilde{\mathcal{C}}, \tilde{D}, \tilde{\mathcal{P}}$ and $\tilde{\mathcal{Q}}$, and on the properties of the nonholonomic projector $\mathcal{P}$, one can prove that

$$\mathrm{Im}\,\tilde{\mathcal{P}}_\kappa = T_\kappa\tilde{\mathcal{C}} \quad \text{and} \quad \mathrm{Im}\,\tilde{\mathcal{Q}}_\kappa = \tilde{D}_\kappa.$$

Summarizing, we may conclude that under the given conditions we have the following decomposition of $T\tilde{Z}$ along $\tilde{\mathcal{C}}$:

$$T_{\tilde{\mathcal{C}}}\tilde{Z} = T\tilde{\mathcal{C}} \oplus \tilde{D}\,.$$

Let $\mathbf{h}$ be the horizontal projector of a connection Υ on π_1 and let Γ denote the vector field on $\tilde{Z}$ defined by (1.27). The composition $\tilde{\mathcal{P}} \circ \Gamma$ then determines a vector field on $\tilde{\mathcal{C}}$, shortly denoted by $\tilde{\mathcal{P}}(\Gamma)$, and it is not difficult to see that it is precisely the vector field associated to the induced connection on $(\pi_1)_{|\mathcal{C}}$ with horizontal projector $\mathbf{h}_\mathcal{P} = \mathcal{P} \circ \mathbf{h}$ (see section 2). We now have the following interesting result.

Lemma 4.4. *There exists a section $\tilde{\alpha}$ of $\tilde{F}$, such that*

$$i_{\tilde{\mathcal{P}}(\Gamma)}\tilde{\Omega}_L = i_\Gamma\tilde{\Omega}_L + \tilde{\alpha}. \tag{6.19}$$

Proof: For $\kappa \in \tilde{\mathcal{C}}$ and $W_\kappa \in T_\kappa\tilde{Z}$, one can deduce from the definition of $\tilde{\Omega}_L$ that

$$(i_{\tilde{\mathcal{P}}(\Gamma)}\tilde{\Omega}_L)(\kappa)(W_\kappa) = \int_M \kappa^*(i_{\tilde{\mathcal{P}}(\Gamma)(\kappa)} i_{W_\kappa}\Omega_L).$$

For the integrand on the right-hand side we have that, with $u \in M$,

$$i_{\tilde{\mathcal{P}}(\Gamma)(\kappa)(u)} i_{W_\kappa(u)}\Omega_L = i_{\Gamma(\kappa)(u)} i_{W_\kappa(u)}\Omega_L - i_{\tilde{\mathcal{Q}}(\Gamma)(\kappa)(u)} i_{W_\kappa(u)}\Omega_L,$$

where $\tilde{\mathcal{Q}}(\Gamma)$ is the vector field associated to $\mathcal{Q} \circ \mathbf{h}$ (note that $\tilde{\mathcal{Q}}(\Gamma)$ is defined along $\tilde{\mathcal{C}}$). Since $\tilde{\mathcal{Q}}(\Gamma)(\kappa)(u)$ is an element of the constraint distribution D, the contraction with Ω_L yields a form $\alpha_{\kappa(u)} \in F_{\kappa(u)}$. Integration over M then gives (6.19). ⋄

We have now collected all ingredients needed to formulate the main result of this section. Consider a constrained Lagrangian field theory, with regular Lagrangian L, with constraints verifying the appropriate conditions and such that the base manifold X admits a global space-time splitting.

Theorem 4.5. *Let* $\mathbf{h}$ *be a solution of the unconstrained De Donder-Weyl equation (1.17) and let* Γ *be the corresponding second-order vector field on* $\tilde{Z}$*. Then, the vector field* $\tilde{\mathcal{P}}(\Gamma)$ *on* $\tilde{\mathcal{C}}$ *satisfies the following relations:*

$$i_{\tilde{\mathcal{P}}(\Gamma)}\tilde{\eta} = 1, \quad i_{\tilde{\mathcal{P}}(\Gamma)}\tilde{\Omega}_L \in \tilde{F} \quad \text{and} \quad \tilde{\mathcal{P}}(\Gamma) \in T\tilde{\mathcal{C}}. \tag{6.20}$$

Proof: The first of these relations can be proved as in theorem 3.10 in chapter 1. If $\mathbf{h}$ satisfies the De Donder-Weyl equation, then the associated vector field Γ is contained in the kernel of $\tilde{\Omega}_L$ (see proposition 3.10). Expression (6.19) then proves the first part of (6.20). The second part follows from the definition of $\tilde{\mathcal{C}}$. ⋄

In addition, we note that $\tilde{\mathcal{P}}(\Gamma)$ is still a vector field of second-order type, due to propositions 2.4 and 3.9.

To conclude, we have shown that under the appropriate assumptions, the Cauchy formalism for nonholonomic field theory induces a vector field of "second-order type" on the infinite-dimensional subspace $\tilde{\mathcal{C}}$ of the space of Cauchy data $\tilde{Z}$. This vector field is a solution of the nonholonomic equations of motion, and can be written as the projection of the second-order vector field on $\tilde{Z}$ associated to the free (unconstrained) Lagrangian system.

Chapter 7

The nonholonomic momentum equation

In this chapter, we study nonholonomic field theories in the presence of symmetry. In this case, symmetries no longer automatically lead to conservation laws as with the Noether theorem. Rather, there exists an equation which describes the evolution of these "conserved currents". This equation was first derived in the context of mechanics in [**13**,**20**]; here we establish a similar result for field theories. Throughout this chapter, $\pi : Y \to X$ will be a fibre bundle as in chapter 1.

We begin by proving a number of additional properties of connections on π_1. In section 2, we then treat the case of field theories where no constraints are present. This should be thought of as more of a warming-up exercise: our main purpose is to review the covariant Noether theorem (see also proposition 2.7 in chapter 1) in a way suitable for generalization to the constrained case. In section 3, we introduce constraints into the framework and we study the implications for the Noether theorem. Finally, in section 4 we break covariance by going to the Cauchy setting to make the link with the geometric structures known from nonholonomic mechanical systems with symmetry.

1. Further properties of connections on π_1

In this section, we will prove a number of straightforward properties of connections on π_1 that will be used later on. The main results here are closely related to those in section 1.3 in chapter 1, but as they are only relevant for the developments in this chapter, we mention them here.

Lemma 1.1. *Let L be a Lagrangian on $J^1\pi$ with associated Cartan form Θ_L. For each semi-holonomic connection Υ on π_1 with horizontal projector $\mathbf{h}$, the following holds: $i_{\mathbf{h}}\Theta_L = n\Theta_L + L\eta$.*

Proof: We give the proof in coordinates. For any connection $\mathbf{h}$, we have

$$i_{\mathbf{h}}\mathrm{d}^{n+1}x = (n+1)\mathrm{d}^{n+1}x \quad \text{and} \quad i_{\mathbf{h}}\mathrm{d}^n x_\mu = n\mathrm{d}^n x_\mu.$$

Therefore,

$$i_{\mathbf{h}}\Theta_L = \frac{\partial L}{\partial y^a_\nu} i_{\mathbf{h}}\theta^a \wedge \mathrm{d}^n x_\nu + n\frac{\partial L}{\partial y^a_\nu}\theta^a \wedge \mathrm{d}^n x_\nu + (n+1)L\mathrm{d}^{n+1}x,$$

where the forms θ^a are the contact 1-forms introduced in section 1.1.3. If $\mathbf{h}$ is semi-holonomic, the first term on the right-hand side is zero and we obtain the desired expression. ◇

This lemma can be viewed as the jet-bundle analogue of the well-known fact in Lagrangian mechanics that $i_X\theta_L = \Delta(L)$ for any second-order vector field X, where θ_L is the Cartan one-form corresponding to L, and Δ the Liouville vector field.

Lemma 1.2. *Let V be a vertical vector field on Y and j^1V its prolongation to $J^1\pi$. If Υ is a semi-holonomic connection on π_1 with horizontal projector $\mathbf{h}$, then the Frölicher-Nijenhuis bracket $[j^1V, \mathbf{h}]$ is a vector-valued one-form taking values in $V\pi_{1,0}$.*

Proof: If $V = V^a\frac{\partial}{\partial y^a}$, then

$$j^1V = V^a\frac{\partial}{\partial y^a} + \left(\frac{\partial V^a}{\partial x^\mu} + \frac{\partial V^a}{\partial y^b}y^b_\mu\right)\frac{\partial}{\partial y^a_\mu}.$$

For the bracket, we have that $[j^1V, \mathbf{h}] = \mathscr{L}_{j^1V}\mathbf{h}$ and a straightforward calculation then shows that this is a semi-basic vector-valued one-form taking values in $V\pi_1$. We now focus on the coefficient of $\mathrm{d}x^\mu \otimes \frac{\partial}{\partial y^a}$, which is just

$$j^1V(\Gamma^a_\mu) - \left(\frac{\partial V^a}{\partial x^\mu} + \Gamma^b_\mu\frac{\partial V^a}{\partial y^b}\right).$$

This coefficient is easily seen to vanish when $\Gamma^a_\mu = y^a_\mu$, *i.e.* when $\mathbf{h}$ is semi-holonomic, which completes the proof. ◇

As a corollary, we note that this lemma implies that the contraction of $[j^1V, \mathbf{h}]$ with a semi-basic form (in particular with Θ_L) vanishes.

2. Symmetry in the absence of nonholonomic constraints

Let $\mathcal{G}$ be a Lie group acting on Y by bundle automorphisms Φ_g over the identity in X. The assumption that $\mathcal{G}$ acts vertically is probably superfluous, but for the sake of clarity we will assume it nevertheless. We recall from section 2.2 in chapter 1 that, for a prolonged action, there always exists a covariant momentum map which is explicitly given by

$$J^L_\xi = j^1\xi_Y \lrcorner\, \Theta_L.$$

The covariant Noether theorem (proposition 2.7 in chapter 1) states that the momentum map J^L is conserved. Here, we give an alternative formulation suitable for generalisation later on.

Proposition 2.1 (Covariant Noether theorem)**.** *Let Υ be a connection on π_1 such that the associated horizontal projector $\mathbf{h}$ is a solution of the unconstrained De Donder-Weyl equation (1.17). For every $\xi \in \mathfrak{g}$, the momentum map J^L_ξ is constant on integral* *sections of* **h***:*

$$\mathrm{d}_{\mathbf{h}} J^L_\xi = 0.$$

Proof: In this proof, we make frequent use of some elementary properties of the Frölicher-Nijenhuis bracket. For the sake of completeness, we have summarized these properties in appendix A.

We have

$$\begin{aligned} \mathrm{d}_{\mathbf{h}} J^L_\xi &= \mathrm{d}_{\mathbf{h}} i_{j^1\xi_Y}\Theta_L \\ &= (i_{\mathbf{h}}\mathrm{d} - \mathrm{d}i_{\mathbf{h}}) i_{j^\xi_Y}\Theta_L \\ &= i_{\mathbf{h}}\mathscr{L}_{j^1\xi_Y}\Theta_L - i_{\mathbf{h}} i_{j^1\xi_Y}\mathrm{d}\Theta_L - \mathrm{d}i_{\mathbf{h}} i_{j^1\xi_Y}\Theta_L. \end{aligned} \tag{7.1}$$

In the last expression, the first term vanishes because of the invariance of the Cartan form (see (1.18)). The second term can be rewritten by using the field equations (note that $\mathbf{h}(j^1\xi_Y) = 0$ as $j^1\xi_Y$ is π_1-vertical):

$$i_{\mathbf{h}} i_{j^1\xi_Y}\mathrm{d}\Theta_L = i_{j^1\xi_Y} i_{\mathbf{h}}\mathrm{d}\Theta_L = -n i_{j^1\xi_Y}\Omega_L,$$

whereas for the last term we have, using lemma 1.2,

$$\begin{aligned} \mathrm{d}i_{\mathbf{h}} i_{j^1\xi_Y}\Theta_L &= \mathrm{d}i_{j^1\xi_Y} i_{\mathbf{h}}\Theta_L \\ &= \mathrm{d}i_{j^1\xi_Y}\left(n\Theta_L + L\eta\right). \end{aligned}$$

Now, $i_{j^1\xi_Y}(L\eta) = 0$ and so we obtain

$$\mathrm{d}_{\mathbf{h}} J_\xi = n i_{j^1\xi_Y}\Omega_L - n\mathrm{d}i_{j^1\xi_Y}\Theta_L = -n\mathscr{L}_{j^1\xi_Y}\Theta_L = 0,$$

again due to the invariance of Θ_L. ◇

Remark 2.2. In chapter 1, we mentioned a slightly different type of Noether theorem (proposition 2.7). Following [**48**, p. 45], we stated that if ϕ is a solution of the field equations, then $\mathrm{d}(j^1\phi)^* J_\xi = 0$. However, it is not hard to prove that, for any k-form α on $J^1\pi$, $(j^1\phi)^*\mathrm{d}_{\mathbf{h}}\alpha = \mathrm{d}(j^1\phi)^*\alpha$ if and only if $j^1\phi$ is an integral section of $\mathbf{h}$. Proposition 2.1 therefore implies that $\mathrm{d}(j^1\phi)^* J_\xi = 0$. The proof of proposition 2.7 in chapter 1 is more straightforward; our proof has the advantage that it will be easily extensible to the case where nonholonomic constraints are present. ◇

3. The constrained momentum map

In this section, we study the case of a constrained field theory, with regular Lagrangian L, constraint submanifold $\mathcal{C}$, and bundle of reaction forces F satisfying the conditions at the beginning of section 1.3 in chapter 6. The constrained De Donder-Weyl equations are then given by (6.9).

Suppose now that in addition to these nonholonomic constraints, there is also a symmetry group $\mathcal{G}$ acting on $J^1\pi$ by prolonged bundle automorphisms as in the previous section, such that L, $\mathcal{C}$, and F are $\mathcal{G}$-invariant, *i.e.*

$$L \circ j^1\Phi_g = L, \quad j^1\Phi_g(\mathcal{C}) \subset \mathcal{C} \quad \text{and} \quad (j^1\Phi_g)^* F \subset F \tag{7.2}$$

for all $g \in \mathcal{G}$. Throughout this section, the action of $\mathcal{G}$ on Y will be assumed to be vertical, as in the previous sections.

In general, as in the case of nonholonomic mechanics (see [**6**, **13**, **20**]), it will no longer be true that these symmetries give rise to conserved quantities; the precise link will be made clear by the *nonholonomic momentum equation* or constrained Noether theorem (theorem 3.1 below). Our treatment extends the one in [**20**]; we refer to that paper, as well as to [**6**, **13**] and the references therein, for more information about the nonholonomic momentum equation in mechanics.

We first introduce the following "distribution":

$$\mathcal{E}(\gamma) = \{v \in T_\gamma J^1\pi : i_v\Phi = 0 \text{ for all } \Phi \in F\} \quad \text{where } \gamma \in \mathcal{C}. \tag{7.3}$$

It is possible that the rank of $\mathcal{E}$ is not constant. For a given $\gamma \in \mathcal{C}$ we consider all elements ξ of the Lie algebra $\mathfrak{g}$ such that $j^1\xi_Y(\gamma) \in \mathcal{E}(\gamma)$. The set of all such ξ we denote by $\mathfrak{g}^\gamma$. We take $\mathfrak{g}^\mathcal{E}$ to be the disjoint union of all these spaces $\mathfrak{g}^\gamma$ and we assume that $\mathfrak{g}^\mathcal{E}$ can be given the structure of a bundle over $\mathcal{C}$.

With these elements in mind, we define the *constrained momentum map* as the map $J^{\text{n.h.}} : \mathcal{C} \to \bigwedge^n(J^1\pi) \otimes \mathfrak{g}^\mathcal{E}$, constructed as follows. With every section $\bar{\xi}$ of $\mathfrak{g}^\mathcal{E}$, one may associate a vector field $\tilde{\xi}$ on $J^1\pi$, along $\mathcal{C}$, by putting $\tilde{\xi}(\gamma) = (\bar{\xi}(\gamma))_{J^1\pi}(\gamma)$. Remark that $\tilde{\xi}$ is a section of $\mathcal{E}$. We then define $J^{\text{n.h.}}_{\bar{\xi}}$ along $\mathcal{C}$ as

$$J^{\text{n.h.}}_{\bar{\xi}} = i_{\tilde{\xi}}\Theta_L.$$

The importance of the nonholonomic momentum map lies in the nonholonomic momentum equation:

Theorem 3.1 (Nonholonomic momentum equation)**.** *Let Υ be a connection on π_1 such that the associated horizontal projector $\mathbf{h}$ is a solution of the constrained De Donder-Weyl equation. Assume furthermore that $\mathcal{G}$ is a Lie group acting vertically on $J^1\pi$ and preserving the Lagrangian. Then the nonholonomic momentum map satisfies the following equation:*

$$\mathrm{d}_\mathbf{h} J^{\text{n.h.}}_{\bar{\xi}} = \mathscr{L}_{\tilde{\xi}}(L\eta) \quad \textit{along } \mathcal{C}. \tag{7.4}$$

Proof: Equation (7.1) from the proof of proposition 2.1 can be used without modification:

$$\begin{aligned}
\mathrm{d}_\mathbf{h} J^{\text{n.h.}}_{\bar{\xi}} &= i_\mathbf{h}\mathscr{L}_{\tilde{\xi}}\Theta_L - i_\mathbf{h} i_{\tilde{\xi}}\mathrm{d}\Theta_L - \mathrm{d}i_\mathbf{h} i_{\tilde{\xi}}\Theta_L \\
&= i_\mathbf{h}\mathscr{L}_{\tilde{\xi}}\Theta_L + i_{\tilde{\xi}}(n\Omega_L + \zeta) - n\mathscr{L}_{\tilde{\xi}}\Theta_L + n i_{\tilde{\xi}}\mathrm{d}\Theta_L.
\end{aligned}$$

In this expression, we have substituted the constrained De Donder-Weyl equation: ζ is an element of $\mathcal{I}(F)$. As ζ can be written as $\zeta = \lambda_{\alpha\mu} \mathrm{d}x^\mu \wedge \Phi^\alpha$ (see the proof of theorem 2.5 in chapter 6), with Φ^α taking values in the bundle F, we may conclude that $i_{\tilde{\xi}}\zeta = 0$. Therefore, we end up with

$$\begin{aligned} \mathrm{d}_{\mathbf{h}} J^{\text{n.h.}}_{\tilde{\xi}} &= i_{\mathbf{h}} \mathscr{L}_{\tilde{\xi}} \Theta_L - n \mathscr{L}_{\tilde{\xi}} \Theta_L \\ &= \mathscr{L}_{\tilde{\xi}} i_{\mathbf{h}} \Theta_L - i_{[\tilde{\xi}, \mathbf{h}]} \Theta_L - n \mathscr{L}_{\tilde{\xi}} \Theta_L \\ &= \mathscr{L}_{\tilde{\xi}} (L\mu), \end{aligned}$$

where we have used the remark following lemma 1.2 to conclude that $i_{[\tilde{\xi}, \mathbf{h}]}\Theta_L = 0$. ⋄

We finish by noting that in the case where $\tilde{\xi}$ can be written as $j^1\xi_Y$ (for example, when $\bar{\xi}$ is a constant section), we may conclude from the $\mathcal{G}$-invariance of L that $\mathrm{d}_{\mathbf{h}} J^{\text{n.h.}}_{\tilde{\xi}} = 0$.

4. The Cauchy formalism

Until now, all of our results have been derived in a purely covariant setting where all of the coordinates on the base space X are treated on an equal footing. We will now break covariance by making the transition to the Cauchy framework. We use the conventions of section 3 in chapter 1.

In the previous chapter, we showed that the nonholonomic field equations on $J^1\pi$ formally induce the equations of motion for a nonholonomic mechanical system on the space of Cauchy data. In this section, we complete that picture by showing that there is a similar transition for some of the symmetry aspects discussed above. In particular, we will show that the covariant Noether theorem induces a version of the nonholonomic momentum lemma (see [**13**, **20**]) on the space of Cauchy data.

We recall from section 3.1 in chapter 1 that a vector field V on $J^1\pi$ induces a vector field $\tilde{V}$ on $\tilde{Z}$ by composition: $\tilde{V}(\kappa) = V \circ \kappa$, and that an $(n+k)$-form α on $J^1\pi$ induces a k-form $\tilde{\alpha}$ on $\tilde{Z}$ by integration as in (1.23). In addition, we recall that the covariant field equations induce a dynamical system Γ on $\tilde{Z}$ whose determining equations are formally identical to those of a time-dependent mechanical system with an infinite-dimensional configuration space

$$i_\Gamma \tilde{\Omega}_L = 0 \quad \text{and} \quad i_\Gamma \tilde{\eta} = 1. \tag{7.5}$$

and that, in the case of nonholonomic field theory, the induced dynamical system on $\tilde{Z}$ is determined by

$$i_\Gamma \tilde{\Omega}_L \big|_{\tilde{\mathcal{C}}} \in \tilde{F} \quad \text{and} \quad \Gamma \in T\tilde{\mathcal{C}}. \tag{7.6}$$

In both cases, the vector field Γ is induced by the corresponding solution of the unconstrained or constrained De Donder-Weyl equation.

In the next sections, we will exhibit the structures on $\tilde{Z}$ induced by the (nonholonomic) momentum map and we will show how the covariant momentum equation give rises to a momentum equation on $\tilde{Z}$ which is formally identical to the one encountered in nonholonomic mechanics (see for example [**13**, **20**]).

The component J_ξ of the covariant momentum map is an n-form on $J^1\pi$. Because of (1.23), it induces a map $\tilde{J}_\xi \in C^\infty(\tilde{Z})$ on the space of Cauchy data:

$$\tilde{J}_\xi(\kappa) = \int_M \kappa^* J_\xi.$$

In the constrained case, there is a similar definition for the map $\tilde{J}_\xi^{\text{n.h.}}$ in the Cauchy formalism, induced by the component $J_\xi^{\text{n.h.}}$ of the constrained momentum map. Note that $J_\xi^{\text{n.h.}}$ is defined along $\mathcal{C}$.

4.1. The unconstrained case. We now turn to proving the analogue of Noether's theorem in the Cauchy framework. There are essentially two ways in which one could approach this problem: either by directly defining the action of $\mathcal{G}$ on $\tilde{Z}$ and using the standard techniques known from mechanics, or by showing that the covariant Noether theorem leads in a straightforward way to a corresponding theorem on the space of Cauchy data. We choose to follow the second approach, as it allows us to postpone to the very end all of the technical matters associated with the calculus on infinite-dimensional manifolds.

Proposition 4.1. *Let Υ be a connection in π_1 such that the associated horizontal projector* $\mathbf{h}$ *is a solution of the De Donder-Weyl equation (1.17). Let $\tilde{J}$ be the momentum map associated to the covariant momentum map J. Then Noether's theorem holds: $\Gamma(\tilde{J}_\xi) = 0$ for all $\xi \in \mathfrak{g}$, where Γ is a solution to the equations of motion (7.5) in the Cauchy formalism.*

Proof: We will use the following characterisation of the exterior derivative $\mathrm{d}\tilde{J}_\xi$ in terms of $\mathrm{d}J_\xi$:

$$\left\langle \tilde{V}, \mathrm{d}\tilde{J}_\xi \right\rangle(\kappa) = \int_M \kappa^*(i_{\tilde{V}} \mathrm{d}J_\xi),$$

for an arbitrary vector field $\tilde{V}$ on $\tilde{Z}$. For a proof, we refer to [**91**, prop. 3.3.9] or to the expressions used in [**49**, lemma 5.1].

The embedding $\kappa : M \hookrightarrow J^1\pi$ can be written as $\kappa = j^1\phi \circ \tau$. Without loss of generality, we may take ϕ to be a solution of the field equations. This lies at the heart of the Cauchy analysis: κ specifies the values of the fields and their derivatives on a hypersurface and due to the (supposed) hyperbolicity of the equations of motion, the subsequent evolution is then fixed (and given by $j^1\phi$). Formally, let $t \mapsto c(t)$ be an integral curve of Γ such that $c(0) = \kappa$. Then $j^1_x\phi = [c(t)](u)$, where $x = \Psi(t, u)$ (and $\Psi : \mathbb{R} \times M \to X$ is a splitting of X as in section 3.1.3 in chapter 1).

We then have, noting that $\mathbf{h}(\mathbf{T}) = Tj^1\phi(\mathbf{T})$,

$$\left\langle \Gamma, \mathrm{d}\tilde{J}_\xi \right\rangle(\kappa) = \int_M \kappa^*(i_\Gamma \mathrm{d}J_\xi) = \int_M \tau^*(j^1\phi)^*(i_{\mathbf{h}(\mathbf{T})}\mathrm{d}J_\xi) = \int_M \tau^* i_{\mathbf{T}}((j^1\phi)^*\mathrm{d}J_\xi).$$

As we pointed out in the remark following proposition 2.1, one can check that $(j^1\phi)^*\mathrm{d}_{\mathbf{h}}\alpha$ is equal to $\mathrm{d}(j^1\phi)^*\alpha$ if and only if $j^1\phi$ is an integral section of $\mathbf{h}$. We conclude that

$$\left\langle \Gamma, \mathrm{d}\tilde{J}_\xi \right\rangle(\kappa) = \int_M \tau^* i_{\mathbf{T}}((j^1\phi)^*\mathrm{d}_{\mathbf{h}}J_\xi). \tag{7.7}$$

As the ξ-component J_ξ of the covariant momentum map satisfies Noether's theorem, *i.e.* $\mathrm{d}_{\mathbf{h}}J_\xi = 0$, we have that $\Gamma(\tilde{J}_\xi) = 0$. This establishes the theorem of Noether in the Cauchy framework. ⋄

4.2. The constrained case. Quite surprisingly, much of the material developed in the preceding section carries over quite naturally to the constrained case. In particular, for the nonholonomic momentum map, equation (7.7) still holds:

$$\left\langle \Gamma, \mathrm{d}\tilde{J}^{\mathrm{n.h.}}_{\bar{\xi}} \right\rangle(\kappa) = \int_M \tau^* i_{\mathbf{T}}((j^1\phi)^*\mathrm{d}_{\mathbf{h}}J^{\mathrm{n.h.}}_{\bar{\xi}}), \quad \text{for } \kappa \in \mathcal{C},$$

where we attribute a similar meaning to all terms involved: $\mathbf{h}$ is a solution of the constrained De Donder-Weyl equation, $j^1\phi$ is an integral section of the corresponding connection and $\Gamma = \mathbf{h}(\mathbf{T})$. Note that Γ is now a solution of (7.6).

Now, if $J^{\mathrm{n.h.}}_{\bar{\xi}}$ satisfies the nonholonomic momentum equation, then

$$\left\langle \Gamma, \mathrm{d}\tilde{J}^{\mathrm{n.h.}}_{\bar{\xi}} \right\rangle(\kappa) = \int_M \tau^* i_{\mathbf{T}}((j^1\phi)^*\mathscr{L}_{\tilde{\xi}}(L\eta)). \tag{7.8}$$

In the following proposition, we further elaborate the right-hand side. We recall that the vector field $\tilde{\xi}$ on $J^1\pi$ naturally induces a vector field $\hat{\xi}$ on $\tilde{Z}$ by putting $\hat{\xi}(\kappa) = \tilde{\xi} \circ \kappa$.

Proposition 4.2. *Let Υ be a connection on π_1 such that along the constraint submanifold $\mathcal{C}$ the associated horizontal projector $\mathbf{h}$ satisfies the constrained De Donder-Weyl equation (6.9). Assume a Lie group $\mathcal{G}$ acts in the way described in (7.2) and let $\tilde{J}^{\mathrm{n.h.}}$ be the momentum map associated to the covariant momentum map $J^{\mathrm{n.h.}}$. Then $\tilde{J}^{\mathrm{n.h.}}$ satisfies the nonholonomic momentum equation: for all $\bar{\xi} \in \mathfrak{g}^{\mathcal{E}}$,*

$$\Gamma(\tilde{J}^{\mathrm{n.h.}}_{\bar{\xi}}) = \hat{\xi}(\tilde{L}) \quad \textit{along } \mathcal{C}.$$

Proof: We rewrite the right-hand side of (7.8) by performing exactly the opposite manipulations as we did to obtain eq. (7.7). This leads to

$$\left\langle \Gamma, \mathrm{d}\tilde{J}^{\mathrm{n.h.}}_{\bar{\xi}} \right\rangle(\kappa) = \int_M \kappa^* i_{\mathbf{h}(\mathbf{T})}\mathscr{L}_{\tilde{\xi}}(L\eta) = \int_M \kappa^*\mathscr{L}_{\tilde{\xi}}(i_{\mathbf{h}(\mathbf{T})}(L\eta)) + \int_M \kappa^* i_{[\mathbf{h}(\mathbf{T}),\tilde{\xi}]}(L\eta).$$

The last term vanishes as $L\eta$ is semi-basic and $[\mathbf{h}(\mathbf{T}), \tilde{\xi}]$ is π_1-vertical ($\tilde{\xi}$ is π_1-vertical). By lemma 3.3.9 of [**91**], we see that the first term on the right-hand side equals

$$\int_M \kappa^* \mathscr{L}_{\tilde{\xi}}(i_{\mathbf{h}(\mathbf{T})}(L\eta)) = \mathscr{L}_{\hat{\xi}}(\tilde{L}),$$

and this proves the momentum equation in the Cauchy formalism. ◇

Chapter 8

Holonomic and affine nonholonomic constraints

In mechanics, a holonomic constraint is a constraint on the configuration space, while a nonholonomic constraint is a nonintegrable constraint on the velocity space. In field theory, there is apparently no such obvious distinction, especially for the case where a canonical splitting of the base space exists (for instance elasticity).

Already in chapter 6 it became clear that it is indeed not so clear to decide whether, for a given constraint, the nonholonomic approach is the right one or not. For instance, a naive generalisation of the definitions from mechanics would lead to the belief that the incompressibility constraint, being nonintegrable, would be a good candidate to be a true nonholonomic constraint. However, if it wasn't for proposition 3.1 in chapter 6, additional terms would have appeared in the nonholonomic field equations which are absent from the traditional equations governing incompressible fluid dynamics.

Even worse, the literature abounds with nonintegrable constraints on $J^1\pi$ for which no analogue of proposition 3.1 can be found, and for these constraints, the nonholonomic method would probably not give the right field equations (see also the discussion in section 1.2 of the next chapter). It seems that we therefore need a more sophisticated criterion to make the distinction between nonholonomic and other constraints. This was also noted in [**81**].

Many of these problems can be understood using the interplay between the covariant and the Cauchy formulation. The incompressibility constraint, and the constraints to which we alluded in the previous paragraph all have one important property in common: they do not involve derivatives with respect to time. As a consequence, as we shall see, they induce holonomic constraints on the space of Cauchy data, meaning that if anyone were to start immediately from the Cauchy formulation, he or she, using conventional wisdom from classical mechanics, would treat these constraints as holonomic (for example, by adding to the Lagrangian a linear combination of the constraints). Therefore any covariant formulation on the jet bundle must ultimately give the same results when making the transition to the Cauchy framework.

This issue is treated below in section 1.3. For constraints that induce holonomic constraints on the space of Cauchy data, which we call *non-covariant holonomic constraints*, this distinction is somewhat artificial. However, this discussion will provide us with a number of criteria for a true nonholonomic constraint in field theory, which will be used in the next chapter to construct a physically relevant example of a nonholonomic field theory.

