NTNU
Norwegian University of
Science and Technology

Discrete Invariant Variational Problems

Geir Bogfjellmo

Master of Science in Physics and Mathematics
Submission date: June 2011
Supervisor: Brynjulf Owren, MATH

Norwegian University of Science and Technology
Department of Mathematical Sciences

Problem Description

Study how moving frames can be used to develop numerical methods for variational problems such that the numerical method inherit symmetries from the continuous problem.

Assignment given:	24 January 2011.
Supervisor:	Brynjulf Owren

ABSTRACT

This thesis studies variational problems invariant under a Lie group transformation, and invariant discretizations of these. In chapters two and three, a general method for creating symplectic integrators preserving certain classes of variational symmetries of first order Lagrangians is developed and demonstrated. In chapters four and five, it is assumed that the discrete Lagrangian is invariant under a certain group action, and the Euler–Lagrange equations for the variational problem are expressed in the invariants of the group action.

PREFACE

This thesis was written for the course TMA4910 – Numerical Mathematics, Master Thesis in the spring term of 2011. It is the final part of my Master's degree in Applied Physics and Mathematics, with Specialization Industrial Mathematics.

The thesis is to a large degree a continuation of my specialization project on symmetry-preserving numerical methods. After studying the techniques for invariantizing numerical schemes, it seemed natural to continue on to variational problems, since variational symmetries are closely connected to conservation laws.

The writing of the thesis would not have been possible without the support of several people. I would like to thank my supervisor Brynjulf Owren, for encouragement, guidance and proof reading. I would further like to thank Elena Celledoni, Tore Halvorsen and Olivier Verdier for additional help and proof reading.

Finally I would like to thank all my friends in Matteland, for countless coffee-, lunch-, dinner- and football breaks, occasionally at bizarre hours.

Geir Bogfjellmo, June 27, 2011

CONTENTS

1 Introduction **1**

2 Group Actions and Moving Frames **3**

2.1 Basic Definitions and Results 3
2.2 Infinitesimals 6
2.3 Prolongation 7
2.4 Differentiation of Invariants 8

3 Mechanical Systems **13**

3.1 Symplectic Manifolds and Symplectic Maps 13
3.2 Lagrangian and Hamiltonian Systems 17
3.3 Symmetries 18
3.4 Discrete Mechanics 20
3.5 Noether's First Theorem 22
3.6 Invariant Discrete Lagrangians 23
3.7 Reparametrization 24

4 The Runge–Lenz Vector **27**

4.1 The Kepler Problem 27
4.2 Stereographic Projection of $\mathbb{S}^2$ 29
4.2.1 $H = -\frac{1}{2}$ 30
4.2.2 $H < 0$ 32
4.3 Numerical Approximation 32
4.4 Numerical Tests 35

5 Invariant Discrete Euler–Lagrange equations **37**

5.1 Discrete Euler–Lagrange equations 37
5.2 Correction Terms and Correction Elements 38

5.2.1 Correction Terms for Differentiation 39
5.2.2 Correction Element for Shifting 42
5.3 Invariant Euler–Lagrange Equations 44
5.4 The Solution in Original Coordinates 46

6 Implementation and Numerical Tests **49**
6.1 $SL(2)$ 49
6.1.1 Numerical Algorithm 51
6.1.2 Analysis and Tests 51
6.2 $SE(2)$ 52
6.2.1 Numerical Algorithm 54
6.2.2 Tests 55

7 Conclusions and Further Work **57**

CHAPTER 1

INTRODUCTION

When studying an object in mathematics or other fields, symmetries, transformations which do not change the object, are often useful for simplifying or reducing a problem. The symmetries of our concern are Lie group actions acting on a manifold. The machinery of moving frames provides tools for calculating and studying objects with such symmetries. We use the moving frame formalism as developed by Fels and Olver [4, 3, 14]. An elementary introduction, which also provides material on the variational problems studied in this thesis is the book by Mansfield [8].

Our main goal is to study variational problems and discretizations of these. The objects to be studied are functions $f : X \to U$, where X and U are smooth manifolds. While the theory in this thesis makes no further assumptions on X and U, the numerical examples will have $X = \mathbb{R}$, $U = \mathbb{R}^n$.

The thesis assumes the reader to be familiar with basic differential topology concepts including the concept of manifolds, tangent bundles and tangent maps. It also assumes the reader is familiar with the basic concepts of Lie groups and Lie algebras.

A few notes on assumptions and notations

- The Einstein summation convention is used. If an index appears twice in a term it is an indication that this index should be summed over.
- Function and manifolds are assumed to be C^∞ smooth, unless otherwise noted.
- We use the two-argument atan2 in some formulas, to avoid the usual quadrant issues of arctan. Mathematically, $\operatorname{atan2}(y, x)$ is the principal argument of the complex number $x + iy$.
- Mostly, low indices are used for elements of a sequence, typically a numerical solution, while high indices are used for coordinates, though his convention has not been followed completely.

CHAPTER 2

GROUP ACTIONS AND MOVING FRAMES

2.1 Basic Definitions and Results

Definition 1. Let G be a Lie group with identity e and M a smooth manifold. Furthermore, let $U \subset G \times M$ be an open set with $\{e\} \times M \subset U$. A smooth *local left group action* is a smooth function $\Psi : U \to M$ satisfying

(a) If $(h, z) \in U$, $(g, \Psi(h, z)) \in U$ and $(gh, z) \in U$, then

$$\Psi(g, \Psi(h, z)) = \Psi(gh, z).$$

(b) For all $z \in M$,

$$\Psi(e, z) = z$$

(c) If $(g, z) \in U$, then also $(g^{-1}, \Psi(g, z)) \in U$ and

$$\Psi(g^{-1}, \Psi(g, z)) = z$$

If $U = G \times M$, we say Ψ is a *global left group action.* In this case, part (c) above follows from parts (a) and (b).

A *local right group action* satisfies part (b) and (c) above and

(a') If $(h, z) \in U$, $(g, \Psi(h, z)) \in U$ and $(hg, z) \in U$, then

$$\Psi(g, \Psi(h, z)) = \Psi(hg, z).$$

We will restrict the study to *connected* group actions, which means that in addition to (a), (b) and (c),

(d) G and M are connected as manifolds,

(e) U is connected,

(f) $G_z = \{g \in G \mid (g, z) \in U\}$ is connected for every $z \in M$.

We will make use of the alternative notations

$$g \cdot z = \Psi(g, z)$$

if the action is left, and

$$z \cdot g = \Psi(g, z)$$

if the action is right. In local coordinates $z = (z^1, \ldots, z^m)$ on the manifold, we will sometimes abuse this notation and write e.g. $g \cdot z^i$ for the ith coordinate of $g \cdot z$.

Additionally we will use the notations

$$\Psi_g(z) = g \cdot z$$
$$\Psi_z(g) = g \cdot z$$

for the group action with either argument fixed.

The difference between local and global group actions is somewhat technical. In this thesis definitions are mostly stated just for the global case. The definitions for local group actions are analogous, but with additional restrictions. Additionally, most group actions of our concern are left.

For theoretical purposes it is sometimes useful to restrict the action to a connected open set $V \in M$, which requires that the domain of U is restricted such that all the axioms above hold.

We will be interested in functions that are invariant or equivariant under a left group action.

Definition 2.

(a) A function $f : M \to \mathbb{R}$ is *invariant* under the group action if

$$f(g \cdot x) = f(x)$$

for all $(g, x) \in G \times M$.

(b) Given two manifolds M and N, and group actions

$$\Psi^M : G \times M \to M$$
$$\Psi^N : G \times N \to N$$

with common Lie group G, a function $F : M \to N$ is *equivariant* if for all $g \in G$

$$F \circ \Psi_g^M = \Psi_g^N \circ F$$

as functions $M \to N$

The right-equivariant moving frames are equivariant maps from M to G, with the left action of G on itself defined by $\Psi_g(h) = hg^{-1}$.

Definition 3. A *right-equivariant moving frame* for a group action is a smooth function $\rho : M \to G$ satisfying the right-equivariance

$$\rho(g \cdot z) = \rho(z) \cdot g^{-1} \tag{2.1}$$

for all $g \in G$, $z \in M$.

Remark: If $z \mapsto \rho(z)$ is a right-equivariant moving frame, $z \mapsto \widetilde{\rho}(z) = \rho(z)^{-1}$ is a *left-equivariant moving frame*, satisfying $\widetilde{\rho}(g \cdot z) = g \cdot \widetilde{\rho}(z)$. While the classic moving frames pioneered by Cartan [1] and others, are left-equivariant, Fels and Olver [4] introduced the right-equivariant version as more practical for computations, and in this thesis, all moving frames are right-equivariant.

Under some assumptions on the group action, a local moving frame, defined in a neighbourhood of an arbitrary point $z \in M$, exists, and can be constructed. To state these assumptions, we need some terminology.

Definition 4.

(a) For a point $z \in M$ the *(group) orbit* is the submanifold $\mathcal{O}(z) = \{g \cdot z \mid g \in G\}$

(b) The *isotropy subgroup* of a point $z \in M$ is $G_z = \{g \in G \mid g \cdot z = z\}$.

(c) A group action is:

- *Locally free* if G_z is a discrete subgroup for all $z \in M$.
- *Free* if $G_z = e$ for all $z \in M$.
- *Regular* if all orbits are of the same dimension, and for each point $x \in M$, there exists a neighbourhood U of x such that the intersection between an orbit and U is either empty or connected.
- *Locally effective* if the *global isotropy subgroup* $G_M^{\star} = \bigcap_{z \in M} G_z$ is discrete.
- *Effective* if $G_M^{\star} = e$.

(d) A submanifold which intersects each group orbit transversally and exactly once, is a *cross-section.*

We note that the adjective "locally" is natural in the sense that if Ψ is locally free, (resp. effective), then by suitably restricting the domain of Ψ, the resulting local group action is free (resp. effective). If the group action is free and regular, then for any point $z \in M$ there is a neighbourhood $U \subset M$, such that there exists a cross-section $\mathcal{K} \subset U$ for the group action restricted to U.

Theorem 1. *Let G act freely and regularly on M, and let $\mathcal{K}$ be a cross-section. Given $z \in M$, let $\rho(z) \in G$ be the unique group element such that $\rho(z) \cdot z \in \mathcal{K}$. Then $\rho : M \to G$ is a (right-equivariant) moving frame.*

Proof. Let $z \in U$, and $h \in G$ such that $h \cdot z \in U$. Since each group orbit intersects $\mathcal{K}$ exactly once, $\rho(h \cdot z) \cdot h \cdot z = \rho(z) \cdot z$. Since the action is free, $\rho(h \cdot z) \cdot h \cdot \rho(z)^{-1} = e$ and the equivariance follows. □

Usually, the cross-section is defined in local coordinates $z = (z^1, \ldots, z^m)$ as the locus of a set of equations

$$Z^i(z) = 0, \qquad i = 1, \ldots, r,$$

and the moving frame can be obtained by solving the equations

$$Z^i(\rho(z) \cdot z), \qquad i = 1, \ldots, r,$$

with respect to the group element $\rho(z)$.

The choice of cross-section $\mathcal{K}$ and moving frame induces an invariantization operator on the space of functions on M. If F is any function on M, its invariantization $\iota(F)$ is invariant under the group action and is defined by

$$\iota(F)(z) = F(\rho(z) \cdot z).$$

Of special importance are the invariantizations of the coordinate functions $z \mapsto z^i$. These are the *fundamental invariants* $I^i = \iota(z^i)$. In coordinates, the invariantization of a function is

$$\iota[F(z^1, \ldots, z^m)] = F(I^1(z), \ldots, I^m(z)).$$

Since the invariantization of an invariant function is the function itself, it follows that any invariant function can be written in terms of the fundamental invariants. This result, known as the replacement theorem, is often useful.

Theorem 2 (Replacement theorem)**.** *If F is an invariant function, then*

$$F(z^1, \ldots, z^m) = \iota[F(z^1, \ldots, z^m)] = F(I^1(z), \ldots, I^m(z)).$$

In particular, the invariantizations of the cross-section equations $\iota\left(Z^i\right)$ are constant. These are the *phantom invariants.*

2.2 Infinitesimals

Let $\mathfrak{g} = T_eG$ be the Lie algebra of the Lie group G. The tangent map at the identity

$$T_e\Psi_z : \mathfrak{g} \to T_zM$$

defines the infinitesimal action

$$\psi : \mathfrak{g} \times M \to TM.$$

If $v_1, \ldots, v_r$ form a basis for the Lie algebra, the corresponding vector fields

$$\mathbf{v}_j(z) = \psi(v_j, z) = \psi_i^j(z)\partial_{z^i}$$

are the *generators* of the group action. The basis will typically be defined by local coordinates $(g^1, \ldots, g^r)$ on G near the identity such that $\mathbf{v}_j(z) = \frac{\partial}{\partial g^j}\Psi(g, z)\Big|_e$.

2.3 Prolongation

For any action Ψ on a manifold M, there are induced actions on M's tangent bundle and more generally, the *jet bundles* $J^n(M,p)$ over M. $J^n(M,p)$ is defined as the set of equivalence classes of p-dimensional submanifolds under the equivalence relation of nth order contact at a single point. The induced action on a jet bundle is known as the prolongation of the action. For the tangent bundle, the prolonged action

$$\Psi^{TM} : G \times TM \to TM$$

is defined by

$$\Psi_g^{TM} = T\Psi_g$$

that is, differentiation with respect to the manifold variable. The actions on higher order jet bundles are defined in a similar manner.

Prolongation has a regularizing effect, if Ψ acts effectively on open subsets, meaning that the only group element fixing every point in an open subset is e, then for a sufficiently large n, the prolongation of Ψ acts locally freely and regularly on an open dense subset of $J^n(M,p)$ [13], so that even if Ψ does not admit a moving frame, $\Psi^{J^n(M,p)}$ for n sufficiently large does.

Locally on $J^n(M,p)$, we can split M into $X \times U$, where X is p-dimensional and consider p-dimensional submanifolds as smooth functions $f : X \to U$. For indexing variables and invariants, it is convenient to use *multi-indices*. With $(x^1, \ldots, x^p)$ local coordinates on X, $(u^1, \ldots, u^{m-p})$ local coordinates on U we use $K = (k_1, \ldots, k_l)$ to index the partial derivatives

$$u_K^\alpha = \frac{\partial^l u^\alpha}{\partial x^{k_1} \cdots \partial x^{k_l}},$$

and use the (x^i, u_K^α) as local coordinates on $J^n(M,p) = X \times U^{(n)}$. The corresponding fundamental invariants are $I_K^\alpha = \iota(u_K^\alpha)$.

We also define the differential operator

$$D_K = \frac{\partial^l}{\partial x^{k_1} \cdots \partial x^{k_l}}.$$

Due to the replacement theorem, any invariant expressed in u^α and derivatives of these can be written in terms of the fundamental invariants. For simpler notation we will often write simply Ψ also for the prolonged action of Ψ, and use the multi-indices to index the elements of $\Psi(g,z)$, where z is an element of the jet bundle.

The generating vector fields on $J^n(M,p)$ are given in terms of the original generators by the prolongation formula. In local coordinates, if the original vector field is

$$\mathbf{v} = \xi^i \partial_{x^i} + \phi^\alpha \partial_{u^\alpha}$$

the prolonged vector field is

$$\mathbf{pr}\,\mathbf{v} = \mathbf{v} + \sum_{\alpha,K} \phi_K^\alpha \partial_{u_K^\alpha}$$

where $\phi_K^\alpha = D_K\left(\phi^\alpha - \xi^i u_i^\alpha\right) + \xi^i u_{K,i}^\alpha$.

A group action on M also induces naturally an action on $M^n = M \times \cdots \times M$

$$\Psi_g^{M^n}(z_1, \ldots, z_n) = (\Psi_g(z_1), \ldots, \Psi_g(z_n)).$$

which we will call the prolongation to M^n. If $\mathbf{v}(z)$ is a generator of the original group action Ψ, the corresponding generator for Ψ^{M^n} is

$$\mathbf{v}^{M^n} = \sum_{i=1}^{n} \mathbf{v}_i(z_i)$$

where $\mathbf{v}_i(z_i) \in T_{z_i} M_i$.

2.4 Differentiation of Invariants

We mainly consider cases where the manifold can be split into $X \times U$ where $x \in X$ represent the independent variables and $u \in U$ the dependent variables. Furthermore, we assume that x is invariant under the group action. It is important to note that invariantization and differentiation do not commute, even if the differentiation is with respect to an invariant variable. However, it is possible to calculate derivatives of the fundamental invariants I_K^α in terms of other fundamental invariants. Combined with the replacement theorem, this can be used to differentiate any invariant function. Fels and Olver [3, Section 13] showed that these expressions can in fact be calculated from the prolonged vector fields and the equations for the cross-section. What follows is a proof of these relations for the case of invariant independent variables.

