NTNU
Innovation and Creativity

Numerical Simulation of Nonholonomic Dynamics

Dag Frohde Evensberget

Master of Science in Physics and Mathematics

Problem Description

Nonholonomic mechanical systems are of great interest in robot technology applications and control, in particular robotic locomotion and robotic grasping. Roughly speaking a mechanical system with nonholonomic constraints is described by a constrained differential equation such that the constrains are involving the velocity of the system and not only the positions. In this project the numerical simulation of some simple nonholonomic mechanical systems will be considered. The geometry behind these problems is beautiful and non trivial. The aim of the project is understanding the basic theoretical features of nonholonomically constrained systems, illustrate them via numerical simulation, and discuss which numerical approaches are best suited for such problems. In particular classical Runge-Kutta methods will be first applied to the problems and then more specific geometric integrators will be also used. A comparison of the performance of the methods will be part of the presented results. The student can choose to explore the more theoretical or the more practical aspects of the considered problems.

Assignment given: 08. February 2006
Supervisor: Elena Celledoni, MATH

Summary

We study the numerical integration of nonholonomic problems. The problems are formulated using Lagrangian and Hamiltonian mechanics. We review briefly the theoretical concepts used in geometric mechanics.

We reconstruct two nonholonomic variational integrators from the monograph of Monforte [15]. We also construct two one-step integrators based on a combination of the continuous Legendre transform and the discrete Legendre transform from an article by Marsden and West [12]. Initially these integrators display promising behavior, but they turn out to be unstable.

The variational integrators are compared with a classical Runge-Kutta method. We compare the methods on three nonholonomic systems: The nonholonomic particle from [15], the nonholonomic system of particles from an article by McLachlan and Perlmutter [14], and a variation of the Chaplygin sleigh from Bloch [3].

Preface

This Master's thesis was carried out at the Department of Mathematical Sciences at NTNU in the first half of 2006. It is written at the level of a fourth year student with basic knowledge of numerics. It should be readable for any final year student in our department.

The aim of this thesis was to study the technique of variational integration, apply the technique to nonholonomic problems from mechanics, and compare the results to the results obtained using standard numerical methods.

To achieve a deep understanding of variational integrators a sound background in classical mechanics is required. There are undoubtedly some oversimplifications of this material in this thesis; some very interesting aspects of the theory have also been glossed over due to the limited time available to the author.

I would like to thank my supervisor Elena Celledoni for rekindling my fascination with the fields of dynamical systems and mechanics, and for introducing me to the equally fascinating nonholonomic dynamical systems and variational integrators. I would also like to thank her for her patience and attention to detail.

Introduction

The aim of this Master's thesis has been to study the theoretical and practical aspects of numerical integration of nonholonomic systems. Nonholonomic systems are physical systems where there are restrictions on the possible velocities of different components. Classical examples are rolling motion and skating motion. Nonholonomic problems are of interest in amongst other things robot technology and the steering of satellites.

Geometric integrators are a type of numerical methods that conserve qualitative aspects of the exact solution in the numerical solution. Because of this, geometric integrators are suited for long-time integration of physical systems, in which the conservation of energy, momentum and other quantities play a key role.

Nonholonomic mechanics fits somewhat uneasily into the framework of Lagrangian and Hamiltonian mechanics, and many of the conserved quantities of the solution to holonomic problems are no longer conserved in the solution of nonholonomic systems. Still, geometric integrators that work well for holonomic problems also work well for nonholonomic problems. The cause of this is that the geometric integrators preserve the structure of evolution of the solution to nonholonomic problems, even when the exact solution quantities are not conserved.

Chapter 1 presents the theoretical background needed in order to model and understand the equations of motion of holonomic systems and nonholonomic systems, as well as of the structures lying beneath the equations of motion. We also present the qualitative aspects of the solution to holonomic and nonholonomic systems that we wish to preserve.

Chapter 2 presents the theoretical background needed in order to construct classical and simple geometrical numerical integrators for nonholonomic systems. We introduce the variational mechanics that lies behind the geometric integrators we construct in this thesis. We present two concrete two-step methods from [15] and we attempt to construct a possible framework for one-step methods.

Chapter 3 presents the implementation of a classical Runge-Kutta method applied to nonholonomic problems, the implementation of the two two-step methods from Chapter 2, and the implementation of two experimental one-step methods related to the one-step methods of Chapter 2. The methods are then applied to 3 different problems. We plot and study a selection of the results.

Chapter 4 is the conclusion to this thesis. We summarize our results and look ahead at possibilities for further work.

It quickly became apparent to the author that in order to study nonholonomic systems and their integrators a healthy knowledge of geometrical mechanics and the theory of holonomic systems is a prerequisite. This is not a part of the curriculum in industrial mathematics at NTNU. The exposition of this thesis is truthful in the sense that it reflects the material that the author has had to learn while working on the thesis. Because of this the thesis may cover more basic theory than others would deem necessary.

Contents

Chapter 1

Mechanics on Manifolds

Galileo is supposed to have said that "the book of nature is written in the language of mathematics". Maybe Galileo had mechanics in mind, the descriptions of classical mechanics given by Newton, Lagrange and Hamilton are certainly very well suited for description in the language of differential equations.

The basic principle we study in this thesis is Newton's second law "*Force = mass ×* *acceleration*". The physical principle of *determinism* states that if we specify the initial position and velocity $(x_0, \dot{x}_0)$ of a particle the particle's movement is determined forever. Newton's second law may be stated as a second order differential equation

$$\begin{cases} \ddot{x}(t) = f(x, \dot{x}) \\ x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0. \end{cases} \tag{1.1}$$

The above equation has a unique solution for any set of initial conditions. The particle will trace out the curve given by this solution $x(t)$ until the end of time.

Since Newton's time, equation (1.1) has been generalized a lot. In modern geometrical mechanics the solution need not represent the path of a particle, but the evolution of a complex multi-body system in the set of possible configurations.

By simply being able to state the equations of motion of a system we can often gain considerable insight into the physical problem. The *flow* of a system, all solutions as a whole, often has properties that we can study without solving the equations of motion.

In this chapter we aim to introduce the mechanics of Lagrange and Hamilton, and the constrained versions (holonomic and nonholonomic) of these formalisms. In the end we will define the flow of a system, and discuss briefly some of the properties of the flow that we desire to retain in the numerical integration of the system.

1.1 Differential Equations on Manifolds

Taking a very general view, the mathematical structures we study in this thesis are differential equations on manifolds. Monforte [15, p. 44] defines a nonholonomic mechanical problem in the following way:

Definition (Nonholonomic Lagrangian System). *A nonholonomic Lagrangian system on a manifold Q consists of a pair (L, M), where $L : TQ \to \mathbb{R}$ is the Lagrangian of the system and M is a submanifold of TQ.*

It is obvious that some mathematical background is needed to appreciate this definition. This section and Section 2.2 aims to provide a brief introduction to the mathematics needed.

Manifolds

A *manifold* is a mathematical object that which locally looks like an open subset of a Euclidean vector space. A general definition of a manifold is given in e.g. [2, p. 77]. We will not give such a general definition of a manifold here, we will consider only submanifolds of $\mathbb{R}^n$. This will be enough for the purposes of this thesis.

Definition 1.1 (Submanifold of a Vector Space). *Consider a vector space $\mathbb{R}^n$ and a function $g : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$, with components $(g_1, g_2, \ldots, g_m)$. The set of points*

$$M = \{x \in \mathbb{R}^n \mid g(x) = 0\}$$

is an $(n - m)$-dimensional submanifold of $\mathbb{R}^n$ if the matrix

$$\nabla g = \left(\frac{\partial g_j}{\partial x_i} \right)_{i,j}$$

has full rank m for all $x \in M$ [2, p. 80].

We say that M is *embedded* in $\mathbb{R}^n$. It can be shown that every manifold can be embedded in an Euclidean space of sufficiently high dimension n [2, p. 80].

When we describe a physical system mathematically, the set of all possible configurations will be the *configuration manifold Q*. Each point in Q will correspond to a particular configuration of the system.

It is always possible to find a local coordinate system around a given point $q \in Q$ such that a neighborhood of the point q will look like a vector space of dimension $\mathbb{R}^{n-m}$. The dimension of this vector space is equal to the *degrees of freedom* of the system. The components of q are called *generalized coordinates*.

Differentiation on Manifolds

Since we are interested in solving differential equations of these manifolds we obviously need to define what we mean by *curves* on manifolds and *tangents* to these curves.

In the same way as we may consider a standard parametric curve in an Euclidean vector space,

$$t \mapsto x(t) \in \mathbb{R}^n,$$

say the solution of (1.1), we may consider a curve

$$t \mapsto x(t) \in M$$

that stays on the manifold; for all times t we have $x(t) \in M$. The *tangent* to a curve $x(t)$ at $t = 0$ is

$$\dot{x} = \lim_{t \to 0} \frac{x(t) - x(0)}{t}.$$

The *tangent space* to a manifold M at a point $x_0 \in M$ is the vector space spanned by the set of tangents $\dot{x}$ of all curves $x(t)$ that go through the point x_0 at $t = 0$ [2, p. 80].

Definition 1.2 (Tangent and Cotangent Space of a Submanifold of $\mathbb{R}^n$)**.** *Let M a sub-manifold of $\mathbb{R}^n$ defined by the function $g(x) = 0$ as in Definition 1.1. For each $x \in M$, the tangent space to M at x is the vector space given by [4, p. 66]*

$$T_x M = \left\{ v \in \mathbb{R}^n \ \middle| \ \frac{\partial g_i}{\partial x} \, v = 0 \right\} \tag{1.2}$$

The cotangent space *of M at x is the vector space of all linear functions from $T_x M$ to $\mathbb{R}$,*

$$T_x^* M = \{ \omega : T_x M \to \mathbb{R} \mid \omega \text{ is linear} \} . \tag{1.3}$$

The cotangent space $T_x^* M$ is the dual space of $T_x M$ [15, p. 14].

If we consider a function $f : M \to \mathbb{R}$, we can differentiate it along a curve $x(t) \in M$ and use the chain rule,

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(x(t)\big) = \frac{\partial f}{\partial x}\big(x(t)\big) \, \dot{x}(t).$$

The cotangent space of a point $x \in M$ is the set of such objects $\frac{\partial f}{\partial x}(x) : T_x M \to \mathbb{R}$.

For the manifold $\mathbb{R}^n$ the tangent and cotangent space are isomorphic to the base space $\mathbb{R}^n$. For submanifolds $M \subset \mathbb{R}^n$ the tangent space at the point x is the set of tangent vectors to curves in $\mathbb{R}^n$ that do not *point out* of the manifold M.

By taking the disjoint union of the tangent spaces and the disjoint union of the cotangent spaces we get the *tangent bundle* and the *cotangent bundle*.

Definition 1.3 (Tangent Bundle and Cotangent Bundle)**.** *The* tangent bundle *of M is a manifold whose underlying set is the union of all the tangent spaces,*

$$TM = \bigcup_{x \in M} T_x M.$$

The tangent bundle may be endowed with a natural manifold structure via the tangent bundle projection $\tau_M : TM \to M$, $\tau_M^{-1}(x) = T_x M$.

The cotangent bundle *$T^* M$ is a manifold whose underlying set is the union of all cotangent spaces $T_x^* M$,*

$$T^* M = \bigcup_{x \in M} T_x^* M.$$

The cotangent bundle may be endowed with a natural manifold structure via the cotangent bundle projection $\pi_M : TM \to M$, $\pi_M^{-1}(x) = T_x M$.

We will specify a point on the tangent bundle TM with coordinates $(x, \dot{x})$. The tangent bundle has twice the dimension of the base manifold. We will not use the manifold structure of TM or T^*M in any explicit way in this thesis, it is stated here to emphasize that the two objects are manifolds.

The tangent bundle TQ to the configuration manifold Q will be given coordinates $(q, \dot{q})$. A vector $\dot{q}$ is called a generalized velocity of the system.

The final object we define in this section is a *smooth distribution*. We will make the same assumption as Monforte [15, p. 44] that D is a submanifold of TM.

Definition 1.4 (Smooth Distribution as Submanifold of TM)**.** *A smooth distribution D on a manifold M is a submanifold of the tangent manifold TM such that for each point $x \in M$ we have that D_x is a subspace of $T_x M$ and that the union $D = \bigcup_{x \in M} D_x$ is a smooth submanifold of TM.*

In this thesis the distribution D will always be generated by a set of functions $\omega : TM \to \mathbb{R}$ in a very similar way to the that of Definition 1.1:

$$D = \{(q, \dot{q}) \in TQ \mid \omega(q, \dot{q}) = 0\} \tag{1.4}$$

The effect of restricting a curve $q(t) \in Q$ to D is to require that that the corresponding curve $(q, \dot{q})(t) \in TQ$ is restricted to $D \subset TQ$. While the curve $q(t)$ may still visit any point of Q, some directions of movement that are inside TQ are prohibited in each point q.

Differential Equations on Manifolds

We are now in a position to state the meaning of a "differential equation on a manifold". A differential equation $\dot{y} = f(y)$ is on the manifold M if $y_0 \in M$ implies that $y(t) \in M \; \forall \, t$. Hairer et al. [8, Section IV.5.2] states the following theorem:

Theorem 1.5 (Differential Equation on Manifold)**.** *Let M be a submanifold of $\mathbb{R}^n$. The problem $\dot{y} = f(y)$ is a differential equation on the manifold M if and only if*

$$f(y) \in T_y M \; \forall \; y \in M.$$

The above theorem has relevance for us. We shall later see that a constrained physical system may often be stated either in $\mathbb{R}^n$ or on a submanifold $M \subset \mathbb{R}^n$ as in in Definition 1.1. The function g that defines M will be the *constraint function*.

1.2   Variational Formulation of Mechanics

In this thesis we will study two *variational formulations* of mechanics, in which Newton's laws are replaced with a variational principle. In both formalisms we solve the equations of motion by finding a curve $q(t)$ on a manifold that extremizes a particular functional or "energy function". The material of this section is taken from [2, 11, 21].

Lagrangian mechanics and the more general Hamiltonian mechanics are formulated on manifolds, and they are independent of any particular coordinate system. This means the complexity of the equations can be reduced by a clever choice of coordinates. It is also a lot easier to *model* a mechanical system in these two formalisms, and more problems may solved analytically.

In this section the Lagrangian and Hamiltonian formalisms are briefly outlined in a constraint-free setting. Later both formalisms will be extended to cater for constrained systems.

The Lagrangian and Hamilton's Principle

Lagrange revolutionized mechanics in 1788 when he reformulated Newtonian mechanics by replacing Newton's laws by a single relation. Lagrangian mechanics has two components: The configuration manifold Q, and the The Lagrangian L.

The *configuration manifold Q* is a set of point in which each point $q \in Q$ represents one possible position or configuration of *all* the elements in the physical system. For instance, the position of a point particle is described by a the vector $x \in \mathbb{R}^3$. To accommodate for several (say two) particles we may take $q = (x_1, x_2) \in \mathbb{R}^6$ where $x_i \in \mathbb{R}^3$. The dimension of Q will always be equal to the number of degrees of freedom in the system. Each scalar component of q is called a *generalized coordinate*.

A point $\dot{q} \in T_q Q$ is called a *generalized velocity*, and a point $(q, \dot{q}) \in TQ$ is called a *state* of the system. Each state represents the totality of information about both position and velocity of all elements of the system. When we know that state of the system at a given time we may find the state at any later time.

The *Lagrangian* is a function

$$L : TQ \to \mathbb{R}. \tag{1.5}$$

For most physical problems the Lagrangian is defined by

$$L(q, \dot{q}) = K(q, \dot{q}) - U(q)$$

where K is the kinetic energy of the system and U is the potential energy of the system.

The variational *Hamilton's principle* now characterizes the evolution of the system and solution of the problem:

Definition 1.6 (Hamilton's principle). *Consider the set of curves $q(t) \in Q$ with fixed endpoints $q_0 = q(t_0)$ and $q_1 = q(t_1)$. Among these curves, the curve that describes the physically correct evolution of the system will extremize the* action functional

$$G(q) = \int_{t_0}^{t_1} L(q(t), \dot{q}(t)) \, \mathrm{d}t \tag{1.6}$$

where $L(q, \dot{q})$ is the Lagrangian of the system.

Hamilton's principle thus allows one, at least in principle, to find the correct evolution of the system if we know the initial configuration q_0 and the final configuration q_1. For pairs (q_0, q_1) that are sufficiently close the action functional has a unique extremizer. By variational calculus we may restate Hamilton's principle in the form of a differential equation in q.:

Theorem 1.7 (Euler-Lagrange equations). *A curve $q(t)$ that satisfies Hamilton's principle also satisfies the Euler-Lagrange equations*

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \tag{1.7}$$

where $L(q, \dot{q})$ again is the Lagrangian of the system.

Proof. The fact that the curve $q(t)$ is a minimum of G means that for any curve $r(t)$ in Q,

$$
\begin{aligned}
0 &= \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} G(q + \varepsilon r), \\
&= \int_0^t \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} L(q + \varepsilon r, \dot{q} + \varepsilon \dot{r}) \, \mathrm{d}t, \\
&= \int_0^t \frac{\partial L}{\partial q}(q + \varepsilon r, \dot{q} + \varepsilon \dot{r})r + \frac{\partial L}{\partial \dot{q}}(q + \varepsilon r, \dot{q} + \varepsilon \dot{r})\dot{r} \, \mathrm{d}t \bigg|_{\varepsilon=0,} \\
&= \int_0^t \frac{\partial L}{\partial q}(q, \dot{q})r + \frac{\partial L}{\partial \dot{q}}(q, \dot{q})\dot{r} \, \mathrm{d}t.
\end{aligned}
$$

By partial integration of the last term and using the fact that $r(0) = r(t) = 0$ we obtain

$$\int_0^t \left(\frac{\partial L}{\partial q}(q, \dot{q}) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}}(q, \dot{q})\right) r(q, \dot{q}) = 0,$$

so since r was arbitrary the integrand must be identically zero. Thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0.$$

$\square$

The Euler-Lagrange equations thus give a fairly easy way of finding differential equation on a manifold whose solution is a curve describing the correct evolution of the physical system. This was Lagrange's revolution, it allows previously unsolvable classes of mechanical problems to be solved with ease and elegance. Examples are given in any book on classical mechanics, e.g. [2].

In order to model nonholonomic constraints we need to extend the Euler-Lagrange equations to include physical forces $F(q, \dot{q})$ that are not derived directly from the potential energy $U(q)$.

Definition 1.8 (Lagrange-d'Alembert Principle)**.** *Let* (L, Q), $L : TQ \to \mathbb{R}$ *be a Lagrangian system, and let the system be acted upon by a resultant force* $F \in T^*Q$. *Consider set of curves* $q(t) \in Q$ *with* fixed *endpoints* $q_0 = q(t_0)$ *and* $q_1 = q(t_1)$. *The curve that describes the physically correct evolution of the system is the solution of*

$$\delta \int_{t_0}^{t_1} L\big(q(t), \dot{q}(t)\big) \, \mathrm{d}t + \int_{t_0}^{t_1} F\big(q(t), \dot{q}(t)\big) \, \delta q \, \mathrm{d}t = 0.$$

This version of the Lagrange-d'Alembert principle is from [3, Chapter 3.4]. It can be shown by variational calculus [11, Chapter 7.8] that the above principle is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = F. \tag{1.8}$$

Equation (1.8) is called the Euler-Lagrange equations with external forces. It is central in the modeling of nonholonomic constraints.

Hamiltonian Mechanics

A third formulation of mechanics was put into its final form by Hamilton in 1834. Much in the same way as the Lagrangian formulation, the Hamiltonian formulation is variational in nature. Hamilton's equations are more general than the Lagrangian formulation, and the Hamiltonian formulation of mechanics is more powerful when it comes to finding analytical solutions. In physics the Hamiltonian formulation is considered the most powerful, and it was instrumental in developing quantum mechanics and modern physics in general.

We will only consider mechanical systems that have both Hamiltonian and Lagrangian formulations. The *Legendre transform* [3, Chapter 3.3] enables us to state a Lagrangian system as a Hamiltonian system.

Definition 1.9 (Hamiltonian formulation of a Lagrangian system)**.** *Let* L *be a Lagrangian on* TQ *and let the* Legendre transformation *be given by*

$$FL : TQ \to T^*Q$$
$$(q, \dot{q}) \mapsto (q, p)$$

where the p *is the* generalized momentum

$$p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}). \tag{1.9}$$

Solving equation (1.9) *to express* $\dot{q}$ *as a function of* (q, p), *the* Hamiltonian *corresponding to the Lagrangian* L *is*

$$H(q, p) = p^{\mathrm{T}} \dot{q}(q, p) - L\big(q, \dot{q}(q, p)\big). \tag{1.10}$$

We will only consider cases where the Legendre transform is bijective. In geometric mechanics points (q, p) lie on the cotangent bundle T^*Q [2, Chapter 3.14], and p is called *generalized momentum*.

The Hamiltonian is a function

$$H : T^*Q \to \mathbb{R}.$$

Unlike the Lagrangian L, the Hamiltonian has a physical interpretation as the total energy of the system,

$$H(q, p) = T(q, p) + U(q),$$

where $T(q, p)$ is the kinetic energy of the system and $U(q)$ is the potential energy of the system expressed in coordinates on T^*Q.

Similarly to the Euler-Lagrange equations in the Lagrangian formalism, *Hamilton's equations* are the equations of motion of a Hamiltonian system. Hamilton's equations are first order.

Theorem 1.10 (Hamilton's Equations). *Consider a Hamiltonian system given by (H, Q), $H : T^*Q \to \mathbb{R}$ which is found by the Legendre transformation of a Lagrangian system (L, Q). The correct evolution of the system is given by* Hamilton's equations

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}. \tag{1.11}$$

When the Lagrangian formulation exists, Hamilton's equations are equivalent to the Euler-Lagrange equations (1.7).

Proof. This proof is taken from [8, Chapter VI.1]. The definition of the Hamiltonian (1.10) and the generalized momentum (1.9) leads to the following:

$$\frac{\partial H}{\partial p} = \dot{q}^{\mathrm{T}} + \underbrace{p^{\mathrm{T}} \frac{\partial \dot{q}}{\partial p} - \frac{\partial L}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial p}}_{=0 \text{ by } (1.9)} = \dot{q},$$

$$\frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q} + \underbrace{p^{\mathrm{T}} \frac{\partial \dot{q}}{\partial q} - \frac{\partial L}{\partial \dot{q}} \frac{\partial \dot{q}}{\partial q}}_{=0 \text{ by } (1.9)} = -\frac{\partial L}{\partial q} = -\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{q}} = -\dot{p}.$$

where the second last equality is by equation (1.7) and the last equality is by (1.9). In the final equalities we have dropped the transpose signs. $\square$

As in the Lagrangian case we need to extend Hamilton's equations to include forces. Marsden and West [12, p. 422] give the *forced Hamilton's equations*

$$\dot{q} = \frac{\partial H}{\partial q}, \quad \dot{p} = -\frac{\partial H}{\partial p} + F \tag{1.12}$$

The forced Hamilton's equations are essential in the formulation of nonholonomic Hamiltonian systems.

1.3   Mechanical Systems with Constraints

In this section we shall consider constrained Lagrangian systems. Most of the material is easily applied to the Hamiltonian case. For simplicity we will assume that the configuration manifold $Q = \mathbb{R}^n$.