In the remainder of this chapter, we then return to the covariant framework for nonholonomic constraints (see chapter 6). We extend the vakonomic approach to the case of constrained field theories, and finally, as a nice side-result, we derive a geometric criterion for the equivalence of the vakonomic and the nonholonomic approaches. This is an extension of a celebrated method of Cortés *et al.* [**27**]. Briefly speaking, we interpret the constraint distribution as the horizontal distribution of an Ehresmann connection: if the curvature of this connection vanishes (which is equivalent to the integrability of the constraints), then the nonholonomic and the vakonomic approach agree.

1. Holonomic constraints

1.1. Covariant holonomic constraints. Let $\pi : Y \to X$ be a fibre bundle as in the preceding chapters. A distribution D on Y is said to be *weakly horizontal (with respect to π)* if there exists a distribution W contained in $V\pi$ such that $D \oplus W = TY$. If W is the whole of $V\pi$, then D is the horizontal distribution of a connection on π. See [**65**] for more information on weakly horizontal distributions.

A weakly horizontal distribution D induces an affine submanifold $\mathcal{C} \hookrightarrow J^1\pi$ defined as follows: γ is an element of $\mathcal{C}$ if γ, viewed as map from T_xX to T_yY (where $x = \pi_1(\gamma)$ and $y = \pi_{1,0}(\gamma)$), takes values in $D(y)$. In coordinates, if the annihilator D° is spanned by the k forms $A^\alpha_a \mathrm{d}y^a + A^\alpha_\mu \mathrm{d}x^\mu$, then $\mathcal{C}$ is the zero level set of the $k(n+1)$ functions $\psi^\alpha_\mu \equiv A^\alpha_a y^a_\mu + A^\alpha_\mu$. Note that weak horizontality implies that the matrix A^α_a has maximal rank k.

If D is integrable, the constraints induced by D are said to be *holonomic*: in that case, $j^1\phi$ takes values in $\mathcal{C}$ if and only if ϕ takes values in a fixed leaf of the foliation induced by D, and we conclude that the constraints can be integrated to constraints on Y.

From a purely covariant point of view, it is therefore a straightforward matter to decide whether a constraint is holonomic or not. However, the matter is more complicated as there exists a large class of constraints for which physical reasoning suggests that a holonomic treatment is appropriate, whereas these constraints are certainly not integrable in the sense indicated above. This will be illustrated in the next section.

Remark 1.1. Constrained field theories with holonomic and affine nonholonomic constraints were also treated in great detail by Krupková and Volný in [**66**]. ◇

1.2. Non-covariant holonomic constraints. Let us consider, as an example, the constraint of incompressibility $\mathcal{J} = \det(y^a_i) - 1$ from section 3 in chapter 6. This constraint does not involve the derivative of the fields with respect to time, and we shall see that it becomes a constraint on the configuration space once we make the transition to the Cauchy framework.

We recall from section 4 in chapter 1 that the dynamics of a fluid in $\mathbb{R}^3$ can be modelled by considering sections of the trivial bundle $\pi : X \times \mathbb{R}^3 \to X$, where $X = \mathbb{R} \times \mathbb{R}^3$. For such a field theory, the space of Dirichlet data is diffeomorphic to $\mathbb{R} \times C^\infty(\mathbb{R}^3, \mathbb{R}^3)$, while the space of Cauchy data is diffeomorphic to $\mathbb{R} \times TC^\infty(\mathbb{R}^3, \mathbb{R}^3)$ (proposition 4.5 in chapter 1): roughly speaking, an element κ of $\tilde{Z}$ can be represented in this case as a triple $[t, \mathrm{Jac}(\phi), \psi]$, where $\phi : \mathbb{R}^3 \to \mathbb{R}^3$ is a smooth map with Jacobian $\mathrm{Jac}(\phi)$ and ψ is a vector field on $\mathbb{R}^3$ along ϕ. The diffeomorphism between $\tilde{Z}$ and $\mathbb{R} \times TC^\infty(\mathbb{R}^3, \mathbb{R}^3)$ is then given by $[t, \mathrm{Jac}(\phi), \psi] \mapsto (t, \psi)$. Note that ϕ, and hence $\mathrm{Jac}(\phi)$, is completely determined by the specification of ψ, because $\phi = \tau_{\mathbb{R}^3} \circ \psi$, where $\tau_{\mathbb{R}^3} : T\mathbb{R}^3 \to \mathbb{R}^3$ is the tangent bundle projection.

Let $\mathcal{C}$ be the hypersurface in $J^1\pi$ determined by the incompressibility constraint, and consider the induced submanifold $\tilde{\mathcal{C}}$ of $\tilde{Z}$ as in (6.17). Using the diffeomorphism between $\tilde{Z}$ and $\mathbb{R} \times TC^\infty(\mathbb{R}^3, \mathbb{R}^3)$ outlined above, $\tilde{\mathcal{C}}$ gives rise to a submanifold of $\mathbb{R} \times TC^\infty(\mathbb{R}^3, \mathbb{R}^3)$, which we also denote by $\tilde{\mathcal{C}}$. However, since the incompressibility constraint does not involve derivatives with respect to time, $\tilde{\mathcal{C}}$ is induced by the submanifold $\mathcal{M}$ of $\mathbb{R} \times C^\infty(\mathbb{R}^3, \mathbb{R}^3)$ consisting of pairs (t, ϕ) such that $\phi^*\mu = \mu$, where μ is the Euclidian volume form on $\mathbb{R}^3$. We conclude that, ultimately, the incompressibility constraint is determined by the manifold $\mathcal{M}$. As $\mathcal{M}$ is a submanifold of the configuration space, we may say that this constraint is *holonomic*. This is a consequence of the fundamental fact that the incompressibility constraint does not involve derivatives with respect to time.

This can be generalized to other classes of constraints as in the following definition, where we use the general setting of elasticity as described in section 4 in chapter 1 (*i.e.* $X = \mathbb{R} \times M$ and $Y = X \times S$). It should be emphasized that this definition only makes sense for field theories whose base space is $\mathbb{R} \times M$: only in that case can a meaningful distinction be made between spatial derivatives and derivatives with respect to time (lemma 4.2 in chapter 1).

Definition 1.2. *Let J be the vertical endomorphism on TS (see (1.1)). A constraint submanifold $\mathcal{C}$ is said to be* non-covariantly holonomic *if for any vector $v \in T(TS)$ in the image of J, $T\Phi^{-1}(v)$ is tangent to $\mathcal{C}$, where Φ is the isomorphism between $J^1\pi$ and $\mathbb{R} \times [J^1(M, S) \times_S TS]$ defined in lemma 4.2 in chapter 1.*

Note that the image of J in $T(TS)$ is spanned by the coordinate vector fields $\frac{\partial}{\partial \dot{y}^a}$, and under the isomorphism Φ these vector fields are pulled back to the coordinate vector fields $\frac{\partial}{\partial y^a_0}$. If we therefore suppose that $\mathcal{C}$ is locally defined by the vanishing of the k

constraint functions φ^α, then this definition is an intrinsic restatement of the fact that

$$\frac{\partial \varphi^\alpha}{\partial y^a_0} = 0,$$

i.e. the constraint functions do not depend on derivatives with respect to time.

In conclusion, for field theories such as elasticity, where a natural splitting of the base space is present, non-covariant holonomic constraints have to be treated as holonomic constraints, even though they may be non-integrable. In the next chapter, we will investigate how constraints should be treated that do involve derivatives with respect to time. We will then establish a new Chetaev principle, which singles out time derivatives, and discards spatial derivatives.

Remark 1.3. We have seen that noncovariant holonomic constraints have to be treated as holonomic. This is done by adding to the Lagrangian L a term $\lambda_\alpha \varphi^\alpha$, and hence corresponds to the vakonomic treatment of constraints (see [**81**]). A few remarks are in order concerning this approach.

(1) The incompressibility constraint is useful to motivate the definition of a noncovariant holonomic constraint, but this example is slightly misleading as well. Recall from proposition 3.1 in chapter 6 that the incompressibility constraint can be written as a divergence. In this rather special case, both the nonholonomic approach of chapter 6 and the vakonomic treatment (see [**81**]) yield the same results.

(2) Truesdell and Noll (see [**100**, sec. 30]) propose the so-called *principle of determinism* to deal with noncovariant holonomic constraints. According to this principle, if the material is hyperelastic, one adds to the stored energy function W a linear combination of the constraints:

$$W \rightsquigarrow W + \lambda_\alpha \varphi^\alpha, \tag{8.1}$$

where the functions λ_α are some Lagrange multipliers. As the Lagrangian for continuum mechanics is $L = \frac{\rho}{2} g(v,v) - W$, this replacement in fact corresponds to the vakonomic treatment of constraints, which is also the method used in [**81**]. ⋄

2. The Skinner-Rusk approach for constrained field theories

The remainder of this chapter will be devoted to the comparison of the vakonomic and the nonholonomic treatment for constrained field theories. We shall assume that these constraints are affine, in the sense of being induced by a weakly horizontal distribution as in section 1.1. In this chapter, we shall follow the work of Cortés *et al.* [**27**], who used the so-called formulation of Skinner and Rusk to recast both models in a form which allows a straightforward geometric comparison.

As the Skinner-Rusk formulation for field theory is not yet universally known, we devote this section to a brief revision of this theory. A full treatment can be found in [**32**,**43**].

In [**97**, **98**], Skinner and Rusk reformulated the equations of motion of a mechanical system as a presymplectic system on $TQ \oplus T^*Q$. Their idea in studying this system was to obtain a common framework for both regular and singular dynamics. Over the years, however, the framework of Skinner and Rusk was extended in many directions. For our purposes, the most important contributions are [**32**, **43**], where the authors reformulated classical field theories as a Skinner-Rusk type system on $J^1\pi \times \bigwedge_2^{n+1} Y$.

2.1. The bundle framework. For notational convenience later on, we henceforth denote the usual system of bundle coordinates on Y as (x^μ, y^A), for $\mu = 1, \ldots, n+1$ and $A = 1, \ldots, m$.

Let $\bigwedge_2^{n+1} Y$ be the bundle of $(n+1)$-forms on Y satisfying the following property:

$$\alpha \in (\textstyle\bigwedge_2^{n+1} Y)_y \quad \text{if} \quad i_v i_w \alpha = 0 \quad \text{for all } v, w \in (V\pi)_y.$$

In coordinates, an element α of $\bigwedge_2^{n+1} Y$ can be represented as $\alpha = p_A^\mu \mathrm{d}y^A \wedge \mathrm{d}^n x_\mu + p \mathrm{d}^{n+1}x$. Hence, on $\bigwedge_2^{n+1} Y$, we have a coordinate system $(x^\mu, y^A; p_A^\mu, p)$. The bundle $\bigwedge_2^{n+1} Y$ is of fundamental interest in classical field theory, because it can be equipped with a natural multisymplectic form, which is the generalisation to higher degree of the symplectic form on a cotangent bundle. If we introduce first the $(n+1)$-form Θ as

$$\Theta(\alpha)(v_1, \ldots, v_{n+1}) = \alpha(T\rho(v_1), \ldots, T\rho(v_{n+1})), \quad \text{where } v_1, \ldots, v_{n+1} \in T_\alpha(\textstyle\bigwedge_2^{n+1} Y)$$

and where $\rho : \bigwedge_2^{n+1} Y \to Y$ is the bundle projection, then this multisymplectic form is defined by setting $\Omega := -\mathrm{d}\Theta$ (see [**22**]).

The central stage for Skinner-Rusk theories is the product bundle $J^1\pi \times \bigwedge_2^{n+1} Y \to Y$. On this bundle, there exists a duality pairing $\langle \cdot, \cdot \rangle : J^1\pi \times \bigwedge_2^{n+1} Y \to \mathbb{R}$, which is reminiscent of the obvious pairing by duality on $TQ \oplus T^*Q$, the bundle originally considered by Skinner and Rusk. This pairing is defined as follows: let $\alpha_y \in (\bigwedge_2^{n+1} Y)_y$ and $j_x^1\phi \in J^1\pi$, such that $\pi_{1,0}(j_x^1\phi) = y$. Now, consider an $(n+1)$-form $\tilde{\alpha}$ on Y extending α_y, *i.e.* such that $\tilde{\alpha}(y) = \alpha_y$. The pullback $(\phi^*\tilde{\alpha})(x)$ is then a form at x of maximal degree, and hence a multiple $a(x)$ of the volume form: $(\phi^*\tilde{\alpha})(x) = a(x)\eta_x$. We now define the duality pairing as

$$\left\langle j_x^1\phi, \alpha \right\rangle := a(x). \tag{8.2}$$

One can easily check that this definition is independent of the extension of α. In coordinates, we have that $a(x) = p_A^\mu y_\mu^A + p$.

3. Skinner-Rusk formulation of vakonomic field theories

Let $\iota : \mathcal{C} \hookrightarrow J^1\pi$ be a constraint submanifold of codimension $k(n+1)$ in $J^1\pi$, locally annihilated by $k(n+1)$ functionally independent constraint functions Ψ_μ^α, where $\alpha =$

$1,\ldots,k$ and $\mu = 1,\ldots,n+1$. Further on, $\mathcal{C}$ will be induced by a weakly horizontal distribution as in section 1, but for now this is not required. As in chapter 6, we assume that $(\pi_{1,0})_{|\mathcal{C}}$ is a fibration, such that it is possible to choose locally an adapted coordinate system $(x^\mu; y^A; y^a_\mu, y^\alpha_\mu)$ on $J^1\pi$, and functions $\Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu)$ such that $\mathcal{C}$ is locally determined by the following set of $k(n+1)$ equations:

$$y^\alpha_\mu - \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu) = 0. \tag{8.3}$$

Hence, $(x^\mu; y^A; y^a_\mu)$ define coordinates on $\mathcal{C}$. We now redefine Ψ^α_μ as $y^\alpha_\mu - \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu)$; note that the zero level set of these functions is still $\mathcal{C}$.

3.1. Direct derivation. The *vakonomic approach* to the constrained problem specified by a Lagrangian L and a constraint manifold $\mathcal{C}$ consists of looking for extremals of the following augmented Lagrangian: $L_{\rm vak} = L + \lambda^\mu_\alpha \Psi^\alpha_\mu$ (see [**81**]), where the functions λ^μ_α are Lagrange multipliers. In other words, we impose the constraints on the space of sections where the action is defined, rather than on the variations, as in nonholonomic field theory.

Let $\tilde{L} := \iota^* L : \mathcal{C} \to \mathbb{R}$ be the induced Lagrangian on $\mathcal{C}$. By looking for extremals of the action associated to $L_{\rm vak}$, and rewriting the resulting extremality conditions in terms of $\tilde{L}$, we obtain the following *vakonomic field equations*:

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial \tilde{L}}{\partial y^a_\mu} - \lambda^\nu_\alpha \frac{\partial \Phi^\alpha_\nu}{\partial y^a_\mu}\right) = \frac{\partial \tilde{L}}{\partial y^a} - \lambda^\nu_\alpha \frac{\partial \Phi^\alpha_\nu}{\partial y^a} \tag{8.4}$$

together with

$$\frac{\mathrm{d}\lambda^\mu_\alpha}{\mathrm{d}x^\mu} = \frac{\partial \tilde{L}}{\partial y^\alpha} - \lambda^\mu_\beta \frac{\partial \Phi^\beta_\mu}{\partial y^\alpha} \quad \text{and} \quad y^\alpha_\mu = \Phi^\alpha_\mu. \tag{8.5}$$

3.2. Skinner-Rusk formulation. Consider now the Cartesian product bundle $\pi_{W_0} : W_0 := \mathcal{C} \times \bigwedge^{n+1}_2 Y \to Y$. Define also the projection $\pi_0 : W_0 \to X$ by putting $\pi_0 = \pi \circ \pi_{W_0}$. The given Lagrangian L induces a function $\mathcal{H}_{\rm vak}$, called *generalized Hamiltonian*, on W_0, defined as follows:

$$\mathcal{H}_{\rm vak}(j^1_x\phi, \alpha) = \left\langle j^1\phi, \alpha \right\rangle - \tilde{L}(j^1_x\phi), \quad \text{for all } (j^1_x\phi, \alpha) \in (W_0)_y, \tag{8.6}$$

where $\langle \cdot, \cdot \rangle$ is the pairing between $J^1\pi$ and $\bigwedge^{n+1}_2 Y$ defined in (8.2), and $\tilde{L} = \iota^* L$ is again the restriction of L to $\mathcal{C}$. In coordinates, we have $\mathcal{H}_{\rm vak} = p^\mu_a y^a_\mu + p^\mu_\alpha \Phi^\alpha_\mu + p - L(x^\mu, y^A, y^a_\mu, \Phi^\alpha_\mu)$.

The multisymplectic form Ω on $\bigwedge^{n+1}_2 Y$ can be used, together with the generalized Hamiltonian $\mathcal{H}_{\rm vak}$, to define a pre-multisymplectic form $\Omega_{\mathcal{H}_{\rm vak}}$ on W_0:

$$\Omega_{\mathcal{H}_{\rm vak}} = \Omega + \mathrm{d}\mathcal{H}_{\rm vak} \wedge \eta.$$

In terms of this form, the Skinner-Rusk field equations are given by

$$i_{\mathbf{h}} \Omega_{\mathcal{H}_{\rm vak}} = n \Omega_{\mathcal{H}_{\rm vak}}, \tag{8.7}$$

where $\mathbf{h}$ is the horizontal projector of a connection on π_0 (see [**32**, **43**]). We will show that these equations are equivalent to the vakonomic field equations (8.4) and (8.5). In brief, we will construct a sequence of submanifolds

$$\ldots \hookrightarrow W_3 \hookrightarrow W_2 \hookrightarrow W_1 \hookrightarrow W_0 = J^1\pi \times \textstyle\bigwedge_2^{n+1} Y.$$

where W_1, W_2 and W_3 admit the following interpretation:

(1) W_1 consists of points where a solution $\mathbf{h}$ of (8.7) exists;
(2) W_2 contains the points of W_1 where the image of the solution $\mathbf{h}$ is tangent to W_1;
(3) W_3 is defined by an additional technical assumption, to be specified later on.

Under a certain regularity condition, W_1 and W_2 coincide and only the manifolds W_0, W_1 and W_3 come into play. In the general case, one needs to apply some form of Gotay's constraint algorithm to formulate the dynamics on a final constraint submanifold W_∞, but this will not be considered here.

Let us now turn to the construction of W_1, W_2, and W_3. Notice that the field equation (8.7) does not necessarily have a solution on the whole of W_0. Hence, we introduce a subset $W_1 \hookrightarrow W_0$, defined as the set of points of W_0 for which there does exist a horizontal projector of a connection on $\pi_0 : \mathcal{C} \times \bigwedge_2^{n+1} Y \to X$ solving equation (8.7). If $\mathbf{h}$ has the following coordinate expression:

$$\mathbf{h} = \mathrm{d}x^\mu \otimes \Big(\frac{\partial}{\partial x^\mu} + A_\mu^A \frac{\partial}{\partial y^A} + B_\mu \frac{\partial}{\partial p} + C_{\mu A}^\nu \frac{\partial}{\partial p_A^\nu} + D_{\mu\nu}^a \frac{\partial}{\partial y_\nu^a} \Big), \tag{8.8}$$

for unknown functions A_μ^A, B_μ, $C_{\mu A}^\nu$, and $D_{\mu\nu}^a$, then a brief coordinate calculation shows that W_1 is determined by the following equations:

$$\begin{aligned} p_a^\mu &= -p_\alpha^\nu \frac{\partial \Phi_\nu^\alpha}{\partial y_\mu^a} + \frac{\partial \tilde{L}}{\partial y_\mu^a} \\ &= -p_\alpha^\nu \frac{\partial \Phi_\nu^\alpha}{\partial y_\mu^a} + \frac{\partial L}{\partial y_\mu^a} + \frac{\partial L}{\partial y_\nu^\alpha} \frac{\partial \Phi_\nu^\alpha}{\partial y_\mu^a}. \end{aligned} \tag{8.9}$$

In addition, the connection coefficients have to satisfy the following constraints:

$$\begin{aligned} A_\mu^\alpha = \Phi_\mu^\alpha, \quad A_\mu^a = y_\mu^a \\ C_{\mu A}^\mu + p_\alpha^\mu \frac{\partial \Phi_\mu^\alpha}{\partial y^A} - \frac{\partial \tilde{L}}{\partial y^A} = 0. \end{aligned} \tag{8.10}$$

Let us now assume that W_1 is a manifold. This is a very restrictive assumption, but for the sake of clarity, we adopt it nevertheless. When dealing with real-world applications, it should be verified by calculations, and it can be expected that interesting behaviour may occur in the points where W_1 fails to be a manifold.

Secondly, we define W_2 as the submanifold of W_1 where the image of the horizontal projector $\mathbf{h}$ solving (8.7) is tangent to W_1. This is expressed by the following equation:

$$\mathbf{h}\left(\frac{\partial}{\partial x^\mu}\right)\left(p_a^\nu - \frac{\partial \tilde{L}}{\partial y_\nu^a} + p_\alpha^\kappa \frac{\partial \Phi_\kappa^\alpha}{\partial y_\nu^a}\right) = 0.$$

In coordinates, this implies the following for the connection coefficients of $\mathbf{h}$:

$$C_{\mu a}^\nu - \mathcal{D}_\mu\left(\frac{\partial \tilde{L}}{\partial y_\nu^a}\right) + C_{\mu\alpha}^\kappa \frac{\partial \Phi_\kappa^\alpha}{\partial y_\nu^a} + p_\alpha^\kappa \mathcal{D}_\mu\left(\frac{\partial \Phi_\kappa^\alpha}{\partial y_\nu^a}\right) = 0, \tag{8.11}$$

where $\mathcal{D}_\mu$ is the operator defined as

$$\mathcal{D}_\mu = \frac{\partial}{\partial x^\mu} + y_\mu^a \frac{\partial}{\partial y^a} + \Phi_\mu^\alpha \frac{\partial}{\partial y^\alpha} + D_{\mu\nu}^a \frac{\partial}{\partial y_\nu^a}.$$

Equation (8.11) uniquely determines the coefficients $D_{\mu\nu}^a$ if the following matrix is nonsingular:

$$\mathcal{C}_{ab}^{\mu\nu} = \frac{\partial^2 \tilde{L}}{\partial y_\mu^a \partial y_\nu^b} - p_\alpha^\kappa \frac{\partial^2 \Phi_\kappa^\alpha}{\partial y_\mu^a \partial y_\nu^b}.$$

This we now assume. Hence, W_2 is the whole of W_1. If $\mathcal{C}_{ab}^{\mu\nu}$ is singular, additional steps in the "constraint algorithm" are necessary. For this procedure, we refer to [**32**].

We end this section by giving a meaning to the coordinate p, and, at the same time, fixing the remaining connection coefficient B_μ. This we do by considering the submanifold W_3 of W_2 defined as

$$W_3 := W_2 \cap \{\mathcal{H}_{\text{vak}}(x^\mu, y^A, y_\mu^a; p_A^\mu) = 0\}.$$

Demanding that a horizontal projector $\mathbf{h}$ on W_2 solving (8.7) is tangent to W_3 leads to the following condition for B_μ:

$$B_\mu + C_{\mu a}^\nu y_\kappa^a + C_{\mu\alpha}^\nu \Phi_\nu^\alpha + D_{\mu\nu}^a p_a^\nu + p_\alpha^\nu \mathcal{D}_\mu(\Phi_\nu^\alpha) - \mathcal{D}_\mu(L) = 0$$

which allows for the determination of B_μ in terms of the other connection coefficients as well as the momenta p_A^μ.

Let us now proceed to derive the vakonomic field equations. On W_3, the Skinner-Rusk equation (8.7) can be locally written as

$$\frac{\mathrm{d}y^a}{\mathrm{d}x^\mu} = \frac{\partial H_{\text{vak}}}{\partial p_a^\mu} \quad \text{and} \quad \frac{\mathrm{d}p_a^\mu}{\mathrm{d}x^\mu} = -\frac{\partial H_{\text{vak}}}{\partial y^a},$$

where H_{vak} is defined on W_3 as $H_{\text{vak}} := -p = p_a^\mu y_\mu^a + p_\alpha^\mu \Phi_\mu^\alpha - \tilde{L}$. By substituting this expression, we finally obtain the following field equations:

$$\frac{\partial \tilde{L}}{\partial y^a} - p_\alpha^\mu \frac{\partial \Phi_\mu^\alpha}{\partial y^a} = \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial \tilde{L}}{\partial y_\mu^a} - p_\alpha^\nu \frac{\partial \Phi_\nu^\alpha}{\partial y_\mu^a}\right)$$

as well as

$$\frac{\mathrm{d}p^\mu_\alpha}{\mathrm{d}x^\mu} = \frac{\partial \tilde{L}}{\partial y^\alpha} - p^\mu_\beta \frac{\partial \Phi^\beta_\mu}{\partial y^a} \quad \text{and} \quad y^\alpha_\mu = \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu).$$

If we identify the momenta p^μ_α with the Lagrange multipliers λ^μ_α, then these equations are precisely the vakonomic field equations (8.4) and (8.5).

Note in passing that, if $\tilde{L}$ is regular, then W_3 is a multisymplectic manifold, with multisymplectic form $\Omega_{W_3} := j^*_{3,0}\Omega_{\mathcal{H}_{\mathrm{vak}}}$, where $j_{3,0} : W_3 \hookrightarrow W_0$ is the canonical injection. This can be verified by a routine coordinate calculation.

Affine constraints. An important simplification occurs when the constraints are affine. In particular, we assume that there exists a fibration $\tau : Y \to Q$ of Y over a new manifold Q, which is fibered in turn over X (see (8.12)). The constraint distribution D will then be taken to be the horizontal distribution of a connection on τ. See the commutative diagram below:

$$\begin{array}{ccc} Y & \xrightarrow{\tau} & Q \\ {\scriptstyle \pi}\downarrow & \swarrow {\scriptstyle \pi'} & \\ X & & \end{array} \tag{8.12}$$

Consider a system of bundle coordinates (x^μ, y^a) on Q, where $\mu = 1, \ldots, n+1$ and $a = 1, \ldots, m-k$, and assume that there exists bundle coordinates on Y adapted to both π and τ, *i.e.* coordinates $(x^\mu; y^a, y^\alpha)$, collectively denoted by (x^μ, y^A), such that τ is locally given by $\tau(x^\mu, y^A) = (x^\mu, y^a)$. In nonholonomic mechanics, a similar setup was studied in [**93**].

Let D be the horizontal distribution of a connection on τ. Since $TY = D \oplus V\tau$, and because $V\tau \subset V\pi$, D is a weakly horizontal distribution on Y, and hence induces a constraint submanifold $\mathcal{C} \hookrightarrow J^1\pi$.

Assume that D is annihilated by the k one-forms $\psi^\alpha = A^\alpha_A \mathrm{d}y^A + A^\alpha_\mu \mathrm{d}x^\mu$ as in section 1.1. Because of the weak horizontality of D, the matrix A^α_A has maximal rank k, and without loss of generality, we can therefore assume that the annihilator D° is locally spanned by the following k forms:

$$\phi^\alpha := \mathrm{d}y^\alpha - B^\alpha_a \mathrm{d}y^a - B^\alpha_\mu \mathrm{d}x^\mu.$$

This basis is generally more suited for our purposes.

In case of affine constraints, the coefficients $D^a_{\mu\nu}$ are determined by the following expression:

$$D^b_{\mu\nu}\frac{\partial^2\tilde L}{\partial y^a_\mu \partial y^b_\nu} = -\frac{\partial^2\tilde L}{\partial x^\mu \partial y^a_\mu} - y^b_\mu \frac{\partial^2\tilde L}{\partial y^b \partial y^a_\mu} - \Phi^b_\mu \frac{\partial^2\tilde L}{\partial y^b \partial y^a_\mu} + \frac{\partial \tilde L}{\partial y^a} + B^\alpha_a \frac{\partial \tilde L}{\partial y^\alpha}$$
$$+ p^\mu_\alpha \left(\frac{\partial B^\alpha_a}{\partial x^\mu} + y^b_\mu \frac{\partial B^\alpha_a}{\partial y^b} + \Phi^\beta_\mu \frac{\partial B^\alpha_a}{\partial y^\beta} - B^\beta_a \frac{\partial \Phi^\beta_\mu}{\partial y^\alpha} - \frac{\partial \Phi^\alpha_\mu}{\partial y^a} \right), \qquad (8.13)$$

where $\Phi^\alpha_\mu = B^\alpha_a y^a_\mu + B^\alpha_\mu$. The expression between brackets in equation (8.13) is closely related to the curvature of D. Indeed, we recall that the curvature R of D is a section of $\bigwedge^2 Y \otimes TY$, locally defined as $R = R^\alpha_{ab}\mathrm{d}y^a \wedge \mathrm{d}y^b \otimes \frac{\partial}{\partial y^\alpha} + R^\alpha_{a\mu}\mathrm{d}y^a \wedge \mathrm{d}x^\mu \otimes \frac{\partial}{\partial y^\alpha}$, where (see (1.7))

$$R^\alpha_{ab} = \frac{\partial B^\alpha_a}{\partial y^b} - \frac{\partial B^\alpha_b}{\partial y^a} + B^\beta_b \frac{\partial B^\alpha_a}{\partial y^\beta} - B^\beta_a \frac{\partial B^\alpha_b}{\partial y^\beta}$$
$$R^\alpha_{a\mu} = \frac{\partial B^\alpha_a}{\partial x^\mu} - \frac{\partial B^\alpha_\mu}{\partial y^a} + B^\beta_\mu \frac{\partial B^\alpha_a}{\partial y^\beta} - B^\beta_a \frac{\partial B^\alpha_\mu}{\partial y^\beta}.$$

(See definition 1.10 in chapter 1). Bearing this in mind, one then obtains for the coefficients $D^a_{\mu\nu}$ the following expression:

$$D^b_{\mu\nu}\frac{\partial^2\tilde L}{\partial y^a_\mu \partial y^b_\nu} = -\frac{\partial^2\tilde L}{\partial x^\mu \partial y^a_\mu} - y^b_\mu \frac{\partial^2\tilde L}{\partial y^b \partial y^a_\mu} - \Phi^b_\mu \frac{\partial^2\tilde L}{\partial y^b \partial y^a_\mu} + \frac{\partial \tilde L}{\partial y^a} + B^\alpha_a \frac{\partial \tilde L}{\partial y^\alpha}$$
$$+ p^\alpha_\mu (R^\alpha_{ab} y^b_\mu + R^\alpha_{a\mu}). \qquad (8.14)$$

These expressions will play an important role in the comparison between vakonomic and nonholonomic dynamics below in section 5.

4. Skinner-Rusk formulation of nonholonomic field theories

A similar, but slightly more involved method can be used to cast the nonholonomic field equations into Skinner-Rusk form. We consider a constraint submanifold $\mathcal{C}$ of codimension $k(n+1)$, determined by similar expressions as in (8.3). The nonholonomic field equations will be recast as a Skinner-Rusk type system on the bundle $\bar\pi_{\bar W_0} : \bar W_0 := J^1\pi \times \bigwedge^{n+1}_2 Y \to Y$.