Lemma 1. *Define $R_g : G \to G$ by $R_g(h) = hg$. Then*

$$\begin{aligned} T_e\Psi_{g\cdot z} &= T_g\Psi_z \circ T_e R_g \\ &= T_g\Psi_z \circ (T_g R_{g^{-1}})^{-1} \end{aligned}$$

Proof. The first equality follows from the identity

$$\Psi_z \circ R_g = \Psi_{g\cdot z}$$

and the chain rule. The second equality follows from $T(R_g \circ R_{g-1}) = \boldsymbol{id}$. □

Lemma 2.

$$T_z\rho = T_{\rho(g\cdot z)} R_g \circ T_{g\cdot z}\rho \circ T_z\Psi_g$$

for all $g \in G$, and specifically, for $g = \rho(z)$

$$T_z\rho = T_e R_{\rho(z)} \circ T_{\rho(z)\cdot z}\rho \circ T_z\Psi_{\rho(z)}$$

Proof. $\rho(z) = \rho(g \cdot z) \cdot g$ for all z, thus $\rho = R_g \circ \rho \circ \Psi_g$. The lemma follows from the chain rule. □

Theorem 3. *Assume that $\Psi : G \times M \to M$ has a cross-section $\mathcal{K}$ with corresponding moving frame $\rho : M \to G$. Let $z = (z^1, \dots, z^m)$ be local coordinates near some point $x \in \mathcal{K}$ and assume that in these coordinates $\mathcal{K}$ is defined as the kernel of the function*

$$Z(z) = (Z^1(z), \dots, Z^r(z)) \in \mathbb{R}^r.$$

Furthermore let $\mathbf{v}_1, \dots, \mathbf{v}_k$, where the $\mathbf{v}_j = \psi_i^j(z)\partial_z^i$ be the generators of the action. Define the matrices

$$\boldsymbol{J}(z) = T_{\rho(z)\cdot z} Z,$$
$$\boldsymbol{\psi}(z) = T_e \Psi_{\rho(z)\cdot z}$$

with entries

$$\boldsymbol{J}(z)_{i,j} = \frac{\partial Z^i}{\partial y^j}\bigg|_{y=\rho(z)\cdot z}$$
$$\boldsymbol{\psi}(z)_{i,j} = \psi_i^j(\rho(z) \cdot z),$$

The tangent map of the invariantization map

$$\iota : z \mapsto \Psi(\rho(z), z)$$

is in local coordinates given by

$$T_z \iota = T_z \Psi_{\rho(z)} - \boldsymbol{\psi}(\boldsymbol{J}\boldsymbol{\psi})^{-1} \boldsymbol{J} T_z \Psi_{\rho(z)}.$$

Proof. By the product rule and lemmas 1 and 2

$$\begin{aligned} T_z \iota &= T_{\rho(z)} \Psi_z \circ T_z \rho + T_z \Psi_{\rho(z)} \\ &= T_e \Psi_{\rho(z)\cdot z} \circ (T_e R_{\rho(z)})^{-1} \circ T_e R_{\rho(z)} \circ T_{\rho(z)\cdot z} \rho \circ T_z \Psi_{\rho(z)} + T_z \Psi_{\rho(z)} \\ &= T_e \Psi_{\rho(z)\cdot z} \circ T_{\rho(z)\cdot z} \rho \circ T_z \Psi_{\rho(z)} + T_z \Psi_{\rho(z)} \end{aligned} \tag{2.2}$$

Assume that $Z^1, \dots, Z^k, z^{k+1}, \dots, z^m$ are functionally independent near x (if not, rearrange the z^i), so that $\eta(z) = (Z^1, \dots, Z^k, z^{k+1}, \dots, z^m)$ form local coordinates. The Jacobian of the transformation map η, at $\rho(z) \cdot z \in K$ is

$$T_{\rho(z)\cdot z} \eta = \begin{bmatrix} \boldsymbol{J} \\ \boldsymbol{A} \end{bmatrix},$$

where $\boldsymbol{J}$ is as defined above, and $\boldsymbol{A}$ plays no further role.

By construction, $Z = \pi \circ \eta$, where π is projection onto the first k coordinates. By $\rho(z) \cdot z \in K$, we have

$$Z \circ \iota = \pi \circ \eta \circ \iota = 0$$

Taking the tangent map of $\pi \circ \eta \circ \iota$ at the cross-section gives

$$T_{\eta(\rho(z)\cdot z)} \pi \circ T_{\rho(z)\cdot z} \eta \circ T_{\rho(z)\cdot z} \iota = 0.$$

Since

$$T\pi = \begin{bmatrix} \boldsymbol{id} & \mathbf{0} \end{bmatrix},$$

we deduce that

$$T_{\rho(z)\cdot z}\eta \circ T_{\rho(z)\cdot z}\iota = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{A} \end{bmatrix}, \tag{2.3}$$

for some $\boldsymbol{A}$ (different from the previous $\boldsymbol{A}$). The top k rows of the matrix equation (2.3) are

$$\boldsymbol{J} \circ T_{\rho(z)\cdot z}\iota = \mathbf{0}.$$

Inserting from equation (2.2), we get

$$\boldsymbol{J}(T_e\Psi_{\rho(z)\cdot z} \circ T_{\rho(z)\cdot z}\rho \circ T_z\Psi_e + T_z\Psi_e) = \mathbf{0}$$

$$\boldsymbol{J\psi} \circ T_{\rho(z)\cdot z}\rho = -\boldsymbol{J},$$

where the last line is due to $T_z\Psi_e = \boldsymbol{id}$. Since the cross-section is assumed to intersect the orbits of the group action transversally, $\boldsymbol{J\psi}$ is of full rank r, so this implies

$$T_{\rho(z)\cdot z}\rho = -(\boldsymbol{J\psi})^{-1}\boldsymbol{J}.$$

Inserting this into equation (2.2) completes the proof. □

In our applications, the Z^i depend on only a few of the z^i, say $\zeta^1, \dots, \zeta^l$. Let J' be the non-zero columns of J

$$J'_{i,j} = \frac{\partial Z^i}{\partial \zeta^j}.$$

We apply the theorem to fundamental invariants of a prolonged action. We let the manifold be $J^n(M,p)$, which we split as before and use local coordinates $z = (u^1, \dots, u^1_1, \dots, u^\alpha_K, \dots)$. Where the indices on matrices in theorem 3 refers to coordinate z^i, we instead index with (α, K).

For a fundamental invariant $I^\alpha_K = \iota\left(u^\alpha_K\right) = \iota(z)_{\alpha,K}$, which we differentiate with respect to the invariant variable x^i, we have by the theorem

$$\begin{aligned}
\frac{\partial}{\partial x^i} I^\alpha_K &= \frac{\partial}{\partial x^i}\left(\iota(z)_{\alpha,K}\right) \\
&= \left(\frac{\partial}{\partial x^i}\iota(z)\right)_{\alpha,K} \\
&= \left(T_z\iota\left(\frac{\partial z}{\partial x^i}\right)\right)_{\alpha,K} \\
&= \left(T_z\Psi_{\rho(z)}\frac{\partial z}{\partial x^i}\right)_{\alpha,K} - \left(\boldsymbol{\psi}(\boldsymbol{J\psi})^{-1}\boldsymbol{J}T_z\Psi_{\rho(z)}\frac{\partial z}{\partial x^i}\right)_{\alpha,K} \\
&= \left(T_z\Psi_{\rho(z)}\frac{\partial z}{\partial x^i}\right)_{\alpha,K} - \boldsymbol{\psi}_{\alpha,K}(\boldsymbol{J\psi})^{-1}\boldsymbol{J}T_z\Psi_{\rho(z)}\frac{\partial z}{\partial x^i},
\end{aligned}$$

where the subscripts on matrices refers to rows. The first term of the final right hand side is equal to $\iota(u^\alpha_{Ki}) = I^\alpha_{Ki}$ by the definition of prolongation. The second term can be simplified by deleting the zero columns of $\boldsymbol{J}$ and the corresponding rows of $\boldsymbol{\psi}$ and elements of $T_z\Psi_{\rho(z)}\frac{\partial}{\partial x^i}z$. The remaining rows of $\boldsymbol{\psi}$ are the coefficients of ∂_{ζ^j} in the generators, evaluated at the cross-section, and the remaining elements in $T_z\Psi_{\rho(z)}\frac{\partial}{\partial x^i}z$ are the fundamental invariants $\iota\left(\frac{\partial\zeta^j}{\partial x^i}\right)$, again by the definition of prolongation.

Let $\boldsymbol{\psi}^{\boldsymbol{\zeta}}$ be the remaining rows of $\boldsymbol{\psi}$, and let

$$\boldsymbol{T}_{i,j} = \iota\left(\frac{\partial\zeta^i}{\partial x^j}\right).$$

and define the *correction matrix*

$$\boldsymbol{C} = -(\boldsymbol{J}'\boldsymbol{\psi}^{\boldsymbol{\zeta}})^{-1}\boldsymbol{J}'\boldsymbol{T}.$$

The correction matrix, which only depends on the ζ^i appearing in the cross-section equations and their first derivatives, provide the correction terms relating invariantization and derivation. We sum up the discussion above in a theorem.

Theorem 4. *The derivatives of the fundamental invariants are*

$$\partial_{x^i} I^\alpha_K = I^\alpha_{Ki} + \boldsymbol{C}_{j,i}\psi^j_{\alpha,K}, \tag{2.4}$$

where $\psi^j_{\alpha,K} = \left.\frac{\partial u^\alpha_K}{\partial g_j}\right|_{g=e}$ *and* $\boldsymbol{C}$ *is as described above.*

The symbolic differential formulas above gives a way to express certain I^α_K as functions of lower-order invariants and derivatives of these. We will call a finite set of invariants $\mathcal{I}_{\text{gen}}$ such that all invariants can be written as a function of invariants in $\mathcal{I}_{\text{gen}}$ and a finite number of their derivatives for a *set of generators*. It is a classic result that a finite set of generators always exists. A more recent result is that if the frame is *minimal*, then the set

$$\left\{\iota(x^j), \iota(u^\alpha), \iota\left(\frac{\partial Z^i}{\partial x^j}\right)\right\}$$

is a set of generators. [7]

CHAPTER 3

MECHANICAL SYSTEMS

We approach classical mechanics as formulated by Lagrange and Hamilton from a geometric point of view. Important sources for the theoretical background in this chapter are a book by Marsden and Ratiu [9], and an article by Marsden and West [10]. We follow these sources and consider Hamiltonian mechanics as vector fields on symplectic manifolds.

3.1 Symplectic Manifolds and Symplectic Maps

We first recall some concepts from differential topology.

Definition 5. Let M be a smooth manifold.

- A *differential k-form* ω over M assigns to any point $q \in M$ a k-linear, alternating map
$$\omega_q : T_qM \times \cdots \times T_qM \to \mathbb{R},$$
in a smooth manner. We identify differential zero-forms with smooth functions $M \to \mathbb{R}$.

- The *cotangent bundle* over M, T^*M is the set of elements of the form (q, p), where $q \in M$ and
$$p : T_qM \to \mathbb{R}$$
is linear. Differential one-forms over M are sections of T^*M.

- The *wedge product* of a k-form ω and an l-form ξ, is the $k + l$ form defined by

$$(\omega \wedge \xi)_q(v_1, \dots v_{k+l}) \\ = \frac{(k+l)!}{k!\, l!} \sum_{\pi} (\operatorname{sgn} \pi) \omega_q \left(v_{\pi(1)}, \dots, v_{\pi(k)}\right) \xi_q \left(v_{\pi(k+1)}, \dots, v_{\pi(k+l)}\right)$$

where all the v_i lie in T_qM and the sum is over all permutations of $k+l$ elements. $\operatorname{sgn} \pi$ is the sign of the permutation,

$$\operatorname{sgn} \pi = \begin{cases} 1 & \text{if } \pi \text{ is even} \\ -1 & \text{if } \pi \text{ is odd.} \end{cases}$$

- If $\mathbf{v}$ is a vector field on M, the *interior derivative* $\mathbf{i}_{\mathbf{v}}$ is a map from $k+1$-forms to k-forms defined by

  $$(\mathbf{i}_{\mathbf{v}}\omega)_q (w_1, \dots, w_k) = \omega_q(\mathbf{v}|_q, w_1, \dots, w_k)$$

  where all the w_i lie in T_pM and $\mathbf{v}|_q$ is the vector field evaluated at q.

- The *exterior derivative* d is a map from k-forms to $(k+1)$-forms satisfying

  (a) If f is a 0-form, i.e. a function. $\mathrm{d}f$ is the normal tangent map.
  (b) d is linear.
  (c) d satisfies the product rule. If ω is a k-form and ξ an l-form then

  $$\mathrm{d}(\omega \wedge \xi) = \mathrm{d}\omega \wedge \xi + (-1)^k \omega \wedge \mathrm{d}\xi$$

  (d) $\mathrm{d}(\mathrm{d}\omega) = 0$ for all ω.

- If $F : M \to N$ is a smooth function between manifolds, its *pull-back* F^* maps differential forms over N to differential forms over M. Let ω be a differential k-form over N, and $q \in M$, and $v_1, \dots, v_k \in T_qM$.

  $$(F^*\omega)_q(v_1, \dots, v_k) = \omega(T_qFv_1, \dots, T_qFv_k).$$

  Pull-backs of functions commute with the wedge product and the exterior derivative [9, Section 4.2].

- If $\mathbf{v}$ a vector field on M and ϕ_t its flow, then the *Lie derivative* of a k-form is

  $$\mathcal{L}_{\mathbf{v}}\omega = \frac{\partial}{\partial t} \phi_t^* \omega \bigg|_{t=0}$$

- If $F : M \to N$ is a diffeomorphism, its *cotangent lift* T^*F is a map

  $$T^*F : T^*N \to T^*M$$

  such that

  $$T^*F(q,p) = (F^{-1}(q), \tilde{p})$$

  where $\tilde{p}$ is the linear function on $T_{F^{-1}(q)}M$ defined by

  $$\tilde{p}(v) = p(T_{F^{-1}(q)}v)$$

We will also need *Cartan's magic formula* [9, Theorem 4.3.3].

$$\mathcal{L}_{\mathbf{v}}\omega = \mathrm{d}(\mathbf{i}_{\mathbf{v}}\omega) + \mathbf{i}_{\mathbf{v}}(\mathrm{d}\omega).$$

A *symplectic manifold* is a smooth manifold M, equipped with a nowhere vanishing two-form Ω. The two-form relates a smooth Hamiltonian $H : M \to \mathbb{R}$ to the Hamiltonian vector field $\mathbf{v}_H$ defined by the relation

$$\mathbf{i}_{\mathbf{v}_H}\Omega = \mathrm{d}H \tag{3.1}$$

The integral curves of $\mathbf{v}_H$ generate the flow of the Hamiltonian ϕ_t. The flow satisfies the two properties

- $\phi_t^* H = H$
- $\phi_t^* \Omega = \Omega$

Proof. Indeed,

$$\begin{aligned}
\frac{\partial}{\partial t}\phi_t^* H\Big|_{t=0} &= \mathcal{L}_{\mathbf{v}_H} H \\
&= \mathrm{d}H\left(\mathbf{v}_H\right) \\
&= \Omega(\mathbf{v}_H, \mathbf{v}_H) \\
&= 0 \\
\frac{\partial}{\partial t}\phi_t^* \Omega\Big|_{t=0} &= \mathcal{L}_{\mathbf{v}_H} \Omega \\
&= \mathbf{i}_{\mathbf{v}_H}\mathrm{d}\Omega + \mathrm{d}(\mathbf{i}_{\mathbf{v}_H}\Omega) \\
&= 0 + \mathrm{d}^2 H \\
&= 0
\end{aligned}$$

□

The last property above implies that the symplectic form is preserved under the pull-back of ϕ_t for all t. Maps with this property are called *symplectic maps*. As this is an important property of Hamiltonian flows, it seems natural to search for numerical schemes which also have this property. Such numerical schemes are known as *symplectic integrators*, and have been widely studied.

In most practical applications, the symplectic manifold M arises as the cotangent bundle over some configuration manifold Q. We write points in T^*Q as (q, p) and write $\langle p, v\rangle$ for pairing of covectors and vectors over Q, and $\omega \cdot u$ to denote the pairing of covectors and vectors over $M = T^*Q$. In both cases, the base point of the covector and vector is assumed to be the same. Additionally, we will sometimes abuse notation and write $\langle (q, p), v\rangle$ for $\langle p, v\rangle$ with base point q.

When $M = T^*Q$, the symplectic form is the *canonical two-form*, defined as follows. Let Θ be the *canonical one-form* on T^*Q. If u is a vector in $T_{(q,p)}(T^*Q)$ then

$$\Theta(q, p) \cdot u = \left\langle p,\ T_{(q,p)}\pi\ u\right\rangle$$

where $\pi : T^*Q \to Q$ is the natural projection. The canonical two-form is then $\Omega = -\mathrm{d}\Theta$.

Maps $F : T^*Q \to T^*S$ such that

$$F^*\Theta_S = \Theta_Q,$$

where Θ_Q and Θ_S are the canonical one-forms on their respective cotangent bundles, are *special symplectic maps*. The special symplectic maps $F : T^*Q \to T^*S$ are exactly the cotangent lifts of diffeomorphisms $f : S \to Q$ [9, Proposition 6.3.2].