In a Lagrangian system (L, Q) a *constraint* is a condition

$$f : TQ \to \mathbb{R}, \quad f(q, \dot{q}) = 0$$

that the solution of the problem myst adhere to. *Holonomic constraints* are well understood. *Nonholonomic constraints* are more general and less well known. They are not as easy to geometrize and historically they have been something of mystery. In both cases we will attempt to use *Lagrangian multipliers* to generalize the Euler-Lagrange equations (Theorem 1.7) and Hamilton's equations (Theorem 1.10) to include constrained problems.

Holonomic Constraints

A *holonomic constraint* is a constraint of motion on the configuration manifold.

Definition 1.11 (Holonomic constraint)**.** *Consider a Lagrangian system (L, Q) with $L : TQ \to \mathbb{R}$. A* holonomic constraint *is a restriction of the possible configurations to a submanifold of Q defined by*

$$g : Q \to \mathbb{R}, \quad g(q) = 0.$$

Holonomic constraints appear when the number of degrees of freedom of the physical system under consideration is smaller than the dimension of the configuration manifold.

The equations of motion of a holonomic Lagrangian systems can be derived from the theory of *Lagrange multipliers* [22, §5.7, §8]. We do not go into that theory here but refer instead to the proof of Theorem 1.12.

Theorem 1.12 (Equivalent Formulations of Holonomic Problems [12, theorem 3.3.2])**.** *Consider a configuration space $\mathbb{R}^n$ and a constraint function $g : \mathbb{R}^n \to \mathbb{R}^m$ that defines a manifold $Q = \{x \in \mathbb{R}^n \mid g(x) = 0\}$, the following is equivalent:*

1. *The curve $q(t)$ is a solution of the Euler-Lagrange equations (from Theorem 1.7) on TQ.*

2. *The curves $q(t)$ and $\lambda(t)$ are a solution of the constrained Euler-Lagrange equations on $\mathbb{R}^n$*

$$\begin{cases} \dfrac{\partial L}{\partial x} - \dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial L}{\partial \dot{x}} = \nabla g(x)\, \lambda, \\ \qquad\qquad\quad g(x) = 0. \end{cases}$$

3. *The curve $\big(q(t), \lambda(t)\big)$ is a solution of the Euler-Lagrange equations for the Lagrangian*

$$\bar{L}(q, \lambda, \dot{q}, \dot{\lambda}) = L(q, \dot{q}) - \nabla g(q)\, \lambda$$

on the manifold $Q \times \mathbb{R}^m$.

The equivalence of Item 2 and Item 3 in Theorem 1.12 show that any Lagrangian system with configuration manifold $Q \subset \mathbb{R}^n$ may be considered to be a holonomically constrained system. This system has configuration manifold $\mathbb{R}^n$ and holonomic constraints $g(x) = 0$, where g is the function that defines Q as a submanifold of $\mathbb{R}^n$. This formulation of a system is called *extrinsic*. Vice versa, all holonomically constrained Lagrangian systems have an equivalent unconstrained Lagrangian system. In this system the constraints instead are part of the definition of the configuration manifold Q. The unconstrained system (L, Q) is called the *intrinsic* formulation.

By comparing the Euler-Lagrange equations with external forces (1.8) with Item 2 in Theorem 1.12 we see by comparing the right hand sides that the constraint g manifests itself as a force

$$F = \nabla g(q)\,\lambda, \quad g(x) = 0 \tag{1.13}$$

on the Lagrangian system. This force is called the *constraint force*. The constraint force ensures that the solution $x(t)$ stays on the manifold Q. The Lagrange multiplier λ takes whatever values are necessary to keep the trajectory of the system $x(t)$ inside Q.

Item 3 in Theorem 1.12 shows that the equations of motion of a holonomic Lagrangian system are themselves derived from a variational principle from a variational principle. This principle is the extremization of the action integral $\bar{G}$ of the so-called *augmented* Lagrangian $\bar{L}$.

In the rest of this thesis we will consider the constraint function $g(q)$ to be vector valued so that each holonomic constraint is a component of the function. If we have m constraints $g_i(q) = 0$, together all the we may write

$$g(q) = \begin{pmatrix} g_1(q) \\ \vdots \\ g_m(q) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

The vector equation $g(q) = 0$ thus contains all the holonomic constraints.

Nonholonomic Constraints

Nonholonomic constraints are constraints on the tangent space TQ. Nonholonomic constraints restricts not only the relative positions of a systems components, but also their velocities. Nonholonomic constraints appear in e.g. rolling motion.

Definition 1.13 (Nonholonomic constraint). *Consider a Lagrangian system (L, Q) with $L : TQ \to \mathbb{R}$. A nonholonomic constraint is a restriction*

$$\omega : TQ \to \mathbb{R}, \quad \omega(q, \dot{q}) = 0$$

on the state space TQ that is not the time derivative $\frac{\mathrm{d}}{\mathrm{d}t} g(x) = 0$ of a holonomic constraint.

A nonholonomic constraint restricts the solution of the equations of motion to a distribution D as given in equation (1.4). It has been shown that the equations of motion for
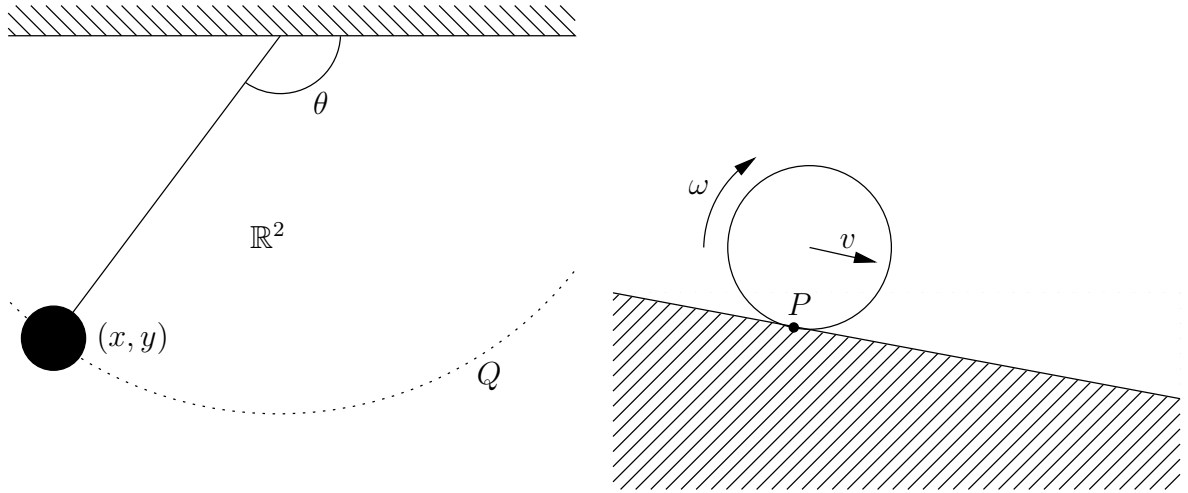
Figure 1.1: *Left:* A holonomic constraint is a restriction to a submanifold of $\mathbb{R}^n$. The pendulum in $\mathbb{R}^2$ is constrained to the manifold $Q \in \mathbb{R}^2$. The configuration manifold is 1-dimensional and may be parameterized by the angle θ. *Right:* A nonholonomic constraint involves velocity. For the cylinder rolling on an incline the velocity of the surface of the cylinder at the point P must be zero. Because of this the angular velocity ω and the linear velocity v are connected through the formula $v = k\omega$ for a constant k.

a nonholonomic Lagrangian system are not possible to derive from a variational principle [15, Chapter 5].

There is no choice of configuration manifold Q such that the nonholonomic constraints on $\mathbb{R}^n$ restricts the solution to Q. In fact a nonholonomic constraint generally does not restrict the configuration manifold at all, it is certain combinations of positions and velocities that are disallowed.

In this thesis we shall only consider nonholonomic constraints that are linear in the velocities $\dot{q}$, so that they can be factored as

$$\omega(q, \dot{q}) = a(q)^{\mathrm{T}} \dot{q}. \tag{1.14}$$

When there is more than one nonholonomic constraint we express them all in a single relation. We let the factor $a_i(q)^{\mathrm{T}}$ of each constraint $\omega_i(q, \dot{q}) = a_i(q)^{\mathrm{T}} \dot{q}$ be a row in the matrix $A(q)$. All the nonholonomic constraints may then be written together as

$$A(q)\,\dot{q} = \begin{pmatrix} \cdots a_1(q)^{\mathrm{T}} \cdots \\ \vdots \\ \cdots a_m(q)^{\mathrm{T}} \cdots \end{pmatrix} \begin{pmatrix} \dot{q}_1 \\ \vdots \\ \dot{q}_m \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = 0.$$

The matrix-vector product $A(q)\,\dot{q} = 0$ contains all the nonholonomic constraints. The nonholonomic constraints now restrict the state space to a submanifold of the tangent space TQ to a smooth distribution defined by

$$D = \{(q, \dot{q}) \in TQ \mid A(q)\,\dot{q} = 0\}. \tag{1.15}$$

A nonholonomic constraint is *not integrable*. This means it may not be derived from a holonomic constraint by differentiation. By differentiating the holonomic constraint in Definition 1.11 respect to time we can obtain a relation on the form of a nonholonomic constraint,

$$\frac{\mathrm{d}}{\mathrm{d}t}g(q) = \nabla g(q)\,\dot{q}. \tag{1.16}$$

We may consider the function $g(q)$ as potential function on Q; this makes ∇g a vector field over Q. Similarly $A(q)$ may be viewed as a vector field over Q. From vector calculus we know that $A(q)$ is integrable (i.e. the derivative of a potential function) if and only if $\nabla \times A(q) = 0$. Thus we see that by differentiation we may obtain a nonholonomic-like constraint from equation (1.16), but it is not possible to obtain a holonomic constraint from equation (1.14).

There is no equivalent counterpart of Theorem 1.12 in the case of nonholonomically constrained systems. The nonholonomic constraint does not restricts the solution of the equations of motion to a submanifold $Q \subset \mathbb{R}^n$. There equations of motion may not be derived from a variational principle.

The holonomic constraint force (1.13) does have an analogy in nonholonomically constrained systems. It is found from the *nonholonomic principle* [3, Chapter 5.1].

Definition 1.14 (Nonholonomic Principle)**.** *Consider a Lagrangian system (L, Q) with linear nonholonomic constraints given by $A(q)\,\dot{q} = 0$. The* nonholonomic principle *is the assumption that*

$$F = \sum_{i=1}^{m} \lambda_i a_i(q)^{\mathrm{T}} = A(q)^{\mathrm{T}}\,\lambda, \tag{1.17}$$

the resultant force F must be a linear combination of the rows of $A(q)$.

The resultant force F thus is perpendicular to the trajectory $x(t)$ of the system in any point. Because of this the work done by the constraint forces is zero, and thus energy is conserved for linear constraints $\omega(q, \dot{q}) = a(q)^{\mathrm{T}}\,\dot{q}$. We can now formulate the equations of motion of a nonholonomic Lagrangian system. There are two formulations:

Intrinsic Equations of Motion for a Nonholonomic Lagrangian System. Consider a Lagrangian system (L, Q) with $L : TQ \to \mathbb{R}$ and a constraint distribution

$$D = \left\{ (q, \dot{q}) \in TQ \mid A(q)\,\dot{q} = 0 \right\}.$$

The correct equations of motion are

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial L}{\partial \dot{q}} - \dfrac{\partial L}{\partial q} = A(q)^{\mathrm{T}}\lambda, \\[2mm] \qquad\quad A(q)\,\dot{q} = 0. \end{cases} \tag{1.18}$$

The formulation above is suited for theoretical considerations. The extrinsic formulation below is more suited for numerical integration:

Extrinsic Equations of Motion for a Nonholonomic Lagrangian System. Consider a Lagrangian system (L, Q) with $L : T\mathbb{R}^n \to \mathbb{R}$ where

$$Q = \{x \in \mathbb{R}^n \mid g(x) = 0\} \,,$$

and a constraint distribution

$$D = \{(q, \dot{q}) \in TQ \mid A(q)\,\dot{q} = 0\} \,.$$

The correct equations of motion is

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial L}{\partial \dot{x}} - \dfrac{\partial L}{\partial x} = \nabla g(x)\,\lambda + A(x)^{\mathrm{T}}\,\mu, \\ \qquad\quad A(x)\,\dot{x} = 0, \\ \qquad\qquad g(x) = 0. \end{cases} \tag{1.19}$$

Bloch [3, Chapter 5] remarks that historically the nonholonomic principle and the nonholonomic equations of motion have been contested several times, because of the non-intuitive behavior of systems. The equations (1.18) have been shown to be equivalent to Newton's laws for fairly general systems, and they agree with experimental results.

1.4 The Geometry of the Phase Space Flow

The *phase space* of a system $\dot{y} = f(y) \in \mathbb{R}^n$ of first order differential equations is the space of possible states y in which the system may find itself. They phase space of a constrained Lagrangian or Hamiltonian system is found by rewriting the system as a set of first order equations in the way that equation (1.1) may be written as

$$\begin{cases} \dot{x}(t) = u, \\ \dot{u}(t) = f(x, u). \end{cases}$$

A holonomic or nonholonomic system may be restated as $\dot{y} = f(y)$, $y \in M$, a set of first order differential equations on a manifold M. For a holonomic Lagrangian system we have $M = TQ$, for a holonomic Hamiltonian system $M = T^*Q$. For a nonholonomic system the phase space is given by a submanifold of M given by the constraint distribution D.

A point $(q, \dot{q}) \in TQ$ completely determines the evolution of the system. A consequence of this is that each point in TQ is on a unique curve in TQ. The evolution of the system is reflected by the movement of the system's point in phase space along this curve.

Definition 1.15 (Flow of a Differential Equation)**.** *Consider a dynamical system with an arbitrary initial state y_0,*

$$\begin{cases} \dot{y} = f(y), \quad y \in M, \\ y(0) = y_0. \end{cases} \tag{1.20}$$

The flow ϕ on the phase space M is a collection of maps such that for all times $t \geq 0$ and for any point $y_0 \in M$ we have that

$$\phi^t(y_0) = y(t),$$

where y(t) is the solution at time t to the differential equation (1.20).

The flow is the concept of *all* solutions to a differential equations at the same time. The flow is a *group* under composition: It is closed and has identity element ϕ^0 and an inverse element $\phi^{-t} \circ \phi^t = \text{id}$.

Flow lines may not intersect, if they could it would mean that the differential equation behind the flow does not have a unique solution. For a systems equations of motion this would break the principle of physical determinism, since the system modeled would have two possible future evolutions.

We shall discuss some further properties of the flow of holonomic and nonholonomic mechanical systems in the rest of this section.

Time-Reversibility

All conservative mechanical systems have the property that if you reverse the orientation of the generalized velocity vector, $\dot{q} \mapsto -\dot{q}$, and keep the generalized position vector you reverse the direction of the flow. Time-reversibility is a particular case of ρ-reversibility as defined in [8, Section V.1].

Definition 1.16 (Reversibility). *Consider a system defined by $\dot{y} = f(y)$, and an invertible linear transformation ρ. The flow ϕ of the system is ρ-reversible if*

$$(\rho \circ \phi^t)(y_0) = (\phi^{-t} \circ \rho)(y_0)$$

For a Hamiltonian system (1.11), the system is time-reversible if it is ρ-reversible for the particular linear transformation $\rho(q, p) = (q, -p)$.

First Integrals and Momentum Maps

Conserved quantities or *first integrals* are quantities that are conserved as the system evolves [8, Chapter IV.1]. Physically this means that by knowing the value of a conserved quantity at the time t_0 we also know it at a later time $t \geq t_0$.

Definition 1.17 (First Integrals). *Consider a system $\dot{y} = f(y)$ with flow ϕ. A first integral is a non-constant function $I(y)$ such that for any point y_0 in the phase space, and for all times t we have that*

$$\frac{\mathrm{d}}{\mathrm{d}t} I\big(\phi^t(y_0)\big) = 0.$$

For Hamiltonian systems we have that the Hamiltonian $H(q, p)$ is a conserved quantity, in the problems we study in this thesis it is equal to the energy of the system. In the case of linear nonholonomic constraints, i.e. constraints that may be expressed by equation (1.15), energy is conserved.

For unconstrained or holonomic system, *Noether's theorem* links symmetries in the Lagrangian with conserved quantities, which are called *momentum maps*. Noether's theorem is given in any book in classical mechanics, e.g. [2, Chapter 20]. Conserved quantities such as momentum and angular momentum are momentum maps and may be found via Noether's theorem. In nonholonomic mechanics Noether's theorem does *not* hold, this is one of the differences between holonomic and nonholonomic systems.

Another type of first integral is very relevant numerically when the equations are given in extrinsic form as in 1.19. For a differential equation $\dot{x} = f(x) \in \mathbb{R}^n$ and a manifold $M \subset \mathbb{R}^n$ as in Definition 1.1, a *weak invariant* is a function $g(x)$ such that for $x_0 \in M$

$$\frac{\mathrm{d}}{\mathrm{d}t} g\big(\phi^t(x_0)\big) = 0 \ \forall \ \phi^t(x) \in M.$$

The function g is only constant if the flow stays in M. If we take g to be the constraint function of a holonomic Lagrangian system on $Q \subset \mathbb{R}^n$, the constraint function is precisely a weak invariant of the system.

Symplecticity

Writing $y = (p, q)$ as in [9, Section I.14], the unconstrained Hamilton's equations (1.11) may be stated as

$$\dot{y} = J^{-1}\nabla H(y), \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

where I is the $(n \times n)$ identity matrix. The flow ϕ of the above equation satisfies

$$\left(\frac{\partial \phi^t}{\partial y_0}\right)^{\mathrm{T}} J \left(\frac{\partial \phi^t}{\partial y_0}\right) = J. \tag{1.21}$$

A mapping that satisfies equation (1.21) is called a *symplectic map*. Holonomic Hamiltonian systems have symplectic flows. A consequence of the symplecticity of the flow is that volume is conserved in the phase space.

Nonholonomic Hamiltonian systems do *not* have symplectic flows. Together with the lack of preservation of momentum maps, this is the most important difference between holonomic and nonholonomic systems from a geometric perspective.

Table 1.1 show that there are major differences between holonomic and nonholonomic systems. The differences are so big that one might ask whether there is anything to be gained by forcing nonholonomic systems into the theoretical framework developed for holonomic systems. The situation is not so bleak as it might appear however, since the symplectic form and momentum maps evolve in a structured way [15, Section 7.5]. Chapter 2 and 3 show that numerical methods designed for holonomic systems may *sometimes* be adapted to the nonholonomic case with great success.

	Holonomic Systems	Nonholonomic Systems
Reversibility	Conserved	Conserved
Energy	Conserved	Conserved for linear constraints
Weak invariants	Conserved	Conserved
Momentum Map	Conserved	Not conserved
Symplectic form	Conserved	Not conserved

Table 1.1: Flow properties that are conserved by holonomic and nonholonomic systems.

Chapter 2

Numerical Integrators

As we stated in Chapter 1, the solution to a differential equation

$$\begin{cases} \dot{y}(t) = f(y), & y \in M \\ y(0) = y_0 \end{cases} \tag{2.1}$$

is a function $y(t)$ that is in agreement with equation (2.1) for all times t. From now on we will refer to this function as the *analytical solution*. For the kind of problems we study in this thesis we will assume that the analytical solution *exists* and is *unique*.

For most differential equations an explicit expression of the analytical solution is difficult or impossible to obtain. By solving a differential equation *numerically* we mean computing an approximation to the analytical solution at a set of time values $\{t_k\}_{k=0}^N$. This set of approximations is called the *numerical solution*.

We denote the numerical solution at a particular time t_k by y_k, and seek a series $\{y_k\}_{k=0}^N$ such that

$$y_k \approx y(t_k), \quad k = 0, 1, \dots, N.$$

The first term of the numerical solution $\{y_k\}_{k=0}^N$ is the initial condition y_0.

The *step length* is denoted by h_k. The step length is the distance between successive elements of $\{t_k\}_k^N$. We always assume that we are solving the equations on the interval $[0, t_{\max}]$. We will consider numerical solutions with constant step length,

$$h_k = h = \frac{t_{\max}}{N-1} \ \forall \ k,$$

so the elements of the series $\{t_k\}_{k=0}^N$ are given by

$$t_k = kh, \quad k = 0, 1, \dots, N.$$

A method of obtaining a numerical solution is called a numerical *method* by mathematicians, an *algorithm* by computer scientists and an *integrator* by physicists. We will use the words method and integrator interchangeably.

Numerical Flow

In the same way that the concept of flow from Section 1.4 facilitates the description of the analytical solution of (2.1) for any initial value y_0, we want to be able to describe of properties that are common to all numerical solutions obtained with a particular integrator.

The *numerical flow*, Φ, of a particular integrator is a set of mappings of the type

$$
\begin{aligned}
\Phi : M \times \cdots \times M &\to M \times \cdots \times M, \\
(y_k, \ldots, y_{k+i}) &\mapsto (y_{k+1}, \ldots, y_{k+i+1})
\end{aligned}
\tag{2.3}
$$

that are used to obtain the numerical solution of that integrator.

If the integrator takes a single copy of M as argument, $\Phi : M \to M$, the integrator is called a *single-step* integrator. A general integrator on the form (2.3) is called a *multi-step* integrator. A i-step integrator requires i initial values to produce an approximation to y_{i+1}. This means that multi-step integrators require several initial values to obtain a numerical solution. This problem is addressed in Chapter 3.

By repeatedly applying Φ we obtain the sequence numerical solution $\{q_k\}_{k=0}^N$ to the particular problem characterized by initial values $y_0, \ldots, y_i$. The elements are given by

$$
y_{i+k} = \Phi^k(y_0, \ldots, y_i), \quad k = 1, \ldots, N - i,
\tag{2.4a}
$$

where

$$
\Phi^k(y_0, \ldots, y_i) = \big(\underbrace{\Phi \circ \Phi \circ \cdots \circ \Phi}_{k \text{ times}} \big)(y_0, \ldots, y_i)
\tag{2.4b}
$$

is the mapping Φ applied k times.

The object of geometric integration is to have the numerical flow Φ approximate the exact flow ϕ in such a way that the numerical flow has some of the same geometric properties as the exact flow. In Section 2.2 we compare the geometric properties of the numerical flow to the properties of the exact flow from Section 1.4.

2.1 Differential-Algebraic Equations and Index Reduction

Rabier and Rheimboldt [18] showed that nonholonomic mechanical systems are suited to be modeled by differential-algebraic equations.

Since we are interested in solving differential equations *on manifolds* we need a way to represent the manifold numerically. We saw in Section 1.1 that the configuration manifold Q may always be embedded into $\mathbb{R}^n$, in which case Q may be defined as by a constraint function so that

$$
Q = \{x \in \mathbb{R}^n \mid g(x) = 0\}, \quad g : \mathbb{R}^n \to \mathbb{R}^m.
$$

and that a nonholonomic problem may be stated in extrinsic form by equation (1.19) which we restate here:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\dfrac{\partial L}{\partial \dot{x}} - \dfrac{\partial L}{\partial x} = \nabla g(x)\,\lambda + A(x)^{\mathrm{T}}\,\mu, \\ \qquad\quad A(x)\,\dot{x} = 0, \\ \qquad\qquad\; g(x) = 0. \end{cases} \tag{2.5}$$

There are no explicit differential equations for λ and μ, they have to be found via the constraint equations $g(x) = 0$ and $A(x)\,\dot{x} = 0$.