Consider first the bundle of constraint forms F spanned by the $(n+1)$-forms Φ^α defined in (6.3) using the Chetaev principle. We again denote by $\mathcal{I}(F)$ the ideal in $\Omega^\bullet(J^1\pi)$ generated by F and we use the same notation to denote the pullback of this ideal to $\bar W_0$.

In the nonholonomic case, the generalized Hamiltonian is defined as

$$\mathcal{H}_{\rm nh} := \langle \mathrm{pr}_1, \mathrm{pr}_2 \rangle - \mathrm{pr}_2^* L.$$

Note that $\mathcal{H}_{\rm nh}$ involves the values of L on the whole of $J^1\pi$ and not just on $\mathcal{C}$ as in the vakonomic approach. The pre-multisymplectic form $\Omega_{\mathcal{H}_{\rm nh}}$ is then defined as in section 3 by putting $\Omega_{\mathcal{H}_{\rm nh}} := \Omega + \mathrm{d}\mathcal{H}_{\rm nh} \wedge \eta$.

The nonholonomic field equations are now

$$(i_{\mathbf{k}}\Omega_{\mathcal{H}_{\rm nh}} - n\Omega_{\mathcal{H}_{\rm nh}})_{|\mathcal{C}\times\bigwedge_2^{n+1}Y} \in \mathcal{I}(F) \quad \text{and} \quad (\mathrm{Im}\,\mathbf{k})_{|\mathcal{C}\times\bigwedge_2^{n+1}Y} \subset T(\mathcal{C}\times \textstyle\bigwedge_2^{n+1}Y) \tag{8.15}$$

for a horizontal projector $\mathbf{k}$ on $\bar\pi_0 := \pi \circ \bar\pi_{\bar W_0}$; notice the similarity between these equations and the nonholonomic field equations (6.9). A similar computation as in section 3 shows us that a horizontal projector, with coordinate expression

$$\mathbf{k} = \mathrm{d}x^\mu \otimes \Big(\frac{\partial}{\partial x^\mu} + A^A_\mu \frac{\partial}{\partial y^A} + B_\mu \frac{\partial}{\partial p} + C^\nu_{\mu A}\frac{\partial}{\partial p^\nu_A} + D^A_{\mu\nu}\frac{\partial}{\partial y^A_\nu}\Big), \tag{8.16}$$

is a solution of the nonholonomic field equations if and only if

$$A^A_\mu = y^A_\mu, \quad p^\mu_A = \frac{\partial L}{\partial y^A_\mu} \quad \text{and} \quad C^\mu_{\mu A} = \frac{\partial L}{\partial y^A} + \lambda^\kappa_{\alpha\mu}\frac{\partial \Psi^\alpha_\kappa}{\partial y^A_\mu}, \tag{8.17}$$

where $\Psi^\alpha_\kappa = y^\alpha_\kappa - \Phi^\alpha_\kappa$ and the $\lambda^\kappa_{\alpha\mu}$ are a set of Lagrange multipliers, to be determined by imposing the second part of (8.15). Let us now define a submanifold $\bar W_1$ of $\bar W_0$, specified by the relations (compare with (8.9)):

$$p^\mu_A = \frac{\partial L}{\partial y^A_\mu} \tag{8.18}$$

Again as with vakonomic dynamics, we define the submanifold $\bar W_2 \hookrightarrow \bar W_1$ as the set of points where the image of the solution $\mathbf{k}$ determined by (8.17) is tangent to W_1. This leads to the following conditions:

$$C^\nu_{\mu A} - \frac{\partial^2 L}{\partial x^\mu \partial y^A_\nu} - y^B_\mu \frac{\partial^2 L}{\partial y^B \partial y^A_\nu} - D^B_{\mu\kappa}\frac{\partial^2 L}{\partial y^B_\kappa \partial y^A_\nu} = 0, \tag{8.19}$$

as well as

$$D^\alpha_{\mu\nu} - \frac{\partial \Phi^\alpha_\nu}{\partial x^\mu} - y^A_\mu \frac{\partial \Phi^\alpha_\nu}{\partial y^A} - D^a_{\mu\kappa}\frac{\partial \Phi^\alpha_\nu}{\partial y^a_\kappa} = 0. \tag{8.20}$$

It is easily seen that, in the case of a regular Lagrangian, these conditions do not restrict the submanifold $\bar W_1$ any further, *i.e.* $\bar W_2 = \bar W_1$.

Finally, we define the submanifold $\bar W_3$ as (compare with the definition of W_3 in the vakonomic case):

$$\bar W_3 := \bar W_2 \cap \{\mathcal{H}_{\rm nh}(x^\mu, y^A, y^a_\mu; p^\mu_A) = 0\}.$$

Demanding that a connection $\mathbf{k}$ whose image is tangent to $\bar W_2$ has an image tangent to $\bar W_3$ imposes an additional condition on the the connection coefficient B_μ:

$$B_\mu + C^\nu_{\mu A} y^A_\nu + D^A_{\mu\nu} p^\nu_A - \left(\frac{\partial L}{\partial x^\mu} + A^A_\mu \frac{\partial L}{\partial y^A} + D^A_{\mu\nu}\frac{\partial L}{\partial y^A_\nu}\right) = 0.$$

If we now define H_{nh} along $\bar{W}_3$ as $H_{\mathrm{nh}} := -p = p^\mu_A y^A_\mu - L$, then the nonholonomic Skinner-Rusk equations (8.15) become

$$\frac{\mathrm{d}y^A}{\mathrm{d}x^\mu} = \frac{\partial H_{\mathrm{nh}}}{\partial p^\mu_A} \quad \text{and} \quad \frac{\mathrm{d}p^\mu_A}{\mathrm{d}x^\mu} = -\frac{\partial H_{\mathrm{nh}}}{\partial y^A} + \lambda^\nu_{\alpha\mu}\frac{\partial \Psi^\alpha_\nu}{\partial y^A_\mu},$$

together with the constraint equations $y^\alpha_\mu = \Phi^\alpha_\mu(x^\nu, y^A, y^a_\nu)$. By using the expression for H_{nh} as well as (8.18), we finally obtain that the Skinner-Rusk equations imply the standard nonholonomic field equations:

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^A_\mu}\right) - \frac{\partial L}{\partial y^A} = \lambda^\kappa_{\alpha\mu}\frac{\partial \Psi^\alpha_\kappa}{\partial y^A_\mu},$$

together with the constraints.

Affine constraints. We now focus on affine constraints, and employ a similar convention for the bundle D of constraint forms as in the vakonomic case. In this case, the third equation of (8.17) splits into two sets of equations,

$$C^\mu_{\mu a} = \frac{\partial L}{\partial y^a} - \lambda^\mu_{\alpha\mu} B^\alpha_a \quad \text{and} \quad C^\mu_{\alpha\mu} = \frac{\partial L}{\partial y^\alpha} + \lambda^\mu_{\alpha\mu}.$$

One can combine these two expressions to eliminate the Lagrange multipliers. In the resulting expression, one can then substitute expression (8.19) to eliminate $C^\mu_{\mu A}$, and expression (8.20) to express $D^\alpha_{\mu\nu}$ in terms of $D^a_{\mu\nu}$. After a long computation, we finally obtain

$$\begin{aligned} D^b_{\mu\nu}\frac{\partial^2 \tilde{L}}{\partial y^a_\mu \partial y^b_\nu} = &-\frac{\partial^2 \tilde{L}}{\partial x^\mu \partial y^a_\mu} - y^b_\mu \frac{\partial^2 \tilde{L}}{\partial y^b \partial y^a_\mu} - \Phi^\beta_\mu \frac{\partial^2 \tilde{L}}{\partial y^\beta \partial y^a_\mu} + \frac{\partial \tilde{L}}{\partial y^a} \\ &+ \frac{\partial L}{\partial y^\alpha_\mu}\left(y^b_\mu \frac{\partial B^\alpha_a}{\partial y^b} + \Phi^\beta_\mu \frac{\partial B^\alpha_a}{\partial y^\beta} + \frac{\partial B^\alpha_a}{\partial x^\mu} - \frac{\partial \Phi^\alpha_\mu}{\partial y^a}\right). \end{aligned} \tag{8.21}$$

5. Comparison between both approaches

Definition 5.1. *Let X be a manifold and consider two fibrations $\pi_C, \pi_D : C, D \to X$. Consider a smooth map $f : C \to D$ and let $\mathbf{h}$ be a connection on π_C, and $\mathbf{k}$ a connection on π_D. These connections are then said to be* f-related *if*

$$Tf \circ \mathbf{h}_p = \mathbf{k}_{f(p)} \circ Tf \quad \textit{for all} \quad p \in C.$$

Consider now the vakonomic and nonholonomic manifolds W_3 and $\bar{W}_3$. There exists an obvious surjective submersion $f : W_3 \to \bar{W}_3$, given in coordinates by $f(x^\mu, y^A, y^a_\mu; p^\mu_\alpha) = (x^\mu, y^A, y^a_\mu)$ (see [**27**, **65**]). The map f can be given an intrinsic meaning by using the Legendre transformation.

In order to study the relation between W_3 and $\bar{W}_3$, and hence the relation between vakonomic and nonholonomic classical field theory, we make use of the following observation of Krupková [**65**] and Cortés *et al.* [**27**]: if **h** and **k** were f-related connections, then any integral section of **h** would project down (under f) to an integral section of **k**. The original theorem concerned integral curves of vector fields, but using definition 5.1 also covers integral sections of connections.

Let **h** be a vakonomic connection (with connection coefficients as determined in section 3) and **k** be a nonholonomic connection (with coefficients as in section 4). By considering the set of points S_1 of W_3 where **h** and **k** are f-related, we obtain a first characterization of the equivalence between **h** and **k**. Let us assume that S_1 is not empty, otherwise both connections are entirely unrelated. A comparison of both sets of connection coefficients then shows the following:

Proposition 5.2. *S_1 is locally determined by the vanishing of the following set of functions on W_3:*

$$\varphi_a = \left(\frac{\partial \tilde{L}}{\partial y^\alpha_\mu} - p^\mu_\alpha \right) (R^\alpha_{ab} y^b_\mu + R^\alpha_{a\mu}).$$

Proof: The local expression for S_1 follows by considering the following contracted difference:

$$\varphi_a = \frac{\partial^2 \tilde{L}}{\partial y^a_\mu \partial y^b_\nu} \left(\check{D}^b_{\mu\nu} - \hat{D}^b_{\mu\nu} \right),$$

where $\check{D}$ is the set of vakonomic connection coefficients (8.14), and $\hat{D}$ is the set of nonholonomic coefficients (8.21). ◇

The submanifold S_1 can be seen as the first stage in a certain *constraint algorithm* (see [**27**, **65**]), the result of which is a final submanifold S_∞ (which might be empty) where the vakonomic and nonholonomic dynamics are equivalent. A general discussion of this constraint algorithm would not differ from the treatment of Krupkova and Cortés *et al.* and is hence omitted. We only wish to point out that, if the constraints are holonomic, and hence $R^\alpha_{ab} = R^\alpha_{a\mu} = 0$, then S_1 is the whole of W_3 and vakonomic and nonholonomic dynamics are everywhere equivalent, by which it is confirmed that the vakonomic and nonholonomic description give the same results for holonomic constraints.

Chapter 9
Nonholonomic kinematic constraints in elasticity

1. Introduction

Having established in the previous chapters a theoretical framework for nonholonomic field theories, we now exhibit a simple, physical example of a field theory with nonholonomic constraints. The basic model is that of a Cosserat rod, a special kind of elastic medium. This rod is allowed to move in a horizontal plane which is supposed to be sufficiently rough, so that the rod rolls without sliding. In this way, we obtain what one could reasonably call a continuum version of the vertical rolling disc.

1.1. Cosserat rods. The theory of *Cosserat rods* constitutes an approximation to the full three-dimensional theory of elastic deformations of rod-like bodies. Originally conceived at the beginning of the twentieth century by the Cosserat brothers, it laid dormant for more than fifty years until it was revived by the pioneers of rational mechanics (see [**100**, §98] for an overview of its history). It is now an important part of modern nonlinear elasticity and its developments are treated in great detail for instance in [**2**], which we follow here.

A Cosserat rod can be visualised as specified by a curve $s \mapsto \mathbf{r}(s)$ in $\mathbb{R}^3$, called the *centerline*, to which is attached a frame $\{\mathbf{d}_1(s), \mathbf{d}_2(s), \mathbf{d}_3(s)\}$, called the *director frame* (models with different numbers of directors are also possible). The rough idea is that the centerline characterizes the configuration of the rod when its thickness is neglected, whereas the directors model the configuration of the laminae transverse to the centerline. In the Cosserat theory, the laminae are assumed to deform homogeneously, and therefore the specification of a director frame in $\mathbb{R}^3$ fixed to a lamina completely specifies the configuration of that lamina. For the remainder of this chapter, we shall assume that the laminae can only effect Euclidian motions. In this case, the director frame can be chosen to be orthogonal.

One is encouraged to think of the director frame at a point as a body frame of a rigid body (the lamina). Indeed, the idea of considering a Cosserat rod as a "curve's worth of rigid bodies", going back to the work of Kirchhoff, is a very fruitful one. We will extend it in this chapter by constructing a continuum's worth of nonholonomic mechanical systems. The basic idea is to consider the motion of a Cosserat rod in a horizontal plane, and to impose on this model the constraint of rolling without sliding.

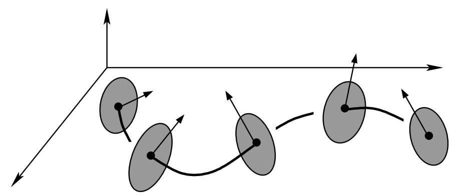


Figure 9.1. Geometry of the constrained rod

This example can be modelled as an elastic medium (in the sense of section 9 in chapter 1). The constraints can be incorporated into this model using the theory of chapter 6, but in order to construct the bundle of reaction forces, some additional care is required. We again have to take into account the distinction between the time derivative and derivatives with respect to the spatial variables, as in chapter 8. The nonholonomic Cosserat rod can also be modelled as the continuum limit of a nonholonomic mechanical system, and as we shall now demonstrate, this is a nice heuristic way of getting an insight into the form of the reaction forces.

Indeed, consider N rigid discs rolling vertically without sliding on a horizontal plane, and assume that these discs are interconnected by flexible beams of length ℓ/N, as in figure 9.1. Now let the number of discs go to infinity, while keeping the total length ℓ fixed: the result is the nonholonomic Cosserat rod.

This mechanical model is interesting for a number of reasons. First of all, the nonholonomic field equations are derived by varying the action with respect to *admissible variations*, and this obviously requires the specification of a bundle of admissible variations, or equivalently, a bundle of reaction forces. In mechanics, this is commonly done by taking recourse to the principle of d'Alembert, which states that the virtual work of the reaction forces is zero. In field theory, this principle can be interpreted in a number of non-equivalent ways, and it is the mechanical model which will eventually determine our choice.

Secondly, our model is a counterexample to the often-held belief that constraints in classical field theories are necessarily vakonomic. In sections 1.2 and 1.3 these two aspects are treated more in detail.

1.2. Relation with other approaches. In a number of papers [**8**, **9**], Bibbona, Fatibene, and Francaviglia contrasted the vakonomic and the nonholonomic treatments for classical field theories, and concluded that for relativistic hydrodynamics only the former gives correct results. Another typical example of a vakonomic constraint is the incompressibility constraint in nonrelativistic fluid dynamics, treated by Marsden *et al.* [**81**]. Many more can be found in Antman's book [**2**] and in the papers by García *et al.* [**45**].

In contrast, our field theory arises as the continuum limit of the vertically rolling disc, a textbook example of a nonholonomic mechanical system. These nonholonomic constraints survive in the continuum limit and hence provide a very strong motivation for the study of nonholonomic techniques in field theories.

It should also be noted that similar theories as ours were explored before by Vignolo and Bruno (see [**109**]). They considered constraints depending only on the time derivatives of the fields, and their resulting analysis is therefore more direct. However, the underlying philosophy is the same: the constraints are "(...) purely kinetic restrictions imposed separately on each point of the continuum".

1.3. Modelling the constraint forces. One important characteristic of the nonholonomic field theories of chapter 6 is that they are fully covariant, *i.e.* no distinction whatsoever is made between the spatial variables and time. In contrast, the nonholonomic Cosserat rod arises as the continuum limit of a mechanical system and for this kind of systems, there is indeed a canonical direction of time.

This noncovariance plays a fundamental role in determining the reaction forces. In this introductory section, we will derive a tentative form of these forces by use of the mechanical analogue, and in subsequent sections, we will then show that these reaction forces can be provided by a "noncovariant" version of the familiar Chetaev principle. Whereas the bundle of reaction forces in chapter 6 was generated by the forms $S^*_\eta(\mathrm{d}\varphi^\alpha)$, we will see that the noncovariant Chetaev reaction forces are linear combinations of the forms $\Phi^\alpha = S^*_{\text{n.c.}}(\mathrm{d}\varphi^\alpha)$, where the noncovariant vertical endomorphism $S_{\text{n.c.}}$ is designed to take into account the difference between space and time.

1.3.1. *Nonholonomic mechanical systems.* The mechanical background is not essential for the description of the continuum theory, but rather serves as a justification for some of our definitions. In particular, it provides a number of valuable clues regarding the type of constraint forces needed to maintain such a nonholonomic constraint. Let S be the configuration space of the vertically rolling disc, so that the configuration space for the entire model, consisting of N discs, is the product space S^N. Denote by $\varphi^\alpha_{(i)}$ the constraints of rolling without sliding imposed on the ith wheel; $\varphi^\alpha_{(i)}$ can be seen as a function on TS^N.

With these conventions, a *motion* of the system is a curve $t \mapsto c(t)$ in S^N, and a variation of such a motion c is then a vector field $(X_1, X_2, \ldots, X_N)$ on S^N along c, *i.e.* a collection of maps $X_i : \mathbb{R} \to TS$ such that $X_i(t) \in T_{c^i(t)}S$ for all $i = 1, \ldots, N$, where $c^i := \mathrm{pr}_i \circ c$.

Let us now consider the one-forms $\Phi^\alpha_{(i)} := J^*(\mathrm{d}\varphi^\alpha_{(i)})$, where J is the vertical endomorphism on TS^N (see (1.1)). In geometric mechanics, linear combinations of these one-forms represent the possible reaction forces at the ith wheel; the bundle F, defined as

$$F := \left\langle \Phi^\alpha_{(1)} \right\rangle \oplus \left\langle \Phi^\alpha_{(2)} \right\rangle \oplus \cdots \oplus \left\langle \Phi^\alpha_{(N)} \right\rangle$$

then represents the totality of all reaction forces along the rod. In coordinates, the one-forms $\Phi^\alpha_{(i)}$ are given by

$$\Phi^\alpha_{(i)} = \frac{\partial \varphi^\alpha_{(i)}}{\partial \dot{y}_{(i)}} \mathrm{d}y_{(i)} \quad \left(= \sum_a \frac{\partial \varphi^\alpha_{(i)}}{\partial \dot{y}^a_{(i)}} \mathrm{d}y^a_{(i)} \right) \quad \text{for all } i = 1, \ldots, N. \tag{9.1}$$

Here, $(y_{(i)}, \dot{y}_{(i)})$ is a coordinate system on the ith factor of TS^N. Note that there is no summation over the index i in (9.1), and that the summation over individual coordinates is implicit, as shown in the term between brackets.

Knowing the precise form of the bundle of reaction forces F is important because the nonholonomic equations of motion are derived by varying the action with respect to admissible variations. Moreover, the principle of d'Alembert tells us that a variation is admissible if it belongs to the annihilator of F, *i.e.* a variation $(X_1, \ldots, X_N)$ of c is admissible if

$$\left\langle \bar{X}_i(t), \alpha(c(t)) \right\rangle = 0, \quad \text{for all } (i,t) \in \{1, \ldots, N\} \times \mathbb{R} \quad \text{and} \quad \alpha \in F, \tag{9.2}$$

where $\bar{X}_i$ is a lift of X_i to $T(TS)$ such that $T\tau_S \circ \bar{X}_i = X_i$. In coordinates, this is equivalent to

$$v_i \frac{\partial \varphi_{(i)}}{\partial \dot{y}_{(i)}} = 0 \quad \text{for all } i = 1, \ldots, N, \tag{9.3}$$

where we have written $X_i = v_i \frac{\partial}{\partial y_{(i)}}$. Note that there is again no summation over i.

In the next paragraph, we will let the number N go to infinity, while keeping the length ℓ constant. The result is a field theory, and a reaction force will be a continuous assignment of a one-form on TS to each point of the centerline of the rod. This definition will be the starting point for our treatment in the main body of the text; from the developments in chapter 6, it follows that once we know the bundle of reaction forces, we can derive the nonholonomic field equations.

1.3.2. *The continuum model.* In the continuum limit, a field is a map ϕ from $\mathbb{R} \times [0, \ell]$ to S, and can be modelled as a section of a trivial bundle π, whose base space is $\mathbb{R} \times [0, \ell]$, and with standard fibre S. The constraints $\varphi_{(i)}$ from the previous paragraph are then replaced by a constraint function φ^α on $J^1\pi$.

As pointed out in section 2.1.2 in chapter 1, an infinitesimal variation of a field ϕ is a map $X : \mathbb{R} \times [0, \ell] \to TS$ with the property that $X(s,t) \in T_{\phi(s,t)}S$; in other words, a vector field along ϕ. Taking our cue from (9.3), we say that a variation X is *admissible* if the following holds (in coordinates):

$$X^a(s,t) \frac{\partial \varphi^\alpha}{\partial y^a_0} = 0,$$

where we have written $X(s,t) = X^a(s,t)\frac{\partial}{\partial y^a}$.

This condition can be rewritten in intrinsic form by using the natural isomorphism between the first jet bundle and the product bundle $\mathbb{R} \times [J^1(M,S) \times_S TS]$ (lemma 4.2 in chapter 1). Now, let J be the vertical endomorphism on TS. This map has a trivial extension to the whole of $\mathbb{R} \times [J^1(M,S) \times_S TS]$, and by using the natural isomorphism with $J^1\pi$, we obtain a map J^* from $T^*(J^1\pi)$ to itself. The bundle F of constraint forces is then generated by the forms $\Phi^\alpha := J^*(\mathrm{d}\varphi^\alpha)$. The similarity with the mechanical case is obvious.

2. Nonholonomic kinematic constraints

Before tackling the Cosserat example, we will sketch in this section an abstract framework for a certain class of classical field theories with nonholonomic constraints, the Cosserat rod being one of these. These field theories are described using a bundle of the type described in section 4 of chapter 1: the base space X is the product $\mathbb{R} \times M$ and the total space Y is $X \times S$. The nonholonomic constraints are "noncovariant", in the sense that they involve derivatives of the fields with respect to time (compare with definition 1.2 in the previous chapter).

The main objective of this section is to propose a suitable definition for the bundle of reaction forces for such constrained field theories. In the next section, this theory will then be applied to the Cosserat rod, and we shall see that the resulting field equations are precisely those which are obtained by taking the continuum limit of the mechanical model described in section 1.3. This serves as a justification for the axiomatic approach that we are about to follow.

2.1. A new vertical endomorphism. Recall the coordinate expression (1.2) of the vertical endomorphism S_η on $J^1\pi$. This tensor field was constructed by Saunders [**94**] using a previously defined map assigning to each one-form ω on X the vector-valued one-form S_ω on $J^1\pi$ given in coordinates by (see [**94**, p. 156])

$$S_\omega = \omega_\mu(\mathrm{d}y^a - y^a_\nu \mathrm{d}x^\nu) \otimes \frac{\partial}{\partial y^a_\mu}, \quad \text{where } \omega = \omega_\mu \mathrm{d}x^\mu. \tag{9.4}$$

Roughly speaking, the vertical endomorphism S_η then arises, once a volume form on X is chosen, by putting

$$S_\eta = S_{\mathrm{d}x^\mu} \dot{\wedge} (\pi_1^* \mathrm{d}^n x_\mu), \tag{9.5}$$

where the wedge operator '$\dot{\wedge}$' is defined as follows: if Φ is a vector-valued k-form on $J^1\pi$, and α is a regular (*i.e.* $\mathbb{R}$-valued) l-form, then $\Phi \dot{\wedge} \alpha$ is the vector-valued $(k+l)$-form given by $\langle \Phi \dot{\wedge} \alpha, \beta \rangle = \langle \Phi, \beta \rangle \wedge \alpha$ for all $\beta \in \Omega^1(J^1\pi)$.

It is obvious that (9.5) is fully covariant, in the sense that no distinction is made between the variables on the base space. However, in elastodynamics, this is not always desirable, as we have a distinguished direction of time. Therefore, we propose the following "non-covariant" vertical endomorphism:

Definition 2.1. *The* non-covariant vertical endomorphism *is the vector-valued* $(n+1)$*-form* $S_{\text{n.c.}}$ *defined as* $S_{\text{n.c.}} := S_{\text{d}t} \dot{\wedge} (\pi_1^* \eta_M)$*, where* $S_{\text{d}t}$ *is the vector-valued one-form associated to* dt *as in (9.4).*

Note that dt is a well defined one-form on $X = \mathbb{R} \times M$; therefore, $S_{\text{n.c.}}$ is an intrinsic object. In coordinates, $S_{\text{n.c.}}$ is given by

$$S_{\text{n.c.}} = (\mathrm{d}y^a - y^a_\mu \mathrm{d}x^\mu) \wedge \mathrm{d}^n x_0 \otimes \frac{\partial}{\partial y^a_0}.$$

Remark 2.2. In section 1.1.1 in chapter 1, we stated that the construction of the vertical endomorphism S_η is related to the affine structure of the bundle $\pi_{1,0} : J^1\pi \to Y$. The noncovariant vertical endomorphism can be understood in a similar vein, by restricting the affine action of $\pi^* T^* X \otimes V\pi$ on $J^1\pi$ to $\pi^* T^* \mathbb{R} \otimes V\pi$. ◇

2.2. The bundle of constraint forces. Let $\iota : \mathcal{C} \hookrightarrow J^1\pi$ be a constraint manifold. As in section 1.1 of chapter 6, in addition to $\mathcal{C}$, we also need to specify a suitable bundle of reaction forces. Let us first describe the general appearance of such a bundle. Later on, we will then see how one can use the noncovariant vertical endomorphism of section 2.1 to give a noncovariant version of the Chetaev principle of chapter 6.

Recall from section 1.1 of chapter 6 that reaction forces are modelled as n-horizontal 1-contact $(n+1)$-forms (defined along $\mathcal{C}$). Forms which satisfy these requirements can be locally expressed as in (6.2). In this section, we attribute a special status to the time coordinate, and therefore we require not only that a reaction force Φ is an n-horizontal one-contact $(n+1)$-form defined along $\mathcal{C}$, but also that the following holds:

$$i_v i_w \Phi = 0$$

for all tangent vectors v, w on $J^1\pi$ such that $T(\mathrm{pr}_1 \circ \pi_1)(v) = T(\mathrm{pr}_1 \circ \pi_1)(w) = 0$, where $\mathrm{pr}_1 : \mathbb{R} \times M \to \mathbb{R}$ is the projection onto the first factor (this condition expresses that v and w do not contain a component proportional to $\frac{\partial}{\partial t}$). In coordinates, this implies that Φ has the following form:

$$\Phi = A_a(\mathrm{d}y^a - y^a_\mu \mathrm{d}x^\mu) \wedge \mathrm{d}^n x_0, \tag{9.6}$$

where the A_a are local functions on $\mathcal{C}$. Compare this with (6.2) and note that only the "time" component is left (*i.e.* the component proportional to $\theta^a \wedge (\frac{\partial}{\partial t} \lrcorner \eta)$).

2.2.1. *The Chetaev principle.* Using the noncovariant vertical endomorphism $S_{\text{n.c.}}$, we may construct a special bundle of reaction forces of the form (9.6). This is the noncovariant Chetaev principle.

Assume that $\mathcal{C}$ is locally defined as the zero set of k functionally independent functions φ^α. Then the associated bundle of reaction forces is the bundle F locally spanned by the following $(n+1)$-forms: $\Phi^\alpha := S^*_{\text{n.c.}}(\mathrm{d}\varphi^\alpha)$, or in coordinates:

$$\Phi^\alpha = \frac{\partial \varphi^\alpha}{\partial y^a_0}(\mathrm{d}y^a - y^a_\mu \mathrm{d}x^\mu) \wedge \mathrm{d}^n x_0. \tag{9.7}$$

The $(n+1)$-form Φ^α is therefore of the form outlined in (9.6), with $A^\alpha_a = \frac{\partial \varphi^\alpha}{\partial y^a_0}$.

Note the similarity with the construction in (6.3); the proof that F is a well defined bundle of $(n+1)$-forms defined along $\mathcal{C}$ proceeds along the same lines as in chapter 6.

2.2.2. *Reaction forces as one-forms.* Using the isomorphism of $J^1\pi$ with $\mathbb{R}\times[J^1(M,S)\times_S TS]$ (lemma 4.2 in chapter 1), we may construct a bundle of one-forms $\bar{F}$ along $\mathcal{C}$ which is closely related to the bundle F defined in the previous section.

The bundle $\bar{F}$ is a k-dimensional codistribution on $J^1\pi$, along $\mathcal{C}$, whose elements are maps $\alpha : \mathcal{C} \to T^*S$ such that $\alpha(\gamma) \in T^*_s S$, where $s = (\mathrm{pr}_2 \circ \pi_{1,0})(\gamma)$. If we denote by $\pi_{TS} : J^1\pi \to TS$ the composition $\pi_{TS} := \mathrm{pr}_3 \circ \Psi$, where Ψ is the isomorphism defined in lemma 4.2 in chapter 1, then the elements of $\bar{F}$ can equivalently be viewed as one-forms along the projection π_{TS}. By pull-back, these one-forms then induce proper one-forms defined along $\mathcal{C}$. In local coordinates, an element α of F can be represented as $\alpha = A_a(x^\mu, y^a, y^a_\mu)\mathrm{d}y^a$, where the A_a are local functions on $\mathcal{C}$.

Now, we establish the Chetaev principle in this context by specifying a bundle of reaction forces starting from the constraint manifold $\mathcal{C}$. The vertical endomorphism J on TS (see (1.1)) extends trivially to a $(1,1)$-tensor $\hat{J}$ on $\mathbb{R}\times[J^1(M,S)\times_S TS]$, defined as

$$\hat{J}(\alpha,\beta,\gamma) := (0,0,J(\gamma)), \tag{9.8}$$

where $\alpha \in T^*_t\mathbb{R}$, $\beta \in T^*_u J^1(M,S)$, and $\gamma \in T^*_v(TS)$ (and where $(t,u,v) \in \mathbb{R}\times[J^1(M,S)\times_S TS]$). We denote the adjoint of this map as $\hat{J}^*$.

Assume again that $\mathcal{C}$ is locally given as the zero set of k constraint functions φ^α. By means of the isomorphism Ψ of lemma 4.2 in chapter 1, these functions induce k functions on $\mathbb{R}\times[J^1(M,S)\times_S TS]$, which we also denote by φ^α.