If

$$q = (q^1, \dots, q^n)$$

are local coordinates on Q, we can expand with

$$p = (p^1, \dots, p^n)$$

to get local coordinates on T^*Q. We will choose the p^i to correspond with the natural basis on TQ, i.e. if $v = v^i\partial_{q^i} \in T_qQ$, then

$$\langle p, v\rangle = p^i v^i.$$

In such coordinates

$$\Theta = p^i \mathrm{d}q^i$$
$$\Omega = \mathrm{d}q^i \wedge \mathrm{d}p^i,$$

where $\mathrm{d}q^i$ and $\mathrm{d}p^i$ are the exterior derivatives of the coordinate functions q^i and p^i

The following theorem, whose proof can be found in [10, Sections 1.4.4-5], shows that the canonical one-form Θ determines symplectic maps.

Theorem 5. *Let*

$$F : T^*Q \to T^*Q$$

be a smooth map and let $\Gamma(F) \subset T^*Q \times T^*Q$ *be its graph. Furthermore let*

$$\pi_{1,2} : T^*Q \times T^*Q \to T^*Q$$

be the projections onto each of the components,

$$i : \Gamma(F) \to T^*Q \times T^*Q$$

the inclusion map, and

$$\hat{\Theta} = \pi_2^*\Theta - \pi_1^*\Theta.$$

Then F is symplectic if and only if there exists, at least locally, a smooth function $S : \Gamma(F) \to \mathbb{R}$ such that $i^*\hat{\Theta} = \mathrm{d}S$. The function S is called the generating function *of F, and we say that F is generated by S.*

A special case occurs if $(\pi_1 \times \pi_2) \circ i : \Gamma(F) \to Q \times Q$ is a diffeomorphism. In this case, S can be written as a function $S : Q \times Q \to \mathbb{R}$ and is a *generating function of the first kind.* In coordinates, the relation between F and S becomes

$$F : (q_1, p_1) \mapsto (q_2, p_2)$$

where
$$p_1 = -D_1 S(q_1, q_2)$$
$$p_2 = D_2 S(q_1, q_2)$$
and D_1 and D_2 are the partial differential operators with respect to q_1 and q_2, respectively.

Generating functions of the first kind have a nice interpretation when combined with Hamilton's principle. To arrive at this interpretation and its consequences for symmetries and discrete mechanics, we first need to recall some classical mechanics.

3.2 Lagrangian and Hamiltonian Systems

In Lagrangian mechanics, one considers mechanical systems on some configuration space Q, for which one can define kinetic energy $T(q, \dot{q})$ and potential energy $U(q)$. The mechanical system follows a path in Q which minimizes the integral of the *Lagrangian* $L\mathrm{d}t = (T-U)\mathrm{d}t$. This is *Hamilton's principle.* By variational calculus, it can be shown that such paths satisfy the *Euler–Lagrange* equations, which in local coordinates $(q^1, \ldots, q^n)$ are
$$E^i(L) = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) + \frac{\partial L}{\partial q^i} = 0.$$

In most of the interesting cases, the Lagrangian formalism is equivalent to the Hamiltonian formalism. In Hamiltonian mechanics, systems are defined by their Hamiltonian function defined on the cotangent bundle of the configuration space. The system follows integral curves of the Hamiltonian vector field, which is defined by (3.1), or in coordinates
$$\mathbf{v}_H = \left(\frac{\partial H}{\partial p^i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p^i}\right).$$

The relation between the Lagrangian and Hamiltonian formulation is the *Legendre transformation* $FL : TQ \to T^*Q$, defined by
$$FL(q, v) = (q, p)$$
such that
$$\langle p, w\rangle = \frac{\partial}{\partial \epsilon} L(q, v + \epsilon w)\Big|_{\epsilon=0}. \tag{3.2}$$
Or in coordinates $p^i = \frac{\partial L}{\partial \dot{q}^i}$.

The Hamiltonian corresponding to the variational problem is then
$$H(q, p) = \langle p, v\rangle - L(q, v) \tag{3.3}$$

We assume that the Lagrangian is *hyperregular*, which means that the Legendre transform is a diffeomorphism with inverse $FH : T^*Q \to TQ$ given by
$$FH(q, p) = (q, v)$$
such that
$$\langle \alpha, v\rangle = \frac{\partial}{\partial \epsilon} H(q, p + \epsilon\alpha)\Big|_{\epsilon=0.} \tag{3.4}$$

3.3 Symmetries

We first explore how invariance under a group action relates to the Lagrangian and Hamiltonian formulations. The cotangent lift enables us to prolong a left Lie group action $\Psi : G \times Q \to Q$ to a right action on $\Psi^{T^*Q} : G \times T^*Q \to T^*Q$, by

$$\Psi_g^{T^*Q} = T^*\Psi_g.$$

The generating vector fields on the cotangent bundle are defined by *cotangent lift*. For a generating vector field $\mathbf{v}$ on Q, let $v \in \mathfrak{g}$ be the corresponding element in the Lie algebra. Then its cotangent lift

$$\mathbf{v}^{T^*Q} = \frac{\partial}{\partial \epsilon} T^*\Psi_{g_\epsilon}\Big|_{\epsilon=0},$$

where g_ϵ is a path in in G with $g_0 = e$ and $\frac{\partial g_\epsilon}{\partial \epsilon}\Big|_{\epsilon=0} = v$.

In coordinates, if $\mathbf{v} = \psi^i \partial_{q^i}$, then the cotangent lift is

$$\mathbf{v}^{T^*Q} = -\psi^i \partial_{q^i} + p^j \frac{\partial \psi^j}{\partial q^i} \partial_{p^i}.$$

This formula is equivalent to the formula in [10, Section 1.4.2], but with terms simplified. We include a proof.

Proof. Let $F_\epsilon = \Psi_{g_\epsilon}$ and $T^*F_\epsilon(q,p) = (\tilde{q}, \tilde{p})$, and use coordinates (q^i, p^i), $(\tilde{q}^i, \tilde{p}^i)$ from the same coordinate chart. From the definition of cotangent lift of the action, $\tilde{q} = F_\epsilon^{-1}(q)$. So for the coefficients of the vector field corresponding to the q^i,

$$\frac{\partial \tilde{q}^i}{\partial \epsilon}\Big|_{\epsilon=0} = \frac{\partial}{\partial \epsilon} \left(F_\epsilon^{-1}(q)\right)^i \Big|_{\epsilon=0} = -\psi^i.$$

For the remaining terms, we let $w = w^i \partial_{\tilde{q}^i}$ be an arbitrary vector with base point $\tilde{q}$. We write $\langle \tilde{p}, w \rangle = \tilde{p}^i w^i$, and similar for p and use the definition of cotangent lift to get

$$\begin{aligned} \langle \tilde{p}, w \rangle &= \langle p, T_{\tilde{q}} F_\epsilon w \rangle \\ \tilde{p}^i w^i &= p^j \frac{\partial \left(F_\epsilon(\tilde{q})\right)^j}{\partial \tilde{q}^i} w^i \end{aligned}$$

Differentiating the individual $\tilde{p}^i$ with respect to ϵ gives the coefficients corresponding to the p^i in the lifted vector field.

$$\begin{aligned} \frac{\partial \tilde{p}^i}{\partial \epsilon}\Big|_{\epsilon=0} &= p^j \frac{\partial^2 \left(F_\epsilon(\tilde{q})\right)^j}{\partial \tilde{q}^i}\Big|_{\epsilon=0} \\ &= p^j \frac{\partial \psi^j}{\partial \tilde{q}^i}\Big|_{\epsilon=0} \end{aligned}$$

Where the $\tilde{q}$ in the last line can be replaced by q since $F_0 = \Psi_e$ is the identity map. $\square$

A Lagrangian L is invariant under the prolongation of a group action Ψ if

$$L(q,v) = L(\Psi_g(q), T_q\Psi_g(v)) \tag{3.5}$$

for all $g \in G$ and $(q,v) \in TQ$.

Theorem 6. *Let the Legendre transform FL be defined as in (3.2) and assume that L is invariant under the prolongation of the group action $\Psi : G \times Q \to Q$. Then the following diagram commutes.*

$$\begin{array}{ccc} TQ & \xrightarrow{FL} & T^*Q \\ {\scriptstyle T\Psi_g}\downarrow & & \uparrow{\scriptstyle T^*\Psi_g} \\ TQ & \xrightarrow[FL]{} & T^*Q \end{array}$$

Proof. Let $(q,v) \in TQ$ and $g \in G$ be arbitrary. We calculate

$$\begin{aligned} \langle T^*\Psi_g \circ FL \circ T\Psi_g(q,v)\ ,\ w\rangle &= \langle (FL \circ T\Psi_g(q,v))\ ,\ T\Psi_g(q,w)\rangle \\ &= \frac{\partial}{\partial\epsilon} L(\Psi_g(q), T_q\Psi_g(v) + \epsilon T_q\Psi_g(w)) \\ &= \frac{\partial}{\partial\epsilon} L(q, v + \epsilon w) \end{aligned}$$

where the last equality is due to the linearity of $T_q\Psi_g$ and the invariance of L. □

Theorem 7. *Let H be defined as in (3.3). H is invariant under the pull-back of the action Ψ if and only if L is invariant under the prolongation of the action.*

Proof. For the first direction, assume that L satisfies (3.5). The Hamiltonian is $H(q,p) = \langle FL(q,v), v\rangle - L(q,v)$, where $(q,p) = FL(q,v)$

$$\begin{aligned} H(T^*\Psi_g(q,p)) &= H(T^*\Psi_g \circ FL(q,v)) \\ &= H(FL \circ T\Psi_{g^{-1}}(q,v)) \\ &= \langle FL \circ T\Psi_{g^{-1}}(q,v)\ ,\ T_q\Psi_{g^{-1}}v\rangle - L(T\Psi_{g^{-1}}(q,v)) \\ &= \langle T^*\Psi_g \circ FL(q,v)\ ,\ T_q\Psi_{g^{-1}}v\rangle - L(q,v) \\ &= \langle FL(q,v)\ ,\ T_{g\cdot q}\Psi_g \circ T_q\Psi_{g^{-1}}v\rangle - L(q,v) \\ &= H(q,p). \end{aligned}$$

For the converse statement, assume that $H(T^*\Psi_{g^{-1}}(q,p)) = H(q,p)$.

$$L(T\Psi_g(q,v)) = L \circ T\Psi \circ FH(q,p)$$

By $FH = FL^{-1}$, and $T^*\Psi_{g^{-1}} = (T^*\Psi_g)^{-1}$, and that the diagram commutes, it follows that this equals

$$L \circ FH \circ T^*\Psi_{g^{-1}}(q,p)$$

and by (3.3) and (3.4)

$$\begin{aligned}
L \circ FH \circ T^*\Psi_{g^{-1}}(q,p) &= \langle T^*\Psi_{g^{-1}}(q,p)\ ,\ FH \circ T^*\Psi_{g^{-1}}(q,p)\rangle - H(T^*\Psi_{g^{-1}}(q,p)) \\
&= \frac{\partial}{\partial \epsilon} H((T^*\Psi_{g^{-1}}(q,(1+\epsilon)p))\Big|_{\epsilon=0} - H(q,p) \\
&= \frac{\partial}{\partial \epsilon} H(q,(1+\epsilon)p)\Big|_{\epsilon=0} - H(q,p) \\
&= \langle p, v\rangle - H(q,p)
\end{aligned}$$

Where $(q,v) = FH(q,p)$. The invariance of L follows. □

3.4 Discrete Mechanics

The Lagrangian and Hamiltonian formulations of mechanics have corresponding formulations in a discrete setting. While the Hamiltonian H has no direct equivalent in a discrete setting, one can define a discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$. Specifically, in discrete Lagrangian mechanics one searches for a sequence of points $(q_1, q_2, \ldots, q_n)$ which minimizes the discrete action

$$\mathcal{A}(q_1, \ldots, q_n) = \sum_{i=1}^{n-1} L_d(q_i, q_{i+1})$$

typically subject to constraints on the endpoints. Differentiating with respect to q_i leads to the *Discrete Euler–Lagrange equations*

$$D_2 L_d(q_{i-1}, q_i) + D_1 L_d(q_i, q_{i+1}) = 0, \tag{3.6}$$

We define the two discrete Legendre transforms $F^{\pm}L_d : Q \times Q \to T^*Q$ by

$$F^+ L_d(q_i, q_{i+1}) = D_2 L_d(q_i, q_{i+1}) \tag{3.7}$$
$$F^- L_d(q_i, q_{i+1}) = -D_1 L_d(q_i, q_{i+1}). \tag{3.8}$$

The discrete Euler–Lagrange equations can be stated in terms of the Legendre transforms as

$$F^+ L_d(q_{i-1}, q_i) = F^- L_d(q_i, q_{i+1}) = p_i.$$

Where the above equation defines the discrete momentum p_i.

Under the assumption that the Legendre transforms are bijective, L_d is the generating function of the first kind of a symplectic map $\tilde{F}_{L_d} : T^*Q \to T^*Q$

$$\tilde{F}_{L_d} = F^+ L_d \circ (F^- L_d)^{-1}$$

alternatively $\tilde{F}_{L_d} : (q_i, p_i) \mapsto (q_{i+1}, p_{i+1})$ where

$$\begin{aligned}
p_i &= -D_1(q_i, q_{i+1}) \\
p_{i+1} &= D_2(q_i, q_{i+1})
\end{aligned}$$

L_d also defines an advancement map $F_{L_d} : Q \times Q \to Q \times Q$ by

$$F_{L_d} = (F^- L_d)^{-1} \circ F^+ L_d.$$

From the definitions of $F^{\pm} L_d$ it follows that

$$F_{L_d}(q_i, q_{i+1}) = (q_{i+1}, q_{i+2})$$

where (q_i, q_{i+1}, q_{i+2}) satisfy the Discrete Euler–Lagrange equations (3.6).

Theorem 8. *Let $L_d : Q \times Q \to \mathbb{R}$, be invariant under the prolongation of the group action $\Psi : G \times Q \to Q$,*

$$L_d(q_0, q_1) = L_d(\Psi_g(q_0), \Psi_g(q_1)). \tag{3.9}$$

The group action commutes with the discrete Legendre transforms in the sense that the following diagram commutes

$$\begin{array}{ccc} Q \times Q & \xrightarrow{F^{\pm} L_d} & T^*Q \\ \Big\downarrow{\scriptstyle \Psi_g^{Q\times Q}} & & \Big\uparrow{\scriptstyle T^*\Psi_g} \\ Q \times Q & \xrightarrow[F^{\pm} L_d]{} & T^*Q \end{array}$$

Proof. For simpler notation, we write Ψ instead of $\Psi^{Q\times Q}$. For the map $F^+ L_d$,

$$\begin{aligned} \langle T^*\Psi_g \circ F^+ L_d \circ \Psi_g(q_i, q_{i+1})\ ,\ v_{q_{i+1}} \rangle &= \langle F^+ L_d(\Psi_g(q_i), \Psi_g(q_{i+1}))\ ,\ T_{q_{i+1}} \Psi_g v_{q_{i+1}} \rangle \\ &= \langle D_2 L_d(\Psi_g(q_i), \Psi_g(q_{i+1}))\ ,\ T_{q_{i+1}} \Psi_g v_{q_{i+1}} \rangle \\ &= \langle F^+ L_d(q_i, q_{i+1})\ ,\ v_{q_{i+1}} \rangle. \end{aligned}$$

where the last line follows from applying the partial differential D_2 to the identity $L = L \circ \Psi_g^{Q\times Q}$ and using that $\Psi(q_1)$ does not depend on q_2. The proof for the map $F^- L_d$ is completely analogous. □

The following theorem gives a sufficient condition for when maps generated from generating functions of the first kind are equivariant under a group action.

Theorem 9. *If $L_d : Q \times Q \to \mathbb{R}$ is invariant under the prolongation of the group action $\Psi : G \times Q \to Q$, then the maps $F_{L_d} : Q \times Q \to Q \times Q$ and $\tilde{F}_{L_d} : T^*Q \times T^*Q$ are equivariant under the prolongation and the cotangent lift of the action, respectively.*

Proof. The statement follows from combining the commuting diagrams for $F^+ L_d$ and $F^- L_d$. □

To relate the discrete and continuous formulations, one introduces the step-length parameter h in the discrete Lagrangian. The discretization of a continuous Lagrangian is a discrete Lagrangian

$$L_d(q_i, q_{i+1}, h) \approx \int_0^h L(q, \dot{q}) \mathrm{d}t, \tag{3.10}$$

where $(q, \dot{q})$ in the integral are the solution to the Euler–Lagrange equations with $q(0) = q_i, q(h) = q_{i+1}$.

The discrete Lagrangian for which equality holds in (3.10), while usually unobtainable, is of some theoretical interest and is known as the *exact discrete Lagrangian* L_d^E. The correspondence between the continuous and discrete Legendre transforms is shown in local coordinates

$$\begin{aligned} F^- L_d^E(q_i, q_{i+1})^\alpha &= -\frac{\partial}{\partial q_i^\alpha}\int_0^h L(q,\dot{q})\mathrm{d}t \\ &= -\int_0^h \frac{\partial L}{\partial q^\beta}\frac{\partial q^\beta(t)}{\partial q_i^\alpha} + \frac{\partial L}{\partial \dot{q}^\beta}\frac{\partial \dot{q}^\beta(t)}{\partial q_i^\alpha}\mathrm{d}t \\ &= -\int_0^h \left[\frac{\partial L}{\partial q^\beta} - D_t\left(\frac{\partial L}{\partial \dot{q}^\beta}\right)\right]\frac{\partial q^\beta(t)}{\partial q_i^\alpha}\mathrm{d}t - \frac{\partial L}{\partial \dot{q}^\beta}\frac{\partial q^\beta}{\partial q_i^\alpha}\Big|_0^T \\ &= \frac{\partial L}{\partial \dot{q}^\alpha}\Big|_{t=0} \\ &= FL(q,\dot{q})\Big|_{t=0}. \end{aligned} \tag{3.11}$$

Similarly $F^+ L_d^E(q_i, q_{i+1}) = FL(q,\dot{q})\Big|_{t=h}$. We see that L_d^E is the generating function of the first kind for the flow of the Hamiltonian vector field.