The problem above is a *differential-algebraic* equation. A differential-algebraic equation is a system of algebraic equations and differential equations that must be solved together. Sometimes we will abbreviate the term differential-algebraic equation to *DAE*.

Index Reduction and the Underlying ODE

The most straightforward approach to solving any differential-algebraic equation is by a technique known as *index reduction*. Index reduction involves reducing the the *differentiation index*, which is the number of times the DAE has to be differentiated to yield an ODE. Hairer et. al [8, Chapter VII.1] sums this up.

Definition 2.1 (Differentiation Index). *An implicit differential equation*

$$F(\dot{y}, y) = 0 \tag{2.6}$$

has differentiation index m *if* m *is the minimal number of analytical differentiations*

$$F(\dot{y}, y) = 0, \quad \frac{\mathrm{d}}{\mathrm{d}t}F(\dot{y}, y) = 0, \quad \cdots, \quad \frac{\mathrm{d}^m}{\mathrm{d}t^m}F(\dot{y}, y) = 0,$$

of (2.6) *such that we may extract by algebraic manipulations an explicit ordinary differential system* $\dot{y} = f(y)$. *This system is called the* underlying ODE.

The underlying ODE is a reformulation of the DAE from which it is derived. They have the same analytical solutions.

The ODE obtained by index reduction will be defined for any set of initial conditions $y_0 \in \mathbb{R}^n$. By differentiation the statement $g(y) = 0$ is equivalent to

$$\frac{\partial g}{\partial y}(y)\,\dot{y} = 0, \quad g(y_0) = 0, \tag{2.7}$$

and vice versa (2.7) guarantees that $g(y(t)) = 0\ \forall\ t$. This situation always occurs in index reduction, the original constraint is turned into a constraint *and* a restriction $g(y_0) = 0$ on the initial condition.

We now show that a constrained Lagrangian system on extrinsic form is a differential-algebraic equation of index 3 when the system holonomic constraints only. When the system has nonholonomic constraints only it has index 2. A system with both types of constraints has index 3.

We consider equation the extrinsic equations of motion of equation (1.19) as a set of first order ordinary differential equations by setting

$$u = \dot{x},$$

and consider a mechanical Lagrangian on the form

$$L(x, u) = T(x, u) - U(x).$$

In this case equation (2.5) may be restated as

$$\dot{x} = u, \tag{2.8a}$$
$$\dot{u} = f(x, u, \lambda, \mu), \tag{2.8b}$$
$$0 = g(x), \tag{2.8c}$$
$$0 = A(x)\, u, \tag{2.8d}$$

where $f(x, u, \lambda, \mu)$ is given by

$$\frac{\partial^2 T}{\partial u^2}(x, u)\, f(x, u, \lambda, \mu) = \frac{\partial T}{\partial x}(x, u) - \frac{\partial U}{\partial x}(x) + \frac{\partial g}{\partial x}(x)\, \lambda + A(x)^{\mathrm{T}}\, \mu - \frac{\partial^2 T}{\partial x \partial u}(x, u).$$

We will now show that one may find the underlying ODE of equation (2.8) by differentiating relevant constraints (2.8c) or (2.8d).

We consider first a Lagrangian system with holonomic constraints only. The system is given by (2.8a–c). By differentiating both sides of the constraint we obtain

$$0 = \frac{\mathrm{d}g}{\mathrm{d}t} = \frac{\partial g}{\partial x}\, u.$$

We suppress the arguments of g to make the formulas more readable. Two differentiations of the holonomic constraint gives

$$0 = \frac{\mathrm{d}^2 g}{\mathrm{d}t^2} = \frac{\partial g}{\partial x}\, f + \left(\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial g}{\partial x} \right) u,$$

where we now suppress the arguments of f and g. The above relation involves λ, since λ is one of the arguments of f. Thus we expect to obtain an ODE in λ in the third differentiation of the holonomic constraint:

$$0 = \frac{\mathrm{d}^3 g}{\mathrm{d}t^3} = \frac{\partial g}{\partial x} \left(u \frac{\partial}{\partial x} + f \frac{\partial}{\partial u} + \dot{\lambda} \frac{\partial}{\partial \lambda} + \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial g}{\partial x} \right) f + \frac{\mathrm{d}}{\mathrm{d}t} \left(\left(\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial g}{\partial x} \right) u \right). \tag{2.8e}$$

Equation (2.8e) is the result of differentiating the holonomic constraint (2.8c) three times. It is an ODE in λ. As long as the matrix

$$\frac{\partial g}{\partial x} \frac{\partial f}{\partial \lambda}$$

is invertible, we have obtained a full set of first order ODEs in (x, u, λ) given by the equations (2.8a–b,e). This system is the underlying ODE of the holonomic system (2.8a–c).

We now consider a Lagrangian system of differential-algebraic equations with nonholonomic constraints only given by the equations (2.8a–b,d). To simplify the derivation we state the constraint on the general form $\omega(x, u) = 0$. By differentiating the nonholonomic constraint we obtain

$$0 = \frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{\partial \omega}{\partial x}\, u + \frac{\partial \omega}{\partial u}\, f,$$

where we have suppressed the arguments of ω and f to make the formulas more readable. The above relation involves μ, since μ is one of the arguments of f. Differentiating the nonholonomic constraint again gives

$$0 = \frac{\mathrm{d}^2\omega}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \omega}{\partial x}\, u\right) + \left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \omega}{\partial u}\right) f + \frac{\partial \omega}{\partial u}\left(u\frac{\partial f}{\partial x} + f\frac{\partial f}{\partial u} + \dot{\mu}\frac{\partial f}{\partial \mu}\right). \qquad (2.8f)$$

The above equation was obtained by differentiating the nonholonomic constraint (2.8d) two times. It is an ODE in μ. As long as the matrix

$$\frac{\partial \omega}{\partial u}\frac{\partial f}{\partial \mu}$$

is invertible, we have obtained a full set of first order ODEs in (x, u, μ) given by equations (2.8a–b,f). This system is the underlying ODE of the nonholonomic Lagrangian system given by (2.8a–b,d).

For a Lagrangian system that has both holonomic and nonholonomic constraints equation (2.8c) must be differentiated three times to give an ODE in λ. Therefore the differentiation index will be 3 as in the holonomic case.

Index Reduction and Numerical Integration

Standard numerical methods such as the ones that are part of MATLAB solve ordinary differential equations or differential-algebraic equations of index 1.

When index reduction is applied to an index 3 problem such as a holonomically constrained mechanical system one can will see that the numerical solution will *drift* away from the constraint manifold Q given by $g(x) = 0$. The error in the constraint will grow quadratically, while the derivative of the constraint grows linearly. This phenomenon is explained in [10, Chapter VII.2]. For index 2 and index 3 DAE the drift is so strong that the long-time integration of such systems is impossible without using techniques other than index reduction.

In the numerical experiments of this thesis two "classical" methods based on index reduction are used to provide a backdrop and a comparison for the mechanical integrators.

Index reduction techniques have the advantage that they are standardized, easy to implement and may be of high order.

2.2   Mechanical Integrators

In a numerical simulation of a physical system it can be argued that some errors are more grave than others. As an example, integrators that do not preserve energy will give totally wrong solutions to the problems they are integrating. In a simulation of the solar system an integrator that does not preserve energy in some sense will fail abysmally either by predicting that all planets fall into the sun, or by predicting that all planets eventually escape the gravity well of the sun and never return. Classical numerical methods are often unusable if one wishes to study the long-time behavior of a physical system.

Geometric integrators [8] or *mechanical integrators* [15] are names used for any numerical integrator that attempts to "do better" than classical methods by preserving qualitative properties of the exact flow in the numerical flow.

A simple form of geometric integrator is a *reversible* integrator. A reversible integrator has a numerical flow Φ which is a reversible map in the sense of definition 1.16. By applying a reversible integrator where the exact flow is reversible, we *preserve* the quantitative trait of reversibility in the numerical solution. This is the general idea of geometric integrators.

As we saw in Section 1.4 and Table 1.1 in that section, the flow of a holonomic Lagrangian system is characterized by the conservation of energy, the conservation of the symplectic form, and the conservation of first integrals.

A property of the exact flow such as the conservation of energy may be enforced by *projecting* the solution in each step onto a manifold defined by $E(x, \dot{x}) = E(x_0, \dot{x}_0)$, however in many cases heavy-handedness such as this can completely wreck the numerical solution instead of improving it! [8, Section IV.4]. This type of problems leads us to rather seek integrators that naturally conserve geometric invariants of the flow.

It has been shown [23] that finding a fixed-time-step integrator that simultaneously preserves energy, the symplectic form and momentum equations actually would be the same as finding the exact flow ϕ of the system [15, Chapter 7.1]. Since we cannot expect to be able to obtain the exact solution by a numerical method, we must make a choice about which quantities of the flow that we desire to preserve.

We shall look at a class geometrical integrators that preserve momentum maps and the symplectic form, *symplectic-momentum* integrators. The numerical flow Φ of symplectic-momentum integrators conserve *exactly* momentum maps and the symplectic form, so energy cannot be exactly conserved at the same time.

The choice of a symplectic-momentum integrator for integrating nonholonomic systems, in which neither the symplectic form nor momentum maps are conserved might seem queer. The reason for making this choice is that the variational symplectic-momentum integrators are able to reflect the structure of the evolution of both the symplectic form and the momentum equations. [15, Chapter 7.5].

These integrators are obtained from a direct discretization Hamilton's principle. We shall eventually see that although these integrators do not conserve energy, they actually conserve energy far better than classical integrators.

2.3 Variational Integrators

There has been many attempts at the development of a discrete mechanics and corresponding integrators [12, p. 360]. We will follow some of the exposition given in [12] (for holonomic systems) and [15, Chapter 7] (for nonholonomic systems).

The discrete flow of systems coming from a discrete variational principle enjoy many of the same characteristics as their continuous counterparts. The aim of this section is to introduce discrete counterparts to many of the constructions of Lagrangian variational mechanics, and to find numerical methods which are based on discretizing these principles themselves instead of the equations of motion that may be derived from e.g. the Lagrangian via Hamilton's principle.

We replace the Lagrangian $L : TQ \to \mathbb{R}$ with the *discrete Lagrangian*

$$L_h : Q \times Q \to \mathbb{R}. \tag{2.9}$$

A *discrete curve* is is defined by a set of points $\{q_k\}_{k=1}^N$; it is the counterpart of a continuous curve $q(t), t \in [t_0, t_N]$. The action integral of a continuous curve is replaced by a sum of values of the discrete Lagrangian evaluated in different points, and Hamilton's principle becomes:

Definition 2.2 (Discrete Hamilton's Principle). *Given a set of points $\{q_k\}_{k=0}^N$. The correct evolution of the discrete curve with fixed endpoints q_0 and q_N is the discrete curve that extremizes the* discrete action sum

$$S(q_0, q_1, \ldots, q_N) = \sum_{k=0}^{N-1} L_h(q_k, q_{k+1}),$$

where L_h is the discrete Lagrangian of the system.

The physically correct discrete curve extremizes the action sum, so we must have that

$$\frac{\partial S}{\partial q_k}(q_0, q_1, \ldots, q_N) = 0, \quad k = 1, 2, \ldots, N - 1. \tag{2.10}$$

Since the discrete Lagrangian is a function of two points only, each partial derivative of (2.10) has two terms only. The above equation is therefore equivalent to the set of equations

$$\frac{\partial}{\partial q_k} L_h(q_{k-1}, q_k) + \frac{\partial}{\partial q_k} L_h(q_k, q_{k+1}) = 0, \quad k = 1, 2, \ldots, N - 1. \tag{2.11}$$

So far nothing has been said about the relation between the discrete Lagrangian and the continuous Lagrangian of Section 1.2, and indeed if one is purely interested in studying discrete mechanics there need not be a connection.

We, however, are interested in turning this approach into a numerical integrator. We consider as in [12, Section 1.6] a particular discrete Lagrangian which we call the *exact discrete Lagrangian*.

Given a continuous Lagrangian $L(q, \dot{q})$, the corresponding exact discrete Lagrangian is

$$L_h^{\mathrm{E}}(q_k, q_{k+1}) = \int_{t_k}^{t_{k+1}} L(q, \dot{q}) \, \mathrm{d}t.$$

For the exact discrete Lagrangian the action sum becomes

$$\begin{aligned}
S(q_0, q_1, \ldots, q_N) &= \sum_{k=0}^{N-1} L_h^{\mathrm{E}}(q_k, q_{k+1}) \\
&= \sum_{k=0}^{N-1} \int_{t_k}^{t_{k+1}} L(q, \dot{q}) \, \mathrm{d}t \\
&= \int_{t_0}^{t_N} L(q, \dot{q}) \, \mathrm{d}t \\
&= G(q).
\end{aligned}$$

The action sum of the exact discrete Lagrangian equals the action integral of the corresponding continuous Lagrangian. This means that a point of the discrete curve $\{q_k\}_{k=0}^{N}$ must lie *on* the corresponding exact solution curve, and that

$$q_k = q(t_k), \quad k = 0, 1, \ldots, N.$$

We now consider as in [12, Section 1.1] a discrete Lagrangian that approximates the action integral on the curve segment of the exact solution given by $q(t)$, $t \in [t_k, t_{k+1}]$:

$$L_h(q_k, q_{k+1}) \approx \int_{t_k}^{t_{k+1}} L(q, \dot{q}) \, \mathrm{d}t.$$

We expect that for small time-steps $h = t_{k+1} - t_k$, the action sum will be approximately equal to the action integral,

$$G(q) \approx S(q_0, q_1, \ldots, q_N).$$

This is the general idea behind variational integrators.

Equation (2.11) leads itself to an integrator

$$\begin{aligned}
\Phi : Q \times Q &\to Q \times Q, \\
(q_{k-1}, q_k) &\mapsto (q_k, q_{k+1}),
\end{aligned} \tag{2.12a}$$

in which q_{k+1} is implicitly given by

$$\frac{\partial L_h}{\partial q_k}(q_k, q_{k+1}) + \frac{\partial L_h}{\partial q_k}(q_{k-1}, q_k) = 0. \tag{2.12b}$$
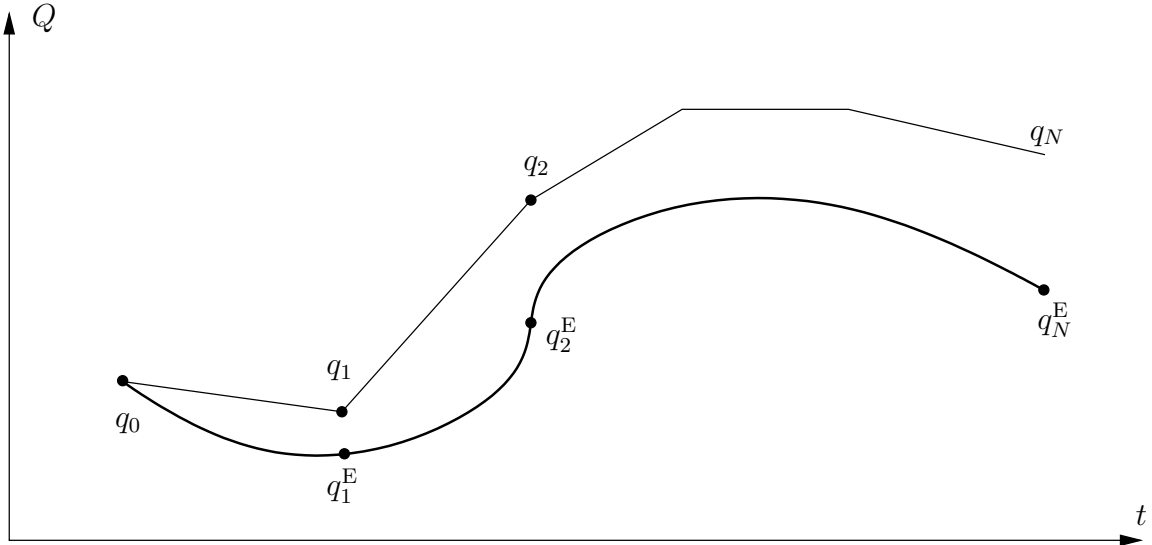
Figure 2.1: The exact solution curve $q(t)$ of the continuous Lagrangian and the sequence $\{q_k^{\mathrm{E}}\}_{k=0}^N$ obtained from the exact discrete Lagrangian coincide in the sense that $q_k^{\mathrm{E}} = q(t_k)$. An approximate solution $\{q_k\}_{k=0}^N$ based on a discrete Lagrangian and the integrator (2.12) is also displayed.

Assuming that equation (2.12) has a solution, we may use it as a numerical integrator. Given the starting values (q_0, q_1) we can find q_{k+1} from equation (2.4). By solving the implicit equation (2.12b) in each time-step we find $\{q_k\}_{k=2}^N$ from $(q_k, q_{k+1}) = \Phi^k(q_0, q_1)$.

Monforte [15, Section 7.4] presents two simple discretizations of a continuous Lagrangian. A simple discrete Lagrangian with a free parameter α is

$$L_h^\alpha(q_k, q_{k+1}) = hL\left((1 - \alpha)q_k + \alpha q_{k+1}, \frac{q_{k+1} - q_k}{h}\right). \tag{2.13}$$

This discretization approximates the action integral on the interval $[t_k, t_{k+1}]$ by evaluating the continuous Lagrangian at a single point. Figure 2.2 gives a graphical interpretation of the approximation.

Another "symmetric" possibility is

$$L_h^{\mathrm{sym}}(q_k, q_{k+1}) = \frac{h}{2}L\left(q_k, \frac{q_{k+1} - q_k}{h}\right) + \frac{h}{2}L\left(q_{k+1}, \frac{q_{k+1} - q_k}{h}\right). \tag{2.14}$$

This discretization approximates the action integral by evaluating the continuous Lagrangian in two points. Figure 2.2 gives a graphical interpretation of the approximation.

So far we have not considered nonholonomically constrained systems. For that we need a discretization D_h of the constraint distribution D and a discretization of the Lagrange-d'Alembert principle as well as a discretization of the Lagrangian. The discrete constraint distribution is given by

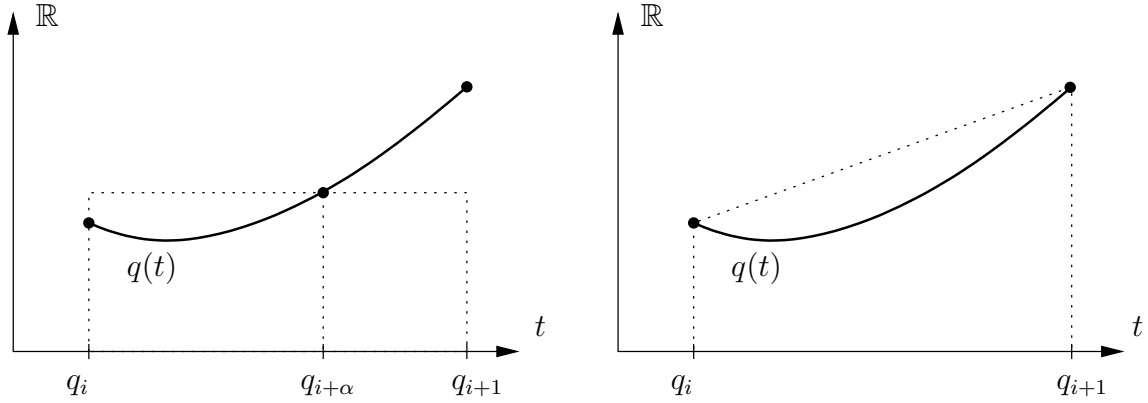$$D_h = \{(q_k, q_{k+1}) \in Q \times Q \mid \omega_h(q_k, q_{k+1}) = 0\}. \tag{2.15}$$

Figure 2.2: Approximation to the action integral $G(q)$ on the interval $[t_k, t_{k+1}]$ by the discrete Lagrangians L_h^α and L_h^{sym}. (Left): L_h^α approximates $q(t)$ by a the constant $q_{i+\alpha} = (1-\alpha)q_k + \alpha q_{k+1}$. (Right): L_h^{sym} approximates $q(t)$ by a linear function through q_k and q_{k+1}. Both method approximate $\dot{q}(t)$ by the constant $(q_{k+1} - q_k)/h$.

We will postpone the precise definition of ω_h to the next section. Monforte [15, Section 7.3] states the discrete Lagrange-d'Alembert principle:

Definition 2.3 (Discrete Lagrange-d'Alembert Principle). *Consider a discrete Lagrangian system (L_h, Q) with $L_h : Q \times Q \to \mathbb{R}$, a smooth distribution D and a discretization D_h of D, and a discrete curve $\{q_k\}_{k=0}^N$ with fixed endpoints q_0 and q_N. The discrete curve that describes the physically correct evolution of the system is the curve that extremizes the action sum*

$$S(q_0, q_1, \ldots, q_N) = \sum_{k=0}^{N-1} L_h(q_k, q_{k+1})$$

amongst the variations that satisfy $\delta q_k \in D$ and $(q_k, q_{k+1}) \in D_h$.

The discrete Lagrange-d'Alembert principle is equivalent to the following set of equations:

$$\begin{cases} \dfrac{\partial L_h}{\partial q_k} L_h(q_k, q_{k+1}) + \dfrac{\partial L_h}{\partial q_k} L_h(q_{k-1}, q_k) = \lambda_k \, \omega(q_k), \\ \qquad\qquad\qquad\qquad \omega_h(q_k, q_{k+1}) = 0. \end{cases} \tag{2.16}$$

The above relation is the basis of all the variational nonholonomic integrators in this thesis. Note that both D and D_h, ω and ω_h are used in Definition 2.3 and equation (2.16).

We shall now take a closer look closer at what is required for the discrete constraint space and the discrete Lagrangian to make a coherent whole.