Let us now define a codistribution $\bar{F}$ on $J^1\pi$, defined along $\mathcal{C}$, locally generated by the one-forms ψ^α given by

$$\psi^\alpha := \Psi^*[\hat{J}^*(\mathrm{d}\varphi^\alpha)].$$

In local coordinates, these forms are given by

$$\psi^\alpha = \frac{\partial \varphi^\alpha}{\partial y^a_0}\mathrm{d}y^a \tag{9.9}$$

and should be compared to the $(n+1)$-forms Φ^α defined in (9.7). The precise link is given in the following proposition:

Proposition 2.3. *Let W be a π_1-vertical vector field on $J^1\pi$. Then $i_W\Phi = 0$ for all $\Phi \in F$ if and only if $i_W\psi = 0$ for all $\psi \in \bar{F}$.*

Proof: If W is given by $W = W^a\frac{\partial}{\partial y^a} + W^a_\mu \frac{\partial}{\partial y^a_\mu}$, then

$$i_W\Phi^\alpha = W^a A^\alpha_a \mathrm{d}^n x_0 \quad \text{while} \quad i_W\psi^\alpha = W^a A^\alpha_a,$$

where Φ^α and ψ^α are the local generators of F and $\bar{F}$, respectively. $\diamond$

Recall that the Euler-Lagrange equations are derived by considering vertical variations only. As we will see later on, admissible variations for the nonholonomic problem are prolongations j^1V of vertical vector fields, such that $\langle j^1V, \alpha\rangle = 0$ for all reaction forces α. Proposition 2.3 then tells us that in deriving the nonholonomic Euler-Lagrange equations, these reaction forces can be modelled either as $(n+1)$-forms (*i.e.* as elements of F) or as one-forms (elements of $\bar{F}$). However, whenever nonvertical variations are considered (for example when applying the nonholonomic momentum lemma to a horizontal symmetry) only the use of F will yield correct results.

2.3. The field equations. Having introduced the constraint manifold $\mathcal{C}$ and the bundle F of constraint forces, we are now ready to derive the field equations. Let $L : J^1\pi \to \mathbb{R}$ be a first-order Lagrangian. From section 2.1, we know that a section ϕ defines an extremum of the action (1.14) associated to L if and only if the Euler-Lagrange equations are satisfied:

$$(j^1\phi)^* i_W \Omega_L = 0 \quad \text{for all}\, W \in \mathfrak{X}(J^1\pi).$$

In the case of nonholonomic constraints, we consider only infinitesimal variations compatible with the constraint, as in the following definition.

Definition 2.4. *A variation V of a field ϕ (taking values in $\mathcal{C}$,* i.e. *such that $j^1\phi \subset \mathcal{C}$) defined over an open subset U with compact closure is* admissible *if*

$$(j^1\phi)^*(j^1V \lrcorner\, \Phi) = 0 \quad \textit{for all}\, \Phi \in F. \tag{9.10}$$

Variations are π-vertical vector fields and therefore (9.10) is equivalent to the fact that $j^1V \lrcorner\, \Phi = 0$ along $\operatorname{Im} j^1\phi$, which is precisely the condition (6.6) in definition 1.3 in chapter 6.

By varying the action S (defined as in (1.14)) with respect to a variation V, we obtain after integrating by parts

$$0 = \frac{\mathrm{d}}{\mathrm{d}\epsilon} S(j^1(\Phi_\epsilon \circ \phi))\Big|_{\epsilon=0} = \int_U \left(\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\frac{\partial L}{\partial y^a_\mu}\right) V^a \mathrm{d}^{n+1}x, \tag{9.11}$$

where Φ_ϵ is a finite variation associated to V. If the variations V were arbitrary, then (9.11) would immediately yield the Euler-Lagrange equations. Because of definition 2.4, however, there are some restrictions to be imposed on the set of variations. The resulting nonholonomic Euler-Lagrange equations are derived below.

Definition 2.5. *A local section ϕ of π, defined on an open subset $U \subset X$ with compact closure, is a* solution *of the nonholonomic problem determined by L, $\mathcal{C}$, and F if $j^1\phi(U) \subset \mathcal{C}$ and (9.11) holds for all admissible variations V of ϕ.*

It follows from (9.11) that a local section ϕ is a solution of the nonholonomic problem if it satisfies the *nonholonomic Euler-Lagrange equations*:

$$\left[\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\frac{\partial L}{\partial y^a_\mu}\right](j^2\phi) = \lambda_\alpha A^\alpha_a(j^1\phi) \quad \text{and} \quad \varphi^\alpha(j^1\phi) = 0. \tag{9.12}$$

Here, λ_α are unknown Lagrange multipliers, to be determined from the constraints. An intrinsic form of these equations is derived below in theorem 2.8.

Lemma 2.6 (lemma 3.2 in [**48**]). *Let W be a vector field on $J^1\pi$. If ϕ is a section of π and if either W is tangent to the image of $j^1\phi$ or if W is $\pi_{1,0}$-vertical, then $(j^1\phi)^*(i_W\Omega_L) = 0$.*

Now, let ϕ be a section such that the image of $j^1\phi$ is a subset of $\mathcal{C}$ and consider a vector field W which is tangent to the image of $j^1\phi$, *i.e.* there exists a vector field w on X such that $Tj^1\phi(w(x)) = W(j^1_x\phi)$ for all $x \in X$. One can follow a similar reasoning as in the proof of lemma 3.2 in [**48**] to show that

$$(j^1\phi)^*(W \lrcorner\, \Phi) = w \lrcorner\,((j^1\phi)^*\Phi)$$

for any $\Phi \in F$. Since Φ is 1-contact, the right-hand side of this expression vanishes. On the other hand, if W is $\pi_{1,0}$-vertical, it follows automatically that $(j^1\phi)^*(W \lrcorner\, \Phi) = 0$. We have therefore proved the following lemma:

Lemma 2.7. *Let ϕ be a section of π such that $j^1_x\phi \in \mathcal{C}$ for all $x \in U \in X$. If either W is tangent to the image of $j^1\phi$ or W is $\pi_{1,0}$-vertical, then $(j^1\phi)^*(W \lrcorner\, \Phi) = 0$ for all $\Phi \in F$.*

Theorem 2.8. *Let ϕ be a section of π. If $\operatorname{Im} j^1\phi \subset \mathcal{C}$, then the following assertions are equivalent:*

(a) *ϕ is a stationary point of the action (1.14) under admissible variations;*
(b) *ϕ satisfies the Euler-Lagrange equations (9.12);*
(c) *for all vector fields W on $J^1\pi$ such that $(j^1\phi)^*(W \lrcorner\, \Phi) = 0$ for all $\Phi \in F$,*

$$(j^1\phi)^*(W \lrcorner\, \Omega_L) = 0. \tag{9.13}$$

Proof: Let us first prove the equivalence of (a) and (c). For arbitrary, not necessarily admissible variations, the following result holds (this is equation 3C.5 in [**48**]):

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} S(\phi_\epsilon)\Big|_{\epsilon=0} = -\int_U (j^1\phi)^*(j^1 V \lrcorner\, \Omega_L).$$

For admissible variations, we have therefore

$$\int_U (j^1\phi)^*(j^1 V \lrcorner\, \Omega_L) = 0.$$

Now, we may multiply V by an arbitrary function on X and this result will still hold true. The fundamental lemma of the calculus of variations therefore shows that

$$(j^1\phi)^*(j^1 V \lrcorner\, \Omega_L) = 0, \tag{9.14}$$

for all admissible variations V defined over U. By using a partition of unity as in [**48**], it can then be shown that (9.14) holds for all π-vertical vector fields V such that $(j^1 V) \lrcorner\, \Phi = 0$ for all $\Phi \in F$. This expression is equivalent to (9.13): to see this, take an arbitrary vector field W on $J^1\pi$ such that $(j^1\phi)^*(W \lrcorner\, \Phi) = 0$ for all $\Phi \in F$. The vector field W can be decomposed as the following sum (to be considered along the image of $j^1\phi$):

$$W = w_\parallel + j^1 V + v_{\pi_{1,0}},$$

where $w_\parallel$ is tangent to the image of $j^1\phi$, $j^1 V$ is the prolongation of a π-vertical vector field V, and $v_{\pi_{1,0}}$ is a $\pi_{1,0}$-vertical vector field. Using lemma 2.7, we have that

$$(j^1\phi)^*(j^1 V \lrcorner\, \Phi) = (j^1\phi)^*(W \lrcorner\, \Phi) = 0,$$

and from lemma 2.6, we get $(j^1\phi)^*(W \lrcorner\, \Omega_L) = (j^1\phi)^*(j^1 V \lrcorner\, \Omega_L)$. The right-hand side of this equation vanishes since $j^1 V$ is admissible, and therefore we conclude that W satisfies (9.13).

The equivalence of (b) and (c) is just a matter of writing out the definitions. In coordinates, the left-hand side of (9.13) reads (for a prolongation of a vertical vector field V)

$$(j^1\phi)^*(j^1 V \lrcorner\, \Omega_L) = V^a \left(\frac{\partial L}{\partial y^a}(j^1\phi) - \frac{\partial}{\partial x^\mu} \frac{\partial L}{\partial y^a_\mu}(j^1\phi) \right) \mathrm{d}^{n+1}x,$$

and this holds for all admissible variations V. Therefore, if ϕ satisfies (9.13), then there exist k functions λ_α such that

$$\left[\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu} \frac{\partial L}{\partial y^a_\mu} \right] (j^2\phi) = \lambda_\alpha A^\alpha_a (j^1\phi).$$

The converse is similar. ◇

We see from the proof of this theorem that only vertical vector fields yield nontrivial results for (9.13). For such vector fields, one can define admissibility in terms of the codistribution $\bar{F}$ defined in section 2.2.2.

Remark 2.9. The nonholonomic field equations can also be cast into De Donder-Weyl form as in chapter 6. Let F be the bundle of reaction forces, generated by the $(n+1)$-forms Φ^α in (9.7). It is easily checked that there exist vector fields Z^α, locally defined along $\mathcal{C}$, such that $Z^\alpha \lrcorner \Omega_L = -\Phi^\alpha$; this is the analogue of proposition 1.2 in chapter 6. In local coordinates, these vector fields are given by

$$Z^\alpha = \frac{\partial \varphi^\alpha}{\partial y^b_0} (\mathcal{H}^{-1})^{\mu 0}_{ab} \frac{\partial}{\partial y^a_\mu},$$

where $\mathcal{H}$ is the Hessian of L. Let D be the distribution along $\mathcal{C}$ spanned by the vector fields Z^α: if $T\mathcal{C} \cap D = \{0\}$, then we have again the decomposition $TJ^1\pi = T\mathcal{C} \oplus D$ (along $\mathcal{C}$), and the solutions of the nonholonomic De Donder-Weyl equation can be obtained through projecting the "free" solutions onto $T\mathcal{C}$.

In the case where L is a Lagrangian of the form (1.31) used in elastodynamics, then the vector fields Z^α take a particularly simple form: one can easily check that

$$Z^\alpha = \sqrt{\det{[G]}} \frac{\partial \varphi^\alpha}{\partial y^a_0} g^{ab} \frac{\partial}{\partial y^b_0},$$

where g_{ab} is the metric on the standard fibre S. These vector fields look very similar to the ones constructed in nonholonomic mechanics. The compatibility condition $T\mathcal{C} \cap D = \{0\}$ is then equivalent to the regularity of the matrix $Z^\alpha(\varphi^\beta)$. ◇

2.4. Noether's theorem. The nonholonomic momentum lemma (theorem 3.1 in chapter 7) can also be applied to the kind of nonholonomic constraints considered here, if we replace the covariant bundle of reaction forces used in chapter 7 by the bundle F described in section 2.2 of this chapter.

Recall that the notation $\bar{\xi}$ is used to designate a section of the bundle $\mathfrak{g}^{\mathcal{E}}$, which is defined as in section 3 in chapter 7: here we have that (compare with (7.3))

$$\mathcal{E}(\gamma) = \{v \in T_\gamma J^1\pi : i_v(S^*_{\text{n.c.}} \mathrm{d}\varphi_\alpha) = 0 \text{ for each } \alpha = 1, \ldots, k\} \quad \text{where } \gamma \in \mathcal{C}.$$

For a given $\gamma \in \mathcal{C}$, we consider the set $\mathfrak{g}^\gamma$ consisting of all $\xi \in \mathfrak{g}$ such that $\xi^{(1)}(\gamma) \in \mathcal{E}(\gamma)$, and then we define the bundle $\mathfrak{g}^{\mathcal{E}}$ over $\mathcal{C}$ as the disjoint union of all $\mathfrak{g}^\gamma$. Recall also that every section $\bar{\xi}$ induces a vector field $\tilde{\xi}$ on $J^1\pi$ defined along $\mathcal{C}$ by putting $\tilde{\xi}(\gamma) = (\bar{\xi}(\gamma))_{J^1\pi}(\gamma)$.

An important corollary of the nonholonomic momentum lemma regards the case where the section $\bar{\xi}$ is "constant", *i.e.* there exists a $\xi \in \mathfrak{g}$ such that $\bar{\xi}(\gamma) = \xi$ for all $\gamma \in \mathcal{C}$. In that case, $\tilde{\xi}$ is just $\xi_{J^1\pi}$ and hence the right-hand side of (7.4) vanishes as L is supposed to be $\mathcal{G}$-invariant. This can easily be proved directly, as we now show. In fact, this theorem is some cases stronger than the nonholonomic momentum equation, which only holds for vertical symmetries, while the symmetries considered here can have a horizontal component as well. An important example is conservation of energy for the

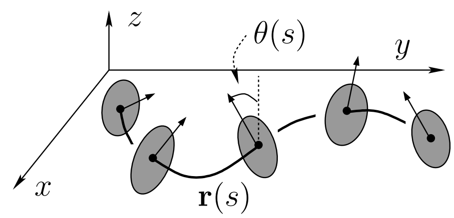


Figure 9.2. Geometry of the constrained rod

nonholonomic Cosserat rod (proposition 3.3 below). In this case, the symmetry vector field is $\frac{\partial}{\partial t}$, which is not vertical.

Theorem 2.10. *Let $\mathcal{L}$ be a $\mathcal{G}$-invariant Lagrangian density. Assume that $\xi \in \mathfrak{g}$ is such that $\xi_{J^1\pi} \lrcorner\, \alpha = 0$ along $\mathcal{C}$ for all $\alpha \in F$. Then the following conservation law holds:*

$$\mathrm{d}[(j^1\phi)^* J^L_\xi] = 0, \tag{9.15}$$

for all sections ϕ of π that are solutions of the nonholonomic field equations.

Proof: Let ξ be an element of $\mathfrak{g}$ such that, along $\mathcal{C}$, $\xi_{J^1\pi} \lrcorner\, \Phi = 0$ for all $\Phi \in F$. Because of the nonholonomic field equations, $(j^1\phi)^* \xi_{J^1\pi} \Omega_L = 0$. Now, according to (1.18), we have

$$0 = \mathscr{L}_{\xi_{J^1\pi}} \Theta_L = \mathrm{d} i_{\xi_{J^1\pi}} \Theta_L + i_{\xi_{J^1\pi}} \mathrm{d}\Theta_L.$$

Upon pulling back this identity along a solution $j^1\phi$ of the nonholonomic field equations, we obtain (9.15). ◇

3. A Cosserat-type model

Recall that in the introduction we considered a Cosserat rod whose laminae are rigid discs. In this case, one can choose the director frame $\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ to be orthogonal with, in addition, $\mathbf{d}_2$ and $\mathbf{d}_3$ of unit length (attached to the laminae) and $\mathbf{d}_1$ aligned with $\mathbf{r}'(s)$, the tangent vector to the centerline. If, in addition, the centerline is assumed to be inextensible, so that we may choose the parameter s to be arc length, $\mathbf{d}_1$ is also of unit length and the director frame is orthonormal. In this case, the specification of, say, $\mathbf{d}_2$ is enough to determine a director frame: putting $\mathbf{d}_1 \equiv \mathbf{r}'$, we then know that $\mathbf{d}_3 = \mathbf{d}_1 \times \mathbf{d}_2$.

Here, we will consider the case of a Cosserat rod with an inextensible centerline and rigid laminae. In addition, we will assume that the centerline is planar, which will allow us to eliminate the director frame almost completely. The result is a Lagrangian field theory of second order, to which the results of section 2 can be applied.

3.1. The planar Cosserat rod. Consider an inextensible Cosserat rod of length ℓ equipped with three directors. If we denote the centerline at time t as $s \mapsto \mathbf{r}(t, s)$, inextensibility allows us to assume that the parameter s is the arc length. Secondly, we can take the director frame $\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$ to be orthonormal, such that $\mathbf{d}_1$ is the unit tangent vector $\mathbf{r}'$. We will not take the effect of gravity into account.

In addition, we now assume that the centerline is a planar curve moving in the horizontal plane, *i.e.* $\mathbf{r}(t, s)$ can be written as $(x(t, s), y(t, s), 0)$. We introduce the *slope* $\varphi(t, s)$ of the centerline as $(\cos\varphi, \sin\varphi) = (x'(t, s), y'(t, s))$. Furthermore, we define the angle $\theta(t, s)$, referred to as the *torsion* of the rod, as the angle subtended between $\mathbf{e}_z$ and $\mathbf{d}_3$. The director frame is completely determined once we know the slope $\varphi(s, t)$ and the torsion $\theta(s, t)$.

The specific constraints imposed on our rod model therefore allow us to eliminate the director frame in favour of the slope φ and the torsion θ. Furthermore, as we shall see, the slope φ is related to the curvature of the centerline. Note that, in formulating the dynamics, we still have to impose the inextensibility condition $(x')^2 + (y')^2 = 1$. Note that this constraint does not involve derivatives with respect to time, and as we saw in chapter 8, it should therefore be modelled as a vakonomic constraint.

Remark 3.1. Note that θ has nothing to do with the usual geometric concept of torsion of a curve in $\mathbb{R}^3$, and neither is θ related to the concept of shear in (for example) the theory of the Timoshenko beam. ◇

3.2. The dynamics. As the director frame is orthonormal, there exists a vector $\mathbf{u}$, defined by $\mathbf{d}_i' = \mathbf{u} \times \mathbf{d}_i$, called the *strain* or *Darboux vector.* With the conventions from the previous section, $\mathbf{u}$ takes the following form:

$$\mathbf{u} = \theta' \mathbf{d}_1 + \varphi' \mathbf{e}_z.$$

($\mathbf{u}$ can be thought of as an "angular momentum" vector, but with time-derivatives replaced by derivatives with respect to arc length.)

The dynamics of our rod model can be derived from a variational principle. The kinetic energy is given by

$$T = \frac{1}{2} \int_0^\ell \left(\rho(s)(\dot{x}^2 + \dot{y}^2) + \alpha \dot{\theta}^2 \right) \mathrm{d}s,$$

where α is an appropriately chosen constant. Here, the mass density is denoted by ρ, and will be assumed constant from now on.

For a hyperelastic rod, the potential energy is of the form

$$V = \int_0^\ell W(u_1, u_2, u_3) \mathrm{d}s,$$

where $W(u_1, u_2, u_3)$ is called the *stored energy density*, and the u_i are the components of $\mathbf{u}$ relative to the director frame: $u_i = \mathbf{u} \cdot \mathbf{d}_i$. In the simplest case, of *linear elasticity*, W is a quadratic function of the strains:

$$W(u_1, u_2, u_3) = \frac{1}{2}\left(K_1 u_1^2 + K_2 u_2^2 + K_3 u_3^2\right). \tag{9.16}$$

We will not dwell on the physical interpretation of the constants K_i any further (in this case, they are related to the moments of inertia of the laminae). If the rod is transversely isotropic, *i.e.* if the laminae are invariant under rotations around $\mathbf{d}_1$, we may take $K_2 = K_3$. The potential energy then becomes

$$V = \frac{1}{2}\int_0^\ell \left(\beta(\theta')^2 + K\kappa^2\right) \mathrm{d}s,$$

where κ is the curvature of the centerline, *i.e.* $\kappa^2 = (\varphi')^2 = (x'')^2 + (y'')^2$, and where we have put $\beta := K_1$ and $K := K_2$. Models with a similar potential energy abound throughout the literature and are generally referred to as the Euler elastica. For more information, see [**67**] and the references therein.

3.3. The second-order model. Having eliminated the derivative of the slope φ from the stored energy density, we end up with a model in which the fields are the coordinates of the centerline $(x(t,s), y(t,s))$ and the torsion angle $\theta(t,s)$. This model fits into the framework developed in section 2.3; the base space X is $\mathbb{R} \times [0, \ell]$, with coordinates (t, s) and the total space Y is $X \times \mathbb{R}^2 \times \mathbb{S}^1$, with fibre coordinates (x, y, θ).

The total Lagrangian now consists of kinetic and potential energy, as well as an additional term enforcing the constraint of inextensibility, and can be written as

$$L = \frac{\rho}{2}(\dot{x}^2 + \dot{y}^2) + \frac{\alpha}{2}\dot{\theta}^2 - \frac{1}{2}\left(\beta(\theta')^2 + K\kappa^2\right) - \frac{1}{2}p\left((x')^2 + (y')^2 - 1\right), \tag{9.17}$$

where p is a Lagrange multiplier associated to the constraint of inextensibility. The field equations associated to this Lagrangian take the following form:

$$\left\{\begin{array}{rcl} \rho\ddot{x} + Kx'''' &=& \frac{\partial}{\partial s}(px') \\ \rho\ddot{y} + Ky'''' &=& \frac{\partial}{\partial s}(py') \\ \alpha\ddot{\theta} - \beta\theta'' &=& 0, \end{array}\right. \tag{9.18}$$

to be supplemented with the inextensibility constraint

$$(x')^2 + (y')^2 = 1, \tag{9.19}$$

which allows to determine the multiplier p. Note in passing that the dynamics of the centerline and the torsion angle θ are completely uncoupled. This will change once we add nonholonomic constraints.

3.4. Field equations and symmetries. We recall the expression (1.19) for the second-order Cartan form. If a Lie group $\mathcal{G}$ is acting on Y by bundle automorphisms, and on $J^3\pi$ by prolonged bundle automorphisms, there is a Lagrangian momentum map $J^L_\xi = \xi_{J^3\pi} \lrcorner \Omega_L$, as described in section 2. We now turn to a brief overview of the symmetries associated to the rod model introduced in the previous section. For an overview of symmetries in the general theory of Cosserat rods, see [**40**].

3.4.1. *Translations in time.* The Lie group $\mathbb{R}$ acts on X by translations in time: $\Phi_\epsilon : (s,t) \mapsto (s, t+\epsilon)$. The Lagrangian is equivariant and the pullback to X (by a solution $j^3\phi$ of the field equations) of the momentum map associated to the infinitesimal generator $\frac{\partial}{\partial t}$ is given by

$$
\begin{aligned}
(j^3\phi)^* J^L_{\frac{\partial}{\partial t}} = &\Big[(px' - Kx''')\dot{x} + (py' - Ky''')\dot{y} + \beta\theta'\dot{\theta} + K(x''\dot{x}' + y''\dot{y}')\Big]\,\mathrm{d}t \\
&+ \Big[\underbrace{\frac{\rho}{2}(\dot{x}^2 + \dot{y}^2) + \frac{\alpha}{2}\dot{\theta}^2 + \frac{K}{2}((x'')^2 + (y'')^2) + \frac{\beta}{2}(\theta')^2 + \frac{p}{2}((x')^2 + (y')^2 - 1)}_{\mathcal{E}}\Big]\mathrm{d}s,
\end{aligned}
\tag{9.20}
$$

where we have introduced the *energy density* $\mathcal{E}$. By taking the exterior derivative of (9.20) and integrating the conservation law $\mathrm{d}[(j^3\phi)^* J^L_{\partial/\partial t}] = 0$ over $[0,\ell] \times [t_0, t_1] \subset \mathbb{R}^2$, we obtain

$$
E(t_1) - E(t_0) = \int_{t_0}^{t_1} \Big[(px' - Kx''')\dot{x} + (py' - Ky''')\dot{y} + \beta\theta'\dot{\theta} + K(x''\dot{x}' + y''\dot{y}')\Big]_0^\ell \mathrm{d}t,
\tag{9.21}
$$

where $E(t) = \int_0^\ell \mathcal{E}\mathrm{d}s$ is the total energy, which is conserved if suitable boundary conditions are imposed. This is the case, for instance, for periodic boundary conditions or when both ends of the rod can move freely, *i.e.* when

$$
px' - Kx''' = py' - Ky''' = 0 \quad \text{and} \quad x'' = y'' = \theta' = 0 \quad \text{at } s = 0, \ell.
$$

3.4.2. *Spatial translations.* Consider the Abelian group $\mathbb{R}^2$ acting on the total space Y by translation, *i.e.* for each $(a,b) \in \mathbb{R}^2$ we consider the map $\Phi_{(a,b)} : (s,t;x,y,\theta) \mapsto (s,t;x+a, y+b, \theta)$. The Lagrangian density is invariant under this action and the associated momentum map is

$$
J^L_{(v_1,v_2)} = -\rho(v_1\dot{x} + v_2\dot{y})\mathrm{d}s + (v_1 px' - v_1 Kx''' + v_2 py' - v_2 Ky''')\mathrm{d}t \quad \text{for all } (v_1, v_2) \in \mathbb{R}^2.
$$

Again, under suitable boundary conditions, $J^L_{(v_1,v_2)}$ gives rise to a conserved quantity, which is the total linear momentum of the rod.

Similarly, $\mathbb{S}^1$ acts on Y by translations in θ, with infinitesimal generators of the form $\alpha\frac{\partial}{\partial\theta}$, and the corresponding momentum map is

$$
J^L_\alpha = -\beta\theta'\mathrm{d}t - \alpha\dot{\theta}\mathrm{d}s.
$$

The ensuing conservation law is given by $\alpha\ddot{\theta} = \beta\theta''$ and, hence, is just the equation of motion for θ.

3.4.3. *Spatial rotations.* Finally, we note that the rotation group $SO(2)$ acts on Y by rotations in the (x, y)-plane. The infinitesimal generator corresponding to $1 \in \mathfrak{so}(2) \cong \mathbb{R}$ is given by $y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$; its prolongation to $J^3\pi$ is

$$\xi_{J^3\pi} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} + \dot{y}\frac{\partial}{\partial \dot{x}} - \dot{x}\frac{\partial}{\partial \dot{y}} + y'\frac{\partial}{\partial x'} - x'\frac{\partial}{\partial y'} + \cdots ,$$

where the dots represent terms involving higher-order derivatives. As Θ_L is semi-basic with respect to $\pi_{3,1}$, these terms make no contribution to the momentum map. The momentum map is given by

$$J_1^L = [-x(-py' + Ky''') + y(-px' + Kx''') - K(x''y' + y''x')]\,\mathrm{d}t + \rho(x\dot{y} - y\dot{x})\mathrm{d}s,$$

leading to the conservation of total angular momentum. Note that the angular momentum does not involve θ, in contrast to the corresponding expression in more general treatments of Cosserat media. This is a consequence of the fact that we defined the action of $SO(2)$ on Y to act trivially on the $\mathbb{S}^1$ factor.

3.5. A nonholonomic model. Consider again a Cosserat rod as in the previous section. The constraint that we are now about to introduce is a generalization of the familiar concept of rolling without sliding in mechanics: we assume that the rod is placed on a horizontal plane, which we take to be perfectly rough, so that each of the laminae rolls without sliding.

However, as the Cosserat rod is also supposed to be incompressible, one must take care that the additional constraints do not become too restrictive.[1] Indeed, a simple argument shows that the model of an incompressible rod which rolls without sliding, and which cannot move transversally, can only move like a rigid body.

There are two immediate solutions: either one relaxes the incompressibility constraint, or one allows the rod to move laterally as well. Either solution introduces a lot of mathematical tedium which greatly obscures the physical background of the system. Here, we will therefore consider a simplified model containing aspects of both models.

In particular, we will *assume* that the motion of the nonholonomic rod is such that the incompressibility constraint is satisfied approximately throughout the motion; this is equivalent to the following assumption:

$$\sqrt{(x')^2 + (y')^2} \cong 1. \tag{9.22}$$

By neglecting the incompressibility constraint in the Lagrangian, a simplified model is then obtained. Of course, this new model is a mathematical simplification of the true physics. However, numerical simulations show that $\sqrt{(x')^2 + (y')^2}$ is bounded throughout the motion, and it's therefore reasonable that the dynamics of this model is close to the true dynamics. One could think of the mathematical model as describing a

[1]This was pointed out to me by W. Tulczyjew and D. Zenkov.

Cosserat rod whose constitutive equation is specified on mathematical grounds, rather than derived from first principles.

The constraints of rolling without sliding are given by (see [**12**, **26**]):

$$\dot{x} + R\dot{\theta}\sin\varphi = 0 \quad \text{and} \quad \dot{y} - R\dot{\theta}\cos\varphi = 0, \tag{9.23}$$

where R is the radius of the laminae. By eliminating the slope φ we then obtain

$$\dot{x} + R\dot{\theta}y' = 0 \quad \text{and} \quad \dot{y} - R\dot{\theta}x' = 0. \tag{9.24}$$

Incidentally, the passage from (9.23) to (9.24) again illustrates why derivatives with respect to time play a fundamentally different role as opposed to the other derivatives.

The Lagrangian density of the nonholonomic rod is still given by (9.17); we recall that it is of second order, as the stored energy function (9.16) is of grade two. The constraints on the other hand are of first order. By requiring that the action be stationary under variations compatible with the given constraints (a similar approach to section 2.3), we obtain the following field equations:

Definition 3.2. *A section ϕ of π is a solution of the nonholonomic problem if and only if $\mathrm{Im}\, j^1\phi \subset \mathcal{C}$, and, along $\mathcal{C}$,*

$$(j^3\phi)^*(W \lrcorner\, \Omega_L) = 0 \tag{9.25}$$

for all π-vertical vector fields V on Y such that $(j^1\phi)^*(W \lrcorner\, \Phi) = 0$ for all $\Phi \in F$.

The left-hand side of (9.25) is just the Euler-Lagrange equation (1.20) for a second-order Lagrangian. As the constraints are first order, they can be treated exactly as in section 2.3. In coordinates, the nonholonomic field equations hence are given by

$$\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\left(\frac{\partial L}{\partial y^a_\mu}\right) + \frac{\mathrm{d}^2}{\mathrm{d}x^\mu \mathrm{d}x^\nu}\left(\frac{\partial L}{\partial y^a_{\mu\nu}}\right) = \lambda_\alpha \frac{\partial \varphi^\alpha}{\partial y^a_0}.$$

By substituting the Lagrangian (9.17) and the constraints (9.24) into the Euler-Lagrange equations, we obtain the following set of nonholonomic field equations:

$$\left\{\begin{array}{rcl} \rho\ddot{x} + Kx'''' & = & \lambda \\ \rho\ddot{y} + Ky'''' & = & \mu \\ \alpha\ddot{\theta} - \beta\theta'' & = & R(\lambda y' - \mu x'), \end{array}\right. \tag{9.26}$$

where λ, μ are Lagrange multipliers associated with the nonholonomic constraints. These equations are to be supplemented by the constraint equations (9.24).

Conservation laws. In the familiar case of the rolling disc, it is well known that energy is conserved. There is a similar conservation law for the nonholonomic rod.