3.5 Noether's First Theorem

Noether's "Satz I" [11] relates symmetries of a Lagrangian to first integrals of the solutions of the variational problem. A similar theorem holds for discrete systems, and we state and prove the simplest case here.

Theorem 10. *Assume that L_d is invariant under an action with generator $\mathbf{v}$. Then the expression*

$$N_i = \langle p_i, \mathbf{v}|_{q_i}\rangle$$

is constant along solutions of the discrete Euler–Lagrange equations.

Proof.

$$\begin{aligned} \mathbf{v}^{Q\times Q}(L_d)q_i, q_{i+1}) &= \langle D_1 L_d(q_i, q_{i+1}), \mathbf{v}|_{q_i}\rangle + \left\langle D_2 L_d(q_i, q_{i+1}), \mathbf{v}|_{q_{i+1}}\right\rangle \\ &= -\langle p_i, \mathbf{v}|_{q_i}\rangle + \left\langle p_{i+1}, \mathbf{v}|_{q_{i+1}}\right\rangle = 0 \end{aligned}$$

□

The most basic case of the continuous Noether's theorem follows by applying the above theorem to the exact discrete Lagrangian L_d^E combined with equation (3.11), namely that if $\mathbf{pr}\,\mathbf{v}(L) = 0$ then the quantity $\langle p, \mathbf{v}|_q\rangle$ is preserved along solutions of the variational problems.

Noether's theorem has an analogue for Hamiltonian systems, which is related to the Poisson structure of a symplectic manifold. We will not need the full power of Poisson manifolds and instead state this theorem in the language of Hamiltonian vector fields.

Theorem 11. *Assume (M, Ω) is a symplectic manifold, and H, f Hamiltonian functions on M. Assume further that $\mathbf{v}_f(H) = 0$. Then f is constant along integral curves of $\mathbf{v}_H$.*

Proof. Indeed

$$\mathbf{v}_H(f) = \mathrm{d}f \cdot \mathbf{v}_H = \Omega(\mathbf{v}_f, \mathbf{v}_H) = -\mathrm{d}H \cdot \mathbf{v}_f = -\mathbf{v}_f(H) = 0$$

□

There is a natural correspondence between the two theorems above when $\mathbf{v}_f$ is the cotangent lift of some vector field on Q. Let $\mathbf{v}$ be a vector field on Q with $\mathbf{pr}\,\mathbf{v}(L) = 0$. Then the cotangent lift $\mathbf{v}^{T^*M}$ is the vector field of the Hamiltonian $H(q,p) = \langle p, \mathbf{v}|_q \rangle$, which is identical to the first integral from Noether's theorem.

3.6 Invariant Discrete Lagrangians

For first order Lagrangian, the problem of finding equivariant integrators is reduced to finding invariant discrete Lagrangians. The invariantization operator defined in Chapter 2 provides a method for creating invariant discrete Lagrangians. Assume that the Lagrangian $L(q, \dot{q})$ is invariant under the prolongation of a group action $\Psi : G \times Q \to Q$, and let $L_d^0(q_1, q_2, h)$ be a discrete approximation. If $\rho : Q \times Q \to G$ is a moving frame for the group action prolonged to $Q \times Q$, the invariantization of L_d^0,

$$L_d(q_i, q_{i+1}, h) \stackrel{\text{def}}{=} L_d^0(\rho(q_i, q_{i+1}) \cdot q_i, \rho(q_i, q_{i+1}) \cdot q_{i+1})$$

is invariant under the group action, and it is still an approximation to the continuous Lagrangian. If $L_d^0(q_i, q_{i+1}, h) = L_d^E(q_i, q_{i+1}, h) + E(q_i, q_{i+1}, h)$, then

$$\begin{aligned} L_d(q_i, q_{i+1}, h) &= L_d^0(\rho \cdot q_i, \rho \cdot q_{i+1}, h) \\ &= L_d^E(\rho \cdot q_i, \rho \cdot q_{i+1}, h) + E(\rho \cdot q_i, \rho \cdot q_{i+1}, h) \\ &= L_d^E(q_i, q_{i+1}, h) + E(\rho \cdot q_i, \rho \cdot q_{i+1}, h) \end{aligned}$$

where $\rho = \rho(q_i, q_{i+1})$.

The Legendre maps can be calculated by setting $z = (q_i, q_{i+1})$, then

$$T_z L_d = T_{\rho \cdot z} L_d^0 \circ T_z \iota \tag{3.12}$$

where ι is the invariantization map, and $T_z \iota$ is as calculated in (3).

3.7 Reparametrization

We have so far only considered group actions which do not transform the invariant variable. It is often useful to also consider variational symmetries transforming the independent variable t as well. In which case the invariance equation is

$$\mathbf{pr}\,\mathbf{v}_j(L) + LD_t\tau^j = 0,$$

where $\mathbf{v}_j = \tau^j\partial_t + \phi_i^j\partial_{q^i}$ are the generators of the group action. However, under a reparametrization $t = t(s)$, where s is invariant, the case of non-invariant independent variables transforms into the case of invariant independent variables, and the reparametrized Lagrangian $\tilde{L}(t, q, t_s, q_s) = L(t, q, \frac{q_s}{t_s})t_s$, satisfies the equation (3.5). Letting $E^i(L)$ be the Euler–Lagrange equations of the original Lagrangian $L(t, q, \dot{q})$, the Euler–Lagrange equations of $\tilde{L}$ are

$$\begin{aligned} E^i(\tilde{L}) &= t_s E^i(L) \\ E^t(\tilde{L}) &= -q_s^i E^i(L) \end{aligned}$$

These equations are underdetermined, and to get a well-determined equation, one usually adds an equation $t_s = f(t, q, q_s)$, fixing the parametrization. This equation is then added to the Lagrangian as a constraint. To maintain the symmetry, the vector field corresponding to the differential equation $t_s = f(t, q, q_s)$ should be invariant under the group action. See [8, Chapter 7] for details on this construction.

There are several possible approaches to the discretization of Lagrangian systems invariant under a group action with a symmetry which transforms the independent variable. One is to make the reparametrization, and then discretion. Another is to let the group action work on the discretized Lagrangian

$$L_d(t_i, q_i, t_{i+1}, q_{i+1})h_i,$$

where $h_i = t_{i+1} - t_i = h(t_i, q_i, q_{i+1}, t_{i-1}, q_{i-1})$. This approach is studied in [2].

A third approach, which we will study in the treatment of the Kepler problem, is to work with an perturbed problem. The motivation for this is that the symmetry in the Kepler problem, which originally is a generalized symmetry depending on derivatives of q, can also be stated as a symmetry of the Hamiltonian of an equivalent system with reparametrized time. The idea generalizes to other systems with this property, and is stated here.

Assume that we have a Hamiltonian $H(q, p)$ and a perturbed Hamiltonian $K(q, p) = \sigma(q)(H(q, p) - H_0)$ where σ is a real function on Q with $\sigma(q) > 0$ for all q. Since

$$\mathrm{d}K = (H(q, p) - H_0)\mathrm{d}\sigma(q) + \sigma(q)\mathrm{d}H(q, p),$$

it follows that on the submanifold $H = H_0$, the Hamiltonian vector fields of H and K are related by

$$\mathbf{v}_K = \sigma(q)\mathbf{v}_H.$$

Thus, their flows are reparametrizations of each other.

Assume further that the vector field $\mathbf{w}$ on Q is such that $\mathbf{w}^{T^*Q}(K) = 0$. Then $\langle p, \mathbf{w}\rangle$ is preserved along the Hamiltonian flows generated by $\mathbf{v}_K$, and along those flows generated by $\mathbf{v}_H$ which lie on the submanifold $H = H_0$. Let L and $\tilde{L}$ be the Lagrangians corresponding to H and K, respectively. Then a quick calculations shows that

$$\tilde{L}(q, q_s) = \left(L\left(q, q_t\right) + H_0\right)\sigma(q)$$

where $q_t = \frac{q_s}{\sigma(q)}$ and $\mathbf{pr}_s\,\mathbf{w}(\tilde{L}) = 0$. [1] The exact discrete Lagrangian is

$$\begin{aligned}
\tilde{L}_d^E(q_i, q_{i+i}, \tilde{h}) &= \int_0^{\tilde{h}} \tilde{L}(q(s)q_s(s))\mathrm{d}s \\
&= \int_0^{\tilde{h}} L(q(s), q_t(s)) + H_0\sigma(q)\mathrm{d}s \\
&= \int_0^{h} L(q(t), q_t(t)) + H_0\mathrm{d}t \\
&= L_d^E(q_i, q_{i+i}, h) + H_0 h
\end{aligned}$$

where $h = \int_0^{\tilde{h}} q(s)\mathrm{d}s$. The H_0 term is constant and does not affect the flow of $\mathbf{v}_H$, so it can be dropped. We propose the following discrete Lagrangian approximating $\tilde{L}$. Let

$$\begin{aligned}
h(q_i, q_{i+i}, \tilde{h}) &\approx \int_0^{\tilde{h}} q(s)\mathrm{d}s \\
L_d^0(q_i, q_{i+i}, h) &\approx L_d^E(q_i, q_{i+i}, h) \\
\tilde{L}_d^0(q_i, q_{i+i}, \tilde{h}) &= L_d^0(q_i, q_{i+i}, h(q_i, q_{i+i}, \tilde{h})).
\end{aligned}$$

Finally we invariantize $\tilde{L}_d^0(q_i, q_{i+i}, \tilde{h})$ as

$$\tilde{L}_d = \tilde{L}_d^0(\rho \cdot q_i, \rho \cdot q_{i+i}, \tilde{h})$$

where $\rho = \rho(q_i, q_{i+1})$.

Now, applying the discrete Noether's theorem, we see that $\langle p, \mathbf{w}\rangle$ is preserved along solutions of the discrete Euler–Lagrange equations (3.6). Since the flows of K and H only coincide on a submanifold of T^*Q, the first integral of the discrete system, $\langle p, \mathbf{w}\rangle$, is generally not equal to the first integral of the continuous system, but only coincides with it on the submanifold. However, the method is symplectic, and for small step sizes, it can be expected that the numerical solution will lie close to this submanifold. Thus, it should be expected that the first integral of the continuous system is at least approximately preserved, although we have not done any rigorous analysis on this.

[1] $\mathbf{pr}_s\,\mathbf{w}$ is the prolongation of $\mathbf{w}$ with s taken as the independent variable.

CHAPTER 4

THE RUNGE–LENZ VECTOR

This chapter started out of the wish to demonstrate the technique described in chapter 3. However, most immediate examples are rather trivial. Lagrangians of the form $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q} - U(q)$ can only have affine symmetries. Furthermore, affine symmetries are preserved for all variational integrators which are based on discrete Lagrangians of the form

$$L_d(q_1, q_2) = \sum_i b_i L(Q_i, \dot{Q}_i).$$

The focus then turned to symmetries for more exotic Lagrangians, and specifically symmetries of the Kepler problem *in momentum space*. These symmetries correspond to components of the *Runge–Lenz vector* which is preserved along trajectories of the Kepler problem. Guillemin and Sternberg has written an excellent book on the Runge–Lenz vector and consequences of the corresponding symmetry for both the classical and quantum Kepler problem [5]. We reproduce the results most relevant to us here.

4.1 The Kepler Problem

We consider the Kepler problem in two dimensions, which is a mechanical problem with Lagrangian

$$L = \frac{1}{2}\|\dot{q}\|^2 + \frac{1}{\|q\|},$$

where $\|\cdot\|$ is the standard 2-norm on $\mathbb{R}^2$. The corresponding Hamiltonian is

$$H = \frac{1}{2}\|p\|^2 - \frac{1}{\|q\|}.$$

with $p = \dot{q}$.

We write

$$q = (q^1, q^2)$$
$$p = (p^1, p^2).$$

The Euler–Lagrange equations of L are the well known equations of motion

$$\ddot{q}^i + \frac{q^i}{\|q\|^3} = 0, \qquad i = 1, 2. \tag{4.1}$$

The first integrals of the Kepler problem

$$E = H(q, p) = \frac{1}{2}\|\dot{p}\|^2 - \frac{1}{\|q\|}$$
$$S = q^1 p^2 - q^2 p^1,$$

which are the energy and the angular momentum, are well known, and corresponds to the variational symmetries generated by

$$\mathbf{v}_E = \partial_t,$$
$$\mathbf{v}_S = x\partial_y - y\partial_x,$$

respectively.

Additionally, the components of the *Runge–Lenz vector* $\mathbf{R} = (R^1, R^2)$

$$\begin{aligned} R^1 &= (p^2)^2 q^1 - p^1 p^2 q^2 - \frac{q^1}{\|q\|} \\ R^2 &= (p^1)^2 q^2 - p^1 p^2 q^1 - \frac{q^2}{\|q\|} \end{aligned} \tag{4.2}$$

are first integrals. We note that the first integrals are not functionally independent, as $(R^1)^2 + (R^2)^2 = 1 + 2ES^2$.

A comparison with the formula for first integral arising under Noether's theorem shows that the components of the Runge–Lenz vector cannot arise from a normal variational symmetry. They do arise from the generalized symmetries with generators [12, Example 5.59]

$$\mathbf{v}_{R^1} = -q^2\dot{q}^2\partial_{q^1} + (2q^1\dot{q}^2 - q^2\dot{q}^1)\partial_{q^2}$$
$$\mathbf{v}_{R^2} = (2q^2\dot{q}^1 - q^1\dot{q}^2)\partial_{q^1} - q^1\dot{q}^1\partial_{q^2},$$

but these cannot directly be used for invariantizing a discrete Lagrangian $L(q_1, q_2)$, as the action can not be calculated without knowledge of the derivatives.

The Euler–Lagrange equations (4.1) have one more symmetry, the scaling symmetry generated by

$$\mathbf{v}_3 = 3t\partial_t + 2q^1\partial_{q^1} + 2q^2\partial_{q^2}$$

and it is in fact possible to derive the preservation of the Runge–Lenz vector from this symmetry [15]. However, it is not a variational symmetry, and cannot be used for invariantizing discrete Lagrangians. Instead, we note that the prolonged action of this symmetry is

$$\Psi_\epsilon \begin{pmatrix} t \\ q \\ p \end{pmatrix} = \begin{pmatrix} \exp(3\epsilon)t \\ \exp(2\epsilon)q \\ \exp(-\epsilon)p \end{pmatrix} = \begin{pmatrix} \lambda^3 t \\ \lambda^2 q \\ \lambda^{-1} p \end{pmatrix} \tag{4.3}$$

where $\lambda = \exp(\epsilon)$, and we have identified $p = \dot{q}$. We note that

$$H \circ \Psi = \lambda^{-2} H.$$

This symmetry means that we can map between solutions with different energy H. Considering only bound states with $H < 0$, we can concentrate on the Kepler problem restricted to the submanifold defined by $H = -\frac{1}{2}$ without loss of generality.

We change coordinates on the phase space $T^*\mathbb{R}^2$ to $w = p, \xi = -q$, and consider w to be the position and ξ to be the momentum. The symplectic form in the new coordinates space is $\tilde{\Omega} = \sum_i \mathrm{d}w^i \wedge \mathrm{d}\xi^i$. The coordinate change is easily seen to be a symplectic transformation. In the new coordinates,

$$H(w, \xi) = \frac{1}{2} \|w\|^2 - \frac{1}{\|\xi\|} \tag{4.4}$$

and the corresponding Lagrangian is

$$L(w, \dot{w}) = 2\sqrt{\|\dot{w}\|} - \frac{1}{2} \|w\|^2 \tag{4.5}$$

with $\dot{w} = \frac{\xi}{\|\xi\|^3}$.

We will show that the Hamiltonian vector field on the submanifold with $H = -\frac{1}{2}$ is related to geodesic flow on the two-sphere. The integral curves of the vector field and the stereographic projections of geodesic curves on the sphere are reparametrizations of each other.