Compatibility of D_h and L_h

The process of obtaining an integrator from from a discrete Lagrangian becomes more obvious if we consider the discrete Lagrangian to be composition of the continuous Lagrangian

L and a set of functions Ψ_h^i so that

$$L_h = \sum_i \beta_i(L \circ \Psi_h^i). \tag{2.17}$$

In the case of a discrete Lagrangian $L_h : Q \times Q \to \mathbb{R}$ the functions Ψ_h^i are on the form

$$\Psi_h : Q \times Q \to TQ,$$

they use the pair $(q_k, q_{k+1}) \in Q \times Q$ to approximate the exact solution $(q, \dot{q}) \in TQ$.
    Considering the function

$$\Psi_h^\alpha(q_k, q_{k+1}) = \left((1 - \alpha)q_k + \alpha q_{k+1}, \frac{q_{k+1} - q_k}{h} \right),$$

the discrete Lagrangians L_h^α and L_h^{sym} may be stated on the form of (2.17) as

$$L_h^\alpha = L \circ \Psi_h^\alpha, \tag{2.18a}$$

and

$$L_h^{\text{sym}} = \frac{1}{2}L \circ \Psi_h^0 + \frac{1}{2}L \circ \Psi_h^1. \tag{2.18b}$$

We want the discrete constraint functions ω_h to reflect the particular discretization of L_h.
    We require that for a general discrete Lagrangian (2.17), the corresponding discretization of the constraint is given by

$$\omega_h = \sum_i \beta_i(\omega \circ \Psi_h^i).$$

This is called the *compatibility condition* between the discretization of the Lagrangian and the discretization of the constraints. The compatibility condition gives

$$\omega_h^\alpha = \omega \circ \Psi_h^\alpha,$$
$$\omega_h^{\text{sym}} = \frac{1}{2}\omega \circ \Psi_h^0 + \frac{1}{2}\omega \circ \Psi_h^1.$$

The compatibility condition ensures that the resulting integrators have the correct order. [15, p. 146].
    We are now in a position to create a variety of two-step integrators based on

$$\begin{cases} \dfrac{\partial L_h}{\partial q_k}(q_k, q_{k+1}) + \dfrac{\partial L_h}{\partial q_k}(q_{k-1}, q_k) = A(q_k)^{\text{T}}\lambda, \\ \qquad\qquad\qquad \omega_h(q_k, q_{k+1}) = 0. \end{cases} \tag{2.19}$$

In Chapter 3 we implement two algorithms based on (2.19) and the discrete Lagrangians L_h^α and L_h^{sym}. They are previously presented in [15, Chapter 7].

One-Step Methods and the Discrete Legendre Transformation

The numerical flow of the multi-step methods that we obtain from (2.19) are determined by pairs of values $(q_0, q_1) \in Q \times Q$, while the trajectories in the flow of the continuous Lagrangian system are given by a single point $(q, \dot{q}) \in TQ$.

In implementations one-step methods are often easier to implement than multi-step methods. It is easier to specify a point in TQ than several points in Q, to specify q_1 with enough accuracy to obtain a high order integrator requires that one uses another integrator to go from $(q_0, \dot{q}_0) \in TQ$ to $(q_0, q_1) \in Q \times Q$, this can defeat some of the purpose of a multi-step integrator.

In [12], Marsden and West use a device known as the *discrete Legendre transformation* to find one-step variational integrators for holonomic Lagrangian systems. The discrete Legendre transformation can be used to restate an integrator on the form (2.12) as a one-step method. The discrete Legendre transform is also mentioned in [8, Chapter VI.6].

Definition 2.4 (Forced Discrete Legendre Transforms)**.** *The forced discrete Legendre transform is given by*

$$
\begin{aligned}
F^+ : Q \times Q &\to T^*Q, \\
(q_k, q_{k+1}) &\mapsto (p_{k+1}, q_{k+1}), \\
p_{k+1} &= \frac{\partial L_h}{\partial q_{k+1}}(q_k, q_{k+1}) + f_h^+(q_k, q_{k+1}),
\end{aligned}
\tag{2.20a}
$$

and

$$
\begin{aligned}
F^- : Q \times Q &\to T^*Q, \\
(q_k, q_{k+1}) &\mapsto (p_k, q_k), \\
p_k &= -\frac{\partial L_h}{\partial q_k}(q_k, q_{k+1}) - f_h^-(q_k, q_{k+1}).
\end{aligned}
\tag{2.20b}
$$

The *discrete forces* f^- and f^+ are defined in [12, Section 3.2]. In our case they are given by

$$
\begin{aligned}
f^- &= A(q_k)^{\mathrm{T}} \lambda_k, \\
f^+ &= A(q_{k+1})^{\mathrm{T}} \lambda_{k+1}.
\end{aligned}
$$

By analogy with the integrator defined in [12, p. 446] and from the nonholonomic principle this gives a possible general one-step integrator

$$
\begin{cases}
p_k = -\dfrac{\partial L_h}{\partial q_k}(q_k, q_{k+1}) + A(q_k)^{\mathrm{T}} \lambda_k, \\[2mm]
p_{k+1} = \dfrac{\partial L_h}{\partial q_{k+1}}(q_k, q_{k+1}) - A(q_{k+1})^{\mathrm{T}} \lambda_{k+1}, \\[2mm]
0 = A(q_{k+1})\dot{q}_{k+1}.
\end{cases}
\tag{2.21}
$$

The first equation of (2.21) must be solved implicitly to find q_{k+1}. The second and third equations must be solved together for p_{k+1} and λ_{k+1}. In the constraint equation we have chosen to use the continuous Legendre transform (1.9) to find $\dot{q}_{k+1}$.

We remark that the article [12] considers holonomic systems exclusively, and we have not seen any attempt to reformulate the two-step nonholonomic integrators of type (2.19) as one-step integrators. In Chapter 3 we construct and test two one-step integrators based on equation (2.21) and the discrete Lagrangians of equation (2.13) and (2.14).

Chapter 3

Experiments

In this chapter we aim to compare the performance of a classical Runge-Kutta method with the variational integrators of Section 2.3. We shall begin by detailing four variational integrators and one classical integrator, as well as the algorithm we will use as the benchmark for the other integrators. We shall compare the performance of these integrators on three problems.

Since we do not know the exact solutions of the problems, it is useful to have a high accuracy numerical solution that we may regard as exact. To obtain this "exact" solution we use another numerical method, run with a very small step-length.

We use MATLAB, version 6.5 (R13), for all computations.

Classical Integrators

We use two classical integrators in this thesis. The first is `ode15s`, which is one of the standard methods of MATLAB:

Algorithm `ode15s`. *This is one of MATLAB's built-in methods for solving ordinary differential equations. **ode15s** integrates ODEs and index 1 DAE. Generally problems must be presented on the form*

$$M(y)\dot{y} = f(t, y). \tag{3.1}$$

The matrix $M(y)$ may be singular—then equation (3.1) is a DAE.

Monforte [15, sect. 7.6] uses another of MATLAB's built-in integrators, `ode113`, as a benchmark with which to compare the nonholonomic integrators. We prefer to use `ode15s` since it accepts index 1 DAE, which leads to less work in the implementation. The integrator has high order and it is run with a very low error tolerance. For the purposes of these experiments we may regard the integrator as exact. The algorithm behind `ode15s` is detailed in [20].

To apply `ode15s` we reduce the index of the system's equations of motion until we get an index 1 DAE. For the problems we study that have only one nonholonomic constraint and no holonomic constraints, this means differentiating the nonholonomic constraint one

31

time. This is elaborated upon in the description of each numerical experiment in the next sections.

Algorithm RK4. *Consider an ordinary differential equation $\dot{y} = f(y) \in \mathbb{R}^n$ and a point $y_k \in \mathbb{R}^n$. Find y_{k+1} by solving the set of explicit equations*

$$
\begin{cases}
g_i = f\left(t_0 + c_i h, y_0 + h \sum_{j=1}^{s} A_{ij} g_j\right) \\
y_1 = y_0 + h \sum_{i=1}^{s} b_i g_i
\end{cases}
\qquad
\begin{array}{c|cccc}
0 & & & & \\
1/2 & 1/2 & & & \\
1/2 & 0 & 1/2 & & \\
1 & 0 & 0 & 1 & \\
\hline
 & 1/6 & 1/3 & 1/3 & 1/6
\end{array}
\tag{3.2}
$$

where the table on the right side is the Butcher-tableaux *containing the coefficients of the method c, A and b^{T}.*

The other classical algorithm used in this thesis is the a standard 4th order Runge-Kutta method. It is explained in any book on the numerical solution of ODEs, e.g. [9, p. 138]. To apply RK4 we reduce the index of the equations of motion of the system until we have the underlying ODE. For the problems we study this means differentiating the nonholonomic constraint twice. We shall show these calculation in the description of each experiment in the next sections.

Multi-Step Variational Integrators

We now present two multi-step algorithms based on equation (2.19) and the discrete Lagrangians of equation (2.13) and (2.14). They are taken from [15, Chapter 7].

Algorithm Ld. *Consider a nonholonomic Lagrangian system $L : TQ \to \mathbb{R}$ with a constraint distribution given by $A(q)\,\dot{q} = 0$. From (2.19) and L_h^α, with $\alpha = 1/2$ from (2.13) we get*

$$
q_{k-1/2} = \frac{q_k + q_{k-1}}{2}, \qquad \dot{q}_{k-1/2} = \frac{q_k - q_{k-1}}{h},
$$

$$
q_{k+1/2} = \frac{q_{k+1} + q_k}{2}, \qquad \dot{q}_{k+1/2} = \frac{q_{k+1} - q_k}{h},
$$

and

$$
\left(\frac{h}{2}\frac{\partial}{\partial q} - \frac{\partial}{\partial \dot{q}}\right) L(q_{k+1/2}, \dot{q}_{k+1/2}) + \left(\frac{h}{2}\frac{\partial}{\partial q} + \frac{\partial}{\partial \dot{q}}\right) L(q_{k-1/2}, \dot{q}_{k-1/2}) = A(q_k)^{\mathrm{T}} \lambda,
$$

$$
A(q_{k+1/2})\,\dot{q}_{k+1/2} = 0.
$$

The above expressions determine the numerical flow $\Phi : Q \times Q \to Q \times Q$. The equations must be solved implicitly for q_{k+1}.

We also construct a symmetric algorithm based on the discrete Lagrangian L_h^{sym}. The algorithm is otherwise similar to Ld.

Algorithm LdS. *Consider a nonholonomic Lagrangian system $L : TQ \to \mathbb{R}$ with a constraint distribution given by $A(q)\,\dot{q} = 0$. From (2.19) and (2.14) we get a numerical flow Φ given by*

$$\dot{q}_{k-1/2} = \frac{q_k - q_{k-1}}{h}, \qquad \dot{q}_{k+1/2} = \frac{q_{k+1} - q_k}{h},$$

and

$$\frac{1}{2}\left(\frac{\partial}{\partial q} - \frac{1}{h}\frac{\partial}{\partial \dot{q}}\right) L(q_k, \dot{q}_{k+1/2}) - \frac{1}{2h}\frac{\partial L}{\partial \dot{q}}(q_{k+1}, \dot{q}_{k+1/2})$$
$$+ \frac{1}{2h}\frac{\partial L}{\partial \dot{q}}(q_{k-1}, \dot{q}_{k-1/2}) + \frac{1}{2}\left(\frac{\partial}{\partial q} + \frac{1}{h}\frac{\partial}{\partial \dot{q}}\right) L(q_k, \dot{q}_{k-1/2}) = A(q_k)^{\mathrm{T}}\lambda,$$
$$\frac{1}{2}\Big(A(q_k)\,\dot{q}_{k+1/2} + A(q_{k+1})\,\dot{q}_{k+1/2}\Big) = 0.$$

The above equations must be solved implicitly for q_{k+1}.

The integrators Ld and LdS are second order accurate if the integrators are initialized with second order accurate values (q_0, q_1). We will use the integrator ode15s with a small time-step to initialize the methods. This is sufficient to obtain second order convergence.

Single-Step Variational Integrators

We now present two one-step methods which are based on equation (2.21) and the discrete Lagrangians $L_h^{1/2}$ and L_h^{sym}.

In the single-step formulations the continuous Legendre transform is also needed to express p as a function of $(q, \dot{q})$. This can be found either by solving the expression of the Legendre transform for $\dot{q}$, or by using the Legendre transform on the Hamiltonian formulation of the system. The relevant equations are equation (1.9) and (1.11):

$$p = \frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \quad \text{and} \quad \dot{q} = \frac{\partial H}{\partial p}(q, p). \tag{3.5}$$

Algorithm Hd. *Consider a nonholonomic Lagrangian system $L : TQ \to \mathbb{R}$ with constraints given by $A(q)\,\dot{q} = 0$. From equation (2.21) and the discrete Lagrangian $L_h^{1/2}$ of (2.13) we get a numerical flow $\Phi : TQ \to TQ$ given by*

$$q_{k+1/2} = \frac{q_{k+1} + q_k}{2}, \qquad \dot{q}_{k+1/2} = \frac{q_{k+1} - q_k}{h},$$

and

$$p_k = \left(-\frac{h}{2}\frac{\partial}{\partial q} + \frac{\partial}{\partial \dot{q}}\right) L(q_{k+1/2}, \dot{q}_{k+1/2}) + A(q_k)^{\mathrm{T}}\lambda_k,$$
$$p_{k+1} = \left(\frac{h}{2}\frac{\partial}{\partial q} + \frac{\partial}{\partial \dot{q}}\right) L(q_{k+1/2}, \dot{q}_{k+1/2}) - A(q_{k+1})^{\mathrm{T}}\lambda_{k+1},$$
$$0 = A(q_{k+1})\,\dot{q}_{k+1}.$$

where $\dot{q}_{k+1}$ is the generalized velocity in the point (q_{k+1}, p_{k+1}), found analytically from the Legendre transform as in equation (3.5). q_{k+1} and $\dot{q}_{k+1}$ are found from solving the above equations.

We also construct a symmetric algorithm based on the discrete Lagrangian L_h^{sym}. The algorithm is otherwise similar to `Hd`:

Algorithm `HdS`. *Consider a nonholonomic Lagrangian system $L : TQ \rightarrow \mathbb{R}$ with constraints given by $A(q)\,\dot{q} = 0$. From equation (2.21) and the discrete Lagrangian L_h^{sym} of (2.14) we get a numerical flow $\Phi : TQ \rightarrow TQ$ given by (2.14),*

$$\dot{q}_{k+1/2} = \frac{q_{k+1} - q_k}{h} \tag{3.7a}$$

and

$$p_k = \left(-\frac{h}{2}\frac{\partial}{\partial q} + \frac{1}{2}\frac{\partial}{\partial \dot{q}} \right) L(q_k, \dot{q}_{k+1/2}) - \frac{1}{2}\frac{\partial L}{\partial \dot{q}}(q_{k+1}, \dot{q}_{k+1/2}) + A(q_k)^{\mathrm{T}}\,\lambda_k, \tag{3.7b}$$

$$p_{k+1} = \frac{1}{2}\frac{\partial L}{\partial \dot{q}}(q_k, \dot{q}_{k+1/2}) + \left(\frac{h}{2}\frac{\partial}{\partial q} + \frac{1}{2}\frac{\partial}{\partial \dot{q}} \right) L(q_{k+1}, \dot{q}_{k+1/2}) - A(q_{k+1})^{\mathrm{T}}\,\lambda_{k+1}, \tag{3.7c}$$

$$0 = A(q_{k+1})\,\dot{q}_{k+1}. \tag{3.7d}$$

where $\dot{q}_{k+1}$ is the generalized velocity in the point (q_{k+1}, p_{k+1}), found analytically from the Legendre transform as in equation (3.5).

q_{k+1} and $\dot{q}_{k+1}$ are found from solving the above equations. equations (3.7c–d) to find p_{k+1} and λ_{k+1}.

Solving Nonlinear Systems of Equations

In the variational algorithms `Ld`, `LdS`, `Hd` and `HdS` the values q_{k+1} and/or $\dot{q}_{k+1}$ are found by solving implicit equation of type $f(y_k, y_{k+1}) = 0$ for y_{k+1}. We use MATLAB's implicit equation solver `fsolve` to solve these equations. `fsolve` is part of MATLAB's "Optimization Toolbox"; its inner workings are beyond the scope of this thesis. We simply state that `fsolve` finds the solution to possibly nonlinear equations on the form $f(y) = 0$ within a prescribed tolerance. In this thesis the tolerance has been set to `tol` $= 10^{-12}$ and we consider the output to be exact. Reasonable variations of `tol` does not affect the results given in this thesis.

The `fsolve` algorithm is described in MATLAB's documentation, the Optimization Toolbox User's Guide [13, p. 7-98–7-111] and the references therein.

3.1 The Nonholonomic Particle

The *nonholonomic particle* is a particularly simple example of a nonholonomic system. We consider a point particle in Euclidean space $\mathbb{R}^3$. The nonholonomic particle is used as an example thorough [15, sect. 4.2, 4.4, 7.6]. We aim here to reproduce the results of [15, sect. 7.6.1].

Because we are in $\mathbb{R}^3$ it is convenient to switch between the generalized coordinate notation and the standard coordinates names, we say that $q = (x, y, z)$ and $\dot{q} = (\dot{x}, \dot{y}, \dot{z})$ and use these notations interchangably.

The nonholonomic particle has kinetic energy which depends only on velocity, and moves in a quadratic potential. The kinetic energy $K(\dot{q})$ and the potential energy $U(q)$ are given by

$$K(\dot{q}) = \frac{1}{2}\dot{q}^{\mathrm{T}}\dot{q}, \qquad U(q) = x^2 + y^2.$$

The system's Lagrangian becomes

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^{\mathrm{T}}\dot{q} - (x^2 + y^2). \tag{3.8a}$$

The particle is subject to a nonholonomic constraint given by

$$0 = \dot{z} - y\dot{x} = \begin{pmatrix} -y & 0 & 1 \end{pmatrix} \dot{q}. \tag{3.8b}$$

Index Reduction

To integrate this system with `RK4` and `ode15s` the equations of motion must first be stated as a differential-algebraic equation. We use the Lagrange-d'Alembert principle to obtain the constrained equations of motion

$$\begin{cases} \ddot{x} = -2x - \lambda y, \\ \ddot{y} = -2y, \\ \ddot{z} = \lambda, \end{cases} \tag{3.9a}$$

where λ is given by the constraint

$$0 = \dot{z} - y\dot{x}. \tag{3.9b}$$

The system (3.9a–b) is an index 2 DAE.

To obtain an index 1 formulation of (3.9) we differentiate (3.9b), which yields a relation in λ,

$$0 = \lambda - \dot{y}\dot{x} + 2xy + \lambda y^2. \tag{3.9c}$$

The equations (3.9a,c) form an index 1 DAE. By using the standard techniqur of rewriting the equations of motion as a first order system we obtain a system on form $M\dot{y} = f(y)$ is

obtained from (3.9a) and (3.9c):

$$
\begin{pmatrix}
1 & & & & & & \\
& 1 & & & & & \\
& & 1 & & & & \\
& & & 1 & & & \\
& & & & 1 & & \\
& & & & & 1 & \\
& & & & & & 0
\end{pmatrix}
\begin{pmatrix}
\dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{u} \\ \dot{v} \\ \dot{w} \\ \dot{\lambda}
\end{pmatrix}
=
\begin{pmatrix}
u \\
v \\
w \\
-2x - \lambda y \\
-2y \\
\lambda \\
\lambda - uv + 2xy + \lambda y^2.
\end{pmatrix}
$$

This system is used in the integrator `ode15s`.

By differentiating (3.9c) we find and ODE in λ given by

$$
0 = (1 + y^2)\dot{\lambda} + 4(\dot{x}y + x\dot{y}) + 3\lambda y \dot{y},
$$

or by solving for $\dot{\lambda}$,

$$
\dot{\lambda} = -\frac{4(\dot{x}y + x\dot{y}) + 3\lambda y \dot{y}}{1 + y^2}. \tag{3.9d}
$$

The equations (3.9a,d) are the underlying ODE of the particle problem. We may now integrate this problem as an index 0 problem. We obtain a set of ODEs from equations (3.9a) and (3.9d):

$$
\begin{pmatrix}
\dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{u} \\ \dot{v} \\ \dot{w} \\ \dot{\lambda}
\end{pmatrix}
=
\begin{pmatrix}
u \\
v \\
w \\
-2x - \lambda y \\
-2y \\
\lambda \\
-\left(4uy + 4xv + 3\lambda yv\right)/(1 + y^2).
\end{pmatrix}
$$

This system is used in the integrator `RK4`.

Variational Integration

The variational integrators `Ld`, `LdS`, `Hd` and `HdS` do not require the system to be formulated as an ODE or DAE. Instead we must provide the partial derivatives of the Lagrangian

$$
\frac{\partial L}{\partial q} = \begin{pmatrix} -2x \\ -2y \\ 0 \end{pmatrix}, \quad \text{and} \quad \frac{\partial L}{\partial \dot{q}} = \dot{q}, \tag{3.10}
$$

and the matrix

$$
A(q) = \begin{pmatrix} -y & 0 & 1 \end{pmatrix} \tag{3.11}
$$

from the constraint $\omega(q, \dot{q}) = A(q)\,\dot{q} = 0$.

The one-step methods also require that we use the Legendre transform to find $\dot{q}$ since the one-step methods store (q, p) but not $\dot{q}$. We find from equation (3.5) that the Legendre transform is given by

$$
p = \dot{q}. \tag{3.12}
$$

Numerical Experiments

We plot some of the results of two numerical experiments on the nonholonomic particle. First we wish to verify the order of the methods.

Experiment Setup 3.1 (Order Verification in the Nonholonomic Particle Problem). *We integrate the nonholonomic particle problem given by* (3.13) *over a constant interval using 20 different step sizes.*
   We pick the initial condidions

$$q_0 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \dot{q}_0 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \lambda_0 = 0,$$

and let the integration take place over the time interval $[t_0, t_{max}]$ *given by*

$$t_0 = 0, \quad t_{max} = 5.$$

We test each integrator with $k = 20$ *different numbers of steps in increments between* $N_1 = 20$ *and* $N_k = 1000$*. The step sizes* h_i *are given by*

$$h_i = \frac{t_{max} - t_0}{N_i - 1}.$$

The numbers of steps N_i *are chosen so that they will appear equispaced on a logarithmic scale as in Figure 3.1.*

$$N_i = \lfloor \alpha^{i-1} N_1 \rfloor, \quad \alpha = \sqrt[k-1]{N_k/N_1}, \quad i = 1, 2, \ldots, k.$$

We run Experiment 3.1 with each of the integrators `Ld`, `LdS`, `Hd`, `HdS` as well as the classical integrator `RK4`. For each step size h_i we compare the generalized coordinates found in the final time-step $q_N \in Q$ with the nearly exact value of $q(t_{\max})$ we obtain from running the `ode15s` run with a tolerance `tol` $= 10^{-12}$. We plot the error of each method in Figure 3.1. The plot shows clearly that for this example all the variational integrators are sencond order. The plot focuses on the work-presicion curves of the variational integrators, so the graph of the Runge-Kutta method `RK4` is not shown in it's entirety.

We remark that `RK4` is of order 4 in *all* the variables $(q, \dot{q}, \lambda)$, and that the one-step methods `Hd` and `HdS` are of second order in both q and p of $(q, p) \in T^*Q$. We still do only plot the error in the generalized coordinates $q \in Q$ of each integrator. This is because the generalized coordinates are the only quantities that are found by the 2-step methods `Ld` and `LdS`. Because of this they are the only quantities that may be used to make comparisons between all the methods.

We next study, as in [3, sect. 7.6], a long-time integration of the nonholonomic particle.