Proposition 3.3. *The total energy (9.21) is conserved for each solution of the nonholonomic field equations (9.26) and constraints (9.24). A fortiori, the solutions of the nonholonomic field equations satisfy the local conservation law $\mathrm{d}[(j^3\phi)^* J^L_1] = 0$, where J^L_1 is the momentum map associated to time translation introduced in (9.20).*

Proof: This follows immediately from proposition 2.10 and the fact that $\frac{\partial}{\partial t}$ (or rather its prolongation to $J^1\pi$) annihilates F along the constraint manifold. Indeed, if we introduce the forms Φ^1 and Φ^2, defined as in (9.7), and explicitly given by

$$\Phi^1 = (\mathrm{d}x - \dot{x}\mathrm{d}t) \wedge \mathrm{d}s + Ry'(\mathrm{d}\theta - \dot{\theta}\mathrm{d}t) \wedge \mathrm{d}s;$$
$$\Phi^2 = (\mathrm{d}y - \dot{y}\mathrm{d}t) \wedge \mathrm{d}s - Rx'(\mathrm{d}\theta - \dot{\theta}\mathrm{d}t) \wedge \mathrm{d}s,$$

then

$$\left(\frac{\partial}{\partial t}\right)_{J^1\pi} \lrcorner\, \Phi^1 = -(\dot{x} + R\dot{\theta}y')\mathrm{d}s,$$

which vanishes when restricted to $\mathcal{C}$. A similar argument shows that the contraction of $\frac{\partial}{\partial t}$ with Φ^2 vanishes. Hence, proposition 2.10 can be applied; the associated momentum map is just (9.20). ◇

To find the nonholonomic counterpart of the other conservation laws in section 3.4, we need to use the full nonholonomic momentum lemma. Here, we only consider the translation action of section 3.4.2, as it appears that the rotational action of section 3.4.3 is not a nonholonomic symmetry.

Here, the Lagrangian is of second order, but the derivation of the nonholonomic momentum lemma proceeds exactly as in chapter 7, up to a few minor modifications: the nonholonomic momentum map $J^{\text{n.h.}}$ is now defined on $J^3\pi$, and the nonholonomic momentum equation hence becomes

$$\mathrm{d}_{\mathbf{h}} J^{\text{n.h.}}_{\tilde{\xi}} = \pi^*_{3,2}\mathscr{L}_{j^2\tilde{\xi}}(L\eta), \tag{9.27}$$

where $\mathbf{h}$ is a solution of the nonholonomic De Donder-Weyl equation.

Consider the action of $\mathbb{R}^2 \times \mathbb{S}^1$ on Y by translations; *i.e.* for each (a, b, φ) we consider the map $\Phi_{(a,b,\varphi)} : (s, t, x, y, \theta) \mapsto (s, t, x + a, y + b, \theta + \varphi)$. Let $\xi = (v_1, v_2, v_\theta)$ be an element of the Lie algebra of $\mathbb{R}^2 \times \mathbb{S}^1$. The corresponding fundamental vector field is given by

$$\xi_Y = v_1\frac{\partial}{\partial x} + v_2\frac{\partial}{\partial y} + v_\theta\frac{\partial}{\partial \theta}.$$

In order for the nonholonomic momentum lemma to apply, $\xi_{J^1\pi}$ must annihilate the bundle of constraint forms F. For the nonholonomic Cosserat rod, this is the case when the fundamental vector field has the following form (considered as a vector field along $\pi_{1,0}$):

$$\tilde{\xi} = -Ry'\frac{\partial}{\partial x} + Rx'\frac{\partial}{\partial y} + \frac{\partial}{\partial \theta}.$$

(Any scalar multiple of the above vector field is also allowed.) This vector field corresponds with the section $\bar{\xi} = (-Ry', Rx', 1)$ of $\mathfrak{g}^{\mathcal{E}}$. By pulling back (9.27) by an integral section $j^3\phi$ of $\mathbf{h}$, we obtain

$$\mathrm{d}((j^3\phi)^* J^{\text{n.h.}}_{\tilde{\xi}}) = j^2\phi^*\mathscr{L}_{\tilde{\xi}}(L\eta), \tag{9.28}$$

where we recall that the left-hand side of (9.27) (pulled back along an integral section of $\mathbf{h}$) can be rewritten as follows (see also remark 2.2 in chapter 7):

$$\mathrm{d}_{\mathbf{h}} J^{\text{n.h.}}_{\tilde{\xi}} = \mathrm{d}[(j^3\phi)^* J^{\text{n.h.}}_{\tilde{\xi}}].$$

The right-hand side of (9.28) is

$$j^2\phi^*\mathscr{L}_{\tilde{\xi}}(L\eta) = \big[- R\rho\dot{y}'\dot{x} + R\rho\dot{x}'\dot{y} - KRx'''y'' + KRy'''x''\big]\eta.$$

The nonholonomic momentum map $J^{\text{n.h.}}$, on the other hand, is given by

$$J^{\text{n.h.}}_{\tilde{\xi}} = -\big[\rho(Rx'\dot{y} - Ry'\dot{x}) + \alpha\dot{\theta}\big]\mathrm{d}s - \big[KR(y'x''' - x'y''') + \beta\theta'\big]\mathrm{d}t,$$

and the nonholonomic momentum equation (9.27) hence becomes

$$Ry'(\rho\ddot{x} + Kx'''') - Rx'(\rho\ddot{y} + Ky'''') = \alpha\ddot{\theta} - \beta\theta''. \tag{9.29}$$

This conservation law can also be derived from the nonholonomic field equations (9.26) by subtracting the second equation multiplied by x' from the first equation multiplied by y', and using the third equation to eliminate the Lagrange multipliers λ and μ. Unfortunately, the knowledge of this nonholonomic conservation law does not help us in solving the field equations (in contrast with the situation for the vertical rolling disc as in [**12**]).

4. Discrete nonholonomic field theories

In this section we present an extension to the case of field theories of the discrete d'Alembert principle described in [**28**]. We also derive an elementary numerical integration scheme aimed at integrating the field equations (9.26).

As in the previous sections, we will consider the trivial bundle π with base space $X = \mathbb{R}\times M$ (where $M = [0,\ell]$), and total space Y the product $X \times S$, where $S = \mathbb{R}^2\times\mathbb{S}^1$. Our discretization scheme is the most straightforward possible, where the base space X is discretized by means of the uniform mesh $\mathbb{Z}\times\mathbb{Z}$.

4.1. Discrete Lagrangian field theories. We begin by giving an overview of discrete Lagrangian field theories, inspired by [**61**,**80**]. In order to discretize the second-order jet bundle, we need to approximate the derivatives of the field (of first and second order). This we do by means of central differences with spatial step k and time step h:

$$\dot{\eta} \approx \frac{\eta_{n+1,i} - \eta_{n,i}}{h}, \quad \eta' \approx \frac{\eta_{n,i+1} - \eta_{n,i-1}}{2k}, \quad \text{and} \quad \eta'' \approx \frac{\eta_{n,i+1} - 2\eta_{n,i} + \eta_{n,i-1}}{k^2}, \tag{9.30}$$

where η stands for either x or y. Other derivatives will not be needed. For θ, we use

$$\dot{\theta} \approx \frac{\theta_{n+1,i} - \theta_{n,i}}{h} \quad \text{and} \quad \theta' \approx \frac{\theta_{n,i+1} - \theta_{n,i}}{k}. \tag{9.31}$$

Let $\mathcal{M}$ be the uniform mesh in $X = \mathbb{R}^2$ whose elements are points with integer coordinates; *i.e.* $\mathcal{M} = \mathbb{Z}\times\mathbb{Z}$. The elements of $\mathcal{M}$ are denoted as (n,i), where the first component refers to time, and the second to the spatial coordinate. We define a 9-*cell* centered at $(n,i) \in \mathcal{M}$, denoted by $[x]_{(n,i)}$, to be a nine-tuple of the form

$$\begin{aligned} [x]_{(n,i)} \;:=\; & \big((n-1,i-1),(n-1,i),(n-1,i);(n,i-1), \\ & (n,i),(n,i+1);(n+1,i-1),(n+1,i),(n+1,i+1)\big) \end{aligned} \tag{9.32}$$

It is clear from the finite difference approximations that a generic second-order jet $j^2_x\phi$ can be approximated by specifying the values of ϕ at the nine points of a cell.

However, in the case of the nonholonomic rod, the Lagrangian depends only on the derivatives whose finite difference approximations were given in (9.30) and (9.31). Therefore, we can simplify our exposition by defining a 6-*cell* at (n,i) to be the six-tuple

$$[x]_{(n,i)} := \big((n,i-1),(n,i),(n,i+1);(n+1,i-1),(n+1,i),(n+1,i+1)\big). \tag{9.33}$$

We will refer to 6-cells simply as *cells*. Let us denote the set of all cells by $\mathbb{X}^6 := \{[x]_{(n,i)} : (n,i) \in \mathcal{M}\}$. We now define the *discrete 2nd order jet bundle* to be $J^2_d\pi := \mathbb{X}^6 \times S^{\times 6}$ (see [**61**, **80**, **108**]). A *discrete section* of π (also referred to as a *discrete field*) is a map $\phi : \mathcal{M} \to S$. Its *second jet extension* is the map $j^2\phi : \mathbb{X}^6 \to J^2_d\pi$ defined as

$$j^2\phi([x]_{(n,i)}) := ([x]_{(n,i)};\phi(x_1),\dots\phi(x_6)),$$

where $x_1,\dots,x_6$ are the vertices that make up $[x]_{(n,i)}$ (ordered as in (9.33)). Given a vector field W on Y, we define its second jet extension to be the vector field j^2W on J^2_d given by

$$j^2W([x];s_1,\dots,s_6) = (W(x_1,s_1),W(x_2,s_2),\dots,W(x_6,s_6)).$$

Let us now assume that a *discrete Lagrangian* $L_d : J^2_d\pi \to \mathbb{R}$ is given. The *action sum* S_d is then defined as

$$S_d(\phi) = \sum_{[x]\in U_F} L_d(j^2\phi([x])), \tag{9.34}$$

where U_F is a finite subset of $\mathbb{X}^6$. Given a vertical vector field V on Y and a discrete field ϕ, we obtain a one-parameter family ϕ_ϵ by composing ϕ with the flow Φ_ϵ of V:

$$\phi_\epsilon([x]) = ([x];\Phi_\epsilon(\phi([x])_1),\dots,\Phi_\epsilon(\phi([x])_6))\,. \tag{9.35}$$

The *variational principle* now consists of seeking discrete fields ϕ that extremize the discrete action sum. The fact that ϕ is an extremum of S under variations of the form (9.35) is expressed by

$$\begin{aligned} \sum_{(n,i)\in\mathcal{M}} \Big\langle & X(\phi_{(n,i)}), D_1L(j^2\phi([x]_{(n,i+1)})) + D_2L(j^2\phi([x]_{(n,i)})) + D_3L(j^2\phi([x]_{(n,i-1)})) \\ & + D_4L(j^2\phi([x]_{(n-1,i+1)})) + D_5L(j^2\phi([x]_{(n-1,i)})) + D_6L(j^2\phi([x]_{(n-1,i-1)}))\Big\rangle = 0. \end{aligned} \tag{9.36}$$

As the variation X is completely arbitrary, we obtain the following set of *discrete Euler-Lagrange field equations*:

$$D_1L(j^2\phi([x]_{(n,i+1)})) + D_2L(j^2\phi([x]_{(n,i)})) + D_3L(j^2\phi([x]_{(n,i-1)})) + \qquad (9.37)$$
$$D_4L(j^2\phi([x]_{(n-1,i+1)})) + D_5L(j^2\phi([x]_{(n-1,i)})) + D_6L(j^2\phi([x]_{(n-1,i-1)})) = 0.$$

for all (n,i). Here, we have denoted the values of the field ϕ at the points (n,i) as $\phi_{n,i}$.

4.2. The discrete d'Alembert principle. Our discrete d'Alembert principle is nothing more than a suitable field-theoretic extension of the discrete Lagrange-d'Alembert principle described in [**28**]. Just as in that paper, in addition to the discrete Lagrangian L_d, two additional ingredients are needed: a discrete constraint manifold $\mathcal{C}_d \subset J^1_d\pi$ and a bundle of constraint forces F_d on $J^2_d\pi$. However, as our constraints (in particular (9.24)) are not linear in the derivatives, as opposed to the case in [**28**], our analysis will be more involved.

The discrete constraint manifold $\mathcal{C}_d \hookrightarrow J^1_d\pi$ will usually be constructed from the continuous constraint manifold $\mathcal{C}$ by subjecting it to the same discretization as used for the discretization of the Lagrangian (*i.e.* (9.30) and (9.31)). To construct the discrete counterpart F_d of the bundle of discrete constraint forces, somewhat more work is needed.

Remark 4.1. For the discretization of the constraint manifold, it would appear that we need a discrete version of the first-order jet bundle as well. A similar procedure as for the discretization of the second-order jet bundle (using the same finite differences as in (9.30) shows that a discrete 1-jet depends on the values of the field at the same four points of a cell as a discrete 2-jet: the difference between $J^1_d\pi$ and $J^2_d\pi$ lies in the way in which the values of the field at these points are combined. Therefore, we can regard the discrete version of $\mathcal{C}$, to be defined below, as a subset of $J^2_d\pi$. ◇

4.2.1. *The bundle of discrete constraint forces.* In this section, we will construct F_d by following a discrete version of the procedure used in section 2.2. Just as in the continuous case, it is here that the difference between spatial and time derivatives will become fundamental. Indeed, we will discretize with respect to space first, and (initially) not with respect to time. It should be noted that the construction outlined in this paragraph is not entirely rigorous but depends strongly on coordinate expressions. Presumably, one would need a sort of *discrete Cauchy analysis* in order to solidify these arguments. For now, we will just accept that this reasoning provides us with the correct form of the constraint forces.

For the sake of convenience, we suppose that $\mathcal{C}$ is given by the vanishing of k independent functions φ^α on $J^1\pi$. By applying the *spatial* discretizations in (9.30) and (9.31) to φ^α, we obtain k functions, denoted as $\varphi^\alpha_{1/2}$, on $J^2_d \times TS$. We define the *semi-discretized constraint submanifold* $\mathcal{C}_{1/2}$ to be the zero level set of the functions $\varphi^\alpha_{1/2}$.

Consider now the forms

$$\Phi^\alpha_{1/2} := J^*(\mathrm{d}\varphi^\alpha_{1/2})$$

(where J is the vertical endomorphism (1.1) on TS); they are the semi-discrete counterparts of the forms Φ^α defined in (9.7). The forms $\Phi^\alpha_{1/2}$ are semi-basic. By discretizing the time derivatives, however, we obtain a set of basic forms on $J^2_d\pi$, which we also denote by $\Phi^\alpha_{1/2}$. An example will make this clearer.

Example 4.2. Consider, for instance, the constraint manifold $\mathcal{C} \hookrightarrow J^2\pi$ defined as the zero level set of the function $\varphi = A_{ab}\, y^a_0 y^b_1 + B_b (y^b_1)^2$, where A_{ab} and B_b are constants. By applying (9.30) and (9.31), it follows that $\mathcal{C}_d$ is given as the zero level set of the function

$$\varphi_d([y]) := A_{ab} \frac{y^a_{n+1,i} - y^a_{n,i}}{h} \frac{y^b_{n,i+1} - y^b_{n,i-1}}{2k} + B_b \left(\frac{y^b_{n,i+1} - y^b_{n,i-1}}{2k} \right)^2 \quad \text{for } [y] \in J^1_d\pi,$$

and $\mathcal{C}_{1/2}$ as the zero level set of the function

$$\varphi_{1/2}([y], v) := A_{ab}\, \dot{v}^a \, \frac{y^b_{n,i+1} - y^b_{n,i-1}}{2k} + B_b \left(\frac{y^b_{n,i+1} - y^b_{n,i-1}}{2k} \right)^2$$

for $[y] \in J^1_d\pi$ and $v \in TS$. The bundle F_d is then generated by the one-form $\Phi_{1/2} := J^*(\mathrm{d}\varphi_{1/2})$, or explicitly,

$$\Phi = A_{ab} \frac{y^b_{n,i+1} - y^b_{n,i-1}}{2k} \mathrm{d}y^a.$$

◇

4.2.2. *The discrete nonholonomic field equations.* Assuming that L_d, $\mathcal{C}_d$ and F_d are given (their construction will be treated in more detail in the next section), the derivation of the discrete nonholonomic field equations is similar to the continuum derivation: we are looking for a discrete field ϕ such that $\operatorname{Im} j^1\phi \subset \mathcal{C}_d$ and such that ϕ is an extremum of (9.34) for all variations *compatible with the constraints*, in the sense that the variation X satisfies, for all (n, i),

$$X(\phi_{(n,i)}) \lrcorner\, \Phi^\alpha_{1/2}(j^2\phi([x]_{(n,i)})) = 0.$$

From (9.36) we then obtain the *discrete nonholonomic field equations*:

$$\begin{aligned} D_1 L(j^2\phi([x]_{(n,i+1)})) + D_2 L(j^2\phi([x]_{(n,i)})) + D_3 L(j^2\phi([x]_{(n,i-1)})) +& \\ D_4 L(j^2\phi([x]_{(n-1,i+1)})) + D_5 L(j^2\phi([x]_{(n-1,i)})) + D_6 L(j^2\phi([x]_{(n-1,i-1)})) =& \\ \lambda_\alpha \Phi^\alpha_{1/2}(j^2\phi([x]_{(n,i)})),& \quad (9.38) \end{aligned}$$

where the Lagrange multipliers λ_α are to be determined from the requirement that $\operatorname{Im} j^1\phi \subset \mathcal{C}_d$.

4.3. An explicit, second-order algorithm. In this section, we briefly present some numerical insights into the nonholonomic field equations of section 3.5. Our aim is twofold: for generic boundary conditions, the nonholonomic field equations (9.26) can probably not be solved analytically and in order to gain insight into the behaviour of our model, we therefore turn to numerical methods. Secondly, in line with the fundamental tenets of geometric integration, we wish to show that the construction of practical integration schemes is strongly guided by geometric principles.

In discretizing our rod model, we effectively replace the continuous rod by N rigid rolling discs interconnected by some potential (see [**5**]). This is again an illustration of the fact that the constraints are truly nonholonomic. Our integrator is just a concatenation of the leapfrog algorithm for the spatial part, and a nonholonomic mechanical integrator for the integration in time.

As a first attempt at integrating (9.26), we present an explicit, second-order algorithm. In the Lagrangian, the derivatives are approximated by

$$\dot{x} \approx \frac{x_{n+1,i} - x_{n,i}}{h} \quad \text{and} \quad x'' \approx \frac{x_{n,i-1} - 2x_{n,i} + x_{n+1,i}}{k^2},$$

where h is the time step, and k is the space step. Similar approximations are used for the derivatives of y, and for θ we use

$$\dot{\theta} \approx \frac{\theta_{n+1,i} - \theta_{n,i}}{h} \quad \text{and} \quad \theta' \approx \frac{\theta_{n,i+1} - \theta_{n,i}}{k}. \tag{9.39}$$

The discrete Lagrangian density can then be found by substituting these approximations into the continuum Lagrangian (9.17). Explicitly, it is given by

$$L_d = \frac{\rho}{2h^2}\left((x_{n+1,i} - x_{n,i})^2 + (y_{n+1,i} - y_{n,i})^2\right) + \frac{\alpha}{2h^2}(\theta_{n+1,i} - \theta_{n,i})^2 - \frac{\beta}{2k^2}(\theta_{n,i+1} - \theta_{n,i})^2$$
$$- \frac{K}{2k^4}(x_{n,i-1} - 2x_{n,i} + x_{n,i+1})^2 - \frac{K}{2k^4}(y_{n,i-1} - 2y_{n,i} + y_{n,i+1})^2. \tag{9.40}$$

Note that L_d only depends on four of the six points in each cell (see (9.33)). The discrete constraint manifold $\mathcal{C}_d$ is found by discretizing the constraint equations (9.24). In order to obtain a second-order accurate approximation, we use central differences:

$$x' \approx \frac{x_{n,i+1} - x_{n,i-1}}{2k},$$

(and similar for $y', \dot{x}, \dot{y}, \dot{\theta}$) and hence we obtain that $\mathcal{C}_d$ is given by

$$x_{n+1,i} - x_{n-1,i} + \frac{R}{2k}(\theta_{n+1,i} - \theta_{n-1,i})(y_{n,i+1} - y_{n,i-1}) = 0, \tag{9.41}$$

and

$$y_{n+1,i} - y_{n-1,i} - \frac{R}{2k}(\theta_{n+1,i} - \theta_{n-1,i})(x_{n,i+1} - x_{n,i-1}) = 0, \tag{9.42}$$

for all (n,i). The semi-discrete constraint manifold $\mathcal{C}_{1/2}$, on the other hand, is given by

$$\dot{x}_{n,i} + \frac{R}{2k}\dot{\theta}_{n,i}(y_{n,i+1} - y_{n,i-1}) = 0,$$

and

$$\dot{y}_{n,i} - \frac{R}{2k}\dot{\theta}_{n,i}(x_{n,i+1} - x_{n,i-1}) = 0,$$

and hence F_d is generated by

$$\Phi^1 = \mathrm{d}x + \frac{R}{2k}(y_{n,i+1} - y_{n,i-1})\mathrm{d}\theta \quad \text{and} \quad \Phi^2 = \mathrm{d}y - \frac{R}{2k}(x_{n,i+1} - x_{n,i-1})\mathrm{d}\theta. \tag{9.43}$$

We conclude that the discrete nonholonomic field equations (9.38) are in this case

$$x_{n+1,i} - 2x_{n,i} + x_{n-1,i} = h^2\lambda_i - \frac{h^2K}{k^4}\Delta^4 x_{n,i} \tag{9.44}$$

and

$$y_{n+1,i} - 2y_{n,i} + y_{n-1,i} = h^2\mu_i - \frac{h^2K}{k^4}\Delta^4 y_{n,i} \tag{9.45}$$

as well as

$$\alpha(\theta_{n+1,i} - 2\theta_{n,i} + \theta_{n-1,i}) = Rh^2\left(\lambda_i\frac{y_{n,i+1} - y_{n,i-1}}{2k} - \mu_i\frac{x_{n,i+1} - x_{n,i-1}}{2k}\right) + \frac{\beta h^2}{k^2}\Delta^2\theta_{n,i},$$

where Δ^2 and Δ^4 are the 2nd and 4th order finite difference operators in the spatial direction, respectively:

$$\Delta^2 f_{n,i} := f_{n,i+1} - 2f_{n,i} + f_{n,i-1}$$

and

$$\Delta^4 f_{n,i} := f_{n,i+2} - 4f_{n,i+1} + 6f_{n,i} - 4f_{n,i-1} + f_{n,i-2}.$$

In order to determine λ_i and μ_i, these equations need to be supplemented by the discrete constraints (9.41) and (9.42).

For the purpose of numerical simulation, the following values were used: $\alpha = 1$, $\beta = 0.8$, $\rho = 1$, $K = 0.7$, $\ell = 4$, and $R = 1$. For the spatial discretization, 32 points were used (corresponding to $k \approx 0.1290$) and the time step was set to $h = 1/8k^2$, a fraction of the maximal allowable time step for the Euler-Bernoulli beam equation (see [**1**]). The ends of the rod were left free and the following initial conditions were used:

$$\mathbf{r}_0(s) = (s, 0), \quad \theta_0(s) = -\frac{\pi}{2}\cos\frac{\pi s}{\ell} \quad \text{and} \quad \dot{\mathbf{r}}_0(s) = (0, 0), \quad \dot{\theta}_0(s) = 0.$$

An mpeg movie (created with Povray, an open source ray tracer) depicting the motion of the nonholonomic Cosserat rod is available from the author's web page[2]. In figure 9.3, an impression of the motion of the rod is given. The arrows represent the director field $\mathbf{d}_3$ and serve as an indication of the torsion. The rod starts from an initially straight, but twisted state and gradually untwists, meanwhile effecting a rotation.

In figure 9.4, the energy of the nonholonomic rod is plotted. Even though our algorithm is by its very nature *not* symplectic (or multi-symplectic – see [**19**]), there is still the similar behaviour of "almost" energy conservation.

---

[2] `http://users.ugent.be/~jvkersch/nonholonomic/`

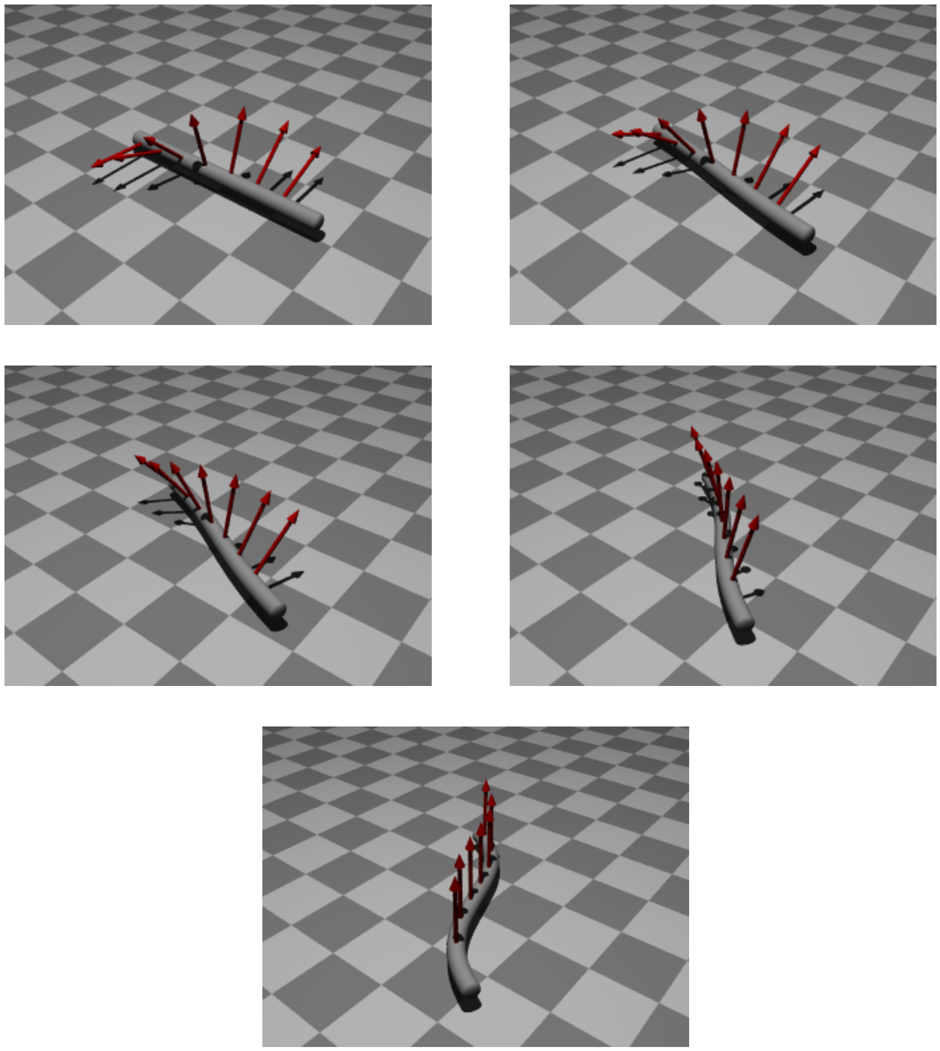


Figure 9.3. Motion of the rod from $t = 0$ to $t \approx 4.5$.

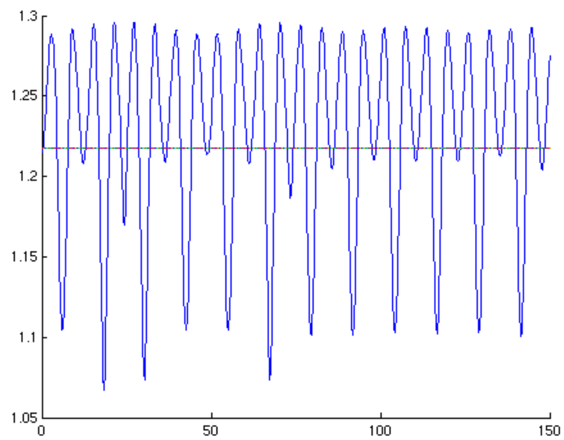


Figure 9.4. Energy behaviour of the integration algorithm.

Chapter 10
Conclusions and future work

1. Discrete jets in higher dimensions

During our treatment of discrete field theories, we have always assumed that the space of independent variables is two-dimensional, and in most cases, we have discretized this space with a regular mesh. However, it is apparent, for example from the derivation of the Euler-Lagrange equations, or the proof of the Euler-Poincaré reduction theorem, that neither is essential for the definition of discrete jets.

For a general base space X, which is not necessarily Euclidian or two-dimensional, we introduce a discretization as follows:

Definition 1.1. *A* discretization *of X is a simplicial complex K together with a homeomorphism $\iota : K \to X$.*

This concept of discretization was studied in detail by many authors; we refer to [**39**, **56**] for an overview of its use in discrete exterior calculus. Note also that such a discretization was the starting point for Whitney's geometric integration theory (see [**111**]).

Discretizations of jets are then easily defined as $(n+2)$-tuples in Y which project down onto an $(n+1)$-simplex in X. This definition is inspired by definition 2.3 in chapter 2, where discrete jets were defined as triples or quadruples projecting down onto a triangle or a quadrangle, respectively.

Definition 1.2. *A* discrete jet *is an $(n+2)$-tuple $(y_0, \ldots, y_{n+1})$ in Y such that the projection $(\pi(y_0), \ldots, \pi(y_{n+1}))$ is an $(n+1)$-simplex in X. The set of all discrete jets is called the* discrete jet bundle*, and is denoted by $J^1_d\pi$.*

The Lie groupoid theory of chapter 3 can also be extended without much difficulty: we define a discrete jet over an $(n+1)$-simplex k to be an assignment of a groupoid element to each edge of k, such that the cyclic multiplication of the groupoid elements on each face of k is a unit. Again, this definition reduces to the familiar case studied in chapter 3 once we take X to be $\mathbb{R}^2$ with its triangular mesh.

Remark 1.3. In the contemporary literature, a few geometric structures have appeared that might be of use in Lie groupoid field theory:

(1) In a recent paper [**113**] on lattice gauge theories, Wise introduces a discretization of space-time by means of a hypothetical "n-graph" structure, which is a list of data $X_0, X_1, X_2, \ldots$, where X_0 is a set of vertices, X_1 a set of edges, and so on, with sets X_i of higher-dimensional objects. These sets have to specify various incidence relations, the nature of which is still not entirely clear. It is probable that this definition could be used instead of definition 1.1. The advantage is that this definition is more in line with the concepts from graph theory which we employed at the beginning of chapter 3.