4.2 Stereographic Projection of $\mathbb{S}^2$

Consider the unit two-sphere, which we embed in $\mathbb{R}^3$ as

$$\mathbb{S}^2 = \left\{ y = (y^0, y^1, y^2) \in \mathbb{R}^3 \mid \|y\| = 1 \right\}.$$

Let $N = (1, 0, 0) \in \mathbb{S}^2$ be the "north pole" and $\mathbb{S}^2_N = \mathbb{S}^2 \setminus N$ The *stereographic projection* away from N is

$$P : \mathbb{S}^2_N \to \mathbb{R}^2$$
$$P : y \mapsto w$$

with coordinate expression

$$w^i = \frac{y^i}{1 - y^0}, \qquad i = 1, 2. \tag{4.6}$$

The inverse $P^{-1} : w \mapsto y$ has the expression

$$\begin{aligned} y^0 &= \frac{\|w\|^2 - 1}{\|w\|^2 + 1} \\ y^i &= \frac{2w^i}{\|w\|^2 + 1}, \qquad i = 1, 2 \end{aligned} \tag{4.7}$$

We can identify $T^*\mathbb{S}^2$ with the $(y, \eta) \in T^*\mathbb{R}^3 \simeq \mathbb{R}^3 \oplus \mathbb{R}^3$ which satisfies $\|y\| = 1$ and $y^i\eta^i = 0$. With these coordinates the canonical one-form on $T^*\mathbb{S}^2$ is the restriction of the canonical one-form $\eta^i \mathrm{d}y^i$ on $T^*\mathbb{R}^3$. Cotangent lifts of diffeomorphisms between the configuration manifolds are special symplectic. [9, Proposition 6.3.2]. Therefore, the cotangent lift of P carries canonical one-forms to canonical one-forms and we can calculate $T^*P(\xi) = \eta$ and its inverse through the identity

$$\eta^i \mathrm{d}y^i = \xi^i \mathrm{d}w^i.$$

By inserting $\mathrm{d}y^0 = \frac{4(w^1\mathrm{d}w^1 + w^2\mathrm{d}w^2)}{(\|w\|^2+1)^2}$ etc. into the identity above, and then using (4.6) to replace the resulting expressions in w, and simplifying by $y^i\eta^i = 0$, we arrive at the formula

$$\xi^i = y^i\eta^0 + (1 - y^0)\eta^i, \qquad i = 1, 2.$$

And, inverted

$$\begin{aligned} \eta^0 &= w^j\xi^j \\ \eta^i &= \frac{1}{2}\left(1 + \|w\|^2\right)\xi^i - (w^j\xi^j)w^i, \qquad i = 1, 2. \end{aligned}$$

4.2.1 $H = -\frac{1}{2}$

Geodesic flow on the sphere is defined by the Hamiltonian

$$G(y, \eta) = \frac{1}{2}\|\eta\|^2$$

which is mapped to the Hamiltonian

$$K(w, \xi) = \frac{1}{8}(1 + \|w\|^2)^2 \|\xi\|^2 .$$

If u is a real function of one variable, the Hamiltonian vector fields corresponding to by K and $u(K)$ are related through $\mathbf{v}_{u(K)} = u'(K)\mathbf{v}_K$, i.e., their integral curves are reparametrizations of each other. With the Hamiltonian $J = u(K) = \sqrt{2K} - 1$, the Hamiltonian vector field

$$\mathbf{v}_J = \frac{1}{\sqrt{2K}} v_K.$$

On the submanifold $K = \frac{1}{2}$, which corresponds to geodesic flow with $\|\eta\| = 1$, the vector fields $\mathbf{v}_J$ and $\mathbf{v}_K$ coincide. Now

$$\begin{aligned} J(w, \xi) &= \frac{(\|w\|^2 + 1)\,\|\xi\|}{2} - 1 \\ &= \|\xi\| \left(H(w, \xi) + \frac{1}{2} \right). \end{aligned} \tag{4.8}$$

And on the submanifold defined by $H = -\frac{1}{2}$ or, equivalently $J = 0$ we have

$$\mathrm{d}J = \|\xi\|\, \mathrm{d}H$$

so on this submanifold

$$\mathbf{v}_J = \|\xi\|\, \mathbf{v}_H = \frac{2}{1 + \|w\|^2} \mathbf{v}_H.$$

The integral curves of J, which are projections of geodesic curves on $\mathbb{S}^2$, are reparametrizations of solutions to the Kepler problem with $H = -\frac{1}{2}$.

We can write solutions to the Kepler problem with $H = -\frac{1}{2}$ as $(t(s), w(s))$, where $w(s) = P \circ y(s)$ and $y(s)$ is a geodesic curve on $\mathbb{S}^2$ with $\|y'(s)\| = 1$, and $t(s)$ satisfies the differential equation

$$t_s = \frac{2}{1 + \|w(s)\|^2}.$$

Rotations of the sphere clearly leaves G invariant. The vector fields corresponding to rotations about each of the three axis y^0, y^1, y^2 have corresponding Hamiltonians

$$\begin{aligned} \sigma^0 &= y^1\eta^2 - y^2\eta^1 \\ \sigma^1 &= y^2\eta^0 - y^0\eta^2 \\ \sigma^2 &= y^0\eta^1 - y^1\eta^0. \end{aligned}$$

These quantities are preserved along the integral curves of G. Their projection onto $\mathbb{R}^2$ are thus preserved along the integral curves of K, and also those of J and H on the submanifold $J = 0$, $H = -\frac{1}{2}$. The projections are, respectively

$$\begin{aligned} \tau^0 &= w^1\xi^2 - w^2\xi^1 \\ \tau^1 &= \frac{1}{2}(1 - \|w\|^2)\xi^2 + (\xi^i w^i)w^2 \\ &= \frac{1}{2}(1 - (w^1)^2 + (w^2)^2)\xi^2 + w^1 w^2 \xi^1 \\ \tau^2 &= -\frac{1}{2}(1 - \|w\|^2)\xi^1 - (\xi^i w^i)w^1 \\ &= -\frac{1}{2}(1 + (w^1)^2 - (w^2)^2)\xi^1 - w^1 w^2 \xi^2. \end{aligned} \tag{4.9}$$

τ^0 is simply angular momentum, while τ^1 and τ^2 are, modulo a constant rotation, equal to the Runge–Lenz vector when $H = \frac{1}{2}\|w\|^2 - \frac{1}{\|\xi\|} = -\frac{1}{2}$. In the coordinates (w, ξ), we have for R as defined in (4.2),

$$\begin{aligned} R^1 &= \left(\frac{1}{\|\xi\|} - (w^2)^2\right)\xi^1 + w^1 w^2 \xi^2 \\ &= \frac{1}{2}\left(1 + \|w\|^2 - 2(w^2)^2\right)\xi^1 + w^1 w^2 \xi^2 \\ &= -\tau^2, \end{aligned}$$

and similarly $R^2 = \tau^1$.

4.2.2 $H < 0$

For general $H < 0$, assume that $(q(t), p(t))$ is a solution to the Kepler problem with $H < 0$. Using the symmetry (4.3) with $\lambda = \sqrt{-2H}$ transforms this to a solution $(\tilde{q}(\tilde{t}), \tilde{p}(\tilde{t}))$, where

$$\begin{aligned} \tilde{q} &= \lambda^2 q \\ \tilde{p} &= \lambda^{-1} p \\ \tilde{t} &= \lambda^3 t \end{aligned}$$

The transformed solution has $H = -\frac{1}{2}$. Setting $w = \tilde{p}, \xi = -\tilde{q}$, the τ^i from (4.9) are preserved. Their expression in terms of the original solution is

$$\begin{aligned} \tau^1 &= \frac{1}{2}\left(\lambda^{-2}\|p\|^2 - 1\right)\lambda^2 q^2 - (q^i p^i)p^2 \\ &= \frac{1}{2}\left(\|p\|^2 + 2H\right)q^2 - (q^i p^i)p^2 \\ &= \left(\|p\|^2 - \frac{1}{\|q\|}\right)q^2 - (q^i p^i)p^2, \end{aligned}$$

which is equal to R^2 as defined in (4.2). Similarly $\tau^2 = -R^1$.

4.3 Numerical Approximation

We will show how to define a variational integrator of the Kepler problem which preserves the symmetry corresponding to rotating the sphere around the y^1 axis, and preserving an approximation to $R^2 = \left(\frac{1}{\|\xi\|} - (w^1)^2\right)\xi^2 + w^1 w^2 \xi^1$. Through the scaling (4.3), we assume that the initial data $w_0 = p_0, \xi_0 = -q_0$ are such that $H(w_0, \xi_0) = -\frac{1}{2}$.

The rotation of the sphere

$$\Psi_\theta^{\mathbb{S}}\begin{pmatrix} y^0 \\ y^1 \\ y^2 \end{pmatrix} = \begin{pmatrix} \cos\theta y^0 + \sin\theta y^2 \\ y^1 \\ -\sin\theta y^0 + \cos\theta y^2 \end{pmatrix}$$

has stereographic projection in $\mathbb{R}^2$

$$\Psi_\theta = P \circ \Psi_\theta^{\mathbb{S}} \circ P^{-1}$$

$$\Psi_\theta \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} = \begin{pmatrix} \frac{2w^1}{1+\|w\|^2 - \cos\theta(\|w\|^2-1) - 2\sin\theta w^2} \\ \frac{-\sin\theta(\|w\|^2-1)+2\cos w^2}{1+\|w\|^2-\cos\theta(\|w\|^2-1)-2\sin\theta w^2}. \end{pmatrix} \tag{4.10}$$

The generating vector field is

$$\mathbf{v} = w^1 w^2 \partial_{w^1} + \frac{1}{2}\left(1 + (w^2)^2 - (w^1)^2\right)\partial_{w^2}. \tag{4.11}$$

This is not a variational symmetry of the Lagrangian defined in (4.5). It is however a symmetry in the Hamiltonian sense of J and of

$$J'(w,\xi) = \frac{2}{1+\|w\|^2}\left(H + \frac{1}{2}\right).$$

which means we are in the setting described in the last part of section 3.7 with $\sigma(w) = \frac{2}{1+\|w\|^2}$.

We prolong the group action to act on $\mathbb{R}^2 \times \mathbb{R}^2$, and denote an element thereof as (w_1, w_2). We define the moving frame by the cross-section

$$Z = (w_1^1)^2 + (w_1^2)^2 - 1 = 0 \tag{4.12}$$

(corresponding to $y^0 = 0$) with the additional requirement $w_1^2 > 0$. The formula for the moving frame is

$$\rho(w_1, w_2) = -\operatorname{atan2}(\|w_1\|^2 - 1, 2w_1^2). \tag{4.13}$$

This is not a global frame, as it is not defined when $w_1 = (\pm 1, 0)$. The unprolonged action has $(\pm 1, 0)$ as fixed points, so it is impossible to define a moving frame on $\mathbb{R}^2$ near these points. It might be possible to define a global frame for the prolonged action by using equations depending on the w_2 for the cross-section, though this has not been done.

Discretize the Lagrangian of J' by setting

$$\begin{aligned} h(w_1, w_2, \tilde{h}) &= \tilde{h}(\sigma(w_1) \\ &= \frac{2\tilde{h}}{1+\|w_1\|^2} \\ L_d^0(w_1, w_2, h) &= \frac{h}{2}\left(L\left(w_1, \frac{w_2 - w_1}{h}\right) + L\left(w_2, \frac{w_2 - w_1}{h}\right)\right) \\ &= 2\sqrt{h}\,\|w_2 - w_1\|^{\frac{1}{2}} - \frac{h}{4}\left(\|w_1\|^2 + \|w_2\|^2\right) \\ \tilde{L}_d^0(w_1, w_2, \tilde{h}) &= L_d^0\left(w_1, w_2, h(w_1, w_2, \tilde{h})\right). \end{aligned}$$

and

$$\tilde{L}_d(w_1, w_2, \tilde{h}) = L_d^0\left(\rho \cdot w_1, \rho \cdot w_2, \tilde{h}\right)$$

The first order approximation to h is to ensure that $\tilde{h} = h$ on the cross section.

We note that $\tilde{L}_d$ should not be calculated explicitly, as the needed tangent maps are given by (3.12).

Setting $z = (w_1^1, w_1^2, w_2^1, w_2^2)$, we calculate the matrices referred to in (2.4) and (3.12).

$$\boldsymbol{\psi} = \begin{pmatrix} I_1^1 I_1^2 \\ \frac{1}{2}\left(1 + (I_1^2)^2 - (I_1^1)^2\right) \\ I_2^1 I_2^2 \\ \frac{1}{2}\left(1 + (I_2^2)^2 - (I_2^1)^2\right) \end{pmatrix}$$

where $I_j^i = \rho(w_1, w_2) \cdot w_j^i$, is obtained by evaluating (4.11) at the cross-section.

$$\boldsymbol{J} = \begin{pmatrix} 2I_1^1 & 2I_1^2 & 0 & 0 \end{pmatrix}$$

is the invariantization of the Jacobian of Z (4.12).

$$T_z \Psi_\theta = \begin{pmatrix} \boldsymbol{T}^1 & 0 \\ 0 & \boldsymbol{T}^2 \end{pmatrix},$$

where $\boldsymbol{T}^i$ are the Jacobians of the map (4.10) applied to w_i, with fixed theta. These expressions are rather lengthy, but trivial to compute, and are not written out here. According to (2.4), we then have

$$T_z \iota = T_z \Psi_{\rho(z)} - \boldsymbol{\psi}\left(\boldsymbol{J}\boldsymbol{\psi}\right)^{-1} T_z \Psi_{\rho(z)}$$

with $\rho(z)$ given by (4.13).

$$\begin{aligned} T_{\rho\cdot z}\tilde{L}_d^0 &= T_{\rho\cdot z} L_d^0 + \frac{\partial L_d^0}{\partial h}\bigg|_{h=\tilde{h}} T_{\rho\cdot z} h \\ &= -\left[-\frac{\sqrt{\tilde{h}}}{\|I_2 - I_1\|^{\frac{3}{2}}}(I_2 - I_1) - \frac{\tilde{h}}{2} I_1 \quad \frac{\sqrt{\tilde{h}}}{\|I_2 - I_1\|^{\frac{3}{2}}}(I_2 - I_1) - \frac{\tilde{h}}{2} I_2 \right] \\ &\quad + \left(\tilde{h}^{-\frac{1}{2}} \|I_2 - I_1\|^{\frac{1}{2}} - \frac{1}{4}(1 + \|I_2\|^2) \right) \begin{bmatrix} -\tilde{h} I_1 & 0 \end{bmatrix} \end{aligned}$$

where I_i are the row vectors (I_i^1, I_i^2). Finally, the symplectic map is defined by

$$\begin{pmatrix} -\xi^1 & \xi^2 \end{pmatrix} = T_{\rho\cdot z}\tilde{L}_d^0 \circ T_z \iota.$$

where ξ^i are the row vectors (ξ_1^i, ξ_2^i).

By the discrete Noether's theorem,

$$\langle \xi, \mathbf{v} \rangle = w_1 w_2 \xi_1 + \frac{1}{2}\left(1 + (w^2)^2 - (w^1)^2\right)\xi_2$$

is preserved along the solutions. By replacing $\frac{1}{\|\xi\|} = \frac{1}{2}\|w\|^2 - H$ in the expression for R^2, we see that

$$\langle \xi, \mathbf{v} \rangle = R^2 + \left(H + \frac{1}{2}\right)\xi^2.$$

So R^2 is "as preserved as" H. The method is a symplectic method applied to the Hamiltonian problem with Hamiltonian J', so one would expect near preservation of J', which with the initial data specified corresponds to near preservation of H.

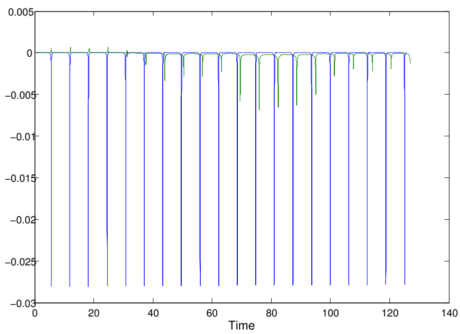
4.4 Numerical Tests

The algorithm described above was tested for initial data $q(0) = -\xi(0) = (0.8, 0.6)$, $p(0) = w(0) = (0.96, 0.28)$, which has for the four first integrals

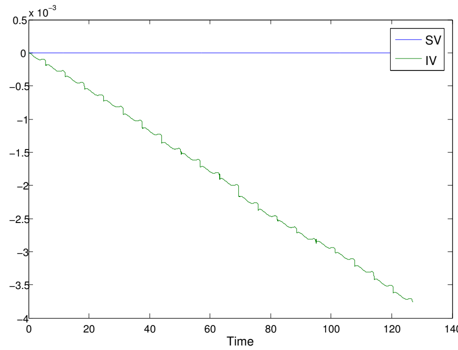
$$\begin{aligned} H(0) &= -0.5 \\ S(0) &= 0.3520 \\ R^1(0) &= -0.26208 \\ R^2(0) &= -0.89856 \end{aligned}$$

and is on a highly eccentric orbit. The algorithm was tested against a simple Störmer–Verlet scheme, and with a Störmer–Verlet scheme with variable step size as described in [6, Chapter VIII] This step size was set to $\|q\|\,\tilde{h}$ to correspond to the step size $h = \frac{2\tilde{h}}{1+\|w\|^2}$ used in the invariant variational (IV) scheme. The other parameters used was $\tilde{h} = \frac{2\pi}{200}$, corresponding to approximately 200 steps per orbit, and the number of steps taken was 4000, or approximately 20 orbits. For this step-size, the simple Störmer–Verlet fails, and the solution of this scheme quickly falls into a hyperbolic trajectory escaping the central force.