Experiment Setup 3.2 (Long-Time Integration of the Nonholonomic Particle Problem). *We integrate the nonholonomic particle problem given by* (3.13) *over a long time interval in order to study the geometric properties of the different integrators.*
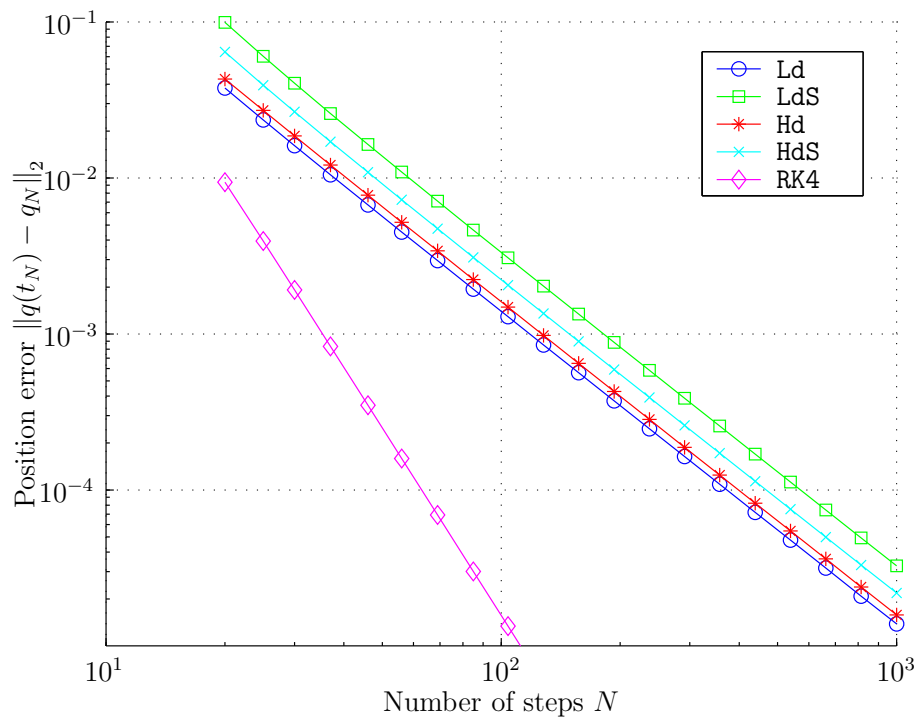
Figure 3.1: Work-Precision plot for the particle problem. The integrators `Ld`, `LdS`, `Hd` and `HdS` are all of second order. The integrator `RK4` is of order 4. The setup behind this plot is explained in Experiment 3.1.

We pick the initial condidions

$$q_0 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \dot{q}_0 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \lambda_0 = 0,$$

and let the integration take place over the time interval $[t_0, t_{max}]$ *given by*

$$t_0 = 0, \quad t_{max} = 250.$$

We set the number of timesteps

$$N = 1250$$

and thereby the step length

$$h = 0.2.$$

We present the results of this experiment in several plots:

Figure 3.2 shows the solution curve in Q obtained from `RK4`. The plot of each component of q shows some structure.

Figure 3.3 shows that the constraint (3.8b) is not exactly conserved by `RK4`, which is to be expected since there is no special mechanism that enforces the constraint in each time-step. Also the conservation of energy along the flow is not preserved.

Figure 3.4 shows the solution obtained from the algorithm `Ld`. This plot shows an apparent periodicity of the solution which was not present in Figure 3.2. The plots of the solution curve obtained from the high-accuracy solution found by `ode15s` and the plots of the solution curves of `LdS`, `Hd` and `HdS` are so similiar to Figure 3.4 that we omit them. Instead we compare the constraint satisfaction and energy conservation properties of the variational integrators.

Figure 3.5 and Figure 3.6 show that the variational integrators conserve the constraint. The small fluctuations seen in the constraint function are noise-like, and they are a consequence of the fact that the stepping equations of the algorithms are solved by an implicit equation solver.

Energy is not exactly conserved by the variational integrators, but we observe that the energy still is well behaved—it remains close to the initial energy.
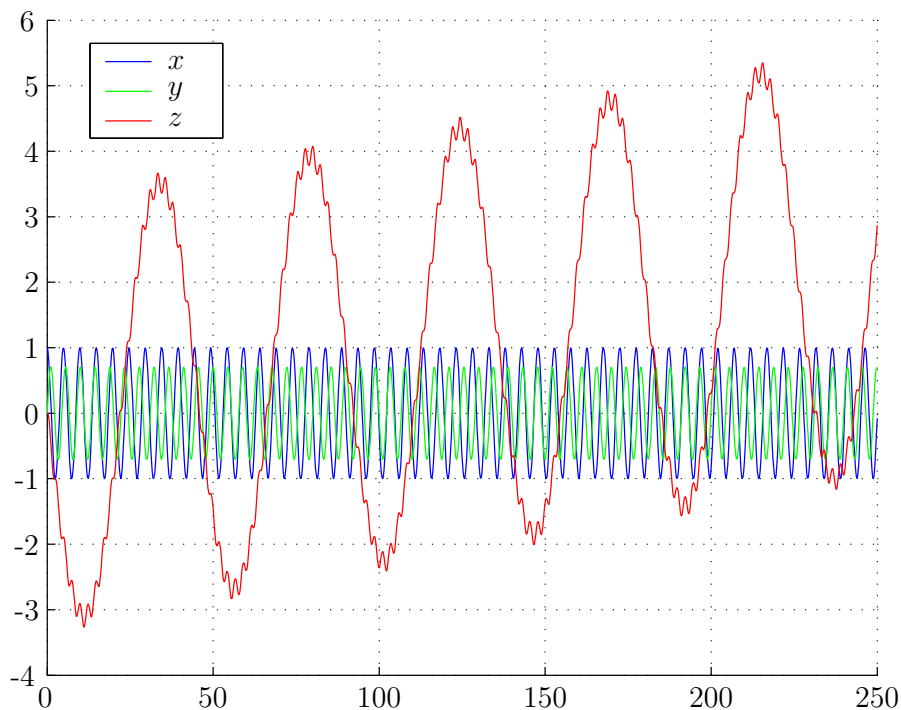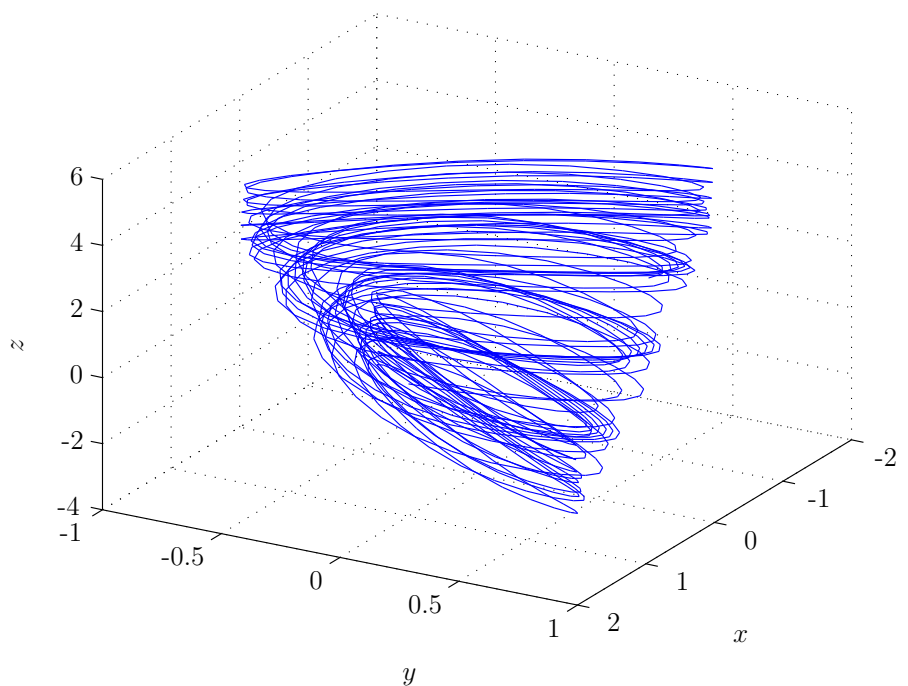
Figure 3.2: Long-time integration of the nonholonomic particle by RK4. (Top): The solution is plotted as a curve in $\mathbb{R}^3$. There is no clear structure apparent. (Bottom): The generalized coordinate componets of the solution plotted individually. A clear structure is apparent. The setup behind this plot is explained in Experiment 3.2.
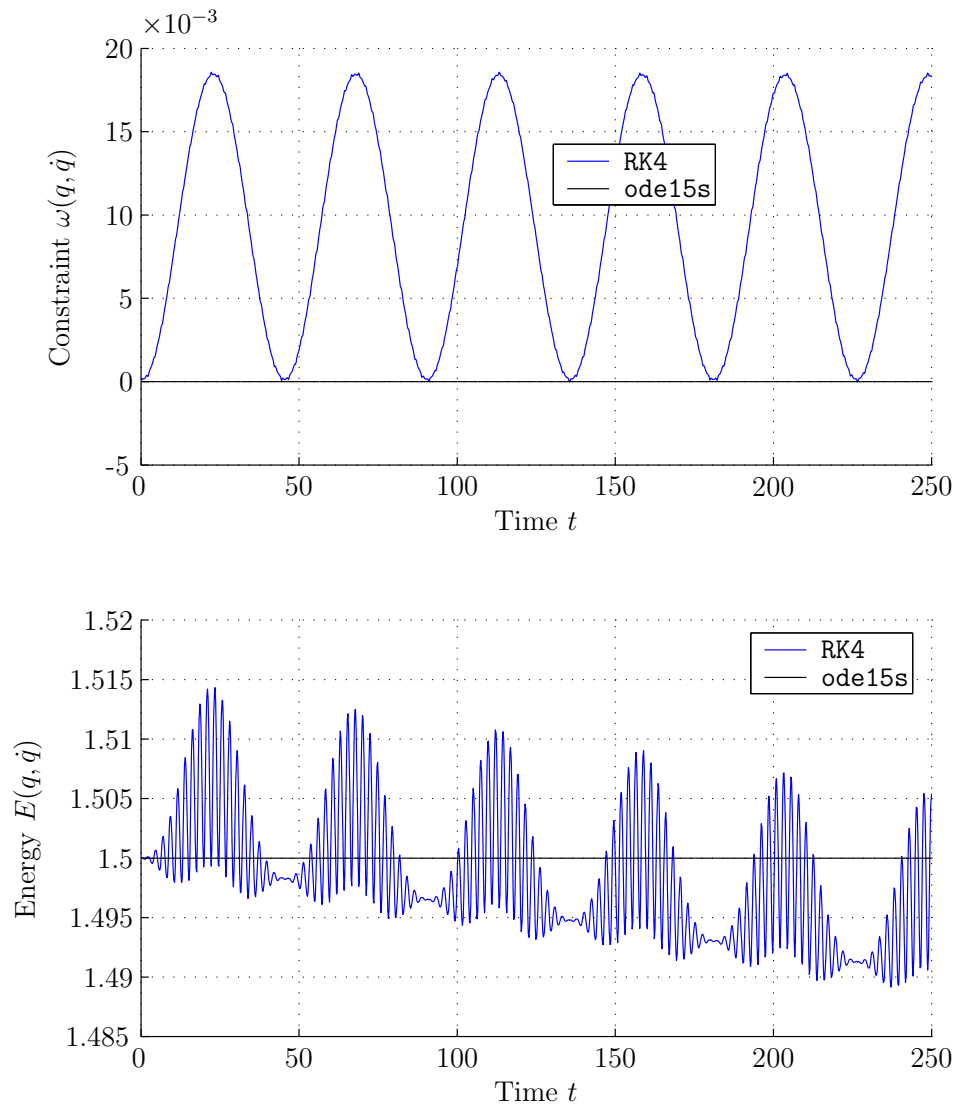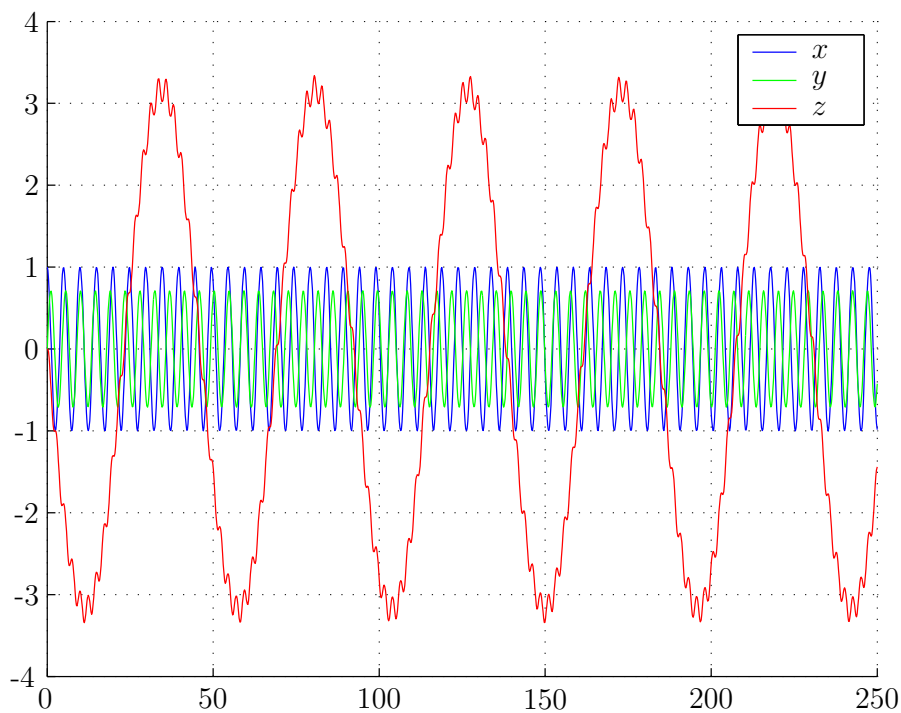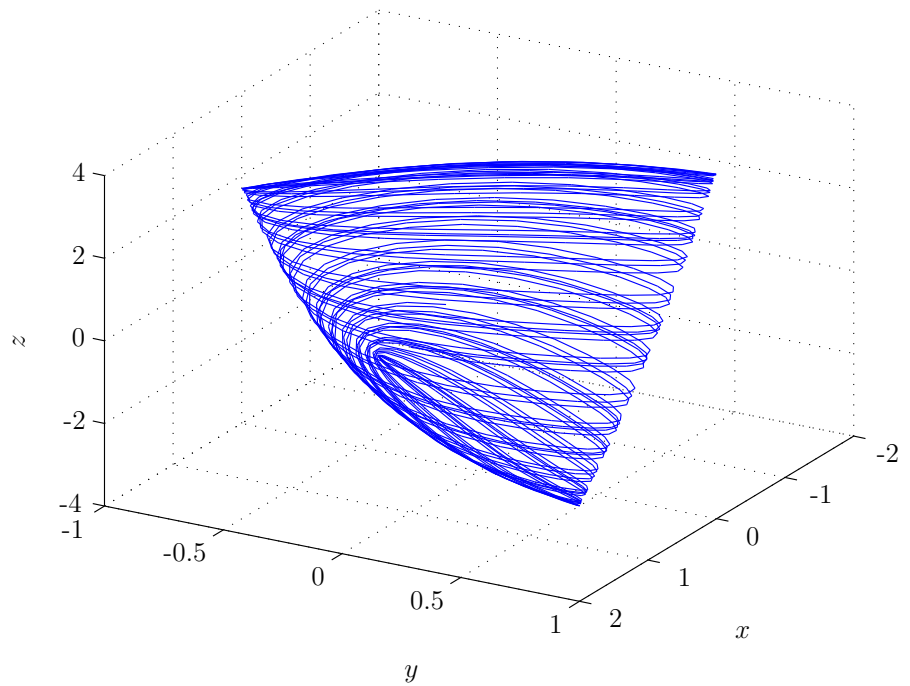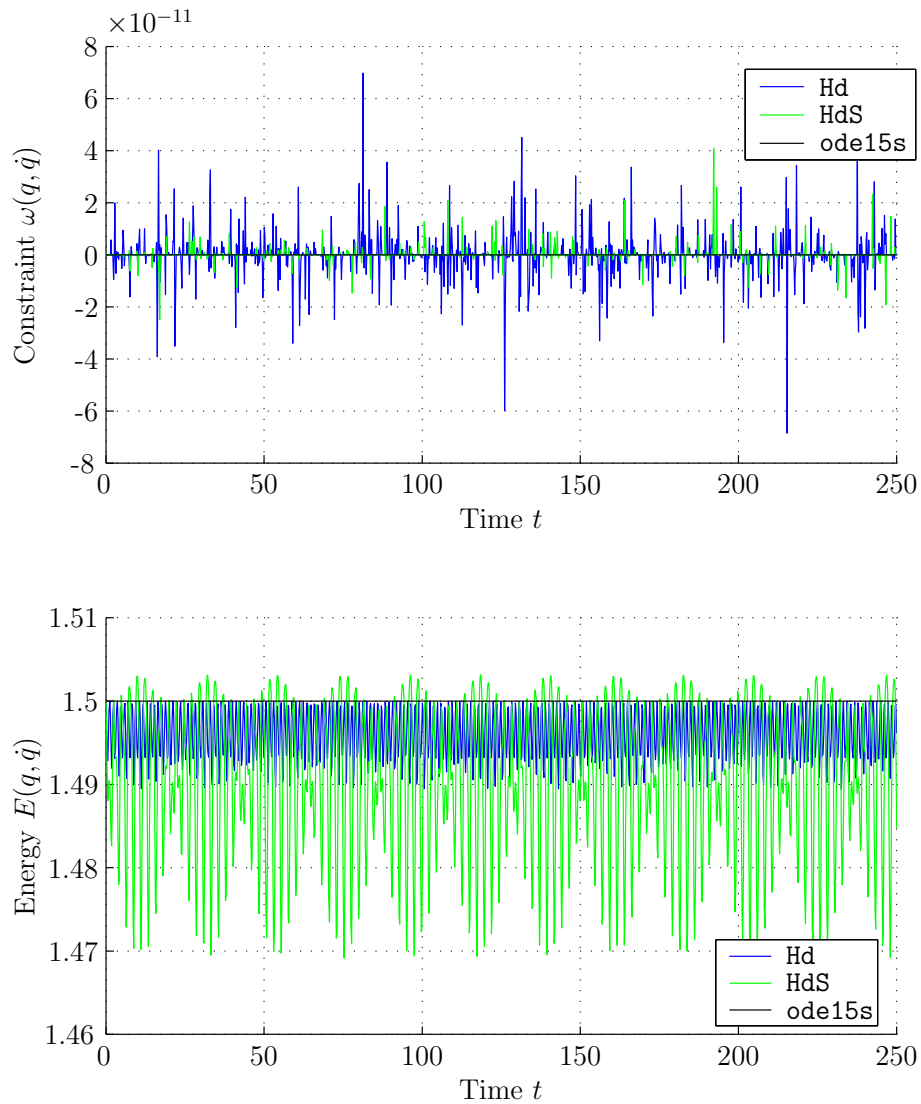
Figure 3.3: Constraint satisfaction and energy by `RK4` of in the long-time integration of the nonholonomic particle. (Top): The nonholonomic constraint is not enforced by the algorithm, it does however appear bounded. (Bottom): Energy is not conserved by the classical algorithm. The energy fluctuates along with the z-component of q, and there is a net loss of energy. The setup behind this plot is explained in Experiment 3.2.

Figure 3.4: Long-time integration of the nonholonomic particle by Ld. (Top): The solution plotted as a curve in $\mathbb{R}^3$. A clear structure is apparent, the solution must remain on a subset of Q determined by the initial conditions. (Bottom): The components of the solution plotted individually. A clear structure is apparent in the plot, and each component of the solution appears bounded. The setup behind this plot is explained in Experiment 3.2.

PSfrag replacements

PSfrag replacements
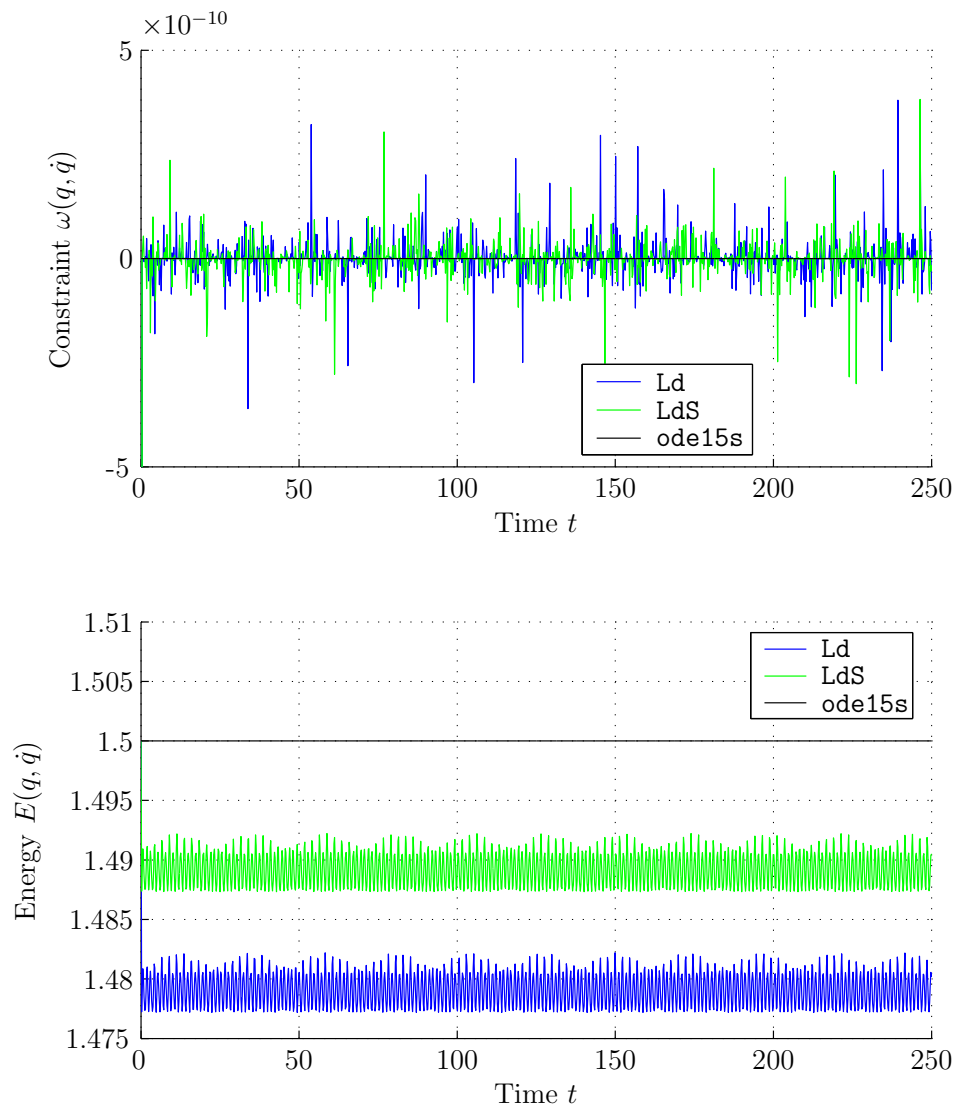
Figure 3.5: Constraint satisfaction and energy by `Ld` and `LdS` in the long-time integration of the nonholonomic particle. (Top): The nonholonomic constraint is enforced in each time-step. The numerical solutions show small and essentially noise-like variations. (Bottom): The variational integrators `Ld` and `LdS` do not conserve energy exactly, but still have good energy conservation properties. Both algorithms display a jump in energy in the first time-step, then they fluctuate around a constant value. This type of behaviour is also seen in [3, chapter 7]. The setup behind this plot is explained in Experiment 3.2.

Figure 3.6: Constraint satisfaction and energy by `Hd` and `HdS` in the long-time integration of the nonholonomic particle. (Top): The nonholonomic constraint is enforced in each time-step. The numerical solutions show small and essentially noise-like variations. (Bottom): The variational integrators do not conserve energy exactly, but still have good energy conservation properties. The energy fluctuates but appears to be bounded. The setup behind this plot is explained in Experiment 3.2.