(2) Recall that in chapter 3 we defined $\mathbb{G}^k$ roughly speaking as k-gons whose edges were labelled by elements of G. In [**7**], the authors consider a similar structure. They construct a simplicial manifold (see [**41**, p. 89] for a definition) whose sets of n-dimensional simplices $\mathbb{G}_n$ are defined as follows: the elements of $\mathbb{G}_n$ are n-simplices whose 1-dimensional faces are labelled by elements of G. Under suitable assumptions, the manifold of k-gons $\mathbb{G}^k$ can be identified with the set of 2-simplices, and we obtain the following sequence:

$$\cdots \mathbb{G}_2 \cong \mathbb{G}^k \substack{\longrightarrow \\ \longrightarrow \\ \longrightarrow} \mathbb{G}_1 \cong G \substack{\longrightarrow \\ \longrightarrow} Q.$$

Presumably, the sets of higher-dimensional simplices would be useful in constructing discretizations of field theories whose base space is of dimension higher than two. ◇

2. Nonholonomic field theories

It is clear that the study of nonholonomic field theories forms a vast subject, and that in this thesis only a few straightforward results could be explored. A first acute problem is the lack of an extensive number of interesting physical examples. The model outlined in section 3.5 of chapter 9 can be extended in a number of ways: first and foremost, one could use the stored energy function for a linear extensible beam instead of the expression used in section 3.5. We did not pursue that road any further, since the resulting field equations quickly became prohibitively unwieldy. Furthermore, to obtain physically interesting results, one should opt for a realistic nonlinear stored-energy function, rather than the quadratic one used in chapter 9.

On the other hand, one could choose to maintain the inextensibility constraint and instead relax the constraints to allow lateral sliding. This model would be somewhat similar to the one studied in chapter 9. From a numerical point of view, the inextensibility constraint could be implemented by an extension of the Shake/Rattle algorithm, as proposed in [**5**].

Another interesting option would be the development of higher-order geometric schemes for nonholonomic field theories. As a first step, one could try to derive higher-order

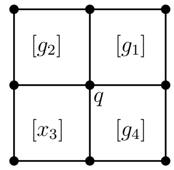


Figure 10.1. A typical element of $\mathbb{G}_2^4$; q is the projection of this element under π_2^4.

integrators for nonholonomic mechanical systems. Such integrators, of arbitrarily high order, can be constructed by composition (see [**84**]) but in order to gain insight into the behaviour of these schemes (in particular the question whether they are amenable to backward error analysis) other classes of integrators should be studied as well.

3. Lie groupoid field theories with nonholonomic constraints

In [**57**], Iglesias *et al.* study mechanical systems with nonholonomic constraints on Lie groupoids. The advantage of this approach is that it incorporates many discretizations of classical systems (including for instance the homogeneous sphere on a rotating horizontal table) which are unrelated at first sight. It would appear that a similar framework can be developed for field theories, thus uniting the two main themes of this thesis, Lie groupoid discretizations and nonholonomic constraints.

As in chapter 3, consider a Lie groupoid G and let $L : \mathbb{G}^k \to \mathbb{R}$ be a discrete Lagrangian. Before introducing nonholonomic constraints into this picture, let us first define a object, denoted by $\mathbb{G}_2^k$ and playing the role of "second-order discrete jet bundle". For the sake of clarity, we henceforth assume that $k = 4$. The elements of $\mathbb{G}_2^4$ are 4-tuples $([g_1], [g_2], [g_3], [g_4]) \in (\mathbb{G}^4)^{\times 4}$ such that (see figure 10.1)

$$[g_1]_4 = [g_2]_2, \quad [g_2]_1 = [g_3]_3, \quad [g_3]_2 = [g_4]_4, \quad \text{and} \quad [g_4]_3 = [g_1]_1.$$

The manifold $\mathbb{G}_2^4$ is equipped with a distinguished projection $\pi_2^4 : \mathbb{G}_2^4 \to Q$, defined as

$$\pi_2^4([g_1], [g_2], [g_3], [g_4]) = \alpha^{(1)}([g_1]) \quad (= \alpha^{(2)}([g_2]) = \alpha^{(3)}([g_3]) = \alpha^{(4)}([g_4])).$$

In addition to L, we also assume that the following objects are given:

(1) a *constraint submanifold* $\mathcal{C}_d \hookrightarrow \mathbb{G}^k$, containing the diagonal Δ in $\mathbb{G}^k$;
(2) a set D_c of *admissible variations*, which is a subbundle of $(\pi_2^4)^* AG$. In other words, the elements of D_c are maps $v : \mathbb{G}_2^4 \to AG$ such that $v([g_1], \ldots, [g_4]) \in A_q G$, where $q = \pi_2^4([g_1], \ldots, [g_4])$. We assume that the rank of D_c is equal to the dimension of $\mathcal{C}_d$.

By varying the action density as in (3.14), but only with respect to admissible variations, we finally obtain the following discrete nonholonomic field equations on the Lie groupoid G (compare with theorem 3.7 in chapter 3):

Theorem 3.1. *Let $\phi : U_E \to G$ be a discrete field defined on a finite set $U_E \subset V \times V$. Then ϕ is an extremum of the action sum (3.9) with respect to admissible variations if and only if ψ takes values in $\mathcal{C}_d$ and the following discrete nonholonomic field equations hold:*

$$v^{(1)}_{[g_1]}(L) + v^{(2)}_{[g_2]}(L) + v^{(3)}_{[g_3]}(L) + v^{(4)}_{[g_4]}(L) = 0 \quad \textit{for all}\, v \in D_c([g_1],[g_2],[g_3],[g_4]) \subset A_qQ,$$

where $([g_1], \ldots, [g_4])$ is defined as in the discussion preceding theorem 3.7 in chapter 3.

Assume that a subbundle D_c° of $(\pi_2^4)^* A^*G$ can be found such that $\langle \Phi, v\rangle = 0$ for all $\Phi \in D_c^\circ([g_1], \ldots, [g_4])$ and all $v \in D_c([g_1], \ldots, [g_4])$, and let Φ^α be a local basis of sections for D_c°. The discrete nonholonomic field equations in theorem 3.1 can be proved to be equivalent to the following set of equations:

$$v^{(1)}_{[g_1]}(L) + v^{(2)}_{[g_2]}(L) + v^{(3)}_{[g_3]}(L) + v^{(4)}_{[g_4]}(L) = \lambda_\alpha \left\langle A^\alpha([g_1],[g_2],[g_3],[g_4]), v\right\rangle, \tag{10.1}$$

for all $v \in A_qQ$.

The discrete field equations for the nonholonomic Cosserat rod can be recovered from (10.1) by taking the Lie groupoid G to be the pair groupoid $Q \times Q$, defining $\mathcal{C}_d$ to be given by (9.41) and (9.42), and taking D_c° to be spanned by (9.43). Some of the definitions above should be adapted since the Lagrangian for the nonholonomic Cosserat rod is of second order.

While it would thus appear that the incorporation of nonholonomic constraints into the framework of Lie groupoid field theories is indeed possible, there are as of yet not enough examples of nonholonomic field theories in order to appreciate the value of such an extension. Judging from the embarrassment of riches with which one is confronted in the case of mechanical systems (see [**57**,**75**]), we are nevertheless confident that Lie groupoid methods will also prove to be useful for discrete field theories with nonholonomic constraints.

Appendix A
Elementary properties of the Frölicher-Nijenhuis bracket

In this appendix, we review some properties of the Frölicher-Nijenhuis bracket and the various derivations associated to vector-valued forms on a manifold. For a detailed treatment of the Frölicher-Nijenhuis bracket, we refer the reader to [**60**, **94**].

Let M be a manifold. A *vector-valued one-form* $\mathbf{h}$ is a section of $TM \otimes T^*M$. Associated to $\mathbf{h}$ is a derivation $i_{\mathbf{h}}$ (of type i_* and degree 0), defined by

$$(i_{\mathbf{h}}\alpha)(v_0, \ldots, v_k) = \sum_{i=0}^{k} (-1)^i \alpha(\mathbf{h}(v_i), v_0, \ldots, \widehat{v_i}, \ldots, v_k) \quad \text{for } \alpha \in \Omega^{k+1}(M). \tag{A.1}$$

We then define $\mathrm{d}_{\mathbf{h}}$ as $\mathrm{d}_{\mathbf{h}} := i_{\mathbf{h}} \circ \mathrm{d} - \mathrm{d} \circ i_{\mathbf{h}}$; this is a derivation of type d_* and degree 1.

Vector-valued forms of higher degree are defined accordingly as sections of the tensor product $TM \otimes \bigwedge^k (T^*M)$. A vector-valued k-form R can easily be seen to give rise to a derivation i_R of degree $k-1$ (by virtue of a generalization of (A.1)) as well as a derivation d_R of degree k. A vector-valued form of degree zero is simply a vector field, and the associated derivations are in this case the contraction i_X and the Lie derivative $\mathscr{L}_X$.

The Frölicher-Nijenhuis bracket of a vector-valued r-form R and a vector-valued s-form S is then defined as the unique vector-valued $(r+s)$-form $[R, S]$ for which

$$\mathrm{d}_R \circ \mathrm{d}_S - (-1)^{rs} \mathrm{d}_S \circ \mathrm{d}_R = \mathrm{d}_{[R,S]}.$$

We have deliberately been vague about the nature of this bracket: most of the time we will only need the bracket of a vector field X with a vector-valued one-form $\mathbf{h}$ (which will be the horizontal projector of a connection). In this case, it is not hard to prove that

$$[X, \mathbf{h}] = \mathscr{L}_X \mathbf{h}.$$

The following lemma collects the properties of the Frölicher-Nijenhuis bracket that we will be needing in the body of the text. They can be suitably generalized and form part of a well-investigated calculus, for which we refer to [**60**].

Lemma 1.2. *Let X be a vector field on M and $\mathbf{h}$ a vector-valued one-form. Then, for any k-form α on M, the following holds:*

(1) $i_X i_{\mathbf{h}} \alpha = i_{\mathbf{h}} i_X \alpha + i_{\mathbf{h}(X)} \alpha$;
(2) $i_{\mathbf{h}} \mathscr{L}_X \alpha = \mathscr{L}_X i_{\mathbf{h}} \alpha - i_{[X,\mathbf{h}]} \alpha$.

Proof: Let α be a 2-form (the case of a k-form α is completely similar) and Y a vector field on M. Then

$$\begin{aligned}(i_X i_{\mathbf{h}} \alpha)(Y) &= \alpha(\mathbf{h}(X), Y) - \alpha(\mathbf{h}(Y), X) \\ &= (i_{\mathbf{h}(X)} \alpha)(Y) + (i_{\mathbf{h}} i_X \alpha)(Y),\end{aligned}$$

which confirms the first property.

The second property (a special case of lemma 8.6 in [**60**]) can be proved directly by noting that a derivation is completely determined by its action on functions and one-forms. For a function f both sides of the relation (2) vanish and for a one-form α we have for the left-hand side

$$\begin{aligned}(i_{\mathbf{h}} \mathscr{L}_X \alpha)(Y) = (\mathscr{L}_X \alpha)(\mathbf{h}(Y)) &= \mathscr{L}_X(\alpha(\mathbf{h}(Y))) - \alpha([X, \mathbf{h}(Y)]) \\ &= \mathscr{L}_X(\alpha(\mathbf{h}(Y))) - \alpha((\mathscr{L}_X \mathbf{h})(Y)) - \alpha(\mathbf{h}([X, Y])).\end{aligned}$$

Taking together the first and third term, we obtain $\mathscr{L}_X(i_{\mathbf{h}} \alpha)(Y)$, whereas the second term is just $i_{[X,\mathbf{h}]} \alpha(Y)$. ◇

Appendix B
Lie groupoids and Lie algebroids

In this section, we recall some of the basic definitions and results from the theory of Lie groupoids and algebroids. It is not our intention to give a detailed introduction to the subject: for a more in-depth overview, the reader is referred to [**72**] and the references therein. We will also recall some of the constructions in [**75**] that will be generalized in the next sections. We note that the definition of a groupoid used here agrees with [**75**,**110**] but differs from [**95**] with respect to the order in which the factors of the product gh are written.

1. Lie groupoids

A *groupoid* is a set G with a partial multiplication m, a subset Q of G whose elements are called *identities*, two surjective maps $\alpha, \beta : G \to Q$ (called *source* and *target* maps respectively), which both equal the identity on Q, and an inversion mapping $i : G \to G$. A pair (g, h) is said to be *composable* if the multiplication $m(g, h)$ is defined; the set of composable pairs will be denoted by G_2. We will denote the multiplication $m(g, h)$ by gh and the inversion $i(g)$ by g^{-1}. In addition, these data must satisfy the following properties, for all $g, h, k \in G$:

(1) the pair (g, h) is composable if and only if $\beta(g) = \alpha(h)$, and then $\alpha(gh) = \alpha(g)$ and $\beta(gh) = \beta(h)$;
(2) if either $(gh)k$ or $g(hk)$ exists, then both do, and they are equal;
(3) $\alpha(g)$ and $\beta(g)$ satisfy $\alpha(g)g = g$ and $g\beta(g) = g$;
(4) the inversion satisfies $g^{-1}g = \beta(g)$ and $gg^{-1} = \alpha(g)$.

On a groupoid, we have a natural notion of left translation l_g, defined as $l_g(h) = gh$, for any $h \in G$ such that $\alpha(h) = \beta(g)$. There is a similar definition for a right translation r_g.

A *morphism of groupoids* is a pair (f, ϕ) of maps $\phi : G \to G'$ and $f : Q \to Q'$ satisfying $\alpha' \circ \phi = f \circ \alpha$, $\beta' \circ \phi = f \circ \beta$ and such that $\phi(gh) = \phi(g)\phi(h)$ whenever (g, h) is composable. Note that $(\phi(g), \phi(h))$ is a composable pair whenever (g, h) is composable.

A *Lie groupoid* is a groupoid for which G and Q are differentiable manifolds, with Q a closed submanifold of G, the maps α, β, m and i are smooth and α and β are submersions. We denote by $\mathcal{F}^\alpha(g)$ the α-fibre through $g \in G$, *i.e.* $\mathcal{F}^\alpha(g) = \alpha^{-1}(\alpha(g))$, with a similar definition for $\mathcal{F}^\beta(g)$. As α and β are submersions, both $\mathcal{F}^\alpha(g)$ and $\mathcal{F}^\beta(g)$ are closed submanifolds of G.

Example 1.1. Any Lie group G can be considered as a Lie groupoid over a singleton $\{e\}$, where the anchors α, β map any element onto e and the multiplication is defined everywhere. Another example of a Lie groupoid is the pair groupoid $Q \times Q$, where $\alpha(q_1, q_2) = q_1$, $\beta(q_1, q_2) = q_2$, and multiplication is defined as $(q_1, q_2) \cdot (q_2, q_3) = (q_1, q_3)$. For other, less trivial examples, we refer to the works mentioned above. $\diamond$

2. Lie algebroids

Definition 2.1. *A* Lie algebroid *over Q is a vector bundle $\tau : E \to Q$ together with a vector bundle map $\rho : E \to TQ$ (called the* anchor map *of the Lie algebroid) and a bracket $[\cdot,\cdot] : \Gamma(\tau) \times \Gamma(\tau) \to \Gamma(\tau)$ defined on the sections of τ, such that*

(1) $\Gamma(\tau)$ *is a real Lie algebra with respect to* $[\cdot,\cdot]$*;*
(2) $\rho([\phi,\psi]) = [\rho(\phi), \rho(\psi)]$*, for all $\phi, \psi \in \Gamma(\tau)$, where the bracket on the right-hand side is the usual Lie bracket of vector fields on Q and we write the composition $\rho \circ \phi$ as $\rho(\phi)$;*
(3) $[\phi, f\psi] = f[\phi,\psi] + \rho(\phi)(f)\psi$*, for all $\phi, \psi \in \Gamma(\tau)$ and $f \in C^\infty(Q)$.*

The Lie algebroid structure allows us to define an exterior differential d_E on the space of sections of $\bigwedge^*(E^*)$, as follows: for functions $f \in C^\infty(Q)$, we put $\mathrm{d}_E f(v) = \rho(v) f$ (where $v \in E$), while for sections θ of $\bigwedge^k(E^*)$, we define $\mathrm{d}_E\theta$ by

$$\mathrm{d}_E\theta(v_0, v_1, \ldots, v_k) = \sum_i \rho(v_i)\theta(v_0, \ldots, \hat{v}_i, \ldots, v_k) + \sum_{i<j} (-1)^{i+j}\theta([v_i, v_j], v_0, \ldots, \hat{v}_i, \ldots, \hat{v}_j, \ldots, v_k).$$

It can be shown that d_E is nilpotent: $\mathrm{d}_E^2 = 0$.

To any Lie groupoid G over Q one can associate a Lie algebroid $\tau : AG \to Q$ as follows. At each point $x \in Q$, the fibre A_xG is the vector space $V_x\alpha = \ker T_x\alpha$ and the anchor map ρ on A_xG is identified with the restriction of $T_x\beta$ to $V_x\alpha$. In order to define the bracket on the space of sections, we note that there exists a bijection between sections of τ and left- and right-invariant vector fields on G. More specifically, if v is a section of τ, then the left- and right-invariant vector fields are denoted as v^L and v^R respectively, and defined by

$$v^L(g) = T_{\beta(g)} l_g(v_{\beta(g)}) \quad \text{and} \quad v^R(g) = T_{\alpha(g)}(r_g \circ i)(v_{\alpha(g)}). \tag{B.1}$$

Let v and w be sections of τ. The bracket $[v, w]$ is then defined by noting that $[v^L, w^L]$ is again a left-invariant vector field, and putting

$$[v, w]^L = [v^L, w^L].$$

We remark that our definition of v^R differs in sign from the one used in [**75**].

Conversely, we say that a Lie algebroid $\tau : E \to Q$ is *integrable* whenever one can find a Lie groupoid such that E is its associated Lie algebroid. It has been known for some years that not all Lie algebroids are integrable. Necessary and sufficient conditions for integrability have been given in [**29**].

Example 2.2. The Lie algebroid of a Lie group G is just its Lie algebra. The Lie algebroid of the pair groupoid $Q \times Q$ is the tangent bundle TQ. ◇

Remark 2.3. For a given section v of τ, we have denoted the corresponding left- and right-invariant vector fields as v^L and v^R, respectively. We will also use this notation for the point-wise operation, by denoting, for v_x an element of A_xG and $g \in \alpha^{-1}(x) \subset G$, the left translated vector $T_x l_g(v_x)$ as $(v_x)^L(g)$, and similarly the right translated vector $T_x(r_g \circ i)(v_x)$ as $(v_x)^R(g)$. ◇

2.1. Lie algebroid morphisms. Consider two vector bundles $\tau' : E' \to Q'$ and $\tau : E \to Q$, and let $\Phi = (\underline{\Phi}, \overline{\Phi})$ be a vector bundle map from τ' to τ. Let θ be a section of $\bigwedge^k(E^*)$. Then the *pullback of* θ *by* Φ is the section $\Phi^\star\theta$ of $\bigwedge^k(E'^*)$ defined as

$$(\Phi^\star\theta)_q(v_1, \ldots, v_k) = \theta_{\underline{\Phi}(q)}(\overline{\Phi}(v_1), \ldots, \overline{\Phi}(v_k)), \quad v_1, \ldots, v_k \in E'_q.$$

Note that we used a "star" $\star$ instead of an "asterisk" $*$ to denote the pullback, which should serve as a reminder that we consider the pullback of θ by a *bundle map* rather than by an arbitrary differentiable map from E' to E.

Now, assume that both τ and τ' are equipped with the structure of a Lie algebroid over Q and Q', respectively, and denote their respective anchor mappings by ρ and ρ'. In this case, a vector bundle map Φ is said to be a *morphism of Lie algebroids* if for each section θ of $\bigwedge^k(E^*)$,

$$\Phi^\star \mathrm{d}_E\theta = \mathrm{d}_{E'}\Phi^\star\theta,$$

where d_E and $\mathrm{d}_{E'}$ are the differentials on E and E', respectively. In other words, Φ is a *chain map*.

A more practical criterion to decide whether a given bundle map is a morphism of Lie algebroids is given below. We say that a section ϕ' of τ' and a section ϕ of τ are Φ-*related* if $\overline{\Phi} \circ \phi' = \phi \circ \underline{\Phi}$.

Proposition 2.4. *Let $\tau' : E' \to Q'$ and $\tau : E \to Q$ be Lie algebroids as in the preceding discussion, and let $\Phi : E' \to E$ be a fibrewise surjective vector bundle morphism. Then Φ is a morphism of Lie algebroids if the following conditions are satisfied:*

(1) $\rho \circ \overline{\Phi} = T\underline{\Phi} \circ \rho'$;
(2) *if* ϕ' *is* Φ*-related with* ϕ*, and* ψ' *with* ψ*, then* $[\phi', \psi']$ *is* Φ*-related with* $[\phi, \psi]$.

Proof: This is proposition 1.5 in [**55**]. ◇

A final class of Lie algebroid morphisms consists of morphisms induced by Lie groupoid morphisms. Consider Lie groupoids G' and G over Q' and Q, respectively. Any groupoid morphism $\Phi = (\underline{\Phi}, \overline{\Phi})$ from G' to G induces a bundle map, denoted by $A\Phi$, from AG' to AG. The base map of $A\Phi$ is just $\underline{\Phi}$, while the total space map is defined as follows:

$$A\Phi(v_{q'}) := T\overline{\Phi}(v_{q'}) \in V_{\underline{\Phi}(q')}(\alpha) = A_{\underline{\Phi}(q')}G.$$

See also theorem 1.7 in [**55**].

3. Prolongations of Lie groupoids and algebroids

3.1. The prolongation of a Lie groupoid over a fibration. Let G be a Lie groupoid over a manifold Q with source and target maps α and β and consider a fibration $\pi : P \to Q$. The *prolongation* $P^\pi G$ is the Lie groupoid over P defined as

$$P^\pi G = \{(g; p_1, p_2) \in G \times P \times P : \pi(p_1) = \alpha(g) \text{ and } \beta(g) = \pi(p_2)\}.$$

Alternatively, $P^\pi G$ is defined by means of the following commutative diagram:

$$\begin{array}{ccc} P^\pi G & \longrightarrow & P \times P \\ \downarrow & & \downarrow{\scriptstyle \pi\times\pi} \\ G & \xrightarrow[(\alpha,\beta)]{} & Q \times Q \end{array} \tag{B.2}$$

It can be shown that $P^\pi G$ is a Lie groupoid over P, with source and target mappings $\alpha^\pi, \beta^\pi : P^\pi G \to P$ defined as

$$\alpha^\pi(g; p_1, p_2) = p_1 \quad \text{and} \quad \beta^\pi(g; p_1, p_2) = p_2,$$

and with multiplication given by

$$(g; p_1, p_2)(h; p_2, p_3) = (gh; p_1, p_3).$$

Note that $\alpha^\pi(h; p_2, p_3) = \beta^\pi(g; p_1, p_2)$ implies that $\alpha(h) = \beta(g)$. Finally, the inversion mapping is defined as

$$i : (g; p_1, p_2) \mapsto (g^{-1}; p_2, p_1),$$

and we can regard P as a subset of $P^\pi G$ via the identification $p \mapsto (\pi(p); p, p)$.

3.1.1. *The prolongation PG.* There is one particular prolongation that will play a significant role in what follows. It is obtained by taking for the fibration $\pi : P \to Q$ in (B.2) the Lie algebroid projection $\tau : AG \to Q$ to obtain

$$P^\tau G \subset G \times AG \times AG$$

which, henceforth, we also simply denote as PG. We recall that PG consists of triples $(g; v_x, w_y)$, where $g \in G$, $v_x \in A_x G$, $w_y \in A_y G$, and $x = \alpha(g)$, $y = \beta(g)$. It is pointed out in [**75**, **95**] that PG is isomorphic as a vector bundle over G to the direct sum $V\beta \oplus V\alpha$, where $V\alpha$ is the subbundle of TG consisting of α-vertical vectors (and similarly for $V\beta$); the isomorphism $\Theta : PG \to V\beta \oplus V\alpha$ is defined by

$$\Theta(g; u_{\alpha(g)}, v_{\beta(g)}) = (T(r_g \circ i)(u_{\alpha(g)}), Tl_g(v_{\beta(g)})). \tag{B.3}$$

It should also be remarked that PG is a vector bundle over G, and in fact, PG can be endowed with the structure of an integrable Lie algebroid over G, where the anchor map $\hat{\rho} : PG \to TG$ is given by

$$\hat{\rho} : (g; u_{\alpha(g)}, v_{\beta(g)}) \mapsto T(r_g \circ i)(u_{\alpha(g)}) + Tl_g(v_{\beta(g)}) = (u_{\alpha(g)})^R(g) + (v_{\beta(g)})^L(g).$$

The bracket of PG is induced by the Lie algebroid structure of AG. Let ϕ be a section of AG; we then define sections $\phi^{(1,0)}$ and $\phi^{(0,1)}$ of PG as follows:

$$\phi^{(1,0)}(g) = (g; \phi(\alpha(g)), 0) \quad \text{and} \quad \phi^{(0,1)}(g) = (g; 0, \phi(\beta(g))).$$

The bracket of sections of PG is then determined by the following relations:

$$[\phi^{(1,0)}, \psi^{(1,0)}]_{PG} = [\phi, \psi]^{(1,0)}, \quad [\phi^{(0,1)}, \psi^{(0,1)}]_{PG} = [\phi, \psi]^{(0,1)}, \quad \text{and} \quad [\phi^{(1,0)}, \psi^{(0,1)}]_{PG} = 0.$$

For more information, see [**75**].

A groupoid morphism $\Phi : (G', Q') \to (G, Q)$ naturally induces a map $P\Phi : PG' \to PG$, defined as $P\Phi(g; u, v) = (\Phi(g); A\Phi(u), A\Phi(v))$. It is easy to see that $P\Phi$ is an algebroid morphism.

3.2. The prolongation of a Lie algebroid over a fibration. Let $\tau : E \to Q$ be a Lie algebroid and consider a fibration $\pi : P \to Q$. The *prolongation* $P^\pi E$ is the Lie algebroid over P defined as

$$P^\pi E = \{(a, v) \in E \times TP : \rho(a) = T\pi(v)\},$$

or by the following commutative diagram as

$$\begin{array}{ccc} P^\pi E & \longrightarrow & TP \\ \downarrow & & \downarrow {\scriptstyle T\pi} \\ E & \xrightarrow[\rho]{} & TQ \end{array} \tag{B.4}$$

We denote by $\hat{\pi} : P^\pi E \to P$ the map defined as $\hat{\pi}(a, v) = \pi_{TP}(v)$, where $\pi_{TP} : TP \to P$ is the tangent bundle projection of P. It can be shown that $\hat{\pi} : P^\pi E \to P$ can be given the structure of a Lie algebroid (see [**55**, **83**, **95**]).

3.2.1. *The prolongations $P^\tau(AG)$ and $P^{\tau^*}(AG)$.* Let G be a Lie groupoid over a manifold Q with Lie algebroid $\tau : AG \to Q$. By taking for the fibration π underlying diagram (B.4) the map τ, we obtain the prolongation $P^\tau(AG)$. It is very useful to think of $P^\tau(AG)$ as a sort of Lie algebroid analogue of the *tangent bundle* to AG. Indeed, $P^\tau(AG)$ can be equipped with geometric objects, such as a Liouville section and a vertical endomorphism, which have their counterpart in tangent bundle geometry.

Similarly, by taking for $\pi : P \to Q$ the dual bundle $\tau^* : A^*G \to Q$, we obtain the prolongation $P^{\tau^*}(AG)$, which is a Lie algebroid over A^*G and should be thought of as the Lie algebroid analogue of the tangent bundle to A^*G. Just as any cotangent bundle is equipped with a canonical one-form, there exists a canonical section

$$\theta : A^*G \to \left[P^{\tau^*}(AG)\right]^*,$$

defined as follows: for $\alpha \in A^*G$ and $(v, X_\alpha) \in (P^{\tau^*}(AG))_\alpha$, we put $\theta_\alpha(v, X_\alpha) = \alpha(v)$. In the case that G is the pair groupoid $Q \times Q$, we have that $A^*G = T^*Q$ and we obtain the usual canonical one-form on T^*Q.

It was shown in [**55**] that $P^\tau(AG)$, the prolongation of the Lie algebroid AG, is isomorphic to $A(PG)$, the Lie algebroid associated to the prolongation Lie groupoid PG.

3.2.2. *The prolongations $P^\alpha(AG)$ and $P^\beta(AG)$.* Associated to the source and target mappings α and β of a groupoid G there are two prolongations $P^\alpha(AG)$ and $P^\beta(AG)$, whose fibres over G are defined as follows: for each $g \in G$, put

$$P^\alpha_g(AG) = \{(v_{\alpha(g)}, X_g) \in A_{\alpha(g)}G \times T_gG : T\beta(v_{\alpha(g)}) = T\alpha(X_g)\}$$

and

$$P^\beta_g(AG) = \{(v_{\beta(g)}, X_g) \in A_{\beta(g)}G \times T_gG : T\beta(v_{\beta(g)}) = T\beta(X_g)\}.$$

Both of these algebroids are integrable: indeed, it follows from the general theory that $P^\alpha(AG)$ is isomorphic to the Lie algebroid of the prolongation $P^\alpha G$, and similarly for $P^\beta(AG)$.

Furthermore, we remark that there are two distinguished mappings from PG (regarded as a Lie *algebroid* over G) into $P^\alpha(AG)$ and $P^\beta(AG)$, given by

$$A\Phi^\alpha : (g; u_{\alpha(g)}, v_{\beta(g)}) \mapsto (g; u_{\alpha(g)}, T(r_g \circ i)(u_{\alpha(g)}) + Tl_g(v_{\beta(g)})) \in P^\alpha(AG)$$

and

$$A\Phi^\beta : (g; u_{\alpha(g)}, v_{\beta(g)}) \mapsto (g; v_{\beta(g)}, T(r_g \circ i)(u_{\alpha(g)}) + Tl_g(v_{\beta(g)})) \in P^\beta(AG).$$

The notations $A\Phi^\alpha$ and $A\Phi^\beta$ serve as a reminder of the fact that these Lie algebroid maps stem from morphisms between the corresponding groupoids (see [**75**]).

Nederlandse samenvatting

In dit proefschrift worden enkele aspecten van de differentiaalmeetkundige theorie van klassieke veldentheorieën behandeld. In het bijzonder geven we een aanzet tot het vinden van een antwoord op de volgende vragen:

(1) Hoe kunnen methoden uit de theorie van Liegroepoïden gebruikt worden om structuurbewarende discretisaties van klassieke veldentheorieën op te stellen?
(2) Bestaan er fysisch relevante voorbeelden van klassieke veldentheorieën met niet-holonome bindingen?

Discrete veldentheorieën met waarden in een Liegroepoïde

Situering. De motivatie voor dit deel van het proefschrift is terug te voeren op het werk van Marsden, Patrick en Shkoller [**80**], die structuurbehoudende discretisaties construeerden voor klassieke veldentheorieën (zie ook [**18**, **68**]). We veralgemenen Liegroepoïde-methodes uit [**75**] om discretisaties op te stellen van klassieke veldentheorieën met waarden in een Liegroepoïde G. Deze Liegroepoïde wordt als gegeven verondersteld en is afhankelijk van de probleemstelling:

- voor de triviale groepoïde $Q \times Q$ bekomen we uiteindelijk het standaardformalisme van Marsden, Patrick en Shkoller (paragraaf 4.1 in hoofdstuk 3);
- in alle andere gevallen bekomen we nieuwe resultaten. Belangrijk is vooral het geval dat in hoofdstuk 5 besproken wordt, waar G een Liegroep is.