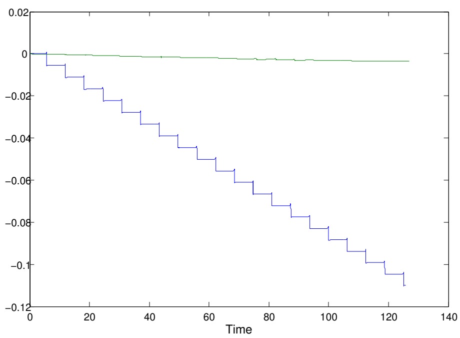
Figure 4.1 shows the error in the first integrals H, S, R^1, R^2 for the IV integrator and Störmer–Verlet with varying step size (SV). As can be seen in the plots, the IV scheme is superior in preserving R^2, as it is designed to. It also compares favourably to Störmer–Verlet with variable step size for preservation of H. On the other hand, while both Störmer–Verlet schemes preserve the angular momentum S up to machine precision, the IV scheme does not preserve this integral, as the invariantization with respect to the group action (4.10) breaks the rotational symmetry.



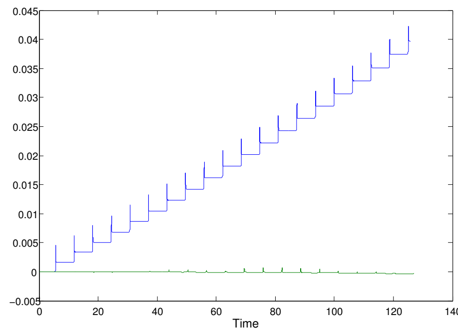
(a) Error in H



(b) Error in S



(c) Error in R^1



(d) Error in R^2

Figure 4.1: Error in the first integrals H, S, R^1, R^2 for the SV and IV schemes

CHAPTER 5

INVARIANT DISCRETE EULER–LAGRANGE EQUATIONS

The theory of Chapters 3 and 4 was aimed at creating numerical integrators for invariant variational problems in the original coordinates. Furthermore, it was limited to first order Lagrangians. While mechanical Lagrangians are usually of this form, Lagrangians of higher order, i.e. depending on higher order derivatives may arise in other settings.

Inspired by [8, Chapter 7], we will try and express the Euler–Lagrange equations for invariant variational problems in the invariant coordinates.

5.1 Discrete Euler–Lagrange equations

We consider sequences of points $\{u_i\} = \ldots, u_{-1}, u_0, u_1 \ldots$ in U. A *discrete Lagrangian* of order n is a smooth, real function depending on $n+1$ successive points of the sequence $\{u_i\}$, that is

$$L(u_i, u_{i+1}, \cdots u_{i+n}).$$

Define the *shifting operator* $\mathbf{S}$ as

$$\mathbf{S} : u_i \mapsto u_{i+1},$$

and the *discrete prolongation* $\mathbf{pr}^n$ as such: If $\{u_i\}$ is a sequence of points in U, its n'th prolongation $\{\mathbf{pr}^n\, u_i\}$ is a sequence in U^{n+1} defined by

$$\begin{aligned}\mathbf{pr}^n\, u_i &= (u_i, u_{i+1}, \ldots, u_{i+n}) \\ &= (u_i, \mathbf{S}u_i, \ldots, \mathbf{S}^n u_i)\end{aligned}$$

We write the discrete Lagrangian

$$L(u, \mathbf{S}u, \dots, \mathbf{S}^n u) = L(\mathbf{pr}^n\, u).$$

To simplify notation, we will usually let the order of prolongation be arbitrary and write simply **pr**.

The discrete variational problem is to find sequences of points $\{u_i\}$ which locally extremize the action sum

$$\mathcal{A}(\{u_i\}) = \sum_i L(\mathbf{pr}\, u_i). \tag{5.1}$$

Equations for the solution of the discrete variational problem (5.1) can be found by introducing the variation variable τ and the arbitrary variation $(u_\tau)_i = (\partial_t u)_i$. A sequence extremizing the sum should satisfy

$$\begin{aligned} 0 &= \partial_\tau \mathcal{A} \\ &= \sum_i \partial_\tau L(\mathbf{pr}\, u_i) \\ &= \sum_i \frac{\partial L}{\partial u_i}(u_\tau)_i + \frac{\partial L}{\partial \mathbf{S}u_i}(\mathbf{S}u_\tau)_i + \cdots \end{aligned}$$

We view $\{L_i\} = \{L(\mathbf{pr}\, u_i)\}$, and its partial derivatives as a sequence indexed by i. Collecting in terms of $(u_\tau)_i$, and using that these are arbitrary yields the *discrete Euler–Lagrange equations*. In local coordinates $u = (u^1, \dots, u^m)$, the discrete Euler–Lagrange equation corresponding to u^α is

$$E^\alpha L = \frac{\partial L}{\partial u^\alpha} + \mathbf{S}^{-1}\left[\frac{\partial L}{\partial \mathbf{S}u^\alpha}\right] + \cdots = 0.$$

We take the equation above to define the *Euler–Lagrange operator* with respect to u^α.

5.2 Correction Terms and Correction Elements

For discrete Lagrangians depending on N points u_i, the Euler–Lagrange equations will typically depend on $2N-1$ points. We will show that if the discrete Lagrangian L is invariant, then one can express the Euler–Lagrange equations invariants of the action, such that they depend on fewer arguments. The derivation is reminiscent of Mansfield treatment of the continuous case in [8, Chapter 7], but as the points are discrete, the shifting operator $\mathbf{S}$ takes the place of the differential operator $\mathcal{D}_x = \partial_x$.

Assume that the group action $\Psi : G \times U \to U$ is such that the discrete Lagrangian

$$L(\mathbf{pr}\, u) = L(u, \mathbf{S}u \dots,)$$

is invariant under the prolonged action, and that the action allows a moving frame $\rho : U^l \to G$ for some l. By the replacement theorem, L can be written as a function of the fundamental invariants of the prolonged action.

For simplicity, we assume that the moving frame is defined by the cross-section equations

$$Z_{j,\alpha} = \mathbf{S}^j u^\alpha = k, \quad (j, \alpha) \in \mathcal{E},$$

where $\mathcal{E}$ is such that $(j > 0 \wedge (\alpha, j) \in \mathcal{E}) \implies \forall \beta, \quad (\beta, j-1) \in \mathcal{E}$. Or in layman's terms, the cross-section is defined by using as few shift operators as possible.

We prolong the action to TU^k in the normal way, and use coordinates

$$u_J^\alpha = (\partial_\tau)^{j_2} \mathbf{S}^{j_1} u^\alpha$$

where

$$J = (\underbrace{1, \ldots, 1}_{j_1}, \underbrace{2, \ldots, 2}_{j_2}),$$

and the fundamental invariants

$$I_J^\alpha = \iota\left(u_J^\alpha\right)$$

We define the right group action on functions $F : U^n \to \mathbb{R}$ by

$$(F \cdot g)(\mathbf{pr}\, u) = F(g \cdot \mathbf{pr}\, u)$$

which is a right action, and invariantization of functions

$$\iota(F)(\mathbf{pr}\, u) = (F \cdot \rho(\mathbf{pr}\, u))(\mathbf{pr}\, u).$$

The operators $\mathbf{S}$ and ∂_τ are invariant and commute, but do not commute with the invariantization operator ι. In order to calculate the invariant discrete Euler–Lagrange-equations, we will need the commutator relations between ι and $\mathbf{S}$, and between ι and ∂_τ.

5.2.1 Correction Terms for Differentiation

For the differential operator ∂_τ, this is exactly the correction terms calculated in Chapter 2.

As examples, and for use later, we calculate the correction terms for two group actions $\Psi : SL(2) \times \mathbb{R} \to \mathbb{R}$, and $\Psi : SE(2) \times \mathbb{R}^2 \to \mathbb{R}^2$.

$SL(2)$

We represent $SL(2)$ as 2×2 real matrices $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with determinant 1 and consider the $SL(2)$ action on $\mathbb{R}$ defined by

$$\Psi\left(\boldsymbol{A}, u\right) = \frac{au + b}{cu + d}, \tag{5.2}$$

which we prolong to $T\mathbb{R}$ in the usual way.

$$\Psi\left(\boldsymbol{A}; u, u_\tau\right) = \left(\frac{au+b}{cu+d}, \frac{u_\tau}{(cu+d)^2}\right)$$

Using (a, b, c) to parametrise $SL(2)$ near the identity, the infinitesimal generators for the prolonged action

$$\begin{aligned}
\mathbf{v}_a &= 2u\partial_u + 2u_\tau \partial_{u_\tau} \\
\mathbf{v}_b &= \partial_u \\
\mathbf{v}_c &= -u^2\partial_u - 2uu_\tau\partial_{u_\tau}.
\end{aligned}$$

And prolonged to TM^n,

$$\begin{aligned}
\mathbf{v}_a &= \sum_J 2u_J\partial_{u_J} + 2u_{J2}\partial_{u_{J2}} \\
\mathbf{v}_b &= \sum_J \partial_{u_J} \\
\mathbf{v}_c &= \sum_J - \left(u_J\right)^2 \partial_{u_J} - 2u_J u_{J2}\partial_{u_{J2}},
\end{aligned}$$

where the sums go over J consisting of only 1's.

Using the cross-section defined by

$$\begin{aligned}
Z_1(g \cdot \mathbf{pr}\, u) &= g \cdot u = -1 \\
Z_2(g \cdot \mathbf{pr}\, u) &= g \cdot \mathbf{S}u = 0 \\
Z_3(g \cdot \mathbf{pr}\, u) &= g \cdot \mathbf{S}^2 u = 1,
\end{aligned}$$

and the notations defined in section 2.4, $\boldsymbol{J'} = \boldsymbol{I}$ and

$$\boldsymbol{\psi}^{\boldsymbol{\zeta}} = \begin{array}{c} \\ I_0 \\ I_1 \\ I_{11} \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 2I_0 & 1 & -I_0^2 \\ 2I_1 & 1 & -I_1^2 \\ 2I_{11} & 1 & -I_{11}^2 \end{pmatrix} \end{array} = \begin{array}{c} \\ I_0 \\ I_1 \\ I_{11} \end{array} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} -2 & 1 & -1 \\ 0 & 1 & 0 \\ 2 & 1 & -1 \end{pmatrix} \end{array}$$

$$\boldsymbol{T} = \begin{array}{c} \\ I_0 \\ \mathbf{S}I_1 \\ \mathbf{S}^2 I_{11} \end{array} \begin{array}{c} \tau \\ \begin{pmatrix} I_2 \\ I_{12} \\ I_{112} \end{pmatrix} \end{array}.$$

The correction matrix is

$$\boldsymbol{C} = -\boldsymbol{\psi}^{-1}\boldsymbol{T} = \begin{pmatrix} \frac{1}{4}(I_2 - I_{112}) \\ -I_{12} \\ \frac{1}{2}(I_{112} - 2I_{12} + I_2) \end{pmatrix}.$$

And the general formula for fundamental invariants is

$$\partial_\tau I_J = I_{J2} + \boldsymbol{C}_i \psi_J^i(\rho(\mathbf{pr}\, u) \cdot \mathbf{pr}\, u)$$

And for integral invariants $\iota(\mathbf{S}^k u)$, where J consists of only 1's,

$$\partial_\tau I_J = I_{J2}\frac{1}{2}(I_{112} - I_2)I_J - I_{12} - \frac{1}{2}(I_{112} - 2I_{12} + I_2)I_J^2. \tag{5.3}$$

$SE(2)$

We consider an $SE(2)$ action on $U = \mathbb{R}^2$. We write $u = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ for elements of $\mathbb{R}^2$. The action is defined by

$$\Psi_{(\theta,a,b)} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x-a \\ y-b \end{pmatrix}.$$

The infinitesimal generators for the prolonged action on TU^k are

$$\begin{aligned} \mathbf{v}_a &= \sum_J (-\partial_{x_J}) \\ \mathbf{v}_b &= \sum_J (-\partial_{y_J}) \\ \mathbf{v}_\theta &= y_J \partial_{x_J} - x_J \partial_{y_J} + y_{J2}\partial_{x_{J2}} - x_{J2}\partial_{y_{J2}} \end{aligned}$$

where the sums go over J consisting of only 1's. We use the cross-section defined by

$$\begin{aligned} Z_1(g \cdot \mathbf{pr}\, u) &= g \cdot x = 0 \\ Z_2(g \cdot \mathbf{pr}\, u) &= g \cdot y = 0 \\ Z_3(g \cdot \mathbf{pr}\, u) &= g \cdot \mathbf{S}y = 0, \end{aligned}$$

with the additional requirement $g \cdot \mathbf{S}x \geq 0$.

Once more $\boldsymbol{J'}$ is the identity, and

$$\boldsymbol{\psi}^{\boldsymbol{\zeta}} = \begin{array}{c} \\ I^x \\ I^y \\ I^y_1 \end{array} \begin{array}{c} \begin{array}{ccc} a & b & \theta \end{array} \\ \begin{pmatrix} -1 & 0 & I^y \\ 0 & -1 & -I^x \\ 0 & 0 & -I^x_1 \end{pmatrix} \end{array} = \begin{array}{c} \\ I^x \\ I^y \\ I^y_1 \end{array} \begin{array}{c} \begin{array}{ccc} a & b & \theta \end{array} \\ \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & -I^x_1 \end{pmatrix} \end{array}.$$

$$\boldsymbol{T} = \begin{array}{c} I^x \\ I^y \\ I^y_1 \end{array} \begin{pmatrix} I^x_2 \\ I^y_2 \\ I^y_{12} \end{pmatrix}.$$

The correction matrix is therefore

$$\boldsymbol{C} = -\boldsymbol{\psi}^{-1}\boldsymbol{T} = \begin{pmatrix} I^x_2 \\ I^y_2 \\ \frac{I^y_{12} - I^y_2}{I^x_1} \end{pmatrix},$$

and the general formula for integral invariants is

$$\begin{aligned} \partial_\tau I^x_J &= I^x_{J2} + \boldsymbol{C}_i \psi^i_{x,J}(\rho(\mathbf{pr}\, u) \cdot \mathbf{pr}\, u) \\ &= I^x_{J2} - I^x_2 + \frac{I^y_{12} - I^y_2}{I^x_1} I^y_J \\ \partial_\tau I^y_J &= I^y_{J2} + \boldsymbol{C}_i \psi^i_{y,J}(\rho(\mathbf{pr}\, u) \cdot \mathbf{pr}\, u) \\ &= I^y_{J2} - I^y_2 - \frac{I^y_{12} - I^y_2}{I^x_1} I^x_J. \end{aligned} \tag{5.4}$$

5.2.2 Correction Element for Shifting

The correction terms described in the section above have an analogue for the discrete operator $\mathbf{S}$. The analogue, the *correction group element*, require somewhat more calculation than the correction terms. Specifically, one has to calculate the group action (not only its infinitesimal generators) and one has to solve the frame equations, though only for special cases.

The expressions for shifting, then invariantizing of a coordinate u^α_J and vice versa are, respectively

$$\begin{aligned}\iota(\mathbf{S}(u^\alpha_J)) &= \iota(u^\alpha_{J1})\\ &= \rho(\mathbf{pr}\,u)\cdot u^\alpha_{J1}\\ \mathbf{S}(\iota(u^\alpha_J)) &= \mathbf{S}(\rho(\mathbf{pr}\,u)\cdot u^\alpha_J)\\ &= \rho(\mathbf{S}\,\mathbf{pr}\,u)\cdot u^\alpha_{J1},\end{aligned}$$

where $\mathbf{S}\,\mathbf{pr}\,u = (\mathbf{S}u, \mathbf{S}^2 u, \dots)$. Thus

$$\begin{aligned}\mathbf{S}(\iota(u^\alpha_J)) &= \rho(\mathbf{S}\,\mathbf{pr}\,u)\rho(\mathbf{pr}\,u)^{-1}\cdot\iota(u^\alpha_{J1})\\ \mathbf{S}(I^\alpha_J) &= P(\mathbf{pr}\,u)\cdot I^\alpha_{J1}.\end{aligned} \tag{5.5}$$

Where $P(\mathbf{pr}\,u) = \rho(\mathbf{S}\,\mathbf{pr}\,u)\rho(\mathbf{pr}\,u)^{-1}$ is the correction group element.

We will now show how $P(\mathbf{pr}\,u)$ can be expressed in the fundamental invariants. Furthermore, only the invariantizations of the coordinates appearing in the equations of the cross-sections and the shifts of these are needed. (Contrast with the continuous case, where the invariantizations of those coordinates and the derivatives of these were needed).

Theorem 12.

$$P(\mathbf{pr}\,u) = \rho(\mathbf{S}\,\mathbf{pr}\,u)\rho(\mathbf{pr}\,u)^{-1} = \rho(\iota(\mathbf{S}\,\mathbf{pr}\,u)) \tag{5.6}$$

Where $\rho(\iota(\mathbf{S}\,\mathbf{pr}\,u))$ means ρ evaluated with I^α_{J1} set in for the arguments u^α_J.

Proof. By the moving frame identity

$$\begin{aligned}\rho\left[\iota(\mathbf{S}\,\mathbf{pr}\,u)\right] &= \rho\left[\rho(\mathbf{pr}\,u)\cdot\mathbf{S}\,\mathbf{pr}\,u\right]\\ &= \rho(\mathbf{S}\,\mathbf{pr}\,u)\rho(\mathbf{pr}\,u)^{-1}\\ &= P(\mathbf{pr}\,u)\end{aligned}$$

□

Again, we calculate the correction expressions in the case of the $SL(2)$ and $SE(2)$ actions described above.