3.2   McLachlan and Perlmutter's Particles

McLachlan and Perlmutter [14] give a higher-dimensional example of a nonholonomic system on Hamiltonian form. The configuration space is $\mathbb{R}^{2n+1}$; we will denote the generalized coordinate components by $q = (x, y_1, \ldots, y_n, z_1, \ldots, z_n)$, and the generalized momentum components by $p = (p_x, p_{y_1}, \ldots, p_{y_n}, p_{z_1}, \ldots, p_{z_n})$. The system is time-reversible and energy is conserved [14].

The energy of the system is given by the Hamiltonian

$$H(q,p) = \frac{1}{2}\left(p^{\mathrm{T}}p + q^{\mathrm{T}}q + z_1^2 z_2^2 + \sum_{i=1}^{n} y_i^2 z_i^2\right). \tag{3.15a}$$

There is a single nonholonomic constraint

$$p_x + \sum_{i=1}^{n} y_i p_{z_i} = \begin{pmatrix} 1 & 0 & \cdots & 0 & y_1 & \cdots & y_n \end{pmatrix} p = 0. \tag{3.15b}$$

In [14] it is shown that the numerical flow integrators that are reversible may exhibit a drift in energy $E(q,p) = H(q,p)$ that is comparable to a diffusion process. By testing the same system with variational nonholonomic integrators we want to show that variational integrators outperform integrators that are only reversible, by approximately conserving the energy over long time intervals.

Index reduction

Hamilton's equations gives the set of first oder equations

$$\begin{aligned}
\dot{q} &= p \\
\dot{p}_x &= -x + \lambda, \\
\dot{p}_{y_i} &= -y_i - y_i z_i^2, & i = 1, \ldots, n \\
\dot{p}_{z_i} &= \begin{cases} -z_1 z_2^2 - z_1 - y_1^2 z_1 + \lambda y_1, & i = 1 \\ -z_1^2 z_2 - z_2 - y_2^2 z_2 + \lambda y_2, & i = 2 \\ -z_i - y_i^2 z_i + \lambda y_i, & i = 3, \ldots, n. \end{cases}
\end{aligned} \tag{3.16a}$$

and the constraint

$$p_x + \sum_{i=1}^{n} y_i p_{z_i} = 0. \tag{3.16b}$$

The system (3.16a–b) is an index 2 DAE.

To obtain an index 1 formulation of (3.16) we differentiate the constraint (3.16b). This yields an expression involving λ,

$$\left(1 + \sum_{i=1}^{n} y_i^2\right)\lambda = x + y_1 z_1 z_2^2 + y_2 z_1^2 z_2 + \sum_{i=1}^{n}\left(y_i^3 z_i + y_i z_i - p_{y_i} p_{z_i}\right). \tag{3.16c}$$

The equations (3.16a,c) form an index 1 DAE which is used in the integrator `ode15s`.
By differentiating (3.16c) we obtain

$$\left(1 + \sum_{i=1}^{n} y_i^2\right)\dot{\lambda} + \left(\sum_{i=1}^{n} 2y_i p_{y_i}\right)\lambda =$$

$$p_x + (z_1 z_2^2 p_{y_1} + y_1 z_2^2 p_{z_1} + 2y_1 z_1 z_2 p_{z_2}) + (z_1^2 z_2 p_{y_2} + 2y_2 z_1 z_2 p_{z_1} + y_2 z_1^2 p_{z_2})$$

$$+ \sum_{i=1}^{n} \left(3y_i^2 z_i p_{y_i} + y_i^3 p_{z_i} - \dot{p}_{y_i} p_{z_i} - p_{y_i} \dot{p}_{z_i} + z_i p_{y_i} + y_i p_{z_i}\right). \quad (3.16d)$$

Equations (3.16a,d) are the index 0 formulation of the problem and are used in the integrator `RK4`.

Variational Integration

The system is given on Hamiltonian form so the Legendre transformation must be used to express the system on Lagrangian form. Because the Legendre transform (3.5) is simple:

$$\dot{q} = \frac{\partial H}{\partial p} = p, \quad (3.17)$$

it is easy to find the Lagrangian formulation of (3.15a):

$$L(q, \dot{q}) = \frac{1}{2}\dot{q}^{\mathrm{T}}\dot{q} - \frac{1}{2}\left(q^{\mathrm{T}}q + z_1^2 z_2^2 + \sum_{i=1}^{n} y_i^2 z_i^2\right). \quad (3.18)$$

The constraint (3.15b) becomes

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & y_1 & \cdots & y_n \end{pmatrix} \dot{q} = 0.$$

The partial derivatives of the Lagrangian are

$$\frac{\partial L}{\partial x} = -x,$$

$$\frac{\partial L}{\partial y_i} = -y_i - y_i z_i^2, \qquad i = 1, \ldots, n$$

$$\frac{\partial L}{\partial z_i} = \begin{cases} -z_1 z_2^2 - z_1 - y_1^2 z_1, & i = 1 \\ -z_1^2 z_2 - z_2 - y_2^2 z_2, & i = 2 \\ \qquad -z_i - y_i^2 z_i, & i = 3, \ldots, n, \end{cases}$$

and

$$\frac{\partial L}{\partial \dot{q}} = \dot{q}$$

for all components of the velocity vector $\dot{q} = (\dot{x}, \dot{y}_1, \ldots, \dot{y}_n, \dot{z}_1, \ldots, \dot{z}_n)$.

Numerical Experiments

As in Section 4.1 we start by verifying the order of the numerical methods. We start with a simple example taking $Q = \mathbb{R}^5$, which is the lowest value accepted by the problem.

Experiment Setup 3.3 (Order Verification in McLachlan & Perlmutter's problem). *We integrate the problem given by (3.15) over a constant interval using 20 different step sizes.*
  We pick the configuration manifold $Q = \mathbb{R}^5$ and initial conditions

$$q_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \dot{q}_0 = \begin{pmatrix} 0 & 1 & -1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}, \quad \lambda_0 = 1.$$

We let the integration take place over the time interval $[t_0, t_{max}]$ given by

$$t_0 = 0, \quad t_{max} = 2.$$

As in Experiment 3.1, we test each integrator with $k = 20$ different numbers of steps in increments between $N_1 = 20$ and $N_k = 1000$. The step sizes h_i are given by

$$h_i = \frac{t_{max} - t_0}{N_i - 1}.$$

The numbers of steps N_i are chosen so that they will appear equispaced on a logarithmic scale (see Figure 3.7).

$$N_i = \lfloor \alpha^{i-1} N_1 \rfloor, \quad \alpha = \sqrt[k-1]{N_k/N_1}, \quad i = 1, 2, \ldots, k.$$

Experiment Setup 3.4 (Integration of McLachlan & Perlmutter's Problem). *We integrate the problem given by (3.15) over a long time interval in order to study the geometric properties of the different integrators. We pick the configuration manifold $Q = \mathbb{R}^5$ and initial conditions*

$$q_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \dot{q}_0 = \begin{pmatrix} 0 & 1 & -1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}, \quad \lambda_0 = 1,$$

and let the integration take place over the time interval $[t_0, t_{max}]$ given by

$$t_0 = 0, \quad t_{max} = 10,$$

We set the number of time-steps

$$N = 1000$$

and thereby the step length

$$h = 0.2.$$

Figure 3.8 show the components of q found by `ode15s`. The solutions found by the other integrators are visually indistinguishable from this plot. We present the results of this experiment in several plots: Figure 3.8 shows the components of each particle. Figure 3.9 shows the energy conservation and constraint satisfaction of the integrator `RK4`. Figure 3.10 shows the constraint satisfaction and energy conservation of the variational integrators.

We now want to test the variational integrators on a complex problem where we expect the variational integrators to show qualitatively correct solutions while the classical integrator `RK4` fails:
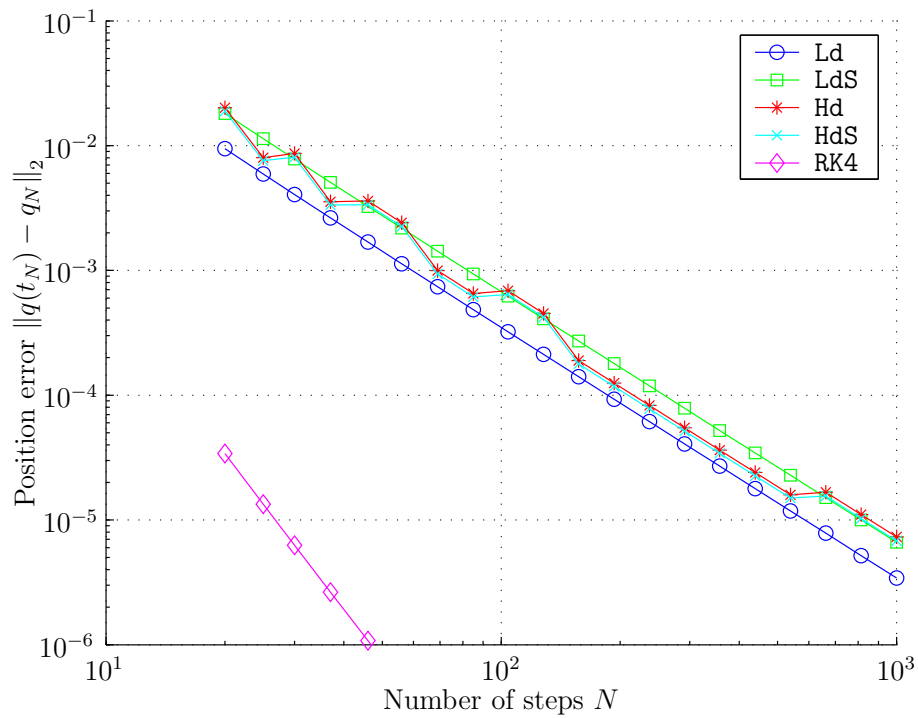
PSfrag replacements

Figure 3.7: Work-Precision plot for McLachlan and Perlmutter's particle problem. The integrators `Ld` and `LdS` are clearly of order 2. The algorithms `Hd` and `HdS` display a curious fluctuation between *two* lines that converge of order 2. The integrator `RK4` is of order 4. The setup behind this plot is explained in Experiment 3.3.
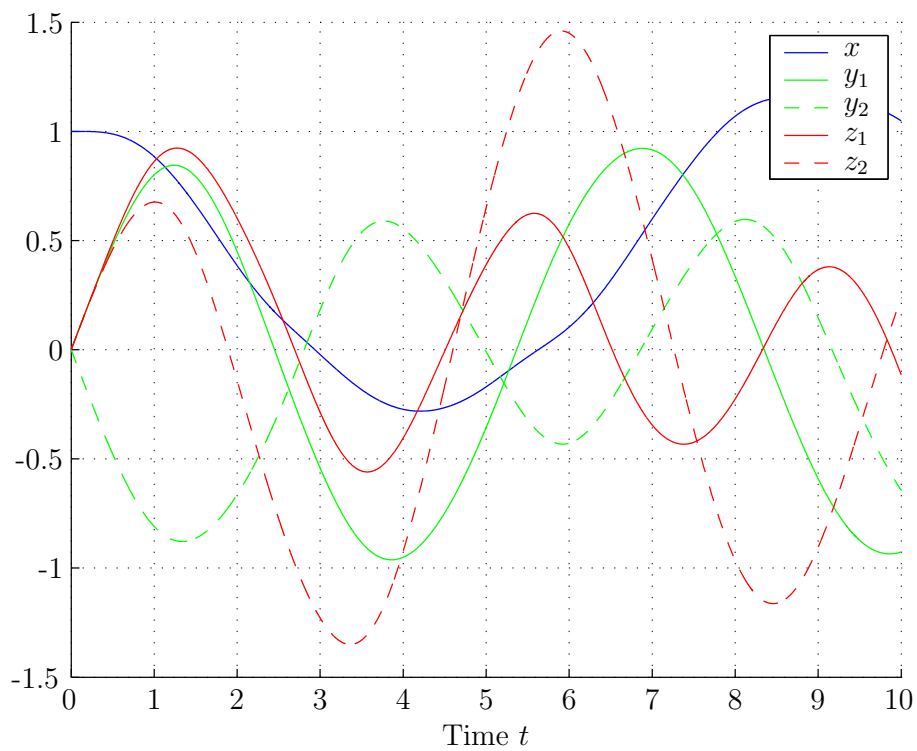
PSfrag replacements

Figure 3.8: Integration of McLachlan & Perlmutter's particles by `ode15s`. The plots of the solution by `Ld`, `LdS`, `Hd` and `HdS` are indistinguishable from this plot. The setup behind this plot is explained in Experiment 3.4.
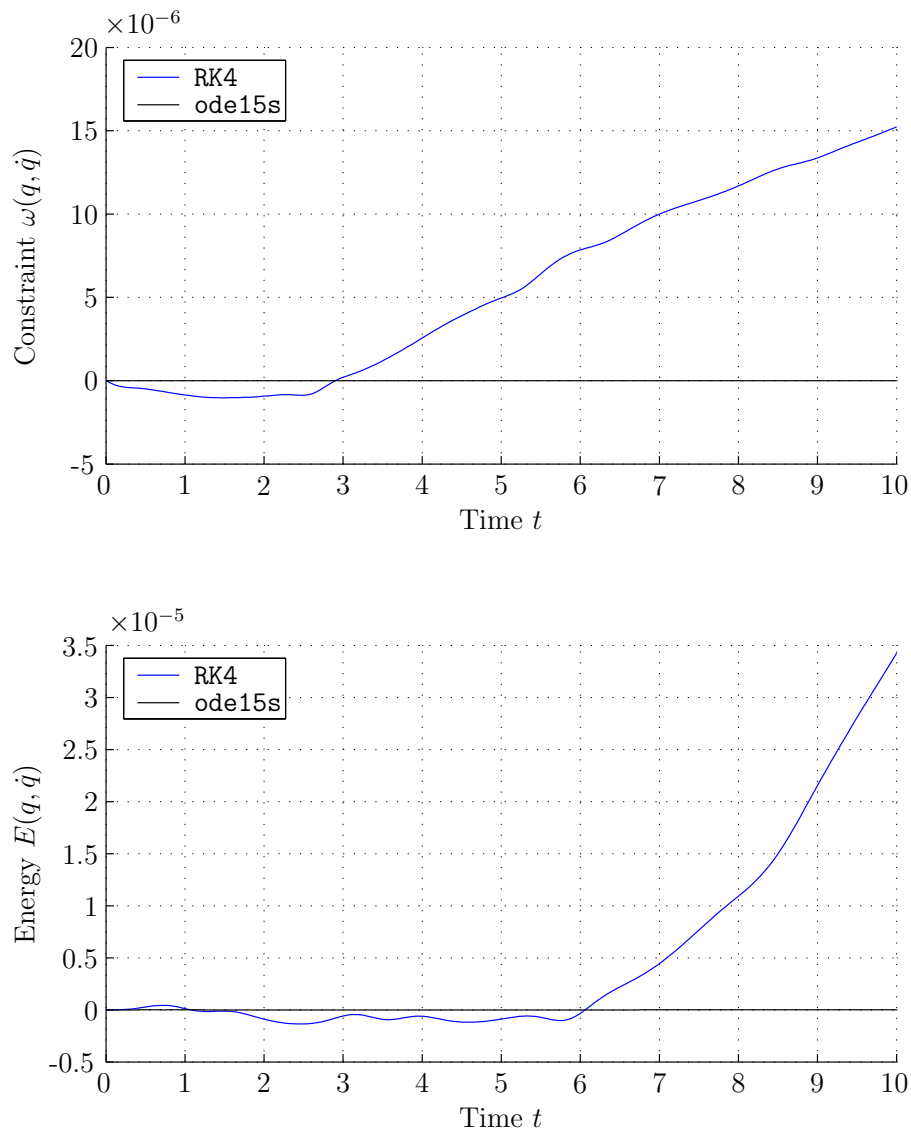
Figure 3.9: Constraint satisfaction and energy conservation by `RK4`. Both the constraint error and the energy error are small but increase rapidly. The setup behind this plot is explained in Experiment 3.4
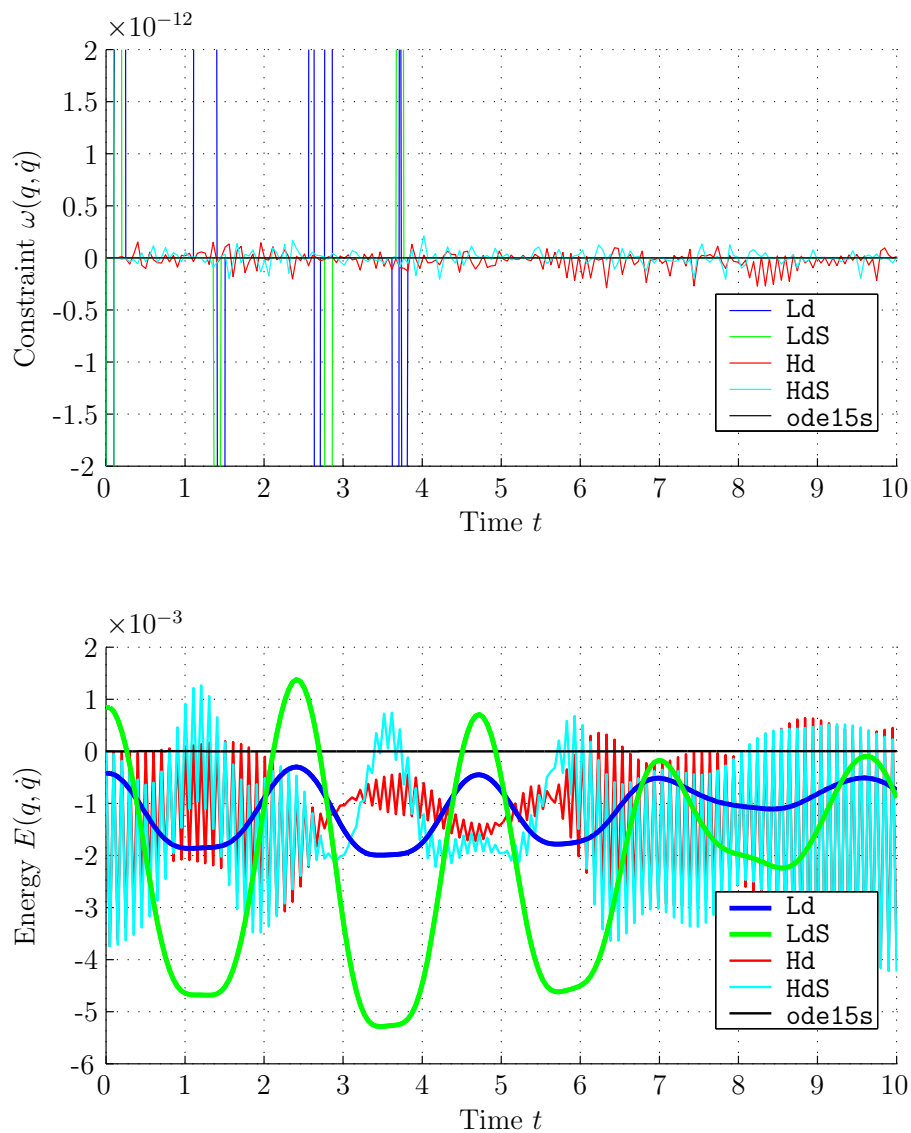
Figure 3.10: Energy and constraint satisfaction of the variational integrators Ld, LdS, Hd and HdS. The constraint is satisfied as should be expected. The energy of the two-step methods Ld and LdS appears bounded. The energy of the methods Hd and HdS oscillate wildly but appear bounded. The setup behind this plot is explained in Experiment 3.4

Experiment Setup 3.5 (Long-Time Integration of 26-Variable McLachlan and Perlmutter's problem). *We integrate the problem given by* (3.15) *over a long time interval in highlight the qualitative properties of the variational integrators.*

We pick the configuration manifold $Q = \mathbb{R}^{13}$ *and initial conditions*

$$x_0 = 1, \quad y_0 = \begin{pmatrix} 0 & -1 & 0 & 1 & -1 & 1 \end{pmatrix}, \quad z_0 = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \end{pmatrix},$$
$$\dot{x}_0 = 0, \quad \dot{y}_0 = \begin{pmatrix} 1 & 0 & -1 & 1 & 0 & -1 \end{pmatrix}, \quad \dot{z}_0 = \begin{pmatrix} 0 & 1 & -1 & 1 & 1 & 1 \end{pmatrix}$$

$$q_0 = \begin{pmatrix} x_0 & y_0 & z_0 \end{pmatrix}^{\mathrm{T}}, \quad \dot{q}_0 = \begin{pmatrix} \dot{x}_0 & \dot{y}_0 & \dot{z}_0 \end{pmatrix}^{\mathrm{T}}, \quad \lambda_0 = 0$$

and let the integration take place over the time interval $[t_0, t_{max}]$ *given by*

$$t_0 = 0, \quad t_{max} = 100,$$

We set the number of time-steps
$$N = 400$$

and thereby the step length
$$h = 0.25.$$

Figure **??** shows all 26 components of q for `RK4` and `HdS`. The numerical solution computed by `RK4` oscillates more and more. For longer integrations the step size has to be reduced, otherwise the method will fail by blowing up. The variational integrators seems unaffected by this kind of blowup.
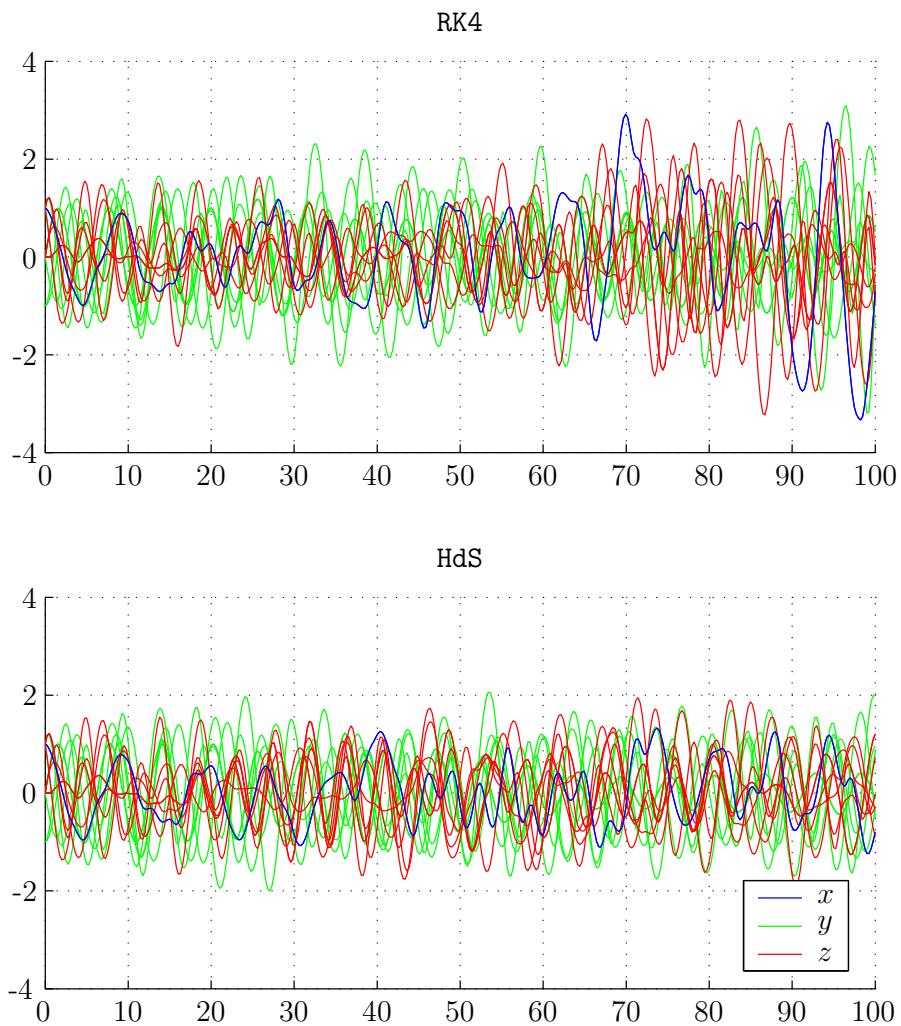
yd
xe
yf