Het belang van dit formalisme is tweevoudig. Allereerst werpt deze beschrijving een nieuw licht op de meetkundige achtergrond van discrete veldentheorieën, zoals we verderop zullen bespreken. Daarenboven duiken Liegroepoïden op natuurlijke wijze op bij symmetriereductie van discrete veldentheorieën: zelfs de reductie van een "triviale" veldentheorie met waarden in $Q \times Q$ geeft aanleiding tot een veldentheorie die waarden aanneemt in de *Atiyah-groepoïde* $(Q \times Q)/\mathcal{G}$ (met $\mathcal{G}$ een Liegroep), waarvoor voorheen geen concrete beschrijving bestond.

Discrete velden. In hoofdstuk 3 stellen we een meetkundig formalisme op voor veldentheorieën die waarden aannemen in een Liegroepoïde G. Naast de specificatie van een geschikte Liegroepoïde G nemen we ook aan dat er een bepaald planair graaf (V, E)

in $\mathbb{R}^2$ gegeven is, dat dienst doet als domein voor de discrete velden. Deze discrete velden worden opgevat als afbeeldingen van de verzameling van bogen van dit graaf naar de Liegroepoïde G. De specificatie van dit graaf is probleem-afhankelijk: met het oog op numerieke integratie zou men er bijvoorbeeld voor kunnen kiezen om de dichtheid van de elementen van V en E te laten toenemen in gebieden van $\mathbb{R}^2$ waar de velden veel variëren. Uiteindelijk bekomen we zo de volgende definitie:

Definitie 1. *Een* discreet veld *is een paar $\phi = (\phi_{(0)}, \phi_{(1)})$, waarbij $\phi_{(0)}$ een afbeelding is van V naar Q en $\phi_{(1)}$ een afbeelding is van E naar G zo dat*

(1) $\alpha(\phi_{(1)}(x,y)) = \phi_{(0)}(x)$ *en* $\beta(\phi_{(1)}(x,y)) = \phi_{(0)}(y)$*;*
(2) *voor elke $(x,y) \in E$ hebben we dat $\phi_{(1)}(y,x) = [\phi_{(1)}(x,y)]^{-1}$.*
(3) *voor alle $x \in V$ geldt er dat $\phi_{(1)}(x,x) = \phi_{(0)}(x)$.*

Centraal in onze beschrijving is de variëteit $\mathbb{G}^k$, waarvan de elementen k-tupels zijn van samenstelbare elementen in G zodat de cyclische vermenigvuldiging een eenheid in G oplevert:

$$(g_1, g_2, \ldots, g_k) \in \mathbb{G}^k \quad \text{als} \quad (g_i, g_{i+1}) \in G_2 \quad (\text{voor } i = 1, \ldots, k) \quad \text{en} \quad g_1 \cdot g_2 \cdots g_k = e_{\alpha(g_1)},$$

waarbij G_2 de verzameling van samenstelbare paren in $G \times G$ voorstelt. In het discrete kader speelt de variëteit $\mathbb{G}^k$ de rol die door de jetbundel $J^1\pi$ in het continue geval vervuld wordt.

Uit stelling 2.2 in hoofdstuk 3 kunnen we dan afleiden dat elk discreet veld ϕ uit te breiden valt tot een afbeelding $\varphi : V \times V \to G$ die daarenboven aan alle eigenschappen voldoet van een morfisme van groepoïdes. Een alternatieve karakterisering wordt gegeven in stelling 2.4: elk morfisme van $V \times V$ naar G kan opgevat worden als een afbeelding van F, de vlakken van de graaf, naar $\mathbb{G}^k$.

Nu induceert de meetkundige structuur van de Liegroepoïde G een aantal interessante structuren op $\mathbb{G}^k$ die gebruikt kunnen worden voor discrete veldentheorie, maar ook op zich de moeite van het bestuderen waard zijn. Over $\mathbb{G}^k$ bestaat er eerst en vooral een bepaalde vectorbundel, die we met $P^k\mathbb{G}$ noteren en gedefinieerd wordt door het onderstaand commutatief diagram:

$$\begin{array}{ccc} P^k\mathbb{G} & \longrightarrow & AG \times \cdots \times AG \\ \downarrow & & \downarrow \\ \mathbb{G}^k & \longrightarrow & Q \times \cdots \times Q \end{array}$$

Hierbij stelt AG de Lie-algebroïde voor die geassocieerd is aan G. In het geval van discrete mechanica (dat immers opgevat kan worden als deelgebied van discrete veldentheorie) valt deze bundel samen met de prolongatie PG van G. Net zoals PG zelf een Lie-algebroïde is, kan men aantonen dat ook $P^k\mathbb{G}$ uitgerust kan worden met de structuur van een Lie-algebroïde. Dit wordt aangetoond in paragraaf 2.1.2. Merk op dat er een aantal canonische projecties van $P^k\mathbb{G}$ op PG bestaan (zie (3.5)); wanneer we $P^k\mathbb{G}$ uitrusten met de Lie-algebroïdestructuur gedefinieerd door (3.8) en PG met de gekende structuur, dan worden deze projecties gepromoveerd tot morfismen van Lie-algebroïdes. Deze eigenschap definieert bovendien de Lie-algebroïdestructuur op $P^k\mathbb{G}$ op unieke wijze.

De Lie-algebroïde $P^k\mathbb{G}$ speelt een cruciale rol in de opbouw van de meetkundige theorie van discrete velden. Een algemene eigenschap uit de theorie van Lie-algebroïdes leert dat de modulen van secties van de duale bundel (hier genoteerd als $P^k\mathbb{G}^*$) en van de antisymmetrische producten $\bigwedge P^k\mathbb{G}^*$ kunnen voorzien worden van een uitwendige afgeleide

$$\mathrm{d}^{(k)} : \bigwedge^n P^k\mathbb{G}^* \to \bigwedge^{n+1} P^k\mathbb{G}^*.$$

De secties van $\bigwedge P^k\mathbb{G}^*$ spelen in deze context dus de rol van differentiaalvormen. Met behulp van de operator $\mathrm{d}^{(k)}$ kunnen we dan k *Poincaré-Cartanvormen* $\theta_L^{(i)}$ definiëren, die secties zijn van $P^k\mathbb{G}^*$ met de volgende kenmerkende eigenschap:

$$\mathrm{d}^{(k)} L = \sum_{i=1}^{k} \theta_L^{(i)}.$$

Hierbij is L een gegeven functie op $\mathbb{G}^k$, die we aanduiden als de *discrete Lagrangiaan.*

Een andere blik op de Poincaré-Cartanvormen wordt geboden door de *discrete Legendretransformaties*, een verzameling van k morfismen van $P^k\mathbb{G}$ naar een nieuwe prolongatiebundel $P^{\tau^*}(AG)$ over A^*G. Deze bundel wordt gedefinieerd in appendix B en de duale bundel ervan is voorzien van een canonische sectie θ (zie [**38**]), die ruwweg overeenkomt met de canonische 1-vorm op een coraakbundel. In stelling 3.9 wordt dan aangetoond dat de pull-back van deze canonische sectie langsheen elk van de k discrete Legendretransformaties precies de verzameling van Poincaré-Cartanvormen oplevert.

In paragraaf 3.2 leiden we uiteindelijk een stelsel van vergelijkingen af die de dynamica van een discreet veld specificeren. De oplossingen van deze veldvergelijkingen zijn de extrema van de actie S, die als volgt gedefinieerd wordt:

$$S(\phi) = \sum L(\psi([x])), \tag{2.5}$$

waarbij de som genomen wordt over een eindige deelverzameling van F, de verzameling van vlakken van het graaf, en $\psi : F \to \mathbb{G}^k$ de afbeelding is geassocieerd met het morfisme ϕ. Zo bekomen we de volgende stelling:

Stelling 2. *Zij $\phi : U_E \to G$ een discreet veld gedefinieerd op een eindige deelverzameling $U_E \subset V \times V$. Dan is ϕ een extremum van de discrete actie (2.5) als en slechts als voor alle $v \in A_qG$ (waar $q = \phi_{(0)}(x_{i,j})$), aan de volgende* discrete Euler-Lagrangevergelijkingen *voldaan is*

$$v^{(1)}_{[g_1]}(L) + v^{(2)}_{[g_2]}(L) + v^{(3)}_{[g_3]}(L) + v^{(4)}_{[g_4]}(L) = 0. \tag{2.6}$$

Hierbij zijn de elementen $[g_i]$, $i = 1, \ldots, 4$ gedefinieerd als op figuur 3.4.

In het geval dat G de triviale groepoïde $Q \times Q$ is, vallen deze veldvergelijkingen samen met de differentievergelijkingen afgeleid door Marsden, Patrick en Shkoller [**80**]. Andere types van discrete veldvergelijkingen worden behandeld in hoofdstuk 5.

Merk op dat de vergelijkingen (2.6) geen naïeve discretisatie zijn van een stel partiële differentiaalvergelijkingen: ze ontstaan daarentegen door het gebruik van discrete versies van de onderliggende meetkundige structuren. In het bijzonder zijn de resulterende vergelijkingen afkomstig van een discrete actie (in dit geval gegeven door (2.5)) en men kan daaruit besluiten dat de oplossingen ervan voldoen aan de speciale behoudswet (3.18). Deze wet, waarnaar men verwijst als de wet van behoud van multisymplecticiteit (zie [**19**]), is het veldentheoretisch analogon van het symplectisch-zijn van de veelgeroemde geometrische integratieschema's voor klassieke mechanica. Aangezien symplectische integratoren over het algemeen een gedrag vertonen dat kwalitatief veel beter is dan dat van traditionele, niet-symplectische integratoren, neemt men aan dat multisymplectische integratoren eveneens superieure resultaten zullen opleveren. Deze verwachting werd ten dele ingelost in [**80**], maar een theoretische onderbouwing ontbreekt vooralsnog.

Discrete veldentheorieën met symmetrie

Zoals reeds voorheen aangehaald werd, levert de symmetriereductie van een discrete veldentheorie met een bepaalde symmetrie een nieuwe discrete veldentheorie op. De "target space" van deze gereduceerde veldentheorie is zelfs in de meest eenvoudige gevallen een niet-triviale Liegroepoïde: het gebruik van het formalisme uit hoofdstuk 3 dringt zich dus op.

In hoofdstuk 4 beschouwen we enkele algemene gevolgen van de aanwezigheid van symmetrie. Dit gegeven wordt vrij abstract geïnterpreteerd als het bestaan van een nieuwe groepoïde G' en een submersief morfisme Φ van G naar G'. Hierbij speelt G' de rol van "gereduceerde Liegroepoïde". Om deze definitie wat concreter te maken, is het aangewezen bijvoorbeeld het geval te beschouwen waarbij G de triviale groupoïde $Q \times Q$ is en $\mathcal{G}$ een Liegroep die regulier werkt op Q. In dat geval werkt $\mathcal{G}$ ook regulier (via de diagonale actie) op $Q \times Q$ en is het morfisme Φ gewoon de quotiëntafbeelding $\Phi : Q \times Q \to (Q \times Q)/\mathcal{G}$.

Het eerste deel van hoofdstuk 4 is gewijd aan enkele algemene eigenschappen van een dergelijk morfisme Φ. Op het niveau van de prolongatiebundels induceert Φ een afbeelding Ψ van $P^k\mathbb{G}$ naar $P^k\mathbb{G}'$, waarvan men kan aantonen (stelling 1.2) dat het een morfisme van Lie-algebroïdes is. Neem dan aan dat L en L' twee discrete Lagrangianen zijn op respectievelijk $\mathbb{G}^k$ en $\mathbb{G}'^k$, die door pull-back gerelateerd zijn: $L = \Psi^\star L'$. In het geval van de situatie geschetst op het eind van de vorige paragraaf wil dit gewoon zeggen dat L $\mathcal{G}$-invariant is, en dat L' de geïnduceerde Lagrangiaan is op de quotiëntruimte.

Zo komen we dan uiteindelijk tot de volgende reductiestelling:

Stelling 3 (Reductie). *Beschouw een submersief morfisme $\Phi : (G, Q) \to (G', Q')$. Zij $L : \mathbb{G}^3 \to \mathbb{R}$ een discrete Lagrangiaan op $\mathbb{G}^3$ en beschouw de geassocieerde Lagrangiaan $L' = \Psi^\star L$ op $\mathbb{G}'^3$.*

Dan is een morfisme $\phi : V \times V \to G$ een oplossing voor de discrete veldvergelijkingen voor L als en slechts als het geïnduceerde morfisme $\Phi \circ \phi : V \times V \to G'$ een oplossing is voor de discrete veldvergelijkingen geassocieerd met L'.

In geval waarbij L en L' door pull-back via Ψ gerelateerd zijn, kan men verder nog aantonen (stelling 1.4) dat ook de Poincaré-Cartanvormen van L en L' op dergelijke manier met elkaar in verband staan. Hieruit kan men afleiden dat een gereduceerde veldentheorie multisymplectisch is in het geval dat de originele veldentheorie dat ook is. Onder reductie worden multisymplectische veldentheorieën dus omgezet in nieuwe veldentheorieën, die eveneens multisymplectisch zijn.

In het tweede deel van hoofdstuk 4 wordt dan een ander aspect van symmetrie besproken, namelijk de aanwezigheid van behoudswetten geassocieerd met een symmetrie (de zogenaamde stelling van Noether). We definiëren (infinitesimale) Noethersymmetrieën als secties van AG die de Lagrangiaan invariant laten, op een aantal specifieke termen na, die niet aan de dynamica bijdragen. Het mag geen verwondering wekken dat ook in het discrete geval een Noethersymmetrie aanleiding geeft tot een bepaalde behoudswet, die in dit geval echter discreet is.

De Euler-Poincarévergelijkingen. De ontwikkelingen in hoofdstuk 4 komen echter pas tot hun volle recht in het daaropvolgende hoofdstuk, waarin de Euler-Poincarévergelijkingen besproken worden. In dit hoofdstuk beschouwen we de Liegroepoïde $\mathcal{G} \times \mathcal{G}$, waarbij $\mathcal{G}$ een Liegroep is, en nemen we aan dat $L : \mathcal{G}^{\times 3} \to \mathbb{R}$ een links-invariante Lagrangiaan is:

$$L(gg_1, gg_2, gg_3) = L(g_1, g_2, g_3) \quad \text{voor alle } g \in \mathcal{G} \text{ en } (g_1, g_2, g_3) \in \mathcal{G}^{\times 3}.$$

In dit geval bestaat er een gereduceerde Lagrangiaan ℓ op de variëteit van k-gonen geassocieerd aan de gereduceerde groupoid $(\mathcal{G} \times \mathcal{G})/\mathcal{G}$ (die isomorf is met $\mathcal{G}$ zelf). Expliciet wordt ℓ gegeven door $\ell(g_1^{-1}g_2, g_1^{-1}g_3) = L(g_1, g_2, g_3)$. Aangezien de quotiëntafbeelding $\Phi : \mathcal{G} \times \mathcal{G} \to (\mathcal{G} \times \mathcal{G})/\mathcal{G} \cong \mathcal{G}$ een surjectieve submersie is, is de reductiestelling (stelling 3) toepasbaar en besluiten we dat een morfisme $\phi : V \times V \to \mathcal{G} \times \mathcal{G}$ een oplossing is van de Euler-Lagrangevergelijkingen van L als en slechts als het morfisme $\Phi \circ \phi : V \times V \to \mathcal{G}$ een oplossing is van het gereduceerde vraagstuk. De Euler-Lagrangevergelijkingen geassocieerd aan de gereduceerde Lagrangiaan ℓ worden aangeduid als de *Euler-Poincarévergelijkingen* en kunnen direct afgeleid worden door het formalisme uit hoofdstuk 3 toe te passen op de Liegroep $\mathcal{G}$ (opgevat als Liegroepoïde).

Stelling 4. *Zij L een $\mathcal{G}$-invariante Lagrangiaan op $\mathcal{G}^{\times 3}$ en beschouw de gereduceerde Lagrangiaan ℓ. Beschouw een discreet veld $\phi : V \times V \to \mathcal{G} \times \mathcal{G}$ en zij $\varphi : V \times V \to \mathcal{G}$ het gereduceerde veld gedefinieerd als $\varphi = \Phi \circ \phi$. Dan zijn de volgende uitspraken equivalent:*

(a) *ϕ is een oplossing van de Euler-Lagrangevergelijkingen voor de Lagrangiaan L;*
(b) *ϕ is een extremum van de actie S voor willekeurige variaties;*
(c) *het gereduceerde morfisme φ is een oplossing van de discrete Euler-Poincarévergelijkingen:*

$$\left[\left(R^*_{u_{i,j}} \mathrm{d}\ell(\cdot, v_{i,j})\right)_e - \left(L^*_{u_{i-1,j}} \mathrm{d}\ell(\cdot, v_{i-1,j})\right)_e\right] + \left[\left(R^*_{v_{i,j}} \mathrm{d}\ell(u_{i,j}, \cdot)\right)_e - \left(L^*_{v_{i,j-1}} \mathrm{d}\ell(u_{i,j-1}, \cdot)\right)_e\right] = 0; \tag{2.7}$$

(d) *het gereduceerde morfisme φ is een extremum van de gereduceerde actie, voor variaties van de volgende vorm:*

$$\delta u_{i,j} = TR_{u_{i,j}}(\theta_{i,j+1}) - TL_{u_{i,j}}(\theta_{i,j}) \in T_{u_{i,j}}\mathcal{G} \tag{2.8}$$

en

$$\delta v_{i,j} = TR_{v_{i,j}}(\theta_{i,j+1}) - TL_{v_{i,j}}(\theta_{i,j}) \in T_{v_{i,j}}\mathcal{G}, \tag{2.9}$$

waarbij $\theta_{i,j} = TL_{\phi^{-1}_{i,j}}(\delta\phi_{i,j}) \in \mathfrak{g}$.

Merk op dat we, in tegenstelling tot de situatie voor mechanische systemen, niet naïefweg kunnen zeggen dat er een bijectief verband bestaat tussen oplossingen van het gereduceerde en het ongereduceerde vraagstuk. Niet alle oplossingen van het gereduceerde vraagstuk kunnen immers geschreven worden in de vorm $\varphi \equiv \Phi \circ \phi$. De vraag stelt zich dus wanneer een oplossing van het gereduceerde vraagstuk afkomstig is van een oplossing van het oorspronkelijke probleem.

Dit reconstructieprobleem kent een eenvoudige oplossing wanneer we discrete velden vanuit een nieuwe hoek bekijken. In hoofdstuk 5 tonen we aan dat discrete velden die waarden aannemen in de Liegroep $\mathcal{G}$ opgevat kunnen worden als *discrete $\mathcal{G}$-connecties* (definitie 1.5), een concept gekend vanuit de studie van roosterijktheorieën[1] (zie [**4**, **51**]) en onafhankelijk daarvan bestudeerd door Novikov [**88**].

Stelling 5 (Reconstructie)**.** *Zij $\varphi : E \to \mathcal{G}$ een oplossing van de Euler-Poincarévergelijkingen. Dan bestaat er een oplossing $\phi : V \to \mathcal{G}$ van de oorspronkelijke veldvergelijkingen als en slechts als φ, opgevat als discrete $\mathcal{G}$-connectie, vlak is. In dat geval is ϕ uniek op de keuze van een element van $\mathcal{G}$ na.*

We eindigen hoofdstuk 5 met een uitbreiding van de vergelijkingen van Moser en Veselov (zie [**85**]). Deze nieuwe discrete vergelijkingen beschrijven de dynamica van een discrete harmonische afbeelding met waarden in een half-enkelvoudige Liegroep $\mathcal{G}$.

[1] Net zoals een ijkveld voor de ijkgroep $\mathcal{G}$ niets anders is dan een connectie op een hoofdvezelbundel met structuurgroep $\mathcal{G}$, is een ijkveld op een rooster op natuurlijke wijze een discrete $\mathcal{G}$-connectie.

Klassieke veldentheorieën met niet-holonome bindingen

Na de behandeling van discrete veldentheorieën met waarden in een Liegroepoïde komen we dan tot het tweede deel van het proefschrift, waarin klassieke veldentheorieën bestudeerd worden die onderhevig zijn aan *niet-holonome bindingen.*

Situering. De studie van mechanische systemen met niet-holonome bindingen gaat terug op het werk van Hertz [**54**] en zijn tijdgenoten (zie ook het overzichtsartikel [**90**] van Poincaré). Na het werk van deze pioniers leefde de studie van niet-holonome systemen vooral in de Sovjetunie verder (zie [**87**]). Met de hernieuwde belangstelling voor mechanica vanuit differentiaalmeetkundige hoek sinds de jaren 80 van de vorige eeuw kende de theorie van niet-holonome systemen echter een sterke opleving, die tot op vandaag voortduurt en waarvan de recente werken [**12**, **26**] getuigen.

Voor klassieke veldentheorieën is de situatie enigszins anders. Vanuit wiskundig standpunt kan men een niet-holonome binding eenvoudig definiëren als een binding die afhangt van de afgeleiden van de velden, en die niet integreerbaar is. Een dergelijke binding kan meetkundig voorgesteld worden als de specificatie van een deelvariëteit $\mathcal{C}$ van de jetbundel $J^1\pi$. Dit was het vertrekpunt voor een aantal theoretische studies (zie [**10**, **66**]) waarin een differentiaalmeetkundig kader geschetst wordt voor klassieke veldentheorieën met niet-holonome bindingen. Ondanks deze elegante beschrijvingen ontbrak een overtuigend voorbeeld van een dergelijke veldentheorie vooralsnog.

In dit proefschrift wordt allereerst de theoretische beschrijving van niet-holonome veldentheorieën verder uitgediept. In de latere hoofdstukken wordt dan een fysisch voorbeeld geconstrueerd van een continuum met een niet-holonome binding, de zogenaamde niet-holonome Cosseratstaaf. Deze ontwikkelingen worden hierna besproken.

Geometrische behandeling. In hoofstuk 6 vatten we de bespreking van veldentheorieën met niet-holonome bindingen aan, uitgaande van de volgende gegevens:

(1) een reguliere Lagrangiaan $L : J^1\pi \to \mathbb{R}$;
(2) een *bindingsoppervlak* $\mathcal{C}$, een deelvariëteit van $J^1\pi$ zodanig dat de restrictie van $\pi_{1,0} : J^1\pi \to Y$ tot $\mathcal{C}$ een deelbundel van $J^1\pi$ bepaalt;
(3) een bundel van $(n+1)$-vormen F, gedefinieerd langs $\mathcal{C}$, die we de *bundel van reactiekrachten* noemen. Hierbij dienen de elementen van F aan twee eigenschappen te voldoen:
    (a) $\Phi \in F$ is *n-horizontaal*, in de zin dat de contractie van Φ met elke twee π_1-verticale vectoren nul geeft;
    (b) Φ is 1-*contact*: voor elke sectie ϕ van π geldt dat $(j^1\phi)^*\Phi = 0$.

Daarnaast nemen we aan dat de rang van F gelijk is aan de codimensie van $\mathcal{C}$, die we met k noteren. Lokaal wordt een bindingsoppervlak gegeven door het nul worden van

k functies φ^α op $J^1\pi$: een 1-jet $j^1_x\phi$ is dus een element van $\mathcal{C}$ als en slechts als lokaal geldt dat

$$\varphi^\alpha\left(x^\mu, \phi^a(x), \frac{\partial \phi^a}{\partial x^\mu}(x)\right) = 0.$$

Let op het optreden van de afgeleiden van de velden ϕ^a met betrekking tot de variabelen x^μ op de basisvariëteit (de onafhankelijke veranderlijken).

De bundel van reactiekrachten F behoeft wat meer uitleg. In het kader van veldentheorie kan een kracht gemodelleerd worden als een $(n+1)$-vorm Φ op $J^1\pi$. Een variatie van een veld ϕ kan immers gezien worden als een verticaal vectorveld V gedefinieerd langs het beeld van $j^1\phi$, en ruwweg gesproken bekomt men dan de globale arbeid op tijdstip t door de n-vorm $V \lrcorner \Phi$ te integreren langs het beeld van $j^1\phi$, gerestringeerd tot het hyperoppervlak van constante tijd t in X. Om een zinvol onderscheid te kunnen maken tussen tijdachtige en ruimtelijke variabelen op X is het Cauchy formalisme (paragraaf 3 in hoofdstuk 1) vereist.

De niet-holonome projector. De bijzondere voorwaarden die aan F opgelegd worden, impliceren het bestaan van een k-dimensionale distributie D op $J^1\pi$ (gedefinieerd langsheen $\mathcal{C}$), die als volgt gedefinieerd wordt:

$$X \in D \quad \text{als en slechts als} \quad i_X\Omega_L \in F,$$

waarbij Ω_L de multisymplectische vorm geassocieerd aan L voorstelt. In tegenstelling tot in het symplectisch geval is het over het algemeen niet mogelijk om voor een gegeven $(n+1)$-vorm Φ een vectorveld X te vinden zodat $X \lrcorner \Omega_L = \Phi$. Omwille van de n-horizontaliteit en de 1-contacteigenschap is dit echter wel het geval voor elementen van F.

Onder bepaalde voorwaarden kunnen we dan de raakbundel aan $J^1\pi$ in punten van $\mathcal{C}$ schrijven als de volgende directe som:

$$T_\gamma J^1\pi = D(\gamma) \oplus T_\gamma\mathcal{C} \quad \text{voor alle } \gamma \in \mathcal{C}.$$

Noteren we nu met $\mathcal{P}$ de projectie van $TJ^1\pi$ op $T\mathcal{C}$, en zij $\mathbf{h}$ de horizontale projector van een connectie op π_1, dan kunnen we aantonen dat de samengestelde afbeelding $\mathcal{P}\circ\mathbf{h}$ opnieuw een connectie bepaalt, maar ditmaal op $(\pi_1)_{|\mathcal{C}}$. Meerbepaald bekomen we zo de volgende hulpstelling:

Hulpstelling 6. *De samenstelling $\mathcal{P} \circ \mathbf{h}_{|T_\mathcal{C}J^1\pi} : T_\mathcal{C}J^1\pi \to T\mathcal{C}\,(\subset T_\mathcal{C}J^1\pi), v \mapsto \mathcal{P}(\mathbf{h}(v))$ is een projectie-afbeelding, waarvan de restrictie $\mathbf{h}_\mathcal{P}$ tot $T\mathcal{C}$ een connectie op $(\pi_1)_{|\mathcal{C}} : \mathcal{C} \to X$ induceert. Bovendien is deze connectie semi-holonoom als* $\mathbf{h}$ *semi-holonoom is.*

De niet-holonome De Donder-Weylvergelijking. Het uiteindelijk hoofdresultaat kan dan als volgt geformuleerd worden. Zij $\mathbf{h}$ een oplossing van de vrije De Donder-Weylvergelijking (1.17): dan is $\mathcal{P} \circ \mathbf{h}$ een oplossing van de De Donder-Weylvergelijking met bindingen gespecificeerd door $\mathcal{C}$ en F.

Stelling 7. *Beschouw een niet-holonome veldentheorie gespecificeerd door een Lagrangiaan L, een bindingsoppervlak $\mathcal{C}$ en een bundel van reactiekrachten F. Neem aan dat de compatibiliteitsvoorwaarde (6.10) geldt. Zij $\mathbf{h}$ een oplossing van de vrije De Donder-Weylvergelijking (1.17) en zij $\mathcal{P}$ de niet-holonome projectie-afbeelding.*

Dan bepaalt de projectie $\mathcal{P} \circ \mathbf{h}$ een oplossing van de Donder-Weylvergelijking (6.9) met bindingen. Daarenboven is $\mathcal{P} \circ \mathbf{h}$ de horizontale projector van een semi-holonome connection op $(\pi_1)_{|\mathcal{C}} : \mathcal{C} \to X$.

Het Cauchy-formalisme. We besluiten hoofdstuk 6 met een blik op het Cauchy-formalisme. In het geval er geen bindingen aanwezig zijn, induceert een oplossing van de vrije De Donder-Weylvergelijking een tweede-orde vectorveld op de ruimte van Cauchy-data dat voldoet aan de bewegingsvergelijkingen voor een tijdsafhankelijk mechanisch systeem. Dit is een klassiek resultaat uit de Cauchy-theorie, dat geometrisch behandeld werd in [**11**, **49**, **91**] en hier uitgebreid wordt naar het geval waar er niet-holonome bindingen aanwezig zijn.

Allereerst tonen we aan dat $\mathcal{C}$ en F corresponderende oneindigdimensionale objecten $\tilde{\mathcal{C}}$ en $\tilde{F}$ induceren op de ruimte van Cauchy-data. Daarna stellen we de vergelijkingen op voor een mechanisch systeem op de ruimte van Cauchy-data met niet-holonome bindingen gespecificeerd door $\tilde{\mathcal{C}}$ en $\tilde{F}$, en bewijzen we dat een oplossing van het niet-holonoom De Donder-Weylprobleem een tweede-orde vectorveld induceert dat een oplossing is van deze vergelijkingen.

Symmetrie en het niet-holonome momentlemma. Zij $\mathcal{G}$ een Liegroep die op de totale variëteit (de ruimte van afhankelijke veranderlijken) werkt en waarvan de actie de Lagrangiaan invariant laat (zodat $\mathcal{G}$ een symmetriegroep is). De gekende stelling van Noether leert dan dat er D behoudswetten bestaan geassocieerd aan die symmetrie-actie, waarbij D de dimensie is van $\mathcal{G}$. Deze behoudswetten kunnen compact neergeschreven worden met behulp van de zogenaamde *momentafbeelding* (Eng. *momentum map*), een $\mathfrak{g}^*$-waardige n-vorm J op $J^1\pi$ gedefinieerd als

$$\langle J, \xi \rangle := J_\xi, \quad \text{waarbij} \quad J_\xi := \xi_{J^1\pi} \lrcorner\, \Theta_L,$$

waarbij $\xi_{J^1\pi}$ de infinitesimale generator is, geassocieerd aan een element ξ van $\mathfrak{g}$. Voor elke oplossing ϕ van de Euler-Lagrangevergelijkingen geldt dan de volgende behoudswet:

$$\mathrm{d}(j^1\phi)^* J_\xi = 0 \quad \text{voor alle} \quad \xi \in \mathfrak{g}.$$

Wanneer een veldentheorie onderworpen is aan niet-holonome bindingen, is de stelling van Noether niet langer geldig. In plaats daarvan heeft men voor niet-holonome mechanische systemen het zogenaamde *niet-holonome momentlemma* (Eng. *momentum lemma*) dat de evolutie beschrijft van de behouden grootheden onder de niet-holonome stroming. Dit lemma werd bewezen in [**13**, **20**] en de veralgemening ervan naar de context van veldentheorie vormt het onderwerp van hoofdstuk 7.

Nemen we aan dat $\mathcal{G}$ een Liegroep is die de data van het niet-holonome probleem (L, $\mathcal{C}$, en F) invariant laat, dan geldt de volgende stelling.

Stelling 8 (Niet-holonome momentlemma)**.** *Zij* $\mathbf{h}$ *de horizontale projector van een connectie op* π_1 *die de oplossing is van de niet-holonome De Donder-Weylvergelijking. Dan voldoet de niet-holonome momentafbeelding aan de volgende vergelijking:*

$$\mathrm{d}_{\mathbf{h}} J_{\xi}^{\mathrm{n.h.}} = \mathscr{L}_{\tilde{\xi}}(L\eta) \quad \textit{langs } \mathcal{C}. \tag{2.10}$$

Hierbij is de niet-holonome momentafbeelding $J_{\xi}^{\mathrm{n.h.}}$ de restrictie van de gewone momentafbeelding tot de vectoren ξ in $\mathfrak{g}$ waarvan de contractie van de infinitesimale generator $\tilde{\xi}$ met de elementen van F nul oplevert. Indien het rechterlid van (2.10) nul wordt, hebben we te maken met een echte behoudswet.