$SL(2)$

To simplify computation, we represent $SL(2)$ by 2×2 real matrices with positive determinant, and consider matrices $\boldsymbol{A}$ and $\lambda\boldsymbol{A}$ to represent the same element of $SL(2)$ The action is still

$$\Psi(\boldsymbol{A}, u) = \frac{au+b}{cu+d},$$

but the prolonged action on u_τ has to be modified,

$$\Psi(\boldsymbol{A}; u, u_\tau) = \left(\frac{au+b}{cu+d}, \frac{(ad-bc)u_\tau}{(cu+d)^2} \right).$$

Let the frame be defined as in section 5.2.1 and $\sigma = \iota(\mathbf{S}^3 u)$ be the lowest order integral invariant. To find P, we have to calculate the moving frame for

$$\begin{aligned} \iota(\mathbf{S}\,\mathbf{pr}\,u) &= (I_1, I_{11}, I_{111}) \\ &= (0, 1, \sigma). \end{aligned}$$

That is, solve the equations

$$\begin{aligned} \frac{a \cdot 0 + b}{c \cdot 0 + d} &= -1 \\ \frac{a \cdot 1 + b}{c \cdot 1 + d} &= 0 \\ \frac{a \cdot \sigma + b}{c \cdot \sigma + d} &= 1. \end{aligned}$$

There is a degree of freedom in the scaling of the group parametres, and one solution is

$$\boldsymbol{P} = \begin{pmatrix} \sigma & -\sigma \\ \sigma - 2 & \sigma \end{pmatrix}$$

which has positive determinant as long as $\sigma > 1$. The inverse of $\boldsymbol{P}$ is, when scaling is ignored,

$$\boldsymbol{P}^{-1} = \begin{pmatrix} \sigma & \sigma \\ 2 - \sigma & \sigma \end{pmatrix}.$$

So the relations are, for the integral invariants (J consists of only 1's)

$$\begin{aligned} \mathbf{S}I_J &= \frac{\sigma I_{J1} - \sigma}{(\sigma - 2)I_{J1} + \sigma} \\ I_{J1} &= \frac{\sigma \mathbf{S}I_J + \sigma}{(2-\sigma)\mathbf{S}I_J + \sigma} \end{aligned} \tag{5.7}$$

and for the invariants $\iota(\partial_\tau \mathbf{S}^k(u))$.

$$\begin{aligned} \mathbf{S}I_{J2} &= \frac{2(\sigma^2 - \sigma)I_{J12}}{((\sigma-2)I_{J1} + \sigma)^2} \\ I_{J12} &= \frac{2(\sigma^2 - \sigma)\mathbf{S}I_{J2}}{((2-\sigma)\mathbf{S}I_J + \sigma)^2}. \end{aligned} \tag{5.8}$$

$SE(2)$

For the $SE(2)$ action and cross-section defined in section 5.2.1, let the lowest order invariants be denoted by $\lambda = \iota(\mathbf{S}x)$ and $\sigma = \iota(\mathbf{S}^2 y)$. To find the group element P, we have to solve the moving frame equations

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \lambda - a \\ 0 - b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
$$-\sin\theta(I^x_{11} - \lambda) + \cos\theta\sigma = 0$$

Which has solution

$$a = \lambda$$
$$b = 0$$
$$\theta = \operatorname{atan2}(\sigma, I^x_{11} - \lambda).$$

The relations between shifting and invariantization for integral invariants are thus

$$\begin{pmatrix} \mathbf{S}I^x_K \\ \mathbf{S}I^y_K \end{pmatrix} = \frac{1}{\sqrt{\sigma^2 + (I^x_{11} - \lambda)^2}} \begin{pmatrix} I^x_{11} - \lambda & \sigma \\ -\sigma & I^x_{11} - \lambda \end{pmatrix} \begin{pmatrix} I^x_{K1} - \lambda \\ I^y_{K1} \end{pmatrix} \tag{5.9}$$

I^x_{11} appearing in the equation above is not one of the lowest order invariants and we wish to replace it with an expression in λ and σ. We set

$$K = 1, \quad I^x_1 = \lambda, \quad I^y_1 = 0, \quad I^y_{11} = \sigma$$

in the above formula and calculate

$$\mathbf{S}\lambda = \sqrt{\sigma^2 + (I^x_{11} - \lambda)^2}.$$

This relation, which simply is the Pythagorean theorem, cannot immediately be inverted to give an expression for I^x_{11} in terms of the lowest order invariants, since there is a sign ambiguity in $I^x_{11} - \lambda$. However, assume that the discretized curve has no acute corners, such that the inequality

$$(\mathbf{S}^2 x - \mathbf{S}x)(\mathbf{S}x - x) + (\mathbf{S}^2 y - \mathbf{S}y)(\mathbf{S}y - y) \geq 0 \tag{5.10}$$

holds. The left hand side of (5.10) is easily seen to be invariant under the group action by applying the prolonged vector field. Thus the invariantized inequality $(I^x_{11} - \lambda)\lambda \geq 0$ also holds, and due to the definition of the frame, $\lambda \geq 0$, the sign ambiguity is resolved and

$$I^x_{11} = \lambda + \sqrt{\mathbf{S}\lambda^2 - \sigma^2}.$$

For the invariants of the first prolonged action, the relations are

$$\begin{pmatrix} \mathbf{S}I^x_{J2} \\ \mathbf{S}I^y_{J2} \end{pmatrix} = \frac{1}{\mathbf{S}\lambda} \begin{pmatrix} \sqrt{\mathbf{S}\lambda^2 - \sigma^2} & \sigma \\ -\sigma & \sqrt{\mathbf{S}\lambda^2 - \sigma^2} \end{pmatrix} \begin{pmatrix} I^x_{J12} \\ I^y_{J12} \end{pmatrix}. \tag{5.11}$$

5.3 Invariant Euler–Lagrange Equations

The relations given above can be combined to give the relations $\sigma_\tau = \mathcal{H}I_2$, where $\mathcal{H}$ is an operator depending on the lowest order invariants and their shifts. This procedure is straight forward and we illustrate it with our two examples of $SE(2)$ and $SL(2)$ actions.

$SL(2)$

From the relation (5.3) we have

$$\begin{aligned}\partial_\tau\sigma = \partial_\tau I_{111} &= I_{1112} - \frac{1}{2}(I_{112} - I_2)I_{111} - I_{12} - \frac{1}{2}(I_{112} - 2I_{12} + I_2)I_{111}^2 \\ &= I_{1112} - \frac{1}{2}(\sigma^2 + \sigma)I_{112} + (\sigma^2 - 1)I_{12} - \frac{1}{2}(\sigma^2 - \sigma)I_2.\end{aligned} \tag{5.12}$$

And from the relations (5.8),

$$\begin{aligned}I_{12} &= \frac{1}{2}\frac{\sigma}{\sigma - 1}\mathbf{S}I_2 \\ I_{112} &= 2\frac{\sigma - 1}{\sigma}\mathbf{S}I_{12} \\ &= \frac{\sigma - 1}{\sigma}\mathbf{S}\left[\frac{\sigma}{\sigma - 1}\right]\mathbf{S}^2 I_2 \\ I_{1112} &= \frac{1}{2}\sigma(\sigma - 1)\mathbf{S}I_{112} \\ &= \frac{1}{2}\sigma(\sigma - 1)\mathbf{S}\left[\frac{\sigma - 1}{\sigma}\right]\mathbf{S}^2\left[\frac{\sigma}{\sigma - 1}\right]\mathbf{S}^3 I_2.\end{aligned} \tag{5.13}$$

Combining the equations (5.12) and (5.13), yields the promised equation

$$\begin{aligned}\partial_\tau\sigma = \frac{1}{2}\Bigg[&\sigma(\sigma - 1)\mathbf{S}\left[\frac{\sigma - 1}{\sigma}\right]\mathbf{S}^2\left[\frac{\sigma}{\sigma - 1}\right]\mathbf{S}^3 \\ &-(\sigma + 1)(\sigma - 1)\mathbf{S}\left[\frac{\sigma}{\sigma - 1}\right]\mathbf{S}^2 + (\sigma + 1)\sigma\mathbf{S} - \sigma(\sigma - 1)\Bigg] I_2 \\ = \mathcal{H}I_2.&\end{aligned}$$

The invariantized Euler–Lagrange equation is thus

$$\begin{aligned}\mathcal{H}^* E^\sigma L =&\frac{1}{2}\Bigg[\mathbf{S}^{-1}\left[\frac{\sigma}{\sigma - 1}\right]\mathbf{S}^{-2}\left[\frac{\sigma - 1}{\sigma}\right]\mathbf{S}^{-3}\left[\sigma(\sigma - 1)\right]\mathbf{S}^{-3}E^\sigma L \\ &- \mathbf{S}^{-1}\left[\frac{\sigma}{\sigma - 1}\right]\mathbf{S}^{-2}\left[(\sigma^2 - 1)\right]\mathbf{S}^{-2}E^\sigma L \\ &+ \mathbf{S}^{-1}\left[(\sigma + 1)\sigma\right]\mathbf{S}^{-1}E^\sigma L \\ &- \sigma(\sigma - 1)E^\sigma L\Bigg] = 0.\end{aligned}$$

where E^σ is the discrete Euler–Lagrange operator, i.e.

$$E^\sigma L = \frac{\partial L}{\partial\sigma} + \mathbf{S}^{-1}\left[\frac{\partial L}{\partial\mathbf{S}\sigma}\right] + \cdots$$

$SE(2)$

Combining the equations (5.9) and (5.11), gives, after some algebraic manipulation, the expression

$$\begin{pmatrix} \lambda_\tau \\ \sigma_\tau \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \begin{pmatrix} I_2^x \\ I_2^y \end{pmatrix} \tag{5.14}$$

where

$$\begin{aligned}
\mathbf{H}_{11} &= \frac{I_{11}^x - \lambda}{\mathbf{S}\lambda}\mathbf{S} - \mathbf{1} \\
\mathbf{H}_{12} &= -\frac{\sigma}{\mathbf{S}\lambda}\mathbf{S} \\
\mathbf{H}_{21} &= \frac{1}{\mathbf{S}\lambda\mathbf{S}^2\lambda}\left(\sigma\mathbf{S}(I_{11}^x - \lambda) + (I_{11}^x - \lambda)\mathbf{S}\sigma\right)\mathbf{S}^2 - \frac{I_{11}^x}{\lambda\mathbf{S}\lambda}\sigma\mathbf{S} \\
\mathbf{H}_{22} &= \frac{1}{\mathbf{S}\lambda\mathbf{S}^2\lambda}\left((I_{11}^x - \lambda)\mathbf{S}(I_{11}^x - \lambda) - \sigma\mathbf{S}\sigma\right)\mathbf{S}^2 - \frac{I_{11}^x(I_{11}^x - \lambda)}{\lambda\mathbf{S}\lambda}\mathbf{S} + \frac{I_{11}^x - \lambda}{\lambda}\mathbf{1}.
\end{aligned}$$

The invariant discrete Euler–Lagrange equations are thus

$$\begin{aligned}
\mathbf{H}_{11}^* E^\lambda L + \mathbf{H}_{21}^* E^\sigma L = 0 \\
\mathbf{H}_{12}^* E^\lambda L + \mathbf{H}_{22}^* E^\sigma L = 0
\end{aligned} \tag{5.15}$$

where $\mathbf{H}_{11}^*$ etc. are the adjoint operators

$$\begin{aligned}
\mathbf{H}_{11}^* &= \frac{\mathbf{S}^{-1}(I_{11}^x - \lambda)}{\lambda}\mathbf{S}^{-1} - \mathbf{1} \\
\mathbf{H}_{12}^* &= -\frac{\mathbf{S}^{-1}\sigma}{\lambda}\mathbf{S}^{-1} \\
\mathbf{H}_{21}^* &= \frac{1}{\lambda\mathbf{S}^{-1}\lambda}\left[\mathbf{S}^{-2}\sigma\mathbf{S}^{-1}\left(I_{11}^x - \lambda\right) + \mathbf{S}^{-2}\left(I_{11}^x - \lambda\right)\mathbf{S}^{-1}\sigma\right]\mathbf{S}^{-2} \\
&\quad - \frac{1}{\lambda\mathbf{S}^{-1}\lambda}\mathbf{S}^{-1}\left(I_{11}^x\sigma\right)\mathbf{S}^{-1} \\
\mathbf{H}_{22}^* &= \frac{1}{\lambda\mathbf{S}^{-1}\lambda}\left[\mathbf{S}^{-2}\left[I_{11}^x - \lambda\right]\mathbf{S}^{-1}\left(I_{11}^x - \lambda\right) - \mathbf{S}^{-2}\sigma\mathbf{S}^{-1}\sigma\right]\mathbf{S}^{-2} \\
&\quad - \frac{1}{\lambda\mathbf{S}^{-1}\lambda}\mathbf{S}^{-1}\left[I_{11}^x\left(I_{11}^x - \lambda\right)\right]\mathbf{S}^{-1} + \frac{I_{11}^x - \lambda}{\lambda}\mathbf{1}.
\end{aligned}$$

5.4 The Solution in Original Coordinates

The sections above show that the solutions to discrete invariant variational problems has to satisfy certain difference equations expressed in the invariants of the symmetry. The remaining question is how these difference equations can be used to solve the original problem. One option is of course to replace the invariants with their expressions in the original coordinates, and solve in terms of these. In some cases, however, a more elegant solution is possible.

Assuming that the boundary conditions are of a such nature that the sequence of invariants is uniquely defined, it is possible to first find the invariants and thereafter recover the original variables. Since $u_i = \rho(\mathbf{pr}\, u_i)^{-1} \cdot (I^u)_i$, the original coordinates can be recovered provided we can find the frame in each step. This can be calculated from the frame at any point using the relation

$$\rho(\mathbf{pr}\, u_{i+1}) = P(\mathbf{pr}\, u_i)\rho(\mathbf{pr}\, u_i), \tag{5.16}$$

where we have already shown that $P(\mathbf{pr}\, u_i)$ can be expressed in terms of the invariants.

CHAPTER 6

IMPLEMENTATION AND NUMERICAL TESTS

6.1 $SL(2)$

We consider the $SL(2)$ action described in (5.2). The lowest order non-phantom invariant of the prolonged action is the Schwarzian derivative

$$\eta = \{u; x\} = \frac{u_{xxx}}{u_x} - \frac{3}{2}\left(\frac{u_{xx}}{u_x}\right)^2. \tag{6.1}$$

We use the Lagrangian

$$\int L \, \mathrm{d}x = \int \eta^2 \, \mathrm{d}x. \tag{6.2}$$

We discretize with constant step size and use the discrete prolongation and moving frame of section 5.2.1. The lowest order non-phantom invariant of the discrete prolonged action is

$$\sigma = I_{111} = \iota(\mathbf{S}^3 u).$$

We construct an approximation for the Schwarzian derivative by taking finite difference approximations of the derivatives in (6.1) and then invariantizing

$$\begin{aligned}\kappa &= \frac{1}{h^2}\iota\left[\frac{\mathbf{S}^3 u - 3\mathbf{S}^2 u + 3\mathbf{S}u - u}{\mathbf{S}u - u} - \frac{3}{2}\left(\frac{\mathbf{S}^2 u - 2\mathbf{S}u + u}{\mathbf{S}u - u}\right)^2\right] \\ &= \frac{1}{h^2}(\sigma - 2).\end{aligned} \tag{6.3}$$

The discrete Lagrangian becomes

$$L_d = \kappa^2 h = \frac{1}{h^3}(\sigma - 2)^2$$

and

$$E^\sigma L_d = \frac{2}{h^3}(\sigma - 2).$$

So the invariant discrete Euler–Lagrange equation is

$$\begin{aligned}\frac{1}{h^3}\Bigg(&\frac{\sigma_{i-1}}{\sigma_{i-1}-1}\frac{\sigma_{i-2}-1}{\sigma_{i-2}}\sigma_{i-3}(\sigma_{i-3}-1)(\sigma_{i-3}-2)\\ &-\frac{\sigma_{i-1}}{\sigma_{i-1}-1}(\sigma_{i-2}^2-1)(\sigma_{i-2}-2)\\ &+(\sigma_{i-1}+1)\sigma_{i-1}(\sigma_{i-1}-2)\\ &-\sigma_i(\sigma_i-1)(\sigma_i-2)\Bigg)=0.\end{aligned} \tag{6.4}$$

Given initial values $(\sigma_0, \sigma_1, \sigma_2)$, one can successively calculate $\sigma_3, \sigma_4, \ldots,$ by solving a third-degree equation at each step in the iteration. The equation may have multiple real roots, but at most one of these satisfy $\sigma_i \geq 1 + \frac{\sqrt{3}}{3}$. Since this corresponds to $\kappa_i = \frac{1}{h^2}(\sigma_i - 2) \geq -\frac{3-\sqrt{3}}{3h^2}$, the other roots can be disregarded for small enough h. $\sigma_0, \sigma_1, \sigma_2$ can be calculated from $u_0, \ldots, u_5$ by solving for the moving frame for the first values. The moving frames equations

$$\begin{aligned}\frac{au_0+b}{cu_0+d} &= -1\\ \frac{au_1+b}{cu_1+d} &= 0\\ \frac{au_2+b}{cu_2+d} &= 1\end{aligned}$$

have solution

$$\rho_0 = \rho(u_0, u_1, u_2) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} u_2 - u_0 & -(u_2-u_0)u_1 \\ u_2 - 2u_1 + u_0 & u_2u_1 - 2u_2u_0 + u_1u_0 \end{pmatrix}. \tag{6.5}$$

and

$$\sigma_0 = \rho(u_0, u_1, u_2) \cdot u_3 = \frac{(u_2-u_0)(u_3-u_1)}{(u_3-u_0)(u_2-u_1)-(u_3-u_2)(u_1-u_0)}. \tag{6.6}$$

The formulae for ρ_i, σ_i is obtained by shifting the indices.