RK4

HdS

Figure 3.11: Long-time integration of the 26-variable McLachlan and Perlmutter's problem. We compare the numerical solutions of RK4 and HdS. It is obvious that the solution by RK4 displays more powerful oscillations as time increases, while HdS does not amplify the oscillations. This indicates that the energy error of RK4 has a significant influence on the solution in this problem. The setup behind this plot is explained in Experiment 3.5.
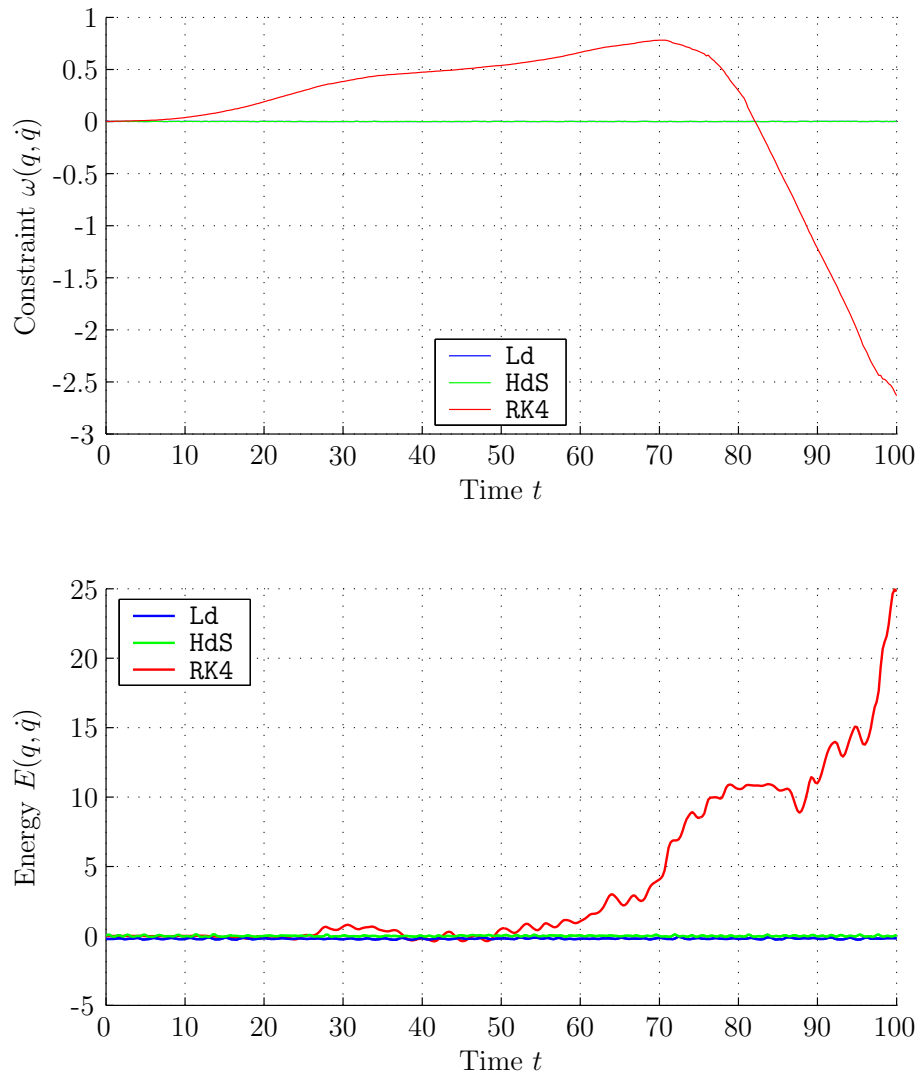
Figure 3.12: Energy error and constraint error by `RK4`, `Ld` and `HdS` in the long-time integration of 26-Variable McLachlan and Perlmutter's problem. This plot compares the performance of `RK4` and the variational integrators in the same scale of the axes. The lack of constraint satisfaction of `RK4` is very clear. The energy plot shows that even though `Ld` and `HdS` do not conserve energy exactly, their energy-conserving properties are far better than `RK4`. The energy conservation of `RK4` has failed totally. The setup behind this plot is explained in Experiment 3.5.

3.3   The Chaplygin Sleigh

The *Chaplygin sleigh* is a classical example in nonholonomic mechanics and is described in e.g. [3, 17]. The sleigh is a rigid body supported on a surface by three points. One of these is a knife edge that only allows motion along the edge, so any skidding motion perpendicular to the knife edge is not allowed. The other two points allow motion in any direction. A drawing of the Chaplygin sleigh is given in [3, p. 26].

The Lagrangian of the Chaplygin sleigh is composed of of momentum and angular momentum terms,

$$L = \frac{1}{2}m(\dot{x}_C^2 + \dot{y}_C^2) + \frac{1}{2}I\dot{\theta}^2 - U(x_C, y_C, \theta),$$

where (x_C, y_C) is the sleigh's centre of mass and θ is the orientation of the sleigh. The constants I, m and a are the sleigh's moment of inertia, the mass of the sleigh, and the distance between the centre of mass and the knife edge.

We follow the approach of [3] by rewriting the Lagrangian in terms of the knife edge coordinates,

$$x = x_C - a\cos\theta,$$
$$y = y_C - a\sin\theta.$$

This gives a Lagrangian with mixed terms,

$$L(q, \dot{q}) = \frac{1}{2}\left(m(\dot{x}^2 + \dot{y}^2) + (I + ma^2)\dot{\theta}^2 + 2ma\dot{\theta}(\dot{y}\cos\theta - \dot{x}\sin\theta)\right) - U(x, y, \theta). \quad (3.22a)$$

The configuration manifold of the sleigh is $Q = \mathbb{R}^2 \times \mathbb{S}^1$. As in equation (3.22a) we will denote a point of $q \in Q$ by (x, y, θ) *or* by q depending on what is more convenient.

The constraint of motion given by the knife edge becomes the nonholonomic constraint

$$\dot{y}\cos\theta - \dot{x}\sin\theta = 0. \quad (3.22b)$$

We consider the sleigh to be on an inclined plane, so the potential energy $U(x_C, y_C)$ becomes

$$U(x_C, y_C, \theta) = U(x + a\cos\theta, y + a\sin\theta, \theta) = mg(y + a\sin\theta). \quad (3.22c)$$

The Chaplygin sleigh is an example of a system with rotational symmetry. Because of the nonholonomic constraint we still do not have conservation of angular momentum. The nonholonomic constraint (3.22b) means that angular momentum and linear momentum interconnected. Taking v as the velocity of the sleigh at the knife edge, a balance of linear and angular momentum gives this momentum equation for angular momentum:

$$\dot{v} = a\dot{\theta}^2,$$
$$\ddot{\theta} = -\frac{ma}{I + ma}v\dot{\theta}.$$

The energy of the system *is* conserved. It is given by

$$E(q, \dot{q}) = \frac{1}{2}\left(m(\dot{x}^2 + \dot{y}^2) + (I + ma^2)\dot{\theta}^2 + 2ma\dot{\theta}(\dot{y}\cos\theta - \dot{x}\sin\theta)\right) + U(x, y, \theta).$$

Index Reduction

The equations of motion of the system may be found through the Lagrange-d'Alembert principle,

$$
\begin{cases}
\ddot{x} - a\cos\theta\dot{\theta}^2 - a\sin\theta\ddot{\theta} = -\lambda/m\sin\theta - U_x, \\
\ddot{y} - a\sin\theta\dot{\theta}^2 + a\cos\theta\ddot{\theta} = \lambda/m\cos\theta - U_y, \\
(I + ma^2)\ddot{\theta} + ma\dot{\theta}(\dot{x}\cos\theta + \dot{y}\sin\theta) = -U_\theta.
\end{cases}
\tag{3.23a}
$$

As always λ is found from the constraint (3.22b), restated here on matrix-vector form,

$$
\begin{pmatrix} -\sin\theta & \cos\theta & 0 \end{pmatrix} \dot{q} = 0.
\tag{3.23b}
$$

The system of equations (3.23a–b) is an index 2 DAE. By differentiating the constraint (3.23b) we get an expression involving λ,

$$
\frac{\lambda}{m} = \frac{I}{I + ma^2}\dot{\theta}\left(\dot{x}\cos\theta + \dot{y}\sin\theta\right) - \frac{a}{I + ma^2}U_\theta + \cos\theta U_y - \sin\theta U_x.
\tag{3.23c}
$$

The equations (3.23a,c) form an index 1 DAE. By using the standard technique of rewriting the equations of motion as a first order system we form an index 1 system on the form $M(y)\dot{y} = f(y)$ from equations (3.23a,c). The system is given by

$$
\begin{pmatrix}
1 & & & & & & \\
& 1 & & & & & \\
& & 1 & & & & \\
& & & 1 & & -a\sin\theta & \\
& & & & 1 & a\cos\theta & \\
& & & & & I + ma^2 & \\
& & & & & a & 0
\end{pmatrix}
\begin{pmatrix}
\dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{u} \\ \dot{v} \\ \dot{\omega} \\ \dot{\lambda}
\end{pmatrix}
=
\begin{pmatrix}
u \\
v \\
\omega \\
-\lambda/m\sin\theta + a\omega^2\cos\theta - U_x \\
\lambda/m\cos\theta + a\omega^2\sin\theta - U_y \\
-ma\omega(u\cos\theta + v\sin\theta) - U_\theta \\
\lambda/m - \omega(u\cos\theta + v\sin\theta) - U_y\cos\theta + U_x\sin\theta.
\end{pmatrix}
$$

These equations are used in the integrator `ode15s`.

By differentiating (3.23c) we find and ODE in λ given by

$$
\dot{\lambda} = \frac{Im}{I + ma^2}\left(\frac{-ma\dot{\theta}(\dot{x}\cos\theta + \dot{y}\sin\theta) - U_\theta}{I + ma^2}\left(\dot{x}\cos\theta + \dot{y}\sin\theta\right)\right)
$$
$$
+ \frac{Im}{I + ma^2}\dot{\theta}\left(a\dot{\theta}^2 - U_x\cos\theta - U_y\sin\theta - \dot{x}\dot{\theta}\sin\theta + \dot{y}\dot{\theta}\cos\theta\right)
$$
$$
+ \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{-am}{I + ma^2}U_\theta + mU_y\cos\theta - mU_x\sin\theta\right).
\tag{3.23d}
$$

The equations (3.23a,d) are the underlying ODE of the equations of motion of the Chaplygin sleigh. We restate the equations as

$$
\begin{pmatrix}
1 & & & & & & \\
& 1 & & & & & \\
& & 1 & & & & \\
& & & 1 & & -a\sin\theta & \\
& & & & 1 & a\cos\theta & \\
& & & & & I+ma^2 & \\
& & & & & & 1
\end{pmatrix}
\begin{pmatrix}
\dot{x} \\ \dot{y} \\ \dot{\theta} \\ \dot{u} \\ \dot{v} \\ \dot{\omega} \\ \dot{\lambda}
\end{pmatrix}
=
\begin{pmatrix}
u \\
v \\
\omega \\
-\lambda/m\sin\theta + a\omega^2\cos\theta - U_x \\
\lambda/m\cos\theta + a\omega^2\sin\theta - U_y \\
-ma\omega(u\cos\theta + v\sin\theta) - U_\theta \\
\text{Equation (3.23d).}
\end{pmatrix}
$$

This system is used in the classical integrator `RK4`. We remark that although this system is stated on the form $M(y)\dot{y} = f(y)$, the matrix $M(y)$ is not singular, and thus we might equivalently state $\dot{y} = M^{-1}(y)f(y)$. The notation above is for convenience only and is due to the length of the formulae.

Variational Integrators

The partial derivatives of the Lagrangian are

$$
\frac{\partial L}{\partial q} = \begin{pmatrix} 0 \\ 0 \\ -ma\dot{\theta}(\dot{x}\cos\theta + \dot{y}\sin\theta) \end{pmatrix}
$$

and

$$
\frac{\partial L}{\partial \dot{q}} = \begin{pmatrix} m(\dot{x} - a\dot{\theta}\sin\theta) \\ m(\dot{y} + a\dot{\theta}\cos\theta) \\ (I+ma^2)\dot{\theta} - ma(\dot{x}\sin\theta - \dot{y}\cos\theta). \end{pmatrix}
$$

The Legendre transform of (3.5) gives

$$
p = \frac{\partial L}{\partial \dot{q}} = \begin{pmatrix} m(\dot{x} - ma\dot{\theta}\sin\theta) \\ m(\dot{y} + ma\dot{\theta}\cos\theta) \\ (I+ma^2)\dot{\theta} - ma(\dot{x}\sin\theta - \dot{y}\cos\theta) \end{pmatrix}
$$

so by solving for $\dot{q}$ we get an expression of the type $\dot{q} = M(q)\,p$ with

$$
M = \frac{1}{Im}\begin{pmatrix}
I+ma^2 - ma^2(\cos\theta)^2 & -ma^2\sin\theta\cos\theta & ma\sin\theta \\
-ma^2\sin\theta\cos\theta & I+ma^2(\cos\theta)^2 & -ma\cos\theta \\
ma\sin\theta & -ma\cos\theta & m.
\end{pmatrix}
\tag{3.24}
$$

The transformation is never singular since the determinant of M is

$$
\det(M) = \frac{I+ma^2 - ma^2(\sin\theta)^2 - ma^2(\cos\theta)^2}{I^2m^2} = \frac{1}{Im^2}.
$$

Numerical Experiments

As with the nonholonomic particle we present two numerical experiments. To verify the order of the methods we integrate over a short time period, and to study the geometric properties of the integrators we integrate over a long time interval.

Experiment Setup 3.6 (Order Verification in the Chaplygin Sleigh Problem). *We integrate the Chaplygin sleigh problem given by (3.22) over a constant interval using 20 different step sizes.*

We pick the initial conditions

$$q_0 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \dot{q}_0 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \lambda_0 = 1/20,$$

and let the integration take place over the time interval $[t_0, t_{max}]$ given by

$$t_0 = 0, \quad t_{max} = 5.$$

We test each integrator with $k = 20$ different numbers of steps in increments between $N_1 = 10$ and $N_k = 1000$. The step sizes h_i are given by

$$h_i = \frac{t_{max} - t_0}{N_i - 1}.$$

The numbers of steps N_i are chosen so that they will appear equispaced on a logarithmic scale,

$$N_i = \lfloor \alpha^{i-1} N_1 \rfloor, \quad \alpha = \sqrt[k-1]{N_k/N_1}, \quad i = 1, 2, \ldots, k.$$

As with the previous two problems we run experiment 3.6 with each of the integrators `Ld`, `LdS`, `Hd`, `HdS` and `RK4`, and compare the solution to the solution of the high-precision `ode15s`. The error $\left\| \Phi_h^N(q_0) - \phi^{t_N}(q_0) \right\|$ of each method against the number of steps N_i. The result is displayed in Figure 3.13. The plot shows that while all the geometric integrators have order 2, the single-step integrators `Hd` and `HdS` do not appear as well behaved as the 2-step integrators. The error of each single-step integrators lies on *two* straight lines, each of which converges with order 2.

To study the geometric properties of the integrators we consider the a system that is again in the initial state given by the next experiment:

Experiment Setup 3.7 (Long-Time Integration of the Chaplygin Sleigh Problem). *We integrate the Chaplygin sleigh problem (3.22) over a long time interval in order to study the geometric properties of the different integrators.*

We pick the initial conditions

$$q_0 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \dot{q}_0 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}, \quad \lambda_0 = 1/20,$$

and let the integration take place over the time interval $[t_0, t_{max}]$ given by
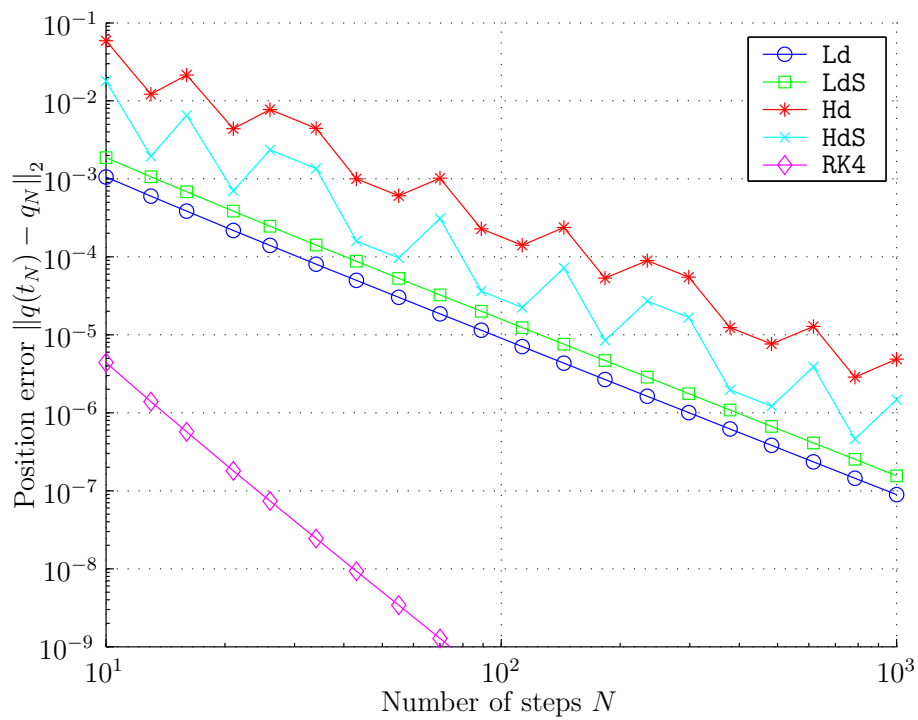
$$t_0 = 0, \quad t_{max} = 30.$$

Figure 3.13: Work-Precision plot for the Chaplygin sleigh. The integrators `Ld` and `LdS` are clearly of order 2. The algorithms `Hd` and `HdS` display a curious fluctuation between *two* lines that converge of order 2. The integrator `RK4` is of order 4. The setup behind this plot is explained in Experiment 3.6.
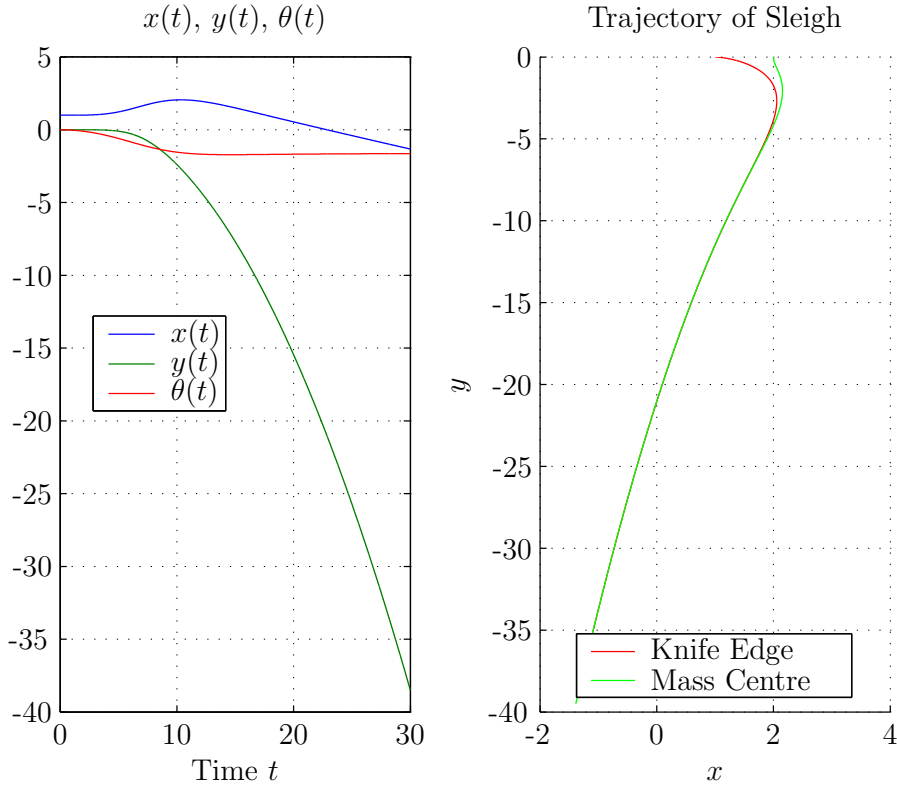
Figure 3.14: Long-time integration of the Chaplygin sleigh by `ode15s` (Left): The generalized coordinate components of the solution plotted individually. (Right): The solution is plotted as a curve in $\mathbb{R}^2$. The down direction of the incline is in the negative y direction. The constraint is clearly visible at the top of the trajectory, where the sleigh initially is oriented perpendicular to the incline direction. After a short time the sleigh goes into a free fall-like trajectory. The setup behind this plot is explained in Experiment 3.7.

We set the number of time-steps

$$N = 250$$

and thereby the step length

$$h = 0.12.$$

We present the results of this experiment in several plots:

Figure 3.14 shows the solution curve in Q obtained from `ode15s`. After a short time the solution settles into a "free fall" trajectory where $\dot{q}$ is increasing. The results from `RK4` and the multi-step integrators `Ld` and `LdS` are indistinguishable from Figure 3.14 so they are omitted.

The single-step integrators display some very surprising behavior. Figure 3.15 shows that `Hd` does not enter a state of stable velocity. Where the exact solution settles down, the numerical solution of `Hd` begins to oscillate. When the oscillations lessen and disappear, the solution is following a completely different trajectory. The behaviour is similar in the

symmetric single-step integrator `HdS`. Reducing the time-step makes the oscillations smaller but the "jump" to a different trajectory still happens.

Figure 3.16 show the constraint error and the energy error of of `Hd` and `HdS`. The constraint error is negligible as we have come to expect from a variational integrator. The energy error is negligible except around the time when the numerical solution undergoes heavy oscillations. Around the oscillations phase of the solutions both integrators display a big dip in energy. Energy error can be used as an indicator of total error in variational integrators, so this plot can be considered a warning that something is wrong.

Figure 3.17 shows that the constraint (3.8b) is not exactly conserved by `RK4`, which we should expect by now since there is no special mechanism that enforces the constraint in each time-step. Also the conservation of energy along the flow is not preserved. After the trajectory of the sleigh has stabilized the energy seems to settle on a constant value, while the constraint error grows linearly.