Bij overgang naar Cauchy-formalisme induceert de niet-holonome momentafbeelding een momentafbeelding (in de zin van symplectische meetkunde) die voldoet aan een niet-holonoom momentlemma (in de zin van [**13**]) op de ruimte van Cauchy data. Deze geïnduceerde momentafbeelding is in feite niets anders dan de vorm $J_{\xi}^{\mathrm{n.h.}}$, geïntegreerd over een hyperoppervlak van constante tijd.

Niet-covariante niet-holonome bindingen

In het laatste deel van dit proefschrift construeren we dan een voorbeeld van een niet-holonome veldentheorie. In hoofdstuk 6 hadden we reeds aangetoond dat het niet-holonome formalisme de correcte veldvergelijkingen oplevert voor de dynamica van een onsamendrukbare vloeistof, alhoewel men op grond van de traditionele aanpak in feite zou verwachten dat dat niet zo is.

Op dit thema komen we terug in hoofdstuk 8, waar we een klasse van bindingen identificeren die op het eerste zicht niet-holonoom zijn, maar toch behandeld moeten worden met het vakonoom formalisme (zie hieronder). Het fundamentele inzicht is dat deze bindingen niet afhangen van de tijdsafgeleiden van de velden, en indien we dan de overgang maken naar het Cauchy-formalisme blijkt dat deze bindingen holonome bindingen induceren op de ruimte van Cauchy-data. Het Cauchy-formalisme is dus van wezenlijk belang in de classificatie van mogelijke bindingen.

Deze opmerking zal van wezenlijk belang blijken te zijn bij onze behandeling van niet-covariante niet-holonome bindingen in hoofdstuk 9.

Het Skinner-Ruskformalisme. In het tweede deel van hoofdstuk 8 beschouwen we dan het Skinner-Ruskformalisme voor veldentheorieën met bindingen. Het voordeel van dit formalisme is dat het een eenvoudige differentiaalmeetkundige vergelijking toelaat tussen twee verschillende, inequivalente modellen voor de dynamica, zijnde:

(1) de niet-holonome methode van hoofdstuk 6;
(2) de *vakonome* methode, waarbij de veldvergelijkingen afgeleid worden door de actiefunctionaal te beperken tot de ruimte van velden die voldoen aan de bindingen.

De aanpak van Skinner en Rusk maakt gebruik van de productbundel $J^1\pi \times \bigwedge_2^{n+1} Y$, waarbij $\bigwedge_2^{n+1} Y$ de bundel van n-horizontale $(n+1)$-vormen (zie boven) op Y is. Op deze bundel kan men een De Donder-Weylvergelijking formuleren met behulp van een zekere pre-multisymplectische vorm. Het oplossen van deze vergelijking komt dan neer op het toepassen van een algoritme zoals dat van Gotay (zie [**32**, **43**]).

Ook de vakonome en de niet-holonome methode kunnen in dit formalisme ingepast worden. In hoofdstuk 8 beschouwen we enkel het geval waarbij de bindingen opgevat kunnen worden als de horizontale deelruimte van een Ehresmann-connectie op een bepaalde fibratie. Uiteindelijk tonen we dan aan dat de vakonome en de niet-holonome dynamica equivalent zijn als en slechts als de kromming van deze connectie verdwijnt. Dit is op zijn beurt equivalent met het integreerbaar zijn van de bindingen.

De niet-holonome Cosseratstaaf. In het laatste hoofdstuk komen alle voorgaande thema's samen bij de constructie van een fysisch voorbeeld van een veldentheorie met niet-holonome bindingen.

Dit voorbeeld dient gesitueerd te worden in de theorie van de Cosserat-media: we beschouwen een Cosseratstaaf, wat erop neerkomt dat we de dynamica van een lange elastische staaf bestuderen door benaderend aan te nemen dat de transversale beweging van de staaf *Euclidisch* is. De staaf kan dus met andere woorden enkel buigen of torsioneel vervormd worden en de dynamica is volledig vastgelegd door het specificeren van de middellijn en van een basis van vectorvelden langs die middellijn (zie figuur 9.1 op pagina 126).

De niet-holonome bindingen duiken op wanneer we zo'n staaf op een vast horizontaal vlak laten bewegen dat voldoende ruw is, zodat de beweging van de staaf van het type "rollen zonder glijden" is. Om dit probleem te analyseren, is het aangewezen om de Cosseratstaaf te interpreteren als een continuum-versie van de verticaal rollende schijf, een standaardvoorbeeld van een niet-holonoom mechanisch systeem (zie [**13**]). Op basis van dit inzicht tonen we aan dat de bundel van reactiekrachten F voor dit systeem ook *niet-covariant* is: F kan uit $\mathcal{C}$ afgeleid worden door te stellen dat

$$F = S^*_{\text{n.c.}}(T^\circ\mathcal{C}),$$

waarbij $S_{\text{n.c.}}$ het zogenaamde niet-covariante verticaal endomorfisme is, gedefinieerd in paragraaf 2.1. Deze vectorwaardige n-vorm heeft de volgende coördinaatgedaante:

$$S_{\text{n.c.}} = (\mathrm{d}y^a - y^a_\mu \mathrm{d}x^\mu) \wedge \mathrm{d}^n x_0 \otimes \frac{\partial}{\partial y^a_0}.$$

Wanneer we deze uitdrukking vergelijkingen met het gewone verticaal endormorfisme (1.2), zien we dat de tijdsafgeleiden hier een bevoorrechte rol spelen, wat het adjectief "niet-covariant" verklaart. In het licht van de ontwikkelingen in hoofdstuk 8 kunnen we aantonen dat deze niet-holonome bindingen ook bij overgang naar het Cauchy-formalisme niet-holonoom blijven.

Veldvergelijkingen. De veldvergelijkingen voor een dergelijke niet-holonome veldentheorie kunnen bekomen worden door middel van het formalisme in hoofdstuk 6. In dit hoofdstuk nemen we echter een alternatieve route en leiden we de veldvergelijkingen direct af door de actie te variëren met betrekking tot *toelaatbare variaties* (variaties die aan de bindingen voldoen). Zo bekomen we de volgende stelling:

Stelling 9. *Zij ϕ een sectie van π. Indien $\operatorname{Im} j^1\phi \subset \mathcal{C}$, dan zijn de volgende beweringen equivalent:*

(a) *ϕ is een extremum van de actie (1.14) onder toelaatbare variaties;*
(b) *ϕ is een oplossing van de niet-holonome Euler-Lagrangevergelijkingen:*

$$\left[\frac{\partial L}{\partial y^a} - \frac{\mathrm{d}}{\mathrm{d}x^\mu}\frac{\partial L}{\partial y^a_\mu}\right](j^2\phi) = \lambda_\alpha A^\alpha_a(j^1\phi) \quad \text{and} \quad \varphi^\alpha(j^1\phi) = 0.$$

(c) *voor alle vectorvelden W op $J^1\pi$ zodat $(j^1\phi)^*(W \lrcorner \Phi) = 0$ voor alle $\Phi \in F$ geldt*

$$(j^1\phi)^*(W \lrcorner \Omega_L) = 0.$$

Voor het geval van de niet-holonome staaf zijn de veldvergelijkingen gegeven door:

$$\begin{cases} \rho\ddot{x} + Kx'''' &= \lambda \\ \rho\ddot{y} + Ky'''' &= \mu \\ \alpha\ddot{\theta} - \beta\theta'' &= R(\lambda y' - \mu x'), \end{cases}$$

waarbij λ en μ Lagrange-multiplicatoren zijn, geassocieerd aan de niet-holonome bindingen:

$$\dot{x} + R\dot{\theta}y' = 0 \quad \text{en} \quad \dot{y} - R\dot{\theta}x' = 0.$$

Symmetrie. De niet-holonome Cosseratstaaf is invariant onder de actie van een aantal symmetriegroepen. Allereerst is er de groep van translaties in de tijd, waarvan de symmetriegenerator het vectorveld $\frac{\partial}{\partial t}$ is. Dit is echter geen verticaal vectorveld en dus is het niet-holonome momentlemma van hoofdstuk 7 niet toepasbaar. Aan de hand van de bewegingsvergelijkingen kunnen we echter direct aantonen dat er met deze symmetrie toch een behoudswet geassocieerd is, waaruit we na integratie kunnen afleiden dat de globale energie behouden is, zoals verwacht.

In het geval dat de Cosseratstaaf niet onderworpen is aan niet-holonome bindingen, kan men gemakkelijk aantonen dat de dynamica eveneens invariant is onder de actie van de Euclidische groep $SE(2)$. De behoudswetten geassocieerd met de actie van dit semidirect product stemmen (na integratie) overeen met behoud van impuls- en draaimoment. Wanneer de niet-holonome bindingen in rekening gebracht worden, blijkt dat enkel de $\mathbb{R}^2$-deelgroep van $SE(2)$ corresponderend met translaties een niet-holonome symmetrie-actie bepaalt. Op deze actie is het niet-holonome momentlemma wel van toepassing en uiteindelijk bekomen we zo de volgende behoudswet:

$$Ry'(\rho\ddot{x} + Kx'''') - Rx'(\rho\ddot{y} + Ky'''') = \alpha\ddot{\theta} - \beta\theta''.$$

Numerieke integratie. Aangezien de veldvergelijkingen van de niet-holonome Cosseratstaaf naar alle waarschijnlijkheid niet integreerbaar zijn, construeren we in paragraaf 4 van hoofdstuk 9 een numeriek integratieschema voor de dynamica. Hiertoe gebruiken we de inzichten uit hoofdstukken 2 en 3 om een discrete versie van het principe van d'Alembert op te stellen.

Voor de Cosseratstaaf leidt dit principe tot een expliciete, tweede-orde geometrische integrator die de bindingen exact bewaart. Naast het behoud van de bindingen vertoont deze integrator ook andere interessante eigenschappen: net zoals bij symplectische integratoren is de energie weliswaar niet exact behouden, maar is de numerieke fout op de energie daarentegen wel begrensd.

Bibliography

[1] W. F. Ames, *Numerical methods for partial differential equations*, Academic Press [Harcourt Brace Jovanovich Publishers], 2nd ed. 1977.

[2] S. S. Antman, *Nonlinear Problems of Elasticity*, Applied Mathematical Sciences, vol. 107, Springer-Verlag, 2005.

[3] U. M. Ascher and R. I. McLachlan, *On symplectic and multisymplectic schemes for the KdV equation*, J. Sci. Comput. **25** (2005), no. 1, 83–104.

[4] J. Baez, *Quantum Gravity seminar: Gauge Theory and Topology (Week 5, Winter 2005)*, Course notes, available from the author's website `http://math.ucr.edu/home/baez/qg-winter2005`.

[5] E. Barth, B. Leimkuhler, and S. Reich, *A time-reversible variable-stepsize integrator for constrained dynamics*, SIAM J. Sci. Comput. **21** (1999), no. 3, 1027–1044.

[6] L. Bates and J. Śniatycki, *Nonholonomic reduction*, Rep. Math. Phys. **32** (1993), no. 1, 99–115.

[7] K. Behrend and P. Xu, *Differentiable Stacks and Gerbes*, arXiv:math.DG/0605694, 2006.

[8] E. Bibbona, L. Fatibene, and M. Francaviglia, *Chetaev vs. vakonomic prescriptions in constrained field theories with parametrized variational calculus*, preprint, math-ph/0608063.

[9] E. Bibbona, L. Fatibene, and M. Francaviglia, *Gauge-natural parameterized variational problems, vakonomic field theories and relativistic hydrodynamics of a charged fluid*, Int. J. Geom. Meth. Mod. Phys. **3** (2006), 1573–1608.

[10] E. Binz, M. de León, D. Martín de Diego, and D. Socolescu, *Nonholonomic Constraints in Classical Field Theories*, Rep. Math. Phys. **49** (2002), 151–166.

[11] E. Binz, J. Śniatycki, and H. Fischer, *Geometry of classical fields*, North-Holland Mathematics Studies, vol. 154, North-Holland Publishing, Amsterdam, 1988.

[12] A. Bloch, *Nonholonomic mechanics and control*, Interdisciplinary Applied Mathematics, vol. 24, Springer-Verlag, Berlin, 2003.

[13] A. Bloch, P. Krishnaprasad, J. E. Marsden, and R. Murray, *Nonholonomic mechanical systems with symmetry*, Arch. Rat. Mech. Anal. **136** (1996), no. 1, 21–99.

[14] A. M. Bloch, J. E. Marsden, and D. V. Zenkov, *Nonholonomic dynamics*, Notices Amer. Math. Soc. **52** (2005), no. 3, 324–333.

[15] A. I. Bobenko and Y. B. Suris, *Discrete time Lagrangian mechanics on Lie groups, with an application to the Lagrange top*, Comm. Math. Phys. **204** (1999), no. 1, 147–188.

[16] A. I. Bobenko and Y. B. Suris, *Discrete Lagrangian reduction, discrete Euler-Poincaré equations, and semidirect products*, Lett. Math. Phys. **49** (1999), no. 1, 79–93.

[17] T. J. Bridges, *Multi-symplectic structures and wave propagation*, Math. Proc. Cambridge Philos. Soc. **121** (1997), no. 1, 147–190.

[18] T. J. Bridges and S. Reich, *Numerical methods for hamiltonian pdes*, J. Phys. A: Math. Gen. **39** (2006), 5287–5320.

[19] T. J. Bridges and S. Reich, *Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity*, Phys. Lett. A **284** (2001), no. 4-5, 184–193.

[20] F. Cantrijn, M. de León, J. C. Marrero, and D. Martín de Diego, *Reduction of nonholonomic mechanical systems with symmetry*, Rep. Math. Phys. **42** (1998), no. 1-2, 25–45, Pacific Institute

of Mathematical Sciences Workshop on Nonholonomic Constraints in Dynamics (Calgary, AB, 1997).

[21] F. Cantrijn, A. Ibort, and M. de León, *On the geometry of multisymplectic manifolds*, J. Austral. Math. Soc. Ser. A **66** (1999), no. 3, 303–330.

[22] J. F. Cariñena, M. Crampin, and L. A. Ibort, *On the multisymplectic formalism for first order field theories*, Diff. Geom. Appl. **1** (1991), no. 4, 345–374.

[23] M. Castrillón López, P. L. García Pérez, and T. S. Ratiu, *Euler-Poincaré reduction on principal bundles*, Lett. Math. Phys. **58** (2001), no. 2, 167–180.

[24] M. Castrillón López and T. S. Ratiu, *Reduction in principal bundles: covariant Lagrange-Poincaré equations*, Comm. Math. Phys. **236** (2003), no. 2, 223–250.

[25] D. Christodoulou, *The Action Principle and Partial Differential Equations*, Annals of Mathematics Studies, vol. 146, Princeton University Press, Princeton (NJ), 2000.

[26] J. Cortés, *Geometric, control and numerical aspects of nonholonomic systems*, Lecture Notes in Mathematics, vol. 1793, Springer-Verlag, Berlin, 2002.

[27] J. Cortés, M. de León, D. Martín de Diego, and S. Martínez, *Geometric description of vakonomic and nonholonomic dynamics. Comparison of solutions*, SIAM J. Control Optim. **41** (2002), no. 5, 1389–1412.

[28] J. Cortés and S. Martínez, *Non-holonomic integrators*, Nonlinearity **14** (2001), no. 5, 1365–1392.

[29] M. Crainic and R. L. Fernandes, *Integrability of Lie brackets*, Ann. Math. (2) **157** (2003), no. 2, 575–620.

[30] M. Crampin, *Tangent bundle geometry for Lagrangian dynamics*, J. Phys. A: Math. Gen. **16** (1983), no. 16, 3755–3772.

[31] M. de León, J. C. Marrero, and J. Marín-Solano, *A geometrical approach to classical field theories: a constraint algorithm for singular theories*, New Developments in Differential Geometry (Debrecen 1994), Math. Appl., vol. 350, Kluwer Acad. Publ. (Dordrecht), 1996, pp. 291–312.

[32] M. de León, J. C. Marrero, and D. Martín de Diego, *A new geometric setting for classical field theories*, Classical and Quantum Integrability (Warsaw, 2001) (Warsaw), Banach Center Publ., vol. 59, Polish Acad. Sci., 2003, pp. 189–209.

[33] M. de León, D. Martín de Diego, and A. Santamaría-Merino, *Tulczyjew's triples and lagrangian submanifolds in classical field theory*, Applied Differential Geometry and Mechanics (Ghent) (W. Sarlet and F. Cantrijn, eds.), Academia Press, 2003, pp. 21–47.

[34] M. de León, D. Martín de Diego, and A. Santamaría-Merino, *Geometric integrators and nonholonomic mechanics*, J. Math. Phys. **45** (2004), no. 3, 1042–1064.

[35] M. de León, D. Martín de Diego, and A. Santamaría-Merino, *Symmetries in classical field theory*, Int. J. Geom. Meth. Mod. Phys. **1** (2004), no. 5, 651–710.

[36] M. de León, M. McLean, L. Norris, A. Roca, and M. Salgado, *Geometric structures in field theory*, preprint, `math-ph/0208036`.

[37] M. de León and D. M. de Diego, *On the geometry of non-holonomic lagrangian systems*, J. Math. Phys. **37** (1996), no. 7, 3389–3414.

[38] M. de León, J. C. Marrero, and E. Martínez, *Lagrangian submanifolds and dynamics on lie algebroids*, J. Phys. A: Math. Gen. **38** (2005), R241–R308.

[39] M. Desbrun, A. N. Hirani, M. Leok, and J. E. Marsden, *Discrete exterior calculus*, preprint (2005), `math.DG/0508341`.

[40] D. J. Dichmann, Y. Li, and J. H. Maddocks, *Hamiltonian formulations and symmetries in rod mechanics*, Mathematical approaches to biomolecular structure and dynamics (Minneapolis, MN, 1994), IMA Vol. Math. Appl., vol. 82, Springer, New York, 1996, pp. 71–113.

[41] J. L. Dupont, *Curvature and characteristic classes*, Springer-Verlag, Berlin, 1978, Lecture Notes in Mathematics, Vol. 640.

[42] D. G. Ebin and J. E. Marsden, *Groups of diffeomorphisms and the notion of an incompressible fluid*, Ann. Math. (2) **92** (1970), no. 2, 102–163.

[43] A. Echeverría-Enríquez, C. López, J. Marín-Solano, M. C. Muñoz-Lecanda, and N. Román-Roy, *Lagrangian-hamiltonian unified formalism for field theory*, J. Math. Phys. **45** (2004), no. 1, 360–380.

[44] M. Epstein and R. Segev, *Differentiable manifolds and the principle of virtual work in continuum mechanics*, J. Math. Phys. **21** (1980), no. 5, 1243–1245.

[45] P. L. García, A. García, and C. Rodrigo, *Cartan forms for first order constrained variational problems*, J. Geom. Phys. **56** (2006), no. 4, 571–610.

[46] P. L. García, *The Poincaré-Cartan invariant in the calculus of variations*, Simposia Mathematica, Vol. XIV (Convegno di Geometria Simplettica e Fisica Matematica, INDAM, Rome, 1973, Academic Press, London, 1974, pp. 219–246.

[47] H. Goldschmidt and S. Sternberg, *The Hamilton-Cartan formalism in the calculus of variations*, Ann. Inst. Fourier (Grenoble) **23** (1973), no. 1, 203–267.

[48] M. Gotay, J. Isenberg, and J. Marsden, *Momentum Maps and Classical Relativistic Fields. Part I: Covariant Field Theory*, preprint, `physics/9801019`.

[49] M. Gotay, J. Isenberg, and J. Marsden, *Momentum Maps and Classical Relativistic Fields. Part II: Canonical Analysis of Field Theories*, preprint, `math-ph/0411032`.

[50] C. Günther, *The polysymplectic Hamiltonian formalism in field theory and calculus of variations. I. The local case*, J. Diff. Geom. **25** (1987), no. 1, 23–53.

[51] R. Gupta, *Introduction to lattice QCD*, Course notes, available as hep-lat/9807028.

[52] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration*, 1st ed., Springer Series in Computational Mathematics, vol. 31, Springer-Verlag, Berlin, 2002.

[53] A. Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002, available electronically from `http://www.math.cornell.edu/∼hatcher/AT/ATpage.html`.

[54] H. Hertz, *The principles of mechanics*, Dover Publications Inc., New York, 1956.

[55] P. Higgins and K. Mackenzie, *Algebraic constructions in the category of Lie algebroids*, J. Algebra **129** (1990), no. 1, 194–230.

[56] A. N. Hirani, *Discrete Exterior Calculus*, Ph.D. thesis, California Institute of Technology, 2003, available electronically from `http://www.cs.uiuc.edu/homes/hirani`.

[57] D. Iglesias, J. C. Marrero, D. Martín de Diego, and E. Martínez, *Discrete nonholonomic lagrangian systems on lie groupoids*, work in progress.

[58] J. Kijowski, *A finite-dimensional canonical formalism in the classical field theory*, Comm. Math. Phys. **30** (1973), 99–128.

[59] S. Kobayashi and K. Nomizu, *Foundations of differential geometry. Vol I*, Interscience Publishers, John Wiley & Sons, 1963.

[60] I. Kolář, P. Michor, and J. Slovák, *Natural operations in differential geometry*, Springer-Verlag, Berlin, 1993.

[61] S. Kouranbaeva and S. Shkoller, *A variational approach to second-order multisymplectic field theory*, J. Geom. Phys. **35** (2000), no. 4, 333–366.

[62] V. V. Kozlov, *The dynamics of systems with nonintegrable constraints. I*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. (1982), no. 3, 92–100.

[63] A. Kriegl and P. Michor, *The Convenient Setting of Global Analysis*, Mathematical Surveys and Monographs, vol. 53, AMS, Providence (RI), 1997.

[64] D. Krupka, *Lepagean forms in higher order variational theory*, Proceedings of the IUTAM-ISIMM symposium on modern developments in analytical mechanics, Vol. I (Torino, 1982), Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur., vol. 117, 1983, pp. 197–238.

[65] O. Krupková, *The Geometry of Ordinary Variational Equations*, Lecture Notes in Mathematics, vol. 1678, Springer-Verlag, 1997.

[66] O. Krupková, *Partial differential equations with differential constraints*, J. Diff. Eq. **220** (2005), no. 2, 354–395.

[67] J. Langer and D. A. Singer, *Lagrangian aspects of the Kirchhoff elastic rod*, SIAM Review **38** (1996), no. 4, 605–618.

[68] B. Leimkuhler and S. Reich, *Simulating Hamiltonian dynamics*, Cambridge Monographs on Applied and Computational Mathematics, vol. 14, Cambridge University Press, Cambridge, 2004.

[69] A. Lew, J. E. Marsden, M. Ortiz, and M. West, *Asynchronous variational integrators*, Arch. Rat. Mech. Anal. **167** (2003), no. 2, 85–146.

[70] A. Lew, J. E. Marsden, M. Ortiz, and M. West, *Variational time integrators*, Int. J. Num. Meth. Eng. **60** (2004), no. 1, 153–212.

[71] A. D. Lewis and R. Murray, *Variational principles for constrained systems: theory and experiment*, Internat. J. Non-Linear. Mech. **30** (1995), no. 6, 793–815.

[72] K. C. H. Mackenzie, *General theory of Lie groupoids and Lie algebroids*, London Mathematical Society Lecture Note Series, vol. 213, Cambridge University Press, Cambridge, 2005.

[73] L. Mangiarotti and G. Sardanashvily, *Connections in classical and quantum field theory*, World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[74] C. M. Marle, *Various approaches to conservative and nonconservative nonholonomic systems*, Rep. Math. Phys. **42** (1998), no. 1-2, 211–229, Pacific Institute of Mathematical Sciences Workshop on Nonholonomic Constraints in Dynamics (Calgary, AB, 1997).

[75] J. C. Marrero, D. Martín de Diego, and E. Martínez, *Discrete Lagrangian and Hamiltonian Mechanics on Lie Groupoids*, Nonlinearity **19** (2006), no. 6, 1313–1348, `math.DG/0506299`.

[76] J. E. Marsden and T. J. Hughes, *Mathematical foundations of elasticity*, Dover Publications Inc., New York, 1994, Corrected reprint of the 1983 original.

[77] J. E. Marsden, S. Pekarsky, and S. Shkoller, *Discrete Euler-Poincaré and Lie-Poisson equations*, Nonlinearity **12** (1999), no. 6, 1647–1662.

[78] J. E. Marsden and S. Shkoller, *Multisymplectic geometry, covariant Hamiltonians and water waves*, Math. Proc. Camb. Phil. Soc. **125** (1999), 553–575.

[79] J. E. Marsden and M. West, *Discrete mechanics and variational integrators*, Acta Numerica **10** (2001), 357–514.

[80] J. E. Marsden, G. W. Patrick, and S. Shkoller, *Multisymplectic geometry, variational integrators, and nonlinear PDEs*, Comm. Math. Phys. **199** (1998), no. 2, 351–395.

[81] J. E. Marsden, S. Pekarsky, S. Shkoller, and M. West, *Variational methods, multisymplectic geometry and continuum mechanics*, J. Geom. Phys. **38** (2001), no. 3-4, 253–284.

[82] J. E. Marsden and T. S. Ratiu, *Introduction to mechanics and symmetry*, Texts in Applied Mathematics, vol. 17, Springer-Verlag, New York, 1994.

[83] E. Martínez, *Lagrangian mechanics on Lie algebroids*, Acta Appl. Math. **67** (2001), no. 3, 295–320.

[84] R. I. McLachlan and M. Perlmutter, *Energy drift in reversible time integration*, J. Phys. A: Math. Gen. **37** (2004), no. 45, L593–L598.

[85] J. Moser and A. Veselov, *Discrete versions of some classical integrable systems and factorization of matrix polynomials*, Comm. Math. Phys. **139** (1991), no. 2, 217–243.

[86] F. Munteanu, A. M. Rey, and M. Salgado, *The Günther's formalism in classical field theory: momentum map and reduction*, J. Math. Phys. **45** (2004), no. 5, 1730–1751.

[87] J. Neimark and N. Fufaev, *Dynamics of nonholonomic systems*, Translations of Mathematical Monographs, vol. 33, American Mathematical Society, Providence, R.I, 1972.

[88] S. P. Novikov, *Discrete connections and linear difference equations*, Tr. Mat. Inst. Steklova **247** (2004), 186–201, `math-ph/0303035`.

[89] P. J. Olver, *Applications of Lie Groups to Differential Equations*, Graduate Texts in Mathematics, vol. 107, Springer-Verlag, New York, 1986.

[90] H. Poincaré, *Les idées de Hertz sur la mécanique*, Revue Générale des Sciences **XII** (1897), 734–743.

[91] A. Santamaría-Merino, *Métodos Geométricos en Teorías Clásicas de Campos e Integración Numerica*, Ph.D. thesis, Universidad Carlos III de Madrid, 2005.

[92] D. J. Saunders, *A note on Legendre transformations*, Diff. Geom. Appl. **1** (1991), no. 2, 109–122.

[93] D. J. Saunders, F. Cantrijn, and W. Sarlet, *Regularity aspects and Hamiltonisation of nonholonomic systems*, J. Phys. A: Math. Gen. **32** (1999), no. 39, 6869–6890.

[94] D. J. Saunders, *The Geometry of Jet Bundles*, London Mathematical Society Lecture Note Series, vol. 142, Cambridge University Press, 1989.

[95] D. Saunders, *Prolongations of Lie groupoids and Lie algebroids*, Houston J. Math. **30** (2004), no. 3, 637–655.

[96] J. C. Simo, J. E. Marsden, and P. S. Krishnaprasad, *The Hamiltonian structure of nonlinear elasticity: the material and convective representations of solids, rods, and plates*, Arch. Rat. Mech. Anal. **104** (1988), no. 2, 125–183.

[97] R. Skinner and R. Rusk, *Generalized Hamiltonian dynamics. I. Formulation on $T^*Q \oplus TQ$.*, J. Math. Phys. **24** (1983), no. 11, 2589–2594.

[98] R. Skinner and R. Rusk, *Generalized Hamiltonian dynamics. II. Gauge transformations*, J. Math. Phys. **24** (1983), no. 11, 2595–2601.

[99] G. Terra and M. H. Kobayashi, *On classical mechanical systems with non-linear constraints*, J. Geom. Phys. **49** (2004), no. 3-4, 385–417.

[100] C. Truesdell and W. Noll, *The non-linear field theories of mechanics*, Handbuch der Physik, Band III/3 (S. Flügge, ed.), Springer-Verlag, 1965.

[101] W. T. Van Est, *Rapport sur les S-atlas (Structure transverse des feuilletages)*, Astérisque **116** (1984), 235–292.

[102] J. Vankerschaver, F. Cantrijn, M. de León, and D. Martín de Diego, *Geometric aspects of nonholonomic field theories*, Rep. Math. Phys. **56** (2005), no. 3, 387–411, `math-ph/0506010`.

[103] J. Vankerschaver, *The momentum map for nonholonomic field theories with symmetry*, Int. J. Geom. Meth. Mod. Phys. **2** (2005), no. 6, 1029–1041, `math-ph/0507059`.

[104] J. Vankerschaver, *Classical field theories with external constraints*, Proceedings of the XIV Fall Workshop on Geometry and Physics, Bilbao (2005), Sept. 14-16, Publ. de la RSME, vol. 10, 2006, pp. 285–291.

[105] J. Vankerschaver, *Euler-Poincaré reduction for discrete field theories*, Accepted for publication in J. Math. Phys., 2006.

[106] J. Vankerschaver, *A class of nonholonomic kinematic constraints in elasticity*, Journal of Physics A: Mathematical and Theoretical **40** (2007), no. 14, 3889–3913.

[107] J. Vankerschaver and F. Cantrijn, *The Skinner-Rusk approach for vakonomic and nonholonomic field theories*, In preparation (2006).

[108] J. Vankerschaver and F. Cantrijn, *Discrete lagrangian field theories on Lie groupoids*, J. Geom. Phys. **57** (2007), no. 2, 665–689, `math-ph/0511080`.

[109] S. Vignolo and D. Bruno, *Iper-ideal constraints in continuum mechanics*, J. Math. Phys. **43** (2002), no. 1, 325–343.

[110] A. Weinstein, *Lagrangian mechanics and groupoids*, Mechanics day (Waterloo, ON, 1992), Fields Inst. Commun., vol. 7, Amer. Math. Soc., Providence, RI, 1996, pp. 207–231.

[111] H. Whitney, *Geometric integration theory*, Princeton University Press, Princeton, N. J., 1957.

[112] S. O. Wilson, *Geometric structures on the cochains of a manifold*, preprint (2005), `math.GT/0505227`.

[113] D. K. Wise, *p-form electromagnetism on discrete spacetimes*, Class. Quantum Grav. **23** (2006), no. 17, 5129–5176.

[114] J. C. Wood, *Harmonic maps into symmetric spaces and integrable systems*, Harmonic maps and integrable systems (Braunschweig), Aspects Math., E23, Vieweg, 1994, pp. 29–55.