From the sequence of σ_i, the solution in the original variables can be calculated from the initial frame ρ_0 by the frame relation

$$\rho_{i+1} = P_i\rho_i = \begin{pmatrix} \sigma_i & -\sigma_i \\ \sigma_i - 2 & \sigma \end{pmatrix}\rho_i \tag{6.7}$$

and the inverse of the moving frame equation,

$$u_i = \rho_i^{-1} \cdot (I_0)_i = \rho_i^{-1} \cdot (-1) \tag{6.8}$$

or the equivalent expressions relating u_{i+1} to $(I_1)_i = 0$ or u_{i+2} to $(I_{11})_i = 1$.

6.1.1 Numerical Algorithm

We summarize the discussion in the above section into a numerical algorithm for solving the variational problem. We assume that $u_0, \ldots, u_5$ are given.

1. Calculate ρ_0, ρ_1, ρ_2 from (6.5) and $\sigma_0, \sigma_1, \sigma_2$ (6.6).
2. For $i = 3, 4, \ldots$ calculate σ_i by solving (6.4).
3. From ρ_2 and (6.7), calculate the sequence $\rho_3, \rho_4, \ldots$.
4. Calculate u_i from (6.8).

6.1.2 Analysis and Tests

Solutions to the continuous variational problem are curves satisfying

$$\eta_{xxx} + 3\eta\eta_x = 0, \tag{6.9}$$

as can be seen by applying the Euler–Lagrange operators directly, or the techniques in [8, Chapter 7]. Furthermore, $u(x) = \frac{\phi^1(x)}{\phi^2(x)}$ where ϕ^1, ϕ^2 are linearly independent solutions of

$$\phi_{xx} + \frac{1}{2}\eta\phi = 0 \tag{6.10}$$

Inserting $\sigma = 2 + h^2\kappa$ and $\kappa_{i-j} = \kappa(-jh)$ into the left hand side of (6.4) yields the Taylor series expansion

$$h^2\left(-2(\kappa_{xxx} + 3\kappa\kappa_x) + 3(\kappa_{xxx} + 3\kappa\kappa_x)_x h + \mathcal{O}(h^2)\right).$$

So the difference scheme is second order accurate for the differential equation in κ.

For testing, we considered the equation on the interval $0 \le x \le 3$. A reference solution of (6.9) was made with MATLAB's ode15 on a fine mesh $h_{\text{ref}} = \frac{1}{64000}$, then two linearly independent solutions of (6.10) with

$$\begin{aligned} \phi^1(0) &= 1 \quad \phi^1(3) = 1 \\ \phi^2(0) &= 1 \quad \phi^2(3) = -1 \end{aligned}$$

were calculated numerically with a standard difference scheme, and $u_{\text{ref}}(x) = \frac{\phi^1(x)}{\phi^2(x)}$ was taken as the reference solution.

Step sizes which were integer multiples of the step sizes of the reference grid $h = Nh_{\text{ref}}$ were used. $u_{\text{ref}}(ih)$ for $i = 0, 1, \ldots, 5$ were taken as input to the algorithm of section 6.1.1. As shown in Figure 6.1, as the step size approaches zero, the error first decreases, then increases. There is apparently an instability. Setting $\sigma = 2 + h^2\kappa$ into (6.4) and ignoring higher order terms gives the linearized difference equation

$$\kappa_{i-3} - 3\kappa_{i-2} + 3\kappa_{i-1} - \kappa_i = 0$$

which is unstable with general solution

$$\kappa_i = A + Bi + Ci^2.$$

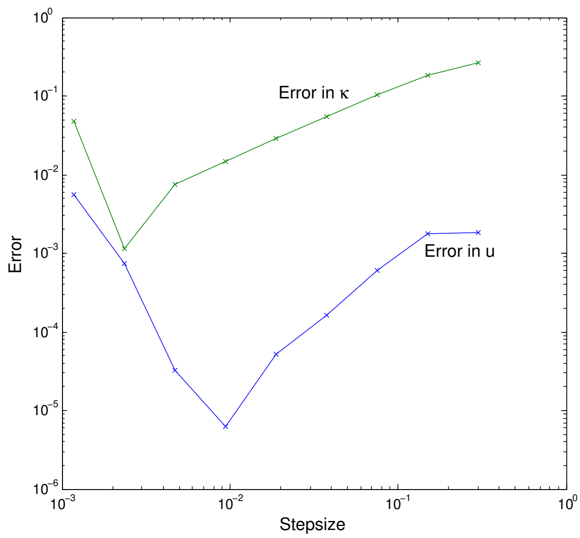


Figure 6.1: Errorplot for the $SL(2)$ invariant Lagrangian

The scheme is unfortunately unstable. The algorithm is accurate for the discrete variational, so it appears that the problem lies in the discretization (6.3). It is not yet clear when an invariant discrete Lagrangian leads to stable schemes.

6.2 $SE(2)$

Setting $u = (x, y)$, we consider curves in the plane minimizing

$$\int \kappa^2 \mathrm{d}s$$

where

$$\kappa = \frac{x_t y_{tt} - y_t x_{tt}}{(x_t^2 + y_t^2)^{\frac{3}{2}}}$$

and $\mathrm{d}s = \sqrt{x_t^2 + y_t^2}\mathrm{d}t$.

Such curves are a special case of elastica curves, known as rectangular elastica. Their solutions satisfy the differential equation

$$\frac{\mathrm{d}^2\kappa}{\mathrm{d}s^2} + \frac{1}{2}\kappa^3 = 0,$$

where

$$\frac{\mathrm{d}}{\mathrm{d}s} = \frac{1}{\sqrt{x_t^2 + y_t^2}} \frac{\mathrm{d}}{\mathrm{d}t}.$$

The lowest order invariants of the $SE(2)$-action on discretized curves are

$$\lambda = I_1^x = \sqrt{(\mathbf{S}x - x)^2 + (\mathbf{S}y - y)^2}$$
$$\sigma = I_{11}^y = \frac{(\mathbf{S}^2 y - y)(\mathbf{S}x - x) - (\mathbf{S}^2 x - x)(\mathbf{S}y - y)}{\lambda}$$

We use the first-order invariant approximation

$$\iota\left[\frac{(\mathbf{S}x - x)(\mathbf{S}^2 y - 2\mathbf{S}y + \mathbf{S}x) - (\mathbf{S}y - y)(\mathbf{S}^2 x - 2\mathbf{S}x + x)}{((\mathbf{S}x - x)^2 + (\mathbf{S}y - y)^2)^{\frac{3}{2}}}\right] = \frac{\sigma}{\lambda^2}. \tag{6.11}$$

We note that the approximation is a second order approximation for $\mathbf{S}\kappa$ when λ is constant.

The invariant discrete Lagrangian is

$$L_d = \frac{\sigma^2}{\lambda^3} \approx \int_{t_i}^{t_{i+1}} \kappa^2 \mathrm{d}s.$$

We force constant step size $\lambda = \Lambda$ by a constraint term to get the constrained Lagrangian

$$L_d^c = \frac{\sigma^2}{\lambda^3} - \mu(\lambda - \Lambda).$$

To calculate the discrete Euler–Lagrange equations of L_d^c, we introduce the preliminary terms

$$E^\lambda L = -3\frac{\sigma^2}{\lambda^4} - \mu$$
$$E^\sigma L = 2\frac{\sigma}{\lambda^3}.$$

By (5.15) with $I_{11}^x = \lambda + \sqrt{\lambda^2 - \sigma^2}$ inserted, the discrete Euler–Lagrange equations are thus

$$\begin{aligned}
&\frac{\sqrt{\lambda^2 - \sigma_{i-1}^2}}{\lambda}\left(-3\frac{\sigma_{i-1}^2}{\lambda^4} - \mu_{i-1}\right) - \left(-3\frac{\sigma_i^2}{\lambda^4} - \mu_i\right) \\
&+\frac{1}{\lambda^2}\left(\sigma_{i-2}\sqrt{\lambda^2 - \sigma_{i-1}^2} + \sigma_{i-1}\sqrt{\lambda^2 - \sigma_{i-2}^2}\right)\frac{2\sigma_{i-2}}{\lambda^3} \\
&-\frac{\lambda + \sqrt{\lambda^2 - \sigma_{i-1}^2}}{\lambda^2}\sigma_{i-1}\frac{2\sigma_{i-1}}{\lambda^3} = 0
\end{aligned} \tag{6.12}$$

$$\begin{aligned}
&\frac{\sigma_{i-1}}{\lambda}\left(3\frac{\sigma_{i-1}^2}{\lambda^4} + \mu_{i-1}\right) \\
&+\frac{1}{\lambda^2}\left[\sqrt{\left(\lambda^2 - \sigma_{i-2}^2\right)\left(\lambda^2 - \sigma_{i-1}^2\right)} - \sigma_{i-2}\sigma_{i-1}\right]\frac{2\sigma_{i-2}}{\lambda^3} \\
&-\frac{1}{\lambda^2}\left[\lambda^2 - \sigma_{i-1}^2 + \lambda\sqrt{\lambda^2 - \sigma_{i-1}^2}\right]\frac{2\sigma_{i-1}}{\lambda^3} + \frac{\sqrt{\lambda^2 - \sigma_i^2}}{\lambda^2}\frac{2\sigma_i}{\lambda^3} = 0.
\end{aligned} \tag{6.13}$$

The constraints are somewhat troublesome to deal with, as they introduce the Lagrange multipliers μ_i as new unknowns. It is possible to solve one of the equations for μ_i, then insert this into the other to get equations for the unknowns σ_i. However, this makes the method a four-step method in stead of a three-step method. In the continuous cases, whether the Euler–Lagrange equations are expressed in invariant or regular coordinates, one of the equations is a total derivative which can be integrated, and the integration constant included in the Lagrangian multiplier. The property in the continuous case depends on the fact that the Lagrangians are invariant under reparametrizations of the curve $(x(t), y(t))$. A similar result for discrete curves has not been found. It is possible that other constraints are more natural choices for the discrete system.

If the equations are solved, however, the solutions in the original variables can be recovered from formula (5.16) and the original frame.

6.2.1 Numerical Algorithm

The four-step method can be implemented as follows:

1. From five initial points
$$(x_0, y_0), \ldots, (x_4, y_4)$$
calculate the first invariants
$$\sigma_0, \sigma_1, \sigma_2, \lambda_0, \ldots, \lambda_3.$$
To fully comply with the formulas (6.12, 6.13) the points should be equidistant. If one keeps track of shift operators on the λ's in (5.15) when calculating the discrete Euler–Lagrange equations (6.12), (6.13), this is not strictly necessary.

2. Calculate μ_1 and μ_2 from (6.13), (6.12) with $i = 2$. respectively.

3. For $i = 3, 4, \ldots$ calculate σ_i from solving (6.13) and μ_i from (6.12).

4. Calculate the sequence $\rho_i = \rho(\mathbf{pr}\, u_i)$ from $\rho_{i+1} = P(\mathbf{pr}\, u_i) \cdot \rho i$.

5. Use $(x_i, y_i) = \rho_i^{-1} \cdot (0, 0)$ to recover the solution in the original coordinates.

When solving for σ_i in step 3 above, one has to solve an equation of the form

$$\sqrt{\lambda^2 - \sigma_i^2}\sigma_i = f(\sigma_i, \sigma_{i-1}, \sigma_{i-2}, \mu_{i-1})$$

This equation will ordinarily have two real solutions. However, it has at most one solution on the interval $[-\frac{\sqrt{2}}{2}\lambda, \frac{\sqrt{2}}{2}\lambda]$. Examining (6.11) we expect that $\sigma = \mathcal{O}(\lambda^2)$ as $\lambda \to 0$, so the solution lying outside the interval can be discarded for small λ.

6.2.2 Tests

The algorithm described above was tested against a reference solution made as follows. $\kappa_{ss} + \frac{1}{2}\kappa^3 = 0$ was solved for the interval $0 < s < 1$ with initial values $\kappa(0) = 0, \kappa_s(0) = 200$ on fine grid (256000 steps) by MATLAB's ode45 with low tolerances. Then, $x(s)$ and $y(s)$ were found by successively integrating the equations

$$\theta_s = \kappa$$
$$x_s = \cos\theta$$
$$y_s = \sin\theta$$

by Simpson's method giving a reference curve of 64001 points.

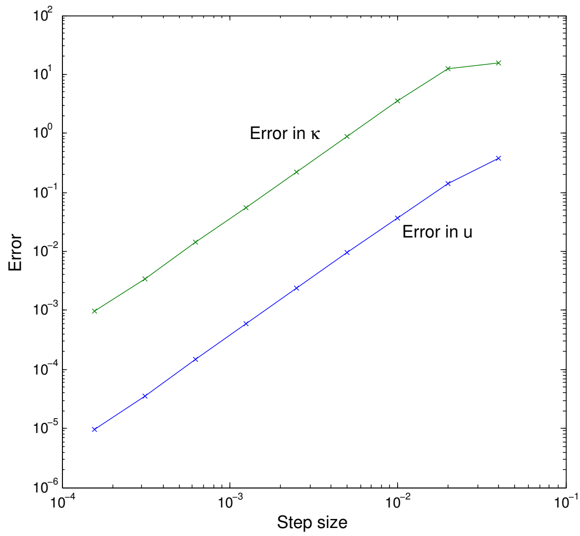


Figure 6.2: Errorplot for the elastica problem

Evenly spaced points on the reference curve were then taken as input to the algorithm described in the previous section with step sizes λ ranging from $\frac{1}{25}$ to $\frac{1}{6400}$. As shown by the error plot in Figure 6.2, the algorithm appears to be converging as $\mathcal{O}\left(\lambda^2\right)$.

CHAPTER 7

CONCLUSIONS AND FURTHER WORK

In this thesis I have studied invariant variational problems and numerical approximations thereof. Two approaches to finding variational methods which inherit symmetries of the original equation were developed. The first approach, detailed in chapters 2 and 3, is only valid for first order Lagrangians. It appears to solve the task it is designed for, but the scope of problems it can be applied to is somewhat narrow. The idea of reparametrizing time used in the example of chapter 3 broadens the scope, though variational problems having known symmetries of this kind are still not in abundance.

The schemes obtained are also in general implicit, and non-linear equations have to be solved for each step, which takes a toll on the performance.

The second approach was to express the discrete Euler–Lagrange equations directly in the invariants. While elegant, at least in the author's opinion, the numerical examples show that this approach has some problems with stability and treating constraints properly. Stability of linear multi step schemes has been studied thoroughly in literature, but the theory here applies only partially to the non-linear multistep methods appearing here.

For both approaches, a large amount of algebraic calculation is required before the schemes for particular ODE's can be implemented. To make the method more general, it would have been interesting to program a computer algebra system to automate some of this calculation.

A question not answered in this thesis is whether the two approaches are equivalent for first order discrete Lagrangians. They should be, as they both solve the same problem, namely solving discrete variational problems when the discrete Lagrangian is invariant under some group transformation, but there has not been enough time to study this thoroughly.

BIBLIOGRAPHY

[1] É. Cartan. *Leçons sur la géométrie projective complexe. La théorie des groupes finis et continus et la géométrie différentielle traitées par la méthode du repère mobile. Leçons sur la théorie des espaces à connexion projective.* Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Éditions Jacques Gabay, Sceaux, 1992. Reprint of the editions of 1931, 1937 and 1937.

[2] V. Dorodnitsyn, R. Kozlov, and P. Winternitz. Symmetries, Lagrangian formalism and integration of second order ordinary difference equations. *J. Nonlinear Math. Phys.*, 10(suppl. 2):41–56, 2003.

[3] M. Fels and P. J. Olver. Moving coframes. I. A practical algorithm. *Acta Appl. Math.*, 51(2):161–213, 1998.

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

[5] V. Guillemin and S. Sternberg. *Variations on a theme by Kepler*, volume 42 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1990.

[6] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[7] E. Hubert. Generation properties of Maurer–Cartan invariants. *INRIA, Preprint*, 2007.

[8] E. L. Mansfield. *A practical guide to the invariant calculus*, volume 26 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2010.

[9] J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1999. A basic exposition of classical mechanical systems.

[10] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numer.*, 10:357–514, 2001.

[11] E. Noether. Invariante Variationsprobleme. *Gött. Nachr*, 235:235–257, 1918.

[12] P. J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.

[13] P. J. Olver. Moving frames and singularities of prolonged group actions. *Selecta Math. (N.S.)*, 6(1):41–77, 2000.

[14] P. J. Olver. Lectures on moving frames. In *Symmetries and Integrability of Difference Equations*, volume 381 of *London Math. Soc. Lecture Note Series*, pages 207–246. Cambridge University Press, 2011.

[15] G. E. Prince and C. J. Eliezer. On the Lie symmetries of the classical Kepler problem. *J. Phys. A*, 14(3):587–596, 1981.