Figure 3.18 shows the constraint satisfaction and energy conservation of the integrators `Ld` and `LdS`. In this integration of the Chaplygin sleigh problem the classical method `RK4` displays far better energy conservation than the variational methods.
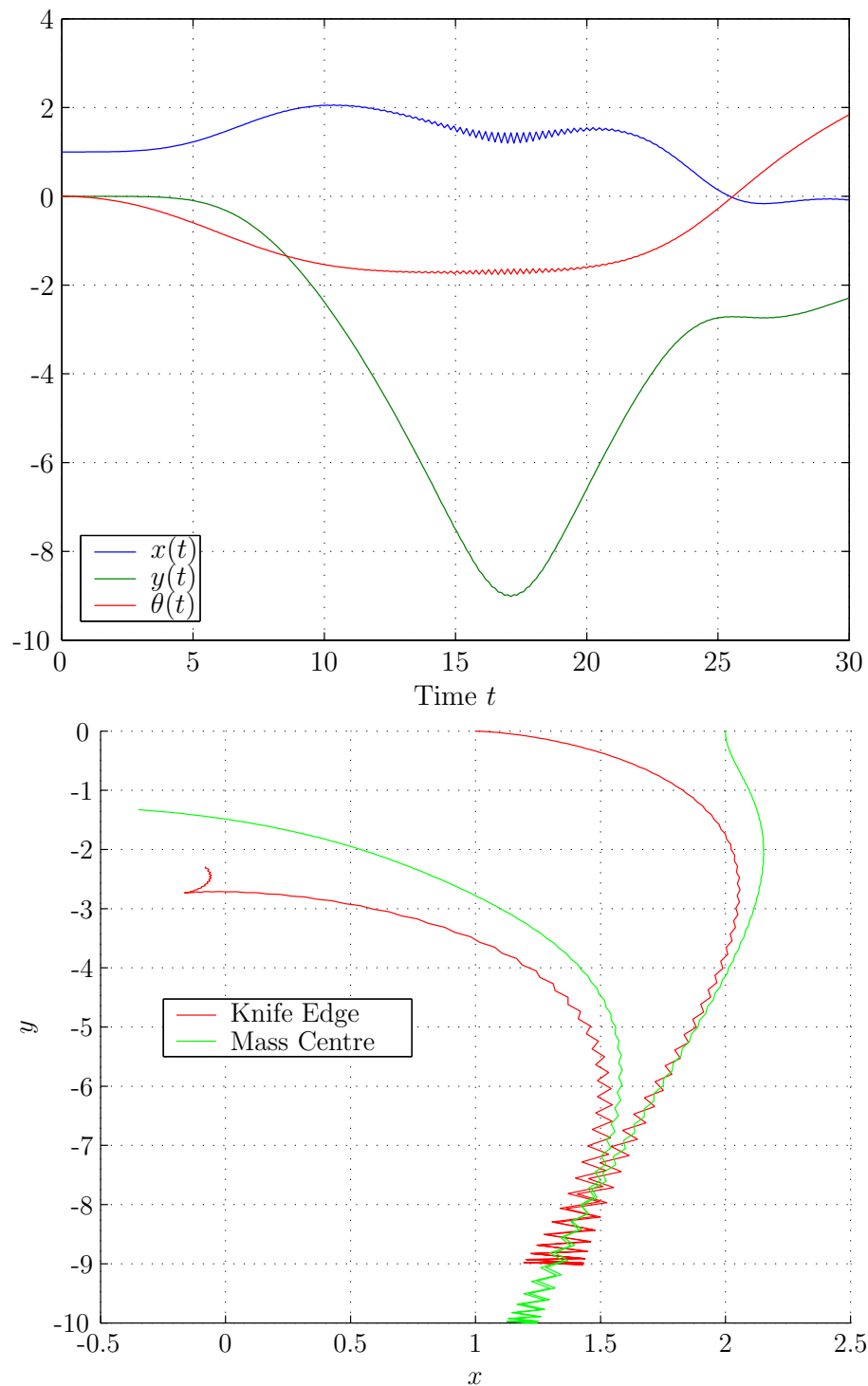
$x(t),\ y(t),\ \theta(t)$

Trajectory of Sleigh

Figure 3.15: Long-time integration of the Chaplygin sleigh by `Hd`. (Left): The generalized coordinate components of the solution plotted individually. (Right): The solution is plotted as a curve in $\mathbb{R}^2$. The down direction of the incline is in the negative y direction. Something very bad obviously happens at $t \approx 16$, the solution curve becomes unstable, oscillates, and eventually settles on a completely different physically correct trajectory around $t \approx 20$. The setup behind this plot is explained in Experiment 3.7.
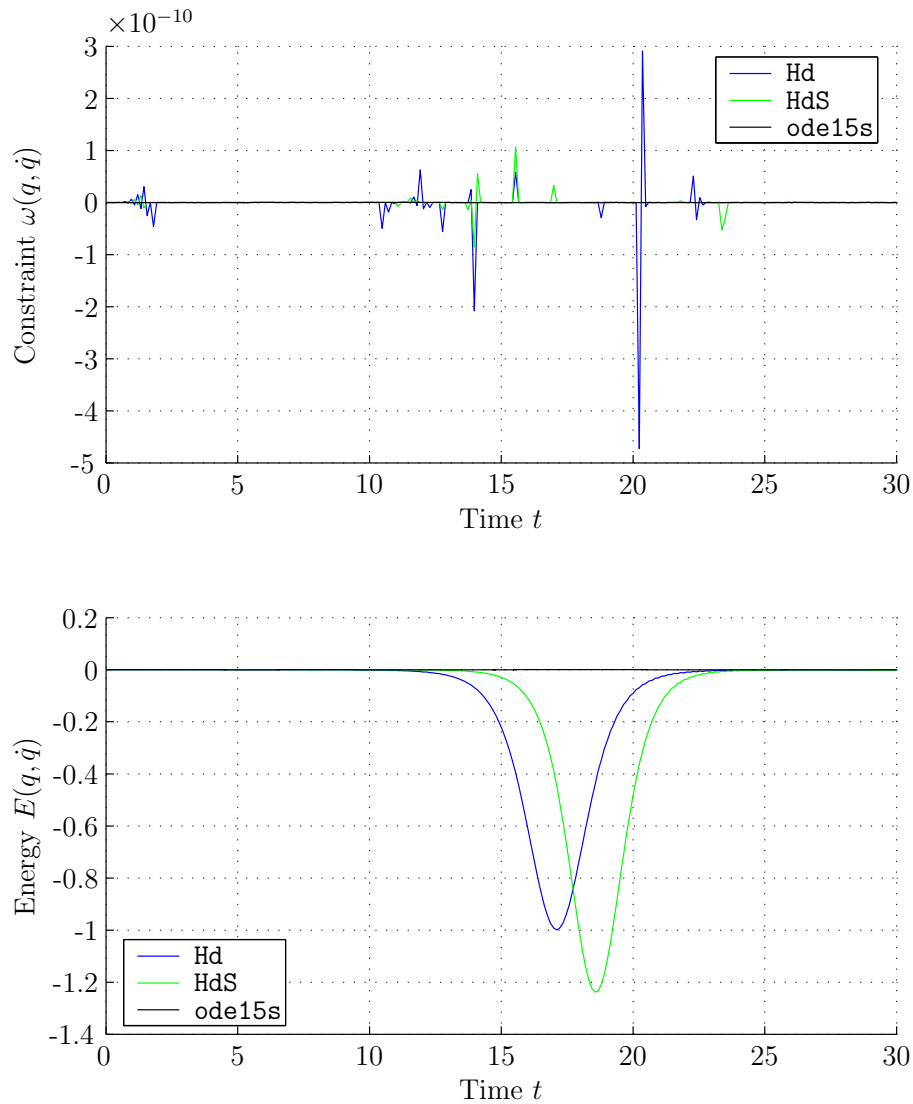
Figure 3.16: Constraint satisfaction and energy by `Hd` and `HdS` in the long-time integration of the Chaplygin sleigh. (Top): The nonholonomic constraint is enforced in each time-step. The numerical solutions show small and essentially noise-like variations. (Bottom): A big dip in the energy error is displayed by both methods in the same time that the numerical solution undergoes oscillations. Anomalies like this in the energy error can be indicators of the fact that something is not right in the solution. When the trajectory of the solution ceases to oscillate, the energy is back to normal. So in fact energy is conserved even when the computed trajectory is completely wrong. The setup behind this plot is explained in Experiment 3.7.
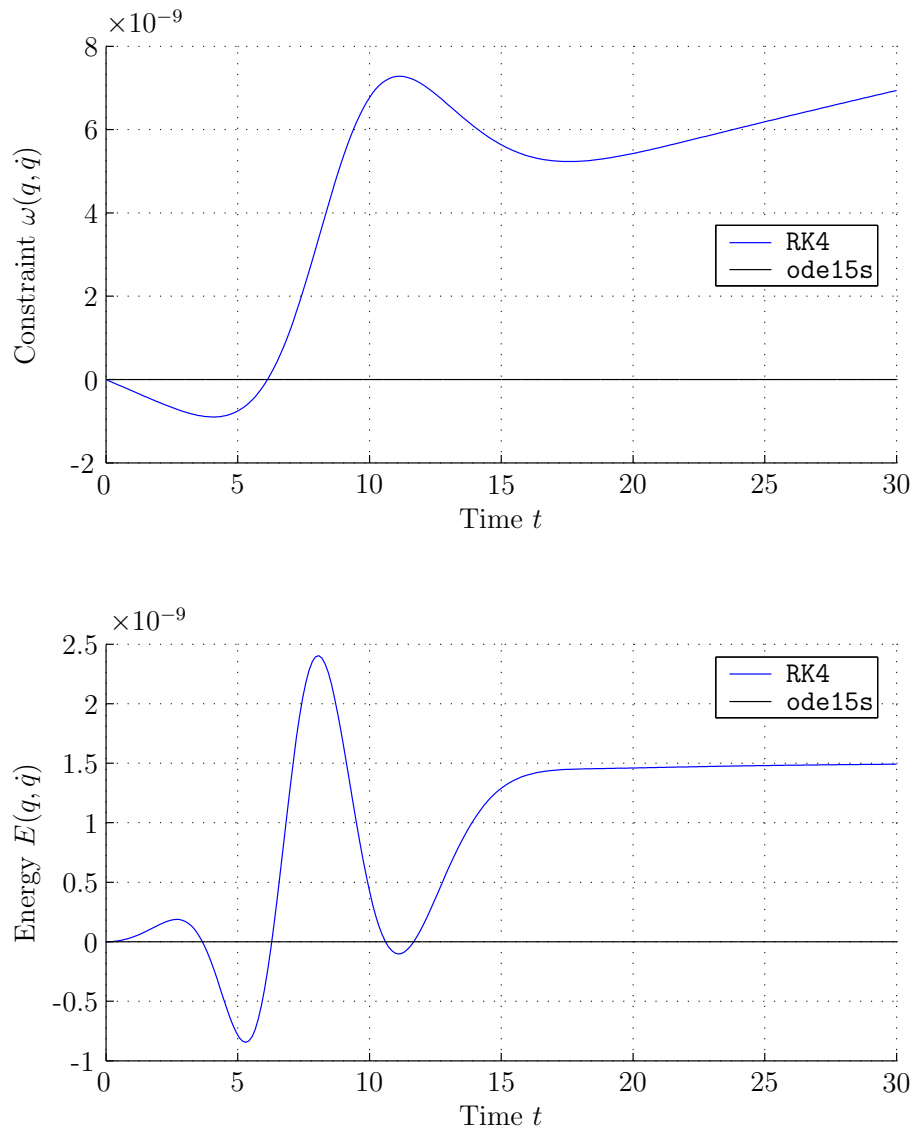
Figure 3.17: Constraint satisfaction and energy by `RK4` of in the long-time integration of the Chaplygin sleigh. (Top): The nonholonomic constraint is not enforced by the algorithm. When the trajectory of the sleigh has stabilized the constraint error appears to grow linearly. (Bottom): As the trajectory of the sleigh stabilizes the energy appears to settle at a constant value. The setup behind this plot is explained in Experiment 3.7.
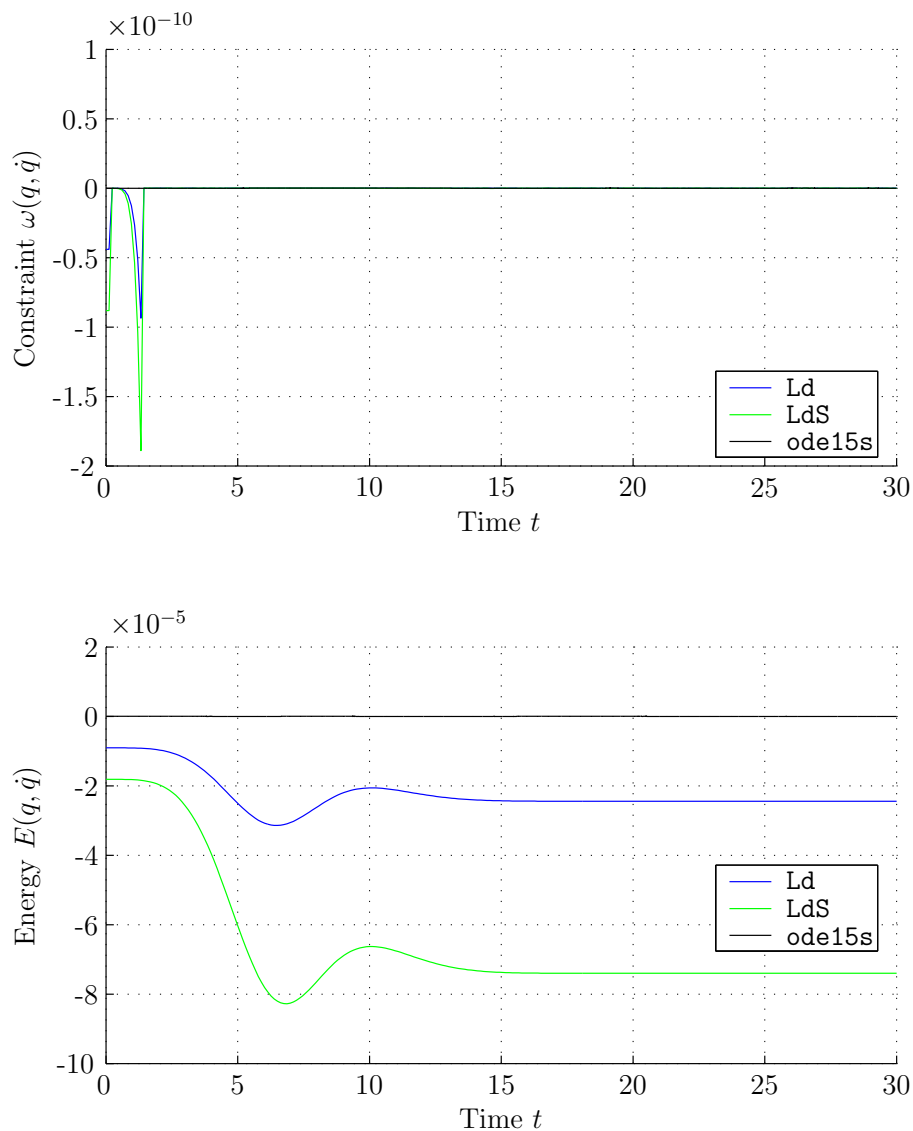
Figure 3.18: Constraint satisfaction and energy by `Ld` and `LdS` in the long-time integration of the Chaplygin sleigh. (Top): The nonholonomic constraint is enforced in each time-step. The constraint error is large early on when the constraint force has the most influence over the solution. (Bottom): The variational integrators `Ld` and `LdS` display the same initial jump as we have seen in the earlier examples. The energy does not oscillate as it did in the previous examples. The setup behind this plot is explained in Experiment 3.7.

Chapter 4

Conclusion

We shall attempt to draw some tentative conclusions and look at some possibilities for further work.

Classical Methods v.s. Two-Step Methods

For short time intervals, the absolute accuracy of the classical integrator outperforms every variational method. One might have expected that the method `RK4` would indeed display convergence of order 4, but that the error would still be much larger than the error of the variational integrators. This is clearly not the case.

For long-time integration the situation is different. While the solution may in one sense be totally wrong for all the integrators, the solution computed by the functioning variational integrators will not exhibit behavior that is totally inconsistent with the underlying physics. Variational integrators do open new problems to long-time numerical integration.

Variational integrators may be more treacherous than classical methods in the sense that the output of a geometric integrators looks correct at a casual glance. Fortunately the behaviour of the energy of the solution can be taken as an indication of the correctness of the method.

The 26-dimensional system of particles integrated in Section 3.2 is the heaviest test of the integrators in this thesis. While it is not possible to say that any of the variational integrators preserve the solution, it is obvious that the lack of energy conservation of the classical method will eventually lead to disaster. Figure 3.11 shows that the amplitude of the solution increases for `RK4`, while the variational integrators do not cause any increase in amplitude. Reducing the size of the time-steps will only postpone this phenomenon. The variational integrators on the other hand show very agreeable energy conservation results.

The Chaplygin sleigh is the only problem where the potential is not quadratic. It appears that a quadratic potential that gives rise to smooth oscillatory motion leads to problems where the variational integrators outperform classical methods. This can be understood intuitively in the sense that quadratic potentials lead to trajectories that constantly curve, in which case classical methods require short step lengths in order to perform

well. The trajectory in the Chaplygin problem, on the other hand, settles into a quasi-stable state where the constraint forces are weak and do not influence the system greatly.

Instability in the One-Step Methods

The Chaplygin sleigh problem uncovers a crippling problem with the one-step integrators we have experimented with in this thesis. Already in the Work-Precision plot in Figure 3.13 there is evidence of instability in the one-step integrators. The fact that the instability in the solution appears for *both* one-step integrators indicate that there is some theoretical difficulty involved in formulating a one-step method for nonholonomic systems, rather than a bug in the implementation.

Furthermore the Work-Precision plot of McLachlan and Perlmutter's particles (Figure 3.7) shows a similar behaviour of the one-step methods as the Work-Precision plot of the Chaplygin sleigh problem (Figure 3.13). This indicates that there may be problems with the one-step integrator in this problem as well.

The instability uncovered by the Chaplygin problem means that the one-step integrator may not be trusted to solve any problem at all correctly. The error in the solution need not always be as spectacular as in the integration of the Chaplygin problem.

The instability in `Hd` and `HdS` takes longer to appear when the step length is decreased, but we found that the instability would always appear in long-time integration. The oscillations would also be smaller when the time-step was smaller, but the "jump" to a different trajectory would be similar in violence.

Further Work

Further work with the material of this thesis may take several directions.

- **Finding the trouble with one-step methods.** Finding out whether the instability of the one-step methods is due to an unknown theoretical difficulty or is due to a bug in the implementation of the methods. This would have been the first priority of the author. It would have required a deeper understanding of discrete mechanics and of stability criteria for numerical integrators.

- **Simulating problems with several holonomic and nonholonomic constraints.** The algorithms constructed require no change to simulate systems with multiple nonholonomic constraints, and very little change in order to accommodate problems with both holonomic and nonholonomic constraints.

- **Optimization of the methods.** It must be said that the classical integrator `RK4` is exceedingly faster than than the variational integrators. Partly this is because of the high accuracy required of the implicit equation solver `fsolve`. It would be interesting to increase the error tolerance of `fsolve` and see how that would affects the geometric properties of the flow. Alternatively Newton's method with a small fixed number of iterations in each time-step could be considered.

Bibliography

[1] Ralph Abraham and Jerrold E. Marsden. *Foundations of Mechanics*. Addison-Wesley, 2nd edition, 1978.

[2] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Number 60 in Graduate Texts in Mathematics. Springer-Verlag, 2nd edition, 1989.

[3] Anthony M. Bloch. *Nonholonomic Mechanics and Control*. Interdiciplinary Applied Mathematics. Springer-Verlag, 2003.

[4] Anthony M. Bloch, Jerrold E. Marsden, and Dmitry V. Zenkov. Nonholonomic dynamics. *Notices of the American Mathematical Society*, 52:324–333, March 2005.

[5] A. V. Borisov and I. S. Mamaev. On the history of the development of the nonholonomic dynamics. *Regular and Chaotic Dynamics*, 7(1):43–47, 2002.

[6] Manuel de León, David Martín de Diego, and Aitor Santamaría Merino. Geometric integrators and nonholonomic mechanics. Preprint IMAFF-CSIC.

[7] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

[8] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric Numerical Integration: Structure Preserving Algorithms for Ordinary Differential Equations*. Number 31 in Series in Computational Mathematics. Springer-Verlag, 2004.

[9] Ernst Hairer, Syvert P. Nørsett, and Gerhard Wanner. *Solving Ordinary Differential Equations 1: Nonstiff Problems*. Number 8 in Series in Computational Mathematics. Springer-Verlag, 2nd edition, 1987.

[10] Ernst Hairer and Gerhard Wanner. *Solving Ordinary Differential Equations 2: Stiff and Differential-Algebraic Problems*. Number 14 in Series in Computational Mathematics. Springer-Verlag, 2nd edition, 1996.

[11] Jerrold E. Marsden and Tudur S. Ratiu. *Introduction to Mechanics and Symmetry*. Number 17 in Texts in Applied Mathematics. Springer-Verlag, 2nd edition, 1999.

[12] Jerrold E. Marsden and Mark West. Discrete mechanics and variational integrators. *Acta Numerica*, 10, 2001.

[13] The Mathworks. *Optimization Toolbox User's Guide*, 3.0.4th edition, March 2006. `www.mathworks.com/access/helpdesk/help/pdf_doc/optim/optim_tb.pdf`.

[14] Robert I. McLachlan and Matthew Perlmutter. Energy drift in reversible time integration. *J. Phys A*, 37(45):L593–L598, 2004.

[15] Jorge Cortés Monforte. *Geometric, Control, and Numerical Aspects of Nonholonomic Systems*. Number 1793 in Lecture Notes in Mathematics. Springer-Verlag, 2002.

[16] James R. Munkrees. *Analysis on Manifolds*. Advanced Book Classics. Westview Press, 1991.

[17] Yuri I. Neimark and Nikolay A. Fufaev. *Dynamics of Nonholonomic Systems*. Number 33 in Translations of Mathematical Monographs. American Mathematical Society, 1972.

[18] Patrick J. Rabier and Werner C. Rheinboldt. *Nonholonomic Motion of Rigid Mechanical Systems from a DAE Viewpoint*. SIAM, 2000.

[19] Sebastian Reich. Symplectic integration of constrained hamiltonian systems by composition methods. *SIAM Journal on Numerical Analysis*, 33(2):475–491, 1996.

[20] Lawrence F. Shampine and Mark W. Reichelt. The MATLAB ODE suite. *SIAM Journal on Scientific Computing*, 18(1):1–22, 1997.

[21] John R. Taylor. *Classical Mechanics*. University Science Books, 2004.

[22] John L. Troutman. *Variational Calculus and Optimal Control: Optimization with Elementary Convexity*. Undergraduate Texts in Mathematics. Springer-Verlag, 2nd edition, 1995.

[23] Ge Zhong and Jerrold E. Marsden. Lie-Poisson Hamilton-Jacobi theory and Lie-Poisson integrators. *Physics Letters A*, 133(3):134–139, 1988.